

OpenURL and SFX Linking

By Jenny Walker

Abstract:

It is only a few short years since the early academic research lead by Herbert Van de Sompel at Ghent University in Belgium that resulted in the formulation of the OpenURL standard and the development of the SFX link server. These complementary components have revolutionized linking for libraries, putting control for linking in the hands of the librarian.

This paper considers briefly the early linking initiatives, explains the purpose and structure of the OpenURL and the role of linking components—the sources, targets, and the link server, with a focus on the SFX link server.

Jenny Walker (jwalker@exlibris-usa.com) is VP Marketing and Business Development for the Information Services Division at Ex Libris (USA) Inc.

OpenURL and SFX Linking

Introduction:

OpenURL and SFX are terms that are now quite familiar to many librarians – though not all may be familiar with the concepts behind these and the benefits they bring. Conferences and conference sessions are dedicated to linking issues and in particular the Open Linking Framework, first proposed by Van de Sompel in his early research papers published in DLib in 1999. Van de Sompel's research led to a new style of linking based on the OpenURL standard and using local link servers such as SFX. This was known as the Open Linking Framework.^{1 2 3 4}

At the 2002 Charleston conference held at the end of October, OpenURL and SFX were mentioned by many of the speakers and treated as de facto standards. Only a year previously many were hearing these terms for the first time, and few had any practical experience of implementing and using linking servers.

SFX linking has now acquired recognition in the marketplace as linking that is based on the OpenURL standard. But SFX is also the link server solution from Ex Libris, based on the technology developed at the Ghent University and acquired by Ex Libris in February 2000. Other OpenURL-based link servers have emerged in the marketplace. Some of these are commercial products whilst others are homegrown solutions from groups such as OhioLink and the Colorado Alliance.

But why are the OpenURL standard and link servers such as SFX so important for libraries today?

First - why link?

But firstly, why link at all? This is probably a question no longer asked by many libraries. Even those libraries that have not yet implemented existing linking technologies now understand the need to do so to best serve their communities. But if there is any doubt, let's take a quick look at the library environment today.

Any library, no matter its size, its user population, or its specialization, will make many different kinds of material available to library users. Materials may originate from different vendors and could be library catalogs, abstracting and indexing databases, full text databases and other data repositories. Services such as interlibrary loan and document delivery provide means to acquire materials that are not hosted by the library and should be integrated into the library environment. Reasons for purchase of disparate materials may vary from subject matter and importance to researchers to cost and accessibility. Because materials are of different types, it is often difficult to organize and present these materials to users under one common interface or even with one research approach. As a result, bibliographic instruction classes have become more focused around vendor-specific pathways rather than common-sense research approaches.

Though they are of different types, materials contained in different places in the library environment – or available on the Web - are often closely related. A citation in an article reference or abstracting and indexing database may have its full-text counterpart available in an electronic journal subscribed to by the library, via an aggregator database, or in a bound volume available in the library stacks. The authors of one article may be cited in other related articles of potential use to a library user. A document delivery service may be available for information not specifically held by the library. Book table of contents, book reviews, book jacket pictures, etc. may be available for book records in the library catalog.

In the library environment users now expect to be able to link from one resource to another; and not having links can be considered a serious drawback.

‘Old-Style’ and Vendor-Specific Linking

It’s not hard to set up an initial set of links from one library resource to another. All it takes is the knowledge of the URL (uniform resource locator, or Web address) of the item you wish to link *to* -- let’s call it a link Target -- and the ability to embed this URL in the item you wish to link *from* – let’s call it a link Source.

Linking by means of embedded URLs is often called static linking and it is quite practical – that is, straightforward in initial setup -- but often not very durable. Journals may move from publisher to publisher, or vendors change the maintenance and makeup of their servers, and these hard-coded URLs may change. Further, to create these static links in a record, the information service provider must have knowledge, at the time of creation of the record of all resources to be interlinked. This approach cannot easily be scaled up to deal with large numbers of resources, may cause delays in the delivery of information and the resulting links may not be reliable.

