10-1-2008

Science Information Without Borders.

Carol Tenopir
University of Tennessee - Knoxville

ONLINE DATABASES
Science Info Without Borders

By Carol Tenopir

ONLINE RESOURCES ARE CRITICAL to the concept of "libraries without borders," the theme of the 2008 World Library and Information Congress at the 74th International Federation of Library Associations and Institutions (IFLA) general conference, held in August in Québec City.

Fee-based online information sources accessed through libraries are an important part of the library's role in bringing high-quality science resources to their patrons, but so are the many scientific products available free of charge on the web.

Unfortunately, it isn't always easier to locate relevant content in free materials than to find it in fee-based tools. Even data within free and open databases gets passed over by web crawlers. This is because databases deliver structured results only in response to queries. Search engine crawlers miss such records and typically index only the homepage or entry portals.

World science data

One solution is WorldWideScience.org, as described to the IFLA Science and Technology Libraries Division audience by Walter Warnick, director of the U.S. Department of Energy's Office of Scientific and Technical Information (OSTI). This search gateway provides free federated search access to science databases from government agencies worldwide.

Many countries already make their government agency–created science information publicly available on the web. However, to find this you first must know it exists, then make multiple searches at local portals.

The federated search architecture used by WorldWideScience.org allows access with a single query. Upon execution, the system contacts the partner sites all over the world, runs the search on each database in real time, and returns the results to the U.S. server.

The system then ranks the returned search results by relevancy and allows the searcher to select which documents to display.

The results screen also includes a link to the Wikipedia entry on the search term to give an overview of the topic, which especially helps laypeople or scientists searching outside their primary area of expertise. Finally, WorldWideScience.org also presents results divided into clusters, so searches can be refined by topic or date.

> Such high-quality information provides "a link between the scholarly world and the real world"

Science.gov the model

Founded in 2007 through an agreement between the British Library and OSTI, WorldWideScience.org is modeled after Science.gov, a portal to 100 million pages of science bulletins from 13 U.S. government agencies. Both are built using Deep Web Technologies, a Santa Fe, NM–based company that calls its product Explorit Research Accelerator.

WorldWideScience.org includes government-sponsored science content from more than 50 member countries and 40 international portals, as well as everything covered by Science.gov. With the addition of China as a member in August, the portal, Warnick says, "will soon reach a billion pages."

For now, all content is in English, although Warnick says translation services are being considered. The WorldWideScience Alliance governing board was officially chartered in June 2008 to make decisions on behalf of the partners regarding future developments and sustainability.

Coverage varies quite a bit by country and region. The main source from Africa is African Journals Online, a collection of 320 journals covering medical, agricultural, and other science topics either by African authors or with an African focus. Brazil offers Scientific Electronic Library Online (SciELO), which includes 211 Brazilian STM journals.

Material available from the United States and Canada is quite extensive, including technical reports, books, conference proceedings, and journals. The Indian Institute of Science provides access to electronic theses and dissertations, though some countries, such as New Zealand, offer (as of August 2008) only historical or limited data.

There is still much to be added by member countries, not to mention by countries that have yet to join the alliance. (Italy, Russia, and Singapore, for example, are not members.) This resource will only get more valuable as the amount of information increases.

The library's role

Attendees at IFLA were reminded of how such resources filter down to users. Don McMillan, librarian at the University of Calgary, described how he works with biology professors at his institution to incorporate free sources such as U.S. patents, PubMed, and bioinformatics data into biology classes at all undergraduate levels. Such high-quality content, he said, "provides a link between the scholarly world and the real world."

Large libraries have a role as contributors and system designers, but even small libraries can participate by incorporating these sources into library instruction, adding links in their catalogs, and finding other ways to alert their patrons to these important tools. This is truly an example of "libraries without borders."

Carol Tenopir (ctenopir@utk.edu) is Professor at the School of Information Sciences, University of Tennessee, Knoxville