Implementing SFX: Impact on Acquisitions and Cataloging

Judith Shelton
*Georgia State University*, jshelton@gsu.edu

William Walsh
*Georgia State University*, wwalsh@gsu.edu

Follow this and additional works at: https://scholarworks.gsu.edu/univ_lib_facpres

Part of the Library and Information Science Commons

**Recommended Citation**
https://scholarworks.gsu.edu/univ_lib_facpres/15

This Presentation is brought to you for free and open access by the Georgia State University Library at ScholarWorks @ Georgia State University. It has been accepted for inclusion in University Library Faculty Presentations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
Implementing SFX: Impact on Acquisitions and Cataloging

Judith Shelton and William Walsh
Georgia State University
COMO 2004
October 14, 2004
What is SFX?

- SFX is a link server from Ex Libris
- SFX:
  - Accepts OpenURL as input from an information resource*
  - Analyzes contents of OpenURL*
  - Evaluates appropriate services based on metadata*
  - Dynamically computes links to target services*

*From SFX implementation documentation
This presentation is not an endorsement (or a critique) of SFX.
We’re focusing on SFX because it’s what we use (and what we’re able to talk about).
There are other commercial link servers on the market…
Other commercial link servers include:

- Article Linker (Serials Solutions)
- Link Finder Plus (Endeavor)
  - LinkSource (EBSCO)
- 1Cate (Openly Informatics)

1Cate = one click access to everything
On Monday, SFX announced the “imminent” release of SFX Version 3.

According to their press release, features of version 3 include:
- OpenURL v1.0 Support
- SFX Menu Customization
- E-journal (A-Z) List
- Enhancements and new tools for the admin. Module

Ex Libris expects to beta test Version 3 for the next two months. They will support Version 2 for six months after Version 3 is officially released.

This presentation deals with SFX Version 2. While the issues discussed will remain the same, some details will change with the release of Version 3.
According to Ex Libris, “the OpenURL standard enables a user who has retrieved an article citation, for example, to obtain immediate access to the ‘most appropriate’ copy of that object through the implementation of extended linking services. The selection of the best copy is based on user and organizational preferences regarding the location of the copy, its cost, and agreements with information suppliers, and similar considerations. This selection occurs without the knowledge of the user; it is made possible by the transport of metadata with the OpenURL link from the source citation to a ‘resolver’ (the link server), which stores the preference information and the links to the appropriate material.”

<http://www.exlibrisgroup.com/sfx_openurl.htm>
What does an OpenURL look like? Let’s just take a quick look at an example, so those who haven’t seen one can see what it looks like in action. This is a screen shot for an author search for Peter Hernon in LISA.
Clicking the button creates the following OpenURL, which is sent to our SFX server:

- [OpenURL] http://sfx.galib.uga.edu/sfx_gsu1?sid=CSA:lisa-set-c&pid=%3cAN%3e131675%3c%2fAN%3e%26%3cAU%3eDugan%2c+R%E3%82%8b+Hemone%2c+P%3c%2fAU%3e&issn=0099-1333&date=2002&volume=28&issue=6&epage=80&spage=376&genre=article&aulast=Dugan&auinit=RE&title=Journal+of+Academic+Librarianship&atitle=Outcomes+assessment%3a+not+synonymous+with+inputs+and+outputs
Source i.d. = LISA: Library and Information Science Abstracts from CSA

- http://sfx.galib.uga.edu/sfx_gsu1?sid=CSA:lisa-set-c&pid=%3cAN%3e131675%3c%2fAN%3e%26%3cAU%3eDugan%2c+R+E%3b+Her non%2c+P%3c%2fAU%3e&issn=0099-1333&date=2002&volume=28&issue=6&epage=80&spage=376&genre=article&aulast=Dugan&auinit=RE&title=Journal+of+Academic+Librarianship&atitle=Outcomes+assessment%3a+not+synonymous+with+inputs+and+outputs

You can see that the sid, or source i.d., equals Library and Information Science Abstracts. This lets SFX know where the user and the citation are coming from.
Accession number = 131675
(Without the URL encoding it would read
“pid=<AN>131675</AN>”)

http://sfx.galib.uga.edu/sfx_gsu1?sid=CSA:lisa-set-
c&pid=%3cAN%3e131675%3c%2fAN%3e%26%3cAU%3eDugan%2c+R+E%3b+
Hernon%2c+P%3c%2fAU%3e&issn=0099-
1333&date=2002&volume=28&issue=6&epage=80&spage=376&genre=article&aulast
=Dugan&auinit=RE&title=Journal+of+Academic+Librarianship&atitle=Outcomes+ass-
essment%a+not+synonymous+with+inputs+and+outputs

pid stands for “private identifier.”

