Southern Cross University
ePublications@SCU

Theses

2007

The emergence of simple business models on the World Wide Web

Michael Featherstone
Southern Cross University

Publication details
Featherstone, M 2007, 'The emergence of simple business models on the World Wide Web', DBA thesis, Southern Cross University, Lismore, NSW.
Copyright M Featherstone 2007

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

The Emergence of Simple Business Models on the World Wide Web

A research thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Business Administration in the Faculty of Business,

Southern Cross University, NSW, Australia

Presented by candidate:

Michael D. Featherstone

Bachelor of Science (Wayne State University, Detroit, Michigan, USA)

Master of Public Administration (Wayne State University)

Submitted: October 2006

ABSTRACT

This research project began with the objective of obtaining a deeper understanding of the conduct of business on the Internet. Research regarding the overall performance of business on the Web has been limited due to several constraining factors.

Firstly, the shear size, volume, and perceived complexity of business on the Web made it an imposing target for research. Secondly, the interdisciplinary nature of the field had an initial dampening effect on research activity. Kuhn (1962) describes this phenomenon as the insufficiency of methodological directives. He writes (p. 3) 'Instructed to examine electrical or chemical phenomena, the man who is ignorant of these fields but who knows what it is to be scientific may legitimately reach any one of a number of incompatible conclusions. Among those legitimate possibilities, the particular conclusions he does arrive at are probably determined by his prior experience in the other fields'. Thirdly, academic interest in Web businesses was somewhat diminished by the bursting of the dot com bubble during 2000-2001 as evidenced by the fact that many academic programs in electronic commerce have been discontinued or significantly reduced (Featherstone, Ellis & Borstorff 2004). Lastly, methodological issues arose which limited the application of previously available sampling methods. The impact of this meant that drawing representative samples of websites in order to examine business behaviour became a more complex endeavour.

As a result, many fundamental questions regarding business on the Web have remained unanswered. For example, there has been no clear answer to the question of what new business models, if any, are evolving within the Web environment. What is the role of entrepreneurship employing the Web? What are the key elements or mechanisms driving business expansion in the Web? Some have suggested that Web use necessitates greater business cooperation than may be necessary in the non-virtual world. Is there evidence that this is so? This thesis addresses these questions. It accomplishes this by reducing the function of business on the Internet to its simplest element, the domain name, and by employing innovative methodologies to explore the business environment of the Web.

Using Complexity and Network Theories as a contextual framework , and based upon a review of existing literature, this thesis develops a series of propositions regarding

both the conduct and attributes of Web business, and proceeds to present evidence of confirmation or refutation of these propositions.

TABLE OF CONTENTS

ABSTRACT ..... **I**

TABLE OF CONTENTS ..... **III**

LIST OF TABLES ..... **VII**

LIST OF FIGURES ..... **VIII**

LIST OF FIGURES ..... **VIII**

CERTIFICATION ..... **X**

ACKNOWLEDGEMENTS ..... **XI**

CHAPTER 1 – MOTIVATION FOR THE RESEARCH **1**

1.0 REGARDING THE FORMAT OF THIS DISSERTATION 1

1.1 INTRODUCTION – WHY THIS AREA OF INQUIRY? 1

1.2 METHODOLOGY EMPLOYED 5

1.3 POINTS OF ACADEMIC INTEREST INCLUDED IN THE THESIS 5

1.4 POINTS OF MANAGERIAL AND PUBLIC POLICY INTEREST INCLUDED IN THE THESIS 6

1.5 OUTLINE OF THE THESIS 7

1.6 DEFINITIONS FOR THIS RESEARCH 9

1.6.1 Definitions of technology centric terms 9

1.6.2 Definitions of research centric terms 14

1.7 CONCLUSION 15

CHAPTER 2 – SUMMARY FINDINGS **16**

2.1 INTRODUCTION 16

2.2 SUMMARY OF QUANTITATIVE RESEARCH FINDINGS 16

2.3 SUMMARY OF QUALITATIVE RESEARCH FINDINGS 19

2.4 CONCLUSION 23

CHAPTER 3 – LITERATURE REVIEW AND RESEARCH PROPOSITIONS **25**

3.1 INTRODUCTION 25

3.2 INTERNET BUSINESS MODELS 27

3.3 COMPLEXITY THEORY 36

3.3.1 Emergence process and complex systems 38

3.3.2 Phase transition in complex systems 40

3.3.3 Summary of elements of complex systems 41

3.4 NETWORK THEORY 42

3.4.1 Summary of network attributes 46

3.5 CONCLUSION 46

CHAPTER 4 – RESEARCH METHODOLOGY 48

4.1 INTRODUCTION 48

4.2 RESEARCH OBJECTIVE 50

4.3 RESEARCH FRAMEWORK 51

4.3.1 Justification for research paradigm 51

4.3.2 Justification for the research methodology 53

4.4 RESEARCH STRATEGY – QUANTITATIVE PHASE 56

4.4.1 Initial sampling attempts 57

4.4.2 Other Web sampling methods 60

4.4.3 Sampling – Top level domain selection 61

4.4.4 Domain zone sampling method 61

4.4.5 Creation of a sample frame 62

4.4.6 Analysis of quantitative data 63

4.5 RESEARCH STRATEGY - QUALITATIVE PHASE 67

4.5.1 Data collection for qualitative analysis 71

4.5.2 Data analysis process 71

4.5.3 Resolving the erroneous classification issue 72

4.5.4 Pattern recognition 74

4.6 ETHICAL CONSIDERATIONS 80

4.7 CONCLUSION 82

CHAPTER 5 – FINDINGS AND DISCUSSION 83

5.1 INTRODUCTION 83

5.2 QUANTITATIVE RESEARCH FINDINGS 84

5.3 QUALITATIVE FINDINGS 87

5.4 DOMAIN NAME MONETISATION STRUCTURES 87

5.4.1 Example 1 - A simple 'level one' business model 89

5.4.2 Example 2 - Simple 'level two' business models 91

5.4.3 Domain name parking strategies 92

5.4.4 The missing link: Advertiser aggregation entities and the impact of key word auctions 99

5.4.5 Example 3 - Simple 'level three' business models 105

5.4.6 Summary of simple business models 110

5.5 RELATING RESEARCH PROPOSITIONS TO THE FINDINGS 111

5.5.1 The first research proposition 111

5.5.2 The second research proposition 113

5.5.3 The third research proposition 113

5.5.4 The fourth research proposition 114

5.5.5 The fifth research proposition 117

5.5.6 The sixth research proposition 118

5.6 CONCLUSION 119

CHAPTER 6 – RESEARCH CONTRIBUTIONS, LIMITATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH 121

6.1 INTRODUCTION 121

6.2 CONTRIBUTIONS OF THE RESEARCH 121

6.3 RESEARCH FINDINGS AND THE LITERATURE 124

6.4 LIMITATIONS OF THIS RESEARCH 124

6.5 SUGGESTIONS FOR FURTHER RESEARCH 125

6.6 CONCLUSION 126

BIBLIOGRAPY 128

APPENDICES 139

APPENDIX A – PROGRAM TO RANDOMLY EXTRACT DOMAIN NAMES FROM LARGE TOP LEVEL DOMANE ZONE FILE 139

APPENDIX B – DOMAIN NAME ACQUISITION FROM THE FINAL SAMPLING FRAME 142

APPENDIX C – INSTRUCTIONS FOR DOMAIN NAME RESOLUTION 143

APPENDIX D – SAMPLE OF INITIAL CATEGORY STRUCTURE 144

APPENDIX E – SAMPLE OF COMPLETED DATA SHEET 145

APPENDIX F - EXAMPLE OF TEXTPAD ENUMERATED DATABASE RECORDS..... 147

APPENDIX G – EXAMPLE OF VERISIGN ZONE FILE ACCESS AGREEMENT..... 148

APPENDIX H - PEER REVIEWED PUBLICATIONS TO WHICH THIS STUDY CONTRIBUTED..... 154

LIST OF TABLES

TABLE 2.1 RESULTS FROM THE SAMPLE OF 'NET' ZONE DOMAIN NAMES 17

TABLE 2.2 RESULTS FROM THE SAMPLE OF 'COM' ZONE DOMAIN NAMES 18

TABLE 3.1 RAYPORT AND JAWORSKI BUSINESS MODELS 30

TABLE 3.2 LAUDON AND TRAVER BUSINESS MODELS 32

TABLE 4.1 DECONSTRUCTION OF THE PHRASE DESCRIBING THE RESEARCH METHODOLOGY 54

TABLE 4.4 FIELD LABELS AND DESCRIPTIONS – DATA CAPTURE SHEET 72

TABLE 5.1 RESULTS FROM THE SAMPLE OF 'COM' ZONE DOMAIN NAMES 84

TABLE 5.2 RESULTS FROM THE SAMPLE OF 'NET' ZONE DOMAIN NAMES 86

LIST OF FIGURES

FIGURE 3.1 ADAM'S SPECTRUM OF BUSINESS USERS OF THE WEB 29

FIGURE 3.2 LINEAR REPRESENTATION OF A POWER LAW DISTRIBUTION SOURCE: (NIELSON 1997) 45

FIGURE 3.3 LOGARITHMIC REPRESENTATION OF A POWER LAW DISTRIBUTION 45

FIGURE 4.1 LISTING OF TEXTPAD ENUMERATED DOMAIN NAMES FROM THE DATABASE OF 30,000 DOMAIN NAMES 64

FIGURE 4.2 THE FIRST EXAMPLE OF WEBSITES WHICH LEAD TO PATTERN RECOGNITION 75

FIGURE 4.3 THE SECOND EXAMPLE OF WEBSITES WHICH LEAD TO PATTERN RECOGNITION 75

FIGURE 4.4 THE 'BLUE HIGHWAY' DESIGN REPRESENTED AN EXAMPLE OF PATTERN DUPLICATION. 76

FIGURE 4.5 DUPLICATE PATTERNS OF WEBSITES WITH UNIQUE DOMAIN NAMES BEGAN TO EMERGE. 77

FIGURE 4.6 GRAPHIC DEPICTION OF THE QUALITATIVE RESEARCH PROCESS (THE HERMENEUTIC CYCLE) 80

FIGURE 5.1 TWO PARKED SITES WITH DIFFERENT DOMAIN NAMES SHARE THE SAME DESIGN 88

FIGURE 5.2 A LEVEL ONE BUSINESS MODEL REPRESENTING THE SPECULATIVE PURCHASE OF A DOMAIN NAME 89

FIGURE 5.3 EXAMPLE OF A PARKED DOMAIN NAME 92

FIGURE 5.4 WEB HOST STRATEGY FOR A NON-DEPLOYED DOMAIN NAME IN 2005 93

FIGURE 5.5 WEB HOST STRATEGY FOR A NON-DEPLOYED DOMAIN NAME IN 2006 NOW INCLUDES REVENUE GENERATING LINKS TO EXTERNAL SITES 93

FIGURE 5.6 THIS REPRESENTS A SUBTLE VARIATION OF A PARKED DOMAIN NAME. THE OWNER RELIES ON DIRECT NAVIGATION AND MISSPELLING TO GENERATE TRAFFIC TO THIS SITE. 95

FIGURE 5.7 LEVEL TWO PARKING SITE REPRESENTING FIRST GENERATION DESIGN TECHNOLOGY 97

FIGURE 5.8 EVOLUTION OF NEW LEVEL TWO PARKING SITE DESIGNS 97

FIGURE 5.9 SIMILARITY OF THE DESIGN MARKS THIS AS A WEBSITE FROM THE SAME DOMAIN NAME PARKING SERVICE PROVIDER AS SHOWN IN FIGURE 5.8. 98

FIGURE 5.10 DEPICTS NEW DESIGN MODELS EVOLVING IN LEVEL TWO SITES. 98

FIGURE 5.11 A GRAPHIC REPRESENTATION OF TYPICAL LEVEL TWO RELATIONSHIPS 101

FIGURE 5.12 GRAPHIC REPRESENTATION OF LEVEL THREE RELATIONSHIPS 107

FIGURE 5.13 AN AFFILIATE SITE REPRESENTS A HIGHER LEVEL OF SOPHISTICATION IN THAT THE DESIGN THOROUGHLY INTEGRATES THE FUNCTION OF THE SITE WHICH IS TO ENCOURAGE VISITORS TO CLICK ON THE ADS LISTED. 108

FIGURE 5.14 SHOWS THAT LEVEL THREE SITES FEATURE MUCH GREATER VARIABILITY IN DESIGN THAN DO LEVEL TWO SITES 109

FIGURE 5.15 LINEAR REPRESENTATION OF DISTRIBUTION OF INBOUND LINKS TO WEBSITES CATEGORISED AS 'BUSINESS' IN THE .COM ZONE 116

FIGURE 5.16 LOGARITHMIC SCALE OF DISTRIBUTION OF INBOUND LINKS TO WEBSITES CATEGORISED AS 'BUSINESS' IN THE .COM ZONE 117

FIGURE 5.17 SUGGESTED EXTENSION TO ADAM'S SPECTRUM OF BUSINESS USERS OF THE WEB 119

CERTIFICATION

I certify that the substance of this thesis has not been submitted for any other degree to date, nor is it currently being submitted for any other degree.

I also certify that, to the best of my knowledge, I have acknowledged all sources used and assistance received in preparing this thesis.

Michael D. Featherstone:

Date: 16-10-2006

ACKNOWLEDGEMENTS

This dissertation represents the culmination of a four year journey. This was a team effort and it is appropriate to acknowledge those who have assisted, mentored and coached me to the finish line.

I thank my advisor Dr. Stewart Adam. At the very time I thought this paper was unlikely to come to fruition, Dr. Adam joined my team. His advice, wisdom, knowledge of the field and, not least, his encouragement were instrumental in the accomplishment of this goal. I consider myself very fortunate to have found not only a superb advisor, but a friend and colleague as well.

Special acknowledgement is due to Dr. Patricia Borstorff of Jacksonville State University. She was convinced I would complete this document long before I was. She was most generous with her time and her advice. She was unyielding in her insistence that I focus on completing the dissertation. As much as anyone, she mentored me through this entire process.

During the first three years of this project, Dr. Paul James provided me with the benefit of his astute insights into the DBA process during weekly telephone conversations. He gave me both a target to shoot for and an incentive to see the project through.

I also wish to acknowledge the contributions of the staff at Southern Cross University, in particular Dr. Peter Miller, Dr. Allan Ellis, Sue White and Susan Riordan. They have had the difficult task of communicating through time and distance, and they have made the journey a very smooth one. Thanks as well to my doctoral candidate colleagues at Southern Cross University. I will miss the annual opportunity to share thoughts, ideas and libation.

I must also thank Chris Sheridan, Pam Sullivan and Rusty Warren of VeriSign Corporation for expediting my research access to the .com and .net Internet domain name data files.

Finally, I thank my wife Judy for her forbearance over the past four years during which, I've spent a disproportionate amount of time closeted in my office, reading or writing. I am grateful for her tolerance and understanding.

> *It grew like a weed without the shaping impact of a consciously defined social contract to determine what the medium should be delivering, for what purpose, and with what value. Perhaps it happened too fast, too unexpectedly for constructive reflection; so that the technology shaped us before we could shape it (Donnelly 1986, p. 21).*

CHAPTER 1 – MOTIVATION FOR THE RESEARCH

1.0 Regarding the format of this dissertation

The format of this dissertation is adapted from one suggested in an article by Uncles (1998). Uncles believes his structure tends to shift the emphasis of the dissertation to the students contributions to knowledge and away from a descriptive recitation of alternate philosophies, methodologies, and research paradigms. The implementation of this format is most noticeable with the inclusion of Chapter 2 which summarises the findings of the research project. Uncles prefers a shorter literature review, and greater emphasis on the findings and discussion, as well as the contributions of the research. Beyond the changes mentioned, the format of this thesis closely follows the suggestions of Perry (2002) with only occasional and subtle influence from Uncles, such as the title of this chapter.

1.1 Introduction – Why this area of inquiry?

One rationale for research in this area of inquiry is provided by Donelly (1986, p. 6):

> An understanding of how the new electronic media will affect our futures depends on a fact-based and involved understanding of the new electronics. … Without some understanding, we may not even be able to march toward the future – it will march over us instead.

The motivation for this particular research project came from a paper which was presented at a conference in 2004. The paper suggested that small businesses did not use the Internet. Though the paper generalised to all small business, what it really said was that small businesses on the main street of any town don't really use the Web to the extent they should (in the eye of the researcher) or could use it (e.g. Dilts & Kahai 2004). This did not seem to represent an accurate description of small business on the Web. There should be an explosion of entrepreneurship brought about by the ease of internet business development and the plethora of business niches

any such new environment produces (Rothschild 1990). Meg Whitman, CEO of E-bay said:

> The surprising thing is how empowering the Net in general has been for small business. Four years ago, everyone thought IBM, Home Depot, Target, and the big companies would dominate the Net like they do in the physical world. And they do well on the Net. But small business has a bigger presence on the Net than you would imagine. (Harbrecht 2003, www.businessweek.com)

The Internet should be a caldron of entrepreneurial activity and innovation (Featherstone & Borstorff 2005; Featherstone & Ellis 2005; Goodin 2002; Weinberger 2002). Therefore, the first of objective of this thesis is to obtain a deeper understanding of the conduct of business on the Internet. Marshall and Rossman (1989, p. 22) write that "Research is worth doing if it builds knowledge". This project aims to accomplish this objective by conducting exploratory, descriptive and explanatory research into the business environment of the World Wide Web and the fundamental building block of any e-business – its domain name. In so doing, it further aims to provide the 'constructive reflection' of which Donnelly speaks. There is a need for this because our view of the Internet – or more precisely the World Wide Web – is very much like a 15th Century European's view of distant lands might have been, incomplete and with a predisposition to impose old familiar labels upon new and unfamiliar entities. It remains, to a very large extent, 'Terra Incognita'.

This is because the Web is a new and relatively unfamiliar environment for the conduct of business as well as for the conduct of business research. See for example, Drew (2002) or Barabási (2002). Webster and Watson (2002) speculate that the youth of the entire field of Information Technology may also explain why there are so few broad theoretical articles regarding this topic. This thesis employs two equally new theoretical frameworks to help explain the business transitions taking place on the Web: Network Theory and its parent discipline, Complexity Theory.

Consider the brief history of research regarding the Web. From its inception in 1991until 1994 little was known outside of the small ring of researchers who were involved in its development. Most notable of these were Sir Tim Berners-Lee the

inventor of the Web and the Hyper-Text Markup Language and Marc Andreesen who developed the Mosaic Web browser (Berners-Lee & Fischetti 1999).

Fledgling businesses sprung up on the Web in 1994, and by 1995 the phenomenon which was labelled the dot com bubble had begun in earnest. Throughout (or because of) the frenzy of the dot com craze and the uproar over the bursting of the dot com bubble in 2001 many fundamental questions about business on the Web have remained unanswered. What types of businesses are on the Web? Is the Web more amiable to large business or to small business? Does the Web consist of mostly entrepreneurial start-ups or mostly companies who have adapted their pre-existing business models to this new environment?

Much of the research regarding business on the Internet has been accomplished from an external perspective or from the outside looking in. For the purpose of this research, an external perspective is defined as one in which data is collected on small businesses in a community (or state or nation), and then, from that data, the researcher ascertains which of small businesses are using the Internet. This process could present a significantly different result from research where one collected data on business Internet websites and from that data, then determined how many or what proportion were small businesses.

Extant research often focuses on businesses and or business models which have only recently migrated to this new environment, while less attention is focused on 'native' Web businesses. Since this environment has only been around since 1994, literally all businesses and business models are relatively recent migrants. A potential problem with regard to business on the Internet may be that researchers could be operating under the assumption that is just another form of business as usual. Little research has been done to confirm that traditional business frameworks are relevant for business content on the Web (see for example Biswas 2004; Porter 2001; Rayport 1999).

The need for fundamental empirical research in this field has been acknowledged by several researchers. Barabási (2002, p. 178) states a compelling rationale for conducting exploratory research in the Web's business environment, 'Our life is increasingly dominated by the Web. Yet we devote remarkably little attention and resources to understanding it'.

Research at the early stages of any phenomenon is necessarily descriptive in nature. It provides a necessary foundation for research that follows. Carlile and Christensen (2005, p. 2) state:

> The phenomena being explored in this [early] stage include not just things such as people, organizations and technologies, but processes as well. These observations can be done anywhere along the continuum from analysis of huge databases on the one end, to field-based, ethnographic observations on the other.
>
> Without insightful description to build upon researchers can find themselves optimizing misleading concepts.

Relative to the extensive literature produced on the importance and potential of the Internet as a tool (see for example: Porter 2001) or as an element of the physical world's business environment, empirical research regarding the demographics of the vast majority of Internet firms themselves, or their marketing and revenue strategies, is limited and sketchy (Colecchia 2000; Constatinides 2004). Drew (2002, p. 21-22) writes 'Many academic empirical investigations and surveys in e-business suffer from small sample sizes, with consequent questions as to the meaning, validity and reliability of findings'.

Because of sheer size and the exponential growth, Internet business and the Internet itself are difficult subjects for analysis (Bauer & Scharl 2000; Kishi et al. 2000; O'Neill, McClain & Lavoie 1998; Rusmevichientong et al. 2001; Yager 2002). For example, by the end of 2004, there were approximately 30 million registered .com domain names and another five million .net names (VeriSign 2004a). The Web is growing at a staggering rate of nearly one million new domain names per month (Sheridan, pers.comm. 2004). This rate of growth coupled with the size of the Web make it a taxing topic for research (Kishi et al. 2000).

Clarke (2000) suggests further obstacles to Web research. For example, he cites its recent emergence, the rapid change that has always characterised the domain, the very substantial variation in behaviour in apparently similar contexts, and the enormous amount of attention paid to it by the populist media and marketing interests, with the inevitable distortion of terminology and data that this engenders. Drew (2002)

identifies additional problems such as the lack of familiarity with e-commerce technologies by many management scholars, and the lack of established instruments and research approaches.

These factors may explain why the preponderance of the existing e-business literature has been fashioned from a perspective which is external to the Web, frequently focusing on material world companies which may or may not have established a Web presence. These business entities present more familiar research paradigms to business researchers.

Another element which is important to this thesis is that, aside from the overall growth of the Web itself (and perhaps more importantly), there has been an increase in the proportion of commercial websites (Brin & Page 1998). The present study suggests that much of this increase is a result of simple business models evolving within the business environment of the Web. Throughout this thesis, the research emphasis is on examination of elementary processes and business changes that would be easy to overlook unless the research methodology employed some Web sampling technique, and thus was conducted from the 'inside out'. The reported examines the significance that these smaller business entities represent in the overall commercial environment of the Web.

1.2 Methodology Employed

As befitting a study exploring and describing a new field and utilising new theoretical frameworks, a relatively new methodological approach is used. This study employs an *empirical multi-phase sequential mixed-methods design.* This methodology is more fully discussed in Chapter 4.

1.3 Points of academic interest included in the thesis

The study reported herein includes the following items of academic interest:

Firstly, this research identifies a methodological process for extracting random samples of websites from Web top level domains. This is important because at this early phase of e-business, without some mechanism for sampling and modelling business activity on the Web, it is difficult to understand what is happening in the environment and to determine trends with regard to business or marketing activities.

One cannot ascertain the viability of any given business model, if one does not recognise the existence of an appropriate model.

Secondly, it identifies new and relatively unexplored business structures and business models which have evolved from within the environment of the Web. The reason these business models have proven difficult to identify is that unless one obtains a viable sample of Web domain names, these models simply do not lend themselves to discovery. For example, one business model actually relies on the user (or Web surfer) to make mistakes typing in domain names in order to generate revenue for the site. These sites would only be encountered by 'direct navigation' techniques (see the definitions in section 1.6 page 9 of this chapter) and then only if one purposefully misspelled common domain names. Thus, this category of website would not be likely to show up in any scheme that uses search engines as a Web navigation mechanism.

Thirdly, it proposes a simple classification scheme for these business structures. These elemental models employ simple business strategies and form identifiable categories exhibiting patterns of increasing complexity and sophistication suggesting a hierarchical classification scheme to distinguish the categories.

Fourthly, this thesis provides confirmatory evidence for prior research estimating the proportion of dynamically generated Internet Protocol addresses.

Finally, this thesis confirms that business entities on the Web form a power law distribution when measured by inbound links to the business sites. The power law distribution of websites was first suggested by Barabási and Albert (1999). However, the applicability of a power law distribution for certain top level domains and business categories had been questioned by Pennock et al. (2002) – a matter for further discussion in this treatise.

1.4 Points of managerial and public policy interest included in the thesis

The four items likely to be of most relevance to management and public policy interests are:

Firstly, it identifies new business models which may be of interest to entrepreneurs and small or rural business enterprises.

Secondly, it identifies these simple business models as a key driver of the overall growth in domain name acquisition on the Web.

Thirdly, contrary to some previous research findings, this thesis finds there is a robust entrepreneurial and small business community on the Web, though this determination could only be made from research involving random sampling methods from which one might generalise to the population of websites in a specific zone, such as the .com domain (Rayport 1999). Further, this research indicates some previous studies may underestimate the actual uptake of e-business amongst entrepreneurs and rural and small businesses. Thus, based on findings reported in this thesis, a review of public policy factors affecting the diffusion and uptake of e-business by entrepreneurs and small business enterprises may be in order.

Fourthly, the thesis submits that the power law distribution found in the .com zone is at least partially explained by preferential attachment mechanisms in place within the business environment of the Web. These mechanisms take the form of search engine marketing strategies and they are becoming ubiquitous on the Web. The results of the study may enable companies to become more familiar with the impact of these marketing strategies in order to more effectively employ them.

1.5 Outline of the thesis

As noted earlier, the adopted thesis format is adapted from a model proposed by Uncles (1998, p. 90-91), who wrote:

> The starting point is an abstract in which the main storyline and key findings are presented succinctly. The reasons for investigating the chosen area are discussed next, so as to engage the attention and interest of the reader. This is followed by a summary of the findings – an elaboration of the material in the abstract. For those readers short of time who only want the main 'take-out' they need read no further. Readers who plan to read on now have a framework for thinking about the research project as a whole - if nothing else this helps to set expectations.

Going forward from this point, the thesis conforms to the more traditional structure advocated by Perry (1998) with two minor exceptions. The third chapter contains the familiar literature review and synthesis with added conceptualisations and research

propositions suggested by the literature. As Perry (2002, p. 7) notes '… some methodologies may be difficult to fit in the structure. For example, grounded theory research may not place as much significance on the Chapter 2 review of the literature and historical research may require different structures'. This thesis also utilises theory building methodologies and incorporates elements of grounded theory. In a field with very little research history, one would expect to find a shorter literature review chapter as is the case in this thesis.

On the other hand, the thesis utilises a mixed-method research methodology, and as a result, the methodology chapter is somewhat longer than one would expect to encounter in a thesis using a single quantitative or qualitative method.

To summarise the organisation of the thesis:

Chapter 1 discusses the motivation for the research project, the rationale for conducting the study, points of academic interest, and points of business and public policy interest.

Chapter 2 presents a summary of the research findings and provides further background for the research.

Chapter 3 presents a review and synthesis of the pertinent literature. At appropriate junctures in the chapter, conceptualisations and research propositions suggested by the literature are defined.

Chapter 4 presents the research questions and summarises the research propositions before discussing the research paradigm directing the study and the research methodologies employed. It includes the aims of the research and the framework to realise the aims. Mixed methodological research projects are generally more time consuming (Creswell 2003; Tashakkori & Teddlie 2003) and, since – by definition – it involves the use of more than one research method, Chapter 4 is divided into two parts, each reflective of a specific research phase. Part one discusses the quantitative research methodology and part two discusses the qualitative research methods.

Chapter 5 presents a detailed description of the findings of the research as well discussion of the relevance of the findings. The chapter revisits the research

propositions and the implications of the findings in relation to the theoretical frameworks of complexity theory and network theory.

Chapter 6 highlights the contributions of this research project with particular attention to points of academic, business and policy contributions. Additionally, this chapter discusses the relevance of the findings of this thesis in the context of the extant literature and Web business typologies. Suggestions for further research and the limitations of this study conclude the chapter.

1.6 Definitions for this research

Definitions adopted by researchers are often not uniform, so key and controversial terms are defined to establish positions taken in DBA research (Perry 1998). Researchers define terms so that readers can clearly understand the researchers precise meanings (Creswell 2003). Defining terms adds an element of precision to the researcher's study. Firestone (1987, p. 17) puts it this way:

> The words of every day language are rich in multiple meanings. Like other symbols, their power comes from the combination of meaning in a specific setting….Scientific language ostensibly strips this multiplicity of meaning from words in the interest of precision. This is the reason common terms are given "technical meanings" for scientific purposes.

The same may be said for definitions of technical terms of reference related to the Web. The entire construct of business on the Web is new and terms and descriptions change and evolve as quickly (or more quickly) as the Web itself. Thus, it is appropriate for this thesis to provide definitions in two separate sets, one with a technology and Internet emphasis and the other with an emphasis on the research terminology employed in this thesis.

1.6.1 Definitions of technology centric terms

Affiliate Model	A business offering financial incentives (in the form of a percentage of revenue) to affiliated partner sites. The affiliates provide purchase-point click-through to the merchant (Rappa 2005). Affiliates may be pay per performance, in which case a percentage of revenue is

	returned to the referrer site if a purchase is transacted by the referee, or pay per click in which a small fixed amount ($.10 for example) is returned to the referrer site.
Business Model	A business model is a description of how a company intends to create value in the marketplace. 'In the most basic sense, a business model is the method of doing business by which a company can sustain itself—that is, generate revenue' (Rappa 2005).
Direct Navigation	Definitions of direct navigation may vary. Some include Web navigation from a 'favorite places' menu as direct navigation. For the purposes of this thesis, direct navigation refers only to the function of physically typing in a Web URL or domain name in the address bar of a browser.
Directed Network	A network where the links point only in one direction. The World Wide Web is an example of a directed network where each link on a given Web page point outward to another node or Web page (Newman 2003).
Domain Name	The unique name that identifies an Internet site. Domain Names always have two or more parts, separated by dots, for example, SCU.EDU. The part on the left is the most specific, and the part on the right is the most general. A given machine may have more than one Domain Name but a given Domain Name points to only one machine. For example, the domain names: scu.edu, mail.scu.edu, workshop.scu.edu can all refer to the same machine, but each domain name can refer to no more than one machine. Domain names function as nodes in the World Wide Web network.

