

2004

Open URL: linking through the maze of online resources

Katie Wilson
Southern Cross University

Publication details

Post-print of: Wilson, K 2004, 'OpenURL: linking through the maze of online resources', *Online Currents*, vol. 19, no. 2, pp. 3-8.

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.

OpenURL: linking through the maze of online resources

Online Currents, vol. 19, no. 2, pp.3 – 8, March 2004

Katie Wilson

We have heard and read much in the last couple of years about reference linking products and capabilities offered by library automation vendors and information providers, making links between library catalogues, citation and full text services and other Web resources. Dig a bit further into the announcements and product descriptions and you find references to a standard called OpenURL. What is OpenURL and where did it come from?

OpenURL is an open NISO (National Information Standards Organization) standard that facilitates linking from a citation or other electronic resource to journal full texts and related online resources which are licensed and accessible through a user's library. It is also known as *dynamic context sensitive reference linking*. What does this mean? It's a seemingly simple but very powerful concept, linking search tools with full texts on-the-fly and bringing them under one umbrella that points to where a user is located (the institutional context), and the resources available via that institution. OpenURL, as I discovered, is a kind of fairy story that came true, of taming the Web non-commercially, of bringing library and user control into the online publishing and research equation.

I first heard about reference linking at the LITA President's Program: **The Distributed Digital Library: putting the pieces together**, during the 2001 American Library Association Conference where Priscilla Caplan and Clifford Lynch gave thought-provoking talks on the topic. I remember thinking how powerful the concept sounded, and that it was going to have a major impact in the library and information world. Soon OpenURL linking emerged as the basis of a product called WebBridge at the library automation vendor Innovative Interfaces, where I was working.

OpenURL extends the Dublin Core metadata or description of a bibliographic citation, carrying it across the Web as a package combined with a user's context (where the search started), to the full text resources within the searcher's *institutional context*. It enriches linking via the Web, which on its own cannot make the link between a user and the library or institutional resources to which the user has access.

A more formal definition of OpenURL is a standardized *syntax* for describing *web*-transportable packages of *bibliographic metadata* and *identifiers* about *information objects* between the systems of different *information providers* (my italics).

These elements can be broken down as follows:

web = http protocol

syntax = elements that describe and create a link between a source and a *link resolver*. OpenURL takes information about the item from its *source* to the link resolver, which passes it to appropriate copies or full text.

bibliographic metadata = OpenURL syntax made up of predefined elements that specify an *object* or item's metatags, such as genre (book, journal, conference, article, preprint, conference paper or proceeding, part of a book), titles, author first and last names, initials, journal title, volume, issue, starting and ending page numbers, date, standard numbers.

META-TAG ::= 'genre' | 'auiast' | 'auiirst' | 'auiunit' | 'auiunit1' | 'auiunitm' | 'coden' | 'issn' | 'eissn' | 'isbn' | 'title' | 'stitle' | 'atitle' | 'volume' | 'part' | 'issue' | 'spage' | 'epage' | 'pages' | 'artnum' | 'sici' | 'bici' | 'ssn' | 'quarter' | 'date'

identifiers: codes that identify the source or citation's vendor SID (service identifier) and institutional private information PID (private identifier).

information objects – journal articles, books, conferences, papers

information providers:

- integrated library system vendors
- abstracting and indexing or citation services
- full text vendors, publishers, journal aggregators
- library online catalogues
- online bookstores

How did OpenURL originate?

During the 1990s libraries began purchasing and providing Web access to full text and electronic journal collections, replacing, or as alternatives to, print journal subscriptions. Commercial online citation or abstracting and indexing databases controlled the links from citations to the electronic full text. Libraries were either locked into using the full text links provided by the citation sources, or offered links to the citation indexes and the corresponding full text resources in separate steps, which was confusing for users. Many diverse overlapping information resources were offered to users. Links were static and not sensitive to a user's context, and libraries were not easily able to direct their users to expensive full text subscriptions. Which copy should a user be linked to (*appropriate copy*)? How to identify that copy?

In response to the situation described above at the University of Ghent, Belgium, research was undertaken by Herbert Van de Sompel and Patrick Hochstenbach from the Central Library Automation Department. They identified what they called *extended links* to full text articles from:

- citations in abstracting and indexing databases
- bibliographic records for journals in library online catalogues
- references within journal articles to a citation in an abstracting and indexing database

Van de Sompel and Hochstenbach also identified situations in which the extended links were controlled by the commercial information providers who owned, published or distributed the full text articles. These links did not take any account of the *context* of the user. The information providers closed off any means for libraries to steer users to the resources or *appropriate copy* they preferred their users to access. For example, an index database may include citations with built in links to e-journals containing the full text of the corresponding articles, but there was no way of letting the library build in a link to its local holdings, full text subscriptions, or ILL services for items not held. Early commercial and non-commercial experiments with linking services showed problems with authentication and depth of linking levels.

