

4-1996

Library Systems: Keeping Up Our Images

Terry Ballard

Gill Library, College of New Rochelle, tballard@cnr.edu

Follow this and additional works at: <http://digitalcommons.cnr.edu/gill-publications>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Ballard, T. (1996). Library Systems: Keeping Up Our Images. *Information Today*, 13(4), 52-54.

This Article is brought to you for free and open access by the Gill Library at Digital Commons @ CNR. It has been accepted for inclusion in Gill Library Publications by an authorized administrator of Digital Commons @ CNR. For more information, please contact lfazzino@cnr.edu.



Library Systems: Keeping Up Our Images

Conversion from "fiber-based" to electronic format is not a pretty picture

I used to work at a university where the librarians put together monthly forums to keep abreast of current trends in information science. A few years ago, someone from the computer center came by to talk about the Internet. In particular, his message was that electronic information had become so pervasive that it would put libraries out of business within a few years. He then went to his com-

puter to log on for a demonstration and experienced a bit of technical difficulty. This difficulty extended for 45 minutes before the whole thing was called on account of rain.

That university library contains more than a million books, heavy in the humanities. The librarians do not fear for their jobs for now, because they know that the Internet will have to develop quite a bit

more depth before it can compete with libraries for anything but the most current information. That may be why librarians all seem to be from the state of Missouri when told about the impending demise of the book. "We've heard all of that before, back in the 1960s," they'd say.

True, electronic information sources are outstanding in delivering access to bibliographic material from the 1970s

forward. Not coincidentally, this is the time period when publishers started keeping their bibliographic data on computer files before printing them on "fiber-based information systems," also known as books. We all know that someday we will have a *Reader's Guide to Periodical Literature* online back to 1900. On the other hand, we are still waiting for retrospective information in most cases.

Studies in Scarlet

Addressing this challenge, the Research Libraries Group instituted a major project in 1995 to digitize a collection of 19th century documents in the field of family law and domestic relationships. This project, known as "Studies in Scarlet," will bring together the collection resources from seven member institutions. They are Harvard University Law School, Princeton University, University of Pennsylvania Law School, New York Public Library, North Carolina State Archives, University of Leeds, and my own New York University School of Law.

I was notified that we were selected to participate a week before I took the job as

It's just a matter of time

For almost half a century, we've been using our R&D staff and the latest technology to meet your serials processing needs. When you needed a fast, easy system for day-to-day serials functions, we developed REMO, the first PC-based serials control system. Then we networked REMO, another industry first.

To help you save time and money in your ordering and claiming, we built the online ROSS system. Then we developed electronic order forms. Electronic invoices. X.12 interfaces for all the major library systems. And, soon, we'll be bringing you BarPro, the bar code technology that will redefine serials control.

Real needs. Real solutions. All designed to make the time you spend managing serials shorter. Easier. And much more productive.

If you're looking for some innovative answers to your serials processing needs, call us today at 1 800-221-3306. Because at Readmore, meeting your needs is just a matter of time.

READMORE AUTOMATION

Meeting every need

22 Cortlandt Street • New York, NY 10007-3194 • 1 800-221-3306

“

The phrase '90 percent accurate' sounds good until you see a converted document that is 90 percent accurate.

”

automation coordinator at the school. From the first day, I have been involved in learning about the issues of scanning, OCR (Optical Character Recognition), marking up, and systematic data storage. I have sat through a number of demonstrations where a vendor will take a microfiche card, put it into a viewer that is tied to a computer, and—ZAP!—capture a perfect image into a TIFF file. It looks beautiful, and you can call it up any time you want to look at it after storing it on your hard disk. The problem is that the image is just a picture. Even though it has words and looks like text, it might as well be a picture of Staten Island. Furthermore, TIFF files cannot be read by Web browsers; you need special software.

To make these images into editable files you may need to run OCR software. Vendors do not like to show you this process because the results are not as pretty. The best of these programs, using the clearest of type, will still leave unresolved patches to clear up manually. The phrase

(continued on next page)

The Systems Librarian

continued from previous page

"90 percent accurate" sounds good until you see a converted document that is 90 percent accurate. For a project as vast as Studies in Scarlet, this kind of cleanup is impractical for doing in-house, so we have spent a lot of time talking to vendors.

Some of these people say that it is easier to skip OCR altogether and have people perform double rekeying from the TIFF image. That means that teams of typists will type every word in thousands of pages, two by two. A quality control person will look at the two sets of pages in a blink machine and correct any discrepancy that shows up. (We were trying to picture this happening in a factory in some developing country, with hundreds of typists in long rows typing words they may not be able to even comprehend. It brings new meaning to the concept "global village.") Anyway, now we have the pages converted into computer files that are 99.95 percent accurate, but we're not done with them yet.