Domain Name Resolution	The process that returns a web page after a name is entered into a browser or sends/receives an email after a valid domain name is entered as an email address (VeriSign 2004c).
Domain Name Server (DNS)	See Name Server, below.
E-Business	Electronic business. The term is used to describe the conduct of business on the Internet, usually on the World Wide Web. Not only buying and selling but also servicing customers and collaborating with business partners.
E-Commerce	Electronic commerce. The buying or selling of goods and/or services on the Internet, usually the World Wide Web. More formal definitions abound in literature; as 'Technically mediated exchanges between parties (individuals or organizations) as well as the electronically based intra- or inter-organisational activities that facilitate such exchanges' (Rayport & Jaworski 2004, p. 4). It is the sales transaction requirement which makes e-commerce an inappropriate term to describe all businesses this project may discover, some websites may only provide descriptive information about the business, and not allow transactions to be completed on the website itself.
Ecosystem	A complex adaptive system, a heterogeneous assemblage of individual agents that interact locally and that are subject to evolution based on outcomes of those interactions (Levin 2005).
Emergence	The process of deriving new and coherent structures, patterns and properties in complex systems.

Internet	A global system of interconnected computer networks using Transmission Control Protocol/Internet Protocol (TCP/IP) to communicate. The Internet is a physical network of computers linked together by optical fibre and other data connections. The Internet grew out the Advanced Research Projects Agency Network (ARPANET), which was developed by the United States Department of Defence in the decade of the 1960's.
IP Address	Internet Protocol Address. An identifier for a computer or device on the Internet. Networks using this protocol route messages based on the IP address of the destination. The format of an IP address is a 32-bit numeric address written as four numbers separated by periods or full stops. Each set is referred to as an octet. Each number can be zero to 255. For example, 1.160.10.240.
IT	Information Technology. The acquisition, processing, storage and distribution of all types of information using computer technology and telecommunication systems.
Meta-tags	Meta-tags are HTML tags that can be used to identify the keywords and description of the page, etc. The most common use of a meta-tag in online marketing is the keyword and description tags, which tell the search engines that index meta-tags what description to use in their search query results.
Name Server **(also known as a DNS or Domain Name Server**	Name servers are setup to maintain host addresses for each sub-domain or zone within the name space as well as to maintain the address of root name server. A DNS server is capable of resolving the IP address for any host in the Internet.

Network	A set of items called vertices or nodes with connections between them called edges or links (Newman 2003). In the specific case of the World Wide Web, nodes are referred to as Web pages and vertices are links.
Pay per Click	Advertising model in which advertisers pay for click-throughs to their website. Ads are served based on keywords or themes. Also a model of website inclusion where you pay an agreed amount every time a user clicks to your website through a search engine or a search engine client. Usually you buy the position you want to rank at, for a particular keyword or keywords and pay the Search Engine every time it generates a hit to your website.
Preferential Attachment	An explanation for the power law distribution of the Internet. Preferential attachment suggests that when a new node joins a network, it is more likely to link itself to a more popular node or one that already enjoys a great popularity. Thus the formation of links is not random but selective.
Redirect	A redirect is an automatic forwarding of a domain name in order to resolve the name.
URL	Uniform Resource Locator. The URL takes the form www.url.com, i.e. www.mensbowties.com. Typically, every page in a website has a unique URL identified as extensions to the domain name. For example: www.mensbowties.com/celebrity.html.
Web business	A company or entrepreneurial start-up organization that has established an Internet presence by publishing a public website for the purposes of conducting business, or engaging in e-commerce or simply providing some level of information

	about the business itself. This is the term of preference the research project uses to describe business or organizations included in the research.
World Wide Web	The portion of the Internet invented by Tim Berners-Lee and enabled in 1991, it is the sum total of all the resources and users on the Internet that are using the Hypertext Transfer Protocol (HTTP). It is a network of Web pages containing information, linked together by hyperlinks from one page to another (Newman 2003). It is commonly thought of as the graphical interface to the Internet.

1.6.2 Definitions of research centric terms

Framework	A logical structure for classifying and organizing complex information. In research, Creswell (2003) identifies three major frameworks as quantitative, qualitative and mixed.
Deductive research	The research process begins with a hypothesis. The research design, data collection, and analysis is specific to the testing of the hypothesis (Saunders, Lewis & Thornhill 2003) .
Inductive research	The research process begins with a description of the phenomenon or event under study and moves to an explanation. That is to say the explanation is induced from the data (Ticehurst & Veal 1999).
Hermeneutics	Hermeneutics is a branch of philosophy concerned with human understanding and the interpretation of texts. Recently the concept of texts has been extended beyond written documents to include, for example, speech, performances, works of art, and even events (http://en.wikipedia.org).
Hermeneutic circle	The concept that the movement of understandings is

(or spiral)	constantly from the whole to the part and back to the whole. The task is to extend in concentric circles the unity of the understood meaning. The harmony of all the details with the whole is the criterion of correct understanding (Gadamer 1976).
Ontology	The fundamental assumptions being made about the primitive elements of reality (specifying what exists) (Parkhe 1993).
Paradigm	A basic set of philosophical positions or beliefs about the nature of the world. It provides guidelines and principles concerning the way research is conducted within the paradigm. The methods and techniques utilised would fall within the guidelines and principles (Ticehurst & Veal 1999, p. 25).
Phenomenology	Research philosophy that sees social phenomena as socially constructed. Phenomenology is particularly concerned with generating meanings and gaining insights into those phenomena (Saunders, Lewis & Thornhill 2003).
Theory	A coherent set of general propositions used as principles of explanation of the apparent relationships of certain observed phenomena (Zikmund 1991).

1.7 Conclusion

This chapter introduced the research project. It included the rationale for the project, a brief description of the methodology employed, and academic as well as managerial points of interest included in the thesis. It also included a general outline of the entire thesis, definitions of technological terms commonly employed in the thesis and definitions of research centric terms. The following chapter offers an overview of the findings of this thesis as suggested by Uncles (1998).

> *The Web has no geography, no landscape. It has no distance. It has nothing natural in it. It has few rules of behaviour and fewer lines of authority. Common sense doesn't hold there, and uncommon sense hasn't yet emerged. No wonder we're having trouble figuring out how to build businesses in this new land (Weinberger 2002, p. 8).*

CHAPTER 2 – SUMMARY FINDINGS

2.1 Introduction

Chapter 2 provides a summary of the findings of this research project. The placement of a summary of findings at this point in the thesis is based upon Uncles (1998) reasoning that most readers would prefer to have the results communicated early on. Uncles bases his reasoning on an earlier article by Ehrenberg (1982), which suggests that readers want to know research findings before learning how they were produced. Nevertheless, in order to best interpret the findings presented, it should be noted that the research employs an empirical multi-phase sequential mixed-methodology utilising both quantitative and qualitative methods. These methods are described more fully in Chapter 4.

2.2 Summary of quantitative research findings

This chapter describes the results of research activities. The research methodology ultimately employed was a multi-phase sequential mixed-methods strategy. The first phase utilised a quantitative simple random sampling technique. A sample of 1286 domain names was drawn from the .com domain. The sample was analysed using the process explained in Chapter 4. This sample gave a first view of the overall penetration of business into the Web environment. The websites were categorised using the following terms:

- Business site.
- Personal site.
- Organisation site (religious, political, governmental, educational, and etc.).
- Other site or site devoid of content or a server directory or not one of the above.
- Site not found error message.

- Anomalous site representing a business type website but one that could not be categorised using extant business model typologies (Note: because of the nature of the site, these anomalies were originally erroneously classified as "Redirect to other site").
- Access to the site denied due to login requirement or password.
- A site reporting to be Under Construction at the time visited.
- A site which classified as 'unable to categorize' due to language issues (See Table 2.1 below for a tabular listing).

The sampling process was conducted so as to optimise the probability of a representative sample. The margin of error was plus or minus 1.8 per cent at a 95 per cent confidence level. To test the sampling process, a sample of 1557 domain names was also extracted from the .net zone. Later in the project, this sample was also used to triangulate both the sampling process as well as the results from the .com zone. Data triangulation essentially uses multiple data sources (i.e. the .net and .com zones) to validate results (Tashakkori & Teddlie 2003). Note that there is a similarity between the two sets of results. One would expect the number of business sites to be slightly higher in the .com zone as that was originally intended to be the primary domain for business sites and hence the top level domain designation of COM, which was a contraction of 'commercial'. The results of the classification were as follows:

Table 2.1 Results from the sample of 'NET' zone domain names

Category	Count	Percentage
Not Found error message	490	31.47
Business site	373	23.96
Redirect to other site	291	18.69
Under construction	134	8.61
Unable to categorize –	89	5.72
Other (empty)	64	4.11
Personal site	61	3.92
Organization site	35	2.25
Access denied	20	1.28
Grand Total	**1557**	**100.00**

Table 2.2 Results from the sample of 'COM' zone domain names

Category	Count	Percentage
Business site	441	34.35
Not Found error message	303	23.60
Redirect to other site	265	20.64
Under construction	75	5.84
Unable to categorize –	73	5.69
Personal site	44	3.43
Other (empty/server directory)	36	2.80
Organization site	31	2.41
Access denied	16	1.25
Grand Total	**1284**	**100.00**

In addition to these initial findings, the research resulted in the establishment of two research databases which, with a high level of confidence, modelled the overall structure as well as the business structure of the World Wide Web to within an error margin of plus or minus 1.8 per cent.

The sites identified as 'Unable to categorise' were sites in foreign languages (primarily Asian) which the researcher was unable to label with adequate assurance. More importantly for this thesis were the sites which were identified as anomalous websites because the functions of these sites were not fully understood by the researcher. Thus, the sites were originally (erroneously) categorised as 'Redirect to other site'. During the initial analysis process the anomalous nature of the websites was recognised and so the sites were disaggregated from the main database and combined in a new database for further study which would involve qualitative research analysis.

Once the quantitative phase of the research project had been completed, the original intention of the project was to use a qualitative research method to further analyse the sites categorised as 'Business site'. However, it was decided to spend more time on the anomalies in an attempt to better understand the attributes of these websites. From a business model perspective, it appeared these websites depended upon an

associative relationship with other Web commercial entities. It was anticipated that further exploration would shed greater light on these associations. Also, the researcher hoped to gain a greater understanding of the revenue generating process or processes these anomalous Web business entities employed.

Carlile and Christensen (2005, p. 4) write that the 'discovery of an anomaly gives researchers the opportunity to revisit the foundation layers in the theory pyramid – to define and measure phenomena more precisely and less ambiguously, or to cut it into alternative categories – so that the anomaly and the prior associations of attributes and outcomes can all be explained.' The balance of this summary of findings discusses conclusions reached with regard to the anomalous sites originally categorised as 'Redirects to other sites'.

2.3 Summary of qualitative research findings

The qualitative phase of the project employed a theory-building interpretivist hermeneutic research design which is discussed in Chapter 4. The disaggregation of the anomalous findings from the quantitative phase of the study resulted in a database of 265 websites in the .com domain. The selection of these anomalous sites had the effect of constraining theoretical variation and thus limiting the qualitative research phase to sites representative of the process under investigation, as suggested by Eisenhardt (1989).

It became apparent these anomalies represented revenue generating sites, but the processes by which revenue was created were not clear. The practice of iterative and frequently concurrent analysis and interpretation of the data suggested by the hermeneutic research methodology led the researcher to gather additional data. After analysis, these sites were determined to represent newly emergent business models which have been largely unreported in the research literature. The investigation of these sites led to a deeper understanding of additional elements in the business environment which have contributed to the emergence of these new business models.

These elements include a significant shift in Web marketing practices which have been initiated by large Web search engine corporations such as Google and Overture (recently acquired by Yahoo). They have provided new entrepreneurial opportunities at the most elementary levels of the Web business venture spectrum. The most visible

evidence of this phenomenon is the increase in the number of new domain names which have been acquired in the most familiar top level domains. Every month over one million domain names are purchased in the .com zone alone (VeriSign 2005).

These newly emergent business models were identified from the anomalous websites in the original sample. These sites represented the initial efforts by entrepreneurs to monetise the acquisition of domain names. Analysis of the revenue generating processes of these sites revealed that this monetisation was accomplished by incorporating algorithmically generated content provided by the search engine marketing companies.

Like the most elementary biological entities in a food chain, these new Web businesses provide symbiotic relationships with larger entities by forwarding relevantly filtered, highly targeted Web users to the larger sites. These larger sites pay the search engine corporations for inclusion in the ad content which is programmatically generated by the search engine companies. The search engines provide portions of this revenue to the smaller websites displaying the ads. The revenue is most frequently paid on a per click basis. This means that whenever an end user clicks on the dynamically generated ads from the website, a small portion of revenue is paid to the domain name owner. In the process of doing so, these sites and the revenue generating mechanisms they employ offer some initial (albeit partial) explanation for the power law distribution of businesses on the Web.

The study reported herein finds that a fundamental shift in the economics of the Web is underway. This shift influences both the entrepreneurial opportunities and the marketing practices of the entire Web environment. This study also finds that new business models are emerging within the business environment of the Web.

The first indication of this was the identification of the anomalous websites from the quantitative phase as ‘parked domain names’. These were websites hosted by specialist service providers. The sites were specifically developed for the purpose of parking domain names, which this study refers to as ‘level two’ business models. These sites are able to generate revenue for domain name owners by establishing symbiotic relationships with advertising aggregators, and displaying these ads on the specially constructed websites. The sites’ pages are dynamically created using

programs and design algorithms provided by the hosting service provider. The domain name owners utilising this business model typically depend on direct navigation to attract end users to these sites. Thus they frequently employ a strategy which features either a very generic domain name, such as 'kitchenfurniture.com' or a common misspelling of a popular term such as 'voyuer.com' or 'plamsatv.com' or the misspelling of a common domain name as in the case of Amazon's acquisition of the domain name 'smazon.com'.

While the qualitative analysis of anomalous websites encountered in the quantitative phase of the research project provided a greater understanding of the nature, and revenue generative capacity of the new business models, it was not until the hermeneutic cycle, and its dependence on the integration of 'knowledge of the part', with 'knowledge of the whole' that understanding of the full ramifications of the emergent phenomena became clear. For example: The revenue generating capacity of the parked domain names for the domain name owner was not discernable until further research was conducted on the revenue mechanisms employed by the domain parking service provider. This led to yet another recently emergent Web entity, the search engine ad aggregator.

Yet the emergence of the search engine aggregator would not necessarily have allowed the prediction of the emergence of the level two websites. Originally these entities (such as Google) utilised the concept of ad aggregation as a resource for internal revenue generation. The idea of partnering with other entities such as the parking site service providers or the myriad of entrepreneurial affiliate (level three) sites came later (Vise & Malseed 2005).

Thus it is the interactions between the various elements of the business environment of the Web which define the business environment itself. Only when these complex interactions between business entities in the system were examined did the role of the original anomalous websites begin to emerge. These new business models which describe the newly emergent level two and level three sites are nearly entirely reliant on linkages to larger entities on the Web which actually produce products and sell them. These large sites are the 'brand names' of the Web. The smaller entities are more likely to link to these sites for two reasons. Firstly, the level two or level three

sites' clients are likely to recognise these brand name entities, and secondly, the search engine ads placed on the site and which are based on content compatibility, assure that the larger more relevant sites are the ones whose ads are most likely to be represented.

So, rather than competing with established business models the larger sites employ, the continued emergence of these simple models have a positive influence on the fortunes of the larger entities. The simple business models form a series of symbiotic relationships, not unlike those seen in the natural biosphere. These relationships are dynamic because they are often programmatically established, and the nature of the program dictates that any business entity willing or capable of bidding the most for key search words has the best chance of having their links imbedded in the simple business model sites. Thus, the emergence of these simple business models may have the effect of fuelling the financial growth of the larger business entities. This is possible because as more of these simple business models emerge there is a concomitant increase in the number of inbound links attaching to the larger sites. In other words, these very small business sites are more likely to have links to larger sites programmatically embedded within their websites as a result of search engine marketing strategies employed by both the very small Web business entity and the very large one.

The results of the analysis of inbound links of websites classified as 'business sites' in the quantitative phase of the study showed these sites represent a nearly perfect power law distribution of inbound links. This is in contrast to a random distribution of links, which would be expected to exhibit a normal Poisson distribution. The attributes of a power law distribution are these: Firstly, a very few entities that rank very high in the number of inbound links are in evidence; secondly, a larger number of entities with close to median numbers of inbound links exist; lastly, a very great number of entities with very few inbound links, referred to as the 'long tail' or a 'fat tailed' distribution are to be seen.

This is important because 'the dynamics of link accumulation – at both global and local granularities – can strongly influence competition and diversity in business and society' (Pennock et al. 2002, p. 5210). This result affirms Barabási and Albert's

(1999) assertion that the Web represents an example of a power law distribution. They suggest that a reason for this type of distribution is the phenomenon of preferential attachment.

Most networks grow by the addition of new nodes to the network. Nodes on the Web may be represented by pages or by domain names. Preferential attachment naturally leads to a power law distribution (Barabási & Albert 1999). Preferential attachment and the resulting power law distribution are a result of the self-organising capacity of the Web (Barabási & Albert 1999; Pennock et al. 2002). For example, a newly created website will be more likely to include links to well-known popular sites which already have a high number of in-bound links attaching to their sites (Barabási & Albert 1999). Barabási and Albert initially suggest that this means older sites increase their connectivity at the expense of younger ones, however, Adamic and Huberman (1999) disconfirmed this by their research which demonstrated no correlation between age of a site and the number of inbound links it has.

To the extent that power laws are a result of the phenomenon of preferential attachment, the present study confirms preferential attachment for business entities on the Web. This is true because this study establishes the validity of a power law distribution for business websites. However exactly how preferential attachment occurs remained unknown until now. It is suggested herein that the revenue generating mechanisms of these emerging business models, such as programmatically derived conjunctive links imbedded by search engine market aggregators, may have an important role to play in the preferential attachment phenomenon. This is suggested as an area for future study.

The anomalous websites originally encountered in the quantitative phase of the research represent the most simplistic embodiment of this phenomenon, and the investigation of these sites has led this thesis to a more robust understanding of the elementary principles underlying search engine marketing and its role in the business environment of the Web.

2.4 Conclusion

The purpose of this chapter was to provide a summary of the findings of this research project. This summary of findings is based upon Uncles (1998) reasoning that most

readers would prefer to have the results communicated early on. Uncles believes an appropriate structure is to start at the traditional end and write the thesis around key findings. He suggests that readers want to know research findings before learning how they were produced. The ensuing chapters are based upon the more traditional model suggested by (Perry 1998), with only moderate adaptations incorporating Uncles model.

> *One reason we see so few theoretical articles in [Information Systems] relates to the youth of the field. Another concerns the complexity of assembling a review in an interdisciplinary field. That is, constructing a review is a challenging process because we often need to draw on theories from a variety of fields. Moreover, we have so few theories of our own (Webster & Watson 2002, p. xiv)*

CHAPTER 3 – LITERATURE REVIEW AND RESEARCH PROPOSITIONS

3.1 Introduction

The purpose of this chapter is to identify research issues worth researching (Perry 1998). It is a 'critical search for an analytical framework which will allow a systematic investigation of an issue' and it should 'avoid the calamities of ignorance and the reinvention of what is already known' (Baker 2000, p. 219-220). Hart (1998) suggests additional purposes for the literature review. Among these are identifying main methodologies and research techniques, synthesising and gaining a new perspective on a topic, rationalising the significance of the problem discussed, relating ideas and theories to applications, and understanding the structure of the subject. This chapter aims to synthesise these goals.

E-business is unique in that it represents a nexus between technology and traditional business and thus cannot be fully understood without some grounding in each realm. It incorporates elements of disciplines which include information technology, information systems, business management, marketing, graphic design and software development. Valovic (2000, p. 39) observes:

> The Internet might be described as part technology, part human interaction. To describe it as one or the other is not quite accurate. Unlike other technologies, it does not do anything in the absence of the human mind – in fact, the human mind is the sole source of its viability. Accordingly, the destiny of the Net will be shaped by the interaction of two types of adaptive agents: The systems and software of the Net and its human users.

E-business and e-commerce represent relatively new business phenomena. The era of electronic business was opened with the invention of the earliest electronic computers in the 1940's. Among the earliest customers for these machines, after government, were business entities. And as computers gained power and reduced size, they became

more ubiquitous in the business environment. The Internet grew out of a coalition of military, government, and universities. Its roots are the Advanced Research Projects Administration's attempts to link computers at the Massachusetts Institute of Technology and the University of California at Berkeley networks in 1965. From that grew the ARPANet which later, in 1982, issued forth the Internet (Leiner et al. 2000).

The term e-business only came into the popular lexicon after Sir Tim Berners-Lee's development of the World Wide Web early in 1992. The Internet and in particular World Wide Web represent entirely new environments for the conduct of business. The Web presents us with a laboratory in which we may conduct business research with of a view of every stage of evolutionary business development, from solo-entrepreneurs inventing entirely new business models to large corporations attempting to adapt their enterprises to the business environment of the Web in order to thrive and prosper (Drew 2002).

There has been a tendency to focus business research on the huge successes (Rothschild 1990), and the trend appears to continue with the emphasis on the large e-commerce enterprises such as Amazon, e-Bay, Google and Yahoo! within e-business/e-commerce research. Diaz (2004, p. 50) explains this as a tendency to emphasise 'macro-phenomena and macro-processes; that is, to those that display themselves as decisive in the unleashing of historical events and political changes. On the other hand, small processes and phenomena, those that each day and over the long term contribute to political changes, are too often overlooked, as if they did not exist and did not have an effective significance.'

Rothschild (1990, p. 82) observes:

> Because of their size and influence, the great corporations – the behemoths of the economy – grab the headlines. But the overwhelming majority of organizations, and organisms, are miniscule. … Like single celled creatures, such firms are incredibly numerous and diverse, but, because of their size, they seem unimportant to casual observers.

An important and significant proportion of Web commerce lies just below the surface of our vision in an abundance of simple, miniscule business models executing small

processes heretofore overlooked in the literature. The next section reviews literature on development of business models in the Web environment.

3.2 Internet business models

The term 'business model' first came into vogue with the widespread use of personal computer and the development of spreadsheet or modelling software by Dan Bricklin, then at MIT. Business models are perhaps the most discussed and least understood aspect of the Web. Lambert (2003) concludes that until the attributes of various business models are agreed upon, useful taxonomies of business models will not be forthcoming. She also laments the lack of an agreed upon definition of the term. For example, current definitions range from 'a general vision or strategy' (Ovans 2000) to a complete business model ontology consisting of the four pillars of; product innovation; infrastructure management; customer relationship; and financial aspects as proposed by Osterwalder and Pigneur (2002). Magretta (2002, p. 87) indicates that business models are:

> … at heart, stories - stories that explain how enterprises work. A good business model answers Peter Drucker's age old questions: Who is the customer? And what does the customer value? It also answers the fundamental questions every manager must ask: How do we make money in this business? What is the underlying business logic that explains how we can deliver value to customers at an appropriate cost?

This thesis employs the succinct definition of a business model provided by Turbin (2004, p. 11) that 'a business model is the method of doing business by which a company can generate revenue to sustain itself'. Using this definition allows one to avoid the temptation to try to 'retrofit' existing business models to new business models evolving in the Web. Such new entities might not employ readily discernable revenue generation schemes. Further, established taxonomies frequently provide insufficient differentiation amongst business Web sites, since it is not uncommon for business sites to incorporate multiple business model categories.

Using revenue generation as a point of departure allows this thesis to view business models from the perspective of the most simplistic method (the "Domain Name Acquisition" model, described in this research as a Level 1 business model) and then

categorise increasingly sophisticated methods of revenue generation as a primary business model determinant. This method has the advantage of offering the researcher greater flexibility to define and differentiate new business entities at the most fundamental levels of e-business.

From a practical perspective, 1994 was the first year that e-commerce revenue generating transactions as we have come know them in the first decade of the 21st Century were conducted with any frequency (Moschovitis 1999). However, there are some reports of primitive commercial electronic transactions using the Web from as early as 1992 (Rayport & Jaworski 2004, p. 27). By 1999, e-business was termed the 'new economy' and e-commerce companies were the darlings of investors and venture capitalists. In January of 2000, the valuation of e-business stocks reached its pinnacle and by the end of 2000 an enormous shakeout of e-business ventures began with what became categorised in the popular media as the bursting of the dot-com bubble. In retrospect, in was clear that many businesses had rushed to participate in this new economy before a full understanding of the real value proposition of the Web had been conceptualised. The brief history of e-business may be summarized as one of incredible early growth and extraordinary valuations of e-business companies followed by an equally precipitous decline.

Today, many see an Internet that is still in its infancy where only a small portion of the total potential capacity is used. E-commerce now seems to be in a second phase, dubbed Web 2.0. This newer phase is one that suggests steady, sustained growth (Hof & Hamm 2002; Patrick 2001; Rayport & Jaworski 2004) as opposed to the earlier perception that e-commerce had withered during 2000 and 2001. With very little fanfare and almost no media hype, the valuation of established Web businesses now approaches or exceeds levels of valuations they reached in the year 2000 (Rappa 2005).

Some feel that e-commerce represents a dramatic paradigm shift from previous business models and thus encompass dramatic changes for business in general (Canzer 2003; Leonhardt 2000; Patrick 2001; Ray 2002). Others feel that e-commerce simply represents a new application or a new tool to be applied to traditional business practices (AACSB 2002; Argyres & McGahan 2002; Porter 2001;

Rayport 1999). In either event there is a growing consensus that e-business is now pervasive and that the term is essentially synonymous with business itself (Patrick 2001; Rayport & Jaworski 2004).

There is much discussion about how the web changes traditional business models. But there is little clear-cut evidence of exactly what this means and there is a gap in the literature with regard to the generation of new business models except for speculation that such new models should exist (Drew 2002; Rappa 2005; Timmers 1998). For example, some maintain that there is no difference between Web business models and real world models (Porter 2001). On the other hand, Canzer (2003) clearly differentiates between businesses that emerged within the Web and those that adapted their existing businesses to the Web environment. He notes 'a fundamental division' between these two business entities and that 'Firms with no history other than one defined on the Internet make their business decisions with a clear focus on the online world. They are not concerned about interfering with other established business activities'. Adam's model of Web use (2002) also differentiates online businesses from natural world businesses by classifying them along a spectrum ranging from 'pure dot com' to 'pure bricks and mortar' as shown in Figure 3.1. But neither the Canzer nor Adam model specifically identified new business models emerging from within the Web.

Figure removed due to copyright restrictions

Source: (Adam 2002, p. 263)

Figure 3.1 Adam's spectrum of business users of the Web

Rappa (2005) suggested that the Internet commerce should give rise to new kinds of business models, while both Drew (2002) and Timmers (1998) believed further research was needed to discover if new business models were actually emerging within the Web. The current study addressed this need.

Web business model typologies have been widely reported in the literature. The following three tables provide a synopsis of major typologies. Table 3.1 presents Rayport and Jaworski's (2004) view of Web business models. This is followed by Table 3.2 which lists business model typologies as described by Laudon and Traver (2007). Finally, Rappa's (2005) view is presented in Table 3.3.

One can discern a pattern of typologies emerging from the tables of business models, but the lack of consensus suggests that there is no clear agreement on exactly what constitutes a Web business model, and even less agreement on which websites best represent a prototypical business model. This is partly explained because the Web is such a new environment for business. How 'real world' business models are employed or how they will mutate to best survive in this new environment is not yet clear. It is also partly explained by the subtle complexity of business models themselves. It may be impossible to 'categorise' Web business models in all but the most generic of terms, just as it would be impossible to categorise individuals in all but general terms. Businesses reflect an infinite variety of structure, process, innovation and technology. For example, business models can range along a very wide spectrum from extremely simple to extremely complex. The questions to be asked are "What would an extremely simple business model look like?"