You can see that the OpenURL created also includes CSA’s accession number for this record.
You get the idea; the info in OpenURL corresponds to…

- http://sfx.galib.uga.edu/sfx_gsu1?sid=CSA:lisa-set-c&pid=%3cAN%3e131675%3c%2fAN%3e%26%3cAU%3eDugan%2c+R-E%3b+Hernon%2c&P%3c%2fAU%3e&issn=0099-1333&date=2002&volume=28&issue=6&epage=80&spage=376&genre=article&aulast=Dugan&auinit=RE&title=Journal+of+Academic+Librarianship&atitle=Outcomes+assessment%3a+not+synonymous+with+inputs+and+outputs

You get the idea, all the information in the OpenURL that is created (Authors, ISSN, volume, date, issue, start page, end page, and title) corresponds to…
...the information from the citation. This metadata is passed along to SFX which determines the services to display to the user in the SFX menu based on that information.
With SFX, we talk about targets and sources

- A source is where the SFX button or link appears. It’s where a user is coming from.
- A target is where a user goes after clicking a choice from the SFX menu.
- For example, a user performs a search in an A&I database like LISA (the source), clicks the SFX button in a record and from the SFX menu is taken to the full text of the article in ScienceDirect (the target).

I want to briefly touch on some of the terminology used with SFX and other link servers.
Out of the box, targets and sources must be activated and configured in SFX.

- For sources:
  - In general, all sources should remain active in SFX—even those not subscribed to.
  - You must request that sources be OpenURL-enabled. This is usually done by: 1) contacting the publisher or 2) entering your link server information in the source’s web admin. module.
  - In the past, not all sources supported a custom SFX button. Most now do, but some sources still will not display a button at all—only a generic link.

CAS ChemPort and EI Village 2 – Compendex are 2 sources that only allow links.

Deactivating SOURCES is generally not recommended, because it may cause unexpected problems, for example, with the OpenURL Generator (which is operated through the DEFAULT source). It is best to leave all SOURCES in their original active state. The active state of SOURCES to which the customer does not subscribe will not affect the contents of the SFX Menu or the KnowledgeBase, because OpenURLs from these sources will not be received by the SFX server.
For targets:
- Activated/configured in the SFX KnowledgeBase Manager
- Can be done manually or in bulk through SFX's DataLoader with a tab-delimited file.
- Easiest targets to configure are those to which you subscribe to everything (e.g., JSTOR Arts & Sciences I). Everything can be turned on at once.
- Most time consuming targets to configure can be those for which you only subscribe to a subset of titles available in the target. If done manually, this can take quite a bit of time. For example, ScienceDirect currently has 2,237 full text titles associated with it. To activate 500 or 1,000 of these by hand takes time—especially if coverage dates have to be adjusted.

At Georgia State, the initial work of activating and configuring targets took Judith and I almost a month. We bulk activated what we could, but there was still quite a bit of manual tweaking involved—especially with coverage dates. Less concentrated, but fairly steady clean up of targets continued for another couple of months, but with only one person doing the work.
Who does what?

- This depends on your organization. Before embarking on SFX implementation, make sure you have people in your library with the necessary skills.
Skills in using the following needed for technical support:

- Telnet
- Unix
- FTP using Unix commands
- HTML tagging in the Unix environment
- Command line editor in the Unix environment
- Graphic editing program such as Photoshop
- Perl

Skill in using the following tools is needed for technical support:

**Telnet** (the protocol for connecting to another computer or server), preferably a secure ‘flavor’ since you will be manipulating the production database.

