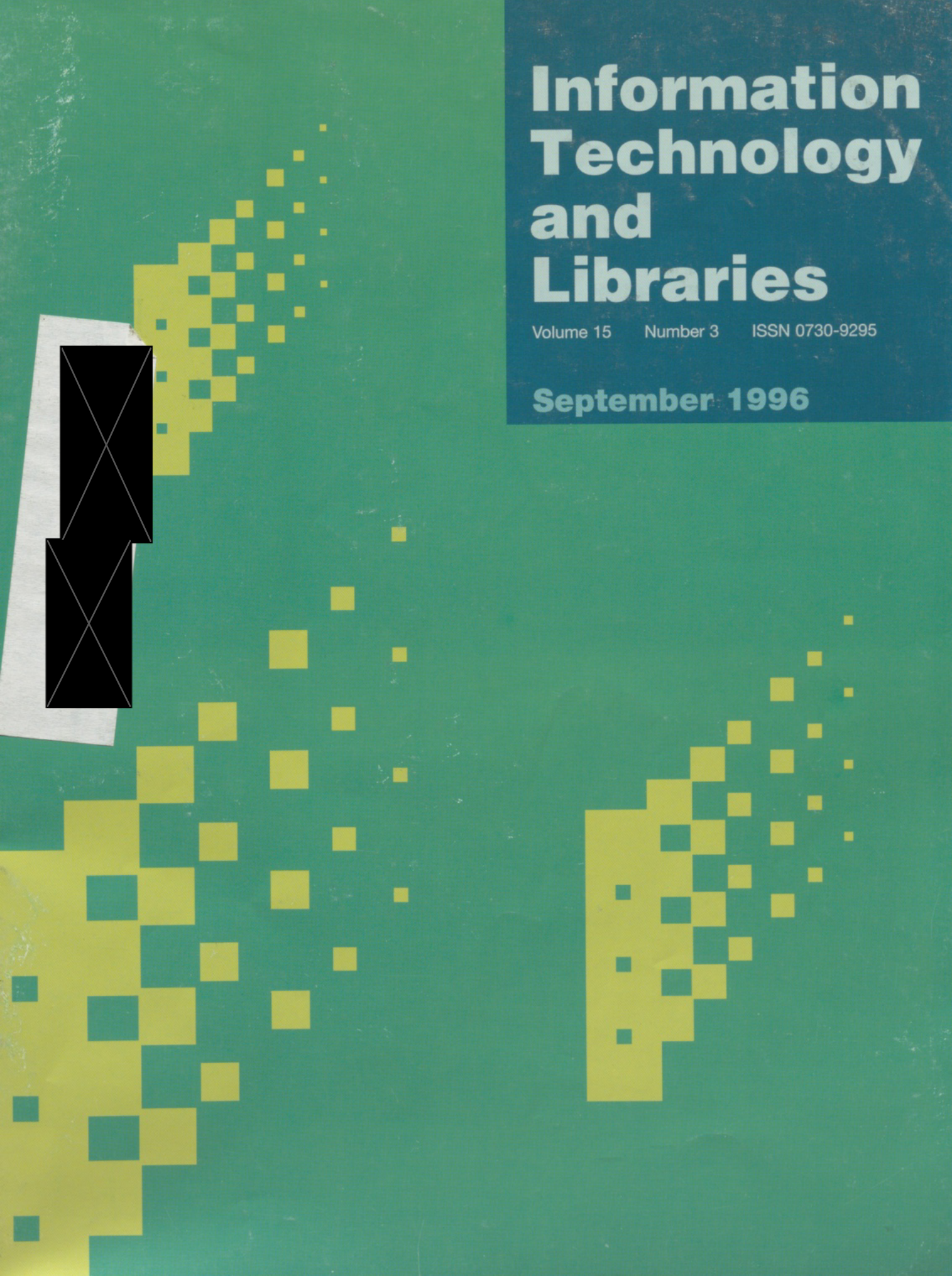


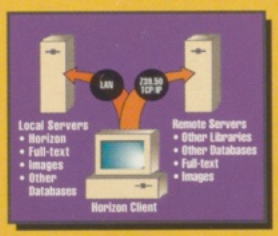
Information Technology and Libraries

Volume 15 Number 3 ISSN 0730-9295

September 1996

The standard by which other client/server systems are judged

Horizon™ is the information management system that sets the standard by giving your library all the benefits of client/server computing.



A true client/server system

The flexibility of an open systems environment means you can add to the system as needs change and adapt to future technologies while leveraging your investment.

Expanded patron access

In addition to your own Horizon database, patrons can use Horizon to gather bibliographic data from any Z39.50-compliant system and connect to global resources over the Internet.

Graphical user interfaces

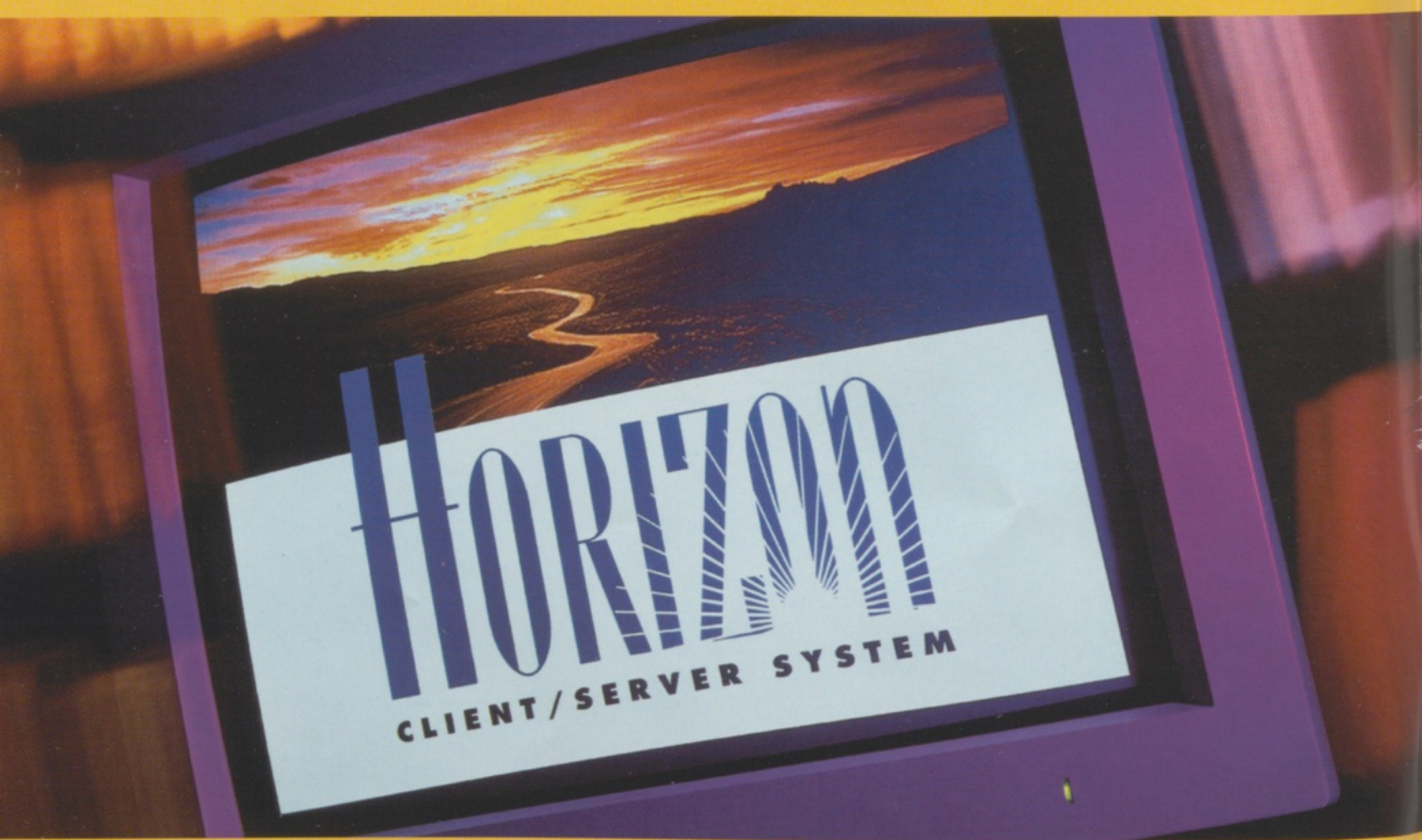
The Horizon client software offers the familiarity of Microsoft Windows,® System 7 (Macintosh®), or OS/2® interfaces. Patrons can search effortlessly with icons, pull-down menus, and point-and-click commands.

Proven by libraries like yours

You need a system with a history of success, one designed by library professionals. Horizon has met those requirements with thorough testing by libraries around the world.

Horizon is available today

You can bring the benefits of Horizon to your library today. Let us show you the leading client/server system.



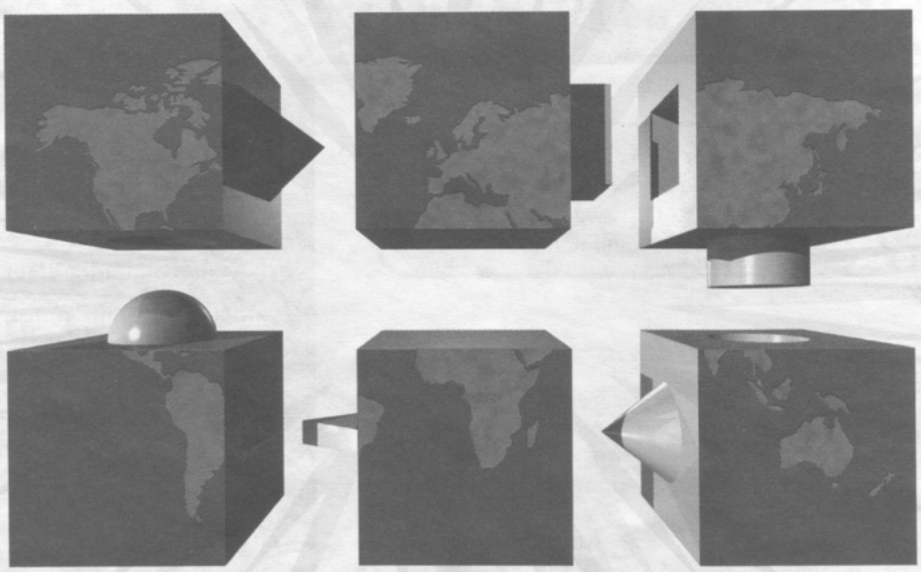
Ameritech Library Services

Academic Division
1007 Church Street
Evanston, IL 60201-3665
(800) 556-6847
Fax (708) 866-0178
academic@amlibs.com

Special & Public Divisions
400 Dynix Drive
Provo, UT 84604-5650
(800) 288-8020
Fax (801) 223-5202
special@amlibs.com
public@amlibs.com

Ameritech

VTLS
Is New Again



VIRTUA
Is Why

◆ Virtua – inherently flexible, intuitive and virtual by design – is built with the newest technological innovations. ● Distributed data warehousing, an elegant and efficient data storage solution, expertly manages data backup and recovery and increases data security. ▼ Future enhancements and new software releases arrive quickly and with the highest quality because we use object-oriented design and rapid application development tools. ■ And three-tier client-server architecture, an evolution of today's systems, enables Virtua to perform for those libraries that demand optimum system and network performance.

Rediscover VTLS Inc., a founder of library automation.

Discover Virtua, *The New Era in Library Automation.*

VTLS

VTLS Inc. 1800 Kraft Drive, Blacksburg, VA 24060 USA • 1-800-468-8857 • WWW.VTLS.COM

Information Technology and Libraries (ISSN 0730-9295) is published quarterly in March, June, September, and December by the American Library Association, 50 E. Huron St., Chicago, IL 60611. It is the official publication of the Library and Information Technology Association, a division of the American Library Association. Subscription price, $22.50, is included in membership dues. Nonmembers may subscribe for $50 per year in the U.S.; $55 in Canada, Mexico, Spain, and other PUAS countries; $60 in other foreign countries. Single copies, $15. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. POSTMASTER: Send address changes to Information Technology and Libraries, 50 E. Huron St., Chicago, IL 60611.

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 8 of the March 1996 issue. Manuscripts of articles, communications, and news items should be addressed to the editor: James J. Kopp, Library Director, Wilson W. Clark Memorial Library, University of Portland, P.O. Box 83017, 5000 N. Willamette Blvd., Portland, OR 97283-0017. Copies of books for review should be addressed to: Barbara Kemp, ITAL Book Reviews, Dewey Graduate Library, University of Albany, SUNY, 135 Western Ave., Albany, NY 12222. Copies of software for review should be addressed to: Tom Jevic, ITAL Software Reviews, University of Illinois at Chicago, Main Library, 861 S. Morgan, Chicago, IL 60607.

As a matter of policy, *Information Technology and Libraries,* as the scholarly organ of LITA, does not review LITA publications. Notice of new publications from LITA will generally be found in the "Other Recent Receipts" column following reviews.

PRODUCTION: ALA Production Services (David Epstein, Bruce Frausto; Christine Squires, Leslie Stella, Laura Tillotson, and Donavan Vicha), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Information Technology and Libraries

Volume 15, Number 3 September 1996 ISSN 0730-9295

Cover Design by Jim Lange
Interior Design by Dianne M. Rooney

Abstracted in *Computer & Information Systems, Computing Reviews, Information Science Abstracts, Library & Information Science Abstracts, Referativnyi Zhurnal, Nauchnaya i Tekhnicheskaya Informatsiya, Otdyelnyi Vypusk,* and *Science Abstracts Publications.* Indexed in *CompuMath Citation Index, Computer Contents, Computer Literature Index, Current Contents/Health Services Administration, Current Contents/Social Behavioral Sciences, Current Index to Journals in Education, Education, Library Literature, Magazine Index, New Search,* and *Social Sciences Citation Index.* Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI 239.48-1992. ∞

Editorial: LITA = Library and Information Technology Advocacy

James J. Kopp

Although the *A* in LITA officially stands for *association*, I have been thinking of the *advocacy* roles of LITA and its members since participating in the orientation and training for LISAN last December. LISAN, as I hope most of you are aware by now, is the Libraries on the Information Superhighway Advocacy Network. Established by LITA and ALA's Office of Information Technology Policy (OITP), with the collaboration of several library associations and initial funding by some of these associations and vendors, LISAN includes more than 100 librarians representing all types of libraries throughout the United States. The objective of the network is to provide a focused means for disseminating information related to issues affecting libraries as the "virtual concrete and asphalt" of the information superhighway is laid, but even more perhaps to get the word out that libraries already possess many of the tools and much of the expertise to assist in constructing that highway.

The establishment and role of LISAN is important for libraries in particular and for the country in general, but we must not lose sight of the fact that each of us is or should be an advocate for libraries and information technology in our respective roles within our institutions, our communities, and even our circle of friends. For advocacy can take on many dimensions, and a role exists for everyone. Much of LISAN's early focus has been on the national issues associated with the Internet such as the Communications Decency Act (in which many LISAN participants played a role in seeking its blockage) and broad economic and political issues such as universal service, but there are issues within your regions, states, cities, and communities that also affect the general direction and understanding of libraries and information technology. LISAN's activities are not just focused on reaction to legislative initiatives but include getting the good word out about libraries. Each of you has many positive stories to tell about how information technology has benefited your library and community.

Like a waterway, a metaphor I prefer to that of the highway, advocacy can work in varying terrain, in varying climates, and in various ways. A waterway often starts as a trickle in the mountains, then becomes a stream that meets with other streams and turns into a rivulet. That waterway joins with others and becomes a river that eventually reaches the sea, where other forces come into play. The water eventually recycles into the atmosphere to once again begin another trickle, most likely on another mountain. Advocacy functions much the same way. It often starts with a trickle of interest and enthusiasm and gradually grows in size and force until there is a sea of support. And then other ideas are recycled back to initiate another stream of advocacy.

Like a waterway, advocacy nourishes the environment along its way. But we also must realize that unimpeded advocacy can be overpowering, even drowning those who become overwhelmed by it (much like unimpeded metaphors!). Thus in some streams of advocacy, channels must be constructed to maintain a proper and efficient flow. LISAN is one method of channeling the energies of advocacy, but there is plenty of space in the library and information landscape for others. We must not assume that one focus of advocacy is all that is needed. Otherwise, we might see sediment build up in our channel or worse, experience an advocacy drought.

Be an advocate for libraries and information technology. Participate in LISAN, state or regional advocacy groups, local activities, or even on a personal level through the one means we all have to be advocates—the vote.

A New Service on the Information Superhighway

If you have been searching for an easy way to authority control your library's current cataloging, try LTI's ***Authority Express*** service.

With ***Authority Express,*** a library uses the Internet to transmit a file of newly cataloged bibliographic records to LTI (via FTP). Overnight, LTI processes the records through its state-of-the-art authority control system. Then, at the library's convenience, it logs into LTI's FTP server to retrieve fully authorized catalog records, along with linked LC name and subject authority records.

Authority Express

- Keeps authority control current at an affordable price
- Integrates easily into existing workflows
- Lowers cost by reducing staff time spent on catalog maintenance
- Provides next-day turn around for up to 5,000 catalog records
- Accepts records for processing even if LTI did not perform the original authority control

"Authority Control for the 21st Century"

LIBRARY TECHNOLOGIES, INC.

2300 Computer Avenue, Suite D-19 Willow Grove, PA 19090

(215) 830-9320 Fax: (215) 830-9422

(800) 795-9504 email: LTI@LibraryTech.Com

Improving Personal-Name Searching in Online Catalogs

Karen M. Drabenstott and Marjorie S. Weller

A sizable proportion of the queries users submit to online catalogs bear personal names. In fact, names figure into the most frequently used search types—author searches and subject searches. Despite findings of online catalog use studies that demonstrate the difficulties users experience entering queries bearing names, the basic system approach to soliciting user queries for personal names and searching for them remains the same as the approaches enlisted by the earliest online catalogs. The purpose of this paper is to examine the performance of user queries that involve personal names and recommend improvements to the basic system approach to soliciting user queries and searching for them. It uses data and analyses from two related research projects to increase understanding of user queries for personal names and to improve the effectiveness of how systems solicit and search for users' queries involving personal names. Improvements include the implementation of a new design for online catalog searching that features search trees to select searching approaches that are likely to yield useful retrievals for the topics, titles, and authors users seek; new methods for soliciting user queries bearing personal names; and enlisting the participation of online catalog users in the evaluation of system prompts, instructions, and messages that request input from them.

A dozen years have passed since the landmark Council on Library Resources-sponsored Nationwide Study of Online Catalogs was conducted.[1] Despite the passing of those dozen years, several dozen follow-up online catalog use and user studies, and hundreds of dozens of new systems that have been deployed in libraries, the basic system approach to soliciting user queries for personal names and searching for them remains the same as the approaches enlisted by the earliest online catalogs. That is, online catalogs prompt users to select a search type from a list that almost always includes author, title, subject, and call number searches. Catalogs then display a dialogue box into which users type their query or instruct users to type their query following a "ready" prompt. Accompanying dialogue boxes or ready prompts where users type their queries are examples of queries or instructions regarding the form that user queries should take. A handful of systems feature *name*, title, subject, and call number searches. Selection of the name search may result in another series of prompts that require users to differentiate between known-item and subject searches for personal names. Alternatively, name searches may induce systems to perform both author searches and subject searches for user queries and distinguish between retrievals that are the result of these two types of searches.

Since the introduction of online catalogs in the early 1980s, librarians, system designers, and researchers have had a very accurate tool to assist them in the study of online catalog performance in the form of transaction logs. Dozens of researchers have studied the access points in these logs, especially points that failed to produce retrievals. Findings from studies on online catalog performance demonstrate that the basic system approach to soliciting user queries and searching for them may result in serious problems that ultimately doom retrieval.

The purpose of this paper is to examine the performance of user queries that involve personal names and recommend improvements to the basic system approach to soliciting user queries and searching for them.

Personal-Name Queries and Online Catalog Searching

Generalizations about Personal-Name Queries

Dozens of online catalog use studies have produced generalizations about online catalog searching and the personal names users enter into known-item and subject searches that suggest problems in the ways in which systems solicit queries from users and search for them. These generalizations are:

1. *Users enter queries using the online catalog's subject search capability that are probably better suited to the catalog's author or title capability and vice versa.*[2-5] Users who enter such queries could be unsure of which searching capability they should use to enter their query (so they use all of them), or they could be surveying results to determine which capability would provide a manageable number of

Karen M. Drabenstott is Associate Professor, School of Information and Library Studies, University of Michigan, Ann Arbor. Her e-mail address is karen.drabenstott@umich.edu. **Marjorie S. Weller** is programmer analyst, Medical Center Information Technologies, University of Michigan, Ann Arbor. Her e-mail address is marjorie.weller@umich.edu.

useful retrievals. Fortunately, such queries are not very common.

2. *Users enter personal names in direct form when systems require the entry of inverted forms and vice versa.*[6–14] Despite systems' repeated attempts to tell users to enter elements of personal-name queries in a specific order, users fail to heed system suggestions and enter elements of their personal-name queries in forms different from the forms that systems expect. In Dickson's study, the percentage of personal names with given names entered first was as high as 37 percent in a system that anticipated names in inverted form.
3. *Users enter personal names with other words.*[15-17] Queries for personal names involve elements in addition to personal-name elements. This is particularly common in subject searches that are accompanied by topics. Examples of such queries are "paintings of pollack" [sic], "descartes future prediction," and "clarence darrow's relegious [sic] views."
4. *The middle names or initials users include in their queries for personal names are sometimes counterproductive in helping them find the heading used in the catalog.*[18-21] There are several reasons why the inclusion of middle names or initials is counterproductive; most center on features of the system into which the query is entered. Here are three examples.
   - The indexed heading had no initial and the user-entered term had an initial; the system placed the user in an alphabetical index of personal names following the name desired by the user and offered no capability for browsing backward in the alphabet.
   - The system submitted the user query bearing spelled-out name(s) and initial(s) to its keyword searches; the desired heading had no initial, so the system failed to produce retrievals in keyword-in-heading or keyword searches.
   - The system submitted the user query bearing spelled out name(s) and one or more incorrect initials to its keyword searches; the desired heading had no or a different initial, so the system failed to produce retrievals in keyword-in-heading or keyword searches.
5. *Users rarely enter personal names queries containing elements that pertain to the name such as dates or titles.*[22-24] Despite the assignment of qualifying elements to personal-name headings such as dates, titles, and qualified names, users almost never include much more than name elements (given name or initial, surname, middle name or initial) and one or more topical elements.

Searching for Personal-Name Queries

All five generalizations have implications for both the ways in which systems solicit queries from users and how systems search for them. With regard to soliciting queries from users, online catalogs do little more than prompt users to select a search type from a list which almost always includes author, title, subject, and call number searches, and to type their query into a dialog box or following a ready prompt. Accompanying dialog boxes or ready prompts where users type their queries are sometimes examples of queries or instructions regarding the form that user queries should take. In summation, systems give users little instruction or directed guidance as to the entry of elements of their personal-name queries and, thus, systems have no knowledge about the elements of users' personal-name queries that they can use to produce useful retrievals.

With regard to searching for personal-name queries, operational systems typically submit user queries to one of three search approaches: (1) alphabetical searches, which result in a list of author-name headings or subject headings in the alphabetical neighborhood of user-entered queries, (2) keyword-in-heading searches, which result in a list of author-name headings or subject headings bearing words in user-entered queries, and (3) keyword searches, which result in bibliographic records bearing words in user-entered queries. Word order is crucial in alphabetical searches because the system's response to user queries in the form of an alphabetical list of headings is based on the initial words in user queries. The two keyword searches are forgiving about word order because systems retrieve headings regardless of the word order in queries. However, spelling is important in keyword searches. If one word is misspelled, keyword searches for the entire query—whether it is one word or a multiword query—will always fail. Alphabetical searches are much more forgiving about misspellings. Misspelled words that occur at or toward the end of a one-word or multiword queries that are submitted to alphabetical searches may result in a display of titles or headings in the alphabetical neighborhood of the desired term despite the spelling error.

