Key Trends Affecting Libraries

July 2004

Acknowledgement

Amy Marino, a staff member of the Office of Policy and Analysis (OP&A) and Amy Nazarov, a staff member of the Smithsonian Center for Education and Museum Studies (SCEMS) explored key issues facing libraries today. The idea arose after OP&A's participation in a strategic planning session with the Smithsonian Institution Libraries' board. While working on this project, researchers were involved in a number of other projects; however, they managed to have frequent conversations with others in the field and to read many books and articles on emerging technologies. They also sent the draft report to several outside libraries to solicit comments. I express my appreciation to the authors and reviewers who provided valuable comments, as well as to Nancy Gwinn, Director, Smithsonian Institution Libraries, for her support and patience.

Carole M. P. Neves
Director
Office of Policy and Analysis

Introduction. The Office of Policy and Analysis here summarizes key trends affecting the management, collection and retrieval of information in libraries, including the 20 libraries in the Smithsonian Institution.

The study team recommends that SIL should consider these themes as the libraries seek to respond to the information-management demands of the 21st century. This report is the first summation of libraries' changing relationships with their patrons largely as responses to virtual technologies, greater information sharing, and enhanced public access. It draws upon a review of relevant literature, and upon conversations with several academic, specialized and public librarians. Each of the following themes could be explored by more in-depth study, particularly as they apply to the SIL's unique environment.

New ways to use libraries. In recent years, public, academic and specialized librarians are redefining their roles, given the vast amount of electronic data resources through which they must guide patrons—and to which, more ominously, some observers say patrons are turning in lieu of a trip to the library. University of Maryland librarian Timothy Hackman points out that with the rise of the Web, libraries have long ceased to be the only source of information. People have *direct* access to information—but with various levels of credibility. Hackman points out:

> *This means that we have to do more in the way of marketing and outreach—taking the library beyond its walls and directly to the user—in order to compete with other info providers. This is different from how we've ever operated and requires a different mindset.*

> *The competition, like "Google Answers," has adapted our traditional model to the new environment. Digital reference services or "chat reference," and more generally the "digital library" – without the boundaries of a physical location – provide all manner of media wherever the user happens to be.*

As librarians and their "customers" alike harness the Net as a tool at their disposal, W. Lee Hisle, librarian and vice president of information services at Connecticut College, points to the challenge at hand:

> *Librarians are aware that an appropriate institutional balance needs to be maintained between traditional library materials and services, and those services (which sometimes overlap) represented by instructional and information technology departments.*

> *Should libraries house campus information commons? Should libraries report through an "information czar," rather than through the traditional academic hierarchy? These are among the challenging questions for the profession and the academy. It is also important, though difficult, to maintain technology currency in the face of decreasing resources, rising*

*costs, and differing views about institutional funding priorities (**College and Research Libraries News**, Nov. 2002.)*

The balancing act Hisle alludes to is echoed in comments made by Hackman and other librarians the study team consulted, who are navigating the confluence of "traditional library materials and services and services represented by instructional and information technology departments."

Many libraries have either completed or are well along in the merging of information services departments with libraries. In fact, it's increasingly common for library directors to wear the mantle of information-services czar (or czarina). As the organizational structure of an institution changes to reflect the changing nature of information access, barriers are chipped away—not without struggle, it must be observed—between the service orientation of the library and the historically inward-looking world of information services. Yet when a library is successful in this collaboration, its continuing viability is assured and patrons are best served.

Consider this example, which may be relevant to SIL: a college librarian with whom OP&A consulted described a strengthened relationship between her staff and members of the college faculty, with whom the library is working on a grant-funded project to increase the level of technology integration in the classroom. One example of this partnership was that of library staff members working with professors to add live data links to online versions of class syllabi. Indices, related Web sites, and other resources for further study were included with the professor's assigned readings and other assignments. In addition, the grant stipulated that faculty members receive a stipend for efforts to incorporate "information literacy" principles into day-to-day coursework.

This partnership helped enhance students' access to a range of information sources. More importantly, perhaps, it resulted in both students and faculty favorably viewing the library as a responsive source of helpful service and relevant materials. Might SIL make a similar service available to SI curators, educators, researchers, scientists and others? The possibilities are intriguing.

