OLAQ

July 2014

Open URLS and Reference Linking: Research and Practical Application in Libraries

Mark Dahl
Lewis and Clark College

OpenURLs and Reference Linking:
Research and Practical Application in Libraries

by Mark Dahl
Library Technology Coordinator
Aubrey R. Watzek Library
Lewis and Clark College

A long-standing problem in library reference work is connecting an article citation found in an index to the article itself. The OpenURL is a new tool that helps address this age-old problem in a world of Web-based information resources.

Recently, there has been much discussion in the library community surrounding OpenURLs. In fact, a new and somewhat confusing vocabulary has arisen around the whole concept. This article will seek to clarify what OpenURLs are and then explain some of the current implementations of OpenURLs in libraries today.

What are OpenURLs?

Currently, most libraries provide indexing and abstracting databases via the Web site of a commercial information vendor such as EBSCO, Gale, or ProQuest. Sometimes the library is given the ability to customize the look and functionality of these databases. But for the most part, libraries are stuck with what the vendor gives them.

Often the citations in the database are not connected directly with the other related information offerings of the library, such as print holdings of the journal in which an article is located, or a document delivery service that could obtain an article. The user of the database must cut and paste pieces of the citation (typically a journal title) into another search tool that can tell them if their library holds the resource.

Some databases do offer the ability to link from a citation to an entry in the local library's catalog that shows the library's holdings of the cited journal. Some databases also offer facilities for linking from a citation to an interlibrary loan request. These are examples of what may be termed *extended services.*

An OpenURL is a device that may be used in indexing and abstracting databases, as well as other places, to link to extended services that are independent of the information provider where the link originates. These extended services may include an article's full text, biographical information about the author of an article, local library holdings information, a Web search on the subject of an article, etc.

In more abstract terms, OpenURLs are URLs attached to resources that are used to link to related resources. The structure of an OpenURL is composed of two parts: first, the base URL is the *service component* or *resolver.* This is typically the address of a library-controlled server that processes the data in the rest of the OpenURL.

The rest of the URL is called the *descriptor.* The descriptor can contain the source of the OpenURL (e.g., the database that created the OpenURL—OCLC FirstSearch, EBSCOhost, etc.), and metadata about the article (or other information object) that the OpenURL describes. In the case of a

Using OCLC FirstSearch Administrative Module to activate OpenURLs. Source: OCLC FirstSearch.
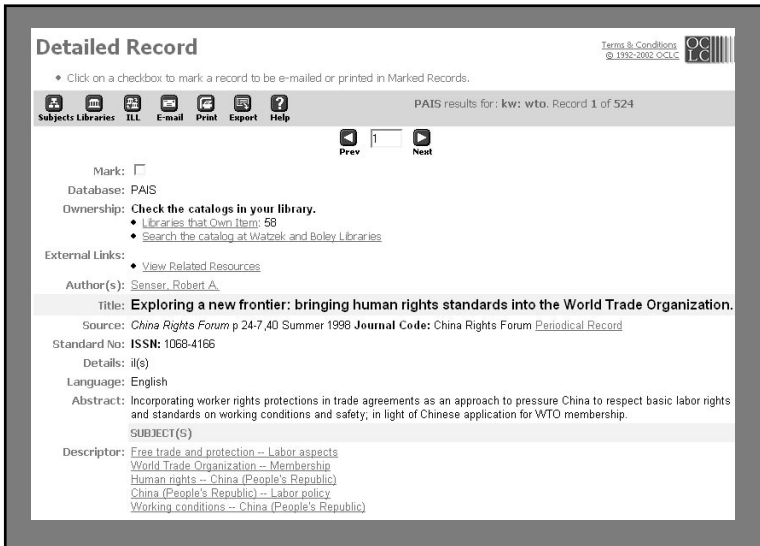
A citation in FirstSearch with a link to an OpenURL resolver. The link is labeled "View Related Resources." Source: OCLC FirstSearch.

journal article, this would include the journal name, its ISSN, the volume and issue numbers, the author's name, the title of the article, etc.