Table 3.1 Rayport and Jaworski business models

Business Model	Revenue Source	Examples
Metamarket Switchboard	• Transaction fees • Product sales • Advertising • Affiliate deals	• Babycenter.com • CarPoint.com • FashionMall.com
Auction	• Transaction fees	• Ebay • Amazon auctions • Covisint

Freshest Information	• Product and service sales • Subscription • Advertising	• NYTimes.com • Salon.com • Business 2.0 • Zagat.com
Highest Quality	• Product, service and information sales • Advertising	• FAO.com • Ashford.com
Widest Assortment	• Product sales	• Amazon.com • CDNow • Sephora.com • EBags.com
Lowest Prices	• Product sales	• Lowest-fare.com • Allbooksforless.com • Outpost.com
Most Personalized	• Product sales • Subscriptions	• Reflect.com • Sevencycles.com • EDiets.com

Source: (Rayport & Jaworski 2004, p. 132)

Table 3.2 Laudon and Traver business models

Business Model	Revenue Source	Examples
Advertising	• Fees from advertisements in exchange for advertisements	• Yahoo.com
Subscription	• Fees from subscriptions in exchange fro access to content or services	• WSJ.com • ConsumerReports.com
Transaction Fee	• Fees commissioned for enabling or executing a transaction	• eBay.com • Etrade.com
Sales	• Sales of goods, information or services	• Amazon.com • LLBean.com • Gap.com • JCPenny.com
Affiliate	• Fees for business referrals	• MyPoints.com

Source: (Laudon & Traver 2007, p. 62)

Table 3.3 Rappa business models

Business Model	Revenue Source	Examples
Brokerage	• Brokers are market-makers: they bring buyers and sellers together and facilitate transactions. Brokers play a frequent role in business-to-business (B2B), business-to-consumer (B2C), or consumer-to-consumer (C2C) markets. Usually a broker charges a fee or commission for each transaction it enables. The formula for fees can vary.	• Orbitz.com • Priceline.com • CarsDirect.com • PayPal.com
Advertising	• The web advertising model is an extension of the traditional media broadcast model. The	• Yahoo.com • Monster.com • Claria.com

Business Model	Revenue Source	Examples
	broadcaster, in this case, a website, provides content (usually, but not necessarily, for free) and services (like email, IM, blogs) mixed with advertising messages in the form of banner ads. The banner ads may be the major or sole source of revenue for the broadcaster. The broadcaster may be a content creator or a distributor of content created elsewhere. The advertising model works best when the volume of viewer traffic is large or highly specialized.	• Google.com • NYTimes digital
Infomediary	• Data about consumers and their consumption habits are valuable. Independently collected data about producers and their products are useful to consumers when considering a purchase. Some firms function as infomediaries (information intermediaries) assisting buyers and/or sellers understand a given market	• Doubleclick • Nielson Netratings • Coolsavings.com • Edmunds
Merchant	• Wholesalers and retailers of goods and services. Sales may be made based on list prices or through auction	• Amazon.com • LandsEnd.com • Barnes & Noble • Apple iTunes
Manufacturer	• The manufacturer or "direct model", it is predicated on the power of the web to allow a manufacturer (i.e., a company that creates a product or service) to reach buyers directly and thereby compress the distribution channel. The manufacturer model can be based on efficiency, improved customer service, and a better understanding of customer preferences	• Dell • Bmwfilms •

Business Model	Revenue Source	Examples
Affiliate	• In contrast to the generalized portal, which seeks to drive a high volume of traffic to one site, the affiliate model, provides purchase opportunities wherever people may be surfing. It does this by offering financial incentives (in the form of a percentage of revenue) to affiliated partner sites. The affiliates provide purchase-point click-through to the merchant. It is a pay-for-performance model—if an affiliate does not generate sales, it represents no cost to the merchant. The affiliate model is inherently well-suited to the web, which explains its popularity. Variations include, banner exchange, pay-per-click, and revenue sharing programs	• Amazon.com • Barnes & Noble
Community	• The viability of the community model is based on user loyalty. Users have a high investment in both time and emotion. Revenue can be based on the sale of ancillary products and services or voluntary contributions; or revenue may be tied to contextual advertising and subscriptions for premium services. The Internet is inherently suited to community business models and today this is one of the more fertile areas of development, as seen in rise of social networking.	• Red Hat • Wikipedia • WCPE.org • Flickr.com
Subscription	• Users are charged a periodic—daily, monthly or annual—fee to subscribe to a service. It is not uncommon for sites to combine free content with "premium" (i.e., subscriber- or member-only) content. Subscription fees	• Listen.com • Netflix.com • Classmates • AOL

Business Model	Revenue Source	Examples
	are incurred irrespective of actual usage rates. Subscription and advertising models are frequently combined	
Utility	• The utility or "on-demand" model is based on metering usage, or a "pay as you go" approach. Metered services are based on actual usage rates. Traditionally, metering has been used for essential services (e.g., electricity water, long-distance telephone services). Internet service providers (ISPs) in some parts of the world operate as utilities, charging customers for connection minutes, as opposed to the subscriber model common in the U.S	• IBM • Slashdot.com

Source: (Rappa 2005)

In order to understand the importance of elemental business models, it is necessary to understand more about the context of the Web itself. The literature abounds with descriptions of the Web as a complex and a chaotic system (Barabási 2002; Barabási & Albert 1999; Buchanan 2002; Dorogovtsev & Mendes 2003; Holland 1998; Johnson 2001; Morowitz 2002; Newman 2003; Schrage 2002; Waldrop 1992; Weinberger 2002; Yook, Jeong & Barabasi 2002). Since a particular aim of the study reported in this thesis is to model business presence in the Web environment, it is appropriate to incorporate a discussion of complex systems in order to better understand the business environment of the Web.

Describing the Web as a complex system provides a framework from which to enhance our understanding of the ways in which we might expect business to evolve. It implies that the Web must exhibit the characteristics which researchers now attribute to all complex systems. For example, complex systems are adaptive, they are dynamic and constantly changing, new entities emerge from complex systems, and such emergence is frequently presaged by a phase transition process. Each of these attributes will be discussed in greater detail in the following sections.

3.3 Complexity theory

Complexity theory is new. It is so new that, like the Web itself, its scientific boundaries are difficult to define (Byrne 1998; Manson 2001; Waldrop 1992). A frequently occurring critique is that complexity theory essentially defines everything and thus defines nothing (Vicsek 2002). Complexity theory's roots are in the natural sciences, most notably, biology and later physics. However, in recent years, theories and applications of complexity have grown exponentially, encompassing the social sciences as well (Manson 2001; Proctor & Larson 2005).

Despite its relative freshness, complexity theory's credentials as multidisciplinary have been firmly established (Berners-Lee 2004; Byrne 1998; Holland 1998; Keller 2005; Levin 2005; Manson 2001; Proctor & Larson 2005; Waldrop 1992). The value of complexity theory lies in its ability to help explain and understand change (Manson 2001), and there are few human endeavours as fraught with change as the conduct of business on the Internet. Thus, complexity theory represents an appropriate framework for research which attempts to explain the business evolution occurring on the Internet.

The term 'complexity' comes from the Latin root 'complexus' or braided together (Gell-Mann 1998). Complexity is a term that is often content dependant, which is to say that complexity rests at least to some extent within the eye of the beholder. In describing the concept of complexity, Gell-Man writes 'But that is just what we mean when we say that the grammar of a certain language is complex, or that a certain conglomerate corporation is a complex organization, or that the plot of a novel is very complex—we mean that the description of the regularities [of the complex object] takes a long time' (Gell-Mann 1998, http://www.ndu.edu/inss).

Complex systems are spontaneously self-organising entities and so they are frequently referred to as complex *adaptive* systems. Indeed, the idea of complex systems may have sprung from the concept of self organization. The modern usage of the term 'Self-organizing systems' dates back to WWII (Keller 2005) when the term was first applied to inanimate objects, but Keller also noted that elements of the term dated as far back as the eighteenth century. For example, Immanuel Kant's application of the term 'self-organized' to living things (p. 1070).

Self-organisation is the property that allows complex systems to change themselves. They change their internal structures in order to better interact with the environment (Manson 2001). The individual entities which make up complex systems propel these changes because they too adapt in order better interact with each other, either in a competitive fashion or in a cooperative fashion.

Complexity theory provides a more dynamic model for understanding the way the world really functions. For example, it helps explain unique historical contingencies such as why one path of development emerges in a business, instead of another (Bollier 1993). A complex adaptive system is one whose properties are not fully explained by an understanding of its component parts. This is because the component parts are often very simple entities. Complex systems consist of a large number of mutually interacting and interwoven parts, entities or agents (Gell-Mann 1998). The complexity of the system results from the interactions of these simple components. Examples are ant colonies, neural networks, the Internet and the Web. These are systems in which the behaviour of the whole is much more complex than the behaviour of the parts (Holland 1998).

It was during the 1990s that the flowering of complexity theory as a scientific discipline reached full blossom. Proctor and Larson (2005, p. 1065) write that this is because of the wide 'applicability of complexity to current research and policy questions.' Technology represents an appropriate setting for application of complexity theory (Berners-Lee 2004; Holland 1998; Manson 2001; Newman 2003; Waldrop 1992). Barabási (2002, p. 149-150) suggests an example of the application of complexity theory and complex adaptive systems to the Internet:

> While entirely of human design, the Internet now lives a life of its own. It has all the characteristics of a complex evolving system, making it more similar to a cell than to a computer chip… Therefore Internet researchers are increasingly morphing from designers into explorers.

Further evidence of the complex nature of technological systems such as the Web, as well as the need for further research in the field is provided by Dorogovtsev and Mendes (2003, p. 2):

> By the middle 1990's, the impact of large growing communications nets with complex architectures, the Internet and the World Wide Web, on our civilization became incredible. However, understanding of their global organization and functioning was absent.

John H. Holland is a Professor at the University of Michigan and a Fellow of the Santa Fe Institute, the USA's think tank for research in complexity theory and complex adaptive systems. He has written extensively on complexity in information systems and in society. Waldrop (1992, p. 145) discusses Holland's views of the aspects of complex adaptive systems:

> First, [Holland] said, each of these systems is a network of many "agents" acting in parallel. … If you were looking at business cycles, the agents might be firms. And if you were looking at international trade, the agents might even be whole nations. But regardless of how you define them, each agent finds itself in an environment produced by its interactions with the other agents in the system. It is constantly acting and reacting to what other agents are doing. And because of that nothing in the environment is fixed.
>
> Furthermore, said Holland, the control of a complex adaptive system tends to be highly dispersed. … If there is to be any coherent behaviour in the system, it had to arise from the competition and cooperation of the agents themselves. …
>
> Second, said Holland, a complex adaptive system has many levels organization, with agents at any one level serving as the building blocks for agents at a higher level. …
>
> Finally, said Holland, complex adaptive systems typically have many niches, each one of which can be exploited by an agent adapted to fill that niche. … Moreover the very act of filling one niche opens up more niches – for more parasites, for new predators and prey, for new symbiotic partners. … In short, complex adaptive systems are characterized by perpetual novelty.

Holland's credentials in computer science lend credence to the application of complexity theory to the business environment of the Web.

3.3.1 Emergence process and complex systems

A key component Complexity theory is the phenomenon of 'emergence'. The term 'emergence' as related to complex adaptive systems refers to the origination of novel

entities 'as in the emergence of a plant out of a seed and the emergence of a visual pattern from the juxtaposition of the tiles in a mosaic' (Bunge 2003, p. 3). In the domain of emergence, the assumption is made that both actual systems as well as models operate by selections from the immense space and variability of the world of the possible. And in carrying out this selection, new and unanticipated properties emerge (Morowitz 2002). Waldrop (1992, p. 88) describes the process of emergent behaviour as beginning with a series of agents:

> These agents might be molecules or neurons or species or consumers or even corporations. But whatever their nature, the agents were constantly organizing and reorganizing themselves into larger structures through the clash of mutual accommodation and mutual rivalry. Thus molecules would form cells, neurons form brains, species would form ecosystems, consumers and corporations would form economies and so on. At each level, new emergent structures would form and engage in new emergent behaviours. Complexity, in other words, was really a science of emergence.

Sir Tim Berners-Lee, the inventor of the World Wide Web as well as the inventor of the lingua franca of the Web, the Hyper Text Mark-up Language (HTML), applied the emergence process directly to the Web in the Keynote speech at the 2004 'Emerging Technologies Conference' at the Massachusetts Institute of Technology:

> Emergence is the process of deriving some new and coherent structures, patterns and properties in a complex system. Emergent phenomena occur due to the interactions between the elements of a system over time. Emergent phenomena are often unexpected, non-trivial results of relatively simple interactions or relatively simple components. What distinguishes a complex system from a merely complicated one is that some behaviours and patterns emerge in complex systems as a result of the patterns of relationship between the elements.

Complex systems consist of a series of simple entities, and the interaction of the entities is responsible for the emergence of newer entities. Holland (1998) believes that in order to fully understand complex systems, we must first comprehend the emergent phenomena that attend them.

There is a persistent pattern to the emergence process. The combinations of previous levels of emergence impact on what emerges at the next level. Thus there is a form of interlocking hierarchical quality to the emergent entities which Holland notes is 'one of the central features of the scientific endeavour' (1998, p. 7-8). The evolution of the conduct of business on the Web emerged from the Web, which in turn emerged from the Internet, which evolved from the early experiments in network architecture undertaken by the Advanced Research Projects Administration which evolved into the ARPANet. In this thesis, the application of the term emergence to the development of new business models provides the most precise descriptor of the phenomena.

3.3.2 Phase transition in complex systems

Phase transition is a term with origins in physics. Examples of phase transitions are the transitions from liquid state to a solid state, for example, when water freezes; or from a liquid state to a gaseous state, when water boils and turns to steam. These represent examples of the most common types of phase transitions, called first-order transitions. Second-order phase transitions are less common in nature. 'They are much less abrupt, largely because the molecules don't have to make that abrupt change from ice to water (Waldrop 1992, p. 229).

A similar phenomenon also occurs in social systems. For example, in describing Christopher Langton's work in artificial intelligence software development, Waldrop (1992, p. 232) writes:

> Phase transitions, complexity and computation were all wrapped together, Langton realized. Or, at least, they were in the von Neuman universe. But Langton was convinced that the connections held true in the real world as well – in everything from social systems to economies to living cells.

Thus, the use of the term "phase transition" in this thesis to describe a specific phenomenon relative to emergence of entities from a sociological or economic perspective is neither novel nor unique. In science (and for the purposes of this thesis) phase transition relates to a specific event. In science the application of heat represents an environmental element that effects the phase transition of water to gas. Barabási (2002) believes phase transitions imply a transition from disorder to order and that there is a relationship between power law distribution (discussed below) and

phase transition phenomena in that power laws also represent a transition to a more orderly environment. Further discussion of phase transition takes place in Chapter 6.

3.3.3 Summary of elements of complex systems

In summary, the elements of complex systems are these:

- They are dynamic in the sense they are constantly changing.
- They are adaptive, which is to say they evolve to benefit themselves and to insure their survival.
- The adaptations are controlled to some extent by the interactions of the agents that comprise the system.
- The control is typically highly dispersed.
- They exhibit many levels of organisation.
- They are comprised of many niches.
- They exhibit emergent mechanisms.
- They are self-organising.
- New elements or entities emerge from complex systems.
- Phase transition is an element of emergence within complex systems.
- The emergent elements are not necessarily predictable from analysis of the individual parts of the system.
- Complex systems are defined by relationships between components more than by describing its constituent parts.

The review of literature on complexity theory and the phenomenon of emergence to this point lead to the statement of the first propositions this thesis examines:

P_1: Complexity theory correctly predicts the emergence of new and coherent structures, patterns and properties (e.g., online business models) in a complex system such as the Web.

P_2: Complexity theory correctly predicts that emergent elements (e.g., Web business models) are not necessarily predictable from analysis of an individual part of the system (e.g., a search engine output).

P_3: As a complex system, the Web is defined by relationships between component websites.

Valovic (2000, p. 39) incorporates the Internet into Complexity theory and suggests some direction for research:

> If the Internet is indeed viewed as a complex adaptive system, then the ways in which it might grow, adapt, and evolve are certainly fascinating areas for exploration. Initially, the more salient and recognizable transformations will likely be centered on the areas of scientific research and business development (although these areas are converging somewhat).

Manson (2001, p. 405) identifies three categories of complexity theory in order to begin to set boundaries. He calls these categories 'Algorithmic complexity', 'Deterministic complexity', and 'Aggregate complexity'. He defines algorithmic complexity as dealing primarily with the field of mathematics and deterministic complexity as dealing with the realms of chaos and catastrophe theories. His discussion of what he calls aggregate theory relates to this study in that it deals with systems of linked components and the focal point of aggregate theory lies in its relationships between components. Manson notes that 'Understanding and tracing the relationships of a single entity is difficult, while tracing them in an entire system is impossible' (2001, p. 409). Manson's characterisation of aggregate complexity as dealing with linked components leads naturally into the discussion of network theory.

3.4 Network theory

We live in a world of networks… physical networks, biological networks, communication networks, transportation networks and business networks. Networks impact our daily lives in a myriad of ways. Both the Internet and the World Wide Web are complex networks. This is important to restate, because so much of what is known about the behaviour of networks has changed as a result of the intense study of attributes of the networks dating from the late 1990's. Until very recently we have imperfectly understood the networks that surround us.

The study of networks is a subset of complexity theory. In an abstract form, the study of networks is also a study of organisation and organising principals (Buchanan

2002). Knowledge of networks went through a stunning transition in the late 1990s. Dorogotsev and Mendes (2003, p. 1) refer to this transition as a 'revolution in network science'. Much of what the world knows about networks today is new information gleaned to a very large extent from research conducted on the Internet and the World Wide Web.

Marc Newman (2003, p. 2) describes a network as 'A set of items called vertices or sometimes nodes with connections between them called edges.' In the case of a network such as the Web, the nodes may be considered websites and the edges or vertices are referred to as links. Thus a website may consist of one or more pages, each page of which may link to another page within the same site (internal links) or to another site on the Web (external links).

Barabási and Albert's (1999) first paper on the attributes and distribution of Web links resulted in a new understanding of Networks in general, but because their research focused on network attributes of the World Wide Web, they moved the theoretical realm of network theory and complexity much closer to the domain of e-commerce and e-business. They wrote, 'But equally complex networks occur in social science, … or in the World Wide Web (WWW), whose vertices are HTML documents connected by links pointing from one page to another. Because of their large size and the complexity of their interactions, the topology of these networks is largely unknown' (Barabási & Albert 1999, p. 510).

Their research indicated that there were fundamental attributes of most networks, including the Internet and the Web, in that they:

- Exhibited forms of preferential attachment.
- Exhibited a power law distribution.
- Exhibited rapid and/or consistent growth.

It should be noted that these attributes are interrelated. The power law distribution of the Web is a consequence of preferential attachments, and preferential attachment is a consequence of growth. Both of these later elements, growth and preferential attachment are needed for a power law distribution to form (Barabási & Albert 1999). Unrelenting growth has been a primary feature of the Internet (Barabási & Albert

1999; Berners-Lee & Fischetti 1999; Brin & Page 1998; Cerf 2002; Colecchia 2000; Edelman 2003; Featherstone & Ellis 2005; Gerbert & Birch 2001; Laudon & Traver 2007). Barabási (2002, p. 86) describes the phenomenon of preferential attachment by writing that:

> In real networks linking is never random. Instead, popularity is attractive. Web pages with more links are more likely to be linked to again, highly connected actors are more often considered for new roles, highly cited papers are more likely to be cited again… Network evolution is governed by the subtle yet unforgiving law of preferential attachment. Guided by it, we unconsciously add links at a higher rate to those nodes that are already heavily linked.

This phenomenon ultimately leads to a network configuration where a few very powerful, extraordinarily well connected sites do the vast majority of business, hence the phenomenon is also characterised as 'rich get richer' or 'popularity is attractive' (Barabási 2002; Barabási & Albert 1999; Dorogovtsev & Mendes 2003).

Power laws are also indicators that complex systems are moving toward order. Nature normally hates power laws. Barabási (2002, p. 77) writes that in ordinary systems, all quantities follow bell curves:

> But all that changes if the system is forced to undergo a phase transition. Then power laws emerge – nature's unmistakable sign that chaos is departing in favour of order. The theory of phase transitions told us loud and clear that the road from disorder to order is maintained by the powerful forces of self-organization and is paved by power laws. It told us that power laws are not just another way of characterizing a system's behaviour. They are the patent signatures of self-organization in complex systems.

This is precisely what has occurred on the Web (Nielson 1997) , where a few very large sites, for example, e-Bay, Amazon or Google have, by far, the largest number of sites linking to them. The ratio of inbound links to these large Web business entities becomes so disproportionate, that a power law configuration results. Figures 3.2 and 3.3 present graphic representations of a power law distribution.

Power laws as related to websites may be verbally represented as:

- a very few entities that rank very high in the number of inbound links;

- a larger number of entities with close to median numbers of inbound links;
- a great number of entities with very few inbound links.

Power law distributions are often presented as logarithmic scales (Figure 3.3) because, as can be seen in Figure 3.2, a linear representation tends to hug the axis on a linear scale (Nielson 1997). The third item in the list above represents what has become known as 'the long tail' (Anderson 2004) or more accurately a wide tail distribution.

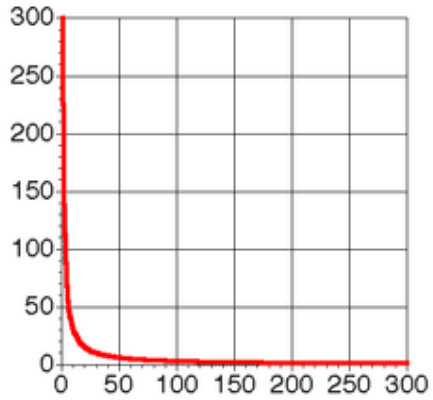


Figure 3.2 Linear representation of a power law distribution Source: (Nielson 1997)

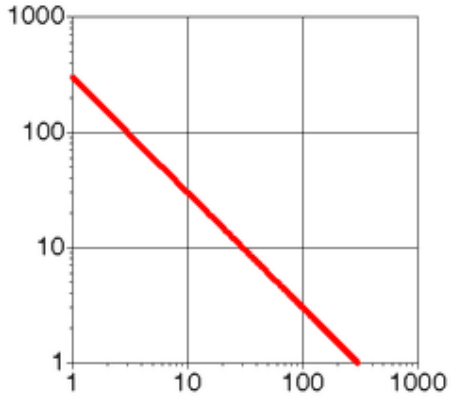


Figure 3.3 Logarithmic representation of a power law distribution

Source: (Nielson 1997)

From the perspective of business on the Web, Figure 3.2 could also be said to represent a scale ranging from a relatively few very complex business models on the left side of the scale to a huge number of simple business models on the right side of

the scale. What we now know about the conduct of business on the Web has been largely gleaned from research studies involving websites on the extreme left side of that scale (see for example Diaz 2004).

Barabási (2002, p. 85-88) writes that the Web's power law distribution is a result of a phenomenon referred to as 'preferential attachment':

> The bottom line is that when deciding where to link on the Web, we follow preferential attachment: When choosing between two pages, one with twice as many links as the other, about twice as many people link to the more connected page. While our individual choices are highly unpredictable, as a group we follow strict patterns.
>
> … Because new nodes prefer to link to the more connected nodes, early nodes with more links will be selected more often and will grow faster than their younger and less connected peers.
>
> … Thus preferential attachment induces a rich-get-richer phenomenon that helps more connected nodes grab a disproportionately large number of links at the expense of the late comers.

3.4.1 Summary of network attributes

The general principles which govern the self-organising processes of networks are that they exhibit growth fuelled by preferential attachment and this results in a power law distribution of nodes. The review of literature covering network theory and the phenomenon of preferential attachment leads to the statement of further propositions this study examines:

P_4: Network theory successfully predicts a power law distribution for business entities on the Web

P_5: Network theory correctly predicts that the emergent entities of the Web exhibit preferential attachment.

3.5 Conclusion

This chapter discussed complexity theory, emergence phenomena, and their relationships to new discoveries in network theory. These theoretical frameworks

have been used help explain the abundance of new entrepreneurial business models emerging in the environment of the Web.

The study of complex systems, emergence, and their relationship to network theory is clearly a recent endeavour, but one which must be considered extremely relevant today. Marc Buchanan wrote, 'The study of emergence in all its forms is one of the most important scientific enterprises of our era, and will remain that way for the next century' (2002, p. 207).

The emergence and the abundance of new business models affirm complexity and network theory as valid frameworks for the study of business on the Internet. These frameworks not only help to understand why there is such an abundance of innovative business models but also help explain the rapid growth of business on the Web. As well, they shed light on the difficulties inherent in identifying these business models using traditional business frameworks. As the next chapter describes, the findings in this study support both the expectations of complexity theory as outlined by Newman (2003) and the power law distribution reported by Barabási and Albert (1999).

> *Led by a new paradigm, scientists adopt new instruments and look in new places. Even more important, during revolutions scientists see new and different things when looking with familiar instruments in places they have looked before. It is rather as if the professional community had been suddenly transported to another planet where familiar objects are seen in a different light and are joined by unfamiliar ones as well (Kuhn 1962, p. 111).*

CHAPTER 4 – RESEARCH METHODOLOGY

4.1 Introduction

This chapter sets forth the research aims of the study which were introduced in Chapter 1. Discussion then turns to the framework selected to examine these issues, including a discussion of the research paradigm, and research methodology. Because the study uses a mixed methods methodology, the discussion and justification of the individual quantitative and qualitative methods employed are to be found in the sections that follow. Section 4.4 details the quantitative method selected and the process employed, while Section 4.5 details the qualitative method selected and its processes.

There is a clear recognition in the literature, not only of the increasing value of mixed methods as a methodology, but also of its escalating stature and increasing acceptance amongst researchers (Creswell 2003; Klein & Myers 1999; Marshall, Kelder & Perry 2005; Perry 2002; Reichertz 2005; Tashakkori & Teddlie 2003; Walsham 1995).

Uma Sekaran (1992, p. 93) defined a seven step research process and framework. Within this seven step process document, she identifies six basic decisions to be considered when determining the appropriate design methodology. Determining the purpose of the study is the initial decision. This study's purposes are defined as 'descriptive.' The goal of the thesis is to identify and describe new business models evolving in the business environment of the Web. In order to accomplish this goal, the researcher faced the problem of the describing the types and categories of sites populating the Web.

Theory building occurs in two stages, the descriptive stage and the normative stage. Early research in a field is necessarily descriptive. For example, early management studies such as Chester Barnard's 'The Functions of the Executive' (1947) were primarily descriptive and are considered classics today. 'Without insightful

description to subsequently build against, researchers can find themselves optimizing misleading concepts' (Carlile & Christensen 2005, p. 2).

The three steps of descriptive research are observation, categorisation and association. This triune system is ubiquitous and it is a process repeated in numerous research methodologies. It is also a method frequently employed as a business process, particularly in the consulting field. There is an iterative process involved. When a new level of understanding is reached, the process is re-iterated. While the labels may change, the process remains essentially the same. Gummesson (2003) refers to the processes as pre-understanding, understanding and explanation with the explanation level of one iteration serving as the pre-understanding phase for the next. Action Research is an interactive, participatory method in which the researcher is immersed in the subject of the study, and frequently acts as a change agent, but the core elements are firstly, planning, secondly action, and thirdly observation and reflection (Carson et al. 2001).

Miles and Huberman (1994) view the qualitative analysis process as one of data reduction (selecting, focusing, simplifying and abstracting), data display (organising and compressing the data) and finally conclusion drawing and verification. They note these three streams are interwoven, which is to say that they may be done iteratively or even concurrently.

Sekaran (1992, p. 97) noted "The goal of a descriptive study, hence is to describe relevant aspects of the phenomena of interest to the researcher from an individual, organizational, industry or other perspective. William Trochim (2002 http://trochim.human.cornell.edu/kb/index.htm) puts it even more succinctly. "When a study is designed primarily to describe what is going on it is designated as descriptive research".

Drew (2002, p. 21) wrote that 'Since the growth of businesses that use the Internet as a core infrastructure is a very recent phenomenon, it is hardly surprising that much investigation in the literature to date is of an exploratory nature'. Exploring new environments is fraught with risks. Firstly, the researcher may not recognize new entities or, in the case of this study, new business models. Secondly, the vocabulary to define these new business models may not exist and the researcher is left with the

option of describing entirely new paradigms by metaphoric references to models which are better understood.

4.2 Research objective

The objective of this research was to obtain a better understanding of the conduct of business on the World Wide Web. This led the researcher to consider the following research question: Are new business models emerging within the business environment of the Web? After a review of extant literature, the research question was rephrased as follows:

RQ_1:	Do complexity theory and network theory correctly and adequately predict business development on the Web?

The following propositions emerged from this question.

P_1:	Complexity theory correctly predicts the emergence of new and coherent structures, patterns and properties (e.g., online business models) in a complex system such as the Web.
P_2:	Complexity theory correctly predicts that emergent elements (e.g., Web business models) are not necessarily predictable from analysis of an individual part of the system (e.g., a search engine output).
P_3:	As a complex system, the Web is defined by relationships between component websites.
P_4:	Network theory successfully predicts a power law distribution for business entities on the Web.
P_5:	Network theory correctly predicts that the components of the Web exhibit preferential attachment.
P_6:	New forms of business websites that commenced in the online environment fall into identifiable classifications in each of the 'pure dot com' and 'almost pure dot com' categories. This extends Adam & Deans (2000).

The balance of this chapter discusses the elements of the research framework, as well as the research paradigm and research methodologies which were employed to complete the project.

4.3 Research framework

Creswell (2003, p. 3) identifies three major elements that comprise a research framework:

- Philosophical assumptions about what actually constitutes knowledge claims.
- General procedures of research called 'strategies of inquiry'.
- Detailed procedures of data collection, analysis, and writing called 'methods'.

The next section (4.3.1) deals with the first of the elements, since Creswell's 'knowledge claims' generally correspond with research paradigms such as post-positivism, and phenomenology. Creswell's 'Strategies of inquiries' are qualitative, quantitative and mixed method approaches. These are discussed in section 4.3.2 below. The third element in Creswell's model framework is to be found in sections 4.4 and 4.6.

4.3.1 Justification for research paradigm

A research paradigm is a philosophical belief that provides guidelines and principles as to which of the various methods and techniques are incorporated into a research study (Saunders, Lewis & Thornhill 2003). A paradigm has been described as a lens through which one views the world (Lynch 2005).

Ticehurst and Veal (1999) identify the Positivist and the Critical Interpretive paradigms as two ends of a spectrum of research philosophies. Tashakkori and Teddlie (2003) classify these paradigms as positivism and constructivism. James (2005) terms these two main paradigms positivistic and phenomenological and lists several features associated with each. For example, he describes the positivist paradigm as tending to produce quantitative data, utilise large samples, be concerned primarily with hypothesis testing, feature data that is highly specific and precise, and to attempt to generalise from a sample to a population.

He also describes the phenomenological paradigm as tending to produce qualitative data, use small samples, be concerned with generating theory, feature data that is rich and subjective, and to generalise from one setting to another.

In addition, positivist (or post-positivist) paradigms imply a deductive approach, while the phenomenological paradigm implies an inductive approach.

The distinction between quantitatively oriented and qualitatively oriented paradigms grew stronger during the decade of the 1980's. Amongst social scientists qualitatively oriented paradigms have become dominant. But so great had the schism become between quantitative advocates and qualitative adherents that Tashakkori and Teddlie (2003, p. 6) refer to this as the era of "paradigm wars".

Most research today falls along a continuum of these paradigms (Creswell 2003; Guba & Lincoln 1994; Gummesson 2003; Saunders, Lewis & Thornhill 2003; Ticehurst & Veal 1999). Miles and Huberman (1994, p. 5) state 'The paradigms for conducting social research seem to shifting beneath our feet, and an increasing number of researchers now see the world with more pragmatic, ecumenical eyes'.