In recent years operational systems have featured two or more searching approaches per search type. For example, systems might give users a choice between alphabetical and keyword searches for their personal-name queries. Unfortunately, systems give users little instruction or guidance as to which search approach is likely to produce useful retrievals for their personal-name queries. The search approach users choose may be based on previous experience, luck, serendipidy, or a combination of all three. It should, however, be based on

the likelihood that the selected approach will produce useful retrievals.

Taking the Content and Structure of Personal-Names into Consideration

The content and structure of assigned headings for personal names must be considered when determining how to improve the performance of user queries involving personal names.

Personal-name headings come from main- and added-entry headings in libraries' bibiographic records. Catalogers establish these headings by following AACR2 rules and guidelines and previous Library of Congress (LC) practice. They refer to LC's Name Authority File (LCNAF) to verify personal names. References are generated from LCNAF records for names used in libraries' bibliographic records. Catalogers also refer to the *Subject Cataloging Manual: Subject Headings* to add subdivisions to personal-name subject headings. The results are single headings bearing elements for personal names, topical subjects, and geographic names. The order of elements, when applicable, in such headings is: (1) surname, (2) given name or initial, (3) middle name or initial, (4) one or more dates, (5) topical subdivisions, and (6) geographic subdivisions. Titles and qualified names (usually given name and middle name, initials, or a combination of the two) may also be elements of personal-name headings.

Generalizations about queries involving personal names demonstrate that users rarely, if ever, express all six elements of personal-name headings in their queries. Furthermore, when two or more elements are included, users do not necessarily enter elements in the sequence prescribed by personal-name headings. When searching for user queries bearing personal-name elements, systems need information from users about the particular elements included in queries to distinguish personal-name elements from nonname elements. Systems can then accommodate the rigid structure of personal-name headings by searching for as many elements as needed to provide useful retrievals and utilize the appropriate search approach in view of the information provided by users.

Personal-Name Queries in Subject Searches

The personal-name queries that are the focus of this paper came from data sets generated in two separate but related sponsored-research projects. The first project, titled "Enhancing a New Subject Access Design to Online Catalogs," was supported by OCLC's Library and Information Science Research Grant Program.[25] We obtained transaction logs from the online catalogs of Syracuse University, the University of California at Los Angeles (UCLA), the University of Kentucky, and the University of Michigan; extracted a total of about two thousand user queries for subjects from the logs; and performed a manual analysis of these queries. The manual analysis required us to categorize user queries according to the types of elements present in them (i.e., topical subjects, corporate names, geographic names, personal names, and combinations of two or more elements), develop subcategories of queries corresponding to the extent to which they matched subject headings and other subject-rich terms in bibliographic records, and identify queries that were neither matches of subject headings nor other subject-rich terms in bibliographic records. The results of the manual analysis demonstrated the extent to which users entered subject queries for personal names and the subject searching approaches that were likely to produce useful retrievals.

The second project, titled "Testing a New Subject Access Design to Online Catalogs," was supported by the Department of Education's College Library Technology and Cooperation Grants.[26] The purpose of this research project was to test a new subject-access design. This design featured an online catalog that had a wide range of subject-searching capabilities and search trees to govern the system's selection of searching capabilities in response to user queries. Search trees asked users to differentiate between subject queries bearing personal names and all other subject queries. On their own, the search trees then determined the extent to which user queries matched subject headings and other subject-rich terms in bibliographic records. This machine-based analysis resulted in the selection of a subject searching approach that was likely to produce useful retrievals in response to user queries.[27]

Despite the fact that the personal-name queries from the two data sets were entered in subject searches, we feel that findings about these queries are generalizable to queries involving names, whether they be known-item or subject searches.

The research questions addressed in this paper are:

1. How can online systems choose search approaches on their own that are likely to produce useful retrievals in response to user-entered queries for personal names?
2. How should online systems solicit queries involving personal names from users?
3. How do users respond to an experimental online catalog that prompts them for the different elements of their personal-name queries?

Producing Useful Retrievals in Response to User Queries

Categorizing Subject Queries Extracted from Transaction Logs

Search trees hold much promise for assuming the burden of determining the subject-searching approach likely to produce useful information for user queries. The designers of the Okapi experimental online catalog first defined search trees as "a set of paths with branches or choices, which enables the system to carry out the most sensible search function at each stage of the search."[28] The search trees they implemented in Okapi "evolved through a process of discussion and trial and error" and placed more emphasis on searching the titles than the subject headings in Okapi's cataloging records because only half of these records contained subject headings.

Some operational online catalogs have subject-searching routines that resemble search trees. For example, the online catalog of the University of Illinois at Urbana-Champaign responds to user queries for subjects with keyword searches of assigned subject headings. When users terminate searches, the system prompts them to continue and gives the results of a title-keyword search.[29] The Illinois online catalog always performs keyword searches of subject-heading fields before title-keyword searches because the former consumes fewer system resources than the latter.

The search trees that we used to subcategorize categorized queries were the result of the empirical study of the subject terms users entered into online catalogs.[30] This study demonstrated that the subject terms users entered into online systems possessed certain characteristics that revealed the subject-searching approaches most likely to succeed in terms of producing assigned subject headings and bibliographic records on the topics users seek. Examples of such characteristics were the number of words in user queries, the extent to which user queries matched controlled vocabulary terms, and their ability to produce retrievals in response to certain subject-searching approaches.

A research team at the University of Michigan selected the initial queries users entered in subject searches from the four libraries' transaction logs. We chose initial queries because subsequent queries might have been unnecessary if catalogs responded to initial queries with useful retrievals. Queries were selected from online catalog terminals searched exclusively by library patrons over a one-month period during these four academic institutions' winter semester or spring quarter. We first categorized queries by the type(s) of elements present in them: (a) topical subjects, (b) corporate names, (c) geographic names, (d) personal names, and (e) combinations of two or more elements (a-d).

We then placed categorized queries into search-tree subcategories using the same series of decisions that an online catalog that was programmed with search trees would make. Drabenstott and Vizine-Goetz[31] discussed search trees in depth and provided flowcharts depicting search-tree decision points; thus, only a brief description of search-tree categories is given here. Our first step was to segregate user queries containing personal names from those lacking personal names. The former queries were subjected to analyses that were different from those performed on the latter. This paper only describes analyses performed on queries involving personal names.

At most, we considered only three elements—last names, first names, and topics—to determine whether keyword-in-heading searches would produce retrievals. If such searches failed, we dropped first names and submitted surname and topical element(s) to keyword-in-heading searches. If keyword-in-heading searches failed altogether, we then considered the same three elements as we considered earlier—last names, first names, and topics—to determine whether keyword searches would produce retrievals. If keyword searches failed to produce retrievals, we ignored given-name elements in a second keyword search. If keyword approaches still failed, we ignored the topic element(s) entirely and performed an alphabetical search using as many name elements as matched assigned subject headings and references in online catalogs.

For example, the three elements "reagan" (surname), "ronald" (given name), and "movies" (topic) might not be successful retrieving single assigned subject headings bearing these words because the topic element is not used in subdivisions under subject headings for personal names. Keyword approaches may yield bibliographic records bearing the topic elements and name element(s) because titles and/or note fields may contain the topic element "movies." Retrieved records may have the assigned subject heading "Reagan, Ronald, 1911– — Career in film." In actual online searches, a feedback capability should follow up keyword searches so that users can find additional information using relevant headings in retrieved records. If users are dissatisfied with the results, systems could continue with the alphabetical approach but omit the topical element of the query to effect a match of a personal-name subject heading.

Categorized Initial Queries

We extracted a total of 1,919 initial queries in subject searches from the transaction logs of online catalogs at

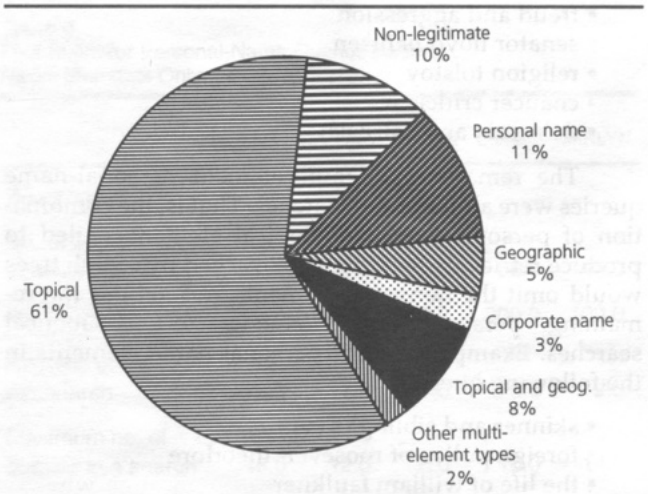


Figure 1
Types of queries
Figure 1. Types of initial queries

Syracuse University (571 queries), UCLA (511 queries), the University of Kentucky (418 queries), and the University of Michigan (419 queries). Figure 1 summarizes the total percentages of types of initial queries across all four libraries.

Of the 1,919 total queries, personal names numbered 219 and accounted for 11 percent of user queries. The most frequent multi-element query contained topical and geographic elements and represented about 8 percent of user queries for subjects. Amongst the 2 percent of other multi-element queries were 32 queries bearing personal-name and topical elements. There were 203 nonlegitimate queries; these were expletives, gibberish, explicit sex terms, known-item searches, etc., and accounted for 10 percent of user queries. Queries for topical subjects were the majority, accounting for over three-fifths of initial queries. When nonlegitimate and non–personal-name queries were discarded from subsequent analyses, a total of 251 subject queries that involved personal names were analyzed.

Personal-Name Queries Bearing Topical Elements

Table 1 is a fact sheet for the 32 user queries for personal names that contained topical and personal-name elements. These queries were candidates for keyword-in-heading and keyword searches.

Queries bearing personal-name and topical elements were not very common in the logs we reviewed. Except for ORION (which contributed only three queries of this type), users entered two or more queries following their initial query. Some searches were quite long.

Table 1
Fact Sheet for Personal-name Queries Bearing Topical and Name Elements

	Total	SULIRS	ORION	LS/2000	MIRLYN
Searches					
No. of search administrations	32.0	13.0	3.0	5.0	11.0
No. of queries	105.0	48.0	4.0	14.0	39.0
Avg. no. of queries per search	3.3	3.7	1.3	2.8	3.5
Maximum no. of queries in a search	13.0	13.0	2.0	9.0	13.0
Initial queries					
Avg. no. of words per query	3.4	2.9	3.0	3.8	4.0
Maximum no. of words in a query	10.0	5.0	4.0	5.0	10.0
Avg. no. of retrievals per query	N/A	8.0*	0.7†	N/A	N/A
Percentage of queries with zero retrievals	N/A	61.5	66.7	N/A	36.4
Percentage of queries with retrievals > 999	N/A	0.0*	0.0†	N/A	N/A

*Number of bibliographic records retrieved. This was an estimate based on substituting 3,000 retrievals for the 99,999 retrievals that was written to logs when number of retrievals exceeded about 1,000.
†Number of assigned subject headings retrieved.

For example, a MIRLYN user entered twelve more queries following his initial query "skinner and sibling\s" (sic). Subsequent queries demonstrated the user's interest in siblings, e.g., "skinner and siblings," "gesell and siblings," "siblings and the oedipal complex," and "theoretical approach to sibling differences."

Queries bearing personal-name and topical elements were quite long; they averaged 3.4 words per query. Of course, this type of query required at least two access points: a name element and a topical element. In SULIRS, postings were quite low. They were even lower in ORION because this system featured a command that was separate from the general subject command for the entry of subject queries for personal names.

We first tested queries to determine if they were candidates for keyword-in-heading searches. This included ignoring first and middle names or initials to effect matches. We discarded queries that were successful producing retrievals through keyword-in-heading searches from subsequent analyses. We submitted remaining queries to keyword searches and, if necessary, ignored first and middle names or initials to effect matches. The alphabetical approach was "the search of last resort." Table 2 shows the result of the analysis.

A total of 32 queries contained both topical and personal-name elements. What was particularly surprising about this analysis was that almost as many queries bearing personal names should have been submitted to search trees for subjects generally as to search trees for personal names. Examples of these queries are:

- keyenesian economics (sic)
- the hapsburg monarchy
- notions of buddha
- myers-briggs type indicator
- hannibal and the battle of carthage
- ras tafari

These and the several other queries that involved personal names presented a problem. They contained personal-name and topic elements, but they would have been satisfied by subject searches in the search trees for subjects generally. End users would not know (and should not *have* to know) that these queries would be satisfied by the search trees for subjects generally. Additional research is needed to help users and systems differentiate queries for subjects generally from queries for personal names.

A handful of personal-name queries were keyword matches. Examples are:

- clinton and poverty
- freud and aggression
- senator lloyd bentsen
- religion tolstoy
- chaucer criticism
- delacroix and colr (sic)

Table 2
Types of Keyword Matches

Type of Match	Personal-Name Searches	Subjects Generally
Keyword-in-record	5	0
Keyword-in-subdivided-heading	2	1
Alphabetical	12	1
Exact	N/A	4
Title	N/A	5
None	0	2
Total	19	13

The remainder of multi-element personal-name queries were alphabetical matches. That is, the combination of personal-name and topical elements failed to produce retrievals, so systems governed by search trees would omit the nonname elements and submit the remaining personal-name elements to alphabetical searches. Examples are the personal-name elements in the following queries:

- skinner and sibling\s (sic)
- foreign policy of roosevelt theodore
- the life of william faulkner
- descartes' future prediction
- paintings of pollock/pollack (sic)
- clarence darrow's relegious views (sic)
- greek mythology's influence on shakespeare

Since alphabetical searches only considered name element(s), users would not necessarily retrieve titles on the specific topics cited in their queries. For example, searches on the "foreign policy of roosevelt theodore" in which the topical elements "foreign" and "policy" were ignored would retrieve titles on many different topics; users would have to persevere to find the several titles that mentioned Roosevelt's foreign policy, e.g., *Roosevelt and the Caribbean, Velvet on Iron: The Diplomacy of Theodore Roosevelt, Roosevelt and the Russo-Japanese War: A Critical Study of American Policy in Eastern Asia in 1902–5,* and *Theodore Roosevelt: Confident Imperialist.*

Personal-Name Queries Bearing Name Elements Only

Of the 1,919 total queries, about 11 percent contained personal-name element(s) only. These queries were candidates for alphabetical searches only. Table 3 is a fact sheet for these user queries.

Searches ranged from one to three queries. Some searches were quite long. For example, a MIRLYN user entered thirteen more queries following his initial query "hart, gary." Subsequent queries demonstrated the user's interest in this person's bid for the presidency, e.g., "extra marital affairs," "sex," "scandals," "sex and presidential candidates," "mass media," "media coverage of elections," "presidential scandals," and "sex in politics."

Personal-name queries bearing personal-name elements were only about two words long. A few, however, were five or six words long. Examples of these long personal-name queries are:

Table 3
Fact Sheet for Personal-Name Queries Bearing Name Elements Only

	Total	SULIRS	ORION	LS/2000	MIRLYN
Searches					
No. of search administrations	219.0	75.0	30.0	66.0	48.0
No. of queries	500.0	143.0	48.0	200.0	109.0
Avg. no. of queries per search	2.3	1.9	1.6	3.0	2.3
Maximum no. of queries in a search	16.0	12.0	6.0	16.0	14.0
Initial queries					
Avg. no. of words per query	1.8	1.7	1.6	1.7	2.1
Maximum no. of words in a query	6.0	5.0	4.0	3.0	6.0
Avg. no. of retrievals per query	N/A	174.5*	1.0†	N/A	N/A
Percentage of queries with zero retrievals	N/A	20.0	86.7	N/A	29.2
Percentage of query with retrievals > 999	N/A	1.3*	0.0†	N/A	N/A

*Number of bibliographic records retrieved. This was an estimate based on substituting 3,000 retrievals for the 99,999 retrievals that was written to logs when number of retrievals exceeded about 1,000.
†Number of assigned subject headings retrieved.

- strunk w oliver william oliver 1901
- mubarak muhammad husni 1928
- prosser walter lee
- van der velde
- dr martin luther king jr
- frank lloyd wright

Form of Name Entered

Subject searching for personal names was handled in three different ways in the four online catalogs from which transaction log data were obtained. MIRLYN and LS/2000 featured two-step approaches that required users to enter inverted forms of names, e.g., "shakespeare, william," or "twain mark," or enter the surnames, e.g., "clinton" or "gauguin." MIRLYN and LS/2000 responded with an alphabetical list of assigned subject headings in alphabetical proximity to the term entered. Users who selected listed subject headings retrieved a display of bibliographic records assigned the selected heading. ORION's keyword-in-subdivided-heading search did not require users to enter inverted forms of personal names. It was a two-step process in which the system responded to user-entered queries for names with a list of assigned subject headings bearing the words in user queries. Users who selected listed subject headings retrieved a display of bibliographic records assigned the selected heading.

MIRLYN's keyword (k=) search and SULIRS' keyword (lc; or sb;) search were keyword-in-record searches; consequently, the order of elements in personal name queries did not matter. Keyword-in-record searches were one-step approaches in which systems retrieved bibliographic records bearing the entered terms. When users entered surnames using terms that were also given names, e.g., "grant," "kelly," "thomas," the number of retrieved records could be quite high because SULIRS and MIRLYN retrieved records bearing these words irrespective of their use as given names or surnames.

Figure 2 summarizes the forms of names entered into the three online catalogs featuring keyword searching—SULIRS, ORION, and MIRLYN—and the two online catalogs featuring alphabetical searching—LS/2000 and MIRLYN.

A little over a third of queries consisted of surnames only. Examples are:

- lawrence
- hitchcock
- fukasa

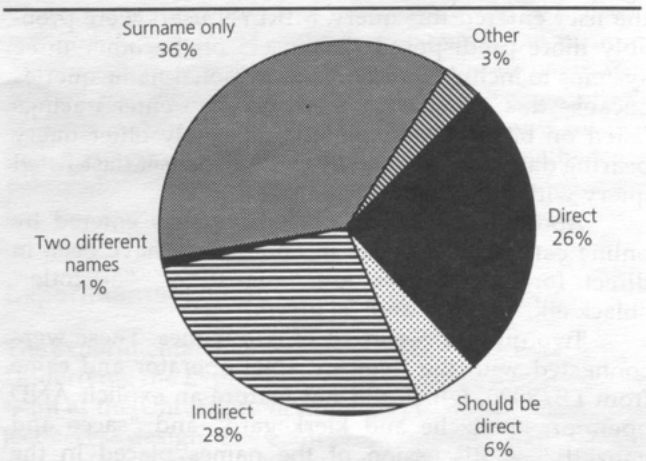


Figure 2
Forms of personal-name queries

- shakespeare
- von sternberg

Indirect forms of names were quite common, accounting for between a quarter and a third of the personal-name queries. Examples are:

- hart, gary
- woolf virginia
- corneille, pierre
- abrahams r d
- wittgenstein, ludwig

Direct forms of names were also quite common, accounting for about a quarter of the queries. Examples are:

- gertrude stein
- william james
- shinon peres (sic)
- jackson pollack (sic)
- reginald wright kauffman

A few names included dates; all but one such name were indirect forms entered into MIRLYN. Examples are:

- strunk w oliver william oliver 1901
- mubarak muhammad husni 1928
- marshall george c george cutlett 1880–1959
- swift, jonathan, 1667–1745
- bach alexander 1813–1893

Queries with dates were so similar to the content, order, and format of assigned subject headings that it was very likely that users entered tracings that were displayed on bibliographic records. For example, the query "strunk w oliver william oliver 1901" was probably the qualified name in the assigned subject heading "Strunk, W. Oliver (William Oliver), 1901– " of a bibliographic record that was displayed on the screen while the user entered this query. MIRLYN users were probably more predisposed than users of the other three systems to include dates in their personal-name queries because this system prompted users to enter tracings listed on bibliographic records. The only other query bearing dates was entered by an ORION user (last-listed query with dates above).

About 6 percent of personal names entered by online catalog users were in and should have been in direct form. Examples are "augustine," "aristotle," "black elk," "ovid," and "el greco."

Two queries consisted of two names. These were connected with the Boolean AND operator and came from LS/2000, which did not feature an explicit AND operator: "nietzche and kierkegard" and "sacco and vanzetti." A discussion of the names placed in the "other" category is interesting. One name consisted only of initials, i.e., "fdr." LCNAF provided a *see* reference under these initials to direct users to the authorized form of name. The given name "vincent" was entered in a search in which the user was probably interested in the artist "Vincent van Gogh." Searches in the LCNAF and several online catalogs failed to verify "n scribner richard," "yassin adj ramadan," and "faisal ibn 'abd alaziz." A *see* reference would have helped the user entering "le roi soliel" to find information on the French King Louis XIV.

Alphabetical Approach for Personal-Name Queries

The alphabetical search was the recommended approach for personal name queries bearing no topical elements. This search placed users in an alphabetical index of personal names in which their entered names were or would be listed in the alphabet. It is important that users enter the correct element for last names because the system's placement at a particular point in the alphabetical index near the desired personal-name heading depends on this element. Table 4 shows the result of our analysis of end-user queries bearing personal name elements only.

Systems governed by search trees would submit a total of 211 queries to the alphabetical approach for personal-name queries. Eight queries bearing personal names should have been submitted to various searches controlled by the search tree for subjects generally. Desired headings would have been listed as a result of the alphabetical search that would have been provoked by search trees for user queries for subjects generally. Examples of these queries and matching topical subject headings are:

- odysseus; matching subject heading: Odysseus (Greek mythology)
- aphrodite; matching subject heading: Aphrodite (Greek deity)
- philoctetess (sic); matching subject heading: Philoctetes (Legendary character)
- oedipus; matching subject heading: Oedipus (Greek mythology)

Table 4
Types of Alphabetical Matches

Type of Match	Personal-Name Searches	Subject Searches Generally
Alphabetical	211	1
Exact	N/A	4
Keyword-in-main-heading	N/A	3
Total	211	8

These names were connected with mythological, fictitious, or legendary characters. Systems that prompt end users for elements of personal-name queries or handle personal-name queries separately from subject queries generally must experiment with the wording of questions that ask users to distinguish between types of queries. An even better solution would be for systems to process subject headings bearing qualifiers such as "(Legendary character)," "(Fictitious character)," "(Greek deity)," and "(Roman deity)" into indexes for both personal names and subjects generally.