Van de Sompel and Hochstenbach's research, published from 1998 to 2000, called for a generic framework to encompass the linking issues they had identified. As a result of the research they developed the SFX (Special Effects) linking server software which integrated the University of Ghent Library's Aleph (Ex Libris) online catalogue with information providers SilverPlatter, Web of Science, and e-journals to which the library subscribed.

Active campaigning by Herbert Van de Sompel, Oren Beit-Arie from the library automation vendor Ex Libris, the Digital Library Foundation, and the International Coalition of Library Consortia (ICOLC) raised awareness of the context sensitive linking issue. In 2000 the OpenURL framework was submitted to NISO and published as OpenURL 0.1 specification in March 2001. In 2000 Ex Libris purchased the SFX linking server software from the University of Ghent, and SFX, the first commercial OpenURL product, was released by Ex Libris in 2001.

OpenURL 0.1 is the version of the standard currently in use. It is a non-proprietary open framework which has been, and continues to be, adopted by major information providers and automation vendors since the acceptance of OpenURL as a standard. OpenURL 1.0, an expanded draft version, is now being trialled. The format specification for OpenURL is at <http://library.caltech.edu/openurl/>, with a syntax description at <http://www.sfxit.com/OpenURL/openurl.html>

How does OpenURL work?

Van de Sompel saw that the key to the OpenURL process was to separate the *linking* service from the description of the work, so that the linking service could be provided by many parties and not limited to the providers of the citation or reference. And if the service component interacted with the user's library collection, the linking service would lead to full text copies that are *appropriate* to the user.

OpenURL is based on the idea that links should lead to full text resources licensed to a library. The OpenURL syntax combines the *base URL* of the institution or library's linking resolution server with the citation or query, made up of metatags. The linking server reads the OpenURL as input and takes action upon it, leading a user to *appropriate* resources.

Figure 1 shows the OpenURL process.

Source or origin: where a search starts. This can be a library online catalogue, an abstract and indexing database, or references within full text journal articles. Sources must be OpenURL compliant and have a vendor **SID** (identifier).

Object: book, article, conference paper etc. Described with OpenURL specified metatags, vendor SID and institutional private identifiers (PID) if required by the vendor.

Transport: of dynamically generated OpenURL syntax including the *base URL*, metadata, and identifiers across the Web.

Link resolver, link server, OpenURL resolver, resolution server: linking software within an institutional context that interprets the *base URL* embedded in the incoming OpenURL syntax, takes the local holdings and access privileges of the institution into account, and resolves the links to appropriate resources. The library or institution maintains its e-journal holdings and determines the appropriate copy, offering a range of locally-configured links and services, to full text, to a local catalogue to check print holdings, to document delivery or ILL services, and to Web-based search engines. In this way libraries are able to control resource links and ensure expensive e-journal resources are used.

Target or destination: linked-to resources. Where the full text is located: e-journals, full text databases, commercial publishers, journal aggregators, online publishers, societies, document delivery services (including local ILL), and Web resources. Targets do not need to be OpenURL compliant but do need to have a persistent URL. Links can be to a journal only or directly to the article. If an object is available from more than one target resource, the library can customize the order of display of targets, and the appropriateness based on library e-holdings data. For example, if a reference is to an e-journal which is out of scope of the library's holdings, it will not display as a target option.

