



# Smart Libraries™

Formerly Library Systems Newsletter™

50 East Huron Street, Chicago, Illinois 60611-2795, USA

February 2003 Volume XXIII Number 2

## IP authentication problems lead to alternatives

In December 2002, JSTOR reported its discovery of a serious breach of its Internet protocol (IP)-based security. Last fall an unauthorized person gained access to the JSTOR archive and attempted to download tens of thousands of journal articles. An investigation showed this crime was a deliberate attack that exploited the existence of open proxies—proxy servers that don't adequately block unauthorized users.

When libraries license commercial electronic resources such as e-journals, access is normally restricted by the contract terms to members of a particular community, such as a university or a public library district. To enforce such restrictions, the library

usually gives the resource provider a list of network (IP) addresses from which users should be allowed access to the service. The resource provider uses IP authentication to block access from computers that are not on the list.

This authentication works best when all legitimate access is likely to come from a definable set of network addresses, such as the public workstations in a library or all computers on a campus network. But libraries often want their patrons to have access wherever they are, whether at home or on the road.

See **Authentication** on page 2

## Bibliographic description debate renewed

Roy Tennant of the California Digital Library has breathed new life into the long debate over the MARC (machine-readable cataloging) establishment and the library's professional attachment to an increasingly arcane system of bibliographic description (*Library Journal*, Oct. 15, 2002 and Nov. 15, 2002).

Several new emerging standards and protocols also are responsible for enlivening this debate. What is less sure is just how libraries—and their vendors—are prepared to address these trends.

### XML vs. MARC

Most catalogers cannot help but laugh or sigh at the growth of XML (Extensible Mark-up Language) in library literature, conferences, and library products and services. Unlike its predecessor, SGML (Standard Generalized Markup Language), XML's relative simplicity and Internet pervasiveness destined it for high impact on an information economy.

After years of being industry talk, XML has arrived. But will it replace

See **Debate** on page 4

## IN THIS ISSUE

NISO Guide Describes RFP Standards  
PAGE 2

The Ties that Bind  
PAGE 2

University Library Weds Portals with Courses  
PAGE 3

Forum Surveys e-book Preferences  
PAGE 4

Ebsco Acquires RoweCom Europe  
PAGE 5

Open-source ILS Releases Version 1.0  
PAGE 6

Ohio Library Selects Koha  
PAGE 6

ILS Veteran Steps Down  
PAGE 6

Sirsi Announces SIP2 Certification  
PAGE 6

ZING Releases Two Specifications  
PAGE 7

New Services Link Databases to Holdings  
PAGE 7

ISSN 1541-8820



## NISO GUIDE DESCRIBES RFP STANDARDS

A new 70-page guide helps librarians avoid vague “the system should support standard x”-type statements in requests for proposals (RFP) in favor of specific requirements that encourage concrete responses from vendors.

In December, NISO Press published *The RFP Writer’s Guide to Standards for Library Systems*, which explains the many standards that can be referenced in RFPs for integrated library systems (ILS). The guide also helps assess vendor responses and provides a general overview of technical standards for building bibliographic systems.

The guide identifies the most important standards in 14 topical areas such as bibliographic formats, character sets, circulation protocols, and Web access. For each standard, its application to libraries is described, sample RFP language is provided, and compliance assessment issues are discussed. A free download of the guide is available from the NISO website.—PC

Contact: [National Information Standards Organization](http://www.niso.org)  
[www.niso.org](http://www.niso.org)



## The TIES that BIND

SF-Systems, creator of the LINCPlus library binding preparation system, has partnered with Ex Libris to incorporate the bindery software into its Aleph 500 integrated library system. The integration allows users to directly query any title record in Aleph by entering the barcode in LINCPlus. All title information is automatically inserted into LINCPlus, reducing the amount of user input to almost none.

Long ignored by major integrated library system vendors, this simple operation of integrating serials and monographic binding with third-party systems is long-awaited in many libraries operating stand-alone binding systems.—AKP

Contact: [Ex Libris-USA](http://www.exlibris-usa.com)  
[www.exlibris-usa.com](http://www.exlibris-usa.com)

### Authentication *from page 1*

To expand access points, institutions use proxy servers. These computers relay messages from one networked computer to another, replacing the address of the originating computer with their own. Typically, a user logs on to the proxy server with a userid and password and is authenticated against a campus directory or library patron file. The user then accesses the licensed resource through the proxy server, whose IP address is known to the resource provider.

### How hackers steal data

Some websites cater to hackers, providing lists of known open proxies and instructions on how to exploit them to gain illicit access to restricted resources. In the case of the JSTOR attack, the attacker had downloaded a list of open proxy addresses and systematically tested each to see if any had access to the JSTOR database.