Browsing in Alphabetical Searches

Over two-thirds of alphabetical matches would lead users to personal-name subject headings on their topics of interest. Examples of satisfactory alphabetical matches of subject headings and references are:

- corneille, pierre; matching heading: Corneille, Pierre, 1606–1684
- laban; matching heading: Laban, Rudolf von, 1879–1958
- twain, mark; matching heading: Twain, Mark, 1835–1910
- augustine; matching heading: Augustine, Saint, Bishop of Hippo
- de genlis; matching heading: De Genlis, Stéphanie Félicité, comtesse, 1746–1830 (reference)

Users had to browse for about 20 percent of their desired subject headings. Some users would have found their desired subject headings on the screen immediately preceding or following an initial screen of a dozen or fewer subject headings for personal names. Other users would have had to browse more than one screen. For example, the user who entered the query "grant" would have had to browse through many names before arriving at "U's," where the desired subject heading "Grant, Ulysses S. (Ulysses Simpson), 1822–1885" resided. Other examples of queries that required browsing to reach the desired personal-name subject headings are:

- lawrence; matching heading: Lawrence, D. H. (David Herbert), 1855–1930
- neera; matching heading: Neera, 1846–1918
- torres; matching heading: many names beginning with "Torres"
- clinton; matching heading: Clinton, Bill, 1946–
- co prbusier (sic); matching heading: Corbusier, 1887–1965 (reference)
- kelly; matching heading: many names beginning with "Kelly"
- jackson pollack (sic); matching heading: Pollock, Jackson, 1912–1956
- farakan louis; matching heading: Farrakhan, Louis
- butler; matching heading: many names beginning with "Butler"

Several personal-name queries were common surnames, e.g., "lawrence," "butler," "kelly." A few personal-name queries that consisted of first- and last-name elements and required much browsing had spelling errors, e.g., "co prbusier," "jackson pollack," and "farakan louis." Such errors typically resulted in the display of an alphabetical list of personal-name subject headings that did not place users close to the desired heading; consequently, they would have had to do a considerable amount of browsing to find the desired headings.

When we were unable to verify names in searches of several online catalogs, we determined that searches for these names were the result of "collection failure." Examples are:

- klagsburn
- steinway henry
- marie teppp
- n richard scribner
- abrahams r d

Testing a New Design for Personal-Name Searching in an Experimental Online Catalog

In the manual analysis of user queries for personal names, we used our judgment to reorder the name elements of direct forms of personal-name queries so that they reflected the order of name elements in assigned subject headings for personal names. This included the analysis of personal name queries bearing elements for both personal names and topics. In operational systems governed by search trees, systems would ask users whether their queries involved a personal name, prompt them for the various elements, and use the surname and given name elements to place users in the appropriate location in the alphabetical name index. We now turn to highlights of a study in which end users searched for personal-name queries in an experimental online catalog that featured search trees for subject searching.

Experimental Online Catalog Development

The experimental online catalog ASTUTE (*A Search Tree Underlying the Experiment*) was developed by a project team at the University of Michigan to test the new subject-access design. The team programmed ASTUTE on a stand-alone Gateway 2000 33MHz 486 IBM-compatible microcomputer with 8 megabytes of RAM and a VGA color monitor. The operating system was MS-DOS ver-

sion 5.0. A dot-matrix printer and a mouse were attached to the microcomputer for use by ASTUTE project staff during development work and end users during online retrieval tests.

The databases of the ASTUTE experimental online catalog were created from two data sources: machine-readable cataloging (MARC) records for bibliographic data from the two participating libraries in selected subject areas of the Library of Congress Classification (LCC) and MARC records for subject-authority data from the CD/MARC Subjects CD-ROM distributed by the Library of Congress. The number and subject areas of MARC bibliographic records were:

1. Mardigian Library of the University of Michigan-Dearborn: 14,686 bibliographic records in Computer Science (QA76's) and Technology (T-TX).
2. Lilly Library of Earlham College: 11,976 bibliographic records in American History (E1-F1199).

We did not combine bibliographic records into a single database. Rather, we used the two libraries' bibliographic records to create separate, searchable databases on computer science and technology for the University of Michigan-Dearborn (UM-D) and on American history for Earlham College, respectively.

Subject Searching in the Experimental Online Catalog

We tested the retrieval effectiveness of the experimental online catalog with search trees by comparing its performance with the performance of an experimental online catalog in which subject-searching approaches were assigned at random. To accomplish this, we designed the ASTUTE experimental online catalog to feature two online catalogs: the Blue System, in which search trees governed the system's selection of a subject-searching capability, and the Pinstripe System, in which the system selected a subject-searching capability randomly. These systems were purposely designed to be very much alike to focus the attention of library patrons and staff on the retrieval of useful information in response to their queries. The Blue and Pinstripe systems had virtually the same interfaces and accessed the same bibliographic and authority databases. Except for the Blue System's enhancement with the search trees, the two systems and their capabilities were the same.

Search trees exemplified the searching strategies used by expert search intermediaries. Intermediaries use controlled vocabulary because it yields relevant output. When controlled vocabulary is not available to express user queries, intermediaries conduct free text searches of titles and abstracts to retrieve a few relevant records, review results to find relevant controlled vocabulary, and then incorporate such vocabulary into the ongoing search. The search trees performed in a similar manner. They invoked searching approaches that looked for matches of user queries in subject-heading fields of cataloging records before enlisting keyword-search approaches that looked for matches in title fields or in a combination of title and subject-heading fields. Search trees for subject queries for personal names effected keyword-in-heading matches of name and topical elements in user queries before ignoring topical elements and displaying an alphabetical browsing list of personal-name subject headings in the alphabetical neighborhood of personal-name elements in user queries. Thus, decisions that the search trees made about responding to user queries with matches of subject headings and words in bibliographic records were very similar to the decisions that judges made about matching user queries in the matching study described in the first half of this paper.

Administering Online Retrieval Tests in Libraries

The project team transported the Gateway computer bearing ASTUTE to two data-collection sites—Mardigian Library at the University of Michigan–Dearborn and Lilly Library at Earlham College. The microcomputer was dedicated to use of ASTUTE. At UM- Dearborn, ASTUTE was located in a quiet study area of the library near the computer science, engineering, and technology stacks. Thus, ASTUTE searchers would not have to go very far to access the library material they retrieved in their searches of the experimental online catalog. At Earlham College, ASTUTE was located in the reference area of the library near the library's MARCIVE CD-ROM-based online catalog and other CD-ROM reference sources. Lilly Library reference staff were also nearby and directed patrons to ASTUTE when they felt patrons would find useful material in the system. At both libraries, signs were placed near ASTUTE to encourage library patrons to use the system.

The ASTUTE experimental online catalog performed recruiting functions on its own. Introductory screens invited users to participate in the experiment; told users how to operate the keyboard and mouse, make selections, and print screens; and asked them to conduct a computer-based search on a topic of their own choosing in the system. ASTUTE told users it was logging their searches, relevance assessments to displayed titles, and responses to closed-ended presearch and postsearch questions regarding system performance. Library users were entirely on their own to read screens, conduct searches, and answer questions.

The data-collection period at the University of Michigan-Dearborn lasted five weeks, from March 12 to April 19, 1993, during which time ASTUTE logged a total of 826 search administrations. At Earlham College, data collection lasted thirteen weeks, from February 23 to May 28, 1993; ASTUTE logged a total of 238 search administrations. Thirty-three of the 1,064 total search administrations involved library staff at the two participating libraries.

Interviewers were not present to monitor system use or conduct postsearch interviews with system users. On its own, ASTUTE recruited users and logged their queries, searches, relevance assessments to retrieved titles, and answers to closed-ended presearch and postsearch questions. Because human monitors were not present, we expected users to enter searches for topics that were not represented in the experimental online catalog. We also expected them to leave the experiment without completing the full search administration, that is, answer presearch and postsearch questionnaires and conduct full-length searches in ASTUTE. Of the 826 UM-D search administrations, half (412) were usable; of the 238 Earlham search administrations, a little under a half (112) were usable. Details about usable and unusable individual search administrations are given in the final report of the project.[32] Our focus in this paper is on the analysis of search administrations bearing user queries with personal names.

Soliciting Personal-Name Queries

This paper's second research question focuses on how online systems can solicit queries involving personal names from users. Based on our analysis of user queries bearing personal names, we recognized that users needed to provide systems with knowledge about the individual elements of their personal-name queries. The techniques that ASTUTE enlisted were the same as the techniques used in the matching study that is the subject of the first half of this paper. That is, ASTUTE's search trees asked users to differentiate their queries for subjects generally from their queries for personal names (figure 3).

The system then prompted users to enter the surname, given name or initial, and topical elements of their queries. Search trees then submitted personal-name queries bearing both personal-name and nonname elements to keyword-in-heading and keyword searches before discarding nonname elements and submitting the remaining elements of queries to alphabetical searches.

Figures 4–7 show the interaction between the experimental system and the user. Figure 4 shows the user typing in the last name "jefferson" in response to a system prompt. When the user hits the Enter key, the system produces a second dialog box and prompts the user to enter the person's first name. The user responds by typing in "thomas" (figure 5). When the user hits the Enter key, the system produces a third dialog box and prompts the user to enter a topic connected with this person. The user responds by typing in "on slavery" (figure 6).

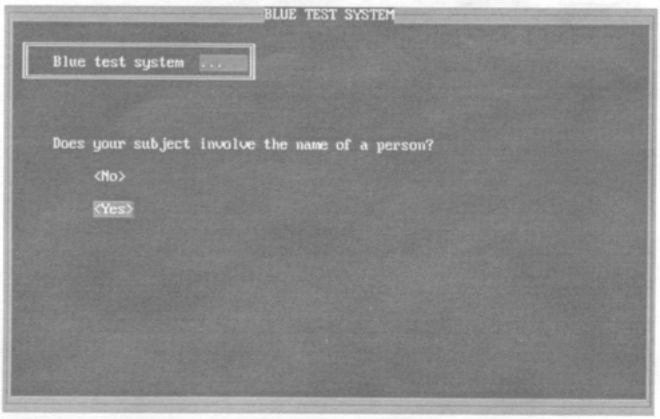


Figure 3
Asking users whether their queries involve names

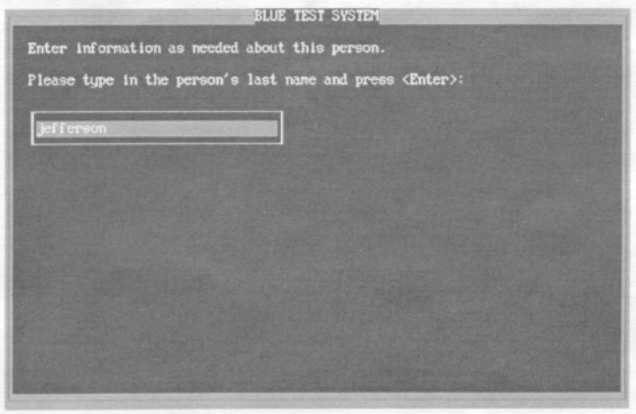


Figure 4
User entry of last name query element

The system submits the entire query to keyword-in-heading searches because this query consists of name and nonname elements. It successfully produces retrievals. Figure 7 shows the first of two retrieved records. This record is retrieved because of matches on the words "thomas," "jefferson," and "slavery" in a single subject heading.

Following the display of retrieved bibliographic records, the system asks users if they want to continue searching. Users who respond positively to such

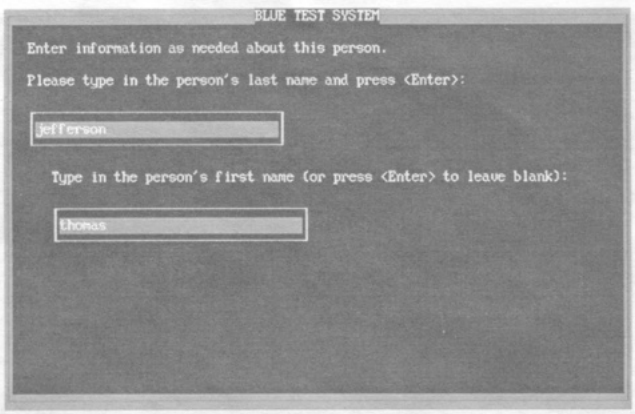


Figure 5
User entry of first name query element

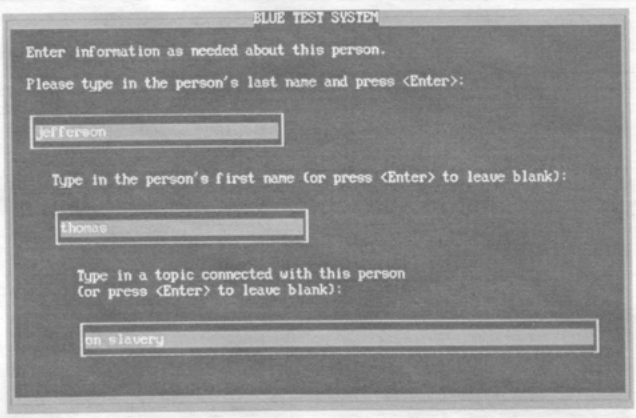


Figure 6
User entry of topical query element

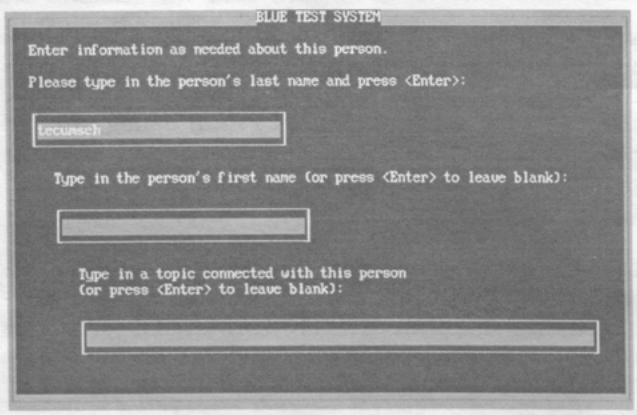


Figure 8
User entry of personal-name query for "tecumseh"

prompts might retrieve additional records through keyword searches and alphabetical searches. Because topical elements are discarded in alphabetical searches, systems are very much likely to retrieve titles that are not specifically on users' topics of interest. In our example, users might retrieve titles on Jefferson's presidency that do not refer to his views on slavery.

When users entered queries bearing name elements only, the experimental online catalog submitted queries to alphabetical searches. In figure 8, the user has entered the name "tecumseh" followed by blank dialog boxes. The system performs an alphabetical search and responds with personal-name subject headings in the alphabetical neighborhood of the entered term (figure 9). The one listed subject heading that refers to "Tecumseh" is the one selected by the user (figure 10). Figure 11 shows the first of ten retrieved bibliographic records.

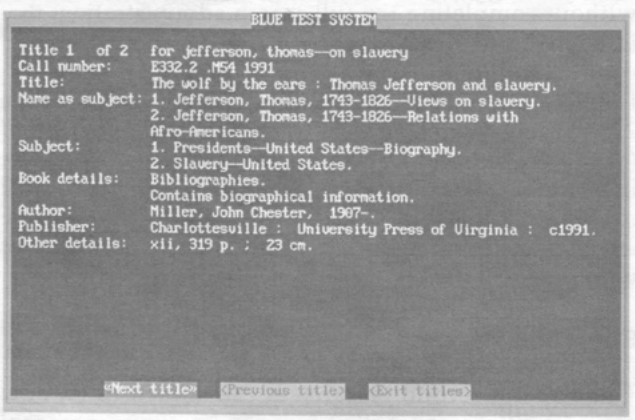


Figure 7
Bibliographic record retrieved in keyword search

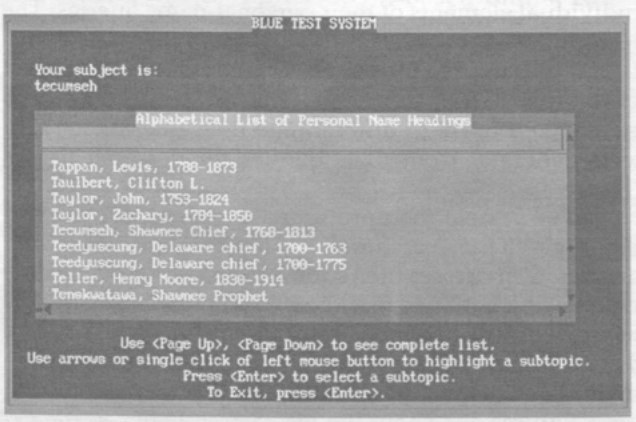


Figure 9
Alphabetical list of personal-name subject headings in response to "tecumseh" query

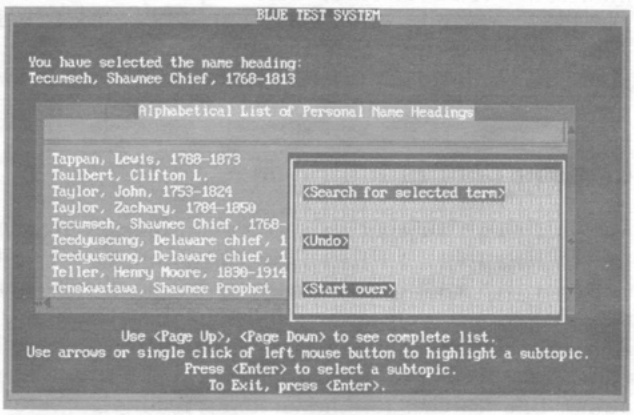


Figure 10
User selection of a personal-name subject heading naming "Tecumseh"

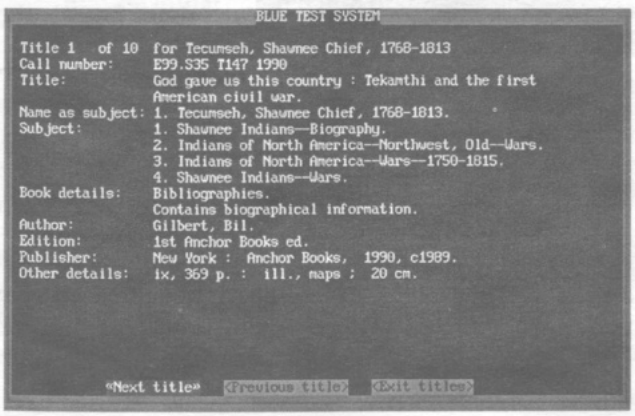


Figure 11
Bibliographic record bearing the subject heading "Tecumseh"

Experimental Online Catalog Performance in Searches for Personal Names

The third research question that the analyses in this paper address focuses on how users respond to an experimental online catalog that prompts them for the various elements of the personal-name queries. This section describes how users entered their personal-name queries into ASTUTE and includes a failure analysis of troublesome searches.

End users entered subject searches for personal names into the experimental online catalogs at UM-Dearborn and Earlham College. The database for the former was composed of library materials in computer science and technology. Few personal names were assigned to UM-D bibliographic records. In fact, the entire alphabetical index contained less than 150 personal-name headings. Of the 826 total comparison search experiments that ASTUTE administered at UM-D, 10 percent involved a personal name.

The database for Earlham College was composed of library materials in American history. Many names were assigned to bibliographic records. The alphabetical index of personal-name subject headings contained almost as many unique names as the alphabetical index of topical subject headings. Of the total 238 Comparison Search Experiments that ASTUTE administered at Earlham, almost 30 percent involved a personal name.

Unusable Queries for Personal Names at UM-D

Despite the infrequency with which personal names were connected with computer science and engineering topics in the UM-D database, some UM-D users responded positively to the system's question about personal names and entered queries bearing personal names. Examples were:

- limbrough [sic] (surname), rush (given name)
- pasteur (surname), louis (given name)
- regan [sic] (surname)
- corniliustacitus [sic] (surname)
- kevorkian (surname), physician assisted suicide (topic)
- perot (surname), ross (given name), petroleum (topic)
- jakeson [sic] (surname), michael (given name), singing muise [sic] (topic)
- ogata (surname), control theory (topic or title)

These eight queries were typical of the personal-name queries for subjects that users entered into the experimental online catalog at UM-D. Most queries were out of scope (first six queries listed above). A few queries were characterized by playing around (seventh-listed query above) or author-title searches (last-listed query above).

Two subject queries for personal names that UM-D users entered into the catalogs deserve mention—"pascal" and "gauss." The former is a name of a computer programming language or a unit of pressure. The latter is an electromagnetic unit of magnetic induction. Both queries were valid *subject* queries. The systems submitted both queries to subject searching approaches for personal names because users responded positively to the system question about personal names being involved with queries and typed these queries into the pop-up boxes that ASTUTE provided for them. In fact, both queries did involve names of persons, Blaise Pascal,

a seventeenth-century French scientist, and Johann Karl Gauss, a nineteenth-century German mathematician.

The question about names in queries that users answered prior to entering their queries was quite effective in categorizing subject queries as personal names or subject queries generally. In view of the two queries discussed above (i.e., "pascal" and "gauss"), it was not entirely effective in effecting a correct categorization. Future computer systems that separate subject searches for personal names or for subject queries generally could experiment with different wording for the question that categorizes subject queries as personal names or subject queries generally.

A handful of UM-D users answered negatively to the question about names in queries. They entered personal names (and sometimes title words) into the catalogs using the system's general subject searching capability. Examples were:

- assad
- hitler
- bush, g
- sinclair upton
- schaum's
- schaum's outline series dynamics
- dijkstra

Some names were out of scope (first four names listed above). Other names were probably attempts at author or author-title searches in the areas of computer science and engineering (last three names listed above). The system failed to yield records for all but queries bearing the word "schaum's" because words and phrases from personal-name subject headings were not processed into searchable indexes of the system's general subject searching capability. They did retrieve records in title-keyword and general keyword searches for queries bearing the word "schaum's" because title words were processed into the searchable indexes of these two search approaches.

Why did users answer negatively to the question about personal names that resulted in the system's submission of these names to subject searching capabilities for subjects generally? In the absence of postsearch interview data, we can only speculate on such reasons. Some users could have changed their minds and entered different queries than the ones they originally intended to enter. Other users could have misunderstood the question about personal names. Still other users could have failed to read the question about personal names.

Future computer systems that separate subject searches for personal names from general subject queries should experiment with different wording to the question that asks users about names in their queries because their responses to this question are critical to information retrieval. Also, users who become frequent users of a particular online catalog could become bored with answering the same question repeatedly. They might need a different method of submitting queries for personal names than infrequent users of the system.

Unusable Queries at Earlham

Personal-name subject headings were quite common in the bibliographic database on American history at Earlham. Yet findings about the entry of queries for personal names were quite similar to those at UM-D. End users at Earlham entered queries for personal names that were out of scope. Some could have been author-title searches for known items. Examples were:

- calvin (surname), john (given name), treatises against the anabaptists (topic)
- chopin (surname), kate (given name)
- poe (surname), edgar (given name)
- schlesinger (surname), arthur (given name), the age of jackson (topic or title)
- arendt (surname), hannah (given name), totalitarianism (topic)
- webster (surname), dictionary (topic or title)

On occasion, end users at Earlham answered negatively to the system's question about personal names in queries. Consequently, the system submitted their queries to subject searching capabilities for subjects generally. Below are examples of out-of-scope queries bearing personal names that were submitted to such capabilities:

- tanner, henry
- marilyn monroe
- rodin

Elements in Usable Queries at Earlham

At Earlham, users entered 27 different queries for personal names that matched the subject contents of the experimental online catalog's database. Some queries contained topical elements:

- jackson (surname), andrew (first name), jacksonian democracy (topic)
- calhoun (surname), john (given name), state rights (topic)
- sumner (surname), charles (given name), reconstruction (topic)

Other queries contained name element(s) only:

- roosevelt (surname), eleanor (given name)
- rustin (surname), bayard (given name)
- fremont (surname)

Of the 27 queries for personal names, a few deserve

special mention because of the ways in which users stated name and topic elements. Both the Blue and Pinstripe systems prompted users for surname, given name, and topic elements. Although they did not prompt users for middle-name elements, four queries contained two elements in the given name pop-up box that represented first- and middle-name elements. These queries were:

- king (surname), martin luther (first and middle names entered into first-name pop-up box), about the bycotts [sic] they had way back in mississippi (topic)
- king (surname), coretta scott (first and middle names entered into first-name pop-up box), learn about what happened in 1927 (topic)
- king (surname), martin luther (first and middle names entered into first-name pop-up box), about in the 1925 (topic)
- washington (surname), booker t (first name and middle initial entered into first-name pop-up box), african american history (topic)

The first three queries were entered successively into the experimental online catalog. They were probably entered by the same individual. In view of our experience of the four queries above, we could speculate that some users may feel compelled to enter the middle name or initial for famous people who are known by their full names. Yet in the following query for an author who is known by her full name (i.e., Harriet Beecher Stowe), the user did not enter the middle name:

- stowe (surname), harriet (first name), uncle tom's cabin (topic)

Two other queries were especially interesting:

- hoover herbert (surname and given name entered into surname pop-up box), herbert (first name), economic (topic)
- robert lafollette (given name and surname entered into surname pop-up box), mcarthy (topic)

With regard to the first query, the user entered both surname and given name in inverted form into the surname pop-up box. He repeated the entry of the given name in the given name pop-up box. Perhaps he realized that he had entered both first and last names into the surname pop-up box after the given-name pop-up box prompted him to enter the given name. When the user was ready to submit the entire query to the system, the system gave the user the opportunity to change entered name and topic elements, but he did not make changes. The Pinstripe System performed a keyword-in-record search for name and topic elements in this query and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements, but it continued searching by placing the user at the personal-name subject heading in an alphabetical list of personal-name subject headings beginning with the name "Hoover, Herbert, 1874–1964." Unfortunately, the success of this Blue System search was marred by user perseverance.