**Unix** (the operating system SFX runs on)

- Ability to do **FTP using Unix commands** to pick up files.
- Knowledge of **HTML syntax** (and possibly Cascading Style Sheets) to customize the look-and-feel of SFX menus.
- Ability to do **HTML tagging in the Unix environment** using a command link editor to customize the SFX menu of services.
- Ability to use a **command line editor in the Unix environment** to change settings in the SFX configuration and other system files.

Graphic editing program for the Web, such as **Photoshop**, so you can add your logo to the menu or customize the SFX button.

Knowledge of **Perl**, the scripting language used by SFX, is useful so you can customize local thresholds, parsers, etc.

**WebAdmin and KBManager.** These are web based tools for setting up and maintaining the local instance. Training will be provided for its use.

**Setup.** The Knowledge base delivered with the SFX server contains preconfigured sources and targets. The biggest part of setting up is customizing the preconfigured resources to reflect local data about targets. To start, you need detailed information about your providers of e-journal subscriptions, coverage dates, and particularly ISSNs. Begin by activating those providers from whom you get everything. The SFX-supplied configuration for these will probably be what you need. Then activate the individual titles from providers from whom you get only selected titles. You may need to adjust coverage dates, called thresholds, for these. Activating sources is usually simply a matter of contacting the information service provider and asking them to add your SFX server address to your institution’s profile.

**Maintenance.** Ongoing maintenance is required to ensure that the links stay current. SFX delivers a monthly knowledge base update. During setup you will have flagged providers (from whom you get everything) to update and activate automatically when the update is loaded. For providers from whom you get only selected titles, you have to activate, deactivate or tweak the configurations individually.

Bill will describe this process in more detail in a few minutes.
Currently, if you want to use your SFX data to add an A-Z journal list, you will also need to get and manipulate files with:

- Access/SQL
- Excel
- Word
- Unicode
- The scripting language used by your web server

(This should change with SFX Version 3, which will have enhanced A-Z capabilities.)

Currently, if you want to use your SFX data to add an A-Z journal list to your web presence, you will also need to get and manipulate files with:

**Access/SQL** (Structured Query Language)
Excel
Word
Unicode
The scripting language used by your web server

This may change in the near future with version 3, when tools for creating the A-Z list will be added to KBManager.
If you want to add bibliographic records to your catalog:

- MARCit! service
- Your system’s batchloading utility
- A program like MarcEdit

If you want to add bibliographic records to your catalog, SFX has a MARC Enhancement Tool, in KBManager, called MARCit! which you can use to download a file of MaRC records for all your activated titles. You then use your system’s batch loading utility to import them into your catalog. Depending on which system you have, you may also need to use a program like MarcEdit to adapt the records before you import them.
Monthly Update package for SFX includes:

- **New Targets**
  - Latest update included the addition of the new ERIC as well as Mary Ann Liebert Publishing

- **New Sources**
  - Latest update added Computing Reviews

- **Updated Documentation**
  - Revised Target and Source Configuration Guides

- **Miscellaneous Changes**
  - Latest update adds option to disable the use of article title metadata in inbound links for targets using the Gale parser.

Ex Libris updates SFX each month. GALILEO currently applies the update to our shared and unique instances of SFX. Our systems staff currently regenerates our Electronic Journal Locator or A-Z list once a month after the update has been applied (or whenever this is necessary due to a large number of changes).

The update includes…

Mary Ann Liebert—Biotechnology titles. ACM Computing Reviews?

Gale parser change:

If you decide to enable this enhancement, any future inbound links to GaleGroup target databases will not use article-title metadata; these inbound links will use ISSN, Volume, Issue, Startpage and author last name. This amount of metadata is usually sufficient to reach the article.

5. The reason your library may wish to enable this option is if your librarians feel that there is an unacceptably high rate of failure creating successful links when the article title is included with inbound links to GaleGroup databases.

6. The results are dependent on what metadata is sent, and what articles in which databases your patrons are attempting to access. Some SFX sites have reported this problem frequently, while many other SFX sites have not reported problems at all. In general, we expect that there will be a drop in the number of unsuccessful searches, while there will be an increase in the number of less exact search-returns (the patron is presented with one or more articles to choose from at the Gale site).

7. If your patrons are not experiencing problems with SFX-to-Gale linking, we do not recommend enabling this option.

20
It’s important to look over the user documentation for the monthly update when it’s published.

Changes to parsers and parse parameters, for example, can cause problems for many users if they aren’t addressed as soon as the update is run.