In fact, Cresswell (2003) identifies pragmatism as one of the knowledge claims or paradigms. Pragmatism, as a research paradigm suggests a concern with applications, an emphasis on the research problem and employing methodologies that work to find solutions, a lack of focus on theoretical origins and an emphasis on theoretical designations. It rejects traditional dualisms such as realism vs. anti-realism and subjectivism vs. objectivism, but endorses more moderate and common sense versions of philosophical dualism based upon how well they work in solving problems. Pragmatism endorses a strong and practical empiricism as the path to determine what works. Its tenets may be summarised as the primacy of the question over the sanctity of the methodology (Creswell 2003; Johnson & Onwuegbuzi 2004; Tashakkori & Teddlie 2003). Pragmatism as a research paradigm implies a pluralistic approach to derive knowledge about a problem, the determination of practical relevance is developed in dialogue with stakeholders, it is non-foundational and its usefulness is an indication of a theory or proposition put forth (Marshall, Kelder & Perry 2005). Rescher (1977, p. 3) wrote:

> It is clear that, with particular regard to methodology at any rate, the pragmatists were surely right: There can be no better or more natural way of

> justifying a method than by establishing that "it works" with respect to the specific appointed tasks that are in view for it.

Pragmatism had it origins in the USA around the turn of the twentieth century. In the present study, it became apparent during the data collection phase that both quantitative and qualitative methods would need to be employed to complete the project. Further research made it apparent that the "Pragmatism Paradigm" best fit the philosophies this thesis sought to represent. Johnson and Onwuegbuzi (2004, p. 16) note that 'pragmatism is the philosophical partner for mixed methods research'. This pointed the researcher into the direction of mixed methods as a research methodology. This is discussed more fully in the next section.

4.3.2 Justification for the research methodology

This study employed an *empirical multi-phase sequential mixed-methods design*. Two issues concern the researcher with regard to the selection of mixed methods. Firstly; is this an appropriate methodology for early career researchers? Secondly; is this a high risk methodology?

With regard to the first issue, mixed methods permit the inclusion of additional methodologies, if the researcher deems such a move appropriate. As the research progressed, the question of appropriateness was always a concern. The pragmatic property of the mixed methods research approach allows greater freedom and experimentation with appropriate methods, and for that reason it is a good choice for the researcher who typically lacks familiarity with a foundational theory building method.

With regard to the second, the employment of such well documented methods as sampling and hermeneutics minimised the risk.

Deconstructing the phrase '*empirical multi-phase sequential mixed-methods design*' is instructive in terms of justification of the methodology and informing the reader of the processes embodied in the method (see Table 4.1).

Mixed methods were justified in this study for the following reasons. The first phase of the research found anomalous data. It became apparent that a second phase would be required to further explore the anomalies

Table 4.1 Deconstruction of the phrase describing the research methodology

Term	Description
Empirical	Activity that uses direct or indirect observation as its test of reality. This study was an empirical research project. Data were collected and analysed specifically for this project.
Multi-phase	The project had two distinct phases. A quantitative data collection and data analysis phase and a second qualitative data selection and data analysis phase. During the quantitative data analysis phase, anomalous data items were discovered. These anomalous data were disaggregated to form the judgment sample data set for the second phase qualitative analysis.
Sequential	The two phases of the project were time discreet. Phase two could not be started until phase one had been completed.
Mixed Methods	Mixed methods research is defined as research in which the investigator collects and analyses data, integrates the findings, and draws inferences using both qualitative and quantitative approaches or methods in a single study or program of inquiry (adapted from Creswell 2003; Johnson & Onwuegbuzi 2004; Tashakkori & Teddlie 2003).

In this situation, Cresswell (2003, p. 215) recommends a 'mixed method sequential explanatory strategy':

> The purpose of the sequential explanatory design typically is to use qualitative results to assist in explaining and interpreting the findings of a primarily quantitative study. It may be especially useful when unexpected results arise for a quantitative study [Morse 1991]. In this case, the qualitative data collection that follows can be used to examine these results in more detail.

A drawback to sequential mixed method research designs is that they are more time consuming since the researcher uses at least two discreet analysis methods over at least two discreet time phases. However, an advantage to mixed methods is that this methodology may provide greater internal validity to the research as it implies a manner of self triangulation. Indeed, several writers suggest that the escalating acceptance of mixed methods amongst the research community was at least partially engendered by its increasing use as a tool for internal validation of research projects (Creswell 2003; Johnson & Onwuegbuzi 2004; Tashakkori & Teddlie 2003).

Thus, two distinct data collection and analysis methods were used. The initial quantitative phase used the simple random sampling technique to generate a sample of Web domain names which could be generalised to the larger population of domain names, in this case the .com domain of the Web. This sample was intended to complete a descriptive study of business categories on the Web. It was during quantitative analysis that anomalous data were discovered. These were websites that proved difficult to categorise immediately, other than to note that the websites clearly had some revenue generating function.

In effect, the aim of the first phase of the research was to extract a random sample of websites using a quantitative sampling methodology. It was thought that this sample would result in a randomly drawn database of website domain names, analysis of which would enable the development of a descriptive model of the relative penetration of business sites within the Web. Further, it was thought that it would be possible to broadly categorise sites in the sample as business, non-business, or unknown. The sample was expected to provide information regarding useful typologies and classifications of internet business models.

The second phase of the research proposed was designed to analyse the domain name data from the first phase employing a qualitative theory building approach. In essence, the aim of the second phase was to further categorise the websites using a theory generating methodology.

Ticehurst and Veal (1999) write, 'Research approaches are not considered to be intrinsically good or bad, but are considered to appropriate or inappropriate for the task in hand'. Given the task of conducting early phase interpretive research of Web

business from which theory building may be initiated, a mixed method approach was deemed most appropriate.

Phase One

4.4 Research Strategy – Quantitative Phase

The first phase was the data generation phase or preparatory phase. The data generation phase relied on quantitative random sampling methods. The objective for this phase was to create a statistically significant probability sample of randomly drawn domain names which would make up a database for additional analysis. Simon (2003, p. 126) put it this way, 'Your aim in taking a sample from a universe (population) is to draw valid conclusions about the universe without the cost or trouble of investigating all the units of in that population'. Since the researcher had access to the entire population, a simple random sample technique was utilised. Simple random sampling requires a list of the elements of the universe in question. 'From this list – the sampling frame – 'a selection is taken with some mechanical technique that seems random … ' (Simon 2003, p. 143).

The sampling method employed consisted of these steps. The first was to define the population to be sampled. As it is later explained, the target population shifted during the initial attempts to extract random samples. The original target population was to be all domain names in all domains on the Web. The actual target population sampled was all domain names in the .com domain, and all domain names in the .net domain. The second step was to create a sample frame from the universe of .com and .net domain names. The third step was to extract samples from the sample frame; in this case a simple random sampling method was employed, as opposed to a stratified sampling technique. In the case of a simple random sample method, all elements of the frame are treated equally and the frame is not artificially subdivided or partitioned (Wright 1979) .

The original goal was to sample the entire Web. This proved to be impractical due to the technology issues soon to be elaborated upon. The target population was shifted to two top level domains of the Web, the entire .net domain and the entire .com domain. These domains are also frequently referred to as 'zones'. Because of the size of each of the domains (30 million in the .com zone, 5 million in the .net zone), a

sample frame of 30,000 names from each of the domains was extracted using a simple random sample technique. To optimise the probability of a random draw for the sampling frame, a program was developed (Appendix A) which randomly extracted the 30,000 domain names from each of the two domains.

The sections that follow describe the initial testing phase of Web sampling and data generation. They document problems with the original design, which was to create a database of Web businesses using random generation of Internet Protocol addresses.

4.4.1 Initial sampling attempts

The first attempt to draw a sample of websites adapted a method based on O'Neill, McClain and Lavoie's (1998) methodology for sampling the World Wide Web. They used randomly generated Internet Protocol (IP) addresses to select Web pages to view and analyse. This project elaborated upon their process in that an algorithm was developed which included testing the IP address to insure it resolved to a Web page. The process is described as follows.

Internet Protocol address validation and database creation process:

- A program was developed to generate random numbers between 0-255 for each of the four octets in Internet Protocol (IP) address, for example, (0-255).(0-255).(0-255).(0-255).
- The program generated a complete IP address (the four octet set), for example, 238.145.7.63.
- The program would then 'ping' the random IP address it generated to determine if the number resolved to a Web Uniform Resource Locator (URL, or domain name), for example, www.jsu.edu . Ping is an acronym for Packet InterNet Groper. It is a utility that forwards data packets (in this case the IP addresses which were randomly generated) to verify the connection resolved to a machine on the Internet.
- If the ping failed to verify a valid address (or in other words if the IP address was not valid), the program generated another random IP address and pinged again.
- If the IP address did resolve to a URL, the URL (domain name) was saved in an Excel database.

Once this process was completed, each of the domain names saved in the database were manually evaluated and the content of the resulting website was analysed. 1200 plus pseudo randomly generated IP addresses were programmatically tested. The initial results were as follows:

- 126 IP addresses resolved to valid URL's. Each of these URL's was individually analysed and categorised.
- 10 were broken links (i.e., 404 errors). This indicated that the website had been allowed to lapse. For example, the owner may have neglected to renew the domain name or failed to pay the Internet hosting fees.
- 16 were access denied. For example, a password was required to access the site. This is not uncommon for protected sites or secure sites.
- 9 were university sites (7 per cent).
- 3 were government sites (2.4 per cent) – note: one govt site was also access denied.
- 1 health enterprise site.
- 1 site listed as under construction.
- 1 association site.
- 86 were classified as business sites (68 per cent).

At this point it was apparent that there was some issue to be resolved due to what appeared to be a disproportionate number of URL's which were classified as 'business sites.' Further analysis of the 86 URL's classified as business sites revealed the following:

- 1 site was unclassified due to a language barrier.
- 1 site was an energy company.
- 1 site was a pornographic material producer.
- 4 sites were retail sites.
- 10 sites were Web portal sites. That is to say, at this time in the research sequence the websites were categorised as portal sites. When a more thorough analysis of the data was conducted, these sites were classified as 'level 2' simple business models.

- 69 sites were classified as Information Technology related sites. Such as Internet hosting sites, telecommunications sites, or network providers. Note that this number represented 80 per cent of the 86 business sites identified.

The number of Information Technology sites relative to all business sites was clearly out of proportion. This finding led to the inevitable conclusion that the IP addresses either were not randomly selected or some other problem had intervened to corrupt the sample.

Further analysis of domain name sites and the identification of Internet Protocol related research articles revealed that the problem was related to the increasing use of dynamic IP addresses. As late as 1999 almost all websites had static IP addresses which would resolve back to the client website URL or domain name. However, after 1999 Internet Service providers and hosting companies began to increasingly assign 'dynamic IP addresses'. That is, when the client site logs onto the Internet an IP number is assigned by the Internet Service Provider. This IP address is in use for only as long as the client or continues their Internet session. When the client logs off, the same IP address may then be assigned to another client of the Internet service provider. When these dynamic IP addresses were generated by the random number generator program, the dynamic IP address resolved back to the Internet Service Provider, not the client site which is currently using the dynamic IP address.

Additional examination of research articles related to the new dynamic protocol for assigning IP addresses revealed that between 45 per cent and 60 per cent of all IP addresses assigned are now dynamic IP addresses (Ricerca 2004). This appears consistent with the results reported in the current study. For example, 80 per cent of the 68 per cent of business domain names generated were IT companies. This implies that about 54 per cent of those domain names would have been resolved from dynamically assigned IP addresses. It was determined that this number represented those companies, groups, organizations, individuals, blogs, or other websites that received short term dynamic IP addresses. This also meant that the process which had been developed would not provide the data needed to complete the exploratory phase of the project.

Nevertheless valuable data was generated. As stated earlier, the findings confirmed estimates derived from other recent research studies regarding the volume of dynamic

IP addressed client sites. And because university websites are universally represented by static IP addresses the proportion of their representation is very likely a good approximation of their actual representation on the Web (7 per cent). The same may be said for government sites (2.4 per cent). However that did not solve the problem encountered which foiled the attempt to generate data regarding the proportion of business websites on the Internet.

4.4.2 Other Web sampling methods

Researchers now posit the notion that random generation of IP addresses is no longer practical because of the ever increasing proliferation of dynamic or 'shared' IP addresses, e.g., (Edelman 2003; Henzinger & Lawrence 2004; Ricerca 2004). This problem limited the research options to attempting to develop a program or method to generate random domain names as opposed to IP addresses or to sample large domain name databases from the agencies that maintain them. Another available option was to purchase domain name databases from private organizations which provide access to them for a fee. However, this alternative was discounted for two reasons. Firstly, it was far too expensive, with the cost of these databases running to five figures. Secondly, there would be no way to assure the quality of the sampling methods employed by these enterprises. Each of these options was pursued and each was found to raise significant issues for the research project.

It was determined that the least potentially time consuming option would be to attempt to randomly or quasi randomly generate domain names (as opposed to IP addresses) to develop the initial database. This process generally depends on the random generation of some number of words (usually 2 to 5) from an online dictionary or a randomly selected vocabulary list. These words are then included in a search process utilising a selected search engine database. The highest ranked page generated using the randomly selected word(s) is presented to the researcher, and the domain name is then saved to an Excel database for analysis. Google, for example, makes its database of over four-billion Web pages available to researchers for such purposes. However, this method was found to consistently result in biased samples, since the search results rendered by the random words used Google's 'page rank' algorithm which returns the most popular, or relevant site, rather than a site which

was randomly selected. Attention then turned to sampling by domain name zones. Here, the project met with greater success.

4.4.3 Sampling – Top level domain selection

The original plan to sample IP addresses and thus draw samples from the entire Web was abandoned and the decision was made to sample top level domains. The next question to be addressed was which top level domain(s) (TLDs) should be sampled. This was not a trivial decision as there are currently over two-hundred sixty-four TLDs available to any enterprise or individual seeking to acquire a domain name. According to VeriSign data (VeriSign 2004b), the largest domain registry is in the .com zone. As of the last quarter of 2005, Forty-six per cent of all domain names registered on the Web are registered in the .com zone. By comparison, thirty-six per cent of domain names are registered in the all of the various country code TLDs, the largest of which is .de (Germany) with eleven per cent. The next largest single category was the .net zone.

It was determined to conduct research in the .com zone, but that research strategies to be employed would initially be tested using the .net zone. This would allow the researcher to 'debug' untried sampling strategies, and also would provide a means of triangulation of the data in the .com zone. Further, these two TLDs represented well over fifty per cent of all domain names and were reasoned to be the domains most likely to accommodate business enterprises, since historically the .com zone was the zone chosen to represent commercial utilisation of the Web and the .net zone was to be reserved to Information Technology and Network type businesses.

4.4.4 Domain zone sampling method

Choosing to sample by Top Level Domain (TLD) changed one aspect of the project. The researcher would not be sampling the entire web, but would attempt to sample by specific TLD, in this case the .com and .net zones. The .com zone remains a preferred naming convention for business and other enterprises. The rest of the chapter describes the process used to obtain the final sample from the .com and .net zones.

VeriSign is the American company charged with managing the .com and .net zones. VeriSign provides university affiliated researchers access to the Zone Files through a relatively simple application and agreement process (See Appendix G).

The requirements at the time the researcher registered with VeriSign included that the machine from which the zone database access was to be conducted must have a static IP address. Most universities employ static IP addresses on a majority of campus machines. VeriSign Corporation granted Jacksonville State University access to the domain name data for research purposes on October 4, 2004. Along with the completed agreement, the university received a specific File Transfer Protocol (FTP) address and a FTP access password.

The next step was to use a FTP program to access the VeriSign Domain Name Database. 'FTP Commander' was selected as the FTP program. A significant factor was the fact that FTP Commander is available as a free download from several sites. Once connected to the VeriSign FTP website, both the .com and .net database files were selected and downloaded to the JSU business research computer. VeriSign provided the databases as highly compressed 'zipped' files. The size of the compressed files as downloaded were 68.6 megabytes for the .net zone and 420 megabytes for the .com zone. The next step was to use a program (WinZip) to extract the full files from the compressed formats. Once the full files were extracted the size of the .net zone was 366 megabytes, the com zone was 2,571 megabytes.

The sheer number of records in the .com zone database (in excess of thirty million records) and the size of the files proved difficult to manage. The file sizes also created data access issues. For example, many text reading files such as MS Word or MS NotePad were simply unable to load files as large as the expanded database. Database enumeration proved to be problematic as well. For example, programs which might easily be used to enumerate each record (such as Micrsoft Excel) could not handle files of that magnitude.

4.4.5 Creation of a sample frame

To address the issue of .com and .net zone file sizes, an algorithm was developed (see Appendix A) which randomly extracted 30,000 domain names from files based on pseudo random numbers seeded by the user. Thus, the program could read through the entire file comprising the universe of domain names in the .com zone, use pseudo random numbers to extract records, and then write the selected records to a new file of domain names. This process resulted in a file of a more manageable size. A file of

30,000 domain names from the .com zone became the working database or sample frame from which the sample pool could be drawn. An editing tool, (TextPad) was employed to load the sample frame. TextPad is a programmer/editor tool which includes the capability to instantly enumerate each record (domain name) in the sample frame.

Next, an additional 1300 random numbers in the range of 1 to 30,000 were generated and pasted to an Excel spreadsheet in sets of 100 numbers. These sets were used to select the correspondingly numbered domain names from the enumerated framing sample in the TextPad file. This resulted in an Excel file of 1300 names drawn from the framing sample. Each record included the domain name and its associated number from the TextPad sample frame. This allowed cross checking of each record in the final Excel file with the associated record in the Random Number file and the TexPad enumerated sample pool.

The population size was 30 million (at the time the .com database was downloaded), and to achieve a confidence interval of plus or minus three per cent at a confidence level of 95 per cent, it was determined that the minimum acceptable sample size needed to be 1,067. The actual sample size employed was 1,286 after duplicate domain names were removed.

4.4.6 Analysis of quantitative data

At the time of the analysis, four popular Web browsers were available to conduct the initial analysis of the database of domain names, Internet Explorer version 6.0, Netscape 7.1, Mozilla Firefox version 1.0, and Opera version 7.54.

For simplicity and to minimize errors, the data analysis regime was divided into several discrete functions. The first of these was the preparation of domain name data sheets of one hundred domain names (Appendix F). This section describes the regime followed to produce these sheets.

A team of three research assistants from the business school at Jacksonville State University assisted in the initial step which was the preparation of data analysis sheets. Each research assistant worked at a computer with a copy of the TextPad text editing software on it. An electronic version of an Excel spreadsheet of one hundred

random numbers in the range of 1 to 30,000 in spreadsheet column one was provided to each of the research assistants.

In a separate window on the respective computers, each research assistant was instructed to open a copy of the database of thirty-thousand enumerated domain names in the TextPad application, (see Figure 4.1 below).

Figure removed due to copyright restrictions

Figure 4.1 Listing of TextPad enumerated domain names from the database of 30,000 domain names

The research assistants were then instructed to observe the first random number on the Excel spreadsheet. Moving to the TextPad application, researchers were asked use the 'Go to' function and type in or paste in the random number from the Excel spreadsheet. This immediately moved the researcher through the database to the equivalent record number in the domain name database. For example: if the random number copied was 20470 the resulting record contained the domain name 'POWERSCOPE.com.'

The research assistants were then instructed to use the TextPad 'Copy' function to copy the domain name. Next the research assistant was instructed to return to the Excel spreadsheet and use the 'Paste' function to paste the domain name into the appropriate column in the Excel spreadsheet.

This process was repeated until all one hundred random numbers on the sheet had the associated domain name in place. The domain names were then turned over the author for actual analysis of the domain names. As the research assistants gained some facility in the process, the average completion time for one sheet of one hundred random numbers was plus or minus thirty minutes.

The next task was to complete the analysis process by examining each domain name individually. A second set of research directives was completed for two reasons. Firstly, in order to insure that the processes were adequately documented, and secondly, in the event the primary researcher was absent or unable to complete a research segment, a clear set of processes were available to assure adherence to the established regime.

The researcher received one of the completed domain name data sheets. Each data sheet was numbered and each number was recorded on a master log. The name of the researcher working on the sheet was recorded on the master log as well. In order to accomplish the analysis process, the researcher was instructed to have two open applications on the Windows desk top, firstly the Excel spreadsheet with the random numbers and domain names (generated in the initial step documented above) and secondly an Internet browser application. The Internet browser most frequently used was Internet Explorer Version 6.0.

The next step was to use the Excel 'Copy' function to copy the first domain name on the list. The researcher was then instructed to use the browser 'Paste' function to paste the domain name into the browser address window. This process was employed to minimize the possibility of misspelling domain names. Next the researcher was instructed to download the resulting website to the browser and to observe the website and to classify the website using the categorization classes previously established. The category listing was provided to both the researcher and the graduate research assistant. The determined category was typed into the specified column on the Excel spreadsheet. As the researcher became more familiar with the process and the

categories, and partly because this was a more exacting process, the average completion time for categorization of one hundred domain names varied between two hours and two hours - thirty minutes. When fourteen such spreadsheets were completed the initial quantitative data analysis process was deemed finished.

The final categories described later in this section came about as a result of extensive involvement with the data. The analysis of the .net zone became a proving ground (or beta test) for the categories to be employed to describe the websites. These categories evolved through the analysis of the 1557 domain names sampled from the .net zone. Written labels for the categories were established once the point of category saturation was reached. Category saturation was the point at which the selection was varied enough to provide meaningful discrimination between sites, and the researcher no longer needed to condense categories.

As each domain name was resolved, the result was categorised based on written guidelines describing each category. This helped to maintain consistency of categorisation class. Periodic random reviews provided an additional means to maintain consistency. The final categories applied to the resolved domain names were as follows:

- Business site.
- Personal site.
- Organization site (religious, political, governmental, educational, and etc.).
- Other site or site devoid of content or a server directory or not one of the above.
- Site not found error message.
- Redirected to another site such as a domain name service provider (more on this below), or VeriSign 'Site Finder' type page, or browser search/portal page, or 'parked' domain name page.
- Access to the site denied due to login requirement or password.
- A site reporting to be Under Construction at the time we visited.
- A site deemed 'unable to categorize'.

See Table 2.1 on page 17 and Table 2.2 on page 18 for a tabular listing of the results.

As the first phase of the research found anomalous data, a second phase was required to further explore the anomalies. The anomalous data from phase one were disaggregated for further analysis employing qualitative methodologies.

Phase Two

4.5 Research Strategy - Qualitative Phase

The discovery of anomalous websites introduced a new phase to the research project. Table 4.3 shows that the disaggregated data formed a pool of some 265 websites representing 20.64 per cent of all websites in the .com zone. The researcher was in a position not unlike the one described by Straus and Corbin (1990, p. 29) when they wrote 'At the same time the researcher has to be puzzled or disturbed about some feature of the data or about their interpretations, so that questions and answers must be raised and sought.'

Kuhn (1962, p. 53) wrote 'Discovery commences with the awareness of anomaly… It then continues with a more or less extended exploration of the area of the anomaly.' The proportion of the anomalous sites encountered raised several questions. First among them, 'What are they', then 'Why are there so many'? These sites, which were originally erroneously classified as 'Redirects', occurred 19.1 per cent of the time in the .net zone also. To gain some perspective on this phenomenon, at the time the analysis was begun (first quarter of 2005), VeriSign estimated some 30,000,000 domain names had been issued in the .com zone and 5,000,000 domain names in the .net zone.

When extrapolating from the data, it was estimated that there would be approximately seven million domain names that would be redirected to other sites (i.e., sites with a different domain name) in the .net and .com zones. The finding that slightly more than one in every five registered domain names redirected to other sites was unexpected, and led to deeper investigation of the phenomenon in search of an alternate explanation.

The question of a formal methodological description is always close to the surface of any research activity. Which of the qualitative methods would be most appropriate for the work to be conducted? Denzin and Lincoln (1994, p. 3) wrote, 'Qualitative

research, as a set of interpretive practices, privileges no single methodology over any other'. The researcher's initial reaction was that grounded theory was the most logical methodological choice, and elements of the grounded theory are clearly evident throughout the balance of the analysis. For example, Straus and Corbin (1990, p. 43) noted that:

> Insight and understanding about a phenomenon increases as you interact with your data. This comes from collecting, and asking questions about the data, making comparisons, thinking about what you see, making a hypothesis, developing small theoretical frameworks (miniframeworks) about concepts and their relationships. In turn the researcher uses these to look at the data.

This suggested the use of an iterative method of looking at data, comparing and analysing the data, coming up with statement of association (theories), and with this new information re-examining the data. Eisenhardt (1989, p. 541), in discussing case study as a methodology, wrote 'The central idea is that researchers constantly compare theory and data – iterating toward a theory which closely fits the data.' This idea is also in keeping with Carlile and Christensen's (2005) theory building model; firstly, observe and describe the phenomenon of interest; secondly, categorise the phenomena based on their attributes; thirdly, make statements of association. The statements of association then form a platform for further research. Holland (1998, p. 7) observed that 'At each level of observation the persistent combinations of the previous level constrain what emerges at the next level. This kind of interlocking hierarchy is one of the central features of the scientific endeavour.'

Holland's explanation is also descriptive of the hermeneutic circle, or what Gummesson (2003) refers to as the hermeneutic spiral. Gummesson sees hermeneutics as an extension of interpretation. The process is one of pre-understanding, understanding, and explanation. He terms this a spiral because the 'explanation' from the first iteration forms the pre-understanding for the following iteration of the research cycle. In summary, Gummesson also views hermeneutics as a general methodology for interpretivist research which is concerned with non-lingual expression of human life (as in text) and which strives to avoid speculation and bias.

The term hermeneutics is defined as 'concerning interpretation, especially of Scripture or literary texts' (The Oxford Concise Dictionary – Online Edition). Hermeneutics

has long been applied to the fields of biblical exegetical work, legal interpretation and linguistic and textual analysis (Addison 1994). Schwandt (1994) provides an historical perspective on hermeneutics. He observed that interpretivism is layered with ideas stemming from the German intellectual tradition of hermeneutics. He sees the goal of the research as grasping or understanding the meaning of social phenomena in the real world.

Like Gummesson (2003), Schwandt (1994) sees hermeneutics as attempting to avoid subjectivity and the error of naïve inquiry by the judicious use of method. Miles and Huberman (1994) also note that hermeneutics has a long pedigree of intellectual history and they place it squarely in the interpretivist and phenomenological realm of inquiry. Hermeneutics' emphasis is on more than just explaining a phenomenon but in achieving understanding of the phenomenon (Gaignon 2002). This is accomplished by a complete immersion in the object of research, and by the continuous interaction with the data, similar to the strategy suggested by Straus and Corbin (1990) in Grounded Theory. Further justification for a hermeneutic methodology came from its extensive use in the examination of textually and graphically rich data (Reichertz 2005).

Hermeneutics has an increasingly vocal and enthusiastic group of advocates in the Internet related field of Information Technology (IT). In the past decade, several authors have called for more interpretivist research articles in order to better represent human thought and action, and to better understand organisational impacts on the field of IT (Boland 1985; Klein & Myers 1999; Marshall, Kelder & Perry 2005; Walsham 1995). For example, Klein and Myers (1999) suggested a prescriptive set of principles for interpretivist researchers in the Information Technology field. Their first principle is one they label The Fundamental Principle of the Hermeneutic Circle:

> This principle suggests that all human understanding is achieved by iterating between considering the interdependent meaning of parts and the whole that they form. This principle of human understanding is fundamental to all the other principles (p. 72).

Walsham (1995, p. 378) observes that 'Boland (1985) used phenomenology and hermeneutics as the philosophical basis for his research. He argues that the use,

design, and study of information systems is best understood as a hermeneutic process.'

The Institute for Hermeneutic Research on Culture and Society (Oevermann 2003, p. 1) observe that hermeneutics is also appropriate in a mixed methodology format:

> The quantitative procedures of empirical social research are considered useful tools for establishing the frequency of a given diagnosis once a diagnosis has been developed but it is the procedures of objective hermeneutics which serve to identify basic structures and types. In this sense "qualitative" or "interpretive" procedures are not considered to be part of an exploratory stage of a type of research which will then proceed in a quantitative logic, but—on the contrary—as its foundation

To summarise the rationale for applying a hermeneutic methodology to qualitative phase of the project, hermeneutics:

- Is perceived as a fundamental research process – particularly from the German and northern European perspective.
- Presents an iterative research process where data collection and data analysis are nearly synchronous activities, punctuated by periods of reflection of the data (The hermeneutic spiral).
- Has a robust history of utilisation in textual and graphic analysis, which makes it particularly appropriate for website research such as this project entailed.
- Has a strong following amongst researchers in the related field of Information Technology.

These aspects suggested that hermeneutics would be the most suitable qualitative methodology to employ in the present study.

Finally, it is appropriate to note that this study also represents an example of an inductive theory building research project. An inductive approach attempts to infer general patterns of order or structure from a set of empirical data (Parkhe 1993). Typically, the empirical, interpretive, inductive research style is involved with the collection of detailed qualitative information and does not attempt to confirm or disconfirm hypotheses (Walsham 1995).

4.5.1 Data collection for qualitative analysis

The sample data set for qualitative analysis resulted from the disaggregation of anomalous data from the original randomly drawn sample. The anomalous data represented websites which the researcher was unable to definitively classify. The selection of only these anomalous sites helped to constrain theoretical variation and limited the qualitative research portion to sites representative of the process under investigation. It reduced extraneous variation and classified the domain of the findings only to websites of interest at this point, and narrowed the focus directly on the phenomenon which was not yet fully understood as suggested by Eisenhardt (1989).

4.5.2 Data analysis process

The data analysis called for some additional preparation of the 265 domain names to be analysed. A specific database of the records that made up the anomalous domain names was established as an Excel spreadsheet. The Excel spreadsheet enabled the researcher to add comments as well as to freely add columns or fields of data as needed. In order to minimise data loss, the domain names were cut and pasted as a group into the new spreadsheet.

Addison (1994, p. 1-3), list five central process related aspects of hermeneutic research strategy. These five aspects are listed here because they provide insight to the analytical goals that were aimed for in the qualitative analysis phase. They relate to this study in the following ways:

- Hermeneutics seeks to uncover what is hidden and illuminate the background, and to find what is missing.
- Hermeneutics place interpretations on a larger social, cultural, historical or economic background. That is to say, researchers learn certain foreground conditions and explore the background conditions which make the foreground conditions occur. As can be seen below, this was an important aspect of this research project. Ascension up the hermeneutic spiral led to opportunities to understand the enabling characteristics operating in the background. These are explored in the next chapter.