In the second query, the user gave the name in direct form. The Pinstripe System performed a keyword-in-record search for name and topic elements and was unsuccessful in retrieving titles. The Blue System was also unable to retrieve titles for name and topic elements, but it continued searching by displaying to the user an alphabetical list of personal-name subject headings that began with the letter *r*. The user browsed backward repeatedly and found the subject heading "LaFollette, Robert Marion, 1855–1925." She selected this subject heading, displayed eight titles, and gave two titles "very useful" ratings.

ASTUTE was not unique in its division of subject searches into subject searches for personal names and for subjects generally. UCLA's ORION online catalog also enlists such a division. ASTUTE was unique in its prompting of users for the name and topic elements of personal-name queries. Additional research is needed to determine how users enter a wide range of names—especially names in which the distinctions among surname, given-name, and middle-name elements are not as clear as the names entered by Earlham College users, for example, foreign names (Vincent Van Gogh, Jan van Eyck, Rogier van der Weyden) Native American names (Tecumseh, Blue Jacket, Geronimo), and Classical names (Hippocrates, Aristotle)—and to determine whether systems can provide satisfactory results.

Failure Analysis

Table 5 summarizes the subject searching capabilities that handled subject queries for personal names in the Blue and Pinstripe systems and characterizes them as successful or unsuccessful.

Users entered a total of 27 personal-name queries. The reason the numbers of searches in table 5 do not total to 27 in the Blue System and 27 in the Pinstripe System is connected with the extent to which users completed search administrations. Some users seached a single system (Blue or Pinstripe) and terminated their use of the experimental online catalog before conducting a search for the same topic in the other catalog. Other users conducted searches of both the Blue and Pinstripe Systems for their personal names of interest.

Between one-third and one-half of subject searches for personal names were successful in Earlham's Blue

Table 5
Subject Searches for Personal Names

Reasons for success/failure	Ealrham Blue (%)	Earlham Pinstripe (%)
Successful searches		
Keyword-in-record	4	8
Alphabetical	8	N/A
Subtotal	12 (52.2%)	8 (38.1%)
Unsuccessful searches		
Perseverance	6	0
Database failure	4	4
Search approach failures	0	9
Query specificity	1	0
Subtotal	11 (47.8%)	13 (61.9%)
Total	23 (100%)	21 (100%)

and Pinstripe Systems, respectively. The most frequent reasons for failure with these systems were perseverance and search-approach failures, respectively. Both systems had their share of database failures.

Of the 27 subject queries for personal names, all but 4 queries featured elements for both names and topics. We expected a large number of personal-name queries to have both elements for names and topics because the experimental online catalogs explicitly prompted users to enter a topical element. Without such prompting, queries consisting of personal-name elements usually outnumber multi-element queries.

Keyword searches failed to retrieve titles for half of these queries because of the topic element. Frequently, this was categorized as a search-approach failure for the Pinstripe System. For example, a user entered the query "sumner" (surname), "charles" (first name), "reconstruction" (topic) into the experimental online catalog. The Pinstripe System submitted this query to a keyword search and failed to retrieve titles. The Blue System submitted this query to keyword searches. When it failed to retrieve titles, the system ignored the topic element and responded with the results of the alphabetical approach. The user selected the personal-name subject heading "Sumner, Charles, 1811–1874" from the alphabetical list, retrieved eight titles, and rated all eight titles "very useful."

The failure of nine subject queries was attributed to search-approach failure in the Pinstripe System. In these nine searches, the Pinstripe System submitted the entire query to keyword searches and failed to yield retrievals. When users also conducted Blue System searches for the same queries, these searches might have succeeded because this system discarded topical elements and presented users with lists of subject headings in the alphabetical neighborhood of personal-name query elements where they found a subject heading naming the person of interest. Selection of a listed subject heading resulted in one or more retrievals that users scanned to determine whether retrieved titles covered the specific topics they had in mind.

Several Blue System searches were successful in retrieving titles through the alphabetical approach to which users gave positive relevance assessments even though the titles did not specifically refer to the topic element of their queries. The failure of a few Blue System searches to retrieve useful titles was attributed to user perseverance, query specificity, or database failure. For example, after the query "whittier" (surname), "john" (given name), and "abolition" (topic) failed to retrieve titles in a keyword search, the Blue System responded with the alphabetical approach, but the alphabetical list of personal-name subject headings did not contain a subject heading for John Whittier. The failure of this Blue System search was attributed to database failure.

On occasion, users were not satisfied with the results of Blue System searches in which the alphabetical approach ignored the topical element of their queries and retrieved titles based only on the name elements of their queries. For example, in response to the query "king" (surname), "coretta scott" (first name), "learn about what happened in 1927" (topic), the Blue System ignored the topical element because of its inability to retrieve titles through a keyword-in-record search and responded with an alphabetical list of personal-name subject headings. The user selected the personal-name heading for Coretta Scott King, retrieved several titles, and gave them all negative relevance assessments. Perhaps this user would have only been satisfied with titles that specifically referred to Mrs. King in 1927.

Discussion

This paper provided data analyses to answer three research questions. The first question sought an answer to the following: how can online systems choose search approaches on their own that are likely to produce useful retrievals in response to user-entered queries for personal names? Search trees hold much promise for assuming the burden of determining the subject-searching approach likely to produce useful information for user queries. We tested search trees that were the result of an empirical study of the subject terms users entered into

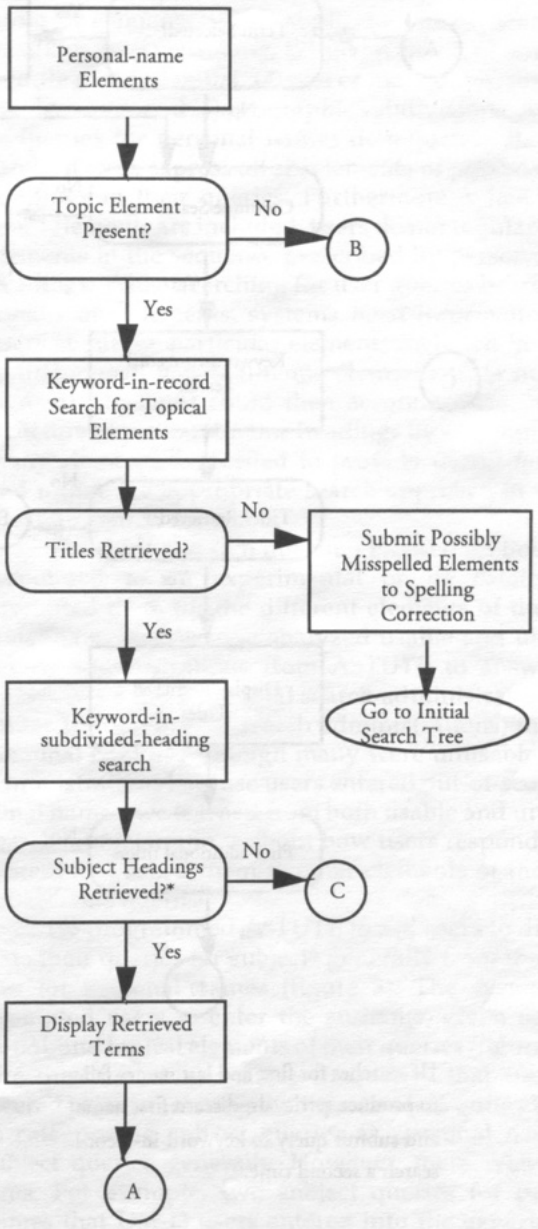
online catalogs. We extracted personal-name queries from the transaction logs of four different online catalogs and submitted them to a manual analysis that simulated how search trees would treat the queries during the retrieval process. The vast majority (87.2 percent) of extracted queries bore personal-name elements only and thus were given to a search tree that featured alphabetical searches for these elements. The small percentage of queries bearing both personal-name and topical elements were first submitted to various keyword searches in attempts to produce retrievals for both element types; failure to produce retrievals in keyword searches forced us to discard topical elements and submit the remaining personal-name elements of queries to alphabetical searches.

A handful of personal-name queries were keyword matches. These matches resulted in the retrieval of single subject headings or bibliographic records bearing personal-name and topical elements of personal-name queries. Most personal-name queries, regardless of the presence of topical names, were submitted to alphabetical searches. Since alphabetical searches only considered name element(s), users whose queries contained topical elements did not necessarily retrieve titles on the specific topics cited in their queries.

The majority of personal-name queries submitted to alphabetical searches led users to personal-name subject headings on their topics of interest. Some users had to browse alphabetical lists to find the desired subject headings. Such browsing ranged from finding the desired subject heading on the screen immediately preceding or following an initial screen of a dozen or fewer subject headings for personal names to browsing dozens of screens of headings to find the desired one. Collection failure also occurred.

Based on our analyses of user queries in two separate but related studies, we recommend subtle but important changes to the original search tree for personal-name queries for subjects to include spelling correction for topical elements of user queries and relevance feedback. We are using the enhanced search trees in figures 12A and 12B to show the changes we recommend.

The search tree in figure 12A submits personal-name queries bearing topical elements to keyword searches. Keyword searches include system-assisted spelling correction on topical elements that fail to produce retrievals. In the process of correcting spelling, users might change their queries entirely. Thus, the personal-name search tree submits "corrected" queries to the initial search tree to redirect "corrected" queries to the appropriate search trees just in case such queries no longer bear personal-name elements. If the two types of keyword searches fail to yield retrievals for personal-name and topical elements of user queries, the search trees discard given- and middle-name elements and



*If searches for first and last names fail to produce retrievals, discard first name and submit query to keyword-in-subdivided heading search a second time.

Figure 12A
Enhanced search tree for personal-name subject queries featuring keyword searches and spelling correction

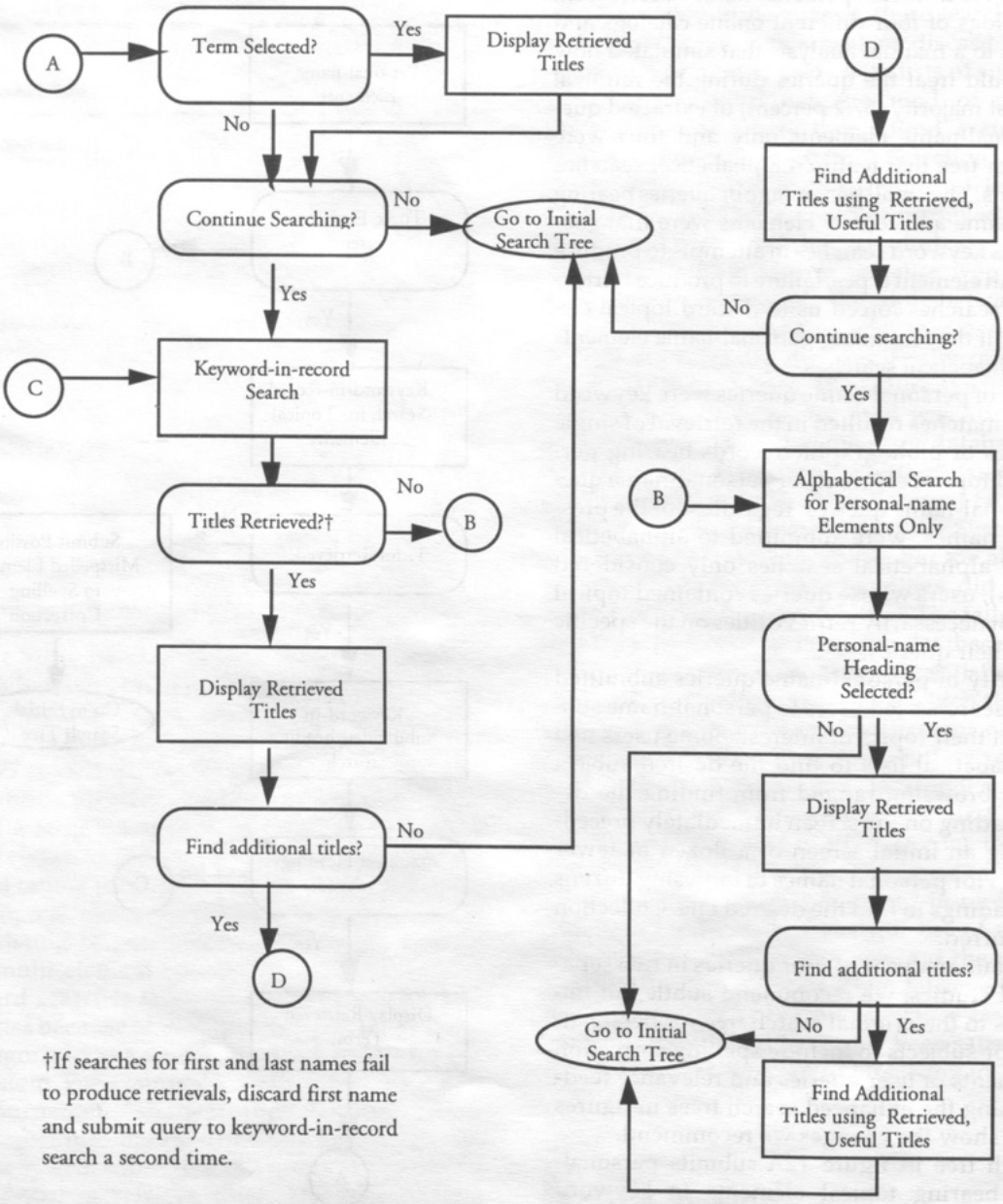


Figure 12B
Enhanced search tree for personal-name subject queries featuring relevance feedback and alphabetical searches

resubmit queries to these keyword searches (figures 12A and 12B). Failure to produce retrievals through keyword searches results in the submission of personal-name elements to alphabetical searches. Alphabetical searches also handle personal-name queries bearing personal-name elements only (figure 12B).

Systems should follow up the results of keyword and alphabetical searches with relevance feedback (figure 12B). They should ask users to rate the usefulness of retrieved titles and use such ratings to identify common subject headings in retrieved, relevant titles to retrieve additional ones with the same subject headings.

The tactics the enhanced search tree for personal-name queries uses to produce retrievals could be extended to author searches, title searches, and combined author-title searches. When users select the author search type, systems would ask users whether the author is a person or organization. In the case of persons, systems would prompt users for the last, first, and middle names and submit queries to keyword-in-heading searches. Failure to produce retrievals would provoke systems to submit queries to alphabetical searches that enable users to browse listed headings in the same neighborhood as the entered names and find author headings that begin with the same or similar names. In the case of organizations, systems would prompt users for the organization name and submit queries to keyword-in-heading searches in an attempt to find corporate- or meeting-name headings bearing words in the user query. If they fail to produce retrievals, they could produce the results of an alphabetical search and show a list of corporate-name headings in the alphabetical neighborhood of the user-entered name. Spelling correction could be included prior to keyword-in-heading searches; if keyword-in-record searches for individual words in organization names fail to yield retrievals, systems could display these words to users and ask them to check their spelling.

The search tree for title searches would be quite similar to that for organization names. When users select the combined author-title search type, systems ask users whether the desired author is a person or organization. In the case of persons, systems prompt users for last, first, and middle names of their desired author, and title words. They submit name elements to keyword-in-heading searches that retrieve author names and title elements to title-keyword searches that retrieve titles, and combine the results to find names and titles in the same record. When systems fail to produce retrievals, they could reduce the number of searchable elements, for example, searching for last names and for low-posted title words, and conduct keyword-in-heading and title-keyword searches for these elements. As a last resort, systems could ignore author-name elements and show titles in the alphabetical neighborhood of user-entered title words, or they could ignore title-word elements and show authors in the alphabetical neighborhood of user-entered name elements through the alphabetical approach.

The second research question addressed how online systems should solicit queries involving personal names from users. The arguments, data, and analyses in this paper suggest that the content and structure of assigned headings for personal names must be considered when determining how to improve the performance of user queries involving personal names. The personal-name headings that catalogers formulate typically result in a predictable order of elements. (This is not true for the formulation of subject headings generally.) The order of elements, when applicable, in personal-name headings is (1) surname, (2) given name or initial, (3) middle name or initial, (4) one or more dates, (5) topical subdivisions, and (6) geographic subdivisions. Analyses of queries for personal names demonstrate that users rarely, if ever, express all six elements of personal-name headings in their queries. Furthermore, when two or more elements are included, users do not regularly enter elements in the sequence prescribed by personal-name headings. When searching for user queries bearing personal-name elements, systems need information from users about the particular elements included in queries to distinguish personal-name elements from nonname elements. Systems could then accommodate the rigid structure of personal-name headings by searching for as many elements as needed to provide useful retrievals and utilize the appropriate search approach in view of the information provided.

The third research question focused on how users responded to an experimental online catalog that prompted them for the different elements of their personal-name queries. We analyzed usable and unusable search administrations from ASTUTE to answer this question. Of the 1,064 total search administrations, just under 15 percent (153 search administrations) involved personal names. Although many were unusable search administrations because users entered out-of-scope personal names, we learned from both usable and unusable search administrations about how users responded to a system that asked them to enter elements of their personal-name queries.

We programmed ASTUTE to ask users to differentiate their queries for subjects generally from their queries for personal names (figure 3). The system then prompted users to enter the surname, given name or initial, and topical elements of their queries (figures 4–7). The question about names in queries that users answered prior to entering their queries was quite effective in categorizing subject queries as personal names or subject queries generally. However, there were problems. For example, two subject queries for personal names that UM-D users entered into the experimental online catalogs contained the words "pascal" and "gauss." The former is a name of a computer programming language or a unit of pressure. The latter is an electromagnetic unit of magnetic induction. Both queries were valid *subject* queries. The systems submitted both queries to subject searching approaches for personal names because users responded positively to the system question about personal names being involved with queries. In fact, both queries did involve names of persons, Blaise Pascal, a seventeenth-century French

scientist, and Johann Karl Gauss, a nineteenth-century German mathematician.

In view of the two queries discussed above (i.e., "pascal" and "gauss"), it can be seen that the systems were not entirely effective in effecting a correct categorization. Future computer systems that separate subject searches for personal names or for subject queries generally could experiment with different wording for the question that categorizes subject queries as personal names or subject queries generally. Another possibility is to process subject headings bearing proper nouns and adjectives into subject indexes for personal names. Other candidates for double posting in general subject and personal-name subject indexes are subject headings that name fictional, mythological, or legendary persons.

In the manual analysis of user queries for personal names, we used our judgment in reordering the name elements of direct forms of personal-name queries to reflect the order of name elements in assigned subject headings for personal names because the four online catalogs that contributed user queries for personal names did not prompt users for the individual elements of their queries. A small percentage of personal-name user queries contained both name and nonname elements. In contrast, ASTUTE explicitly prompted users for name and nonname elements. Of the 27 valid subject queries for personal names, all but four queries featured elements for both names and topics. We should expect systems that prompt users for query elements in subject and known-item searches to generate a large number of personal-name queries having both elements for names and topics or names and titles. Without such prompting, queries consisting of a single element will usually outnumber multi-element queries.

We need to test the wording of prompts systems use to solicit queries and to differentiate between different types of queries and search types. Such tests should use several methods—user and system performance in quasicontrolled tests using predetermined search tasks, analysis of logged searches from publicly accessible online catalog terminals, and protocol analysis in system walk-throughs with end users.

Conclusion

This paper examined the performance of user queries that involved personal names and recommended improvements to the basic system approach to soliciting personal-name queries and searching for them. It used data and analyses from two related research projects to increase understanding of user queries for personal names by answering the following three research questions:

- How can online systems choose search approaches on their own that are likely to produce useful retrievals in response to user-entered queries for personal names?
- How should online systems solicit queries involving personal names from users?
- How do users respond to an experimental online catalog that prompts system users for the different elements of their personal-name queries?

Search trees hold much promise for assuming the burden of determining the subject-searching approach likely to produce useful information for user queries. Search trees are a key component of a new design for subject access to online catalogs that also includes a wide range of subject-searching functionality to respond to the wide variety of user queries for subjects. We programmed the ASTUTE experimental online catalog with responses that were controlled by search trees and tested its effectiveness handling user queries for subjects. Failure analyses of end-user searches led us to conclude that online catalogs governed by search trees were more effective in selecting a subject-searching approach that would produce useful information for the personal names users seek than users would select on their own.

The new subject searching design enlists an entirely different approach to soliciting user queries for personal names that takes into consideration the content and structure of personal-name subject headings. We programmed ASTUTE to respond with this entirely different approach and evaluated its effectiveness in online retrieval tests. Although we concluded that this approach was basically sound, there were still some troublesome queries—for example, queries that involved names for fictional, legendary, and mythological characters and proper nouns and adjectives, and queries that involved personal names that would be better served through general subject searches because the library catalog's controlled vocabulary used topical subject headings to represent these subjects and thus the search trees directed user queries to the incorrect type of search. We made recommendations for double posting of subject headings in name and general subject indexes; unfortunately, solving the problem of proper nouns and adjectives in subject headings may not be easy because systems cannot automatically identify proper nouns and adjectives in subject headings without explicit coding that identifies words as proper nouns or adjectives.

We presented an enhanced search tree for handling personal-name queries for subjects and generalized findings about subject searching for names in known-item searching. Search trees for author searches, title searches, and combined author-title searches can enlist the same search approaches as subject searching. When

searches result in few or no retrievals, systems controlled by search trees can enlist more than one search approach to find retrievals instead of placing the burden of determining which search approach produces useful retrievals on end users. While search trees are effective at reducing search approach failures, system prompts require users to correctly differentiate between different types of queries, choose the right search types, and accurately identify search elements in their queries. We also need to test the wording of prompts using several different methods to ensure that systems are collecting the desired information because user responses to such are critical to the success of their searches.