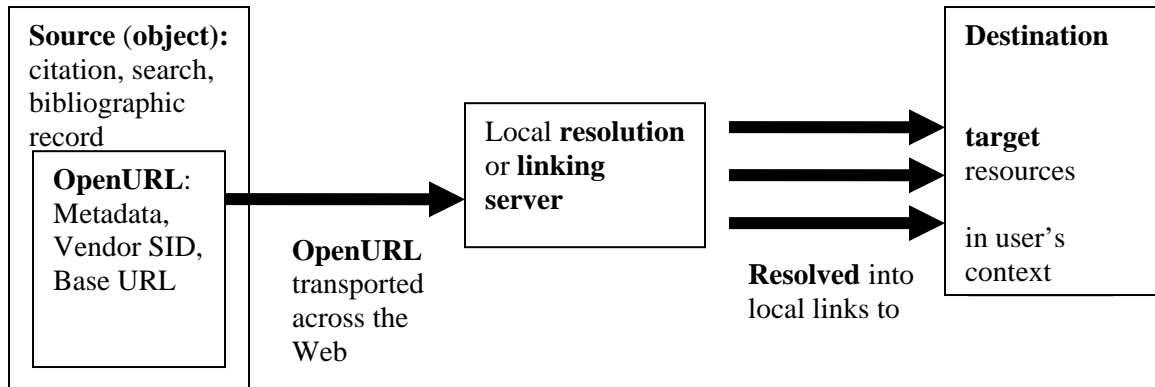


Figure 1: Diagram of OpenURL process

The target resources need to be able to differentiate between a user that has institutional access to the resource, and a user that does not. To achieve this, authentication of users by the institution or library can be accomplished using a range of methods such as IP address, logins, and proxy verification. The *base URL* of the linking server can also be stored in a cookie in the user's Web browser, or created by a "cookiepusher" script at the target information resource. The cookiepusher reads the *base URL* from the incoming OpenURL and pushes the *base URL* into the user's browser, or stores it in the user's profile. The user is then linked to the target resource's URL.

As an example, Figure 2 shows linking from an external information provider:

- WebSpirs citation database result with OpenURL link
- user selects link and linking server software resolves OpenURL syntax
- appropriate target resources display
- user selects ProQuest

