

Moving Z39.50 to the Web

OCLC researcher develops open-source software to blend an important library standard into the WWW infrastructure and help solve Web searching problems

By Tom Storey

As a result of digital technology, the library community increasingly overlaps with other online communities, such as museums and other cultural organizations, and the Web development community itself. These communities have substantial intellectual resources, but often lack extensive experience with discovery metadata and information retrieval from text-based databases. Libraries have been working on these issues for decades, and one tool they have developed is Z39.50, a client/server standard for searching local and distributed databases.

Ralph LeVan, an OCLC Research Scientist who has been involved with Z39.50 since 1983, is helping create the next generation of Z39.50 software. Called Search and Retrieve Web/Search and Retrieve URLs (SRW/U), the new Z39.50 is a Web service that uses several Z39.50 features and combines them with simple, robust and modern Web technology. Mr. LeVan has written an open-source version of the software, which libraries and other organizations can download at no charge from the OCLC Web site.

“The SRW/U initiative is part of an international collaborative effort to develop a standard, Web-based, text-searching interface,” says Mr. LeVan. “It is ‘plumbing’ that uses different mechanisms for transmitting requests and presenting data, and it draws heavily on the abstract models and functionality of Z39.50, yet removes much of the complexity.”

Mr. LeVan says that the Z39.50 standard solved most of the problems associated with text searching of large, distributed text databases—different interfaces, indexing, record formats and sizes. Now that expertise needs to be brought to the Web where there are few widely supported standards for searching databases. This makes searching less efficient for the system provider, the content provider and ultimately the end-user.

SRW/U brings the Z39.50 experience to the Web community using tools they are comfortable with—Web Service Description Language, Simple Object Access Protocol and Extensible Markup Language, says Mr. LeVan. It provides a basis for the Web community to solve some of their searching problems, and the potential for Web applications to access Z39.50 resources.

“Developing SRW interfaces to data repositories is significantly easier than for Z39.50,” he says.

One example of where SRW could be a solution for the Web community is in metasearching. Metasearch engines are Web sites that search a number of search engines and database systems and combine the results, such as dogpile.com and highway61.com. Mr. LeVan says that developers for these sites would love to have a standard that they could expect database providers to implement.

Initially developed before the advent of the Web, Z39.50 is a client/server-based protocol that makes databases easier to use by establishing uniform standards so different systems can communicate with one another in a way that is transparent to users. It is one of the premier examples of library cooperation and a standard that large parts of the library and archive world rely on to exchange records and share information.

However, Mr. LeVan says that for the newer generation of developers, Z39.50 is far too big and unwieldy to read, understand and implement. Z39.50 uses protocols for communications, search language, information retrieval and data structure that are different from those used with the Web.



OCLC Research Scientist Ralph LeVan is bringing the Z39.50 experience to the Web with SRW/U.

It also is not very popular due to its complexity: 11 native services, seven extended services and 150 pages of documentation, for example.

By contrast, SRW/U is one service—search and retrieve—that uses common Web development tools.

Mr. LeVan notes that a search and retrieve Web service is a component called for in most digital repository architectures. “The distinctive advantage of SRW over classic Z39.50 is that the Web community is already committed to services running over HTTP,” he says. “This is a great opportunity to leverage library experience with text searching in local and distributed databases, and to provide increased interoperability with other communities on the Web.”

To learn more about SRW/U, please see: <http://www.oclc.org/research/projects/webservices/default.htm>

From OCLC Research Labs

FRBR Bookmarklets: A way to weave your library's collection into Web bookstores

OCLC Research is integrating a technology called “bookmarklets” with its FRBR algorithm to create an experimental service that lets Web surfers instantly search a local library catalog while browsing an online bookseller. FRBR Bookmarklets expands on work originally done by Jon Udell at InfoWorld and makes a library's holdings more visible and accessible on the Internet.

Here's how it works.

Written in JavaScript code, FRBR Bookmarklets captures the ISBN displayed in the URL on the bookseller's page or from the page itself, combines it with associated ISBNs for the same work, and sends the search to the library's online public access catalog. If the library owns the book, or any associated edition, a record or list of records appears on the screen complete with shelf status information.

In addition to Web booksellers, the bookmarklets also work from other sources of bibliographic information containing ISBNs, such as the OCLC FirstSearch service.

FRBR Bookmarklets uses WorldCat as its source for ISBNs. The FRBR Algorithm, which OCLC Research developed to convert library catalogs to the FRBR model, was used to build tables of ISBNs for all intellectual works represented in WorldCat. FRBR Bookmarklets queries the tables and receives a string of all ISBNs for the work, and then sends the string in a Boolean OR search to the library catalog.

To date, there are 20 online public access catalogs profiled for the experimental service. To activate any one of them as a FRBR Bookmarklet, users need to drag the OPAC bookmarklet to their browser's links bar. Then, when viewing detailed book information at a Web bookseller or another source, users click on the bookmarklet in their links bar to see if their local library holds it.

Chief Scientist Thom Hickey, along with Software Architects Jeffrey A. Young and Jennifer Toves and Consulting Product Support Specialist Eric Childress, developed the experimental service and believe they can enhance it using library holdings information in WorldCat. “We want to take advantage of the fact that WorldCat knows which edition of a work a library holds and place that edition at the top of the results set,” says Mr. Young.

New libraries can be added to the profile list by filling out a request form on the FRBR Bookmarklets Web site.

To learn more about FRBR Bookmarklets, visit: www.oclc.org/research/researchworks/bookmarklets/