All Words Are Not Created Equal

The eventual users of these documents will need to have meaningful access to them. Even though the files will be mounted in a database and searchable, all of the words are not equal. To be truly useful to researchers, these pages will need to be analyzed. Fields need to be assigned to distinguish authors, titles,

chapter headings, and so on. This marking-up process will be done in SGML (standard generalized markup language), which is better suited to textual research than is the more widely known HTML (hypertext markup language).

The Scarlet project is being administered by two member services officers of the Research Libraries Group, Win-Shin Chiang and Ricky Erway. "Since this is the first project of its kind, we all expect to learn what's involved in putting together a major imaging project, so nobody is expected to have all of the answers," said Chiang. In view of this, they put together a small listserv for all of the participants, and we were able to share information and try out all of the puns on the name

Scarlet (frankly, my dear . . . , scarlet fever, *Scarlet Letter*, etc.). At the end of three months, we all know a lot more about imaging and marking up than we did before. Now that we are almost ready to start sending pages to the vendors, we will have a spring meeting in New York for all of us to compare notes and to develop a master strategy.

Making Progress?

Imaging seems to be taking hold everywhere in the field. I just spoke with Cathy Hajo from the Sanger Archives here at NYU. The archive is part of the history department, and it is planning to image the personal papers of birth control

pioneer Margaret Sanger. NYU Law's contribution to the Scarlet project is related to the Comstock Laws, which outlawed sending information about birth control through the mails because it was considered obscene.

If these things seem a bit esoteric, bear in mind that the Comstock Laws were invoked in the portion of the recent Telecommunications Act that is trying to censor the Internet. The more things change, the more they stay the same.

Terry Ballard is automation coordinator at the New York University School of Law Library. He can be reached at ballardt@turing.law.nyu.edu.

THE EVOLVING VIRTUAL LIBRARY:

Visions and Case Studies

Edited by Laverna M. Saunders

A transformation is taking place in all types of libraries. Just a few years ago, libraries were occupied primarily with developing local automated systems and collections of print materials. Although these

functions are still being performed, external forces have created the need for change. Budget cutbacks and the escalating cost of materials, especially scholarly journals, have caused librarians to examine more carefully the collection needs of their users. The concept of access to remote resources which supplement or substitute for items owned locally has become more acceptable due to the availability of worldwide telecommunications networks, appropriate hardware and software, and machine-readable databases. The "virtual library" allows a user at a computer terminal to search bibliographic citations, databases, electronic publications, and other types of information in digital format. And as the technology changes and evolves, so does the virtual library.

This book addresses many of the practical issues involved in developing the virtual library. It provides background on the growth and development of electronic resources and also examines innovative pilot projects that are occurring in all different types of library environments. There is no doubt that a dramatic change is taking place. This book will help you deal with that change and position you and your organization for the future.

1996/211 pp/hardbound/ISBN: 1-57387-007-2/Price: \$39.50

Express Order Service: Call 1-800-300-9868 or 609-654-6266

Fax Order Service: Call 609-654-4309

Mail Orders: Complete & Mail Form Below to:



Information Today, Inc.

143 Old Marlton Pike, Medford NJ 08055

ORDER FORM

Please enter my order for _____ copy(ies) of *The Evolving Virtual Library* at \$39.50 each. Add \$3.95 shipping and handling for the first book and add \$3.25 for each additional book.

Name _____

Title _____

Organization _____

Department _____

Address _____

City/State/Zip _____

Daytime Phone _____

Method of Payment:

Bill me. PO# _____

Check enclosed.

Charge my: MasterCard Visa AMEX

Account # _____

Exp. Date _____

Signature _____

174/96

SIRS Mandarin Launches Version 2.71

SIRS Mandarin, a library technology vendor and educational publisher in Boca Raton, Florida, has announced the release of Mandarin Library Automation System, Version 2.71. This PC-based, fully integrated library automation system includes several software enhancements.

The new features include a sorting option in the Online Public Access Catalog that allows users to create a sorted bibliography in one easy step, eight new Mitinet custom templates to aid librarians in creating their own MARC records, and updated subject cross references that guide OPAC users to related search topics.

The Mandarin System is priced at \$3,500. Customers who maintain an annual Service & Update Agreement automatically receive the 2.71 update, along with 24-hour toll-free and online technical support. To receive a SIRS Mandarin Catalog and more information on Version 2.71, contact SIRS Customer Service at 800/232-SIRS, via e-mail at custserve@sirs.com, or visit the SIRS Web site at <http://www.sirs.com>.