To address the constraints imposed by static linking, some vendors adopted dynamic linking solutions whereby the links are calculated at the time the user requests the link. However, these vendor-specific solutions tend to be closed solutions, with control of the linking being in the hands of the information service providers who are focused on maximizing the link traffic to their sites, rather than in the hands of the librarian who would like to maximize usage of the varied resources for which the institution has already subscribed. Further, the scope of the links offered is limited typically to full-text retrieval, or to holdings lookups in online catalogs; although we are now increasingly seeing a richer range of extended service offerings such as links to citation databases, related Web searches, and online bookstores. Lately vendors tend to allow libraries some localization options for the links available from their product, but what this means in practical terms is that the library must continually update the information for *each* vendor. Local holdings change too often for this task to be manageable. If changes are not applied regularly, users may find that they are barred from access to the service offered to them. This leads to user frustration and in some cases may result in the user purchasing a service for which the institution has already paid, but via an alternate provider.

The OpenURL – the way forward

Under the static linking scenario, a Source database contains an embedded or hard link to a Target item. Resolution of links – making sure the user is sent where they ought to go - is the job of the Source database, meaning that if the link changes, it must be updated *in the Source record*. Also, each hard link by its very nature can go only to one Target: one link, one Target.

The introduction of vendor-specific dynamic linking scenarios enabled some diversity of links with local customization; but under the overall control of the vendor of the source database.

How about a new scenario whereby the resolution of the links is no longer the job of the source database but rather is placed fully under the libraries' control? This is achieved through the introduction of a third component placed in the middle of the transaction, between the source and the target: a local link server. The end user still sees a link in the source database and is prompted to click on it. However, instead of being a URL that describes the target document, this link is special: it is an *OpenURL* that describes the item or citation at which the user is looking.

This is the new way to provide links.

The OpenURL, already a de facto standard for linking, is well on the path to NISO adoption. It is a standardized way to pass metadata describing a record, from the interface that displays that record to a link server that is under a library's control. The library, therefore, can determine what links should be offered to the user. Adopting the OpenURL linking framework, libraries are able to set up links between the resources to which they subscribe, as the librarians deem helpful for their users, not in the manner determined by the vendors. Once implemented, the OpenURL and link servers make it possible for users to navigate seamlessly and in a meaningful way among the resources that librarians have selected for them.

For vendors this is an effective mechanism to ensure that their databases become fully incorporated in the library's interlinked environment. According to Walt Crawford, Senior Analyst at RLG,

If properly implemented, OpenURL is a win-win situation. Good abstracting and indexing services become valuable by linking to local resources. Licensed resources and print holdings see more use because the link from the identification to holdings is fast and easy. None of this requires fancy new numbers; the information is already there – ISSN, journal and article titles, year, volume, and so on.⁵

The OpenURL has two key components:

1. The BASE-URL that determines the address of the link server to which the OpenURL should be sent. Sites must provide this information to the information

- service providers for each link source. Notifying the source database of the BASE-URL can be automated.
2. The QUERY or “content”. This can be descriptive metadata elements and/or identifiers that are used to gather the metadata.

(Insert Table 1 – SFX Article Table 1.doc)

An important illustration of the ongoing penetration of the OpenURL linking framework, is its integration with the DOI/CrossRef linking solution, providing a solution to the problem with DOI/CrossRef whereby the user’s context is not taken into consideration in the delivery of links and hence leading to immense user frustration. This integration has the support of important groups such as PILA (Publishers International Linking Association, Inc.), IDF (International DOI Foundation), DLF (US Digital Library Federation), and CNRI (The Corporation for National Research Initiatives), all of whom participated in early proof of concept prototypes. ⁶

OpenURL version 1.0 is in its final stages of development by NISO and is due to be available for testing by vendors and libraries from April 2003. It builds on the existing syntax description by introducing further mechanisms for extensibility – including further information about who is the user, and from where they are coming. It will accommodate metadata fields that cover a broader range of information types beyond just textual scholarly information.

More information about NISO committee AX and the OpenURL itself can be found at: <http://library.caltech.edu/openurl>

Metadata for linking

Through use of the OpenURL the metadata that exists in a record, for example in a citation database or in a reference in a bibliography, can be sent to a local link server and used by that link server to supply creative types of services that are under the control of the local library. Here are a few examples showing the relationship between metadata available in database records and *extended services* that libraries may wish to create:

- ISSN, year, volume, issue, and start page can be used to create a link from the citation to the cited article at a publisher’s Web site or in an aggregator database;
- ISSN or journal name can be used to check for print holdings in the library catalog, whether or not electronic full-text is available;
- Author names can be used to look up the authors in a citation database to see other articles they have written or how well-cited they are;
- Subject terms from the original citation can be re-used in other related databases or to link to Web sites that librarians judge potentially useful;

- ISSN or journal name can be used to look up the journal in a serials directory to find out more general information about the journal – its publication schedule, where it is indexed, publisher information, etc.