References

1. Joseph R. Matthews and others, *Using Online Catalogs: A Nationwide Survey* (New York: Neal-Schuman, 1983).
2. Thomas A. Peters, "When Smart People Fail: An Analysis of the Transaction Log of an Online Public Access Catalog," *Journal of Academic Librarianship* 15, no. 5 (Nov. 1989): 270.
3. Rhonda N. Hunter, "Successes and Failures of Patrons Searching the Online Catalog at a Large Academic Library: A Transaction Log Analysis," *RQ* 30, no. 3 (Spring 1991): 400.
4. Steven D. Zink, "Monitoring User Search Success through Transaction Log Analysis: The WolfPAC Example," *Reference Services Review* 19, no. 1 (Spring 1991): 53.
5. Karen M. Drabenstott and Diane Vizine-Goetz, *Using Subject Headings for Online Retrieval: Theory, Practice, and Potential* (San Diego: Academic, 1994), 159–61.
6. Jean Dickson, "An Analysis of User Errors in Searching an Online Catalog," *Cataloging & Classification Quarterly* 4, no. 3 (Spring 1984): 32.
7. Arlene G. Taylor, "Authority Files in Online Catalogs: An Investigation of their Value," *Cataloging & Classification Quarterly* 4, no. 3 (Spring 1984): 8.
8. Margaret Henty, "The User at the Online Catalogue: A Record of Unsuccessful Keyword Searches," *LASIE* 17, no. 2 (Sept./Oct. 1986): 51.
9. Peters, "When Smart People Fail," 270.
10. Hunter, "Successes and Failures of Patrons Searching," 400.
11. Zink, "Monitoring User Search Success," 52.
12. David Holmes and Derrick Bulger, "A Day in the Life of a Public Terminal—A Transaction Analysis of an Online Catalogue Terminal in a Bilingual Environment," *Canadian Journal for Information Science* 13, nos. 3/4 (1988): 27.
13. Terry Ballard and Jim Smith, "The Human Interface: An Ongoing Study of OPAC Usage at Adelphi University," in *Advances in Online Public Access Catalogs*, ed. Marsha Ra (Westport, Conn.: Meckler, 1992), 65, 67.
14. Drabenstott and Vizine-Goetz, *Using Subject Headings for Online Retrieval*, 224.
15. Taylor, "Authority Files in Online Catalogs," 8.
16. Marilyn Ann Lester, "Coincidence of User Vocabulary and Library of Congress Subject Headings: Experiments to Improve Subject Access in Academic Library Online Catalogs" (Ph.D. diss., Univ. of Ill., 1989), 186.
17. Drabenstott and Vizine-Goetz, *Using Subject Headings for Online Retrieval*, 227–29.
18. Dickson, "An Analysis of User Errors," 32.
19. Dennis R. Walter, "The User at the Online Catalogue: A Record of Unsuccessful Keyword Searches—Another Case Study," *LASIE* 18, no. 3 (Nov./Dec. 1987): 80.
20. Peters, "When Smart People Fail," 270.
21. Hunter, "Successes and Failures of Patrons Searching," 400.
22. Dickson, "An Analysis of User Errors."
23. Taylor, "Authority Files in Online Catalogs."
24. Drabenstott and Vizine-Goetz, *Using Subject Headings for Online Retrieval*, 224.
25. Karen M. Drabenstott, *Enhancing a New Design for Subject Access to Online Catalogs* (Ann Arbor, Mich.: School of Information and Library Studies, University of Michigan, 1994).
26. Karen M. Drabenstott and Marjorie S. Weller, *Testing a New Design for Subject Access to Online Catalogs* (Ann Arbor, Mich.: School of Information and Library Studies, University of Michigan, 1995).
27. For further analysis of the data collected in these two projects see Karen M. Drabenstott and Marjorie S. Weller, "Handling Spelling Errors in Online Catalog Searches," *Library Resources & Technical Services* 40, no. 2 (Apr. 1996), 113–32.
28. Nathalie Mitev, Gillian Venner, and Stephen Walker, *Designing an Online Public Access Catalog*, Library and Information Research Report (London: British Library), 94.
29. Charles R. Hildreth, *Intelligent Interfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems* (Washington, D.C.: Library of Congress), 85–88.
30. Drabenstott and Vizine-Goetz, *Using Subject Headings for Online Retrieval*.
31. Ibid., 301–29.
32. Drabenstott and Weller, *Testing a New Design*, 130–39, 143–48.

Training for Automated Systems in Libraries

Julie Hallmark and
C. Rebecca Garcia

In a series of interviews, the automation administrators of forty-nine libraries that had recently installed automated systems discussed staff and user training for their new systems. The interviews focused on training objectives and procedures, timing and effectiveness of the training, and problems encountered. Eight vendors then presented their points of view as they responded to issues raised by the librarians. The candid comments from both groups on the difficulties they confronted, and their advice for ensuring successful training, should prove valuable to others about to institute training.

The installation of a new automated system in a library may cause feelings of stress and confusion among staff members. Such feelings arise from problems with the timing of the installation, difficulties in communication with those affected, the uncertainties of change, and the exigencies of training. The authors' past research on migration to new automated systems revealed some especially vexing training problems and motivated the present study.[1] Many librarians who participated in the earlier investigation felt very strongly, either positively or negatively, about their training experiences. Their concerns are mirrored in much recent literature, which features extensive discussions about the quality, cost, scheduling and duration of training, as well as about the role of vendors in training.[2]

The present study concentrated on training from both librarians' and vendors' points of view. School, academic, special, and public librarians participated, generously sharing their experiences and suggestions for improvement. Trainers employed by major vendors also contributed their opinions. The authors hope that the investigation will improve communication between librarians and vendors, and clarify issues and tradeoffs in the training process.

We begin with a description of our methodology and an overview of training procedures for library staff and users. We then enumerate problems with training that participants encountered and pass along their advice for those undertaking installation of automated systems. Finally, we present comments from vendors who participated in the project.

Methodology

The authors contacted ten vendors, explained the purpose of the study, and asked the vendors to identify customers who had recently installed their systems. Eight vendors agreed to participate. After each vendor submitted the names of system administrators at some six or seven sites (fifty-four in all), the investigators spoke with these individuals to explain the project and request that they join in the study. A total of forty-nine libraries and consortia made up the final group, as five did not meet the criteria of having completed both installation and training. The breakdown of participating libraries was as follows:

Public	31 percent
School	20 percent
Academic	29 percent
Special	9 percent
Consortia	11 percent

Sixteen of the libraries had migrated from other systems.

The interview schedule for the librarians included questions in these areas:

- Training objectives and the number of staff trained;
- Methodology (when, who, how, where);
- Effectiveness of the training;
- Problems encountered; and
- Training library users.

A copy of the schedule appears in appendix A. In lengthy telephone interviews participants described their rationale for training staff and users as well as their experiences, both good and bad. Freely offering advice for colleagues, system administrators forthrightly discussed their mistakes and how they would do things differently if they were to repeat the process.

In analyzing the reactions of customers to vendor training, the investigators isolated nine areas librarians found particularly problematic and over which vendors could reasonably be expected to exercise some control. (Other problems surfaced but were internal to the particular library.) We enumerate these areas in the section entitled "Problems and Pitfalls." In telephone interviews, vendor trainers responded to these problems and complaints and explained the vendors' points of view.

Staff Training

Depending on the size of the library, between one and 256 people required training. Except in libraries that needed training for no more than twelve people, the vendors trained selected staff members who in turn

Julie Hallmark is Professor at the Graduate School of Library and Information Science, University of Texas at Austin. **C. Rebecca Garcia** is Product Manager for Ameritech Library Services.

trained others. Staff members selected for the vendor training typically constituted an "implementation team" made up of a representative from each department plus the library director.

Everyone was trained at the library site except for one vendor's customers, who received instruction at the vendor's facilities. Typically, a larger number of staff needed training in circulation and public access; a smaller number needed instruction in more specialized modules such as acquisitions or cataloging. The majority of school libraries had only one professional and that individual received training in all software modules purchased.

Library/vendor communication gaps occurred when the library could not or did not provide advance instructions to the vendor concerning the level and type of training desired. This situation may have arisen from the fact that very few libraries establish specific training objectives. Only two libraries in the present study did so, both of which were migrating libraries. (The ongoing transition in migrating libraries often imposes unique constraints and time requirements that may encourage the setting of objectives.) Other libraries simply hoped to get everyone trained or to ensure that everyone felt comfortable with the system and accepted it.

Pretraining techniques such as test databases or training newsletters added to the success of the training. Some vendors encouraged librarians to read training manuals and system documentation in advance. In some cases, the system administrator rewrote training and documentation materials in order to ensure clarity and successful training.

One or two days of training for each module was typical; however, training occasionally lasted as long as four or five days. An individual staff member ordinarily received less than one day of training, an amount in agreement with earlier findings.[3] Of course, libraries always have the option of requesting that the vendor return to conduct further training after the staff has had a chance to use the system in its everyday work. In the learning stages, telephone support was critical, and the ability to contact the vendor easily was very important to success in using the system. Having personal access to a local vendor representative was also very helpful. Several participants described how much they appreciated their vendor's personal, understanding approach, which made them feel like one of the "family."

The skill and effectiveness of trainers varied greatly, even within the same company. In one case the trainer had helped design the system and was very knowledgeable. Others seemed to know little about the modules they were giving instruction in, or about the type of library being trained. If the vendor was geared to academic libraries, for example, the trainer may not have understood problems and terminology unique to public libraries such as the distinction between "hold" and "reserve." Some libraries later discovered in day-to-day operation of the system that key parts of modules had not been covered.

Some migrating libraries chose to create flip charts and handouts to facilitate comparisons of new system operations to old. These libraries found that the results justified the time expended. On the other hand, one respondent related how their implementation committee had discussed the merits of comparing the old with the new and finally said, "to hell with it"—staff and library users could just learn the new system rather than constantly making comparisons.

During the rush of the implementation phase, many libraries failed to revise documentation and training materials so they would be specially suited for their users. Even many large organizations that used staff trainers provided only the vendor's handouts and documentation. Other libraries revised vendor material, convinced that such rewriting of documentation was as much a part of the training process as the hands-on sessions. Whatever the approach, librarians stressed the importance of learning to use the documentation during training.

Ongoing Staff Training

What happens after the trainers go home? Most libraries offered ongoing training, local users' group meetings, or the chance to attend refresher courses taught by the vendor. During local user group meetings, attendees shared handouts, tips on using the system, and new aspects they had encountered. School libraries frequently took advantage of regional vendor courses during the summer. These courses focused on "how to use the system more effectively" rather than on features of a new release or upgrade tape. As one librarian pointed out, "The training never stops."

Training Library Users

In this section we define "training" as a process more formal than ad hoc one-on-one assistance by the library staff or the provision of written instructions next to the online public access catalog (OPAC) stations. User training varied enormously, ranging from required attendance at small classes to providing a video for campus departments to screen.

Of the institutions we studied, 93 percent of learning resource centers provided formal training for users,

compared with 54 percent of academic libraries, 44 percent of public libraries, and 33 percent of special libraries and consortia.

In learning resource centers the librarian typically offered training to English classes or in conjunction with specific projects. Starting at the second-grade level, most librarians emphasized browse/alphabetical searching and advanced to keyword and Boolean concepts in later years. Study participants suggested that students and other users grasped concepts best when they were able to put them to immediate use. Librarians in learning resource centers reported the same enthusiasm and motivation among students as that described by Adams.[4] Training teachers, however, proved to be more challenging. One librarian, who called students competent and enthusiastic, described faculty as resistant, uninterested in Boolean search strategy, and still dependent on the card catalog. Schools that offered their faculty formalized one-on-one training—which included information on accessing the system from their home or office—were most successful.

Training offered by academic libraries was usually in the form of classes associated with their ongoing bibliographic instruction (BI) program. Special OPAC sessions offered to freshman English classes and others had high attendance when required; when voluntary, "often no one came." Faculty tended to be apathetic about training, and few took advantage of classes.

Instead of offering extensive training, a majority of public libraries relied on the user-friendliness of their system, as well as the proximity of the reference desk to the OPAC stations. One participant mentioned his library's "roving volunteers"—Friends of the Library members who wandered through the OPAC stations during the busiest periods to assist users.

Relatively few special librarians trained their users. Special libraries usually serve a smaller client base than other types of libraries, and a more informal approach seems desirable. Special librarians interviewed were generally of the opinion that one-on-one assistance was adequate and probably preferable.

Although some libraries developed training materials for their users, the majority depended on the ease of using straightforward, menu-driven systems. Procrastination, along with no specific training plan, often resulted in libraries' not having produced the training materials by opening day. Many libraries said they were still planning on developing these materials months after the system was installed and in use.

Most librarians made no formal attempt to evaluate training effectiveness; rather, they offered intuitive observations. A couple of participants suggested that users' continuing to access the card catalog indicates less-than-successful OPAC training. One librarian reported:

> Effectiveness was hard to determine. People *use* the system, but are they getting what they want? . . . We looked at our transaction logs and found about one-third zero hits, but that might be about par for the course.

In summary, several factors affected the popularity and extent of user training:

- The type of library;
- Whether other libraries in the vicinity had the same automated system, in which case many users would already be familiar with the system;
- The enthusiasm for and commitment of the administration—library, city, or university—to training (teachers at one secondary school, for example, received "technical credit" as an incentive for training, which included free computers if enough credit were accumulated);
- the perceived user-friendliness of the software and the effectiveness of printed documentation available at the OPAC stations (one librarian described their system as a "piece of cake—no harder to use than an ATM machine");
- whether users were required to attend the training sessions; and
- the perceived sophistication and computer literacy of the users.

Problems and Pitfalls

In this section we describe nine specific problems that librarians encountered during training, along with some solutions proposed by those we interviewed.

Too great a time lapse may occur between training for a specific module and actual use of the module. Training should occur immediately prior to the implementation of a module; otherwise, details are forgotten. In one instance, circulation training occurred far in advance of system installation in order to take advantage of summer slack time. In another, training occurred before Christmas; by January those trained had forgotten much of what they had learned. At the opposite extreme, one library staff received training while the system was going up for public use.

Too few days of vendor training can lead to difficulties in absorbing the required information effectively. Trainers should allow plenty of time for questions and answers. Consider additional follow-up training on site after the staff has had a chance to get familiar with the system. Build in funds for training on a system just as you would budget money for that slick new computer equipment. It's all critical to the success of your project.

Inadequate indexing of documentation and out-of-date manuals cause frustration and anger. Examine vendor documentation carefully in the contract negotiation stage.

Factors that may affect installation can unexpectedly cause difficulties. One trainer could not deal with the complexities of the library's Novell LAN.

Outside circumstances such as site preparation delays and telecommunications/networking difficulties can adversely impact training. Have these details worked out before training begins.

Poor training logistics can result in a training group that is too large. Sharing workstations is not as effective as having one terminal per trainee, and one trainer may have difficulty dealing with a large group. Inadequate physical quarters may inhibit learning in other ways as well.

Training not suited to present levels of staff expertise is frustrating. In the case of one migrating library, the reference staff didn't want the "nitty-gritty," which they already knew. Instead, they wanted more sophisticated instruction about subjects such as typical user errors they might expect. One librarian commented:

> [The trainer] didn't need to belabor the basic system components for so long. We should have freed up the extremely expensive cost of training to cover more sophisticated aspects. The vendor needs to learn the audience and cut the basics.

Flexibility is critical. When training a large organization, for example, one can determine what works well after an initial session with the first group and revise further training accordingly. As one librarian put it, "Don't just trot out the boiler plate." The library should request a detailed training plan that is tailored to meet the specific needs of the library and that includes specific and well-defined objectives. Explain to the vendor what is expected.

Lack of centralized control over training by the system administrator can result in uneven results in multisite installations. Some locations may receive excellent instruction, whereas others may receive little or no assistance.

Ingredients for Successful Training

In addition to the problems discussed in the previous section, librarians offered suggestions for successful training. Observing the following hints should allow the reader to refrain from characterizing training, as did one library, as "days of infamy" that were "intense, hurried, and humorless."

- *Ensure that support service after initial training* includes excellent telephone support, a local back-up vendor representative, or both.
- *In a multisite system that will be implemented over time, carefully choose the first sites to be installed and trained.* In a county system or school district, for example, the earliest installation might be considered the "trailblazer" whose success or failure is closely watched by others. If the initial automation occurs in libraries whose personnel or other local circumstances do not fit the enthusiastic trailblazer model, negative attitudes toward the project can result. As one school media coordinator put it:

  > The first schools should have been chosen more carefully. I should have stood my ground with the principals. Politics dictated who was automated when.

- *No additional responsibilities should be assigned to the system administrator* during the implementation phase of the system. Timing of the training should allow the system administrator to become trained on all modules. As one participant said:

  > I was so busy bringing up the system that I couldn't attend all the training. Consequently, I wasn't careful enough about data elements that the staff were entering since I didn't know enough at the time to say anything.

In the case of school or special librarians who may be alone in the library, allow them additional time to work with the system without having to serve library users simultaneously.

- *In fact, everyone should be given official time to work with the system after initial training.* "Make using the system a priority," one administrator suggested. "If they don't have a chance to use the system, they can't learn it."
- *Ascertain in advance the expertise and experience of the trainer* responsible for a particular module. He or she should be an expert on that module. Ask other customers about their experiences with specific trainers. Trainers who have worked in libraries are generally more successful than those with a sales background. And just because someone knows the software well doesn't mean he or she can train. As one participant noted, "The best guy to help you in an emergency is [sometimes] the worst guy to train you."
- *Utilize options for advance preparation* in order to become somewhat familiar with the system before training begins. Staff should be given release time

to take advantage of pretraining tools. As one librarian pointed out:

> Training from a position of ignorance wastes a lot of time. It's like someone explaining the pieces of a jigsaw puzzle; until you work with it, it doesn't make any sense.

For one library, training began in the form of a newsletter, which was both funny and educational and offered multiple-choice questions and other training tidbits. One administrator enthusiastically described the opportunity given her staff to become familiar with a new system in advance of formal training by "playing around" with a large test database of the library's own records. The vendor provided the script, and the staff could check out functionality and call the vendor in case something did not work. Each staff person scheduled time to try out the system using the basic script; since they were on a test database, making mistakes such as deleting a book didn't matter. Another administrator at a library with no advance preparation commented:

> Vendors should develop a computer-based tutorial for pre-training. Staff could get familiar with the basics early. We had to make decisions on screen design before we had the opportunity to play around with the system. We should have been able to try options in advance.

- *The automation administrator must be positive.* And if other staff are doing the training, they must have a positive attitude too, even if it means putting aside earlier reservations they had about the system.
- *Expect higher stress levels and outright resistance* on the part of a few staff members with the advent of a new system. Smith has reviewed types of negative behavior which may arise to challenge the automation administrator and suggests some solutions.[5] Nuckolls reminds us that communication with staff is the key to ensuring cooperation and positive attitudes.[6] In the present study one librarian reported an unexpected outcome:

  > We tried to group some of the faster people with some slower ones, thinking that they could help each other. The problem was that after the training the older people felt that the pages were ready to run the library.

In another case, former staff "experts" on the old system suddenly found themselves at the same level as everyone else, which resulted in their having a negative attitude toward the new system.

- *Give staff an overview* of the entire system before focusing on specific details. As one participant explained:

  > They need to know how it all fits together and what effects what. Our clerical staff said, "Just tell me what I need to know," but if they don't know the framework, they won't understand the system.

- *Choose local trainers carefully.* Although a majority of participants believed that those trained directly by the vendor received better training, local training geared to specific situations can be highly effective. Participants agreed that library staff trainers definitely need to know a great deal about teaching and training; they concurred with Clayton's analysis that simply letting a "local computer whiz" conduct the training can backfire:

  > Creating an in-house program is both the best and most difficult way to train staff. . . . A common pitfall of internal instructional programs is to select individuals to do the training who already have the knowledge, regardless of their teaching ability. If a poor choice is made, this expedient solution can exact a price over time that can range from staff dissatisfaction with the training (at best) to a training program that actually repels staff from learning the topic. If one were to err on one side or the other, clearly it would be better to select an individual capable of distilling a topic to its essentials and presenting it clearly and understandably over someone who knows the topic.[7]

Finally, keep in mind that staff trainers will need release time for their new assignment.

- *Ensure the availability of excellent documentation.* Participants consistently emphasized the necessity of thorough, up-to-date, meticulously indexed, and well-organized documentation. A public librarian offered a typical opinion:

  > I think user support consists of three major areas: training, documentation, and direct customer interface (via phone, fax or in person); but the greatest of these is *documentation.* Documentation is available twenty-four hours a day, seven days a week. But in order for documentation to be truly useful, it has to be clear, thorough and well-organized. Otherwise, people will not use it.

In this library, training included a lecture by library staff who discussed a series of scenarios geared to local circulation policies and procedures. These scenarios were included in a manual kept at the circulation desk, where it was available for staff use when needed.

A variety of types of documentation are available, both for initial training and for staying current with new technology.[8]

Vendors' Views

In interviews lasting approximately forty-five minutes, trainers from the eight vendors participating in the

study discussed their views of the following problem areas brought up by librarians:

Noncustomized, "Boilerplate" Training, not Geared to Specific Situations and Present Levels of Competence

Some boilerplate training is unavoidable. With a wide variety of trainers and complicated, rapidly evolving software, it's a challenge to keep training up to date. Trainers are constantly faced with new materials, and standardization is necessary. "Of course, we have some boilerplate," one vendor said. "We must maintain a general plan that everyone follows." However, the same vendor noted that their company will skip modules or aspects at the customer's request. Nevertheless, the customer must "sign off" on omitted items, since "We don't want to be hit with a we-weren't-trained-on-this complaint."

Most vendors will alter training for migrating libraries if requested to do so. One vendor offers a choice of beginning, intermediate, advanced, or customized training. For example, the company presented a customized two-day training session for faculty in a school system to teach them how to use the software in teaching.

All vendors urge customers to let them know special needs or specific questions *in advance*. At the same time they point out that when libraries put *everyone* in training, instruction must then be paced to meet the needs of slower people, new hires, etc. Pleasing everyone on a site is difficult. By utilizing pretraining interviews and trips (including a walk-through), vendors attempt to vary training according to what actually happens in the library. It may be more appropriate to make follow-up training tailored for the local situation.

Inadequate Time Devoted to Training

Vendors agree that too little time for training is definitely a problem. Often they tell the customer what is needed, but the customer cuts back to save money. Sometimes acquiring training funds is difficult and the library must use in-service money. One vendor offered to "stay a year" if necessary:

> Our sales consultants do "due diligence" up front and may recommend extra days. It then falls on the library to make the decision. Librarians need to budget for training; the total training cost is low compared to the overall cost of the system.

New software platforms are sophisticated and complex, and users need adequate time to learn. Training days should be spaced with intervening time to practice using the system, as otherwise there is too much to absorb. "We are happy to come back if requested." Ideally, users should be grouped according to levels of competence which also becomes more expensive.

Trainer Goes Too Fast

This can definitely be a problem because the trainer often has a great deal to cover in a small amount of time. In any group, experience and abilities vary considerably; smaller groups of participants with roughly the same abilities are more effective than larger groups of people with widely differing skill levels. Some vendors ask everyone to complete an evaluation and believe that if they were proceeding too fast, previous participants would have let them know.

Variation among Trainers in Terms of Skills, Effectiveness, and Familiarity with a Module

Trainers use software every day, but since each trainer comes from a particular background and may be more skilled on one system than another, it is hard for them to stay equally up to date in all areas. Trainers undergo intensive in-house instruction; experienced trainers observe new trainers and ask them the same questions librarians are likely to pose.

Vendors try to keep their trainers at the same skill level, identifying and remedying each instructor's weaknesses. A vendor may have trainers read customer service logs to see what questions the customer is asking after training, thus providing them with useful feedback.

Inadequate Documentation

All the vendors consulted said that they are constantly working to improve and enhance their documentation and indexes, often a thankless task. Some vendors have sent modifications and corrections of documentation to the library's contact person only to find out later that the information was not passed on to staff.

Vendors revise documentation and correct errors through updates, directives, service bulletins, and user surveys. Vendors ask users to provide a variety of information, ranging from mistakes that were made to new indexing entries being used. One vendor consults users on such issues as whether all system information should be in a single manual or divided into separate volumes by function.

Another vendor described offering documentation online with hypertext links. In this interesting application users can input local policy and procedures, thus customizing documentation for their own environment.