According to a number of librarians, for all of the media coverage generated by Google and its like, the search engine has not replaced the librarian wholesale—far from it. In fact, while "a lot of the literature says that reference librarians are going away, [the numbers of patrons we serve] have been going up," said one.

Others concurred, noting that the questions they now receive tend to be far more narrow and precise than in the past, thus demanding librarians' awareness of the breadth of specialized information resources that are most suited to address the query at hand. Perhaps the refinements come from library customers having begun their information quest independently, by visiting a search engine.
In sum, librarians have used search engines and other Internet technologies for years to better meet patrons' needs. With the explosion of the World Wide Web as commercial,

information, social and cultural phenomenon, librarians have not seen their role as guides through information resources obviated, but altered.

The move from broad to focused searching. The Pew Internet and American Life Project reported in August 2003 that "the Net has quickly become the place to go to answer a question, no matter what the subject. Consequently, this activity has become one of the most popular on the Web, and more than 75 percent of users nationally have gone online at one time or another to look for an answer."

Moreover, the place where many millions of information seekers first go is Google, which according to *The New York Times* receives more than 200 million hits a day and accounts for 34.9 percent of all Internet searches. With the Web entrenched as a place to go to launch an information search, more content providers seek to set themselves apart by offering in-depth searching abilities divided by vertical industry, topic or other subject area. The ability to refer patrons to these resources can further increase libraries' value as a partner in navigating the Web.

Here are a few examples of the focused searching that not only is becoming more common in the academic, non-profit, federal and business worlds, but that offers models SIL might consider adopting.

> **SIL—and SI—as content providers**. SIL's Galaxy of Knowledge (http://www.sil.si.edu/) contains several public databases of texts and documents culled from the libraries' own collections; users can search science, art and design, business and industry and American history texts and documents. Likewise, the Electronic Biologia Centrali-Americana site gives viewers a marvelous opportunity to examine close up a rare scientific text and view its plates in stunning detail. However, the frequency with which these resources appear in search-engine results lists is unknown. (For better or worse, algorithms can be tweaked by programmers who want their company, organization or product to appear at the top of a hit list of search-engine results. On a related note, Nutch (http://www.nutch.org/docs/en/), an open-source search engine currently under development, is garnering interest for its impartiality; its designers claim that Nutch will order users' query results with no influence from the proprietary ranking that most search engines offer companies for a fee.)
>
> Yahoo, the Web portal and search-engine developer, hopes to engage the Smithsonian as a content-provider partner in a new venture. This spring, Yahoo severed its long-time partnership with Google in the hopes of making its own online search technology a viable competitor. To that end, Yahoo officials told Smithsonian employees in a recent presentation that the company is partnering with content providers such as the Library of Congress, the New York Public Library and National Public Radio in what Yahoo calls its Content Acquisition Program, or CAP. Yahoo claims that

CAP will bring the rich resources of what Yahoo terms the "hidden Web" to a much larger audience of information-seekers. Partners provide links to their data holdings to Yahoo, which makes them available online in a premium search feature available to paying clients. Content providers like the Smithsonian, however, would pay nothing to participate in the service.

Yahoo officials claim that CAP data will be vetted for quality—mostly by the content providers themselves, but also by Yahoo's own editorial team—and will result in libraries, museums, and other entities' data being made available to a much larger swath of users. An application such as this has the potential to enable SIL's collections to be far more broadly known. The concept is simple, but whether it will require the integration of a complex set of factors including leadership willing to embrace change and resources to implement and evaluate the project remains to be seen.

Data aggregation. Many universities, libraries, archives and comparable institutions are devising ways to pool their data resources. The term "federated search" has come into broad use as a means of accessing data from numerous digital and physical resources from one accessible site.

In addition to SIL's efforts mentioned above, there are hundreds of examples of libraries and similar institutions seeking to aggregate their data in this manner. Perhaps the best established is the Online Computer Library Center's WorldCat (http://www.oclc.org/worldcat/), which gives users access to the library catalogs of over 9,000 worldwide members. In fact, the OCLC WorldCat concluded in June a year-long pilot program to make library resources available from non-library Web sites; over 12,000 public, school and academic libraries are participating.