- Hermeneutics decree that the research moves back and forth between understanding and interpretation. This aspect is similar to the call for immersion in the data and then reflection upon the meaning of the data.
- The hermeneutics researcher should practice conscious self-reflection in order to maintain an objective view. The researcher should constantly endeavour to limit personal biases from colouring the analysis process.
- The researcher should allow for new possibilities, change and growth. As the researcher understands more and more about the phenomenon under study, it becomes feasible to envision new possibilities of explanation based upon external phenomena. This particular aspect played its part in the present study and is explored further later in the chapter.

This exegetical process is based upon the first variant of text explanation (Reichertz 2005, p. 573), which is referred to as the 'detailed analysis strategy. …The first variant initially found many adherents within qualitative research, not least because in its most important elements it is formalised and therefore easy to learn'.

The analysis of individual domain names continued the process from the initial examination in which each of the domain names was cut and pasted into the browser address window. The researcher then typed in the leading 'www' and the trailing 'com'. When the Web address was resolved and the page loaded, the resulting site would be analysed using a hermeneutics strategy designed for textual analysis.

The first thing which became apparent was that once the 265 anomalous domain names were disaggregated from the original 1284 domain names sampled from the .com zone, the structure and patterns inherent in these websites became more readily apparent. The first question that was addressed was the error in originally classifying these names as "Redirects to other sites".

4.5.3 Resolving the erroneous classification issue

The data classification sheet which the research project employed helped to resolve the issue of erroneous classification of the anomalous sites. The field labels of the data capture sheet are described in Table 4.4:

Table 4.4 Field Labels and Descriptions – Data Capture Sheet

rn:	The original random number corresponding to the domain name's ranking in the VeriSign .com database.
Dn:	The domain name.
category:	The original category assigned during the quantitative phase.
type:	The actual type (which was determined as the qualitative research project progressed).
subtype:	This became the field to be assigned to the domain park host site as the patterns became clearer or from the 'Whois' registry.
Comment:	Used for any comment or for the actual Web address to which the domain name resolved.

When data analysis was commenced and the websites were reviewed, a rationale for the original classification became more apparent. For example, the actual domain name entered was;

www.certifiedmortageprovider.com.

But the domain name that ultimately appeared in the browser address window was;

http://apps5.oingo.com/apps/domainpark/domainpark.cgi?client=netw8744&s=www.certifiedmortgageprovider.com.

In effect this was a re-direct, but not in the sense that the domain name could not be resolved. In fact the name was redirecting to a specific type of host site. Only 40 per cent of the domain names in the qualitative sample actually resulted in a new name appearing in the browser window. However during the quantitative phase, the sites that had similar attributes were also classified as redirects. These discrepancies were not apparent in the original database but were readily apparent in the smaller qualitative dataset because so many similar names appeared. For example, many of the original domain names finally resolved to a name beginning with:

http://apps5.oingo.com/apps/domainpark/domainpark.cgi.

The singular phrase in the name above, “domainpark” was an initial indication that this was a specific type of phenomena, which deserved deeper investigation.

4.5.4 Pattern recognition

A fundamental aim of research on data such as this is to begin to identify patterns amongst the data items. The hermeneutic method, which calls for constant immersion in the data, led to the discernment of patterns amongst the websites analysed.

The recurring pattern of so many domain names resolving to a similar address represented one of the first instances of pattern recognition. Another was the general design characteristics of the sites themselves. While difficult to verbalise, these patterns become quite apparent with a visual representation of the websites. Figure 4.2 and Figure 4.3 indicate the consistency of design elements and textual content which led to the initial pattern identification.

These sites, though dissimilar in subject and thus topical content, show a clear similarity in their textual makeup. As noted above, pattern recognition was much more readily discernable once the analysis was limited to the qualitative database. As the analysis process progressed, one of three strategies was employed to maintain the information.

These sites, though dissimilar in subject and thus topical content, show a clear similarity in their textual makeup. As noted above, pattern recognition was much more readily discernable once the analysis was limited to the qualitative database. As the analysis process progressed, one of three strategies was employed to maintain the information.

Figure removed due to copyright restrictions

Figure 4.2 The first example of websites which lead to pattern recognition

Figure removed due to copyright restrictions

Figure 4.3 The second example of websites which lead to pattern recognition

- Firstly, an additional data field (column) was added to the spreadsheet. This additional data field might contain a simple phrase which was descriptive of the overall design of the site, or a notation of a specific design element of the site. For example, one recurring design element was an image of a highway with mountains, a blue sky and a highway sign as shown by Figure 4.4. In the additional field on the data sheet, any site featuring this design element was

labelled 'blue highway'. Thus these previously viewed sites could be easily referred to for further review and comparison purposes.

- Secondly, a comment field might be added to an individual data cell. This was most commonly employed when the individual domain name appeared to have some specific attribute not common to other domains, or when a specific time and date may need to have been recorded for some event related to the domain name.
- Thirdly, in some instances an index card was completed and the information filed.

Figure removed due to copyright restrictions

Figure 4.4 The 'Blue Highway' design represented an example of pattern duplication.

At a macro level the process took on the attributes of Gummesson's (2003) description of hermeneutics' three stages of pre-understanding, understanding, and explanation. With each cycle through the list of sites, greater understanding of the phenomenon was achieved. For example, subsequent cycles through the database of domain names led to clear identification of sites which were, in effect exact duplicates of each other, the only differentiation being the domain name itself and the subsequent links which made up the content of these sites. Figure 4.5 presents two sites which illustrate the occurrence duplicate design.

However, the identification of these duplicate sites led to further questions. For example, if these sites employed duplicate site designs, did this imply that they were owned by the same individual who was essentially deploying one design over many sites, or was this an indication that another Web business entity was providing the service of site design? This new finding immediately became the most demanding variant, and as such, came to represent a perfect functional example of deeper understanding leading to further questions, the answers to which led to an even deeper understanding, i.e., the hermeneutic spiral.

Figure removed due to copyright restrictions

Figure 4.5 Duplicate patterns of websites with unique domain names began to emerge.

To answer these questions, a greater range of tools was required. The Internet Corporation for Assigned Names and Numbers (ICANN) requires website registry services to maintain a list of website registrants. This facility is known as the "Whois" registry. In most instances, one may visit a central online registry to learn the name of the domain name registrar, and at the registrar's website ascertain the name of the owner of the domain name.

This process led to the determination that the websites were independently owned, which meant that some other Web service entity was responsible for the design similarity. This understanding led to an additional level of questions. What were these entities? What was their function? Were these merely Web design studios?

Further cycles through the database revealed additional and more subtle patterns amongst the domain names themselves. In many instances there were more subtle similarities between domain names of similarly patterned sites. These led to new and even more sophisticated implementations of business structures. Many of these names had a generic quality about them. These were not sites with what are typically understood to be specific company names, such as 'mikesPCsoftwaresales.com'. More likely the names would be 'kitchenfurniture.com' or 'furniture.com'. It emerged that the names began to fall into one of a few categories, e.g. domain names that were generic in nature such as:

- Prophotographers.com;
- Gunbuyer.com;
- Globalmember.com;
- Managedfund.com.

A second category which came into focus was domain names that were close approximations of actual entities, but which often entailed a misspelling, such as the following:

- Plamsatv.com (approximates plasmatv).
- Tai-che.com (approximates tai-chi).
- Giajoe.com (GIjoe).
- Computervest.com (computerfest).

The consistent replication of these patterns suggested that these might represent a new business structure which was emerging (or had recently emerged) within the Web business environment.

The process discerned thus far reflected the cycle of data collection and data analysis. This process led to a new level of understanding or new clarity of the entities analysed. This in turn led to the identification of new external entities which it transpired had a direct impact on the classification of previous entities. This required a new data collection cycle aimed at gathering an understanding of these new entities, as well as their role in the emerging phenomenon, as diagrammed in Figure 4.6. The discovery process described thus far is an example of Klein and Myers' (1999, p. 72) first principle for interpretive field research: This principle suggests that all human understanding is achieved by iterating between considering the interdependent meaning of the parts and the whole that they form.

The new data collection cycle variously involved a search through an ownership database as well as the deeper analysis of host websites in order to better understand the service being provided. As the analysis developed further, in order to understand the business structures involved, it became necessary for the researcher to acquire domain names and participate more fully in the process.

In the hermeneutic process, the analysis period (Figure 4.6) is the time in which the researcher considers not only the individual components, but also the impact these components had on the entire environment, and the impact the environment may have had upon the components.

The explanations for the patterns discerned and the business processes and revenue generating capacities from this phenomenon are more fully discussed in Chapter 5 which follows.

Concurrent with the analytical process – the search for an explanation of these phenomena was taking the researcher deeper into the study of complexity theory and network theory. These seemed to offer the best theoretical rationale and contextual framework for what was uncovered. These disciplines suggest the possibility that the phenomenon being researched might represent an example of emergent phenomena within the Web environment. If so, this explanation would apply not only to these

anomalous websites, but also to the Web entities that were providing some level of service for them.

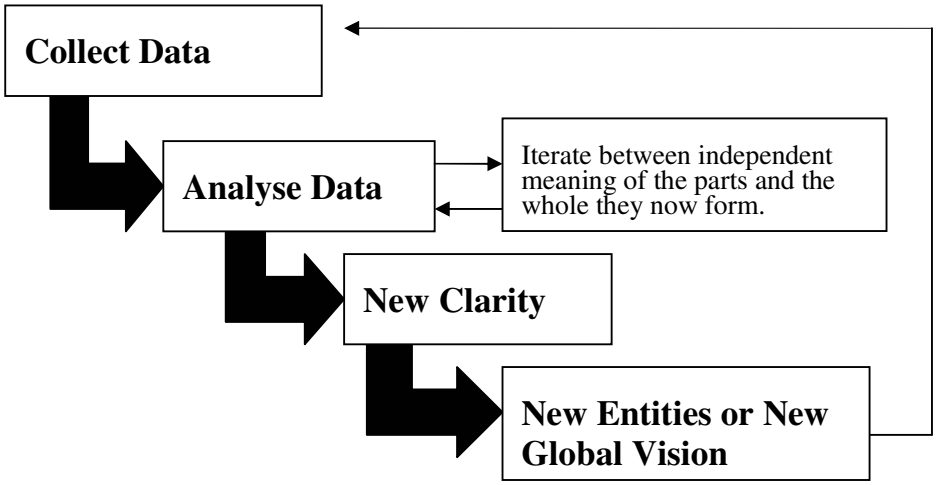


Figure 4.6 Graphic depiction of the qualitative research process (the hermeneutic cycle)

This was not a preconceived theoretical framework for the explanation of these events. The research was not begun with the view that complexity theory could or should be applied to the business environment of the Web. Rather, from this point forward, the further up the hermeneutic spiral the research travelled, the more fully complexity theory became integrated into the explanatory framework.

4.6 Ethical Considerations

Ethical considerations are a growing concern for business researchers. Most ethical concerns centre on responsibilities to respondents. The primary admonition is to do no harm to the respondent, and to insure the respondents take part freely and that their participation is based upon informed consent. (Cavaye et al. 2002).

Zikmund (2000) identifies three stakeholders whose rights and obligations are to be considered. They are the subjects (respondents), the researchers, and the client (or sponsoring agency). The 'subjects' of this research are public Internet business websites. This project did not involve research on human or animal subjects, and thus did not fall under the approval requirements of the 'Southern Cross University

Guide/s for Research Involving the Use of Human Participants'. Specifically, no part of the proposed research fell under the province of Section 3 of the Guide; 'Activities Which Require Ethical Approval'.

Ethical issues in Internet research are a relatively new and rapidly expanding topic for consideration. These issues primarily focus, once again, on research which involves human participants as subjects and much of the literature on Internet ethics is generated from the field of psychology (Peden 2004).

Authorities agree that Internet research should maintain adherence to a code of research ethics (Leedy & Ormrod 2001; Nancarrow, Pallister & Brace 2001; Zikmund 2000). Nancarrow, Pallister and Brace, in particular, address issues surrounding Internet research methods. They identify seven ethically questionable practices directly related to Internet research. These practices express concern that the nature of Internet research (i.e., it's speed and availability to a wide range of researchers) would result in a disaffected universe of Internet users unwilling to participate in research activities if such activities are not checked by adherence to a strong ethical code. They also note that the European Society for Opinion and Marketing Research (ESOMAR) 'has developed a set of Internet guidelines to try to ensure respondents understand how the industry and the research organisation with which a respondent comes into contact view privacy amongst other things' (p. 145). The research activities which were conducted to prepare this thesis adhered to the ESOMAR Internet Research guidelines.

Klein and Myers (1999) principles provide a set of fundamental ideas regarding interpretive (and particularly hermeneutic) research methods. While not intended to specifically address ethical issues, nevertheless adherence to the principles they suggest would at least act as a check to ensure interpretivist researchers consider their judgements during the course of the project. For example, their seventh principle is 'The Principle of Suspicion' (p. 72) which challenges researchers to be sensitive to their own biases.

Cresswell (2003, p. 67) observes that ethical issues do not stop with data collection and analysis but extend to writing as well when he states that 'Other ethical issues in writing the research will involve the potential of suppressing, falsifying, or inventing findings to meet a researcher's or an audience's needs'. In a similar vein, James

Thomas (1999, p. 5) writes 'I argue that we need not invent new ethical rules for online research. We need only increase our awareness of and commitment to established ethical principles.'

The researcher agrees with this position and posits that in the final analysis ethical performance is contingent upon the integrity of the researcher(s) and their commitment to do the right thing and to base actions on contemporary societal norms.

4.7 Conclusion

This chapter presented the research objective, the research framework and justifications for the research paradigm and methodology. The chapter also discussed the sequential phases of the project in order of occurrence. Phase one dealt with the quantitative method in detail, whilst phase two described the qualitative analysis process. The chapter closed by discussing ethical considerations in research involving the Web.

By the middle of the 1990's, the impact of large growing communications nets with complex architectures, the Internet and the World Wide Web, on our civilization became incredible. However, understanding of their global organization and functioning was absent (Dorogovtsev & Mendes 2003, p. 2)

CHAPTER 5 – FINDINGS AND DISCUSSION

5.1 Introduction

Chapter 4 discussed the mixed method research models and processes employed to complete the analysis of phase one and phase two of the project.

A primary objective of this study was to conduct research in order to gain a more comprehensive understanding of the conduct of business in the environment of the World Wide Web. This chapter discusses the findings of the research. These findings deal with the fundamental building blocks of businesses on the Web, domain names as they relate to complex systems. In this regard, Waldrop (1992, p. 279) writes:

> The most surprising lesson we have learned from simulating complex physical systems is that complex behaviour need not have complex roots … Indeed tremendously interesting and beguiling complex behaviour can emerge from collections of extremely simple components.

This chapter labels several of the findings regarding the emergence of new business structures as 'simple business models', but the use of the term 'simple' should not be construed so as to imply that these business models do not represent the components of a very complex system.

As noted in Chapter 3 at page 27, this thesis employs the succinct definition of a business model provided by Turbin (2004, p. 11) that 'a business model is the method of doing business by which a company can generate revenue to sustain itself'. This definition reduced the tendency to try to force emerging business models into pre-existing taxonomies, since any such new entities might not employ readily discernable revenue generation mechanisms.

Using revenue generation modalities as a point of departure allowed the researcher to view business models from the perspective of the fundamental building block of any Web based business endeavour, the domain name itself, and then use categorise

increasingly sophisticated methods of revenue generation as a primary business model determinant. This method has the advantage of offering the researcher greater flexibility to define and differentiate new business entities at the core of Internet business.

The findings of the quantitative phase are presented first. These are relatively simplistic quantitative measures of classes of sites populating the Web. An original purpose of the research was to sub-categorise business websites. However the discovery of the anomalous websites was an unexpected finding that has led to a richer study.

The findings of the qualitative phase flowed from twelve months of analysis of many websites, domain name acquisition processes, domain name ownership databases, and new methods of hosting domain names.

Finally, this chapter relates the research propositions established in Chapter 4 to the findings and discusses the appropriate role of complexity and network theory as a conceptual framework. In this section, the format is to re-introduce the individual propositions followed by some discussion regarding the affirmation or disconfirmation of the proposition.

5.2 Quantitative Research Findings

The .com count total of 1,284 represents the total of domain names after duplicate domain names (from duplicate random numbers) were removed. Table 5.1 summarises the results of the initial domain name analysis.

Table 5.1 Results from the sample of 'COM' zone domain names

Category	Count	Percentage
Business site	441	34.35
Not Found error message	303	23.60
Redirect to other site	265	20.64
Under construction	75	5.84
Unable to categorize –	73	5.69
Personal site	44	3.43
Other (empty)	36	2.80

Organization site	31	2.41
Access denied	16	1.25
Grand Total	**1284**	**100.00**

The findings indicate that the largest single category identified is the 'business site' item. 'Not founds' also represented a high proportion of the domain names investigated. These were partly explained as instances where an owned domain name had been allowed to lapse or the owner simply chose not to renew the ownership of the domain name. The third major category was 'Redirects to other sites'. Redirects take place when one types in a domain name (or links to a domain name) and is 'redirected' to another site in order to resolve the domain name.

Sites categorised as 'Redirects to other sites' were subjected to deeper analysis for several reasons. Firstly, it was not clear exactly what these sites represented. The 'Redirect to other sites' label was applied as a matter of convenience until the entire subset could be more deeply analysed. Secondly these sites seemed to present a distinct pattern as described in Chapter 4. For example, while the sites often had diverse content characteristics, they appeared to fall into distinct sub-categories. One group clearly redirected back to the original domain name supplier i.e., the site from which the owner of the domain name originally purchased it. These sites often provided some notification that the site was 'under construction' or was 'coming soon'. A second pattern to emerge involved sites that redirected to what appeared to be a version of a Web 'portal site'. Since these were ultimately recognised as anomalous results the discussion of this category continues in the qualitative findings that follow.

The 'Other' category included sites which were originally categorised as 'Empty'. These were sites where the domain name resolved to the active site, but the content was a blank page or a page with only a few words or a sentence. For example, the text might proclaim 'test site'. The category 'Unable to categorize' was used for sites which could not definitively be categorised. This was most often due to language issues but also occurred when the domain name resolved to a site consisting of unrecognizable fonts and/or characters.

The findings from analysis of both the .net and .com zone are discussed in this chapter. It should be noted that the .net zone served as a preliminary test zone to ensure the sampling technique was appropriate. It also served to refine category identification for the researcher. The 'not found' category was found to be somewhat higher in the .net zone and the 'business site' category somewhat lower (Featherstone & Borstorff 2005). The research led to the conclusion that sites which were classified as 'redirects' accounted for about twenty per cent of all domain names in both the .net and the .com zones.

A somewhat surprising aspect of the quantitative findings research was how closely the .com zone results followed earlier results quantified from the .net zone sites. It was anticipated there would be greater differentiation between various categories within the two domains. This is because the various top level domains were assigned for specific types on Web entities. The .com zone was to be applied to entities with a commercial interest. The .net zone was to be assigned to entities whose primary function was in the information technology arena or network suppliers. The present study demonstrates that, in the years since the inception of the existing domain name system, the two zones have reached the point where the distribution of websites by function is very nearly the same, as shown in Figure 5.2.

Table 5.2 Results from the sample of 'NET' zone domain names

Category	Count	Percentage
Not Found error message	490	31.47
Business site	373	23.96
Redirect to other site	291	18.69
Under construction	134	8.61
Unable to categorize –	89	5.72
Other (empty)	64	4.11
Personal site	61	3.92
Organization site	35	2.25
Access denied	20	1.28
Grand Total	**1557**	**100.00**

The results from the .net zone were ultimately used as a validation device to confirm results from the .com zone. The anomalous data was extracted and became the basis for establishment of a new database and formed the foundation for the qualitative findings discussed next.

5.3 Qualitative Findings

The previous chapter discussed the various patterns of websites and the manor in which these patterns led to consideration of the total environment in which the pattern had emerged. It is important to understand the nature of the patterns and the manor in which both the similarities and variations led to identification of emergent business structures. As interpretivist hermeneutics had suggested, periods of contemplation of the nature of these sites led ultimately to the fundamental building blocks of business on the Web – Internet domain names.

5.4 Domain name monetisation structures

The first indication that a specific business structure was behind these sites was the comparison of two almost identical sites from a design perspective, each with content consisting primarily of links to external sites. Further, many of these links were generically related to the domain name itself. Figure 5.1 illustrates this. The emphasis here is on the similarity of the overall pattern, rather than textual content.

The Whois registry indicated these domain names were independently owned. Thus a third party site was providing the site design. Since the site content was so similar, it seemed likely the business entity which designed the site may have provided content as well. These third party service providers presented the researcher with the first indication of a newly emerging Web business model, domain name parking.

Once a person or a company purchases a domain name several options present themselves. The common assumption has been that the domain name will immediately put to into service as a website. This project has shown that in a large number of instances, this is not the case. In fact there are several choices available to a domain name owner.

The first choice a domain name owner has is to do nothing at all. The domain name owner is under no compulsion to create a website, simply because a domain name has

been purchased. This actually represents a very simple business model. The owner may purchase a domain name in the same way an investor collector purchases a postage stamp or a coin or a painting. There is the expectation that the item (domain name in this instance) will appreciate in value and will be resold at a later date. This is not a new business model, but it has been unrecognised and unreported in the e-business research literature.

Figure removed due to copyright restrictions

Figure 5.1 Two parked sites with different domain names share the same design

VeriSign now sells nearly one million domain names every month in the .com zone alone. Simple business models, particularly those identified as level two and level three in this study are now considered by VeriSign to be the primary driver of domain name sales (VeriSign 2005).

As a consequence of the qualitative research, it was often possible to discern a simple business strategy at even the most elementary level.

5.4.1 Example 1 - A simple 'level one' business model

At this point in the project, an elementary classification scheme was employed in order to label and categorise the typologies the research revealed. The scheme which was employed was to simply label the business models incrementally, starting with one as the most elementary of models and increasing the level number as the perceived sophistication of the business model increased. An example of a 'level one' simple business model is:

Someone types in a domain name such as www.mikefeatherstone.com. This person expects to arrive at a website telling them something about Mike Featherstone, but instead the domain name is resolved to the Web page shown as shown in Figure 5.2.

Figure removed due to copyright restrictions

Figure 5.2 A level one business model representing the speculative purchase of a domain name

Figure 5.2 points to one reason why these sites were originally labelled as 'redirects'. In this instance, the name was resolved and redirected to one of a series of 'GoDaddy' websites which have been developed to handle domain names purchased where no further action was taken by the domain name owner. Note that this instance of

domain name resolution represents the state of the technology and business acumen regarding the value of domain names as of June, 2005.

GoDaddy, Inc was the domain name registrar (or provider) from whom the owner purchased the domain name. The fact that someone entering the domain name was redirected to that site might mean that the owner just purchased the name and had not yet deployed a website. It might also mean the owner purchased the name, and had no intention of building a website, but merely wished to hold ownership of the name until such time as some business or individual desired to make an offer to purchase the domain name from the owner. This strategy is as old as the Web itself. It represents one of the very early elemental business models of the Web. Originally called 'domain name mining', this strategy has more recently been referred to as 'cyber squatting'.

The researcher identifies this strategy as the most elementary business model on the Web. This is so because, firstly, it requires little technical skill. Secondly, the cost of purchasing a speculative domain name is about US$ 9.00. The strategy behind this simple business model is to acquire a desirable domain name and then resell it. During the course of this research project, one instance of a successful execution of this strategy was transacted in 2005 by the sale of the domain name www.voyuer.com (sic) for US$ 100,000.00.

The analysis of the anomalous websites found eighty-six instances (6.70 percent of the total sample of 1284; or 32 per cent of the total sample of 265 anomalous sites) where domain names redirected to what were clearly domain name registrars or service providers. If this percentage is extrapolated to the entire .com zone it suggests that over two million domain names would meet this criterion.

If the simple business model of domain name speculation had continued, one might have expected to see this percentage rise. However, the emergence of a new business model called domain name parking has meant that the acquisition of domain names solely for the purpose of resale is an all but extinct strategy. This phenomenon is discussed in the next section.

5.4.2 Example 2 - Simple 'level two' business models

The second example of a simple business model represents a new finding. It was labelled a level two business model because it represents a markedly more sophisticated model from a revenue generation perspective. Figure 5.3 illustrates a level two website. Note that the design of this site corresponds to the very anomaly that ultimately triggered this research project. A level two business model represents a different option available to the domain name purchaser. In this instance the domain name was purchased and, similar to the instance above, the owner may have had no intention of building a website. Instead the strategy was to 'park' the domain name at a specialised type of service provider. The term park is descriptive but somewhat misleading. Parking involves informing the domain name registrar that the domain name server resolving your site domain name will be the parking service provider, often naming the domain name server in a format provided by the service, i.e. ns1.dnserviceprovider.com.

The process is as follows:

After the domain name is purchased the owner registers with a specialised hosting site called a domain name park (or parking) site. The terms of service providers vary, but almost all offer to provide a Web presence at no charge to the domain name owner. In order to monetise the domain name, the owner must insure that the parking service will share revenues earned from the website with the domain name owner. At this point the name server for the domain name is specified, usually by physically typing in the name server address at the registrar site.

From this point forward, the domain parking site would take over most aspects of management of the domain name site. It provides the design for the site, which is often an automated process under programmatic control. It populates the site with a list of advertisers who pay either the domain name parking site or a third party advertiser aggregator (Google, for example) with whom the parking site has a business agreement. Anyone visiting the site who clicks on any of the links served up by the hosting company creates revenue in the amount of anywhere from US$ .05 to US$ .50 per click. The host providing the service may take a percentage of the total revenue, ten per cent to fifty per cent for example, as their revenue portion.

Complexity theory suggests that when new components emerge in an environment, numerous variations of those components may emerge as well. This is true both of the services providing domain name parking and of the strategies domain name owners employ for their sites.

The site presented in Figure 5.3 had all the appearances of a business portal site, and indeed, to some extent it is a business portal site. But the owner of the domain name www.toptraveltours.com neither developed, nor paid for development of the website. The owner simply chose to 'park' the domain name with a domain name service provider, in this case, Sedo.com.

Figure removed due to copyright restrictions

Figure 5.3 Example of a parked domain name

This study has shown that 13.8 per cent of all .com domain names are in fact 'parked domain names'. The margin of error is plus or minus 1.8 per cent at a 95 per cent confidence level. This extrapolates to over four million websites in the .com zone alone.

5.4.3 Domain name parking strategies

Partly because this represents a newly emerging business model, and partly because the Web represents a relentlessly changing business environment, the strategies employed are rapidly evolving phenomena. These strategies impact not only the domain name owner, but also the hosting site, service provider or even the domain name registrar. For example, Figure 5.4 illustrates a domain purchased at GoDaddy in July 2005. This domain name was not deployed as an active website. No website

was developed for it. In this instance, the domain name registrar has provided a substitute screen as a placeholder for the Web presence represented by the domain name.

Figure removed due to copyright restrictions

Figure 5.4 Web host strategy for a non-deployed domain name in 2005

The content of this placeholder screen indicates that this website that is 'coming soon' and the registrar populates the rest of the page with advertisements for their firm. This provides a clear example of the thinking regarding the decision to acquire a domain name in 2005. That is, that indeed a website would be 'coming soon'. As this paper will show, by 2006 this was no longer the assumption. When comparing this to Figure 5.5 one can discern an example of a newly emergent revenue strategy for displaying a non-deployed domain name purchased in July 2006.

Figure removed due to copyright restrictions

Figure 5.5 Web host strategy for a non-deployed domain name in 2006 now includes revenue generating links to external sites

It can be seen from Figure 5.5 that the registrar GoDaddy's strategy has changed. In its design pattern, this page is very much like the sites pictured in Figure 5.1. While the domain name provider still exhibits advertising specifically describing their services. However, in addition, note that the Web page now includes numerous category links leading to external sites (arrow A). Unlike the domain park specific services, the revenue generated from click throughs to the links listed on the domain name registrar's site is not shared with the domain name owner. In this case, the revenue is exclusively retained by the domain name service provider.

The emergence of the revenue generating domain name parking model has resulted in a new opportunity for revenue generation for domain name registrars. One which allows the registrar to offer 'free parking' for a domain name owner willing to let the domain name simply reside with the registrar. This model relies on the relative naivety of domain name owners who may not be aware of the option to park a domain name at a host site which is willing to share the revenues generated from such click throughs.

During the analysis phase, a subtle variant of domain names was noted. This was the domain name variant that was identified as purposeful misspellings of a generic name. In complex biological systems in nature such subtle mutations appear and they have been shown to lead to greater differentiation of species (Levin 2005). In the environment of the Web such mutations appear when strategic revenue generation models reward subtle differentiation and enhanced sophistication of business models. For example, Figure 5.6 depicts a website that demonstrates such differentiation – www.plamsatv.com.

The present study has found that many of the level two sites rely on 'direct navigation' in order for visitors to arrive at the site. Direct navigation refers to the function of physically typing in a Web Uniform Resource Locator (URL) or domain name in the address bar of a browser. Purposefully misspelled domain names represent a variant of level two sites, and such names frequently land visitors using direct navigation on a level two domain name parking website.

Figure removed due to copyright restrictions

Figure 5.6 This represents a subtle variation of a parked domain name. The owner relies on direct navigation and misspelling to generate traffic to this site.

The owner of the domain name illustrated in Figure 5.6 appears to have made a strategic decision to misspell the word 'plasma' in hope that this would result in visits to the parked domain name from direct navigation users who inadvertently misspell the term. Once on the site, the visitor might be expected to then click on one of the many links populating the site. As with most level two sites, the majority of the links would take to visitor to an e-commerce site which sells or services products similar to those implied by the domain name itself, in this case plasma television sets. The links which populate the parked site provide revenue to the domain name owner with every click-through. VeriSign (2004a) reports that the use of direct navigation techniques by Web users results in twenty million such misspelled domain names daily.