The important thing to remember is that the links created in these situations are *appropriate* – meaning that local librarians take an active role in their creation; they appear or do not appear subject to local decisions; *and* they link to material that is really available to the users of that library.

Citation databases are an obvious good choice for link sources because they usually contain no full text of their own. However, many other databases are equally applicable as link sources. If integrated library system vendors implement the OpenURL, then the library OPAC could also serve as a link source and there would be no need to maintain 856 links since the link upkeep would be the job of the link server.

Use of the OpenURL for linking is not limited to the vendor world; local databases containing local faculty citations or company product literature can be incorporated into the interlinked environment through the implementation of OpenURLs.

Local Link Servers

The OpenURL does not offer much value on its own. Its power is realized through a link server such as SFX. Link servers are configured by librarians to determine what types of links their users should see, and to ensure that the links are accurate and effective. These link servers come in different flavors, from different vendors, often with different capabilities and features and management options.

Local Link servers, also known as link resolvers, resolve the metadata received on the OpenURL to appropriate services, creating new pathways for their users.

In some cases the OpenURL metadata received from the link source is incomplete and needs to be supplemented in order to create effective links to a target service. Further, in some cases the source and target follow different rules. The source database may use a print ISSN in an article citation that is then passed on the OpenURL; it may even use the journal name or abbreviated name rather than ISSN. However, to link to the electronic full text version of the article an eISSN may be required. A link server therefore “adds natural intelligence to the computer-to-computer process to navigate complexity” says Crawford and continues that it “offers more power than a pure computer-to-computer protocol – and, well, it makes more fun.”⁵

The many advantages of a link server for library stakeholders include:

- For user services librarians: it allows for standardization of link services across resources, reducing training needs. Users will see standard buttons appearing in many, if not all, their resources. Clicking on an SFX button is intuitive.⁷

- ❑ For technical services librarians: a link server offers a single point of administration of the various services across different data sources. The link server is configured with “knowledge” about all the resources to be interlinked and the link server then delivers up the links. No more uploading holdings files to many different information service providers to enable linking.
- ❑ For library managers: a link server can capture and report on a wealth of statistics on the use of distributed resources as has never before been available.

The SFX Link Server

SFX, the link server from Ex Libris, is typically hosted by a local library, but whether hosted locally or remotely (through the Ex Libris-hosted subscription service), it remains under the control of the library to determine what types of links should be offered to their users and to where the links should resolve.

For SFX to work successfully, it must have access to information about the library’s subscriptions, it must know what types of SFX services are required and finally, it must know how the library wishes to present such services, both in terms of the interface and in determining the conditions for presentation. This information is held within an underlying repository, known as the SFX KnowledgeBase.

- Information about the local collections: the SFX KnowledgeBase contains information on different resources that is *generally* applicable. This includes information on possible SFX sources - databases that can deliver an OpenURL; and SFX targets - e-journals, aggregator databases, catalogs, Web sites, etc. to which links can be made. Librarians must localize their SFX KnowledgeBase by activating the items that match their local subscriptions. For example, if a library subscribes to some, but not all, of the journals from ScienceDirect, it must activate only those journals in the SFX KnowledgeBase.
- Definition of *potential* services: when a library implements SFX, the local librarians decide what kinds of services they would like to present to a user “in a perfect world” – that is, if all metadata elements are present -- and configure the SFX server accordingly. Potential services could include links to the full text, holdings and TOC; a document delivery form; and a link to further searching using a Web search engine, etc. SFX will determine the relevancy of a certain service based on the metadata provided for the specific object of interest.
- Rules supporting a decision on the relevance of services: These are rules placed on individual items or groups of items (e.g., all journals of a certain publisher) in the database to ensure that the services returned on the SFX menu will be optimal. Rules could be set to determine the availability of full text service for a particular e-journal published in a certain time period, or to suppress a document delivery option if the full text is available electronically. Rules could determine that a link should be made directly to the full text of an article or ebook if available electronically or if a menu of options should be presented. Further, if an article

were available to a user from a number of different targets, rules can be applied to determine the library's preference for presentation to the user.