The JSTOR type of attack can occur on any resource provider that uses IP-based authentication. Preventing the existence of open proxy servers is difficult; any Web server on a network can function as an open proxy if it is not properly configured.

As a short-term precaution, JSTOR recommends that subscribing institutions more strictly control access. Instead of authorizing all addresses on the campus network, a library can authorize a limited number of closely monitored proxy servers through which all traffic to the resource provider will pass.

In the long term, however, libraries and vendors will find developing alternatives to IP-based authentication more advantageous.

### Digital certificates are costly

Digital certificates, also known as PKI (public key infrastructure) certificates, are widely used in government, health, and business applications for user authentication on the Internet. Libraries can use them to control access to licensed commercial resources. A digital certificate is an encrypted electronic document issued by an agency called a certificate authority that verifies the identity of the certificate owner. The JSTOR system has used institutionally issued certifi-

cates by checking that the signer of a presented certificate is an institution subscribing to JSTOR.

Certificates, however, are not widely seen as a solution to the authentication problem for licensed commercial resources. Problems include the expense of issuing certificates, the varied requirements of resource owners, and a desire to protect the identity of individual users.

## Shibboleth answers ID needs

The Shibboleth Project, an Internet 2 initiative, is developing an authentication architecture geared toward restricted educational resources and providing open-source software to support it. (According to *Judges* in the Bible, the word *shibboleth* was used to distinguish the Ephraimites, who could not pronounce *sh*, from the Gileadites.)

In Shibboleth, responsibility for authentication is divided between the origin site (where the user is) and the target site (where the resource is). The origin authenticates the browser user and maintains information about the user's attributes; the target maintains the access control policy.

In most cases the identity of the user does not need to be known to the target. Only the attributes need be known, such as that the user is a registered student. In some cases the origin can actually calculate whether the user is authorized to access a resource under the license and simply return a *yes* or *no* to the target.

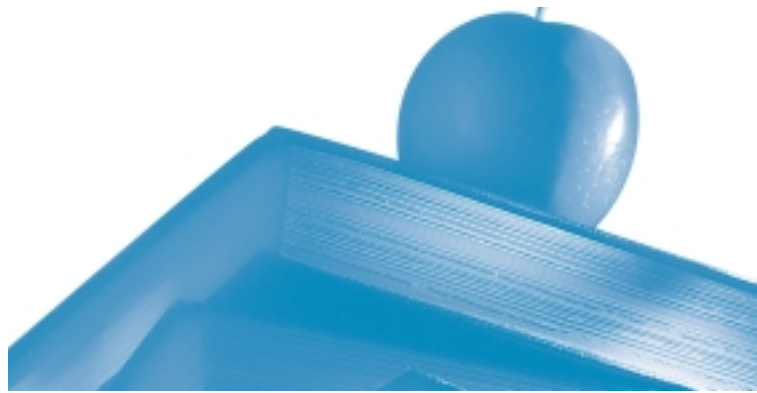
Shibboleth can be used for cross-institutional authentication as well as authentication between educational institutions and their commercial vendors. Several campuses are piloting implementations of origin software, including Ohio State, Penn State, University of California, and University of Michigan. Interested vendors include online resource hosts such as Ebsco and OCLC, makers of course management systems such as WebCT, and vendors of linking systems such as Ex Libris' SFX.—*Priscilla Caplan*

Contact: JSTOR

[www.jstor.org/resources/openproxies.html](http://www.jstor.org/resources/openproxies.html)

Internet 2

<http://shibboleth.internet2.edu>



## University library weds portals with courses

Academic libraries constantly struggle to establish a meaningful relationship with academic course offerings, usually through course reserve systems. Similarly, library portal systems have attempted to offer faculty and students resources specific to their particular discipline or subject interest.

The University of Rochester, River Campus Libraries has wedded the two with its new CoURse Resources System created by its Digital Initiatives Unit. Part of the library's new database-driven website (created with Cold Fusion), the CoURse Resources System presents course and instructor information, along with links to course reserve materials entered in the Voyager integrated library system (ILS).

Course or faculty searches also present the name, e-mail, and phone number of librarians for that particular subject. Like library subject portals, the system also directs users to related websites, articles, and books. The database presents a photograph of the library subject specialist teaching the course, as well as class meeting times and a link to the course homepage if one exists.