About 125 academic institutions use DSpace (http://www.dspace.org/), archiving software written by the Massachusetts Institute of Technology and Hewlett-Packard, to provide digital storage for academic papers, technical reports and other work by university professors. Librarians OP&A interviewed cited DSpace as one of the models of academic data aggregation, in fact. The DSpace concept, known as "superarchiving," refers to the creation of a searchable digital depository that is available free to libraries and other institutions.

Another initiative with an academic bent is the Scholars Portal (http://www.arl.org/access/scholarsportal/), which seven universities and the Association of Research Libraries joined forces to develop. This "academic Google" would perform a search across many different digital databases, including many of the subscriptions that the library currently has. Project participants include Dartmouth, the University of Arizona, the University of Southern California and four others.

Other projects that attempt to aggregate information resources of different libraries and archives include ArchivesHub (http://www.archiveshub.ac.uk/), which gives users information about thousands of archives in the United Kingdom; Bibilotheca Universalis (http://www.bl.uk/gabriel/bibliotheca-universalis/en/bibliotheca_universalis_collections.htm), an international effort spearheaded by the G-7 nations to digitize library holdings around the world and make them accessible online; SearchLight (http://searchlight.cdlib.org/cgi-bin/searchlight), which lists books and other resources available throughout the University of California campus system; and Google's partnership with MIT, 16 other worldwide universities around the world and OCLC to tap into scholarly papers as a Web resource, possibly to be developed as a component of Google's advanced search page. There are many others.

Searching with greater precision. Public search engines like Teoma (www.teoma.com) and Amazon's emerging (albeit commercial) A9 project (http://a9.com) are breaking new ground by attempting to not only deliver "hit lists" to users but to automatically generate possible refinements to a user' search, lists of related Web links, lists of past searches and other value-added features. Other focused search engines include Aesop (www.aesop.com); JoeAnt (www.joeant.com), which is managed by volunteers; and SurfWax (www.surfwax.com).

Vertical search engines—which respond to users' queries by hunting through hundreds of thousands of documents related to a given field—offer users more control over their searching, and may be the most attractive model for SIL to emulate. A basic example is Google's "UncleSam" interface which automatically limits a search to government Web sites and documents. Another more sophisticated example is Scirus (www.scirus.com) – a site for scientific information resources only. The U.S. Department of Justice's Virtual Library makes available to its employees information pertaining to cases and other legal and judicial matters. Dartmouth College offers an "engineers' portal" that performs a similar function for the engineering and science concentrations. Many other academic and business entities have designed custom search mechanisms for their particular topics.

Filtering down from business. Libraries in general, and SIL specifically, would be wise to watch carefully as the business world—which in recent years has been sifting through a multitude of nascent data-management technologies like XML and data-integration tools—begins to place its bets.

A common language. The established and widespread adoption in the business world of Extensible Markup Language, or XML, is paving the way for a more consistent way to manage and integrate data. Documents,

records and other information created using the XML format may be stored anywhere on a computer network but retrieved and processed at any workstation tied into the network, since it was created in a common language.

Industry's acceptance of XML addresses a fundamental problem in information management. Over many years, banks, credit-card companies, insurance firms and other corporations have aggregated data about clients and products in numerous "silos" of information that are maintained independently of one another; they may run on different computer operating systems, be administrated by different network managers and be developed in different languages. Since these systems were not designed to "talk" to one another, customers and employees using them may never gain a complete picture of the collective services an institution offers its customers.

In the same vein, SIL's collections are also housed in disparate physical and virtual rooms, buildings and networks. The seeker of data housed in SIL is rarely exposed to the comprehensive richness of the collections.

Worth noting: the OpenURL standard, developed for use in academic environments, seeks to use information tags assigned to documents to deliver the correct version of an item to the end user. Its development may assist in the creation of open, searchable databases of information, which could eventually enhance SIL patrons' use of resources.

Information storage. Data storage is another key issue facing business, and one that will affect libraries as they convert more holdings, such as decades' worth of back issues of periodicals, to a searchable online format.

The biotechnology industry, for example, reportedly collects so much information in its research and testing processes that the amount of raw data doubles every six to 12 months. While biotech firms are bound by federal regulations for data collection and maintenance, there are parallels to the information-management challenges libraries face in the coming years and decades.