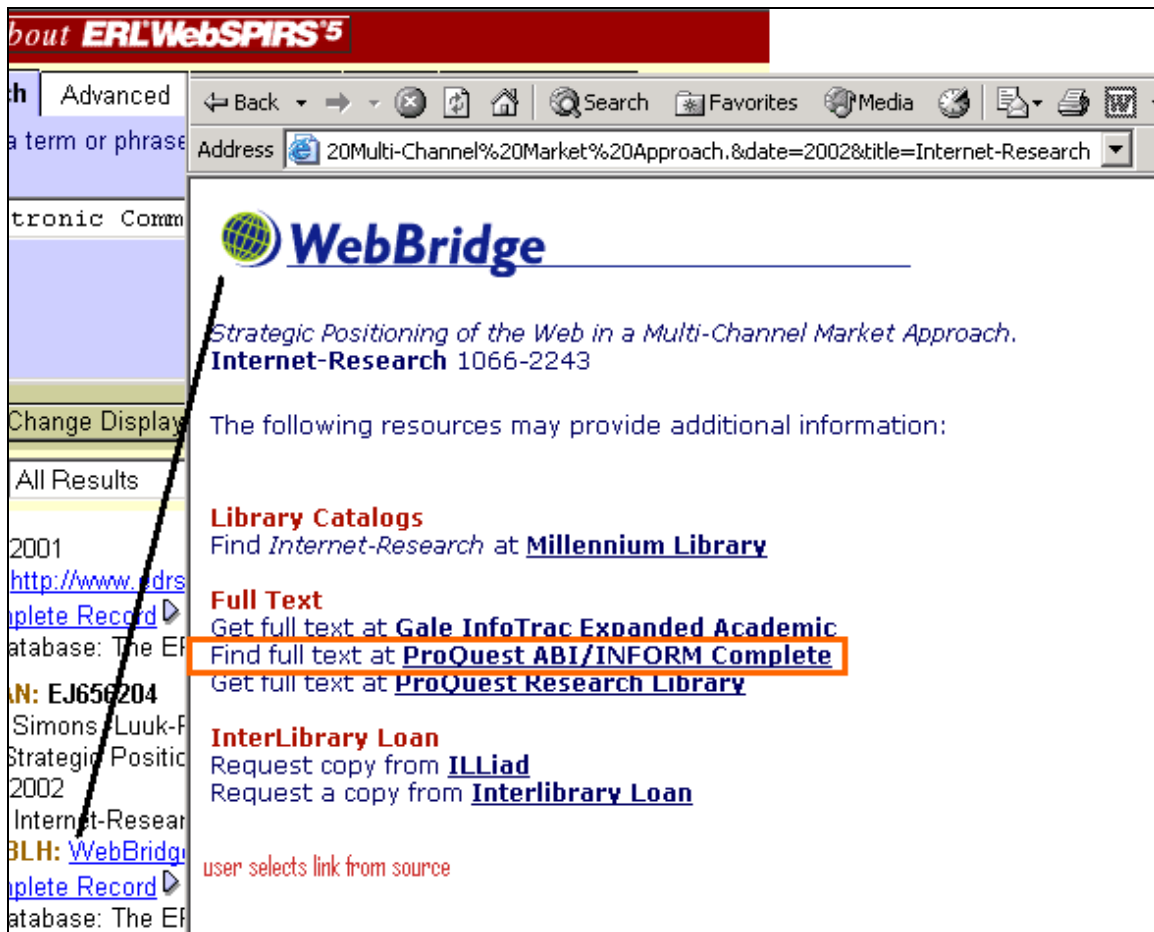


Figure 2: Linking from an external provider

The OpenURL link is activated when a user selects a button or named link displaying on the source. The button or link is dynamically generated by Javascript in the Web page, or by the Web server as part of the search. Sites implementing OpenURL can name the service as they wish, and customize the buttons accordingly. Each click or selection of an OpenURL link is recorded, allowing institutions to gather statistics, and to build a list of preferred resources based on the user selections.

The example in Figure 3 from the University of Iowa Library breaks down as follows:

- a journal title search is in the OPAC for **Ethnohistory**
- the resulting bibliographic record displays the OpenURL button **InfoLink**
- selecting the **InfoLink** button displays the OpenURL resource window with the institutional holdings and other targets which the library wishes to offer to the user
- if the user fills in the year, volume, issue and start page the link goes to article level. If those boxes are left blank the link is only to the journal title

The screenshot shows the InfoLink interface. On the left, there is a sidebar with a search result for 'Ethnohistory /'. The main content area displays the source 'Ethnohistory [0014-1801] JSTOR (Organization)'. Below this, there are several links and search options:

- Full text available via **Academic Search Elite**. Availability: from 1985. Availability: Most recent 1 year not available. year: volume: issue: start page:
- Full text available via **JSTOR Arts and Sciences 2**. Availability: from 1954 to 1999. year: volume: issue: start page:
- Full text available via **Project Muse**. Availability: from 1999 volume 46 issue 4. year: volume: issue: start page:
- View this journal's **Impact Factor Trend Graph in JCR SocSci**
- Additional publication information from **Ulrich's**
- Click here to **Ask a librarian for help**
- What you should know about **using electronic information resources**. Plagiarism, copyright, fair use, and terms of licenses

A red arrow points from the 'Academic Search Elite' link to the 'user selects link from source' message at the bottom left of the interface.

Figure 3: Internal catalogue link

The dynamically generated OpenURL syntax for this example is:

[http://infolink.lib.uiowa.edu:8080/sfx_local?sid=ALEPH:IOW01&genre=&isbn=&issn=0014-1801&date=999&volume=&issue=&spage=&aulast=JSTOR%20\(Organization\)&aufirst=&aunit=&title=Ethnohistory.&atitle=&pid=DocNumber=000781707,Ip=ocean.lib.uiowa.edu,Port=9929](http://infolink.lib.uiowa.edu:8080/sfx_local?sid=ALEPH:IOW01&genre=&isbn=&issn=0014-1801&date=999&volume=&issue=&spage=&aulast=JSTOR%20(Organization)&aufirst=&aunit=&title=Ethnohistory.&atitle=&pid=DocNumber=000781707,Ip=ocean.lib.uiowa.edu,Port=9929)

In addition to citation links, OpenURL servers can carry a search from any OpenURL enabled source -- an online catalogue, integrated library system staff modules -- to a range of Web targets such as other catalogues, indexes, online bookstores, book reviews, interlibrary loan services, search engines.

In Figure 4, the example from Suffolk County Public Libraries shows:

- a subject search in the OPAC on **ethnohistory**
- retrieval of a bibliographic record with OpenURL button **Other Resources**
- selection of the **Other Resources** button displays links carries the subject headings from the bibliographic record to other target resources specified by the library