Tying disciplinary library resources to specific courses and instructors also eases the burden placed on subject specialists to populate library portal systems with resources. Moreover, the course system could eliminate the redundant efforts of reference subject specialists working directly with faculty to create library subject guides for other parts of a library website. A fully integrated system of library resources and course management, the River Campus Library's system is an excellent model for academic libraries willing to both manage courses and maintain disciplinary portal systems.—*AKP*

Contact: University of Rochester, River Campus Libraries

[www.lib.rochester.edu/index.cfm?page=10](http://www.lib.rochester.edu/index.cfm?page=10)

Debate *from page 1*

MARC as a sensible, usable, and easily adoptable means of bibliographic description? Many experts think so.

A couple of years ago, thinking of MARC and XML as complements to each other was more productive than thinking of them as competitors. XMLMARC and other crosswalk conversion programs have been around for some time, but usually are employed to output bibliographic data for other uses, rather than as a method for converting databases to XML instead of MARC.

MARC deserves attention, if not a complete redress, with the assumption that this conversion trend may soon take on more meaning (that is, the ultimate conversion away from MARC) and with growing acceptance of XML as the Internet's preferred descriptive language.

### The standard-bearers

The Library of Congress (LC)—not exactly the owner of MARC, but certainly the leader in establishing the changes to AACR (Anglo-American Cataloging Rules) and the MARC standard itself—is developing an XML schema that will port MARC records to XML for many purposes, including use in library applications.

The Metadata Object Description Schema (MODS), an XML schema, will carry selected data from existing MARC 21 records

and enable the creation of original resource description records. A subset of MARC fields comprises the schema, which uses language-based tags rather than numeric ones (such as MARC). In some cases, MODS regroups elements from the MARC 21 bibliographic format.

Specifically created for encoding descriptive, structural, and administrative metadata for digital library objects, MODS is an extension to METS, the Metadata Encoding & Transmission Standard. METS, an initiative of the Digital Library Federation, is (like MODS) maintained in the Network Development and MARC Standards Office of LC. So although LC and the majority of catalogers rally around MARC, at best even MARC's biggest proponents appear to be hedging their bets. Nevertheless, MODS elements still inherit the semantics of MARC, which could result in continued incompatibility with other metadata schemes while maintaining compatibility with more traditional library data.

Added to this alphabet soup of new standards is work on the international front. To give bibliographic description an international flavor and to correct the hierarchical shortcomings of MARC, the Functional Requirements for Bibliographic Records (FRBR) standard also has emerged.

Sponsored by the International Federation of Library Associations and Institutions (IFLA), FRBR sets out to restructure catalog databases to reflect the conceptual structure of information resources. Unlike the single flat record concept underlying the MARC record, FRBR includes four descriptive



## FORUM SURVEYS E-BOOK PREFERENCES

The Open eBook Forum (OeBF) surveyed consumer preferences for printed and electronic books at the literary festival, New York is Book Country, in October 2002. The survey, titled *Consumer Survey on Electronic Books*, measures attitudes toward e-books by people who read books in print.

Event participants voluntarily completed 263 surveys. Survey group demographics were similar to U.S. demographics in age, gender, and income level, but the survey group had a significantly higher level of education. The findings include:

- 31% of respondents had read an e-book on their computer, and 15% had purchased an e-book.
- People who had read articles on e-books were more likely to have read or purchased an e-book than people who had not.
- No significant correlation exists between level of computer skills and reading or buying e-books.
- People who used the Internet daily or who read magazine or news articles daily on their computers were not more likely to have read or bought an e-book.



levels: work, expression, manifestation, and item. Though the original FRBR report came out almost four years ago—before full acceptance of XML—and is based on research that predates the prominence of the Web, FRBR does a lot to inform the metadata needs of any future system that might replace MARC.

From some perspectives, FRBR presents more challenge to AACR and Western cataloging practice than it does to the MARC record. If a library or integrated library system vendor (ILS) decided to switch from AACR/MARC to FRBR, that switch would represent the biggest change to cataloging practice since the advent of the MARC record.

## Vendor reaction

This discussion is academic without knowledge of how ILS vendors, publishers, and online content distributors are reacting to these changes. Despite rapid development of digital library software, multisearch modules, and other new areas for ILS companies, XML remains mostly a conceptual framework to which new products might adapt.

MARC is still the accepted storage, communication, and presentation standard for libraries' holdings. Until that standard changes, ILS vendors are unlikely to build new functionality into legacy library modules. Some vendors are, however, already adept at XML output and crosswalking from MARC records. Development in that area remains more theoretical than practical. There isn't any readily available use for such features.

- Most people agreed e-books should be priced the same as paperback books; there was no agreement on pricing e-books the same as hardbacks, DVDs, or music CDs.
- 62% of respondents agreed they would want to read an e-book at a library.

