XML and MARC: The New Digital Divide

Dick Miller, Systems Librarian and Head of Technical Services at Stanford's Lane Medical Library, sees XML (eXtensible Markup Language) as the "lingua franca of the Information Age." In a recent article, he offers an excellent tutorial on the origins and uses of XML for representing information content, optimized for Web delivery.

According to Miller, libraries are creating a new "digital divide" from the Web's users since the vast majority of records for library resources are currently stored in the proprietary formats of integrated library system (ILS) vendors and in the "complex and arcane" MARC formats which even Systems Librarians are "hard-pressed" to extract. At Lane Medical, catalog records have now been converted from ILS and MARC formats into XML formats using locally created "XMLMARC" conversion software. The new XML records are ready for integration with other Web resources. Lane is offering its conversion programs free on the Web at xnlmarc.stanford.edu. BiblioML, which converts Unimarc to XML, has been released by a French government agency and is available at www.culture.fr/BiblioML. Miller notes that ILS vendors, with Endeavor taking a leadership role, are beginning to incorporate XML into their systems but feels that a fully XML-based ILS is still three to five years away.

Excerpted from "XML: Libraries' Strategic Opportunity," in Netconnect, Summer 2000. Full text of this article is available online at www.jidigital.com/xml.asp.

The Leaders Speak: Report from the Oxford 2000 Collection Development Retreat

Over 75 key publishers, librarians, and suppliers, from the U.S., U.K., Europe, Scandinavia, the Middle East, and Australia gathered at Keble College, Oxford University, from July 20-22, 2000, to debate and discuss the "new world order in collection development." The meeting was the 2nd in the Fiesole Retreat Series begun in 1999 in Fiesole, Italy, dedicated to the memory of Mario Casalini of Casalini Libri, and sponsored by The Charleston Company and Against the Grain.

Discussions were lively, as participants reacted to 15 papers presented by their colleagues as part of three panels. Major themes included end users and what they want from us, the impact of digitization and automation on publishers, librarians, and suppliers, and a review of creative emerging business models. Opening keynote speeches were given by Reg Carr, Director of Oxford University Library Services and Chairman of the Board of the Research Libraries Group, and Herman Pabbruwe, CEO, Walters Kluwer International Health & Science; the wrap-up was offered by Anthony Watkinson, a widely respected publishing consultant.

Among the conclusions emerging from the conference were:

- The format of information has become a service issue. Navigational layers will link to content; content will point to objects; and objects will point to other information resources.
- Links are the key enabling technology and most highly valued by users. The journal article can now serve as a portal, providing an access point to relevant data in a wide variety of other formats and locations.
- The journal will become a managed database—"print, mail, and forget" will not work any longer.
- Peer review is the process of validation and can be separated from publication, as BioMed Central is trying to prove with "open peer review."
- The library, as evidenced by mission statements studied by one speaker, has embraced a "user-centered" strategy and now is actively seeking to understand user needs (as are all parties).
- The "serials problem" is not a shortage of money but a need to get more for the money available. OhioLink's e-journal licensing program proves that demand exists for titles libraries did not acquire through traditional selection.
- Selection of e-journals and e-books takes on a new definition in the electronic environment: the library's job moves from choosing individual titles to acquiring rights to general collections and profiles.
- As the library's role in selection moves toward licensing, the supplier takes on many of the tasks previously handled by the library, thus facilitating the entire process.
- The new mission in collection development for librarians is to expand access to as much reasonable material as possible, thus allowing the user to be the true selector, choosing on an as-needed basis from wide storehouses of licensed information.

Full conference proceedings from both the 1999 and 2000 retreats are being published by Casalini Libri; many papers are already available at www.digital.casalini.it and printed editions are expected to be available later this year. The Charleston Report will offer complete coverage of selected papers from the retreat in future issues.
Libraries on the Move

The Seattle Public Library has raised more than $43 million, against a goal of $60 million, toward a major program that will renovate or replace all 22 branch libraries and build three new branches and a new central library. Monies raised will also be used for other capital improvements, including new technology, and for library materials and staff training. The target for reaching the total funding goal is 2003. Library Hotline, June 5, 2000.

The University of Virginia's Institute for Advanced Technology in the Humanities has received a $1 million grant from the Mellon Foundation for a three-year project to evaluate scholarly research based on digitized primary resources. The project will be conducted jointly with the Library at UVA, which will use both existing and new materials collections. ARL Bimonthly Report on Research Library Issues and Actions, June 2000.