Here is an example of a citation and its corresponding OpenURL:

Hagedorn, George A.: Raising and lowering operators for semiclassical wave packets. *Ann. Physics* 269 (1998), no.1, 77–104

    http://sfx.caltech.edu:8088/
    caltech?sid=Caltech:
    WOS&genre=article&atitle=Raising
    %20and%20lowering%20operators
    %20for%20semiclassical%20wave%
    20packets%2E&title=ANN
    %20PHYSICS&issn=00034916
    &date=1998&volume=269&issue=
    1&spage=77&epage=
   104&aulast=Hagedorn&auinit=GA

The OpenURL's first line, "http://sfx.caltech.edu:8088," is the resolver, which can also be known as a link server. The rest of the URL is the descriptor, which contains the information about the article that the link server needs. The format in which the descriptor portion of the URL (with the familiar question marks, ampersands, equals signs, etc.) resides is the common "query" format for embedding data in URLs. This data describes the object for which the link server needs to find related links. The name-value pairs in which the data, or more appropriately, metadata, is given (issn=00034916, date=1998, etc.) are part of the OpenURL standard.

Typically, a user would start out using an electronic index, find an article of interest, then click on the OpenURL attached to that article to find out what kinds of extended services related to that article were offered by the user's library.

In the background, by clicking on the OpenURL, the user would be taken off of the vendor's Web server and on to a server controlled by the local library. This would be the OpenURL resolver or link server. The library would have configured the resolver to provide the user with the appropriate services relevant to that article. For example, it might offer a link to the article in a full text database to which the library subscribes and/or offer a link to local print holdings of the journal. Given the information in the OpenURL, it might be smart enough to know that the article was not available at the local library and automati-
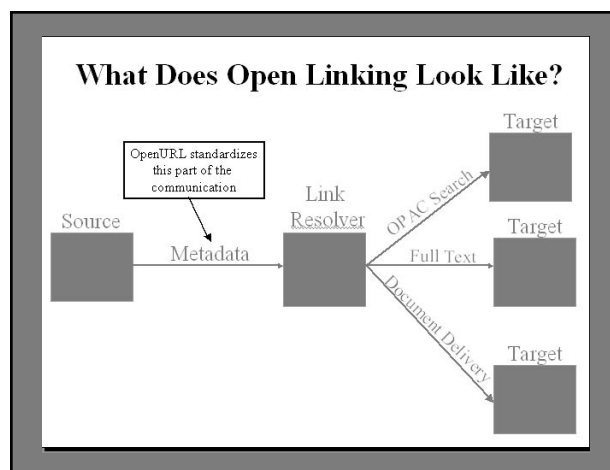
This demonstrates the flow of information in the OpenURL framework. Source: Van De Velde and Samuels, 2001

cally set up an interlibrary loan request.

OpenURLs also can be used for journal article citations within electronic journals and for monographic citations. They can be used to link books, articles, reference sources, Web sites, and other information objects together.

The Research and the Standard

OpenURLs were conceived and developed by Herbert Van de Sompel and his colleagues at Ghent University in Belgium. In the late 1990s, Van de Sompel recognized that information providers (i.e., database vendors) were providing links with their citations, but that often those links only pointed to resources within a vendor's own realm. Perhaps they provided a link to full text of the resource, but only through that particular vendor's own fee-for-use service.

Van de Sompel described these vendor-provided links as non-context sensitive (meaning they did not take into account who was accessing the resource or where they were coming from), limited in scope (meaning they often included only the information universe associated with the vendor) and closed (meaning that they did not allow the library to supplement these links with other resources).

Responding to these shortcomings, Van

de Sompel embarked on what he and his colleagues termed "SFX research" (SFX for special effects). They sought a linking solution that was context-sensitive, unlimited in scope, and open. Over the course of several experiments, they developed the OpenURL framework described above.

The OpenURL framework is context sensitive because it allows the library to control the link server that provides related resources. If the library so desires, it can tailor the related resources that it provides by the IP address or other means. It can point a user to the most inexpensive copy of a resource given its own contracts with information providers, thus solving what has been termed the "appropriate copy problem." Library control over the link resolver and the non-proprietary nature of the OpenURL protocol makes for limitless scope and completely open application. Any database can link into the resolver; and given a description of an article (or another type of information object) from an OpenURL, the resolver can relate that metadata to any number of other resources (Van de Sompel and Beit-Arie, 2001).