These results challenge the industry assumption that links computer skills and Internet use with e-book usage. The results also indicate the majority of the reading public may be willing to read e-books at libraries, which adds to a growing corpus of findings on e-book acceptance by library patrons. A summary of the survey is available at the OeBF website. The full survey is available only to OeBF members.—PC

Contact: [Open Book Forum](http://Open Book Forum)  
[www.openebook.org](http://www.openebook.org)

VTLS is the only ILS vendor that complies with the new FRBR hierarchical relationship model. OCLC has begun extensive research into the new format. Other vendors are beginning a slow embrace of XML for content markup, or of dumbed-down versions of bibliographic description such as ONIX, and of XML DTD, developed by the book industry. But with MARC still firmly in place as the dominant communication format, all other standards must at least be portable to the accepted standard.

## Is MARC dying?

MARC's days in the sun are over. As issues of Web presentation and data communication outside the world of libraries continue to dominate library development, a more flexible, extensible, and up-to-date standard must emerge to take the place of the aging MARC format. Such a transition must only take place, however, with the benefit of 40 years of complex development. Issues such as circulation interchange protocols are often discussed outside the context of bibliographic formats or the changes presented by FRBR.

Bona fide objections over moving to a descriptive format favor the English language required in the current XML schemas relating to MARC (as opposed to MARC's language-neutral, numeric markup). Forward-thinking digital libraries embracing new ways to describe traditional library materials is not enough; ILS vendors also must be serious for any format to have widespread impact on the library world.—Andrew K. Pace

## Ebsco acquires RoweCom Europe

To extend its already major presence in Europe, Ebsco Industries, Inc. confirmed signing a letter of intent to acquire the European operations of RoweCom, Inc., a subsidiary of divine, inc. The transaction is not final and requires French regulatory approval. divine, inc., simultaneously announced its intent to divest itself of its content subscription business.

For libraries, which continue to watch subscription rates, this content owner merger represents yet another in a long list of such information conglomerates.—AKP

Contact: [Ebsco Information Services](http://Ebsco Information Services)  
[www.ebsco.com](http://www.ebsco.com)

## Avanti

## Open-source ILS releases Version 1.0

One year in the making and four years after its inception, Version 1.0 of the Avanti MicroLCS open-source library computing system became available in mid-January. Avanti MicroLCS, a compact, self-contained client-server library system, is written in Java.

Version 1.0 includes basic cataloging and OPAC functionality but lacks a completed circulation system. Its documentation and access to all system-created data is via a well-written application programming interface (API).  
—AKP

Contact: [Avanti](http://www.avantilibrarysystems.com)  
[www.avantilibrarysystems.com](http://www.avantilibrarysystems.com)

## Koha

## Ohio library selects Koha

Koha, another open-source library system, announced its Version 1.3 in October 2002. The announcement was preceded in August 2002 by the announcement of its first planned U.S. implementation at Nelsonville Public Library in Nelsonville, Ohio. Long considered an integrated library system solution for only tiny library operations, its growing adoption by larger (yet still relatively small) libraries is a significant change in the library system marketplace.—AKP

Contact: [Koha](http://www.koha.org)  
[www.koha.org](http://www.koha.org)

## ILS veteran steps down

In a surprising announcement on Dec. 26, 2002, Endeavor Information Systems announced Jane Burke would be stepping down from her senior post as president and CEO. An industry veteran, Burke joined Endeavor in 1995 for the launch of its Voyager system.

Though Endeavor quickly established itself as the darling of the academic library world, it has recently lost ground due to several factors, including growing competition from newcomer Ex Libris, Innovative Interfaces's new Millennium system, and the acquisition of a large customer base in the Sirsi buyout of DRA.

A company that prided itself on employee ownership, Endeavor has been criticized since March 2000, when it became a wholly owned subsidiary of giant Dutch publishing group, Elsevier. Roland Dietz, managing director of global sales and marketing for the science and technology division of Elsevier, succeeds Burke as Endeavor president and CEO.—AKP

Contact: [Endeavor Information Systems, Inc.](http://www.endinfosys.com)  
[www.endinfosys.com](http://www.endinfosys.com)



## Sirsi announces SIP2 certification

Sirsi launched its certification program for the standard interchange protocol, Version 2 (SIP2), Dec. 9, 2002. The program enables third-party vendors to ensure their SIP2-compliant products integrate seamlessly with Sirsi systems.

The SIP2 protocol is used by many circulation-based library products, such as self-checkout machines. Sirsi plans to begin evaluation and certification in the first quarter of 2003.

The certification program will improve product development and ultimately create efficiencies for library staff and their third-party product selections. This standard certification program is the only one of its kind offered by an ILS vendor.