Level two business models such as these required a somewhat higher level of technical expertise than is required with the simple level one model, at least to the extent that the domain name owner possessed the knowledge required to advise the domain name registrar that a new domain name server would be used to resolve the domain name. If individual level two sites are to thrive and profit in the Web environment, the owner or the parking service need to apply additional skills. This is because, unlike a level one model, the domain name owner now has a vested interest in driving more client traffic to the site. Thus, marketing skills and the ability to use tools such as Google AdWords, which give the domain owner some control over directing client traffic to the site, become important to the success of the enterprise.

Level two models may seem trivial when considered as individual domain names or websites. However, as with many business models, when a domain name owner, whether an individual or corporate entity, accrues larger numbers of domain names, the effect of multiples begins to accrue significant revenue. For example, the research found that one of the most active enterprises internationally in the acquisition and deployment of parked domain names is a company in Brisbane. This firm suggests that domain name values should grow in proportion to the overall growth in both e-commerce and the Internet (Moore 2005a, p. 3). The firm has a domain name portfolio approaching 500,000 domain names, which generates revenue in the form of parked domain names of AU$ 22,950,000 (estimated 2005). The annual cost per domain name ranges from AU$ 8.00 to $12.00 amounting to estimated maintenance expenses of AU$ 5,400,000.

Complexity theory suggests that each business entity finds itself in an environment produced by its interactions with the other business entities in the system. This is clearly the case with these level two business models. Each entity is constantly acting and reacting to what the others are doing. And because of this fact, nothing in the environment is fixed (Waldrop 1992). This includes the evolution of new structures in the environment based on previous entities, and this is true with regard to the level two parking sites. Figure 5.7 represents the state to which level two sites had evolved by 2005.

Revenue is dependent on clicking on one of the links populating the site. Newer generations of these level two sites are evolving. As the purpose of and familiarity with the earlier generation becomes more recognizable, end users may become less inclined to click on the links. Almost simultaneously, new level two sites are emerging that more closely replicate the attributes of a site employing a more sophisticated business model. Figure 5.7 and Figure 5.8 are representative of this new generation of domain parking website. The parking service providers maintain a similar theme and just as in the earlier varieties, these sites are programmatically generated.

Figure removed due to copyright restrictions

Figure 5.7 Level two parking site representing first generation design technology

Note the consistency of design amongst the two sites. This is indicative of websites deployed by the same parking service provider. Despite some obvious differentiation, there is also an overall similarity in the design shown in figure 5.10, which is a page deployed by a competing parking site enterprise.

Figure removed due to copyright restrictions

Figure 5.8 Evolution of new level two parking site designs

Figure removed due to copyright restrictions

Figure 5.9 Similarity of the design marks this as a website from the same domain name parking service provider as shown in figure 5.8.

Figure removed due to copyright restrictions

Figure 5.10 Depicts new design models evolving in level two sites.

The research project had now reached a plateau where the situation once again suggested that it would be appropriate to examine the wider environment concurrently whilst interacting with the data, as recommended by Strauss and Corbin (1990, p. 43):

> Insight and understanding about a phenomenon increase as you interact with your data. … Often one idea or insight sparks another, directing you to look more closely at the data, to give meaning to words that seemed previously not to have meaning, and to look for situations that might explain what is happening here. This increasing sensitivity to concepts, their meanings, and relationships is why it is so important to interweave data selection with data analysis. Each feeds into the other thereby increasing insight and recognition of the parameters of the evolving theory.

It was now clear that the key to revenue generation (and thus the emergence of the new business model) was the ability to populate these parked pages with links to websites willing to pay for sending end users to their websites. But how was this accomplished? How were so many hosting entities able to aggregate enough advertisers to populate the vast array of parked domain names this research study indicated existed? What mechanisms were operating behind the processes powering this phenomenon?

Returning to published research on complexity theory and network theory provided guideposts to the answers these findings reveal. Complexity theory suggested the answer had to be relatively simple in order to engender such a high level of spontaneous organisation. This study had revealed that nearly fourteen percent of all websites in the sample had employed this novel but simple business model. Sir Tim Berners-Lee stated (2004, http://mitworld.mit.edu/video/236), 'Emergent phenomena occur due to the interactions between the elements of a system over time. Emergent phenomena are often unexpected, non-trivial results of relatively simple interactions of relatively simple components'.

5.4.4 The missing link: Advertiser aggregation entities and the impact of key word auctions

The business environment of the Web abounds in symbiotic relationships. From the perspective of the domain name owner, the revenue model for each of the websites

(now understood to be domain name parked sites) represented a symbiotic relationship with other external Web enterprises such as the domain name parking site, the advertisers with which domain name service populated the site, and the ad aggregator.

At this point the direction of the research turned toward understanding the key aspects of these relationships from a business perspective. This process was reminiscent of Addison's (1994) process related prescriptions.

- Hermeneutics seeks to uncover what is hidden and illuminate the background, and to find what is missing.
- Place interpretations on a larger social, cultural, historical or economic background. That is to say, researchers learn certain foreground conditions and explore the background conditions which make the foreground conditions occur.

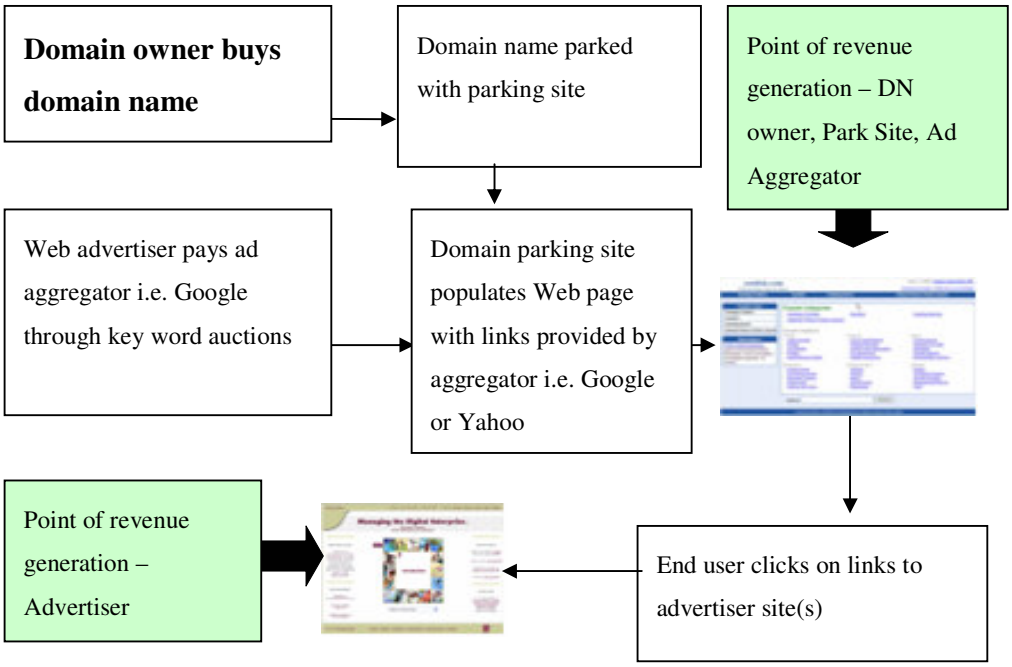
To investigate the economic background of this phenomenon, the research investigation turned to the 'whois' registry. Many of the domain name parking services were easily identified by their names on the Web pages generated for their clients, the domain name owners. Others were more difficult to identify and required multiple searches through multiple whois registries.

Access to corporate websites led to further generalised insights regarding such items as revenue distribution, money flows and revenue projections (Moore 2005a;, 2005b). Such information proved invaluable in piecing together a generic overview of the domain name parking business model. Figure 5.11, provides a graphic representation of these relationships.

In the initial steps in the process, that is the purchase of the domain name and the process for parking the domain name, the role of the advertising aggregator became clear.

Mason (2001, p. 410) writes, 'Emergence is a function of synergism'. The synergistic relationship between the four business entities that now are seen to comprise the level two domain name parking structure came into focus at this point: Firstly, the domain name owner; secondly, the domain name parking service provider; thirdly, the

advertiser aggregator; and fourthly, the advertiser. A fifth entity, the end user, powers the system. It is also of note that in Figure 5.11, the revenue generating action for three of the business entities is the end user's decision to click on the ad link which the advertiser paid to have placed on the site.



Source: Adapted from Sloan (2005)

Figure 5.11 A graphic representation of typical level two relationships

The advent of the search engine as a marketing partner is a relatively recent phenomenon and only became a factor in 2003 (Sloan 2005). Ad aggregators have been visible on the Web for some time, but a new species of ad aggregator came into being in the period of 2001-2002, and it is search engine based. The process has come to be known as 'search engine marketing' (Vise & Malseed 2005).

When search engine enterprises began acting as advertiser aggregators, they began to offer two distinct differentiating advantages. Firstly, they were able to open the Web advertising market to even the smallest of Web enterprises. Secondly, the search engines were able to offer the advertisers an introduction to a narrowly focused

demographic market segment explicitly oriented to the advertisers' product. These two advantages are further elaborated upon in the following sections.

Opening the Web advertising market to small enterprises

Search engine advertising aggregators achieved this by allowing the individual enterprise to bid on key words as a form of enabling Web ads. For example, imagine a domain name owner has purchased a domain name such as 'GardeningInAustralia.com'. The domain name owner now seeks to monetise the domain name by parking the name with a domain name parking service provider. Once the domain name has been parked, the owner may recognise the need to drive more end users to the site in order to increase the revenue potential. To best accomplish this, the domain name owner decides to register at a search engine site, develop an appropriate ad, and participate in a key word auction in order to determine the placement of the ad as well as the frequency of placement. The owner decides the most favourable key words for the site. For example, gardening, bush garden, garden supplies, flowers, and perennials.

The owner (now an advertiser as well) is willing to pay $ 0.07 every time someone clicks on the ad for www.gardeninginaustralia.com. The amount the owner is willing to pay will influence the visibility of the ad to end users on the Web. Some other large gardening companies may be willing to pay $1.00 per click to induce end users to visit their site. This means their ad would be placed higher on the search engine result site, and thus would have a greater likelihood that an end user would click on it. So, if an end user were to go to the search engine site, type in 'gardening' as a term to search, the ad for the large garden company might appear on the first page of the search result. The link to the advertiser's site might appear either as a link in the organic listing or as an ad in a separate but adjacent listing depending on which of the particular search engine ad aggregators the enterprise chose to associate with. Links located in the organic listing are referred to as 'paid inclusions' since the link to the advertiser's site appears included as just another search result. Links appearing as ads in a separate but adjacent area of the search results page are sometimes referred to as 'sponsored links'.

Similarly, the domain name owner's website may be five pages into in the search listing since the smaller entity chose to bid less on the key word 'garden'. Nevertheless, the ad will appear, and inevitably, some number of end users will find it and click through to it.

The fact that any website owner can bid any amount, from as low as one cent to as high as is reasonable to the owner, now means that any site can participate in a search engine driven marketing campaign. The owner has the ability to limit the daily budget for clicks as well. Thus the domain name owner might be willing to increase his bid to 10 cents per click, but limit the daily budget to a total of $1.00 of accumulated clicks.

This process is worthy of note because it completely reverses traditional advertising schemes. The advertisers or buyers now determine the price they are willing to pay to advertise their message, and as will be shown below the message is delivered to a targeted market. The ease of access to the process, and its 'do it yourself' character provides a significant advantage to small Web entrepreneurs who might not be able to afford to reach the widely distributed markets that television or radio advertising offers.

Narrowly focused, product oriented demographic

Search engines search the Web. Google, the leading search engine, was once an emergent phenomenon whose key differentiator was its ability to deliver relevant search results. Google was the first search engine to base its search results on the relevance of the key word or phrase for which the end user was searching. Prior to the ascendance of Google, search engines marketed the value of their search results as a function of the total number of links returned without regard to the relevance of the links to the key word search.

Google's key insight was to leverage the algorithms originally developed to deliver relevant search results ads to the search engine advertising market as well. Google then used the algorithms to also deliver relevant ad content to websites using Google's advertising model. For a client website such as 'www.gardeninginaustralia.com',

Google would typically populate the ad portion of a page to display ads related to some aspect of gardening.

Two elements influence this. Firstly, the parking service provider executes an agreement with the domain name aggregator to display the aggregator's (for example, Google or Yahoo) ads on their client's Web pages. The aggregator's Web crawl programs pick up elements such as the domain title, domain name, page content, and site meta-tags and use this to interpret appropriate ad content to be displayed. Secondly, the domain name owner often has the option to select descriptive key words to the domain name parking service provider, and these key words are also used to influence the content of the Web ads with which the page is populated.

In addition, the same algorithms that search engines use to determine the relevance and thus the rank of the content of the site in a normal search, can be used to determine the relevance and rank for marketing purposes. For example, in the case of Google, the amount the advertiser bids for a given key ad word is not the sole determinant of the placement of the ad nor the frequency of placement. If advertiser 'A" bid 50 cents for a given key word, and advertiser 'B' bid 35 cents, advertiser 'B' might still receive a preferred ranking if more end users had clicked on the ad for advertiser 'B' (Vise & Malseed 2005, p. 117) .

This process has the advantage of delivering a pre-screened demographic to the advertiser's site. The pre-screening mechanism is the search phrase the end user selected in the first instance. This is important, since the advertiser's opportunity for revenue generation only begins at the point at which the end user arrives at the advertiser's site.

This discussion provides a fitting introduction to the level three business model, the affiliate site. Like level one models, affiliate sites do not represent a new business model per se. Rather, the inclusion of affiliate sites helps to conceptualise two important theoretical elements, both of which are emerging factors in Web commerce. Firstly, these models help to clarify complexity theory's phase transition attribute, and secondly, the models embody a key element of network theory – preferential attachment. It is the symbiotic relationship offered by search engine marketing to

very small Web business entities that now powers a great deal of Web business growth, and contributes to the power law distribution the Web exhibits.

5.4.5 Example 3 - Simple 'level three' business models

Affiliate business models are not new to the Web. Nevertheless, the present study explores a new species of affiliate site which is emerging within the Web, i.e. pure affiliate sites. They do not sell products directly. The only function these sites perform is to serve as a referral site for larger Web enterprises. The goal seems to be to attract a community of Web users by whatever strategy, and then to monetise the user base by providing an array of search engine supplied ads. These sites are also differentiated from level two sites in that they frequently present branded ads in addition to search engine provided ads. These branded ads are usually contextually related to the site's content as well. This is due to the fact that they are individually selected by the entrepreneur or site owner for inclusion.

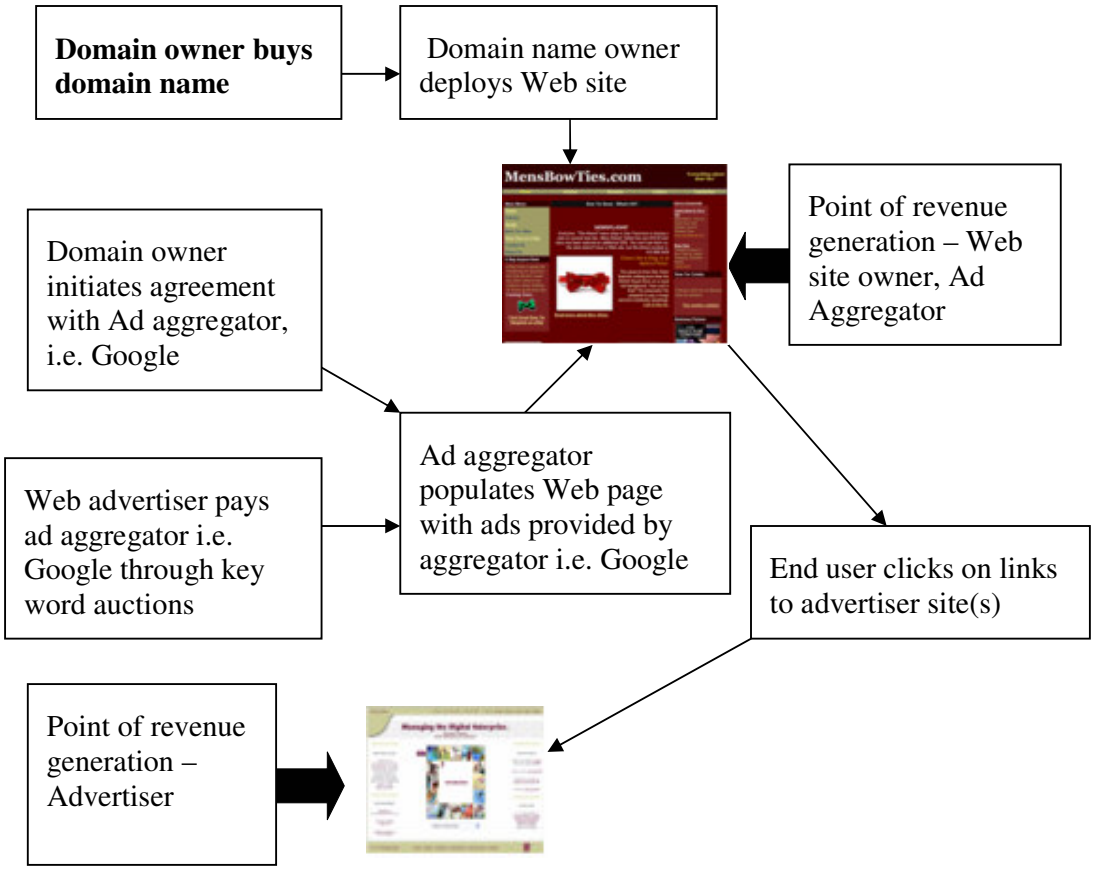
The environment in which affiliate sites operate in the first decade of the Twenty-first Century is indicative of an environment in a constant state of flux. This is partly because of the emerging symbiotic relationships which allow the most elementary level of websites access to powerful global marketing strategies (Vise & Malseed 2005). This suggests that level three business models are emerging in greater abundance within the business environment of the Web, due in part to the ease with which new business interactions with a host of higher level organisations may be established. Many of the largest organisations on the Web, such as Amazon, E-bay, and Google encourage interaction among even very simple and lower level organisations by making code which links to their sites readily available to them. This symbiotic relationship provides opportunities for the new entities to grow and prosper, and it also provides a source of inbound links to the larger organizations. These links represent a key element in the continued growth of the Web behemoths, particularly since inbound links are an element in determining site rankings in search engine organic result listings (Brin & Page 1998).

Levin's (2005, p. 1077) description of ecosystemic patterns is both analogous and instructive:

> Ecosystems and the biosphere are complex adaptive systems, heterogeneous assemblages of individual agents that interact locally and that are subject to evolution based on outcomes of those interactions. …
>
> The patterns that characterize ecosystems – the distributions and abundance of species, and their spatial organization, size structure distributions, and the patterns of nutrient use [for example, revenue in economic systems] – all can be realized as emergent from the selection forces operating at much lower levels of organization, not for the benefit of the whole system, but within the framework of well-established principles of evolutionary change. Indeed, the methods are so powerful that they are also providing insights into the organization of societies and economies in terms of actions and rewards …

Level three business models represent a step up the Web business evolutionary ladder for domain name owners. Unlike level two models, level three sites require a Web presence in which the domain name owner provides for both the design and content of the Web page. Level three business models take the concepts employed in level two models and place them fully in the hands of the domain name owner and in return these models present the opportunity for enhanced revenue generation. Owners of level two models rely primarily on pay per click revenue sources provided by an external resource – the domain name parking service, by way of the search engine ad aggregator. Owners of level three sites interact directly with the ad aggregator and this creates a different revenue generation dynamic. This model is one in which the owner has much more control over the style of ad content and ad placement on the site itself.

Level three models move the study away from sites which were originally classified as 'redirects' and into the realm of sites the study would have originally categorized as 'business sites'. But these may still be referred to as simple business models, because level three sites do not engage in commerce per se. That is to say, level three sites do not incorporate shopping carts or retail sales on the site pages themselves. Rather, level three sites generate revenue from direct affiliations with the larger enterprises. Figure 5.12 provides an overview of the revenue relationships of level three affiliate sites.



Source: Developed for this thesis

Figure 5.12 Graphic Representation of Level Three relationships

The level three business model's goals are also to create symbiotic relationships with these larger e-business enterprises. The level three sites generate traffic and deliver a demographic expressly targeted to the e-commerce enterprise site's content. In return, the enterprise sites provide the code and even the graphic images needed to display the links to the target enterprise site. Affiliate sites also lead into the realm of business models more commonly identified in the e-commerce literature. Figures 5.13 and 5.14 present examples of affiliate sites.

Site design can vary significantly on level three sites, much more so than in level two sites. This is explained because in level two sites, the domain owner has little input into site content or design. The service provider handles these details on level two websites. The process of populating the page with content is usually automated, with

the result that many of these level two sites present a very similar appearance on the Web. This is the case even though they may represent several different domain name service providers.

Figure removed due to copyright restrictions

Figure 5.13 An affiliate site represents a higher level of sophistication in that the design thoroughly integrates the function of the site which is to encourage visitors to click on the ads listed.

In contrast, level two sites, like amoebic life forms, have a very simple function. They exist solely to populate a standard Web page with links of pay per click enterprises for site visitors to click on.

The level three sites also represent a significantly more diverse pattern of revenue generation, including pay per click, and pay per sale commission structures. Thus, greater diversity in level three websites might be expected, and this appears to be the case. For example, in design, colour palette, style, and ad integration, the site in Figure 5.14 differs from the site shown in Figure 5.13.

Level three simple business models can offer an infinite variety of site and content design ranging from silly and amateurish to edgy and professional. They exhibit varying levels of community content from primitive to erudite, but all have in common the fact that they do not create a product, but rather attempt to aggregate a community of users in order to insure long term survival of the enterprise revenue stream.

Figure removed due to copyright restrictions

Figure 5.14 shows that level three sites feature much greater variability in design than do level two sites

They also have in common, to a greater or lesser degree, the inclusion of search engine marketing provided ads. Figure 5.14 clearly defines the search engine marketing provided ads as the light coloured (beige boxes) with blue text. These ads are relevant to the site content and are integrated into the site by the site owner (enterprise or individual) with code fully provided by the search engine ad aggregator. The code is integrated into the HTML code of the site itself. Once this code is integrated, it allows the ad aggregator to populate the site with relevant ad content. The site (or domain name) owner then generates revenue with every click through to an advertiser site.

The findings of the present study suggest the inclusion of search engine marketing ads marks a phase transition amongst many sites from a merely personal site, a hobby site, or even a blog, to a legitimate, though elementary business structure. Once these search engine provided ads appear on the individual websites, the process of preferential attachment and the incumbent power law distribution are inevitably reinforced. This is because the placement of ads is determined by one or both of two elements. The first element is the advertising budget of the ad purchaser. Larger organisations are more likely to bid more in the auction for key words and thus their ads are likely to appear on more sites incorporating the ads leading to more links to their site; the second element is the algorithmically determined relevance of the ad client's site. A key element in the algorithms that determine relevance for search engines, in fact, is the number of inbound links (Brin & Page 1998). And so, as more links are generated from the ads, the relevance of the advertisers site is also enhanced. This process fuels the 'preferential attachments' to the larger sites which, in turn, continues to reinforce the power law distribution already evident on the Web.

Still more sophisticated skills and knowledge are required of level three domain owners since they must now interpret site analytics, the measurements which inform the domain name owner as to who is visiting the site and with what frequency. They must also possess even more advanced marketing skills and if site traffic grows, eventually manage others in a growing e-commerce network of businesses.

5.4.6 Summary of simple business models

The dynamic nature of conducting business on the Web is important to anyone who utilizes this medium. It was posited in this chapter that new business models are evolving in this environment. Using Emergence Theory as an explanation of this development, it is suggested that there is an evolutionary nature to the development of the simple business models found on the Web.

The findings suggest a hierarchical order within these models. For example, an individual requires little skill to participate in an e-business venture at the simplest level and the risk involved is minimal. An awareness of the evolving models might lead an individual to become a level one Web entrepreneur. At the second level, one might require an advanced beginner skill set. This would include the ability to

transfer the domain name(s) to a selected parking host site. If the owner of this level of business wants to increase the domain name profit potential, they would need some marketing, advertising, and market segmentation skills to drive business to the site. More sophisticated skills are also required because the entrepreneur assumes responsibility for site design and content, and there is the need to establish linkages with successful Web businesses in an attempt to generate additional revenue and earnings.

However, an important point from an entrepreneurial perspective is that even at this level, the new venture risk is minimal, while the new venture reward potential is high. This turns upside down the old adage, 'the greater the risk, the greater the return.' The owner has the potential of making a profit with a minimum investment of $8.00 to $10.00 for the initial domain name purchase and minimal annual web hosting cost of $48.00 at level three. This opportunity to have a Web presence as a potential business owner for virtually no investment, no accountant, and no attorney is a new phenomenon. Following the evolution of these simple business models over time should help explain entrepreneurial aspects of the Web and suggests a path for future research. The next section relates these findings to the research propositions introduced in Chapter 3 and discusses the role of both complexity and network theories in the process.

5.5 Relating research propositions to the findings

Research propositions of this thesis were introduced in Chapter 3 at appropriate points in the review of literature, and they were further discussed in Chapter 4. This section describes findings regarding the propositions and also presents an opportunity for further discussion of the properties of the simple business models this study introduced.

5.5.1 The first research proposition

P_1:	Complexity theory correctly predicts the emergence of new and coherent structures, patterns and properties (e.g., online business models) in a complex system such as the Web.

This proposition is affirmed. The present study has shown that new business models are emerging within the business environment of the Web.

The first indication of this was the identification of the anomalous websites as parked domain names hosted by sites specifically developed for the purpose of parking domain names, referred to as 'level two' business models. These sites were able to generate revenue for domain name owners by establishing symbiotic relationships with advertising aggregators, and displaying these ads on specially constructed websites. The sites were dynamically constructed using programs and design algorithms provided by the host company. The domain name owners utilising this business model frequently depend on direct navigation to attract end users to these sites and so, typically, try to employ a strategy which features either a very generic domain name, such as 'kitchenfurniture.com' or a common misspelling of a popular term such as 'voyuer.com' or 'plamsatv.com'.

Further, the research suggests there is a good probability that these new business models, in turn, generate even more new Web business models which may serve a support role to the emerging models discussed in Chapter 5 or which leverage these models to form evolutionary or generational deviations. Though the titles may be similar to previously identified business models, these business structures represent significant species differentiation. The monetisation of domain names through the expedient of the level two business model discussed above is a phenomenon which has gone relatively unreported in the scholarly journals. The importance of the phenomenon is the fundamental role it plays in the business ecology of the Web.

The new business models employ elemental mechanisms, such as search engine marketing advertising links which are programmatically imbedded in the content of the site; or by direct branded links the owner chooses to include on site pages. These links act as vehicles to drive a narrowly focussed customer base to larger websites which pay for the links and for the traffic the links generate. The links themselves become the devices which influence the business structure of the Web. The structure itself is discussed relative to propositions three and four below.

5.5.2 The second research proposition

P_2:	Complexity theory correctly predicts that emergent elements (e.g., Web business models) are not necessarily predictable from analysis of an individual part of the system (e.g., a search engine output).

This proposition is affirmed. While qualitative analysis of anomalous websites encountered in the quantitative phase of the research project provided a greater understanding of the nature, and revenue generative capacity of the new business models, it was not until the hermeneutic cycle, and its dependence on the integration of knowledge of the part, with knowledge of the whole, that understanding of the full ramifications of the emergent phenomena became clear. Each cycle of understanding led to further questions regarding the constitution of the entire ecosystem. For example, the revenue generating capacity of the parked domain names for the domain name owner was not discernable until further research was conducted on the revenue mechanisms employed by the domain parking service provider. This led to yet another recently emergent Web entity, the search engine ad aggregator.

The emergence of the search engine aggregator might not necessarily have led to the prediction of the emergence of the level two websites. Originally these entities (such as Google) utilised the concept of ad aggregation as a resource for internal revenue generation. The idea of partnering with other entities such as the parking site service providers or the myriad of entrepreneurial affiliate (level three) sites came later (Vise & Malseed 2005).

5.5.3 The third research proposition

P_3:	As a complex system, the Web is defined by relationships between component websites.

This proposition is affirmed. The interactions between the various elements of the business environment of the Web define the business environment itself. Only once the complex interactions between business entities within the system were examined did the role of the original anomalous websites begin to emerge. These new business models which describe the newly emergent level two and level three sites are nearly

entirely reliant on linkages to larger entities on the Web which actually produce products and sell them. These are the 'brand names' of the Web. The smaller entities are more likely to link to these sites for two reasons. Firstly, the level two or level three site's clients are likely to recognise these brand name entities, and secondly, the search engine ads are placed on the site based on content compatibility. This assures that only the more relevant of the larger sites are the ones whose ads are presented.

Thus, rather than competing with established business models that the larger sites employ, the continued emergence of the simple models has a positive symbiotic impact on the larger entities. These simple business models form a series of symbiotic relationships, not unlike those seen in the natural biosphere. These relationships are dynamic because they are often programmatically established, and the nature of the program dictates that the business entity willing or capable of bidding the highest amount for contextually relevant key words has the best chance of having their links imbedded within the simple business model sites. Therefore, it may be stated that the continued emergence of these simple business models has the impact of fuelling the growth of the larger business entities.

5.5.4 The fourth research proposition

P_4:	Network theory successfully predicts a power law distribution of inbound links for business entities on the Web

Some suggest that, while the Web itself exhibits a power law distribution, the distributions of a power law configuration on the Web is not inevitable in all domains or in specific sub-categories of top level domains. For example, Pennock et al. (2002, p. 5207) found

> '…a richer structure [of inbound links] among subsets of Web pages in the same category. We find that these category-specific distributions exhibit very large deviations from power law scaling, with the magnitude of deviation varying from category to category'.