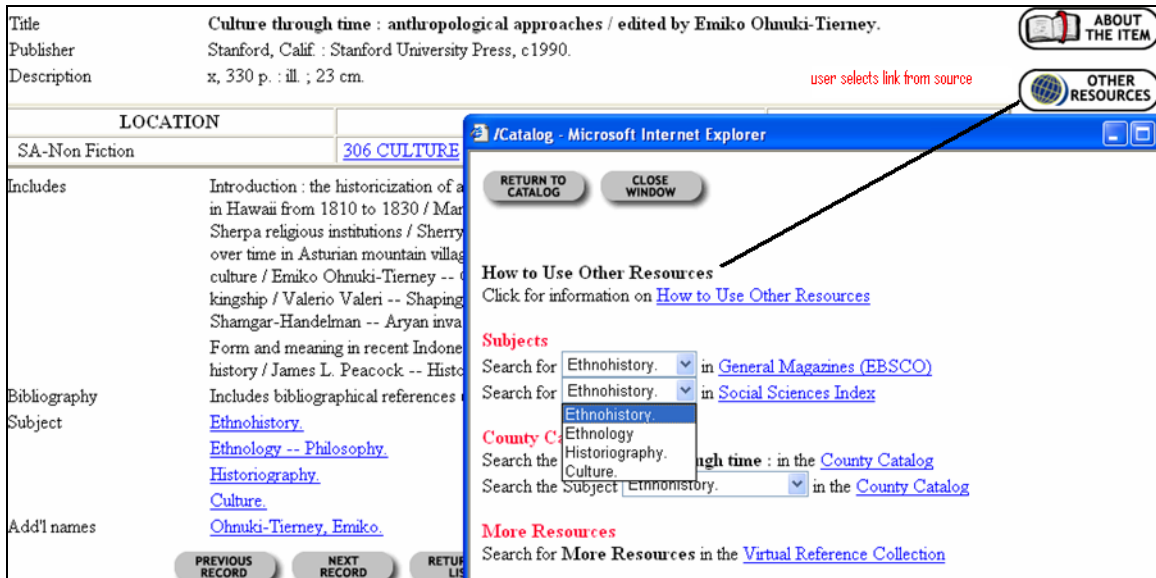


Figure 4: Sample of search results linking to resources

Implementing OpenURL

Key benefits of an OpenURL linking service are that it optimises the use of an institution's licensed full text e-journals, and enables libraries to expand their online catalogue services to offer multiple Web resources. How do you go about implementing it in your institution?

OpenURL products are now available from a range of library automation vendors and information providers including IGate (Openly Informatic), JournalLinker (Serials Solutions), Horizon Link Resolver, LinkFinder*Plus* (Endeavor), LinkSource (Ebsco), OL2 (Fretwell Downing), SFX (ExLibris), Sirsi Resolver, WebBridge (Innovative Interfaces).

The product can be purchased as part of an integrated online system, or as a separate standalone product to integrate with the system. Other libraries have developed do-it-yourself OpenURL systems, e.g. OhioLink Libraries, UKOLN Link resolver. UKOLN's OpenResolver web site includes a demonstration of OpenURL that shows linking from citations:

<http://www.ukoln.ac.uk/distributed-systems/openurl/>

A number of Australian university and research libraries are implementing the SFX OpenURL product as part of the government funded AARLIN research portal. (For more information, see **http://www.aarlin.edu.au/about_phase2.shtml**).

Setting up an OpenURL linking service will vary according to the product and how it is designed, but the fundamental steps are as follows:

1. Identify **sources** – where to offer the OpenURL link from. Online catalogues, abstracting and indexing services, library system modules. They must be OpenURL compliant. A recent list is at **<http://www.sfxit.com/sources.html>**
2. Obtain SID (service identifier) from vendors. An SID can include a database name as well as vendor, e.g.
 SID=Ovid:Medline
 SID=Ebsco:MFA

- Create, update the link syntax for sources, including vendor SID and PID (private subscription identifier) if necessary.
3. Advise external sources of your local institutional resolution server URL to enable them to place OpenURL links from citations back to your server and resources.
 4. Identify **target** resources for your community – e-journals, full text providers, online catalogues, search engines, online bookstores. Any Web resource with a predictable URL can be a target.
 5. Maintain and customize your library's collection of e-journal holdings (knowledge base) for linking to targets; add local subscription and licensing data. Some linking products include adding and updating A-Z journal lists into the knowledge base.
 6. Specify the metatags and conditions for which parts of citations and searches are to be carried across to linking services, and from which sources.

Information providers can become OpenURL compliant sources and offer institutional links from references or citations by implementing the OpenURL mechanism called a *hook*. When added to resources the hook enables the insertion of linking server URLs which take users back to their institutions and a list of target resources.

OpenURL integration with CrossRef and DOI concepts

Libraries can link to CrossRef as part of OpenURL linking and a number of information providers and researchers have experimented with incorporating the two processes. CrossRef (<http://www.crossref.org>) is a collaborative network that acts as a registry of full text articles for over 170 publishers. It facilitates linking through the use of a DOI (Digital Object Identifier), an open standard. CrossRef DOIs link to publisher response pages, which include the full bibliographic citation and abstract, as well as full text access (for authenticated users or at no charge, as determined by the publisher). The publisher response page may include other linking options, such as pay-per-view access, journal table of contents and homepage. A user following a link based on a DOI will be taken to the publisher's site. If the user or their institution does not have a registered subscription they may be asked to pay a fee for access. CrossRef tries to solve the *appropriate copy* question by providing a reverse DOI lookup to identify a title.