The State of Colorado has made $2 million available to libraries for the purchase of education resources, including books and periodicals. Computers and Internet access are excluded from the grants, which will be awarded in amounts of $3,000 minimum. Libraries applying for the grants must implement one of three “filtering” approaches to the Internet. The State Library of Colorado has been established as the program administrator. Press Release, May 26, 2000.

Job Hunting on the Internet

Here are the top six Internet job sites for May 2000, according to a recent Nielsen/Net Rating (number of unique visitors/time spent online follows):

1. monster.com (1,357,009 visitors/19:08 minutes on average)
2. careerpath.com (641,004 / 13:04)
3. careermosaic.com (438,003 / 8:00)
4. headhunter.net (339,002 / 38:02)
5. employment.classifieds.yahoo.com (329,002 / 11:33)
6. usajobs.opm.gov (316,002 / 18:36)

Not surprisingly, computers/software resumes top monster.com’s database of resumes. Sales and engineering top the list of total number of job searches. USA Today, June 23, 2000.

Charleston Advisor Vendor Showcase Filling Fast . . .

The Charleston Advisor Vendor Showcase on November 1, 2000, in Charleston, S.C., offers you the chance to show your latest electronic products to a key group of over 500 library selectors and buyers. If you have not yet received your exhibitor’s sign-up kit, be sure to contact us right away at 303-282-9706 or email us at rlenzini@charlestonco.com. Don’t delay—space is limited.

Recent Rulings Appear to Favor Writers

A series of court rulings over the last few months seem to indicate growing judicial support for writers and other copyright holders. In September 1999, a federal appeals court overturned the 1997 federal court finding in Tasini vs. New York Times that allowed publishers to include freelance articles in their online and CD-ROM products (see TCR, v.4, no.2 for more background). In July of this year, a settlement in the Ryan vs. CARL UnCover document delivery lawsuit was announced. MP3 earlier had also announced a settlement with Warner Music and BMG in its copyright infringement suit. And, of course, much press has been devoted to the July 26, 2000, ruling against Napster.

In the case of Tasini, the Second Circuit Court of Appeals ruled that the reuse of freelance work on databases and CD-ROMs without the authors' express permission constitutes copyright infringement. The original ruling favored the New York Times and was based on the argument that electronic products were simply another “version” of a work already published.

While no ruling was made by the court in Ryan vs. CARL, which was heard in the Federal District Court for Northern California, the $7.25 million settlement requires UnCover to expand its permissions system to grant royalties to authors as well as publishers. The class-action copyright suit was brought by a group of freelance authors who claimed that UnCover sold copies of articles to which they, not the magazine publishers, held the rights.

Music shareware service Napster is still alive thanks to two judges from the 9th Circuit Court of Appeals who issued a temporary stay of federal Judge Marilyn Hall Patel's injunction that would have shut down the site. Napster now claims more than 22 million users and is said to be seeking a settlement, which would allow it to continue to operate.

Most in the industry see plenty of parallels between music and published information and are monitoring all these developments closely. The general conclusion appears to be that the new models of distribution are just what the consuming public wants, and we in the industry are charged with creating new systems to allow easy copyright compliance.

Ed Note: Don’t miss the chance to see Jonathan Tasini, President of the National Writer’s Union, in action at this year’s Charleston Preconference on “Copyright: Who Owns What in Cyberspace,” set for November 1, 2000, from 1:00 – 4:00 p.m. at the Lightsey Center, College of Charleston. For registration information, logon to www.cofc.edu/library/conference.
ALA Hot Topics

The American Library Association’s Annual Meeting took place in Chicago from July 6-12, drawing healthy attendance of 24,427, including 5,338 exhibitors. Libraries by and large are enjoying a period of stable funding, while many of the themes that were hot at previous conferences remain in the spotlight. Popular topics included:

- **E-books.** There is a scramble to create “new business models” that will affect access and pricing. Sessions devoted to this area were standing room only, as libraries continue to wrestle with the logistics of purchasing and distributing this new medium. Of note, the netLibrary model based on “check-out/check-in” is unpopular. Libraries want to license e-book collections for open access by their users.