Among the categories they classified were 'company home pages', a category which would be inclusive of the category 'business sites' used in this study. As mentioned earlier in this chapter, the 'business sites' category would also be inclusive of level

three simple business models – the affiliate sites. Pennock et.al. selected companies from an investor guide website listing company's by stock categories (2002, p. 5208). This suggested that further inquiry was required to affirm proposition six, since this study contradicts certain elements of Barabási and Albert (1999). It was determined that for the purposes of affirming or disconfirming Proposition Four, it would be instructive to ascertain the number of inbound links from the 441 sites categorised as 'business sites' in Table 5.1 earlier in this chapter. From this analysis it would be possible to determine if a power law distribution resulted, as Barabási and Albert predicted, or whether a more normative distribution would result, as suggested by Pennock et. al (2002).

The 441 samples examined did not yield a truly large business of the magnitude of an Amazon, IBM, Dell or e-Bay. This fact suggested that a power law distribution might be at work in the .com zone, since a power law distribution would predict few very large sites given the size of the sample examined relative to the size of the entire .com zone. In order to determine whether business sites in the .com zone reflected a power law distribution, 436 of the sites categorised as 'business sites' were subjected to an inbound link analysis (in the other five instances the researcher was unable to establish a link analysis connection).

The inbound link analysis was performed utilising the software program 'Link Survey (version 1.4)' which performs actions similar to a Web crawler. One enters the domain name or uniform resource locater for a given website or a series of websites. The software then performs a survey of all sites linking to the domain name(s) of interest. These are referred to as 'Inbound Links'. The resulting data counts were entered into an Excel spreadsheet and SPSS was used to analyse the data. The data in Figure 5.15 were the result of a simple scattergram of the number of websites which exhibited a given number of inbound links.

Figure 5.15 displays the numbers of inbound links of the sites categorised as business along the Y axis. The X axis displays the count, or the total number of business sites which had the corresponding number of inbound links. The average number of inbound links for the business sites was 155, and the median was 76.

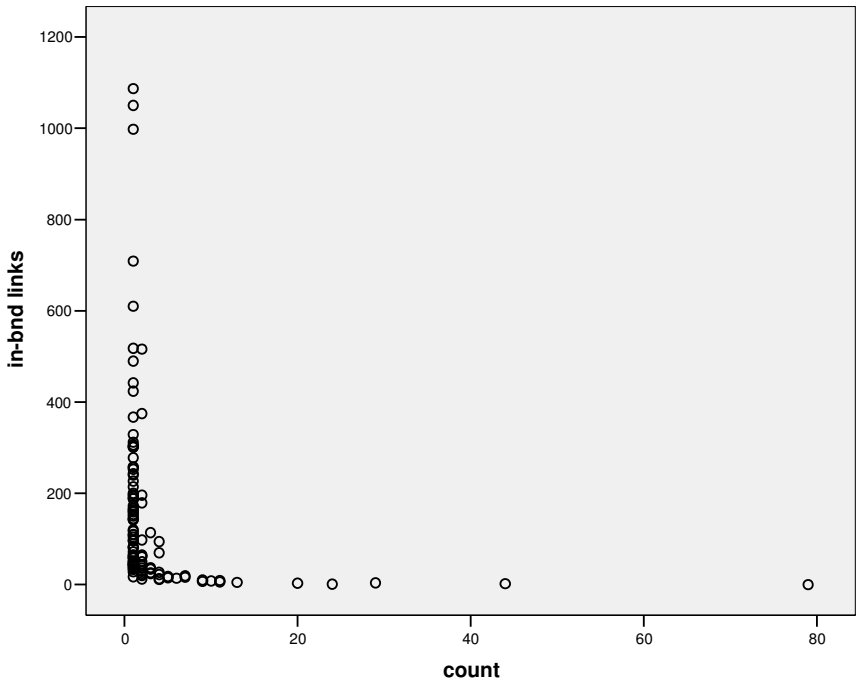


Figure 5.15 Linear representation of distribution of inbound links to websites categorised as 'Business' in the .com zone

The results of the analysis represent a nearly perfect power law distribution. This is confirmed by comparing Figure 5.15 to Figure 3.2 in Chapter 3 page 45. The power law distribution is in contrast to a random distribution of links, which would be expected to exhibit a normal Poisson distribution. The attributes of a power law distribution are these: Firstly, a very few entities that rank very high in the number of inbound links; secondly, a large number of entities with close to median numbers of inbound links; and lastly, a very large number of entities with few inbound links, referred to as the 'wide tail' or a 'fat tailed' distribution, are evidenced.

Power law distributions are often presented as logarithmic scales as is shown in Figure 5.16. This is because the linear representation (Figure 5.15) tends to hug the axis on a linear scale (Nielson 1997). When Figure 5.16 is compared to Figure 3.3 in Chapter 3 page 45, the power law distribution is further confirmed.

This result affirms proposition six and also affirms Barabási and Albert's (1999) assertion that the Web represents an example of a power law distribution. They

suggest that a reason for this type of distribution is the phenomenon of preferential attachment. This is addressed further in the discussion of the next proposition.

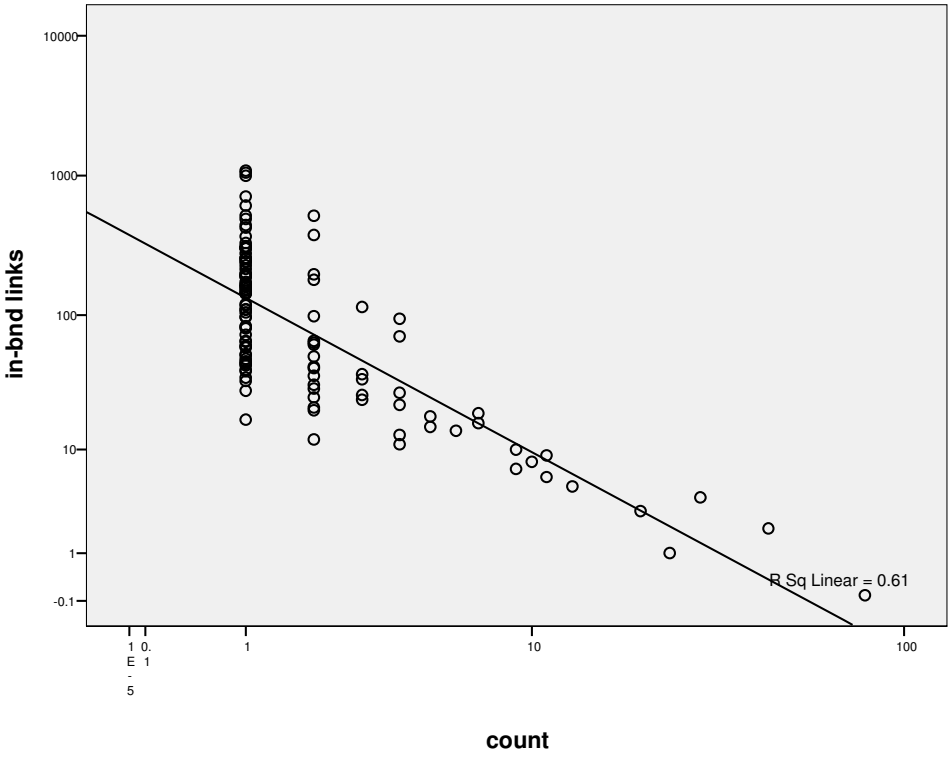


Figure 5.16 Logarithmic scale of distribution of inbound links to websites categorised as 'Business' in the .com zone

5.5.5 The fifth research proposition

P_5:	Network theory correctly predicts that the components of the Web exhibit preferential attachment.

Most networks grow by the addition of new nodes to the network. Nodes on the Web may be represented by domain names or by pages. Preferential attachment naturally leads to a power law distribution. Preferential attachment and the resulting power law distribution are a result of the self-organising capacity of the Web (Barabási & Albert 1999; Pennock et al. 2002). For example, a newly created website will be more likely to include links to well-known popular sites which already have a high number of in-

bound links attaching to their sites (Barabási & Albert 1999). Barabási and Albert initially suggested that this means older sites increase their connectivity at the expense of younger ones, however, Adamic and Huberman (1999) disconfirmed this by their research which demonstrated no correlation between age of a site and the number of inbound links it has.

To the extent that power laws are a result of the phenomenon of preferential attachment, the present study confirms preferential attachment for business entities on the Web (see figure 5.15, for example). However, exactly how preferential attachment occurs remains unknown, although this study suggests that the revenue generating mechanisms of these emerging business models, such as programmatically derived conjunctive links imbedded by search engine market aggregators, may play an important role in the preferential attachment phenomenon. This suggests an area for future research.

5.5.6 The sixth research proposition

P_6:	New forms of business websites that commenced in the online environment fall into identifiable classifications in each of the 'pure dot com' and 'almost pure dot com' categories (this extends Adam 2002).

As Figure 5.17 below illustrates, this thesis adds an additional dimension to Adam's 'spectrum of business users on the Web' (see Figure 3.1 on page 29). As a result of the research, one can now observe that there has been development across two planes rather than one. On the horizontal plane, there is the representation of traditional businesses that have embraced the Web and presented themselves in both the physical world and on the Web. At the other end of the horizontal continuum, new business models are in evidence.

Adam (2002) suggested that when examining websites, it is possible to place them into the categories shown on the horizontal continuum in Figure 5.17. There are those that may be termed pure dot.coms (e.g., infomediaries such as eBay.com), almost pure dot.coms (e.g., sites such as Amazon.com which still distribute in the physical world), those that exist in both domains (e.g., Borders bookstores) and at the other end of the

continuum, those that have yet to integrate the Web into their business processes to any great degree, if at all.

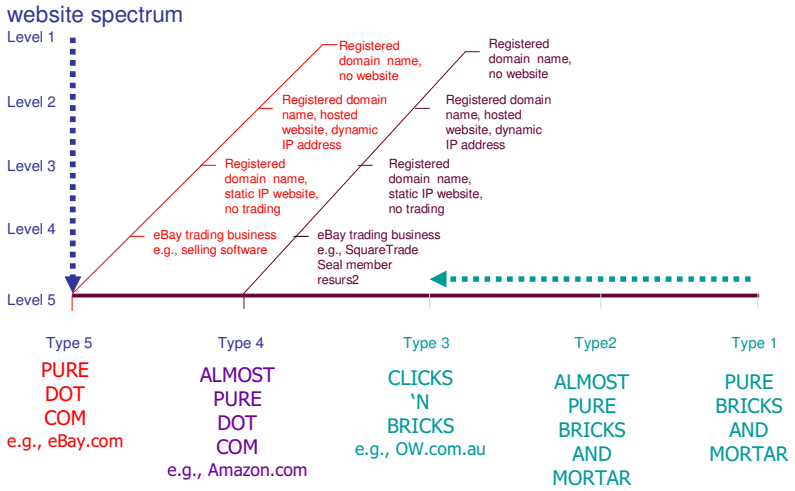


Figure 5.17 Suggested extension to Adam's spectrum of business users of the Web

Source: Adapted from (Adam 2002, p. 263)

The new dimension suggests an incremental level of sophistication in the implementation of Web business models. It varies from extremely simple business models that required neither technological nor business acumen to any great extent, to more sophisticated models which may incorporate warehousing and the sale of products.

5.6 Conclusion

This chapter discussed the findings of the research project in two distinct phases, the quantitative research and the qualitative phase. The original purpose of the quantitative data was to give high level business demographic information, however the data resulted in the identification of numerous anomalous websites that appeared to be revenue generating in nature, but which not easily classifiable into extent

business model typologies. This lead to an extensive qualitative phase of the research which identified these entities as relatively simple business models utilising the most elementary building blocks of the Web business environment – domain names. The chapter described the processes by which these domain names were monetised. Further, it was suggested that these models form a hierarchical classification scheme, ranging from level one – the mere acquisition of the domain name for resale, through level two models which rely on complex symbiotic relationships within the Web business environment, and finally to level three models resembling more commonly recognised extant business entities.

Next, the chapter related the findings to a series of propositions which were first introduced in Chapter 3. These propositions dealt with the application of complexity theory and network theory to the findings of the study. The chapter illustrated important attributes of both complexity theory and network theory which were found amongst the business models evolving in the business environment of the Web and further, it illustrated that these theoretical frameworks provided a sound basis for examining the emergence and evolution of Web business models.

The next chapter is the concluding chapter of this thesis. It provides a listing of the contributions this thesis makes to Web business theory and practice. It also discusses the limitations of the thesis and provides suggestions for further research projects related to the evolution of business models in the business environment of the Web.

> *If the Internet is indeed viewed as a complex adaptive system, then the ways in which it might grow and evolve are certainly fascinating areas for exploration (Valovic 2000, p. 39)*

CHAPTER 6 – RESEARCH CONTRIBUTIONS, LIMITATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH

6.1 Introduction

This chapter concludes the thesis. Previous chapters have discussed the methodologies employed, the findings, and implications of the findings. This chapter discusses the contributions this study makes to both practice and theory, limitations of the study, and finally offers suggestions for future research.

This study aimed to provide a deeper understanding of the conduct of business on the Internet and its graphical face, the World Wide Web. It accomplished this by reducing the function of Web business to its simplest element, the domain name, and by applying complexity theory as a conceptual framework to enhance understanding of the emergence and evolution of new methods of conducting business on the Web. Additionally, the study explored and explained the fundamental processes underlying business on the Web. In doing so, it contributed to both theory and practice.

6.2 Contributions of the research

In summary the research reported in this thesis makes the following contributions:

Firstly, it identifies new business structures which have evolved from within the environment of the Web.

At the simplest level these structures merely involve the acquisition of a domain name. It is in an effort to monetise this acquisition that new structures, such as domain name auctions and parking services, have evolved. That these business structures have been overlooked in Internet business literature is, arguably, not surprising. At more complex levels this study identifies the existence of new business models that provide extensive support for these elementary models, such as domain name server hosts, which provide revenue generating 'parking services'.

Secondly, it proposes a simple classification scheme for these business structures.

A three tiered classification scheme is proposed for these elementary new business models: Level one, domain name acquisition for resale, level two monetisation of acquired domain names by employing a domain name parking service provider; and level three, the deployment of affiliate websites which obtain revenues solely through the process of referral to higher level websites.

Thirdly, it identifies these simple business structures as a key driver of the overall growth of domain name acquisition on the Web.

Since this thesis was begun, the total number of domain names acquired in the .com domain alone has risen to 46 million (Rhoads 2006). It is reiterated that over one-million domains names are sold in the dot com zone alone each month. The use of online advertising as described in Chapter 5 and the ease of implementing parked domain name websites are identified as key drivers in the growth of domain name sales. This confirms VeriSign corporation's conjecture that the increasing acquisition of domain names is primarily driven by new methods of monetisation emerging in the Web environment (VeriSign 2005). It is not the rush of external businesses seeking to adopt their business models to the Web that explains the phenomenon. Rather, it is the level one and level two (and to a somewhat lesser extent, level three) business models which contribute an explanation for the surge in the acquisition of domain names. These elemental business structures are still emerging, evolving and extending the business environment of the Web.

Fourthly it identifies a methodological process for random sampling of Web top level domains.

The study reported in this thesis introduces a methodology for randomly sampling the Web by top level domain. The suggested process involves obtaining access to the domain or domains of interest through the top level domain manager. A program (Appendix A) was developed which used a quasi-random number generation process to extract domain names from the entire domain listing provided by the domain manager. This step allowed the creation of a data set of manageable proportions in which common tools such as Microsoft's Excel or SPSS could be used to further analyse and manipulate the data.

Fifthly, it provides confirmation that business websites form a power law distribution.

The study provides further confirmation of Barabási and Albert's theory that the Web represents a power law distribution. In particular, this study confirms the power law distribution of business websites. This is important because Pennock et al.'s (2002, p. 5210) research suggests that some categories of commercial sites do not follow a pure power law distribution, but a show a distribution 'mostly dominated by the "rich get richer" process of preferential attachment' which resulted in a distribution representing a partial power law. The research reported in this thesis suggests that the commercial sites sampled represent a nearly pure power law distribution as described by both Barabási and Albert (1999).

Finally, the present research provides an example of the application of complexity theory to the realm of business within the ecosystem of the Internet and the World Wide Web.

This application is not simply a metaphorical one to be applied to the management of enterprises. In this instance, Complexity Theory provides an explanation for the very existence of these simple business models which is consistent with the theory's application across other scientific boundaries.

Further studies of these simple business models are likely to reveal ever increasing numbers of mutations of these basic forms and thus an increasing variety of ever more complex 'sub-species'. The patterns that led to the identification of these simple business models were derived from research conducted within the environment of the Internet. Attempting to discern these models by viewing businesses external to the Web would have been a difficult proposition. This thesis suggests that continued analysis of these simple models is best conducted from the perspective of the environment in which they thrive.

Perhaps most importantly, this study has confirmed the utilisation of complexity theory and network theory as valid theoretical frameworks to describe the emergence and evolution of new business structures, as well as providing explanatory theory for the growth of business on the Internet.

6.3 Research findings and the literature

This research refutes many of the studies which claim that small business is underrepresented on the Web (Anonymous 2001; Dilts & Kahai 2004; Dini 2003; Lidsky 2003; Lockett & Brown 2003; Nachira 2002; Pool, Parnell & Carraher 2004; Walczuch, Braven & Lundgren 2000). The implication drawn from these articles was that wider use of the Web by the population of small businesses would result in a greater probability of success in growing the business, or would lessen the incidence of failure among small businesses in general.

While these suppositions might be true in some geographical areas, this study suggests that they have only been shown to be true for small businesses that were external to the Web. These businesses would typically fall into Type 1, Type 2, or Type 3 categories of Adam's website spectrum (Figure 5.17).

This study has shown that there exists a whole class of small businesses and entrepreneurs using the Web that were not included in the extant literature. It also suggested that one reason they may have been overlooked was that unless one actually drew a sample of website entities from the Web environment, these businesses would be very unlikely to have been 'discovered' and thus could not have been included in studies purporting to document the proportion of small businesses conducting business in the Web environment. In other words, these small businesses were 'not counted' because the only way one could count them is to draw a representative sample of sites on the Web.

This study demonstrates that there is, in fact, a robust and rapidly growing community of small businesses and/or entrepreneurs who are using the Web and developing innovative business models. The findings further suggest that the proportion of small businesses reported as 'not having a Web presence' may be somewhat exaggerated simply because previous studies have failed to detect the large number of these simple business models operating just below the surface of the commonly visible Web.

6.4 Limitations of this research

When this study began, an initial goal was to as to analyse the conduct of business on the entire World Wide Web. However, initial sampling issues limited the reach of

this study to two domains and the emphasis of the study was limited to one domain, the .com zone, since it represents the largest top level domain on the Web. Since the study was accomplished using data primarily from the .com generic top level domain, one may not generalize these findings to other generic top level domains such as .info, .biz, .name, or country code top level domains such as .au, .uk, or .tv.

It is of note that earlier research has employed Web crawlers which reach the entire Internet (Barabási 2002; Barabási & Albert 1999). These studies concentrated on reporting all Web pages. For example, a large corporate entity may have literally tens or even hundreds of thousands of pages within its website domain. Since the sample for this study was drawn from a database of all domain names in the .com zone, this means that each Web site is generally represented by one page only – its home page. Thus, a site for a corporate giant such as IBM had the same chance as being drawn for the sample as the tiniest one page site in the .com zone. This may be viewed as a limitation, or it may represent an advantage depending on the orientation and goals of the particular business research study.

6.5 Suggestions for further research

This study is a relatively early attempt to better understand the business environment of the Web and has such it was expected to generate many questions regarding the conduct of business in this complex environment.

What is interesting to note is that affiliate sites may represent, by far, the largest category of e-commerce business model. This represents one area where further research is suggested. For example, a study examining the numbers or proportions of simple affiliate sites may shed additional light on the entrepreneurial aspects of Web businesses.

It would be instructive to establish whether significant differentiation exists among the .com, .net, .org, zones. Similarly, it seems particularly important to examine differentiation amongst country code Top Level Domains such as .com.UK, .com.AU, .com.DE and so on as this might lead to important insights as to the intensity of entrepreneurial activity amongst these TLDs.

The line continues to blur between what is and what is not a business in the traditional sense. Blogs (a contraction of the term Web Logs) are the most recent phenomenon on the Internet and their growth is the subject of much speculation in the popular media. This is a fertile area for further research. For example; this study mentions the phase transition represented by the inclusion of search engine ads in affiliate and personal websites, and which explains a large portion of the growth in the acquisition of domain names. One might surmise a similar phenomenon could partially explain the rapid growth in the number of blogs. Further research along this line is called for, especially since such a phase transition enables the monetisation of an otherwise content-only site.

Further studies of indigenous Web business models are likely to reveal ever increasing numbers of mutations of these basic forms and thus an increasing variety of ever more complex 'sub-species' as witnessed by the recent proliferation in 2006 of pixel level ad websites.

It is suggested that some of the revenue generating mechanisms of these emerging business models, such as programmatically derived conjunctive links embedded by search engine market aggregators, may have an important role to play in the phenomenon of preferential attachment. This topic represents an important and potentially productive topic for future study as well.

6.6 Conclusion

The dynamic nature of conducting business on the Web is important to anyone who utilizes this medium. This is especially true since, increasingly, the term e-business is becoming synonymous with the term business. This research study has posited that new business models are evolving in this environment. Identifying complexity theory as an explanatory framework applicable to this phenomenon, it suggests that there is an evolutionary nature to the development of the simple indigenous business models found on the Web. They have emerged from within the Web, they are self-organized, and they are self-adaptive. They represent the 'long tail' of the power-law distribution ascribed to the Web by Barabási and Albert (1999) and elaborated upon by Anderson (2004)

Analogous to life on our planet, the simplest business models are the most abundant and prolific. Simple models are the most likely ones to emerge in any new business environment, and like the microscopic life forms they emulate, many of these simple 'one cell' business models thrive just under the surface of our unaided vision, requiring a magnified view to fully appreciate their abundance, or to understand their revenue generating capacities, or to discern the important functions they serve relative to larger business entities. As these business models continue to evolve and increase in complexity, they become ever more visible and visibly more diverse.

Similar to biological environments, the simple indigenous business models identified and discussed in this study, depend on symbiosis with higher level enterprises for their continued existence. Without domain name registrars, the level one model would cease to exist. Without more sophisticated domain service providers level two models would not exist. Without large Web enterprises which offer both pay per click and pay per sale revenue streams level three models would not exist.

The continued 'viral like' growth of these simple but progressively complex business forms will have an impact on existing large business enterprises within the environment of the Web and they will provide entrepreneurial opportunities for new generations of business professionals. It is only by studying these phenomena from within the context of the Web itself that we can hope to discover and understand the new forms of businesses emerging there.

Kuhn (1962) asserted that scientific advancement is not a smooth continuous process but instead it is frequently punctuated by radical intellectual or technological revolutions. Business is in the midst of just such phenomenon, labelled by Kuhn as a new paradigm. The impact of the Web on culture, society, and business is not yet clear to the research community. But what is true of all human endeavours is exponentially true for the Web. Reality, meaning and knowledge change over time. Websites – like organisms and humans – are constantly adapting to new situations in a relentlessly changing environment. What is true about this study today will almost certainly be less true tomorrow. This fact makes future Web research both exciting and frustrating... and as inevitable as change itself.

BIBLIOGRAPY

Anonymous 2001, *The Other Digital Divide: E-commerce and Australian Small Business*, SETEL Small Enterprise Telecommunications Centre Limited [Online], Available: http://www.setel.com.au/site.php?id=79, [Accessed April 30, 2006].

AACSB 2002, *Management Education at Risk: Report of the Management Education Task Force*, St. Louis, AACSB, pp, 1-31. http://www.aacsb.edu/members/metf/default.asp.

Adam, S. 2002, 'A Model of Web Use in Direct and Online Marketing Strategy', *EM - Electronic Markets: International Journal of Electronic Markets* [Online], vol., http://www.electronicmarkets.org/modules/pub/view.php/electronicmarkets-418, Accessed 4

Adamic, L. A., and Huberman, B. A. 1999, 'Technical comment to 'Emergence of Scaling in Random Networks'', *Science*, vol. 286, no. 5439, pp. 509-512.

Addison, R. B. 1994, Balint Leadership and Hermeneutic Research: Proceedings of the 9th International Balint Federation Congress, Charleston, South Carolina.

Anderson, C. 2004, 'The Long Tail', *Wired* [Online], vol., http://www.wired.com/wired/archive/12.10/tail.html, Accessed 10

Argyres, N., and McGahan, A. M. 2002, 'An Interview with Michael Porter', *Academy of Management Executive*, vol. 16, no. 3, pp. 43-52.

Baker, M. J. 2000, 'Writing a Literature Review', *TheMarketingReview.com* [Online], vol., http://www.themarketingreview.com/browse.cfm, Accessed 2

Barabási, A.-L. 2002, *Linked: The new science of networks*, Perseus Pub., Cambridge, Mass.

Barabási, A.-L., and Albert, R. 1999, 'Emergence of Scaling in Random Networks', *Science*, vol. 286, no. 5439, pp. 509-512.

Barnard, C. I. 1947, *The Functions of the Executive*, Harvard University Press, Cambridge.

Bauer, C., and Scharl, A. 2000, 'Quantitive evaluation of Web site content and structure', *Internet Research*, vol. 10, no. 1, pp. 31-43.

Berners-Lee, T. 2004, *Opening Keynote - The Semantic Web*, MIT Technology Review [Online], Available: http://mitworld.mit.edu/video/236, [Accessed June 7, 2005].

Berners-Lee, T., and Fischetti, M. 1999, *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*, Harper, San Francisco.

Biswas, D. 2004, 'Economics of information in the Web economy: Towards a new theory?' *Journal of Business Research*, vol. 57, no. 7, pp. 724-733.

Boland, R. J. 1985, Phenomenology: A preferred approach to research on information systems, *in Research Methods in Infomation Systems,* E. Mumford, R. Hirschhiem, G. Fitzgerald, and T. Wood-Harper, eds.: New York, North-Holland, pp. 193-201.

Bollier, D. 1993, *The Promise and Peril of Emerging Technologies: A report on the second annual roundtable on information technologies*, Aspen, The Aspen Institute.

Brin, S., and Page, L. 1998, 'The Anatomy of a Large-Scale Hypertextual Web Search Engine', *WWW7/Computer Networks*, vol. 30, no. 1-7, pp. 101-117.

Buchanan, M. 2002, *Nexus: Small worlds and the groundbreaking science of networks*, W.W. Norton, New York ; London.

Bunge, M. A. 2003, *Emergence and convergence : qualitative novelty and the unity of knowledge*: Toronto studies in philosophy, University of Toronto Press, Toronto.

Byrne, D. S. 1998, *Complexity theory and the social sciences an introduction*, Routledge, London ; New York.

Canzer, B. 2003, *E-Business Strategic Thinking and Practice*, Houghton Mifflin Company, Boston.

Carlile, P. R., and Christensen, C. M. 2005, *The Cycles of Theory Building in Management Research*, Innosight LLC [Online], Available: http://www.innosight.com/documents/Theory%20Building.pdf, [Accessed February 18, 2006].

Carson, D., Gilmore, A., Perry, C., and K., G. 2001, *Qualitative Research Marketing*, London.

Cavaye, A., Perry, C., Evans, M., and Sankaran, S. 2002, *MNG03047 Business Research Methods*, Southern Cross University, East Lismore, NSW, 2480.

[Online]. [Accessed

Clarke, R. 2000, *Appropriate Research Methods for Electronic Commerce* [Online], Available: http://www.anu.edu.au/people/Roger.Clarke/EC/ResMeth.html, [Accessed February 18, 2006].

Colecchia, A. 2000, *Defining and Measuring Electronic Commerce: Towards the development of an OECD methodology*, Statistics Singapore [Online], Available: http://www.singstat.gov.sg/conferences/ec/d8.pdf, [Accessed December 7, 2003].

Constatinides, E. 2004, 'Strategies for Surviving the Internet Meltdown', *Management Decision*, vol. 42, no. 1, pp. 89-107.

Creswell, J. W. 2003, *Research Design: Qualitative, quantitative, and mixed methods approaches*, Sage Publications, Thousand Oaks, Calif.

Denzin, N. K., and Lincoln, Y. S. 1994, Entering the Field of Qualitative Research, *in Handbook of Qualitative Research,* N. K. Denzin, and Y. S. Lincoln, eds.: Thousand Oaks, California, Sage Publications Inc., p. 643.

Diaz, C. J. D. 2004, 'The Political Significance of Small Things', *E:CO*, vol. 6, no. 1-2, pp. 49-54.

Dilts, J., and Kahai, P. S. 2004, 'Taking a Small Business Online: A systematic approach', *Journal of Business & Entrepreneurship*, vol. 16, no. 1, pp. 29-45.

[Online]. [Accessed

Donnelly, W. J. 1986, *The Confetti Generation: How the new communications technology is fragmenting America*, Holt, New York.

Dorogovtsev, S. N., and Mendes, J. F. F. 2003, *Evolution of networks : from biological nets to the Internet and WWW*, Oxford University Press, Oxford.

Drew, S. 2002, 'E-Business Research Practice: Towards an Agenda', *Electronic Journal of Business Research Methods* [Online], vol., http://www.ejbrm.com/, Accessed 1

[Online]. [Accessed

Ehrenberg, A. 1982, 'Writing Technical Papers or Reports', *The American Statistition*, vol. 36, no. 4, pp. 326-329.

Eisenhardt, K. 1989, 'Building Theories From Case Study Research', *Academy of Management Review*, vol. 14, no. 4, pp. 532-550.

Featherstone, M., and Borstorff, P. 2005, 'O, What a Tangled Web We Weave When First We Practice to Perceive: An exploratory study of the conduct of business on the World Wide Web', *E-Business Review*, vol. 5, no. 1, pp. 54-57.

Featherstone, M., and Ellis, A. 2005, *The Abundance of Simple Business Models on the World Wide Web*, Proceedings of the 11th Australasian World Wide Web Conference [Online], Available: http://ausweb.scu.edu.au/aw05/papers/refereed/featherstone/index.html, [Accessed October 14, 2005].

Featherstone, M., Ellis, A., and Borstorff, P. 2004, 'The Role of E-Commerce in Business Education', *Australasian Journal of Business and Social Enquiry*, vol. 2, no. 2, pp. 1-12.

Firestone, W. A. 1987, 'Meaning in method: The rhetoric of quantitative and qualitative research', *Educational Researcher*, vol. 16, no. 7, pp. 16-21.