The OpenURL framework is non-proprietary and is being reviewed for acceptance as a NISO standard. If accepted as a standard, more information providers will, hopefully,
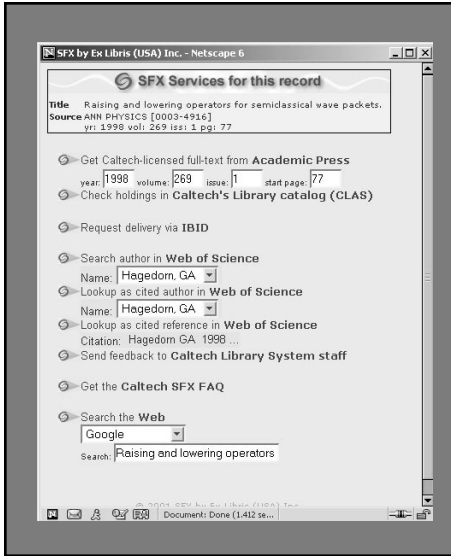
This is a sample set of extended services offered by a link server, in this case an SFX server.
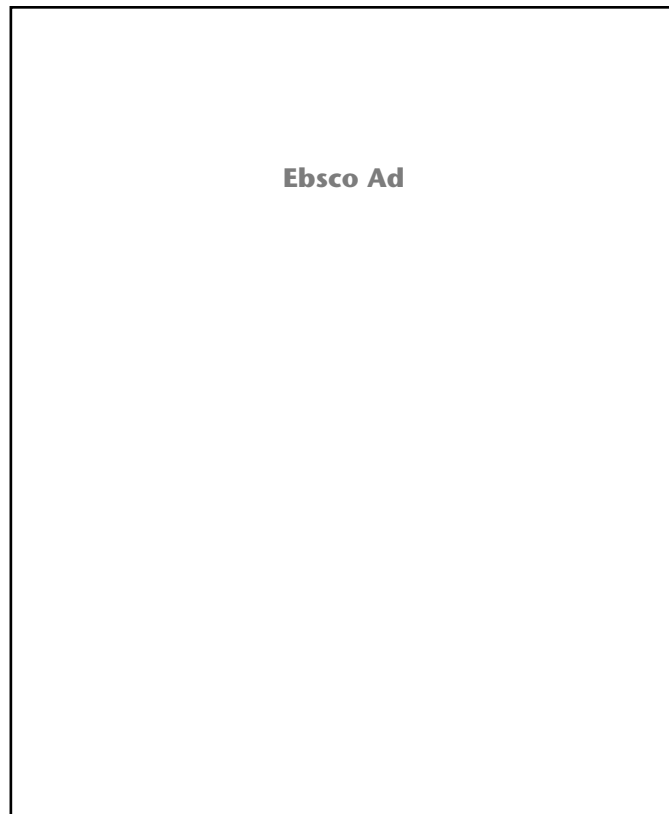Source: <http://library.caltech.edu/openurl/Documents/Demo.htm>

provide OpenURLs with their resources. In addition, any vendor or library may produce an OpenURL resolver. Many information providers and library automation vendors have already begun to design or modify their products around the standard.

Current working papers on the proposed OpenURL standard can be found at: http://library.caltech.edu/openurl/Working_Documents.htm.

Application of the OpenURL Framework

In order to use the OpenURL standard, libraries need to have OpenURL links attached to the citations that their patrons are viewing in their abstracting and indexing databases. They might also want such links to appear in the text of journal articles, in Web-based subject bibliographies, and next to bibliographic entries in their public catalogs.

Currently, a number of database vendors including EBSCO, ProQuest, Gale

Group, HW Wilson, and OCLC, offer the option of including OpenURLs alongside their citations. Electronic journal vendors, notably Ingenta, also are beginning to offer OpenURLs. CrossRef, a nonprofit initiative to link citations between electronic journal articles of various publishers, has now integrated OpenURLs into their linking framework. See http://www.crossref.org.

Setting up an electronic index to provide OpenURL links with citations need not be difficult. For example, to activate OpenURLs in OCLC FirstSearch, libraries need only login to the FirstSearch Administrative Module, and turn the feature on. From the Administrative Module, a librarian may simply enter the address of the library's local resolver. Once OpenURLs are turned on in FirstSearch, every citation viewed has an OpenURL link offered.

Purchasing and configuring a link server to deal with these OpenURLs is, perhaps, a more complex process. Library automa-

tion vendors have stepped up to the task of providing link servers. As OpenURL resolvers are designed to manage the resources of a particular library—much like an integrated library system—library system vendors may believe that it is a natural progression to move into the area of link management. At some point, a link server may be as essential an ingredient of an integrated library system as a circulation or acquisitions subsystem.