Lack of a Detailed Training Plan

Most vendors believe they are providing adequate advance information. They send outlines and detailed day-to-day schedules to their contact person before training begins. And contracts may contain details of training that include hour-by-hour descriptions of instruction for each module. If the library so requests, most vendors say they are happy to provide more details, including ad-

vance copies of training hand-outs. At the beginning of the training they explain their overall agenda and timing of the training.

Vendors suggest that in these areas and others relevant information is sometimes not passed on to the people affected. "Maybe we should reproduce the material and send it directly to the department involved," one vendor said.

Failure to Present the "Big Picture" Before Focusing on Details

One vendor who always gives an overall view of the training at the outset pointed out that libraries will send a staff member for "only the part he or she needs" and that the trainer can't start over again for each newcomer. For example, one person only learned how to handle fines and missed the procedure for charging out books!

Another vendor sends customers a packet called "How to Prepare for Your Training." The packet contains material that customers need to read so they can make decisions about the training before it begins. But the vendor often finds that participants haven't read the packet when they start the training. Librarians must take more responsibility in this arena.

Lack of Input from the Library Concerning Timing of the Training

Librarians must request a training date in advance, especially for popular periods like August that fill up fast. Only a finite number of libraries can be trained in a given month, so the more notice the better. If the vendor's calendar is already full, the library must take another date. Sometimes the library staff being trained for a specific module doesn't set the date; rather, someone higher up in the library schedules the training. He or she should consult the people who are actually being trained.

Sometimes the library is eager to get started with training right away. One library insisted on being trained immediately (in June) and then complained that staff had forgotten what they learned by September. Many customers want to receive training and then "play around" with the system for a couple of months before coming up live. One vendor argued strongly against this approach, however, recommending that training occur when modules come up live.

Steps to Ensure Successful Training

We asked trainers what advice they might have for librarians who want to ensure that their training is successful. Trainers emphasized:

- Take the advice of the vendor;
- Communicate clearly what you want to learn; and
- Know that training will take time and cost money.

Training is a necessity. The library must budget for the appropriate amount of training and should plan carefully for training in consultation with the vendor's training representative. Libraries don't value training highly enough. Remember that when the group is diverse (seven librarians from seven different schools for example) training will proceed more slowly, less efficiently, and less in-depth. The "happiest" libraries, one trainer noted, are those which have a full-time trainer on staff and spend a lot of money on training.

Find a physically appropriate venue for training, away from the normal work environment with as little noise and traffic as possible. People attending the training should not feel pressured to continue their regular jobs. Some trainers said they taught in the middle of the public access terminals and, in one case, in a "closet." The ideal area would allow for three terminals, two people per terminal, with space behind and between them for the trainer to help. "Hands-on training is critical; we have to be able to show them physically," one vendor insisted.

Pick in-house trainers carefully: Choose a popular person who is smart and an excellent communicator; don't opt automatically for computer folks. According to one vendor, "Computer whizzes will not wait for the trainer; they are impatient and go off on their own. We look at their CRT and wonder how in the world they got where they are!" Occasionally librarian trainers will not really want to disseminate information, but hold onto it to protect and enhance their own turf.

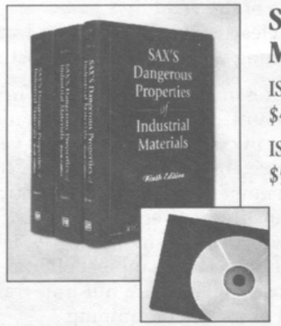
Some libraries have employed professional teachers to train their staff; it's critical to use trainers who have experience in making presentations and who have good communication skills.

The Future of Training

Most trainers believe there will be more computer-based instruction in the future, and some are exploring software that will ease the process, including multimedia options. Such an approach would require the addition of several new staff members for design, testing, updating, etc. Using the Internet for training—particularly for small groups that require unique, specialized instruction—is becoming increasingly attractive as options such as desktop video become more widely available.

Teleconferencing through interactive video is also a likely possibility, but most libraries don't have access to the necessary equipment and communication links;

Distinct, Comprehensive, Powerful...

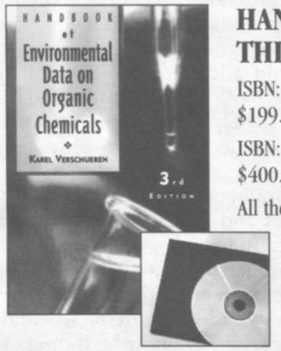


SAX'S DANGEROUS PROPERTIES OF INDUSTRIAL MATERIALS, NINTH EDITION by Richard J. Lewis, Sr.

ISBN: 0-442-02025-2 (Three-Volume Set) **Available now**
$499.95 ©1996

ISBN: 0-442-02335-9 (CD-ROM) **Available now**
$500.00 ©1996

A vast, reliable source of critical information on health and safety data, regulatory standards, toxicity, carcinogenicity, and physical properties of over 20,000 chemical substances. New: • 2,000 completely new chemicals, drugs, and biological agents • 14,000 updated entries • New or additional physical data added for 9,000 entries • CAS Registry Numbers for more than 18,000 entries • 100,000 domestic and international synonyms • 3 Instant Locator Indexes • Streamlined toxicity data • OSHA, ACGIH, MAK exposure limits • NIOSH REL workplace air limits • New advanced safety profiles for handling spills, accidental exposures, and fires • Carcinogens and reproductive effects • New European Data. Also available in CD-ROM.

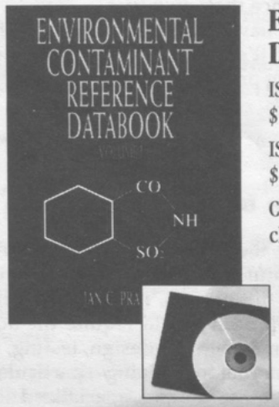


HANDBOOK OF ENVIRONMENTAL DATA ON ORGANIC CHEMICALS, THIRD EDITION by Karel Verschueren

ISBN: 0-442-01916-5 (Print Version) **Available April 1996**
$199.95 ©1996

ISBN: 0-442-02421-5 (CD-ROM) **Available June 1996**
$400.00 ©1996

All the information you need to prudently use potentially dangerous chemicals is in this fully revised edition of our classic handbook. Pesticides, detergents, phthalates, polynuclear aromatics, and polychlorinated biphenyls are all investigated in detail. Includes physical/chemical properties; air, soil, and water pollution factors; aquatic toxicity, and biological effects; odor thresholds; sampling and analysis data; and structural formulas of over 3,000 chemicals. Tables have been refined to focus on environmentally related materials. Also available in CD-ROM.



ENVIRONMENTAL CONTAMINANT REFERENCE DATABOOK, VOLUMES I AND II by Jan C. Prager

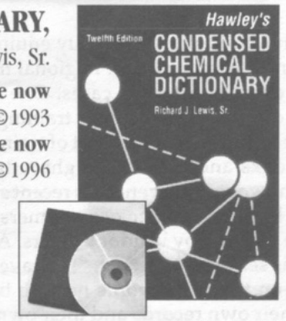
ISBN: 0-442-01918-1 (Vol. I), ISBN: 0-442-01969-6 (Vol. II) **Available now**
$129.95 (each volume) ©1995, 1996

ISBN: 0-442-02420-7 (CD-ROM) **Available now**
$400.00 ©1996

Comprehensive coverage of environmental effects of regulated chemicals. Identifies regulated chemicals and chemicals of special interest to environmental and safety professionals due to widespread contamination or considerable hazard. A reference compendium of physical, chemical, and biological dangers of environmental chemicals and selected simple and complex mixtures. Listings include CAS number; SAX number; sampling, analysis, and detection limits; environmental transport, fate and effects; structural formula; molecular formula; common uses; manufacturers; odor thresholds; regulatory jurisdictions and authorities; standards; reactions; monitoring methods; international data; synonym index. Also available in CD-ROM.

Order from your local wholesaler or order direct from:
Donald O'Connor, National Accounts Manager, Van Nostrand Reinhold
212-780-6129 (Phone) • 212-254-9499 (Fax)
doconnor@vnr.com (E:mail)

that's what sets VNR's chemical and environmental references above the rest (Available in print and CD-ROM)

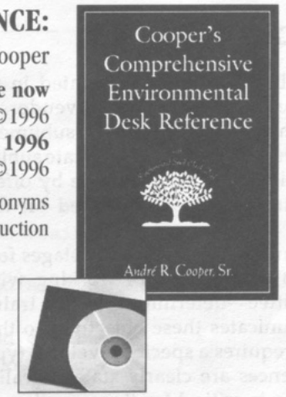


HAWLEY'S CONDENSED CHEMICAL DICTIONARY, TWELFTH EDITION by Richard J. Lewis, Sr.

ISBN: 0-442-01131-8 (Print Version) **Available now**
$79.95 ©1993

ISBN: 0-442-02419-3 (CD-ROM) **Available now**
$300.00 ©1996

The most widely recognized dictionary of industrial chemicals, terms, processes, reactions, and related terminology, Hawley's is an essential reference for all levels of staff in industries where chemicals are used — from management and administrative people to marketing and sales to technical and scientific staff. Hawley's presents three distinct types of information: (1)descriptions of chemicals, raw materials, manufacturing processes, and equipment; (2)expanded definitions of chemical entities, phenomena, and terminology used in every phase of engineering and technical development; and (3)descriptions or identifications of a wide range of trademarked products. Also available in CD-ROM.

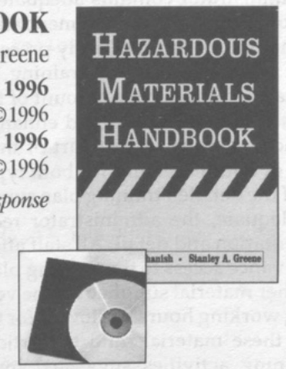


COOPER'S COMPREHENSIVE ENVIRONMENTAL DESK REFERENCE: WITH SUPPLEMENTAL SPELL CHECK DISK by André R. Cooper

ISBN: 0-442-02159-3 (Print Version) **Available now**
$99.95 ©1996

ISBN: 0-442-02161-5 (CD-ROM) **Available August 1996**
$200.00 ©1996

A unique comprehensive reference with more than 1,000 pages of key terms, over 5,000 acronyms and useful, detailed information on: • Superfund • Environmental engineering • Biology • Construction • Wetlands • OSHA • Threatened and endangered flora and fauna • and International organizations and treaties. Information is supplemented by maps, charts, tables, and diagrams, as well as crucial agency contact information to give the professional access to environmental officials who can offer advice and assistance. Destined to become a classic work, essential for all environmental, hazardous waste and safety engineers and managers, industrial hygienists, and environmental lawyers, as well as for students in these fields. Also available in CD-ROM.



HAZARDOUS MATERIALS HANDBOOK

by Richard P. Pohanish and Stanley A. Greene

ISBN: 0-442-02212-3 (Print Version) **Available July 1996**
$149.95 ©1996

ISBN: 0-442-02422-3 (CD-ROM) **Available September 1996**
$400.00 ©1996

This corrected, updated and expanded version of the U.S. Coast Guard's *Chemical Hazards Response Information System (CHRIS) Manual* features widely used and transported industrial materials including, practical hands-on data and technical data as well as chemical properties. This book is a portable and easier to use version of the *CHRIS Manual.* This new resource covers more than 1,240 substances, while also improving on the original by including a CAS index, synonym index, and a helpful identification section with new fields and vital ID information moved to the front of each record for easy accessibility. Also available in CD-ROM.

Van Nostrand Reinhold
115 Fifth Avenue, New York, NY 10003

the vendor would have to buy equipment and ship it to the library being trained. Regional training centers may prove effective in some cases. One vendor plans to increase offerings of advanced training at conferences and in user groups. Development of training programs using workbooks and tutorials might reduce on-site training costs; however, one vendor's recent survey revealed that a very high percentage of customers preferred personal on-site training by vendor trainers. Another vendor echoed this sentiment, saying, "We have no plans to replace hands-on training. People need to be on their own site with their own records and their own terminals with the trainer physically present."

Conclusion

One librarian who participated in the study said her dealings with two different vendors over the last eight years had ranged "from the sublime to the ridiculous." In the hope that we can facilitate sublime experiences for our colleagues, we conclude by offering the following ideal training scenario based on the findings of our study.

In the initial planning stages for staff training, the system administrator—working with the automation committee—determines specific training objectives and communicates these objectives to the vendor. If the library requires a specific level and type of training, these preferences are clearly stated. Realizing that excellent training is critical for the overall success of the project, the administrator commits adequate funds for training based on the vendor's recommendations. Initial training for a module will immediately precede implementation of that module. Follow-up training, held after the staff have had an appropriate amount of time to actually use the system, supplements and enhances initial training. Advance planning on the part of the library ensures a choice of optimum dates for both types of training.

If the detailed training plan supplied by the vendor is inadequate, the administrator requests appropriate augmentation and detail. All staff affected have convenient advance access to the training plan, documentation, and other material supplied by the vendor. Enough time during working hours is allowed for the staff to read and digest these materials and to participate in any other pretraining activities suggested by the vendor. Staff members benefit greatly from these efforts, since the system administrator ascertained in advance the quality of vendor documentation and pretraining material. Throughout the training process, thorough, ongoing communication with staff has the highest priority.

During implementation, the system administrator is able to devote full time to the automation project. The administrator designates a quiet, relatively secluded training venue, and adequate space permits the trainer to interact easily with trainees. The venue also has an adequate supply of terminals so that no more than one or two trainees are assigned to each one. Both administrator and trainer emphasize to staff the necessity of thoroughly understanding the documentation and learning how to use it. The learning process is enhanced by allowing plenty of time for questions and comments from staff.

Carefully selected local trainers continue staff training after vendor training has been completed. They, as well as other staff, can rely on competent, knowledgeable telephone support from the vendor. Looking back on the installation of the new system, all concerned agree that training was an overwhelming success.

References

1. Julie Hallmark and C. Rebecca Garcia, "System Migration: Experiences from the Field," *Information Technology and Libraries* 11 (Dec. 1992): 345–58.

2. For example, Charles A. Litchfield, "Vendor Training: A Question of Commitment to User Success," *Journal of Library Administration* 12, no. 2 (1990): 3–12; Stuart Glogoff, ed., "Staff Training in the Automated Library Environment: A Symposium," *Library Hi Tech* 7, no. 4 (1989): 61–83; Margie Epple, Judy Gardner, and Robert T. Warwick, "Staff Training and Automated Systems: 20 Tips for Success," *The Journal of Academic Librarianship* 18 (May 1992): 87–89; Brenda Turner, "Staff Training for a NOTIS OPAC: Methodologies and Assessment," *Library Software Review* 11, no. 6 (Nov./Dec. 1992): 12–23; Byron Anderson and Samuel T. Huang, "Impact of New Library Technology on Training Paraprofessional Staff," *The Reference Librarian* No. 39 (1993): 21–29.

3. *Training of Technical Services Staff in the Automated Environment: SPEC Kit 171* (Washington, D.C.: Systems and Procedures Exchange Center, OMS, Association of Research Libraries, 1991).

4. Helen Adams, "Media Magic: Automating a K–12 Library Program in a Rural District," *Emergency Librarian* 21, no. 5 (May–June 1994): 24–29.

5. Kitty Smith, "Toward the New Millennium: The Human Side of Library Automation (Revisited)," *Information Technology and Libraries* 12 (June 1993): 209–16.

6. Karen Nuckolls, "Humanism and Automation: Working with People in the Library Automation Process," *The Reference Librarian* No. 38 (1992): 109–14.

7. Marlene Clayton with Chris Batt, *Managing Library Automation*, 2d ed. (Brookfield, Vt.: Ashgate, 1992).

8. Roy Tennant, "The Virtual Library Foundation: Staff Training and Support," *Information Technology and Libraries* 14 (Mar. 1995): 46–49.

Appendix A
Interview Schedule for Participating Librarians

Librarian's Name _____

Title _____

Name of Library _____

City and State _____

Type of Library _____

Name of Automation System: _____

First Time? _____

Migrating? From which system? _____

I. Training the Staff
   A. How many staff were to be trained in all?
      What were your main objectives in training staff?
   B. Methodology
      1. When?
      2. Who?
         Who determined the schedule and manner in which training would be conducted?
         Did the vendor train all staff or only selected staff?
      3. How?
         Were staff trained all at once or over a period of time?
         Was there advance preparation for staff (pre-training)?
      4. Where?
         Local site or vendor site?
   C. How effective was the training?
      Please describe your successes and problems in the following areas.
      1. Cataloging
      2. Circulation
      3. Acquisitions
      4. Serials
      5. System Administration
      6. Other
   D. Is there anything else that you would do differently if you had it to do over again?

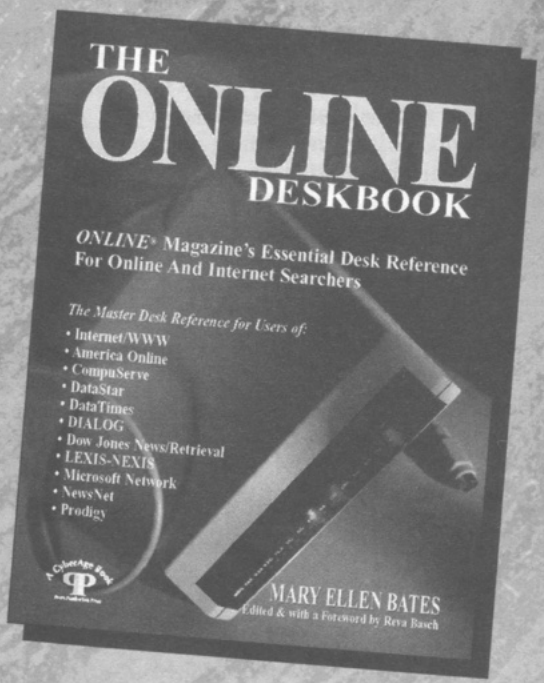
II. Training Users
   A. Did you train users?
   B. Methodology
      Did you have different types of training for different age levels or user groups?
      How did you train users?
   C. How effective was the training?
   D. Please describe problems.

The Ultimate Reference for Anyone Who Searches Online Services and the Internet

THE ONLINE DESKBOOK

ONLINE Magazine's essential desk reference for online and Internet Searchers

By Mary Ellen Bates
Edited and with a foreword by Reva Basch



- ✔ Provides the essential tools you need to identify important online information sources and put them to immediate use.
- ✔ Loaded with command charts, shortcuts, troubleshooting guides, tips, techniques, and answers to frequently asked questions.
- ✔ Contains the nuts and bolts to get up and running online…log-on information, system requirements, Internet connectivity, pricing, directory/finding tools; and more.
- ✔ Covers *all* the major online services and the Internet including:

FREE ONLINE UPDATES
WORLD WIDE WEB
LINKS TO USEFUL SITES

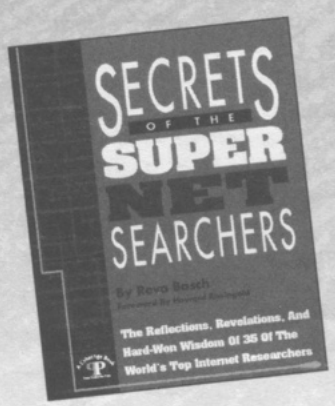
- Internet/WWW
- America Online
- CompuServe
- DataStar
- DataTimes
- DIALOG
- Dow Jones News/Retrieval
- LEXIS-NEXIS
- Microsoft Network
- NewsNet
- Prodigy

256 PAGES • 7 x 9¼ • PAPER
$29.95 • ISBN 0-910965-19-6

Available at your local bookstore.
To order or for a free catalog,
email: booksales@onlineinc.com or
call 800/248-8466, Ext. 517.
Outside the U.S. call 203/761-1466.
Fax: 203/761-1444

"Wish I'd written it...wish I'd had it long ago."
Barbara Quint
Editor, Searcher magazine

And Coming this Fall…**Secrets of the Super Net Searchers**



The Reflections, Revelations, And Hard-Won Wisdom Of 35 Of The World's Top Internet Researchers

By Reva Basch
Foreword By Howard Rheingold

Topics covered include:

- Distinguishing cyber-gems from cyber-junk
- Finding experts and enlisting their help
- Search strategies that really work
- How not to get "Lost In Cyberspace"
- How to avoid "Internet Overload"
- Keeping up-to-date on new Internet resources

From the publisher of… ONLINE database online user

Pemberton Press
an imprint of Online Inc.

OPAC Missing Record Retrieval

Karl E. Johnson

A study was conducted to determine whether a single library's cataloging records known to be missing from a consortium OPAC database could be identified by using the database search features. Attempts to create lists of bibliographic records held by other libraries in the consortium by using the Boolean search features of the database failed due to search feature limitations. Samples of search logic were created; collections of records based on this logic were assembled manually and then compared to the card catalog of the single library. The results suggested that use of the Boolean operator OR to conduct the broadest possible search could find approximately 56,000 of the single library's missing records that were held by the other libraries. Use of the Boolean AND to conduct the narrowest search possible found that approximately 85,000 of the single library's missing records were held by the other libraries. A specific library search made of the records of the most likely consortium library to have overlaid the single library's holdings, found that approximately 80,000 of the single library's missing records were held by the specific library. Use of the Boolean operator AND was found to give the highest rate of return (50 percent) of the missing records.

In 1984 the University of Rhode Island (URI), Roger Williams University (RWU), Rhode Island College (RIC), and the Community College of Rhode Island (CCRI) formed a consortium called the Higher Education Library Information Network (HELIN). (Providence College (PC) subsequently joined the consortium in March 1992.) An automated system produced by Innovative Interfaces, Inc. (III) was chosen for the consortium. In March 1991 the URI library's new online catalog became accessible to staff, nine months before students gained access. It soon became obvious that data had somehow been lost or corrupted. Bibliographic records attached to item records of affiliated institutions often indicated that URI, whose records had been loaded first, was the originator of the bibliographic records; yet URI item records were missing. An investigation of the problem concluded that approximately 10 percent of the URI monographic records were missing from the database.[1] Many records had been overlaid by records of other consortium institutions—namely CCRI, RIC, and RWU—due to the fact that a database overlay setting was in effect during tape loading of records, and many URI bibliographic records were not accompanied by item records.

Methodology

Because many of the missing records were for items shared with other consortium members, the author decided to attempt to identify the missing records by using the OPAC's Boolean logic search features rather than by comparing the HELIN database to the URI library's more than 700,000-card shelflist. Bibliographic location codes for consortium member institutions other than URI were to be "ORed" or "ANDed" to create lists of records held by one or all consortium members other than URI. In addition, a count of the number of records held by individual libraries as well as the order in which their records had been loaded was made in an attempt to identify the most likely single library candidate for holding records for items also held by URI. As the capacity of the OPAC Boolean list makers available was not large enough to count total records for individual consortium libraries or for shared records, estimates were calculated as the result of the percentage of records found in 40,000 record lists. The search logic follows:

Broad Search
LOCATION = CCRI and LOCATION not URI
or LOCATION = PC and LOCATION not URI
or LOCATION = RIC and LOCATION not URI
or LOCATION = RWC+ and LOCATION not URI
or LOCATION = RWU and LOCATION not URI
+ Earlier name for Roger Williams University (Roger Williams College) under which earlier records are coded/identified

Narrow Search
LOCATION = CCRI and LOCATION not URI
and LOCATION = PC and LOCATION not URI
and LOCATION = RIC and LOCATION not URI
and LOCATION = RWC and LOCATION not URI
and LOCATION = RWU and LOCATION not URI

Library Specific Search
LOCATION = RIC and LOCATION not URI

Rhode Island College was believed to be the most likely candidate to have overlaid URI records, since its records had been loaded second into the OPAC database and its holdings were second largest (see table 1).