A DOI is an alphanumeric name that identifies digital content, such as a book or journal article. The DOI is paired with the object's electronic address, or URL, in an updateable central directory, and is published in place of the URL in order to avoid broken links while allowing the content to move as needed. DOIs are distributed by publishers and by CrossRef, and there is no end-user charge associated with their use. An OpenURL can include a DOI as an identifier. For more information on DOI see <http://www.doi.org/>.

Developments

While the initial focus of OpenURL has been in the scholarly research information environment, the NISO AX Committee is charged with generalizing and extending the standard beyond the bibliographic citation or record into a wider information arena. NISO votes to accept OpenURL 1.0 as a standard in early 2004.

Some OpenURL information providers are experimenting with Shibboleth, the patron verification security project being developed by Internet2. Shibboleth is developing open standards solutions for authorizing users to access information resources, as an alternative to IP address

authentication, login names and passwords and proxy servers. For more information see <http://shibboleth.internet2.edu/>.

OpenURL is a key library and information standard, like MARC and Z39.50 that has been adopted by vendors, publishers, and libraries. But this one goes further: OpenURL encourages co-operation between commercial and non-commercial information providers, delivering a much-needed solution to the overwhelming array of Web-published resources. Such events don't happen without vision, effort, and lobbying by individuals and organisations. It is significant that the pioneering research came from a recognised need within a library, and we have Herbert Van de Sompel and Patrick Hochstenbach to thank for identifying the initial problem and delivering such an elegant solution.

Further information:

Blake, Miriam. Implementation of the OpenURL and the SFX Architecture in the production environment of a digital library. *Victorian Association for Library Automation Biennial Conference*, Melbourne, Australia, 6-8 February, 2002.
<http://www.vala.org.au/vala2002/2002pdf/39Blake.pdf>

NISO workshop: The next generation of access: Metasearch and OpenURL, October 29 – 30, 2003, Washington DC. http://www.niso.org/news/events_workshops/MS-2003_ppts.html

OpenURL NISO standard
<http://library.caltech.edu/openurl/>
<http://alcme.oclc.org/openurl/servlet/OAIHandler?verb=ListSets>

Powell, Andy. "OpenResolver: A Simple OpenURL Resolver." *Ariadne* Issue 28, June 2001
<http://www.ariadne.ac.uk/issue28/resolver/>

Robertson, Wendy C., Paul Soderdahl. "Everything you always wanted to know about SFX but were afraid to ask." *Serials Librarian*, 47(1/2), 2004.
<http://staffweb.lib.uiowa.edu/psoderdahl/papers/sfx.pdf>

SFX overview (Ex Libris)
<http://www.sfxit.com/OpenURL/>

Van de Sompel, Herbert and Patrick Hochstenbach. "Reference Linking in a Hybrid Library Environment. Part 1: Frameworks for Linking." *D-Lib Magazine*, 5(4), April 1999.
http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html

Van de Sompel, Herbert and Patrick Hochstenbach. "Reference Linking in a Hybrid Library Environment. Part 2: SFX, a Generic Linking Solution." *D-Lib Magazine*, 5(4), April 1999.
http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt2.html

Van de Sompel, Herbert and Patrick Hochstenbach. "Reference Linking in a Hybrid Library Environment. Part 3: Generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment." *D-Lib Magazine*. 5(10), October, 1999.

http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html

Van de Sompel, Herbert and Oren Beit-Arie. "Open Linking in the Scholarly Information Environment Using the **OpenURL** Framework." *D-Lib Magazine*, 7(3), March 2001.

<http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html>

Van de Sompel, Herbert, and Oren Beit-Arie. "Generalizing the OpenURL Framework beyond References to Scholarly Works: The Bison-Fute Model." *D-Lib Magazine*, 7 (7/8), July/August 2001.

<http://www.dlib.org/dlib/july01/vandesompel/07vandesompel.html>

Van de Sompel, Herbert, Patrick Hochstenbach, and Oren Beit-Arie. "OpenURL Syntax description." November, 2001.

http://library.caltech.edu/openurl/Record_Documents/OpenURL_Version_0.1.mht