Rights to Van de Sompel's SFX server, a link server that his team designed in the course of their research, were purchased by Ex Libris, a library automation vendor in the United States. Ex Libris then developed their SFX server product, an OpenURL resolver. While Van de Sompel and his colleagues conceived the OpenURL standard as open, they recognized that the resolvers used to deal with OpenURLs could and would be proprietary. The SFX server may be purchased independently from Ex Libris' other library automation products. See http://www.sfxit.com/ for more information.

Endeavor Information Systems offers an OpenURL resolver product called LinkFinderPlus, also available as a stand-alone product. LinkFinderPlus comes with a pre-populated database of information providers and content. Presumably, the library selects the appropriate full-text databases and electronic journal aggregators, and LinkFinderPlus provides the patron with an appropriate list of full-text content options for a given citation. Such an approach could eliminate the need for special journal title databases populated from aggregator title lists or electronic journal management services such as Serials Solutions. See http://www.endinfosys.com/prods/linkfinderplus.htm.

Innovative Interfaces offers a product called WebBridge that appears to be more focused on creating links from citations in an OPAC to external resources. It allows libraries to add related links to records in their catalogs. Its promotional literature claims that it has the capability to resolve OpenURL links external to the catalog but does not explain the process of resolving such links (Davidson, 2001). See http://www.iii.com/html/products/p_map.shtml.

In general, library automation vendors are hoping to use OpenURL resolution as well as other linking initiatives to keep themselves at the center of libraries' information resources. In addition, there are other OpenURL servers available besides those offered by traditional library automation vendors, notably 1cate from Openly Informatics. See http://www.openly.com.

Libraries also have the option of creating their own OpenURL resolvers. The data embedded in an OpenURL is easy to extract using a CGI-script or Web scripting language like PHP. I recently OpenURL-enabled the Web-based journal title database at Lewis and Clark College. When pointed to by an OpenURL, this journal title database checks for any journals with an ISSN matching that specified in the OpenURL parameter. It then returns the appropriate journal title record.

This implementation of the OpenURL protocol is not fully realized, as it is only using a small part of the metadata available (the ISSN) and is only offering journal holdings information—no other related resources. But it is a start. A future enhancement could be a link to an interlibrary loan request—already filled out with citation details—should the journal title search fail. Libraries could create more advanced OpenURL resolvers with any number of features.

Perhaps more interesting than the development of OpenURL technologies is the new role that reference-linking brings to librarianship. Librarians now will have the ability establish rules for offering related resources within the configurations of their link servers. Determining what to offer could be a complicated affair. Given

Also see the Department's Oregon Community Profiles (nice descriptions of local cities with a host of information on economy, climate, festivals, sites, recreation, etc.). http://159.121.111.9/profile.htm

References

American Statistics Index. Congressional Information Service, Washington, D.C.

Batten D. (ed.), 2000. *Guide to U.S. Government Publications.* Gale, Detroit.

Rand McNally Commercial Atlas & Marketing Guide. Rand McNally, Chicago. (Annual.)

United States Bureau of the Census, 1976. *Historical Statistics of the United States, Colonial Times to 1970.* Basic Books, New York.

Open URLs
Continued from page 6

a citation from a science database, what additional resources make sense? Science-specific biographical sources, citation indexes, search engines, other reference resources? What different types of resources are important for a humanities citation?

Reference linking using OpenURLs has gone from the theoretical to the possible. But will it really be useful to library patrons? I think librarians can rightly assume that given a citation, users will want to get to the full text of the article. OpenURLs definitely will help with this issue. But will users also appreciate an abundance of other related links attached to a citation? This remains to be seen.

References

Davidson L., 2001. Integrating access to digital collections: MAP from Innovative Interfaces. http://www.iii.com/pdf/map_wp.pdf

Powell A., 2001. OpenResolver: a simple OpenURL resolver. *Ariadne* 28. http://www.ariadne.ac.uk/issue28/resolver/

Van de Sompel H., and Beit-Arie O., 2001. Open linking in the scholarly information environment. *D-Lib Magazine* 7(3). http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html

Van de Velde E., and Samuels H.E., 2001. Open linking and the OpenURL standard. Presented at the CNI Fall Task Force, San Antonio, November 2001.

Walker J., 2001. Open linking for libraries: the OpenURL framework. *New Library World.* 1: 127–133. http://www.emerald-library.com/ft