Acquisitions currently reviews updates to check for any needed changes to be made (this can include activating new targets, etc.)
An example of this is a recent change to Project Muse

Very shortly after a monthly update was applied, a patron reported receiving an error when attempting to access an article from Project Muse.
When the clicked the Muse link on the SFX menu, they received a File not found error.
But the problem was easily fixed. With the monthly update, SFX had changed the Project Muse target service so that it offered a choice of 2 URL’s (one for the U.S. and one for Australia) to access Muse (host 1 and host 2).
Until either a 1 or 2 was entered in the Flag Value field (choosing access through either the U.S. or Australian site), the error message displayed when anyone attempted to access Project Muse through SFX. Adding the 1 solved the problem. We just hadn't gotten that far along through the update when the problem was reported. It's best to read through the update documentation as soon as it's available and to complete the update as quickly as possible.
- In addition to updated user documentation, Ex Libris also provides monthly Update Reports.

- Although they are available from a directory on the server of the local instance, these reports can also be run from the SFX Admin Center.
There are 3 types of reports: Changed Items, Deleted Items and New Items.
- These tab delimited reports may be opened in Excel

- The reports “contain detailed information about all new, changed and deleted items in the SFX instance. This information is specific to the SFX instance in which the update is run.”
This is an example of a cleaned-up update report opened in Excel. This is the changed object portfolios report from the August 2004 update. Although it contains more than 23,000 lines, not all of these changes need to be looked at. As an example, the last update in the list shows SFX changing the threshold of a title to indicate coverage ends after 2003.
Ex Libris recommended procedure for processing the “changed” and “deleted” reports:

- Check those items not changed or deleted automatically. Manually change or delete if necessary.

- In general, SFX does not automatically changed or deleted items that have been customized locally.

a) The option 'Automatic Update', overrides the rule in the standard update procedure, which states that database items with local changes will not be updated or deleted automatically.

b) The option 'Automatic Update + Automatic Activation will override the following two rules in the standard update:

database items with local changes will not be updated or deleted automatically.
database items like TARGETS, TARGET_SERVICES and PORTFOLIOs which are added to the SFX database during the KBupdate are added as inactive (to prevent these new items from showing up in the SFX menu screen).

These two options are often used for aggregator databases such as EbscoHOST or Gale, because the library typically subscribes to the whole database (i.e. all journals within a particular TARGET_SERVICE), and the content in these databases changes frequently. The two options reduce the amount of manual work necessary after the KBupdate (activation of new database items and manual updates of locally changed database items).
Out of 4,700 plus deleted object portfolios in this monthly update, only 4 were not deleted automatically—so following the Ex Libris recommended procedure, only a few items would need to be checked.
- In practice, you probably do need to look at more than just those items that weren’t automatically changed or deleted.
- For example, Georgia State currently looks (at least briefly) at any change made to an object portfolio that isn’t associated with an aggregator. In particular, we look at titles that are subscribed to individually (like titles in ScienceDirect, Wiley InterScience, etc.).
In the following update, SFX removed the volume and number information for the start date of coverage to the journal Urology in ScienceDirect.
Since this information is helpful and normally included, we added a local threshold to the object portfolio to replace this information.
In this next example, the monthly SFX update coverage for the Journal of Addictive Diseases moved from a starting date of 1993 back to 1991.
But some investigation showed that our coverage really does only go back to 1993.
In this case, a local threshold was added to change starting date of coverage back to 1993. SFX was also notified so they can look at and fix the global threshold.
Update changed threshold to add an end date of 2002 to coverage...