Sirsi also is active with the NCIP (Circulation Interchange Protocol Z39.83) standard, the presumed successor to SIP and SIP2. Sirsi's own Mark Needleman, product manager, standards, serves as the NISO Standards Development Committee liaison for NCIP. The NCIP standard governs lending and borrowing transactions among library systems.—AKP

Contact: [Sirsi Corp.](http://www.sirsi.com)  
[www.sirsi.com](http://www.sirsi.com)

# ZING releases two specifications

Although Z39.50 concepts remain powerful and relevant, the technologies used in the protocol date from the 1980s and are no longer attractive in today's Web-based environment. To make the protocol more appealing, the Z39.50 International Next Generation (ZING), an initiative of the Z39.50 maintenance agency, released two specifications in November 2002 for nine months of trial use:

- Search/Retrieve Web service protocol (SRW) allows a limited set of standard Z39.50 functions to be carried out using common Web communications protocols and XML schema.
- Common Query Language (CQL) is a formal standard for expressing queries to information retrieval systems in a simple, human-readable format.

Many major information providers have refused to install Z39.50 servers, which has led to alternative approaches to cross-system search. These techniques, however, require reformulating searches into the native syntax of each online service and may require screen scraping to capture results. ZING developers hope more information providers will adopt Z39.50 if the query language, record formats, and communications protocols used are updated to Web-friendly standards.—PC

**Contact:** Library of Congress International Standard Maintenance Agency  
<http://lcweb.loc.gov/z3950/agency>

## NEW SERVICES LINK DATABASES TO HOLDINGS

Patrons can link from an entry in an abstracting or indexing database to a list of all the library's holdings of the cited journal using a free service called Journal Linker, initiated in December 2002 by Serials Solutions. From the link, patrons can list journal issues or other information, depending on the vendor. The service optionally includes print and microform holdings.

Library patrons can use Journal Linker with many content providers, including ProQuest, Ebsco, Gale/Infotrac, OCLC, Cambridge Scientific Abstracts, Ovid, and Silver Platter. The service is hosted and maintained at Serials Solutions as part of the company's A-to-Z title lists and full MARC record services. Development and maintenance costs were offset by a Jan. 1 price increase on A-to-Z title lists.

A second product, Article Linker, is a full-featured OpenURL resolver scheduled for release in early 2003.

Article Linker will accept OpenURLs and provide direct links to articles as well as extended services such as inter-library loan. Like Journal Linker, Article Linker will take advantage of Serials Solutions' existing database of journals available at its customer institutions. Both the linking server and the data will be maintained at Serials Solutions, so customer sites will not have to install hardware or software, or maintain knowledgebases all of their journal holdings.

Although Journal Linker is a no-cost, low-effort way to direct patrons to library holdings, Article Linker will compete with services, such as Openly Informatics' 1Cate, and linking products, such as Ex Libris' SFX.—PC

**Contact:** SerialsSolutions  
[www.serialssolutions.com](http://www.serialssolutions.com)



**ALA TechSource**  
www.techsource.ala.org

Smart Libraries Newsletter  
American Library Association  
50 East Huron Street  
Chicago, IL 60611-2795 USA

---

NON PROFIT  
US POSTAGE  
PAID  
PERMIT 1479  
ROCHESTER, NY

---

## February 2003 IP security breached through open proxies

### Smart Libraries Newsletter An innovative overview of library automation

Smart Libraries Newsletter delivers hard data and innovative insights about the world of library technology, every month.

**Publisher**  
Miriam Wuensch  
800-545-2433, ext. 3252  
312-280-3252  
mwuensch@ala.org

**Editor**  
Chris Santilli  
630-495-9863  
chris@wordcrafting.com

**Contributing Editors**  
Priscilla Caplan  
352-392-9020, ext. 324  
pcaplan@ufl.edu

Judy Luther  
610-645-7546  
judy.luther@informedstrategies.com

Andrew K. Pace  
919-515-3087  
apace@unity.ncsu.edu

**Administrative Assistant**  
Judy Foley  
800-545-2433, ext. 4272  
312-280-4272  
jfoley@ala.org

Production and design by Angela Hanshaw,  
American Library Association Production  
Services.

Smart Libraries Newsletter is published monthly  
by ALA TechSource, a unit of the publishing  
division of the American Library Association.  
Copyright American Library Association 2003.  
All rights reserved.

**TO SUBSCRIBE**

To reserve your subscription, contact the Customer Service Center at **800-545-2433**,  
press 5 for assistance, or visit [www.techsource.ala.org](http://www.techsource.ala.org).  
The 2003 subscription price? Just \$79 US.