The original intent was to create 1,000 record samples for each search logic and then compare the resulting samples to the URI card catalog. Missing-item call numbers would then be obtained from the catalog card for

Karl E. Johnson is former Monographic Cataloger, University of Rhode Island.

Table 1

Order of Loading	Approximate Number of Records	
1 URI	715,725	(66% of 1,084,423)
2 RIC	383,885	(35.4% of 1,084,423)
3 CCRI	112,195	(10.3% of 1,084,423)
4 RWC/RWU	169,170	(15.6% of 1,084,423)
5 PC	177,845	(16.4% of 1,084,423)

subsequent OPAC database item creation; the addition of barcodes would require another step in the process. Unfortunately, the database Boolean logic feature was unable to separate the bibliographic location URI from other consortium member records. Therefore, the author removed URI records manually from the sample lists. This necessitated acceptance of smaller samples in order to draw some conclusions as to record recovery returns if and when the Boolean search logic feature became fully operative. The resulting lists, after comparison to records in the URI card catalog, gave the following rates of return for URI missing records.

Broad Search
87/468 = 18.6%

Narrow Search
95/190 = 50%

Specific Search
133/635 = 21%

The total size of the electronic database was 1,084,423 records as of January 1995. The approximate number of "ORed" records obtained by the broad search was 28 percent of the total database, or 303,638 records. As the percentage of URI records retrievable from the ORed records was 18.6 percent, the total number of URI records retrievable by this method calculates out to 56,477 records. The approximate number of ANDed records obtained by the narrow search was 15.8 percent of the total database, or 171,339 records. As the percentage of URI records retrievable from the "ANDed" records was 50 percent, the total number of URI records retrievable by this method calculates out to 85,669 records. The approximate number of RIC records obtained by the specific search was 35.4 percent of the total database, or 383,886 records. As the percentage of URI records retrievable from the RIC records was 21 percent, the total number of records retrievable by this method calculates out to 80,616 records. The approximate number of missing and retrievable URI records for the three searches are summarized as follows:

Broad Search
56,477 records

Narrow Search
85,669 records

Specific Search
80,616 records

Summary

Final calculations suggest that approximately 56,000 of URIs missing cataloging records could be retrieved by the Boolean OR process whereas approximately 85,000 could be retrieved via the Boolean AND process. The results of an examination of the second-largest and second-loaded RIC records suggest that approximately 80,000 of URIs missing records could be retrieved by this method. As expected, these numbers are higher than the total number of shared (ORed) records calculated for monographs only by Baer, Barnett, and Johnson (54,000).[2] But the total number of shared (ORed) records of 311,229, as indicated by the broader search, is significantly higher than the 207,000 reported by the above. It must be noted that the database is undergoing constant change; not only is it growing in size as more records are added, but correction and reconstruction of records are ongoing as well. A project to reconstruct the Reference Collection records began in the spring of 1994 and was approximately 75 percent complete one year later.

Conclusions

Many but not all of the URI records, destroyed by unwise merging of the four original databases, should be retrievable without examining the URI shelflist if and when the database Boolean logic search feature becomes fully operational. The total number of records, monographic and serial, retrievable from the OPAC database by these methods appears to be between 56,000 and 85,000 records. Staff would have to examine between 171,000 and 384,000 records, depending upon the method chosen, in order to identify the missing records. This task is certainly an arduous one but somewhat less daunting than examining the more than 700,000 records in the shelflist. Obviously, the narrow search appears the

method of choice as it promises the highest rate of return. However, the calculations upon which these judgments are based are statistically weak due to the difficulty in obtaining adequate samples—especially for the broad and the narrow searches. More accurate studies are recommended strongly if and when a fully workable OPAC database Boolean logic search feature is implemented.

References

1. Nadine L. Baer, James A. Barnett, and Karl E. Johnson, "OPAC Database Creation Problems," *Information Technology and Libraries* 14, no. 3 (Sept. 1995): 179–85.
2. Ibid., 184.

We found a new way to update Tulane's library serial holdings and, unfortunately, not spend a night in New Orleans.

When the TALX team was told they would be updating Tulane's local library system, they literally jumped for joy.

When they were told they'd be doing it with innovative technology on the Internet, and not spending time in New Orleans, you could've heard a pin drop. The program was to be done online from St. Louis.

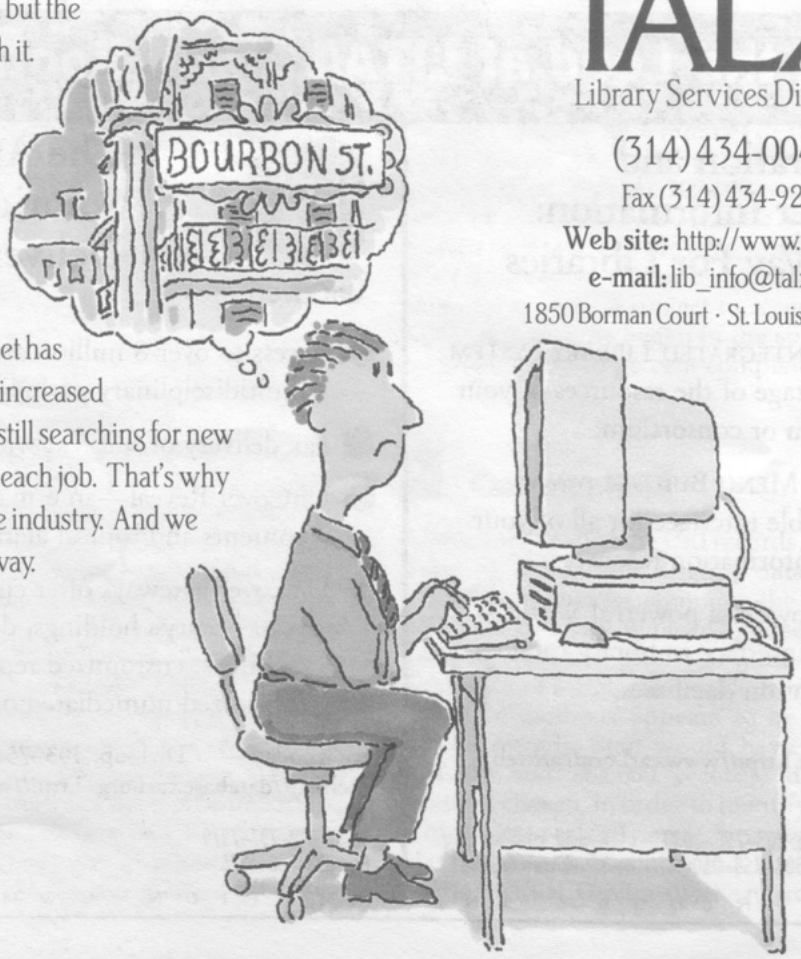
The TALX team had lost an opportunity to visit a fun city on a regular basis, but the results would be worth it.

By going online, Tulane's local library system is having their serial holdings database updated while remaining in service.

While the Internet has opened new doors to increased productivity, TALX is still searching for new and better ways to do each job. That's why TALX is a leader in the industry. And we intend to keep it that way.

Today our services cover a wide range. We offer Retrospective Conversion, Authority Control, Bar Coding, online MARC record upgrade via the Internet, and additional complementary services.

We feel that TALX knows library cataloging inside and out. And we'd like to talk to you about getting your library in line. And maybe even online.



TALX
Library Services Division
(314) 434-0046
Fax (314) 434-9205
Web site: http://www.talx.com
e-mail: lib_info@talx.com
1850 Borman Court · St. Louis, MO 63146

The Many Faces of a Catalog Record: A Snapshot of Bibliographic Display Practices for Monographs in Online Catalogs

Gregory Wool

Full-level displays in thirty-six North American online catalogs of bibliographic records for five monographs were analyzed and compared for both layout and content with the International Standard Bibliographic Description (ISBD) and traditional catalog-card display practice for headings and tracings. This report presents a taxonomy of display characteristics found in the sample, including completeness, visual layout, data sequence, label specificity, and integration of description and access points. The relationship of display technology and practice to current standards for creating catalog records is discussed, with emphasis on the need for content standards reflecting the modular nature of machine-readable records.

One of the hallmarks of online catalog technology is the variety of forms a bibliographic record can now take. Not only are there dozens of catalog software packages on the market, each offering its own style of interface, but most of these packages also enable libraries to customize their bibliographic displays by mapping MARC fields to locally defined labels and determining which fields to display and in what order. As a result, records created under AACR2 and printed in the citation-style format prescribed by the ISBD often look very different online, in content as well as format.

Previous research by Wool et al. characterized this customization process as "machine translation," identified problems resulting from this process in one catalog (at the University of Nebraska-Lincoln, or UNL), and explored possible implications for cataloging practice and standards. Issues highlighted include the blurring of the AACR2 distinction between description and access points, the potential for inaccurate labeling of data, and the emergence of data redundancies.[1] To obtain a sense of how widely relevant such issues might be, the present study was devised to examine bibliographic record displays in research-level online catalogs throughout North America. The focus of this study was on records for monographs, since these represent the vast majority of library catalog records (over 80 percent of records in the OCLC database are in the Books format)[2] and thus the dominant data structure within catalogs.

Literature Review

Wool et al. were able to identify only one previous study of online bibliographic record reformatting, and that one did not address the relationship of such reformatting to cataloging standards and practice, concentrating instead on visual appeal and convenience.[3] It was found to be supplemented by a number of articles offering recommendations and principles for online record-display design.[4] Wool et al. also made reference to a more extensive body of literature on issues related to record reformatting such as (1) the role of authority-controlled headings in general, and the main-entry heading in particular, in online catalogs; (2) the nature, role, and future of ISBD; and (3) the possible need for a new cataloging code.[5] In addition, they noted evidence of a dawning awareness of the relationship between machine reformatting and cataloging practice, mostly in listserv postings, conference presentations, and a task force charge.[6]

Discussions of bibliographic single-record displays continue to be scarce in the literature of OPAC interface design and—for the most part—all too brief. A possible justification for this can be found in Curwen's survey of the background and development of ISBD, which he maintains was intended to serve as a standard for the exchange of information, not for its display to end users.[7] He sees the rearrangement and labeling of record elements in OPACs as an issue needing discussion, but nevertheless views it as a welcome move toward user-friendliness.[8] Brunt, on the other hand, sees in the adoption of new display formats a gradual abandonment of AACR2 and proposes changes to the rules to make them more relevant and functional in the online environment.[9] Shires and Olszak include records in a comprehensive discussion of screen display principles, advocating selectivity and multiple levels of fullness in display, along with other features promoting attractive layout and user convenience.[10] O'Brien cites the structure of the bibliographic record as an impediment to the use of advances in database design.[11] In a case study of a specialized bibliographic database, Pickens argues that field definition should follow closely the underlying structure of the information to enable searching on as many information facets as possible.[12] He even makes a case for data repetition within a record, noting that presenting citation information in a single field enables quick identification, while repeating the same information in separate, specialized fields makes it indexable.[13] Troutman notes the major arguments for and against selectivity and labeling of data, while recommending a nearly complete display for music materials.[14] Weston suggests a need for OPACs to offer a choice of display

Gregory Wool is Monographs/Special Projects Cataloger, Parks Library, Iowa State University. His e-mail address is gwool@iastate.edu.

styles, since many scholars might want to see records in a familiar bibliographic citation style such as MLA or Chicago.[15]

Occasional suggestions continue to appear for changing cataloging to take better advantage of OPAC capabilities. These include bringing together multiple editions of a work in a single record (Ayres), ending the use of abbreviations in the physical-description portion of a record (Warrick), abandoning name headings and rule-based descriptions in favor of straight title-page transcriptions to facilitate keyword-style known-item searching and slash processing costs (Kilgour), and redesigning the bibliographic record based on the demonstrated needs and preferences of catalog users (Wildemuth and O'Neill).[16] Winke and Bierbaum add to the chorus of voices calling for the elimination of main entry.[17]

Mandel, in advocating a paradigm shift in the way catalogers approach their work in order to deal with a number of pressing issues, does not address bibliographic standards or the display of records.[18] At the same time though, she notes trenchantly: "While the intellectual process of original cataloging and its end product have adapted admirably to the production of machine-readable records, they have remained surprisingly unaffected by the environment in which these records are used."[19] Jeffreys and Bourne are more explicit in characterizing AACR2 as oriented to card-style displays.[20] Leazer applies the conceptual framework of database design to an analysis of the USMARC formats and calls for a rethinking of all cataloging standards using this framework.[21] Heaney answers the call with a proposal based on object-oriented modeling.[22]

Several comprehensive surveys of online catalog features and capabilities at the library level were conducted during the early 1980s (when the number of online catalogs was still small). Those of Hildreth and Salmon included consideration of single-record display formats.[23] Hildreth found thirty-seven different formats (including brief displays) in the ten catalogs he studied and reported briefly and generally on vocabulary, spacing, typography, punctuation, content, and data sequence.[24] Salmon noted an overwhelming prevalence of labeled, tabular displays.[25] Much more recently, Cherry et al. developed a checklist of features for a comparative survey of twelve OPACs in Canadian academic libraries, including the following aspects of record display: use of labeling, use of upper- and lowercase typography, availability of more than one level of fullness, and user specification of fields for display.[26] Crawford's book-length survey of online catalog features includes numerous examples of record displays, but discussion of them, and of the topic in general, is cursory.[27]

The reformatting of bibliographic records for online display has been a discussion topic on the AUTOCAT listserv several times (fall 1993, March-April 1994, January 1995, early May 1995, mid-June 1995) and has figured as well in AUTOCAT discussions of such other matters as main entry, ISBD punctuation, and order of notes.[28] Johnson has discussed its impact on record integrity and usefulness in a brief article.[29] More recently, the possible need for a new cataloging code was the topic of a daylong seminar at the American Library Association Annual Conference, with papers by Lynch, Levy, Graham, Hensen, Yee, Tillett, Gorman, and Howarth presenting a wide variety of arguments pro and con.[30]

Methodology

This study compares bibliographic record displays from a variety of library catalogs to one another and to the card-style "original." To this end, five records from the sample used in the above-mentioned UNL study were chosen for analysis. They describe and provide access to the following monographs:

American Apostasy: The Triumph of "Other" Gospels, edited by Richard John Neuhaus

Selected Letters of Witter Bynner, edited by James Kraft

Social Control: The Production of Deviance in the Modern State by Nanette J. Davis and Bo Anderson

Faulkner's Country Matters: Folklore and Fable in Yoknapatawpha by Daniel Hoffman

The Bieberbach Conjecture: Proceedings of the Symposium on the Occasion of the Proof (Purdue University, 1985)

Figures 1–5 show the OCLC/MARC records for these items.

The records were chosen less for subject variety than to ensure coverage of such common bibliographic elements as multiple authors, series statements traced differently, editors, uniform titles, and corporate main entries. Moreover, the records chosen were each created by the Library of Congress (LC), ensuring a high standard of AACR2 compliance and a low level of variance at the MARC-record level (although that level of variance turned out to be significantly higher than anticipated).[31]

The pool of catalogs to be studied was then created by determining, as of August 1994, which online catalogs included all five records. Examination of holdings records on OCLC produced twenty-one libraries meeting this criterion (three of which were eliminated from the study because their catalogs were not Internet-accessible). To reach libraries doing significant amounts of cataloging outside OCLC, as well as consortium-based catalogs with a common record interface, the author

```
OCLC:  18560494          Rec stat:  p
Entered:  19880919       Replaced:  19890624      Used: 19931006
Type: a        Bib lvl: m        Source:          Lang:  eng
Repr:          Enc lvl:          Conf pub: 1      Ctry:  miu
Indx: 0        Mod rec:          Govt pub:        Cont:  b
Desc: a        Int lvl:          Festschr: 0      Illus:
               F/B:     0        Dat tp:   s      Dates: 1989,
  1  010      88-30148
  2  040      DLC |c DLC
  3  020      0802802109
  4  050 0    BT1315.2 |b .A45 1989
  5  082 0    273.9 |2 19
  6  245 00   American apostasy : |b the triumph of "other" gospels / |c essays by Peter L.
              Berger ... [et al.] and The story of an encounter by Paul T. Stallsworth ;
              edited and with a foreward by Richard John Neuhaus.
  7  260      Grand Rapids, Mich. : |b Eerdmans Pub. Co. in cooperation with the Rockford
              Institute Center on Religion & Society, |c c1989.
  8  300      ix, 137 p. ; |c 21 cm.
  9  490 1    Encounter series ; |v 10
 10  500      Papers presented at a conference sponsored by the Rockford Institute Center on
              Religion & Society, New York, Jan. 23-24, 1987.
 11  504      Includes bibliographical references.
 12  505 0    Different gospels / Peter L. Berger — Secular justifications of truth-claims /
              James Turner — Gospel, church, and politics / Avery Dulles — A Protestant
              constructive response to Christian unbelief / Robert W. Jenson — The story of
              an encounter / Paul T. Stallsworth.
 13  650  0   Apostasy |x Congresses.
 14  700 10   Berger, Peter L.
 15  700 10   Neuhaus, Richard John.
 16  710 20   Center on Religion & Society (New York, N.Y.)
 17  830  0   Encounter series (Grand Rapids, Mich.) ; |v 10.
```

Figure 1
Record for *American Apostasy*, OCLC/MARC format

catalogs with a common record interface, the author combed the list of RLIN contributors and lists of library catalogs at several gopher sites for additional catalogs to search. In the end, thirty-six Internet-accessible catalogs in North America were found to contain all five records. They are

- Arizona State University (CARL)
- Auburn University (NOTIS)
- Brigham Young University (NOTIS)
- Carleton University (in-house)
- CLIC—Cooperating Libraries in Consortium (Minnesota) (Dynix)
- DRA catalog of Library of Congress records (DRA)
- Duke University (DRA Atlas)
- Harvard University (in-house)
- ILLINET—State of Illinois union catalog (in-house)
- Indiana State University (NOTIS)
- Indiana University at Bloomington (NOTIS)
- Library of Congress—SCORPIO LOCI file (in-house)
- Marquette University (INNOPAC)
- MELVYL—University of California system (in-house)
- MSUS—Minnesota State University System (PALS)
- Northwestern University (NOTIS)
- Ohio State University (in-house)
- Rutgers University system (Geac)
- Southern Methodist University (SMU) (NOTIS)
- State University of New York (SUNY) at Buffalo (NOTIS)
- University of Alabama at Tuscaloosa (NOTIS)
- University of California at Santa Barbara (NOTIS)
- University of Hawaii at Manoa (CARL)
- University of Kansas (in-house)
- University of Nebraska at Lincoln (INNOPAC)
- University of New Brunswick (in-house)
- University of North Carolina at Chapel Hill (DRA Atlas)
- University of Rochester (Geac)
- University of Toledo (NOTIS)
- University of Toronto (DRA Information Gateway)
- University of Virginia (NOTIS)

```
OCLC:   7276463          Rec stat:  c
Entered:  19810305       Replaced:  19900210     Used:  19931118
Type: a        Bib lvl: m       Source:          Lang:  eng
Repr:          Enc lvl:         Conf pub: 0      Ctry:  nyu
Indx: 1        Mod rec:         Govt pub:        Cont:  b
Desc: a        Int lvl:         Festschr: 0      Illus: c
               F/B:     0a      Dat tp:   s      Dates: 1981,
  1  010      81-2087//r90
  2  040      DLC |c DLC
  3  020      0374185042 : |c $30.00
  4  050 00   PS3503.Y45 |b Z48 1981
  5  082 00   811/.52 |a B |2 19
  6  100 1    Bynner, Witter, |d 1881-1968.
  7  240 10   Correspondence. |k Selections
  8  245 10   Selected letters / |c edited, and with an
              introduction, by James Kraft.
  9  260      New York : |b Farrar, Straus, Giroux, |c c1981.
 10  300      xxiv, 275 p. : |b port. ; |c 22 cm.
 11  490 1    The Works of Witter Bynner
 12  504      Includes bibliographical references and index.
 13  600 10   Bynner, Witter, |d 1881-1968 |x Correspondence.
 14  650  0   Poets, American |y 20th century |x Correspondence.
 15  700 10   Kraft, James.
 16  800 1    Bynner, Witter, |d 1881-1968. |t Works. |f 1978.
```

Figure 2
Record for Witter Bynner's *Selected Letters*, OCLC/MARC format

```
OCLC:   9219477          Rec stat:  p
Entered:  19830124       Replaced:  19870425     Used:  19930721
Type: a        Bib lvl: m       Source:          Lang:  eng
Repr:          Enc lvl:         Conf pub: 0      Ctry:  nyu
Indx: 1        Mod rec:         Govt pub:        Cont:  b
Desc: a        Int lvl:         Festschr: 0      Illus:
               F/B:     0       Dat tp:   s      Dates: 1983,
  1  010      83-136
  2  040      DLC |c DLC |d m/c
  3  020      082900727X : |c $18.95
  4  050 0    HM291 |b .D339 1983
  5  082 0    303.3/3 |2 19
  6  100 1    Davis, Nanette J. |w cn
  7  245 10   Social control : |b the production of deviance in the modern state / |c Nanette
              J. Davis and Bo Anderson.
  8  260      New York, N.Y. : |b Irvington Publishers, |c c1983.
  9  300      xxi, 365 p. ; |c 24 cm.
 10  504      Bibliography: p. [329]-356.
 11  500      Includes indexes.
 12  650  0   Social control.
 13  650  0   Social structure.
 14  650  0   Deviant behavior.
 15  650  0   Self-help groups.
 16  650  0   Social policy.
 17  700 10   Anderson, Bo. |w cn
```

Figure 3
Record for Davis and Anderson's *Social Control*, OCLC/MARC format

```
OCLC:  19556051          Rec stat:  c
Entered:  19890321       Replaced:  19910914      Used:  19931129
Type: a        Bib lvl: m        Source:          Lang:  eng
Repr:          Enc lvl:          Conf pub: 0      Ctry:  lau
Indx: 1        Mod rec:          Govt pub: s      Cont:  b
Desc: a        Int lvl:          Festschr: 0      Illus:
               F/B:    0         Dat tp:   s      Dates: 1989,
  1  010      89-32484//r91
  2  040      DLC |c DLC
  3  020      0807115622 (alk. paper)
  4  043      n-usu---
  5  050 00   PS3511.A86 |b Z789 1989
  6  082 00   813/.52 |2 20
  7  100 1    Hoffman, Daniel, |d 1923-
  8  245 10   Faulkner's country matters : |b folklore and fable in Yoknapatawpha / |c Daniel
              Hoffman.
  9  260      Baton Rouge : |b Louisiana State University Press, |c c1989.
 10  300      xviii, 181 p. ; |c 23 cm.
 11  440  0   Southern literary studies
 12  504      Includes bibliographical references and index.
 13  600 10   Faulkner, William, |d 1897-1962 |x Knowledge |x Folklore.
 14  650  0   Fables, American |z Southern States |x History and criticism.
 15  650  0   Literature and folklore |z Southern States.
 16  650  0   Yoknapatawpha County (Imaginary place)
 17  650  0   Southern States in literature.
 18  650  0   Oral tradition in literature.
 19  650  0   Country life in literature.
 20  650  0   Folklore in literature.
```

Figure 4
Record for Daniel Hoffman's *Faulkner's Country Matters*, OCLC/MARC format

```
OCLC:  13703242          Rec stat:  c
Entered:  19860519       Replaced:  19870627      Used:  19931108
Type: a        Bib lvl: m        Source:          Lang:  eng
Repr:          Enc lvl:          Conf pub: 1      Ctry:  riu
Indx: 0        Mod rec:          Govt pub:        Cont:  b
Desc: a        Int lvl:          Festschr: 1      Illus: af
               F/B:    0         Dat tp:   s      Dates: 1986,
  1  010      86-10843//r86
  2  040      DLC |c DLC |d m/c
  3  020      0821815210 (alk. paper)
  4  050 0    QA331 |b .S947 1985
  5  082 0    515 |2 19
  6  111 2    Symposium on the Occasion of the Proof |d (1985 : |c Purdue University) |w cn
  7  245 14   The Bieberbach conjecture : |b proceedings of the Symposium on the Occasion of
              the Proof / |c Albert Baernstein II ... [et al.], editors.
  8  260      Providence, R.I. : |b American Mathematical Society, |c c1986.
  9  300      xvi, 218 p., [2] p. of plates : |b ill. ; |c 26 cm.
 10  440  0   Mathematical surveys and monographs, |x 0076-5376 ; |v no. 21
 11  504      Includes bibliographies.
 12  650  0   Geometric function theory |x Congresses.
 13  600 10   Bieberbach, Ludwig, |d 1886- |x Congresses. |w cn
 14  700 10   Baernstein, A. |q (Albert), |d 1941- |w cn
 15  710 20   American Mathematical Society. |w cn
```

Figure 5
Record for *The Bieberbach Conjecture*, OCLC/MARC format

- Vanderbilt University (NOTIS)
- Wake Forest University (Dynix)
- Washington & Lee University (INNOPAC)
- Yale University (NOTIS)
- York University (NOTIS)

System/vendor information, in parentheses, was drawn either from catalog screens or from the HYTELNET utility for Internet access to catalogs.[32] Catalog systems for which no "brand" information was available are assumed to be locally developed. The breakdown by vendor is as follows:

- NOTIS: 14
- DRA: 4
- INNOPAC: 3
- CARL: 2
- Dynix: 2
- Geac: 2
- PALS: 1
- in-house: 8

Admittedly, as a representative sample of North American OPACs, this set leaves much to be desired. The small size alone is a problem. Large and medium-sized universities dominate the list. Consortia are oversampled, as are universities with theology schools. Public libraries, special libraries, school libraries, community colleges, and small liberal arts colleges are represented only through consortia such as ILLINET and Minnesota's Cooperating Libraries in Consortium (CLIC); the local OPACs of such institutions are absent altogether.