This next example is a case where the SFX monthly update added an ending date for coverage to the title Computers & Chemistry. It’s especially important to take a look at changes like this—they often are the result of a title change.
These are often updated due to a title change

Here we see ScienceDirect indicate that Computers & Chemistry changed title after volume 26, issue 5 to Computational Biology and Chemistry
Currently, when we learn of title changes, we add a note to the object portfolio in SFX...
The note displays on the SFX menu to aid patrons interested in additional volumes. Depending on local procedures for handling title changes, libraries may obviously need to do some additional checking/editing to title lists, etc.
This next example is a problem one of our Reference librarians submitted to the HelpDesk about a title missing from our e-journal locator. In this case, the title Biochemica et Biophysica Acta. General Subjects actually was there, but was listed only as General Subjects.
These are the OCLC records for the title.
Ex Libris was actually only using the information from the 245. We needed to request that they add the clarifying information found in the varying form of the title.
This is another example of a HelpDesk problem submitted by a reference librarian over a title missing from our E-Journal Locator. In this case, the person who submitted the problem quickly noted that the title they were looking for (The Times of London) actually was there, but listed as only The Times.
Here’s how that title displays on the SFX menu.
Again, in this case SFX is only using the information found in the 245. It actually might be more helpful for patrons for them to display the uniform title.
In this example, the monthly update added a closing date to the title Biomedical Applications in ScienceDirect. Again, this type of update will often be in response to a title change.
In researching this title, we noted that searching this ISSN in the citation linker without a date attached brings up two ScienceDirect choices—one for this title and one for the title that continues it.
Two menu choices display because SFX has associated these ISSN’s with each other.

This was because SFX had included a related ISSN field to the object portfolio that connects the record for this title and the subsequent title. When someone looks for one, the menu includes both choices. This can be confusing and is inconsistent with how SFX generally handles title changes.

To make this even more confusing, SFX includes another related ISSN on this record. This 3rd ScienceDirect choice does not display on the menu because this object isn’t active in our instance (it’s actually a dead end—this ISSN isn’t used in ScienceDirect and an active menu link for this ISSN takes the user to an error screen).
Just as an example, this is how SFX generally handles title changes for ScienceDirect. Here we have the title Clinical Immunology and Immunopathology.

It is continued after December 1998 by the title Clinical Immunology.
Here’s the ScienceDirect page for Clinical Immunology; it notes that it was formerly known as Clinical Immunology and Immunopathology.
We’ve added a note to the earlier title to indicate that it’s continued as Clinical Immunology.
Here--unlike the previous example--a patron looking for this title in our E Journal Locator will only be presented with one ScienceDirect choice.
And the same is true for the title that continues it…
Problem: Georgia State access to the Journal of Protein Chemistry no longer reflected in SFX.

In this example, a problem was submitted to the HelpDesk because our E Journal Locator was no longer displaying access to the Journal of Protein Chemistry even though we knew we subscribed to it.
Monthly update shows SFX record for Journal of Protein Chemistry being deleted.

Here, the SFX monthly update actually deleted the record for the Journal of Protein Chemistry.
We can see from the OCLC record that Journal of Protein Chemistry changed title to Protein Journal after 2003 and that the online issues for the Journal of Protein Chemistry are actually listed under Protein Journal.
And at kluwer online, we can see that both titles are listed under Protein Journal and the ISSN for Protein Journal. This is the opposite of how title changes were handled with ScienceDirect, which had separate records and ISSN’s for each title. Until the monthly update deleted the earlier title’s record, however, SFX treated them separately.
Monthly update shows record for *The Histochemical Journal* being deleted from SFX.

This last example, is the same thing. The SFX monthly update deletes the record for *The Histochemical Journal*. 
The Histochemical Journal changed its title to Journal of Molecular History after v.34, no.11/12 in 2002.
kluwer online places issues for The Histochemical Journal back to 1997 with issues of Journal of Molecular Histology, under that title’s ISSN.

As before kluwer online places issues for both titles under the Journal of Molecular History and the ISSN for Journal of Molecular History. Since it deleted the earlier title’s record, SFX is now handling the titles in the same way.
Cautionary Tales

Begin Judith’s section.
All of the idiosyncrasies to which print serials have traditionally been subject are still present in the online environment and then some.

Single Entry Lists
- Author/Title Proper vs. Uniform Title vs. Title Proper: SFX uses Title Proper in all cases
- "Minor" title change treated as a variant in the MARC record will not be selected for the A to Z list at all, even though the provider now considers it the current title.

ISSN issues and ‘Aggregator neutral records’

Consistent recognition and handling of entry changes

Back files and supplements to back files

---

All of the idiosyncrasies to which print serials have traditionally been subject are still present in the online environment. They still split, merge, change title, change publisher, change frequency and so forth. They also now migrate from one format to the other and change from serial to integrating resource.