A library is not limited to what is already configured in the SFX KnowledgeBase; it is also possible to set up original, local targets in SFX. Links to such targets work best if there is some kind of link-to or search syntax that can be used by the SFX server. As there is not yet a prescribed standard for a link-to syntax -- or inbound linking -- the way there is for the OpenURL or outbound linking, configuration of targets needs to adapt the link to the specific syntax expected by the target.

To date SFX customers have defined and implemented many different and interesting target services e.g. "Look for other books on this subject in Books in Print", "Look up this movie in the Internet Movie Database", "Look up this journal in Journal Citation Reports," etc. With OpenURLs delivering metadata contained within databases and available link-to syntaxes enabling re-use of that metadata, libraries can potentially link from anywhere to anywhere.

The underlying SFX KnowledgeBase in each server is regularly updated by Ex Libris with changes in aggregator content, journal moves from publisher to publisher and when new journal backfiles are made available from the publisher. New targets are also added on a regular basis.

Ex Libris provides Web-based tools for local administrators to manage the KnowledgeBase. Local collections can be reflected one-by-one in the KnowledgeBase, or by batch loading if local subscription information, particularly on e-journals, is available elsewhere, e.g. in an Excel spreadsheet or other database, or from your subscription agent or other provider such as SerialsSolutions.

The SFX Flow

The SFX linking process itself involves a number of steps:

1. A user accesses an information resource and is presented with an SFX button for a particular citation or bibliographic reference.
2. When the user clicks on the SFX button a metadata string is assembled for the citation or reference and this is sent in a standard syntax – the OpenURL – to the user's local link server.
3. The SFX server, on receiving the OpenURL, extracts the metadata elements, retrieves or "fetches" additional metadata for this record if required and compares the metadata elements with those in the KnowledgeBase. A metadata "fetch" may be required if, for example, the only element on the OpenURL is a DOI or a PubMed Id. For a DOI SFX must fetch the descriptive metadata from the CrossRef database and for a PubMed Id from the PubMed database. SFX KnowledgeBase determines if and when a "fetch" is required.
4. SFX evaluates appropriate services to be shown on the SFX Menu based on metadata received and contents of the localized KnowledgeBase. For example, a full-text link can be presented if the metadata contains an ISSN – or equivalent

- title, abbreviated title, or eISSN – that exists in the rules database and the volume and date are within the ranges set for that ISSN/eISSN.
5. SFX presents an SFX Menu to the user with a number of options.
 6. When the user clicks on any of the menu items, SFX dynamically computes a link to the chosen Target Service. Once again the KnowledgeBase is used to determine the rules for creating the actual URL to the selected target. Further, SFX can replace metadata in the incoming OpenURL, in order to comply with the link-to syntax of the Target: for instance, it can replace a print ISSN, with the corresponding eISSN, extracted from the KnowledgeBase, to link to the full text of an article. As no standards exist for linking *to* targets (the OpenURL standardizes the linking *from* the source), the specific syntax for each target service must be held in the KnowledgeBase.

Metadata, whilst not always perfect, is key throughout this linking process. The richer it is, the more services the link server can provide to end users. As noted by Blake and Knudson, “Metadata, however different in its content and markup, is the one thing available in all text-based systems”. Blake and Knudson note that this increasing reliance on metadata for linking may lead towards more useful systems, and metadata is not only becoming more than a descriptive resource, but rather a tool in and of itself.⁸

(Insert Figure 1 – see jenny-411.jpg)

Fig. 1. User Workflow for SFX linking – from an SFX *Link Source* via an SFX *Link Services* Menu to an SFX *Link Target*, in this case the full text of the article.

Extended uses of a link server

Once a library has localized their link server, this information is quite likely the most accurate information on electronic journal holdings available to users of that library, and this information can be re-used in other ways, including:

- To generate a journal title list. Many libraries offer a browseable list of e-journals on their Web sites and these are maintained as static lists. With SFX, these can be generated directly from the KnowledgeBase. Furthermore, the links provided for each e-journal on the list are not static links, but links that are generated by the SFX server, thus enabling the library to gather statistical information for usage analysis.
- To pre-qualify document delivery requests. The library’s InterLibrary Loan Web form can be modified – or replaced by an SFX-supplied supplied form – which will first check for local holdings before accepting a document delivery request, thereby preventing unnecessary requests, particularly in an unmediated environment.