Consider the evolution from serials' microfilm and microfiche editions to the possibility of storing them in digital formats. The transition from microfilm-to-digital alone represents the obsolescence of old equipment and the need to purchase new, scalable hardware and software with sufficient capacity to hold yesterday's back issues and tomorrow's forthcoming ones.

As data-storage evolves further, future moves to still more advanced technology could necessitate other capital outlays and increase the demand for technological expertise. What's more, libraries face a unique conundrum as they "go digital:" what becomes of the publications themselves when they are converted into digital format? Should they be discarded or stored for posterity? Where will the resources to store decades' worth of periodicals come from? What will guide budgetary and acquisition decisions for print versus electronic investments? Will electronic format publications be accessible in the future? These and other questions must be explored as libraries move toward a greater digital presence.

Also on the horizon: some technology vendors are developing strategies for what they call "Information Life-Cycle Management," or ILM. The term refers to organizations' plans for creation, management, distribution, archival and disposal of data. This approach, too, could be a useful one for SIL to explore further.

Handling different data formats. Aside from the theoretical and philosophical issues surrounding the storage of library holdings, it's necessary to examine the demands of storing pictures in books and magazines, as well as text.

SIL's holdings are rich with photographs, lithographs, sketches, prints and other imagery related to the creation and expansion of a collection of printed matter documenting American art, history, science and culture. There are opportunities for SIL and others to expand library/information research into non-traditional areas such as image retrieval and video indexing. Images such as these call for additional storage space in the form of databases that can recognize and handle data in both structured and unstructured formats, as well as scalable hardware capable of running an enterprise-level database.

To attempt to address these issues, companies like Avaki, BEA Systems and IBM have developed products in a category known as "data-integration tools." These offerings let users simultaneously query a group of disparate data sources and retrieve all related structured (such as listings found in a card catalog) and unstructured (such as text files) data as though it were stored in a single place. **Network World** magazine reports that annual sales of data-integration software, tallied now at about $1 billion, could reach $1.5 billion by 2007, according to International Data Corp., a research and consulting firm.

Portals. Just as the corporate world has adopted the use of Web portals to offer customers a unified online place where they can get information or complete transactions, it may be useful to consider this model for libraries.

Utility companies are but one vertical industry that has adopted the portal approach, in part to offload customer-service representative duties onto the customers themselves, who are given the tools to check their account balances or pay bills on the Web, and in part to move closer to the much touted concept of the "paperless office." Companies like Computer Associates sell portal software that permits some of the data-aggregation capabilities offered by the data-integration tools described above; the product is used by financial-services companies such as J.P. Morgan/American Century. The means to digitally renew books or receive automated notification of the arrival of a book of interest are just two of the features such portals would make available to SIL patrons; the ability to conduct a virtual search of all collections is perhaps an even richer capability such portals would allow.

The data aggregation trend described above is exemplified in the library portal created by UNESCO (http://portal.unesco.org/ci/ev.php?URL_ID=6513&URL_DO=DO_TOPI C&URL_SECTION=201&reload=1085154325). As libraries, companies or other organizations invest additional resources into the development of portals, they can consider adding workflow and information-sharing features that have the potential to significantly increase employee productivity. Organizations are also opting to mass discrete Web pages behind a single portal, to give users a consolidated point at which to start their data-retrieval quest. Again, this model for information retrieval, storage, and sharing has broad application for different kinds of organizations, including libraries.

Timothy Hackman shared an example from the University of Maryland – an interface called "Research Port" enables users to search multiple proprietary databases, along with the Libraries' online catalog, from a single interface. "Find It" links within databases (via SFX) connect item records to full text wherever it appears (full text database, electronic journal, free-access web site), to the correct page in the UM Catalog, or to the Interlibrary Loan form.

Wireless. Another key factor for SIL to consider is the boom in wireless services. Both telecommuters and on-site workers in manufacturing, retail, health-care and other industries use handheld devices to send, retrieve, store and view relevant data.

As these cutting-edge users advance in their adoption of wireless devices, it's reasonable to assume that library users and staff may also find relevant applications for their data-management, search and retrieval tasks.