Gadamer, H.-G. 1976, The Historicity of Understanding, *in Critical Sociology, Selected Readings,* P. Connerton, ed.: Harmondsworth, UK, Peguin Books, pp. 117-133.

Gaignon, C. 2002, Truth in Interpretation: A Hermeneutic Approach, *in Is There a Single Right Interpretation,* M. Krausz, ed.: University Park, PA, The Pennsylvania State University Press, pp. 264-284.

[Online]. [Accessed

Gerbert, P., and Birch, A. 2001, *Digital Storm: Fresh business strategies from the electronic marketplace*, Capstone, Oxford.

Goodin, S. 2002, 'Survival is Not Enough', *Fast Company*, vol. no. 54, pp. 90-91.

Guba, E. G., and Lincoln, Y. S. 1994, Competing Paradigms in Qualitative Research, *in Handbook of qualitative research,* N. K. Denzin, and Y. S. Lincoln, eds.: Thousand Oaks, Sage Publications, pp. xii, 643.

Gummesson, E. 2003, 'All Research is Interpretive', *Journal of Business & Industrial Marketing*, vol. 18, no. 6/7, pp. 482-492.

Harbrecht, D. 2003, *Meg Whitman on eBay's Self-Regulation*, Time Magazine [Online], Available: http://www.businessweek.com/bwdaily/dnflash/aug2003/nf20030818_1844_db049.htm, [Accessed April 7, 2004].

Hart, C. 1998, *Doing a Literature Review*, Sage, London.

Henzinger, M., and Lawrence, S. 2004, 'Extracting knowledge from the World Wide Web', *Proceedings of the National Academy of Science*, vol. 101, no. S-1, pp. 5186-5191.

[Online]. [Accessed

Holland, J. H. 1998, *Emergence: From Chaos to Order*, Addison-Wesley, Reading, Mass.

James, P. 2005, *Knowledge Asset Management: The strategic management and knowledge management nexus*, Southern Cross University.

Johnson, R. B., and Onwuegbuzi, A. J. 2004, 'Mixed Methods Research: A reasearch paradigm whose time has come', *Educational Researcher*, vol. 33, no. 7, pp. 14-26.

Johnson, S. 2001, *Emergence: The connected lives of ants, brains, cities, and software*, Scribner, New York.

Keller, E. F. 2005, 'Ecosystems, Organisms, and Machines', *Bioscience*, vol. 55, no. 12, pp. 1069-1074.

[Online]. [Accessed

Klein, H. K., and Myers, M. D. 1999, 'A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems', *MIS Quarterly*, vol. 23, no. 1, pp. 67-94.

Kuhn, T. S. 1962, *The Structure of Scientific Revolution*, The University of Chicago Press, Chicago, Illinois.

[Online]. [Accessed

Laudon, K. C., and Traver, C. G. 2007, *E-Commerce: Business, technology, society*, Prentice Hall, Upper Saddle River, NJ.

Leedy, P. D., and Ormrod, J. E. 2001, *Practical research : planning and design*, Merrill Prentice Hall, Upper Saddle River, N.J.

Leiner, B. M., Cerf, V., Clark, D. D., Kahn, R. E., Kleintock, L., Lynch, D. C., Postel, J., Roberts, L. G., and Wolff, S. 2000, *A Brief History of the Internet* [Online],

Available: http://www.isoc.org/internet/history/brief.shtml#darpa, [Accessed March 18, 2006].

Leonhardt, D. 2000, 'At Graduate Schools, a Great Divide over E-Business Studies', *New York Times*, Business, January 16, 2000.

Levin, S. A. 2005, 'Self-organization and the Emergence of Complexity in Ecological Systems', *Bioscience*, vol. 55, no. 12, pp. 1075-1079.

Lidsky, D. 2003, 'State of the Small Business Web', *Fortune* [Online], vol., http://www.fortune.com/fortune/smallbusiness/skeptic/0,15704,411307,00.html, Accessed

Lockett, N. J., and Brown, D. H. 2003, *Innovations affecting SMEs and E-Business with reference to Strategic Networks, Aggregation, & Intermediaries*, Lancaster University Management School [Online], Available: http://www.lums.lancs.ac.uk/publications/viewpdf/000161/, [Accessed December 12, 2005].

Lynch, B. K. 2005, *The Paradigm Debate* [Online], Available: http://www.iltaonline.com/newsletter/01-2005may/latedialog-lynch.htm, [Accessed May 29, 2006].

Magretta, J. 2002, 'Why Business Models Matter', *Harvard Business Review*, vol. 80, no. 5, pp. 86-92.

Manson, S. M. 2001, 'Simplifying Complexity: A review of complexity theory', *Geoforum*, vol. 32, no., pp. 405-414.

Marshall, C., and Rossman, G. B. 1989, *Designing Qualitative Research*, Sage Publications, Newbury Park, Calif.

Marshall, P., Kelder, J.-A., and Perry, A. 2005, Social Constructivism with a Twist of Pragmatism: A suitable cocktail for Information Systems research: Proceedings of the 16th Conference in Information Systems, Sydney, NSWhttp://www.utas.edu.au/infosys/publications/research/phil_research/Social_Constructionism_and_Pragmatism.pdf.

Miles, M. B., and Huberman, A. M. 1994, *Qualitative Data Analysis: An expanded sourcebook*, Sage Publications, Thousand Oaks, Calif.

Moore, R. 2005a, *Domain Name Valuation Considerations*, Brisbane, Dark Blue Sea Ltd., p 32.

Moore, R. 2005b, *Investor Presentation*, Brisbane, Dark Blue Sea Ltd., p 32. http://affiliate.battletrolls.com/announcements/117_announcement_re_december_investor_presentation.pdf.

Morowitz, H. J. 2002, *The Emergence of Everything: How the world became complex*, Oxford University Press, New York.

Moschovitis, C. J. P. 1999, *History of the Internet : a chronology, 1843 to the present*, Abc-Clio, Santa Barbara, Calif.

Nachira, F. 2002, *Towards a Network of Digital Business Ecosystems Fostering the Local Development*, European Commission - Information Society Technology [Online], Available: http://www.digital-ecosystem.org/html/repository/dbe_discussionpaper.pdf, [Accessed October 14, 2004].

Nancarrow, C., Pallister, J., and Brace, I. 2001, 'A New Research Medium, New Research Populations and Seven Deadly Sins for Internet Researchers', *Qualitative Market Research: An International Journal*, vol. 4, no. 3, pp. 136-149.

Newman, M. E. J. 2003, 'The Structure and Function of Complex Networks', *SIAM Review*, vol. 45, no. 2, pp. 167-256.

Nielson, J. 1997, *Zipf Curves and Website Popularity*, Jakob Nielson [Online], Available: http://www.useit.com/alertbox/zipf.html, [Accessed May 29, 2006].

Oevermann, U. 2003, *About the Institute for Hermeneutic Research on Culture and Society*, The Institute for Hermeneutic Research on Culture and Society [Online], Available: http://www.ihsk.de/en/institute.htm, [Accessed June 8, 2005].

O'Neill, E., McClain, P. D., and Lavoie, B. F. 1998, *A Methodology for Sampling the World Wide Web*, Online Computer Library Center [Online], Available: http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000003447, [Accessed July 28, 2004].

[Online]. [Accessed

Ovans, A. 2000, 'Can You Patent Your Business Model?' *Harvard Business Review*, vol. 78, no. 4.

Parkhe, A. 1993, '"Messy" Research, Methodological Predispositions, and Theory Development in International Joint Ventures', *Academy of Management Review*, vol. 18, no. 2, pp. 227-268.

Patrick, J. R. 2001, *Net attitude : what it is, how to get it, and why your company can't survive without it*, Perseus Pub., Cambridge, MA.

Peden, B. F. 2004, Virtual Research Ethics: A Content Analysis of Surveys and Experiments, *in Readings in Virtual Research Ethics: Issues and Controversies,* E. A. Buchanan, ed., vol. 1: Hershey, PA, Information Science Publishing, pp. 1-26.

Pennock, D. M., Flake, G., Lawrence, S., Glover, E., and Giles, C. L. 2002, 'Winners don't take all: Characterizing the competition for links on the web', *Proceedings of the National Academy of Science*, vol. 99, no. 8, pp. 5207-5211.

Perry, C. 1998, 'A structured approach for presenting research thesis', *Australasian Marketing Journal*, vol. 6, no. 1, pp. 63-86.

Perry, C. 2002, A Structured Approach to Presenting Thesis: Notes for Students and Their Supervisors, *in DBA Thesis Units, MNG03051 to MNG3066: DBA Thesis Stage,* C. Perry, and A. Cavaye, eds., vol. III: Lismore, NSW, Australia, Southern Cross University.

Pool, P. W., Parnell, J. A., and Carraher, S. 2004, Utilizing E-Commerce in Small Business: Proceedings of the 27th Annual Conference of the Association for Small Business & Entrepreneurship, Albuquerque, NM.

Porter, M. 2001, 'Strategy and the Internet', *Harvard Business Review*, vol. 79, no. 3, pp. 63-78.

Proctor, J. D., and Larson, B. M. H. 2005, 'Ecology, Complexity, and Metaphor', *Bioscience*, vol. 55, no. 12, pp. 1065-1068.

Rappa, M. 2005, *Managing the Digital Enterprise*, North Carolina State University [Online], Available: http://digitalenterprise.org/index.html, [Accessed January 5, 2001].

Ray, R. 2002, E-Commerce: Curriculum, Pedagogy, Outlook: Proceedings of the Paper presented to the Academy of Business Education Conference, Key West, FL, September 19-21, Key West, Florida Villanova Univesity.

[Online]. [Accessed

Rayport, J. F., and Jaworski, B. J. 2004, *Introduction to e-Commerce*, McGraw-Hill, New York.

[Online]. [Accessed

Rescher, N. 1977, *Methodological pragmatism : a systems-theoretic approach to the theory of knowledge*, Blackwell, Oxford.

Rhoads, C. 2006, 'Plan for Adult Area Sparks a Fight On Control of Web', *Wall Street Journal*, 1, May 10, 2006, p. 1.

Ricerca, J. 2004, *Search Engine Optimization: Static IP vs. dynamic IP addresses*, Circle ID [Online], Available: http://www.circleid.com/print/610_0_1_0/, [Accessed September 12, 2004].

Rothschild, M. L. 1990, *Bionomics: The Inevitability of Capitalism*, H. Holt, New York.

Rusmevichientong, P., Pennock, D., Lawrence, S., and Giles, C. L. 2001, 'Methods for Sampling Pages Uniformly from the World Wide Web', *Proceedings of the AAA1 Fall Symposium on Using Uncertainty Within Computing*, no., pp. 121-128.

Saunders, M. N. K., Lewis, P., and Thornhill, A. 2003, *Research methods for business students*, Financial Times/Prentice Hall, Harlow.

Schrage, M. 2002, 'Network Theory's New Math', *Strategy+Business*, vol. 29, no. Fourth Quarter 2002, pp. 41-46.

Schwandt, T. A. 1994, Constructivist, Interpretivist Approaches to Human Inquirey, *in Handbook of Qualitative Research,* N. K. Denzin, and Y. S. Lincoln, eds.: Thousand Oaks, California, Sage Publications Inc., pp. 118-136.

Sekaran, U. 1992, *Research methods for business: a skill-building approach*, Wiley, New York.

Sheridan, C. 2004 VeriSign Corporation, Personal communication with M. Featherstone, September 27, 2004.

Simon, J. L. 2003, *The Art of Empirical Investigation*, Transaction Publishers, New Brunswick, NJ.

Sloan, P. 2005, 'Master of Their Domains', *Business 2.0*, vol. no. 12, December, 2005.

Strauss, A. L., and Corbin, J. M. 1990, *Basics of Qualitative Research: Grounded theory procedures and techniques*, Sage Publications, Newbury Park, Calif.

Tashakkori, A., and Teddlie, C. 2003, *Handbook of mixed methods in social & behavioral research*, SAGE Publications, Thousand Oaks, Calif.

Thomas, J. 1999, 'Balancing the Ethical Conundrums of Internet Research: An existentialist view from the trenches', *Iowa Journal of Communication*, vol. 31, no., pp. 8-20.

Ticehurst, G. W., and Veal, A. J. 1999, *Business research methods: a managerial approach*, Longman, Sydney.

Timmers, P. 1998, 'Business models for E-commerce', *Electronic Markets*, vol. 8, no. 2, pp. 3-7.

Trochim, W. M. 2002, *Research Methods Knowledge Base*, Cornell University [Online], Available: http://trochim.human.cornell.edu/kb/index.htm, [Accessed October 17, 2002].

Turban, E. 2004, *Electronic Commerce: A managerial perspective*, Prentice Hall, Upper Saddle River, NJ.

Uncles, M. 1998, 'A Structured Approach to the Presentation of Research Thesis', *Australasian Marketing Journal*, vol. 6, no. 1, pp. 87-94.

Valovic, T. S. 2000, *Digital Mythologies: The hidden complexities of the Internet*, Rutgers University Press, New Brunswick, N.J.

VeriSign 2004a, *COM NET Registry - TLD Zone Access Program*, VeriSign [Online], Available: http://www.verisign.com/products-services/naming-and-directory-services/naming-services/com-net-registry/page_001052.html, [Accessed September 3, 2004].

[Online]. [Accessed

[Online]. [Accessed

[Online]. [Accessed

Vicsek, T. 2002, 'The Bigger Picture', *Nature*, vol. 418, no., p. 131.

Vise, D. A., and Malseed, M. 2005, *The Google Story*, Delacorte Press, New York.

Walczuch, R., Braven, G. d., and Lundgren, H. 2000, 'Internet adoption barriers for small firms in the Netherlands', *European Management Journal*, vol. 18, no. 5, pp. 561-572.

Waldrop, M. M. 1992, *Complexity: The Emerging Science at the Edge of Order and Chaos*, Simon & Schuster, New York.

Walsham, G. 1995, 'The Emergence of Interpretivism in IS Research', *Information Systems Research*, vol. 6, no. 4, pp. 376-394.

Webster, J., and Watson, R. T. 2002, 'Analyzing the Past to Prepare for the Future: Writing a Literature Review', *MIS Quarterly*, vol. 26, no. 2, pp. xiii-xxiii.

Weinberger, D. 2002, *Small Pieces Loosely Joined: A unified theory of the Web*, Perseus Pub., Cambridge, MA.

Wright, S. R. 1979, *Quantitative Methods and Statistics: A guide to social research*, Sage Publications, Beverly Hills.

Yager, L. 2002, *International Electronic Commerce: Definitions and policy implications*, United States Government Accounting Office, p 98. http://www.gao.gov/new.items/d02404.pdf.

Yook, S.-H., Jeong, H., and Barabasi, A.-L. 2002, 'Modeling the Internet's large scale topology', *Proceedings of the National Academy of Science*, vol. 99, no. 21, pp. 13382-13386.

Zikmund, W. G. 1991, *Business Research Methods*, Dryden Press, Chicago.

Zikmund, W. G. 2000, *Business research methods*: The Dryden Press series in management, Dryden Press, Fort Worth, TX.

APPENDICES

Appendix A – Program to randomly extract domain names from large Top Level Domane zone file

```
import java.io.BufferedReader;
import java.io.BufferedWriter;
import java.io.File;
import java.io.FileInputStream;
import java.io.FileOutputStream;
import java.io.InputStreamReader;
import java.io.OutputStreamWriter;
import java.util.Random;
public class RandomLineExtractor
{
File _inputFile = null;
File _outputFile = null;
int _numberOfLinesInput = 0;
int _numberOfLinesOutput = 0;
int _randomSeed = 23333333;
public static void main( String[] args )
        {
if ( args.length < 4 ) System.out.println( "This application requires 4 parameters" );
RandomLineExtractor rle = new RandomLineExtractor( args[0],
args[1],
Integer.parseInt( args[2] ),
Integer.parseInt(args[3] ));
if ( args.length > 4 ) rle._randomSeed = Integer.parseInt(args[4] );
try {
rle.extractFile();
} catch (Exception e) {

                }
        }

public RandomLineExtractor( String inputFilename,
String outputFilename,
int inputLines,
int outputLines )
        {
_inputFile = new File( inputFilename );
_outputFile = new File( outputFilename );
_numberOfLinesInput = inputLines;
_numberOfLinesOutput = outputLines;

        }
```

```
protected void extractFile() throws Exception
        {
if ( ! _inputFile.exists() )
throw new Exception( “Input File not found.” );
FileInputStream fis = null;
BufferedReader reader = null;

FileOutputStream fos = null;
BufferedWriter writer = null;
try
                {
fis = new FileInputStream( _inputFile );
reader = new BufferedReader(new InputStreamReader(fis));
fos = new FileOutputStream( _outputFile );
writer =  new BufferedWriter( new OutputStreamWriter( fos ) );
String lineData = null;
Random random = new Random(_randomSeed);
int lineNumber = 0;
int lastPrintout = 0;
int numberOfRandomLines=0;
while((lineData=reader.readLine())!=null)
                        {
if ( random.nextInt( _numberOfLinesInput ) <= _numberOfLinesOutput )
                                {
writer.write( lineData );
writer.newLine();
numberOfRandomLines++;
writer.flush();
                                }
lineNumber++;
if ( (lineNumber - lastPrintout) > 1000 )
                                {
lastPrintout = lineNumber;
System.out.println( “Current Line number = “ + lineNumber );

                                }
                        }

System.out.println( “Number of Lines =” + lineNumber );
System.out.println( “Number of Random Lines =” + numberOfRandomLines );
                }
catch( Exception e )
                {
System.out.println( “Error: “ + e.getMessage() );
                }
finally
                {
reader.close();
```

```
fis.close();

writer.close();
fos.close();
                }
        }
}
```

Appendix B – Domain Name acquisition from the final sampling frame

Data Acquisition Process

Software: TextPad

Data Access: 30 NET Domain Names

Process:

1. You will receive dated Random number Exel Spreadsheets.
2. Open the TextPad program and load the NET Domain names file
3. Go the View option and make sure the "Line Number" function is enabled.
4. Open the current Excel Random Number Spreadsheet.
5. Observe the first open RN (random number) on the Excel sheet.
6. From the TextPad - Use Alt G to bring up the "goto" function.
7. In the widow, type in the RN from the Excel sheet.
8. Copy the domain name in the TextPad file and paste it next to the RN in the Excel sheet.
9. Go to step 5 and repeat the process until the Excel sheet is full (this will typically be 100 domain names.
10. Move the completed Excel spreadsheet to the Digital Drop box and send me an e-mail notifying me you've completed a sheet paste it next to the RN in the Excel sheet.

Appendix C – Instructions for domain name resolution

Process:

1. You will receive Excel Files with the Date, RN, and Domain Name Fields completed.
2. Each file is numbered and will be assigned to a specific student. WORK ONLY ON FILES ASSIGNED TO YOU.
3. Open the Excel File and a browser.
4. Copy the first DN (domain name) available into the browser and add the Top Level Domain extension - in this case ".net". It is important that you ENSURE you've type in the appropriate TLD as the site visited will not be correct if you leave it blank or type in ".com".
5. The first priority is to complete column D "categories". A code list is available in the Data Analysis Team Room in the File Transfer section. Second priority is Column F "type". An initial Type Code list will be available in the File Transfer area as well.
6. Put completed files in the digital drop box.

We all recognize that questions will arise regarding interpretation of the site. We will arrange team meetings to discuss these issues, or you may call me if you have a pressing issue.

Appendix D – Sample of initial category structure

Category	Description
NF	Not found
Empty	No data
Biz	Business site
DNS	Domain Name System site
Redirect	Redirected to another site (save site name)
UnderCons	Site is under construction
?	Cannot determine if it is business site
Political	Political site

Appendix E – Sample of completed data sheet

Date	rn	Dn	Category	language	type
<<		Key Cells		>>	Optional
14-Oct-04	3512	BODYSOLIDOUTLET	Nf		no dns
14-Oct-04	8583	ESNN	biz		
14-Oct-04	23589	SMALLBIZSTRATEGIES	Undercons		future biz
14-Oct-04	22271	SACREDTOURS	biz		
14-Oct-04	18593	NTEEN	Nf		no dns
14-Oct-04	515	999CASINO-ONLINE	biz		casino
14-Oct-04	16792	MICROWAVEPOPCORNONASTRING	dns		www.godaddy.com
14-Oct-04	19278	OVERSEAMAIL	Nf		no dns
14-Oct-04	17455	MUCIC	?		search page
14-Oct-04	9191	FEVO	Nf		no dns
14-Oct-	4744	CETIAN	underco		personal page?

04			ns		
14-Oct-04	6054	CREDIBILITYBASEDCOMMUNICATIONS	dns		www.godaddy.com
14-Oct-04	14716	KVZPRESTICE	?	no idea	
14-Oct-04	23991	SPAIN-PROPERTY-FOR-SALE	biz		real estate
14-Oct-04	27839	WAYFON	nf		no dns
14-Oct-04	9864	FRESHBOOKMARKS	blank		
14-Oct-04	4829	CHARLOTTEODONNELL	nf		no dns
14-Oct-04	7340	DON-ORG	dns		expired
14-Oct-04	23388	SIMPLESIMONVIDEOS	underconns		biz
14-Oct-04	5719	CONFLATOR	nf		blank
14-Oct-04	28853	XN—B1ADEF0BAN2H	dns		search page

Appendix F - Example of TextPad enumerated database records

Figure removed due to copyright restrictions

Appendix G – Example of VeriSign Zone File access agreement

VERISIGN ZONE FILE ACCESS AGREEMENT

1. PARTIES

The User named in this Agreement hereby contracts with VeriSign, Inc. ("VeriSign") for a non-exclusive, non-transferable, limited right to access an Internet host server or servers designated by VeriSign from time to time, and to transfer a copy of the described Data to the User's Internet host machine specified below, under the terms of this Agreement. Upon execution of this Agreement by VeriSign, VeriSign will return a copy of this Agreement to you for your records with your UserID and Password entered in the spaces set forth below.

2. USER INFORMATION

(a) User: _____

(b) Contact Person: _____

(c) Street Address: _____

(d) City, State or Province: _____

(e) Country and Postal Code: _____

(f) Telephone Number: _____

(including area/country code)

(g) Fax Number: _____

(including area/country code)

(h) E-Mail Address: _____

(i) Specific Internet host machine which will be used to access VeriSign's server to transfer copies of the Data:

Name: _____

IP Address: _____

(j) Purpose(s) for which the Data will be used: During the term of this Agreement, you may use the data for any legal purpose, not prohibited under Section 4 below. You may incorporate some or all of the Data in your own products or services, and distribute those products or services for a purpose not prohibited under Section 4 below.

3. TERM

This Agreement is effective for a period of three (3) months from the date of execution by VeriSign (the "Initial Term"). Upon conclusion of the Initial Term this Agreement will automatically renew for successive three-month renewal terms (each a "Renewal Term") until terminated by either party as set forth in Section 12 of this Agreement or one party provides the other party with a written notice of termination at least seven (7) days prior to the end of the Initial Term or the then current Renewal Term.

NOTICE TO USER: CAREFULLY READ THE FOLLOWING TERMS AND CONDITIONS. YOU MAY USE THE USER ID AND ASSOCIATED PASSWORD PROVIDED IN CONJUNCTION WITH

THIS AGREEMENT ONLY TO OBTAIN A COPY OF .COM/.NET TOP-LEVEL DOMAIN ("TLD") ZONE FILES, AND ANY ASSOCIATED ENCRYPTED CHECKSUM FILES (COLLECTIVELY THE

"DATA"), VIA THE FILE TRANSFER PROTOCOL ("FTP") OR HYPERTEXT TRANSFER PROTOCOL ("HTTP") PURSUANT TO THESE TERMS.

4. GRANT OF ACCESS

VeriSign grants to you a non-exclusive, non-transferable, limited right to access an Internet host server or servers designated by VeriSign from time to time, and to transfer a copy of the Data to the Internet host machine identified in Section 2 of this Agreement no more than once per 24 hour period using FTP or HTTP for the purposes described in this Section 4. You agree that you will:

(a) use this Data only for lawful purposes but that under no circumstances will you use this Data to:

(1) allow, enable, or otherwise support the transmission by e-mail, telephone, or facsimile of mass unsolicited, commercial advertising or solicitations to entities other than your own existing customers; or

(2) enable high volume, automated, electronic processes that send queries or data to the systems of VeriSign or any ICANN-accredited registrar, except as reasonably necessary to register domain names or modify existing registrations. VeriSign reserves the right, with the approval of the Internet Corporation for Assigned Names and Numbers ("ICANN"), to specify additional specific categories of prohibited uses by giving you reasonable written notice at any time and upon receiving such notice you shall not make

such prohibited use of the Data you obtain under this Agreement.

(b) copy the Data you obtain under this Agreement into a machine-readable or printed form only as necessary to use it in accordance with this Agreement in support of your use of the Data.

(c) comply with all applicable laws and regulations governing the use of the Data.

(d) not distribute the Data you obtained under this Agreement or any copy thereof to any other party without the express prior written consent of VeriSign, except that you may redistribute the Data insofar as it has been incorporated by you into a value-added product or service that does not permit the extraction of a substantial portion of the Data from the value-added product or service, provided you prohibit the recipient of the Data from using the Data in a manner contrary to Section 4(a).

(e) take all reasonable steps to protect against unauthorized access to, use, and disclosure of the Data you obtain under this Agreement.

5. FEE

You agree to remit in advance to VeriSign a quarterly fee of $0 (USD) for the right to access the files during either the Initial Term or Renewal Term of this Agreement. VeriSign reserves the right to adjust, with the approval of ICANN, this fee on thirty days' prior notice to reflect a change in the cost of

providing access to the files.

6. PROPRIETARY RIGHTS

You agree that no ownership rights in the Data are transferred to you under this Agreement. You agree that any copies of the Data that you make will contain the same notice that appears on and in the Data obtained under this Agreement.

7. METHOD OF ACCESS

VeriSign reserves the right, with the approval of ICANN, to change the method of access to the Data at any time. You also agree that, in the event of significant degradation of system processing or other emergency, VeriSign may, in its sole discretion, temporarily suspend access under this Agreement in order to minimize threats to the operational stability and security of the Internet.

8. NO WARRANTIES

The Data is being provided "as-is." VeriSign disclaims all warranties with respect to the Data, either expressed or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, and non-infringement of third party rights. Some jurisdictions do not allow the exclusion of implied warranties or the exclusion or limitation of incidental or consequential damages, so the above limitations or exclusions may not apply to you.

9. SEVERABILITY

In the event of invalidity of any provision of this Agreement, the parties agree that such invalidity shall not affect the validity of the remaining provisions of this Agreement.

10. NO CONSEQUENTIAL DAMAGES

In no event shall VeriSign be liable to you for any consequential, special, incidental or indirect damages of any kind arising out of the use of the Data or the termination of this Agreement, even if VeriSign has been advised of the possibility of such damages.

11. GOVERNING LAW

This Agreement shall be governed and construed in accordance with the laws of the Virginia, USA. You agree that any legal action or other legal proceeding relating to this Agreement or the enforcement of any

provision of this Agreement shall be brought or otherwise commenced in the state or federal court in Virginia, USA. You expressly and irrevocably agree and consent to the personal jurisdiction and venue of the federal and states courts located Virginia, USA (and each appellate court located therein) for maters arising in connection with this Agreement or your obtaining, use, or distribution of the Data. The United Nations Convention on Contracts for the International Sale of Goods is specifically disclaimed.

12. TERMINATION

You may terminate this Agreement at any time by erasing the Data you obtained under this Agreement from your Internet host machine together with all copies of the Data and providing written notice of your termination to VeriSign at 21345 Ridgetop Circle, Dulles, Virginia 20166, attn: Director, Business Affairs. VeriSign has the right to terminate this Agreement immediately if you fail to comply with any term or condition of this agreement. You agree upon receiving notice of such termination of this Agreement by VeriSign or expiration of this Agreement to erase the Data you obtained under this Agreement together with all copies of the Data.

13. DEFINITION

"Data" means all data contained in a DNS zone file for the Registry TLD as provided to TLD nameservers on the Internet.

14. ENTIRE AGREEMENT

This is the entire agreement between you and VeriSign concerning access and use of the Data, and it supersedes any prior agreements or understandings, whether written or oral, relating to access and use of the Data.

VERISIGN, INC.

USER: _____

By: _____

By: _____

(sign)

(sign)

Name: _____

Name: _____

(print)

(print)

Title: _____

Title: _____

Date: _____

Date: _____

ASSIGNED USERID AND PASSWORD

(To be assigned by VeriSign upon execution of this Agreement):

USERID: _____

*PASSWORD:*_____

Appendix H - Peer reviewed publications to which this study contributed

articles in professional journals:

March, 2006 "Sampling the Internet" M. Featherstone and P. Borstoff, E-Business Review, vol. VI, no. 1, pp. 47-49

April, 2005 "Oh what a tangled web we weave when first we practice to perceive: An exploratory study of the conduct of business on the World Wide Web" M. Featherstone and P. Borstorff *E-Business Review, vol. V, no. 1, pp. 54-57*

June, 2004 "The Role of E-Commerce in Business Education" M. Featherstone, P. Borstorff, and A. Ellis, *Australasian Journal of Business and Social Inquiry. vol.2, no. 2 [online]* Available: http://ausweb.scu.edu.au/aw05/papers/refereed/featherstone/index.html .

Articles Included in Edited Collections:

July, 2005 "The Abundance of Simple Business Models on the World Wide Web" M. Featherstone and A. Ellis in *"Managing the Digital Enterprise"*, Editor, Dr. Michael Rappa, Publisher - North Carolina State University *2 [online]* Available: http://digitalenterprise.org/models

Proceedings:

March, 2006 "Employing Simple Business Models in E-Commerce Education" M. Featherstone and P. Borstoff, Paper presented at *International Academy of E-Business Conference, Orlando,FL*

Presentations:

July, 2005 "The Abundance of Simple Business Models on the World Wide Web" Featherstone and Ellis *11th Australasian World Wide Web conference, Brisbane, QLD*

Sept, 2003 "E-Commerce Framework for Business Education" Featherstone *Academy of Business Education, San Francisco, CA*