**Single entry lists:**

**Main entry.** People have been trying to declare the concept of ‘main entry’ a dead issue for decades, but once again it raises its ugly head now that we all want A to Z journal lists on our web sites. In compiling a single-entry list, it is necessary to choose an appropriate entry. SFX uses the CONSER database as their source of ‘authoritative records’ but the programming isn’t sophisticated enough to choose the most useful form of entry. SFX considers the CONSER Title Proper to be the authoritative title in all cases, but there are frequently occasions when an Author/Title or Uniform Title would be the better choice.

**Minor changes.** This is further exacerbated by AACR2 rule 21.2.C2.b.ix which says that “the addition, deletion, or rearrangement anywhere in the title of words that indicate the type of resource such as “magazine,” “journal,” or “newsletter” or their equivalent in other languages” is a minor change. This can be especially inconvenient if one of these words is added or dropped at the beginning of the title. In any case, if the current title of a journal, according to the provider, is being treated as a variant in the Marc record, it will not be selected by SFX for the A to Z list at all.

**ISSN issues and ‘Aggregator neutral records’:**

ISSNs are commonly viewed as unique identifiers, and they have frequently been co-opted for use as a hook to connect citations to records of holdings. They were not designed for this purpose, and have a number of problems, not least that the ISSN network considers them to be carrier specific while the rest of the world wants them to be linked to content regardless of carrier. But, for lack of anything better, service providers are using them this way. Although they are not article specific, they are now being used as a building block in constructing links to full text online.

Most SFX objects use the ISSN for the print version of a title. There are separate fields for the eISSN and related ISSNs, but not all objects have them. Only if they are present can the link resolver cope with a search using these other ISSNs. Because the objects use the print ISSN, the MARCit! product provides an “enhanced” version of the Marc record for the print, in order to preserve SFX linking services that are based on the print ISSN. It seems ironic to me that after all the agonizing over single vs. separate records, and the wonderful compromise resulting in the creation of aggregator-neutral records, the only aggregator neutral records you are likely to receive from SFX are for electronic-only titles.

**Consistent recognition and handling of entry changes:** You can see from Bill’s examples that SFX is somewhat at the mercy of the providers when it comes to consistent recognition and handling of entry changes, which vary, depending on whether the provider uses a latest entry approach on their web site or not, and on how they set up the various titles. We are able, as Bill demonstrated, to make adjustments in the knowledge base to compensate for these problems in linking, but we haven’t worked out what to do about the records in the catalog.

**Back files and supplements to back files:**

A question was recently posted to the SFX list about how Elsevier back files would be handled but no definitive answer has been posted. The question is whether those who purchased one back file but choose not to purchase supplements will have to go from an activate-all situation to a selected-titles-in target situation.
- MARCit! is **still being developed**. We have found that there is quite a bit of manipulation needed before the MARC records can be loaded into the catalog.

**MARCit! is still being developed.** We have found that there is quite a bit of manipulation needed before the Marc records can be loaded into the catalog.

Additions, changes and deletions are delivered in one file. The records are coded new, change or delete in the fixed field. Other systems may have better batch loading utilities that can handle this file as delivered, but Voyager can’t. The file needs to be split because the bulk import process is different for each category.

Ex Libris uses the 090 tag for their Object ID/SFX control no. For libraries that still use 090 for locally assigned LC call nos. this needs to be moved to another local tag. We move it to 024 and use it for matching when updates or deletes are distributed.

We also find and delete 050s with pseudo call nos. such as

- **NOT IN LC**
- **ISSN RECORD**
- **IN PROCESS**
- **Microfilm**
- **CURRENT ISSUES ONLY**
- **LAW**
- **WMLC [accession no.]**

Each batch of records delivered contains a few that lack the 245 field.

Change 856 ‡u to insert our proxy address and ISSN, and change 856 ‡y to 856 ‡z (because Endeavor hasn’t implemented the use of ‡y for link text).
SFX is still a young product. It delivers a very exciting and useful service as it is, but there is still room for improvement. We look forward to the opportunity to influence how it addresses the various problems.
Questions?

Judith Shelton, libjms@langate.gsu.edu

William Walsh, wwalsh@gsu.edu