Professional development groups such as the American Association of Law Libraries are advising members to pay attention to the spreading ubiquity of wireless access points. Wireless connections are already enhancing the museum visitor's experience. Hand-held devices serve as both tour guide and omniscient curator, if you will; using a wireless device, a visitor can call up more information about the collections on demand and take it home with them. A visitor may also learn about the library's holdings on the subject. In this way, the wireless device serves as a starting point for the user's adventure and search for information.

Coffee shops, universities, businesses, trains, airports, ferries and libraries large and small, are installing wireless access points to meet consumer demands. In 2004, the UK Museums, Libraries and Archives Council invested £60,000 in 10 rural libraries to install fixed wireless equipment in their buildings. It cost the provider relatively little to establish the wireless access point. Consumers benefit from real-time access to information on demand without being tied to a computer terminal or kiosk. In many cases, consumers can access desired information and transfer it to their laptop, PDA, or mobile phone – all wirelessly. At home, the user can then download the information to their home network which may link audio, video or voice equipment.

In Maine, an ambitious wireless pilot project is underway. Funding from the Maine Telecommunications Wireless Access Fund has been earmarked to support the deployment of wireless connectivity in about fifty of the state's public libraries. The sites are scheduled to be ready for public use in June 2004, and more libraries than were originally foreseen may be added, according to the project's Web site (http://www.maine.gov/tools/whatsnew/index.php?topic=Library+News&id=2385&v=Article).

Priming the Library for Change. To conclude, the computer networks that link all corners of the globe simultaneously simplifies and complicates librarians' roles. Libraries of all sizes are faced with pressing questions on how best to embrace the new opportunities and burden posed by changing technology. Where will financial investments be most rewarding? What staffing adjustments are necessary in order to maintain advanced computing expertise? The answers may vary by user expectations, the library's role and strategic plan.

The part librarians play in their interactions with patrons is made simpler by the ease with which countless data sources may be accessed, either on the Web or via traditional use of books and periodicals. Even so, the Web is not all-knowing;

librarians told the study team that some patrons with whom they work are frustrated to find that not every book and article ever written is available online. Users have come to expect that anything they want to know will be found on the Web, and while it's disappointing for them to learn that some information simply won't be found there, herein lies the great opportunity for librarians. Not only does the Web pave the way for library holdings to be more readily located and shared among users around the country or around the world, but the worth of the work of librarians—as guides through an ever expanding wilderness of information sources—actually increases. Hackman agreed:

> *Even if our job changes somewhat …since we're no longer the sole providers/keepers of information the importance shifts to helping users learn how to use effectively the resources directly available to them and how to discriminate between authoritative resources and the rest of the noise.*

More than ever, said another librarian with whom the study team consulted, "librarians are needed to help patrons distinguish between authoritative sources of information and the vast quantities of unmediated materials available on the Net, much of it superficial and of dubious value, if not completely erroneous."

Many online and face-to-face courses are available to librarians to help them hone their information navigation skills and learn from their peers. The Online Computer Library Center, commercial companies such as Amigos, and other entities offer training sessions on using OCLC uber-catalogs like WorldCat, digitizing serial holdings, using metadata to organize digitized collections, and many other subjects. A portal for librarians, WebJunction (www.webjunction.com), was developed by OCLC with support from the Bill and Melinda Gates Foundation to serve as a virtual community where librarians can ask questions of peers about serving patrons and integrating technology into their day-to-day work.

Librarians' jobs have changed as the Net has morphed from a test project developed by the Defense Department into a commercial entity with unprecedented global reach. According to *Toward Equality of Access: The Role of Public Libraries in Addressing the Digital Divide*, a report issued by the Gates Foundation, librarians have responded favorably to the new demands placed upon them.

> *The role of librarians in [local] communities as both technology provider and community educator was crucial for public access computing to be successful. Today, research confirms that these librarians have embraced their new responsibilities enthusiastically. Though the computers bring new stresses, the librarians report an increase in job satisfaction since the arrival of computers.*

As Smithsonian founder James Smithson envisioned when he gave to Congress the funding that would form the nucleus of the Institution, it's appropriate that the Internet—a tool for "the dissemination and diffusion of knowledge" like no other—is the vehicle that makes the Smithsonian Institution Libraries and their specialized holdings more broadly available to a wider public.

-end-