- **Article linking.** In the realm of e-journals, sessions featuring CrossRef, JAKE, and SFX all drew packed houses. Executive Director Ed Pentz calls CrossRef’s service the “phone book” of the electronic journal world. JAKE’s creators reside at the Yale Medical Library, where they have created a grass-roots title-level index to aggregated journal collections. SFX, developed at Ghent University, is now being marketed by ExLibris under a licensing agreement. Libraries will expect e-journal services as well as indexing and abstracting services to take advantage of all three of these important efforts.

- **Harry Potter and the Goblet of Fire.** The fourth in this very popular series was released on Saturday July 8th, the first morning of the show. The record-breaking advance sales gave evidence that the print book is still very much alive. Also of note, the audiobook on both cassette and CD was published simultaneously.

- **The virtual library.** This concept is expanding to encompass not only the collection but also services and programming.

- **DVDs.** This format is rapidly overtaking VHS as the medium for videos.

On the whole, librarians anticipate there will be breakthrough announcements on products coming soon, but most say they didn’t see many of them at this conference. On a positive note for publishers and suppliers, a number of public libraries are expanding with new buildings and will be purchasing corresponding opening day collections.

(Ed Note: For more information about CrossRef, be sure to read The Charleston Advisor’s interview with Ed Pentz in our July 2000 issue. Also watch for our “In the Field” profile of JAKE in the upcoming October 2000 issue.)

By the Numbers

$1,632 . . . is the price the median ARL library will pay for a journal subscription in 2020, assuming that current trends continue. ($107 is the price for a monograph in the same period.) *ARL Bimonthly Report*, June 2000.

2 million . . . is the number of hits to Gnutella’s site the morning following Judge Patel’s initial ruling against Napster. *Wall Street Journal*, July 28, 2000.

216 books . . . for every 1,000 service area residents must be purchased annually according to a new UK standard for the country’s public libraries. The standard is set to begin in 2004. *Library Journal*, June 15, 2000.

400% . . . is the increase in moms on the Internet, from 4.5 million in 1997 to 16.4 million in 1999. Kids are the primary reason for the increase. Roughly 2/3 of families in the U.S. now have home computers; 46% of families are hooked to the Internet. *St. Petersburg Times*, June 12, 2000.

$34,000 . . . is the cost of the technology package in Memphis/Shelby County Public Library’s new “InfoBUS,” the high tech replacement for the traditional bookmobile. *Netconnect*, Summer 2000.

24 months . . . is the target for profitability for Barnes & Noble.com, according to Steve Riggio, vice chairman. For the year, B&N.com expects sales to reach $335-350 million, compared to $202 million in 1999. *Wall Street Journal*, August 1, 2000.

Marketing Tips: Knowledgeable Staff at the Booth Needed!

At least two important consultants who talked to *TCR* noted that staff at the exhibit booths at ALA could not answer important questions about their products. In fact, some booth staffers admitted that they were really just “actors” who know how to talk into headsets and give scripted demos. This pattern is a huge turn-off for potential buyers who’ve made a big effort to come to the conference and are looking for decision-level information.

How People Search the Web

In a recent interview, Chris Sherman, Internet search expert and the creator of the About.com Web search site [http://websearch.about.com](http://websearch.about.com), observed that “People don’t use my site to learn how to be expert searchers. They don’t want to learn tricks or techniques to help them retrieve information more skillfully. They’re really using the Web to solve a problem or satisfy a need.”

The most popular destination on About.com is the section devoted to MP3 music files, followed by articles and links on how to find people using the Web, images, maps and directions, and health information. Sherman says “Searching the Web is still too hard, even for people with relatively well-developed skills.”

If he had the opportunity to build the perfect search tool, Sherman says that, like Apple, he’d “think different,” and would avoid defining searching as an information retrieval problem. Instead he would look at problem-solving behaviors and techniques used by experts in various fields and would design these approaches into the search engine.

From an interview in *Searcher: the Magazine for Database Professionals*, July/August 2000.

How About It?

Since 1995, Dolly Parton has been an active supporter of children’s reading programs through her nonprofit Dolly Foundation and her own “Imagination Library” project. This year, she is offering children from birth until age five a free book every month by mail. “Imagination Library” gives a 60-volume library and bookcase to preschool children in order to encourage reading at an early age. To date, more than 100,000 books have been distributed by the program. *Library Hotline*, May 29, 2000.
Short Takes

Frequently Asked Questions (FAQ’s) about e-books and technical processing are now available on the Algonquin (Illinois) Area Public Library’s homepage: www.nlsilsius.org/ alhome. Look under Library Services. AAPL has also established an e-book discussion forum. To subscribe, send a message to ebook-talk-subscribe@egroups.com. For more information, contact neeljune@yahoo.com. Library Hotline, June 5, 2000.