New Directions in Linking

So much has happened in the world of linking in so short a time, but what next?

While SFX puts linking under the control of the library, in the real world linking can be under the control of any – or all – of a number of different stakeholders, and not just the local library. Link servers are now developing to accommodate a range of different environments in which rights evaluation can be determined in a distributed manner by the appropriate authority. Such distributed authorities could include the local library, the library consortium and potentially other third party providers with whom the library has contracted. Work is already underway whereby two SFX servers are communicating with each other – server-to-server – to evaluate the generation of appropriate services based on the user's rights as a member of a local institution as well as the rights the user has as a member of a much larger consortium.

Standards developments such as OpenURL v1.0 and Shibboleth, the Internet2/Mace development relating to attribute-based authentication will continue to transform the linking landscape in the coming months and years. ^{9 10 11}

Summary

OpenURL and SFX linking have been variously described as “A Linkalicious Service”, “It's Magic”, and “It's fun”. ^{5 12 13} The very name SFX – special effects – was given by Van de Sompel who believed that SFX linking would offer to scholarly communication some of that magic offered by special effects in the film industry. I suspect it will not be long before these “special effects” become the norm and resources that cannot provide such enhanced access may not be successful.

It is a long way from those early SFX research days in Ghent. Van de Sompel must indeed be proud.

References

1. Herbert Van de Sompel and Patrick Hochstenbach, “Reference Linking in a Hybrid Library Environment. Part 1: Frameworks for Linking,” *D-Lib Magazine* 5, no. 4 (1999): #. Available from http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html (July 2002).
2. Herbert Van de Sompel and Patrick Hochstenbach, “Reference Linking in a Hybrid Library Environment. Part 2: SFX, a Generic Linking Solution,” *D-Lib Magazine* 5, no.4 (1999): #. Available from http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt2.html (July 2002).
3. Herbert Van de Sompel 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 no. 10 (1999):

- #. Available from
http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html (July 2002).
4. Herbert Van de Sompel and Oren Beit-Arie, "Open Linking in the Scholarly Information Environment Using the OpenURL Framework," *D-Lib Magazine* 7, no. 3 (2001): #. Available from
<http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html> (July 2002).
5. Walt Crawford, "OpenURL: Standards Can Be Fun!" *American Libraries* 33, no.7 (August 2002): #. Available from <http://www.ala.org/aloonline/crawford/cf802.html> (August 2002).
- 6 Oren Beit-Arie et al., "Linking to the Appropriate Copy," *D-Lib Magazine*. 9, no. 3 (2001): #. Available from <http://www.dlib.org/dlib/sept01/caplan/09caplan.html> (July 2002).
- 7 Miriam Blake, "SFX at Los Alamos," Presentation by author at American Library Association Conference, June 2001. Available from
<http://www.sfxit.com/publications/presentations/SFXALAJun2001Mblake.zip>
(November 2002).
8. M.E. Blake and F.L. Knudson, "Metadata and Reference Linking," *Library Collections, Acquisitions, & Technical Services* 26, no. 3 (2002): 219-230.
9. Herbert Van de Sompel and Oren Beit-Arie, "Generalizing the OpenURL Framework Beyond References to Scholarly Works: The Bison Futé Model." *D-Lib Magazine* 7, no. 3 (2001): #. Available from
<http://www.dlib.org/dlib/july01/vandesompel/07vandesompel.html> (July 2002).
10. NISO Committee AX. 2002. *Development of an OpenURL Standard*. Available from <http://library.caltech.edu/openurl> (March 2002).
11. Internet2/MACE. 2002. *Shibboleth Project*. Available from
<http://middleware.internet2.edu/shibboleth/> (November 2002).

12. K. G. Schneider, "SFX: a Linkalicious Service," *American Libraries* 32, no. 6 (2001): 118. Available from <http://www.ala.org/online/netlib/il601.html> (November 2002).

13. Judi Briden, "Implementing SFX @ University of Rochester,". Presentation by author at NERCOMP Conference, September 2001. Available from http://www.sfxit.com/publications/presentations/SFXNRCMP_share.ppt (November 2002).