More books in Spanish are needed, according to librarians at this year’s Book Expo in Chicago. Also requested: more books for African-Americans and help to sort through all the information now being published in e-formats. Publishers Weekly, June 12, 2000.

Roughly 40% of e-book borrowers expected to read 50 or fewer pages, according to a study of online book usage conducted at Columbia University over a three-year period. Most students intended to take notes and photocopy or read just a small part of the book. The final report of the study is available at www.columbia.edu/cu/libraries/digital/ obdocs/finalreport.html.

Charleston Advisor Guide to Electronic Products

The Charleston Advisor is pleased to announce that it will be introducing a new Web-based “Guide to Electronic Products,” targeted to information professionals who need to locate and contact publishers and vendors of these important products. The Guide will be free to users on the Web and will include complete product descriptions, contact information, hot links and more. For more information on the Guide, be sure to read the flyer enclosed with this issue of TCR.

A Less Appreciative View of Amazon.com

Ed Note: The following comments were sent to us in reaction to our article in the previous issue of TCR, v.4, no.6, entitled “Small Publishers Appreciate Amazon.” Read on for another point of view.

We are delighted to sell books to Amazon.com—but only on a prepayment basis! Contrary to the experience of the small publisher quoted in the May/June Charleston Report, for whom Amazon.com “leveled the playing field,” we found the online bookseller to be extremely slow (as in 180 days or more) to pay. So we switched to a prepayment system, and Amazon.com’s purchase orders now include a credit card number. Recently their credit card was declined by the bank. We called and got a second card number—which was declined! We called again and got a third number—which was declined! Finally, on the fourth try, we had a card that worked.

Given this experience, I question the wisdom of sending books to Amazon.com without invoices, of expecting automatic payment at the beginning of every month, and of the previous author’s statement that he had “no need to deal with an accounts payable department.” Caveat seller!

100 Most Frequently Challenged Books of the Decade

ALA’s Office for Intellectual Freedom (OIF) has published its list of the 100 most frequently challenged books of the decade just ended. Topping the list: Scary Stories (series) by Alvin Schwartz, deemed “too scary” and “unsuited to age group,” followed by Daddy’s Roommate by Michael Willhoite, accused of “promoting homosexuality as a normal lifestyle.” The list was compiled from 5,716 challenges to library materials reported to or recorded by the OIF from 1990-1999. Reasons most often cited:

1. “Sexually explicit” (1,446 challenges)
2. “Offensive language” (1,262 challenges)
3. “Unsuited to age group” (1,167 challenges)
4. “Occult theme or promoting the occult or Satanism” (773 challenges)
5. “Violent” (630 challenges)
6. “Homosexual theme or promoting homosexuality” (497 challenges)
7. “Promoting a religious viewpoint” (397 challenges)
8. “Nudity” (297 challenges)
9. “Racism” (245 challenges)
10. “Sex Education” (217 challenges)
11. “Anti-family” (193 challenges)

Banned Books Week, which is sponsored by ALA and the American Booksellers Association, is set for September 23-30, 2000, and features the theme “Fish in the River of Knowledge: Celebrate Your Freedom to Read.” For more on BBW 2000, see www.al.org/bbooks/index.html. The entire list of the top 100 challenged books can be found at www.al.org/alaorg/oif/ top100bannedbooks.html. ALA News, June 2000.

ARL Launches Aggressive Campaign

The Association of Research Libraries (ARL), together with the Association of College and Research Libraries (ACRL) and SPARC (ARL’s alternative publishing affiliate), has begun a campaign entitled “Create Change” which intends to inspire academic campuses to alter radically the current publish-ormiser cycle and “take back” scholarly communication. “Create Change” aims to:

• educate faculty on the serials pricing issue by providing information, tools and skills;
• instruct faculty on the ways they can advocate for and undertake change in scholarly communications;
• nurture would-be faculty leaders in this arena;
• support library directors and staff who have been leading the effort to date.

The effort is being strongly promoted to academic faculty, administrators, and library directors during the course of the summer. To learn more, logon to www.createchange.org/


Coming in Future Issues

— E-journal Usage Statistics from OhioLink
— Consortial Buying and Its Impact on Library Suppliers
— Charleston Conference Preview