Statistical generalization, however, was not an objective in this study. What was sought, rather, was a snapshot of display practices among libraries and consortia with broad collections of scholarly materials. Such libraries naturally have an incentive to make full bibliographic data available to their users. Moreover, the apparent oversampling of NOTIS sites produced a cluster useful for studying variations in display practice among users of the same OPAC software. (NOTIS-specific analyses, though, were not prepared for this article.)

Printouts of the 180 record displays were made at the time of searching (August 1994) and were examined later. For each printout, the data labels, their display order, and the fields and subfields displayed (as found in the OCLC-MARC record) were noted. The data across the five records were compiled into composite display formats for each catalog. With only five records studied and the MARC records from the individual libraries not examined, the resulting information is necessarily incomplete (and in a few instances, conjectural) with regard to the entire subset of monograph records in each catalog. A surprisingly large number of data variations from the (LC-created) OCLC records made analysis more difficult. Finally, a small number of printouts proved to be incomplete (missing the second screen of data), and several others showed some garbling of data, due most likely to a poor Internet connection. These records, nevertheless, were kept in the study, and their available data were used in building the composite formats, since the missing or garbled data did not call into question patterns established in the other examined records. Two printouts turned out to be false matches and were excluded from the study.

Results and Analysis

Composite patterns of MARC field display and labeling for each of the thirty-six catalogs are presented in Appendix A. Comparative analysis of these display formats as representations of ISBD and as surrogates for the familiar LC catalog cards revealed not only a wide variety of features, but also a range of values in five aspects pertinent to record structure: completeness, visual layout, sequence, label specificity, and integration.

Completeness

Each of the record displays studied represents the "full display" as defined for its catalog. How much of the data recorded in the MARC record are actually included? What gets left, so to speak, on the cutting room floor?

Most complete: LC's record displays included every one of the variable fields in the MARC records, including the geographic area code (field 043—labeled GEOG. AREA CODE) and the Dewey classification number (field 082—labeled DEWEY DEC). The DRA version of LC's records omitted only these two fields, while Harvard left out the LCCN as well. The Harvard records were noteworthy for displaying main entries and controlled series statements twice—the second time in a cluster with added entries.

Most abridged: Wake Forest, Minnesota State University System (MSUS), and Carleton stood out slightly from the crowd in this regard. In addition to the half-dozen or so commonly omitted fields, Wake Forest's records did not include bibliography or contents notes or added-entry series titles (fields 504, 505, and 830). MSUS appeared to omit general notes and name-added entries (fields 500, 700, and 710). Carleton also omitted the 504 and 505 note fields as well as all series fields except for 440. Since the MARC records at these institutions were not examined, it is not clear whether variance from the LC master records, rather than programming decisions, caused any of these omissions.

The middle ground: A remarkable consensus regarding inclusion or omission of data elements in a "full" record display is evident in this sample. Of the twenty-two MARC variable fields used in the sample records, seven (fields 100, 245, 260, 300, 440, 600, and 650) are displayed by all libraries, and eight others (111, 500, 504, 505, 700, 710, 800, and 830) are displayed in no fewer than thirty-one of the thirty-six catalogs studied.

Virtually all libraries omit the geographic area code (field 043) and the Dewey classification number (082). A large majority (twenty-four libraries) do not display the Library of Congress Control Number (010). A larger majority (twenty-seven libraries) display the LC classification number (050), not as part of the bibliographic record, but in a separate holdings/location information display. None of these fields, to be sure, are part of the bibliographic description as defined in AACR2 and the ISBDs, but all except the geographic area code (043) have traditionally appeared on catalog cards (at least the widely-distributed LC cards).

Two elements of bibliographic description and a type of heading, however, were found to be omitted from the OPAC displays of many catalogs. The generic uniform title (field 240) did not appear in twenty-nine libraries' displays of the record that included it (Witter Bynner's *Selected Letters*). Nineteen libraries (more than half of the sample) did not display the ISBN (field 020) in their records. Eleven (nearly one-third of the sample) omitted the uncontrolled series statement (field 490) from their displays, and one library (SMU) omitted it from the sample record, which also contained field 830, while displaying it in the record containing a field 800 (which in turn was omitted from display). In general, the OPACs in the study tended to display the full complement of bibliographic description elements and access points while omitting fields considered redundant and less informative (e.g., artificial titles and uncontrolled series statements), as well as control numbers such as the ISBN.

Visual Layout

The extent to which bibliographic displays resemble the layout of LC catalog cards not only has psychological implications (expressing continuity or a break with the past), it also reflects vendor or local-library decisions about record structure and the interrelationship of data elements.

Most cardlike: The hands-down "winner" in this category was Ohio State's LCS system, in which the bibliographic description was presented in two unlabeled, paragraph-style blocks (but without much of the ISBD punctuation that separates the data elements; only that punctuation input in the MARC record was included). Above these were the LC call number on one line and the author main entry (if applicable) on the next, both also unlabeled. Below, however, subjects, name-added entries, and control numbers were discreetly labeled.

The Library of Congress system presented the main entry, title and publisher statements, and physical description in the familiar card style, but other data elements were labeled. Arizona State and Hawaii, which used the CARL system, presented the publisher statement, physical description, series statement, notes, and a control number in an unlabeled block of separate lines (one field per line).

Most listlike: The dominant visual paradigm in this sample is the two-column list or table of attributes, with labels on the left and the corresponding data elements on the right. All catalogs in this sample, except those at Ohio State and (technically, at least) the Library of Congress, followed this paradigm.[33] (The CARL libraries, Arizona State and Hawaii, followed it with minimal use of labels.) Variations within this paradigm involved (1) the use of spacing between lines or groups of lines, (2) repetition or nonrepetition of labels, and (3) clustering of similar fields in paragraph-style blocks. The most listlike displays presented individual fields on separate lines and used either no spacing between lines or spacing between each field or group of fields bearing the same label. Catalogs with such displays were found at Yale, Marquette, Nebraska, Washington & Lee, MSUS, Carleton, Kansas, and New Brunswick (no spacing); Auburn, SUNY, Alabama, and Wake Forest (spacing between each field or labeled group).

The middle ground: The majority of catalogs in the study, while adopting the two-column tabular format for their records, gave a small nod to the card style of display by using line spacing selectively to create logical "clusters" of data categories. The configuration of these clusters varied widely from catalog to catalog but often would include title/author clustering (citation block), publication/physical description clustering (item information block), or subjects/other authors clustering (tracings block). In another nod to card style, the Geac catalogs in the study (Rutgers, Rochester) used a modified paragraph style in displaying multiple fields beside a label, with spacing and an asterisk (*) to separate individual fields. Similarly, MELVYL displayed the title, statement of responsibility, and publishing information (fields 245 and 260) in a single citation-style paragraph beside the label "Title:".

Sequence of Data

Card-catalog records created following AACR2 and long-established conventions for placement of headings

and tracings show a standard sequence of data elements which is largely followed in the variable fields of the USMARC formats. (On an LC card, the standard sequence also encompasses assigned classification and control numbers.) Online technology, however, enables libraries to present the elements of a catalog record in any order.

Most cardlike: None of the displays studied followed completely the sequence of data elements one would find on an LC card (call number, author main entry, title statement, publishing statement, physical description, series statement, notes, ISBN, subject headings, author tracings, series tracings, utility control number, LCCN). However, nearly half of them came very close. Two of the DRA sites (DRA/LC, Toronto) displayed the ISBN at the end of the record with the other control numbers. Ohio State omitted the ISBN and presented the OCLC number and the LCCN in reverse order. The two DRA Atlas catalogs (Duke, North Carolina), the two CARL sites (Arizona State, Hawaii), and New Brunswick all followed the sequence faithfully but omitted the various control numbers. The MSUS records, in addition, did not include author tracings. MELVYL omitted the control numbers and placed the LC classification number at the end.

Just a step below this level of faithfulness to card data sequence we found a half-dozen catalogs (Alabama, Carleton, Marquette, Northwestern, Washington & Lee, and Yale) where, besides omission or out-of-sequence display of classification and control numbers, the major innovation in display order was to consolidate series information in one place on the record.

Least cardlike: To determine which libraries departed most radically from the standard catalog-card sequence of data display, certain common changes and changes of marginal significance were first noted. Over two-thirds of the catalogs studied consolidated series information, usually by displaying the authority-controlled series heading (either field 440 or 830) and omitting the uncontrolled series statement, but sometimes by displaying all series fields in a single location. Likewise, nearly all libraries in the sample either omitted the ISBN or shifted it away from its catalog-card position immediately following the notes. Sequence shifts of marginal significance included those involving classification and control numbers as well as notes within an unbroken sequence.

Six of the catalogs studied (Auburn, ILLINET, Kansas, Nebraska, SMU, and SUNY) included at least two further shifts in data sequence. Such shifts included displaying all author headings (main entry and tracings) in a single sequence (Auburn, Nebraska, SMU, SUNY); placing the title statement ahead of the author main entry (Nebraska, SMU, SUNY); moving a generic uniform title ahead of the author main entry (Nebraska, SUNY); displaying the publishing date beside its own label, below the physical-description field (ILLINET); moving subjects and author tracings ahead of the physical description (Kansas); putting the subject headings ahead of the notes (Auburn, ILLINET, Kansas, SUNY); and placing series information between the notes and the subjects (Nebraska).

The middle ground: As noted above, most of the catalogs studied displayed series information (statements and/or tracings) in a single location or sequence, and those displaying classification or control numbers within the record tended not to place them in a sequence analogous to their positions on a catalog card. Fourteen of the catalogs made one further shift in data sequence but otherwise retained the standard catalog-card order. Here, the more popular shifts were subject and author tracings ahead of notes (Brigham Young, California, Toledo); author tracings ahead of subjects (Rochester, Rutgers, York); author tracings ahead of series information, notes, and subjects (Indiana, Indiana State); and the contents note ahead of other notes (Harvard, Library of Congress). Other changes included author tracings together with author main entry (Virginia), title portion of author-title series heading (800 $t) immediately following title statement (CLIC), notes ahead of series information (Wake Forest), and author tracings ahead of notes (Vanderbilt).

Despite all the variations in data sequencing, at least one display principle held true for all catalogs: The traditional citation information (author, title, publication), while variously defined, was always displayed together at (or very near) the head of the record.

Level of Detail in Labeling

While no catalog display achieved a strict one-to-one correspondence in mapping either MARC fields or the bibliographic concepts they represent, to labels, some libraries were more likely than others to make fine conceptual distinctions in their data displays. Label specificity in each library's composite display format was measured by comparing the number of labels used to both the number of MARC fields and the number of Distinct Bibliographic Concepts (DBCs) represented by those labels. (DBCs were defined for this study as the smallest nonrepeatable segments of an LC card record filling a function apparent to the practiced catalog user. They are listed in table 1.) Across all catalogs (except Ohio State, for which missing data precluded measurement), the number of MARC fields per label ranged from 1.273 to 2.500, with an average of 1.685; the number of DBCs per label ranged from 1.000 to 2.125, with an average of 1.671. (Tables 2–4 list figures for individual catalogs.)

Table 1
DBCs Found in Sampled Records
(in order of appearance on an LC card)

Concept	MARC Field or Subfield
call number	050
author main entry	100, 111
uniform title	240
title	245 $a
subtitle	245 $b
statement of responsibility	245 $c
place of publication	260 $a
publisher	260 $b
date of publication	260 $c, fixed field
extent	300 $a
illustration statement	300 $b
size	300 $c
series statement	440, 490
notes	500, 504
contents note	505
ISBN	020
subject tracings	600, 650
author tracings	700, 710
series tracings	800, 830
alternate call number	082
LCCN	010

Most detailed: The lowest fields-to-labels and concepts-to-labels ratios in this study tended to be associated with factors such as display of control numbers, call numbers, and/or location information (Library of Congress, ILLINET, MSUS, New Brunswick); omission of common descriptive elements (Carleton, Auburn); use of MARC tags and indicators as labels (Carleton); and unlabeled display of some elements (Library of Congress). The catalog which gave the more commonly displayed elements the most detailed labeling was that of Washington & Lee, which provided separate labels for series-statement and series-tracing fields, as well as a separate label for bibliography notes. ILLINET deserves mention as well for giving the date of publication its own label.

Most compressed: The highest fields-to-labels ratios (and second-highest DBCs-to-labels) belonged to the CARL catalogs of Arizona State and Hawaii, where most of the bibliographic description was unlabeled and all the tracings fields received the single label OTHER ENTRIES. Other high-scoring catalogs in this area displayed main- and added-entry authors beside a single label (Nebraska, Virginia); title statement and uniform title beside a single label (Nebraska); title statement and publishing information beside a single label (MELVYL); physical description and notes other than the contents note beside a single label (Virginia); the contents note with other notes beside a single label (MELVYL, Kansas); and author and series tracings beside a single label (MELVYL).

The middle ground: Again, a broad consensus on label definition emerged from this study. Of the thirty-six catalogs examined

- thirty-one treated the title statement (MARC field 245) as a single, complete labeled category.
- thirty treated publishing information (field 260) as a single, complete labeled category.
- thirty treated physical-description information (field 300) as a single, complete labeled category.
- thirty-four treated subject headings collectively as a distinctive labeled category.
- thirty-four assigned the same labels to personal and corporate author headings.
- thirty-one gave separate labels to main-entry and added-entry author headings.
- thirty-one gave separate labels to author and series headings.
- twenty used a single label for all series data displayed.
- twenty-seven assigned general and bibliography notes the same label.
- twenty-six treated the contents note (field 505) as a separate labeled category.

Twelve catalogs (including nine of the fourteen NOTIS sites) shared all ten of the above characteristics.

Relationship of Description and Access Points

One of the basic principles underlying ISBD (and therefore AACR2) is the functional separation of description from access. With two significant exceptions (the title proper and—conditionally—the series statement), the components of an ISBD bibliographic description do not serve as index entries or "access points" in a card catalog. Rather, data from the description are reformulated into standardized "headings" that facilitate searching

Table 2
Ratios of MARC Fields to Labels and DBCs to Labels by Catalog (NOTIS sites)

Library	Labels	MARC Fields	DBCs	Fields per Label	Concepts per Label
Auburn	10	16	16	1.600	1.600
Brigham Young	9	15	16	1.667	1.778
Indiana State	11	17	17	1.545	1.545
Indiana U.	11	17	17	1.545	1.545
Northwestern	10	17	17	1.700	1.700
SMU	10	16	17	1.600	1.700
SUNY	10	16	17	1.600	1.700
U/Alabama	10	16	17	1.600	1.700
UC-Santa Barbara	9	15	15	1.667	1.667
U/Toledo	9	15	15	1.667	1.667
U/Virginia	8	16	17	2.000	2.125
Vanderbilt	10	16	16	1.600	1.600
Yale	9	15	16	1.667	1.778
York	10	17	17	1.700	1.700

Table 3
Ratios of MARC Fields to Labels and DBCs to Labels by Catalog (DRA, Innopac, CARL, Dynix, Geac, and PALS sites)

Library or Consortium	Labels	MARC Fields	DBCs	Fields per Label	Concepts per Label
DRA/LC	13*	20	20	1.538	1.538
Duke	10*	17	17	1.700	1.700
North Carolina	9*	16	16	1.778	1.778
Toronto	12*	17	18	1.417	1.500
Marquette	9	16	16	1.778	1.778
Nebraska	8	17	17	2.125	2.125
Washington & Lee	11	16	16	1.455	1.455
Arizona State	4	10†	8†	2.500	2.000
Hawaii	4	10†	8†	2.500	2.000
CLIC	11	16	17	1.455	1.545
Wake Forest	8	13	13	1.625	1.625
Rutgers	11*	17	18	1.545	1.636
U/Rochester	11*	17	17	1.545	1.545
MSUS	10*	14	16	1.400	1.600

*Not including labels pointing only to nonDBC data.
†Not including nonlabeled data.

Table 4
Ratios of MARC Fields to Labels and DBCs to Labels by Catalog (In-House Systems)

Library or Consortium	Labels	MARC Fields	DBCs	Fields per Label	Concepts per Label
Carleton	11	14	13	1.273	1.182
Harvard	12	19	19	1.583	1.583
ILLINET	11	16	17	1.455	1.545
Library of Congress	10‡	15*†	10	1.500	1.000
MELVYL	9	19	18	2.111	2.000
Ohio State	5‡	XXXXXX	XXXX	XXXXXXXX	XXXXXXXXX
U/Kansas	8	16	16	2.000	2.000
U/New Brunswick	11	17	17	1.545	1.545

*Not including labels pointing only to nonDBC data.
†Not including fields for nonDBC data.
‡Not including nonlabeled data.
Ohio State not included due to missing data.

and collocation by names of persons, organizations, events, and works associated with bibliographic items. In AACR2, descriptive data and access points are defined under separate sets of rules, the former emphasizing transcription of information from the item itself, the latter a kind of vocabulary control (commonly referred to as "authority control").

In the card environment, the distinction between description and access data is, for the most part, readily visible. On catalog cards such as those printed by the Library of Congress and OCLC, the main-entry heading (either an authority-controled author heading or the title proper) appears at the beginning of the description. All other access points, including subject headings, are recorded in a separate block (known as the "tracings") beyond the end of the description, primarily as a kind of inventory control for the multiple copies of a record in the catalog (the card set).

The MARC format, however, treats elements of description and headings/tracings as so many pieces of data within a single record. This appears to have encouraged online catalog designers to regard headings/tracings information (properly labeled, of course) as part of the description. The present study examined the various catalogs to gauge the extent to which authority-controled data have been integrated into the bibliographic description presented to the user. Aspects of "record integration" found include the positioning of headings/tracings information between elements of the description, the display of author-added entries adjacent to the author main-entry heading, the mapping of description and access fields to the same label, and the suppression of data considered redundant in a unified record (either a heading/tracing or the descriptive data justifying a heading/tracing).

Most separated: None of the above-mentioned aspects of description/access integration were found in the following eight catalogs: DRA/LC, the two DRA Atlas sites (Duke and North Carolina), the two CARL sites (Arizona State and Hawaii), MELVYL, Ohio State, and New Brunswick. In all, the data sequence traditionally found on cards was preserved. In the CARL and Ohio State catalogs, the separation was somewhat more visible in that most of the description was unlabeled, and two or all three of the categories of headings shared a single label. In most of the catalogs, the wording of the labels (e.g., ADDED ENTRY, OTHER ENTRIES) emphasized the function of the data as providing access rather than description.

Most integrated: A half-dozen catalogs (Auburn, SMU, SUNY, California, Toledo, and Nebraska) exhibited four or more significant instances of description/access data integration, some of the more unusual of which include: bringing all author headings/tracings together (Auburn, SMU, SUNY, Nebraska); shifting the author main-entry heading to a position following the title statement (SMU, SUNY, Nebraska); displaying an author/uniform title main-entry heading as two separate, unconnected pieces of data (SUNY, Nebraska); and suppressing the "statement of responsibility" (MARC 245 $c) from display (Auburn, California, Toledo).

The middle ground: A more limited integration of

headings/tracings data into the description characterized the majority of catalogs in this study. Common display practices of this sort included

- displaying series tracings ahead of notes (sixteen catalogs).
- defining a single label for series information (16).
- displaying tracings ahead of the ISBN (12).
- suppressing field 490 (non-traced series statement) from display (11).
- displaying author tracings ahead of notes, but still separately from the author main-entry heading (9).
- displaying subject tracings ahead of notes (7).

In general, integration of series-data display and repositioning of tracings to display ahead of some parts of the description are fairly common in this sample, integration of author data much less so.

Discussion

If nothing else, this "Summer of '94" group portrait of bibliographic display practices reveals a decisive break with card-style ISBD displays on the part of all thirty-six catalogs depicted. Every catalog labels at least some of its data, and even within unlabeled blocks of data the ISBD punctuation separating one field from another is missing. ISBD punctuation within fields appears in all catalogs, but only, it would seem, because catalogers following AACR2 continue to input such punctuation in MARC records. For the most part, the functions of ISBD punctuation have been taken over by line breaks and labels.

In some catalogs, however, this break with the past has been more radical than in others. While most use a table format of data elements with labels, a few still employ a paragraph/citation style with a separate tracings block reminiscent of catalog cards. While most catalogs have taken advantage of the opportunities online technology affords to suppress some data from display and rearrange the rest (either to focus the user's attention on certain data elements or to make name and subject headings more fully a part of the bibliographic description), relatively few have done so to more than a limited extent. While consolidation of series data display is relatively common, consolidation of author data is much less so. Most systems suppress control numbers from display, while few suppress notes.

Even so, these thirty-six catalogs show a remarkable variety in their record display formats. Aside from those of Indiana and Indiana State, no two composite displays are exactly alike. This pair, and the two sets of near-identical twins in the study (Arizona State and Hawaii, and Duke and North Carolina), probably owe their unusual level of similarity to, in varying degrees, common system software and consortial relationships. Labeled attribute lists contrast with cardlike and even (in one instance) MARC-like displays. A large number of display features are unique to one catalog, and most (if not all) catalogs can claim at least one feature unique within the study sample, highlighting both the immense range of possibilities for record display and a widespread readiness to experiment.

At the same time, a broad consensus on many display issues is evident, one which reflects a fairly conservative approach to data suppression and rearrangement, to consolidation and rationalization of data categories, to integration of description and access points. This probably reflects an often-noted time lag between introduction of a technology and realization of its full potential, as a recent paper on a related topic has noted: "Interface developers have tended to develop products that are based on the way we use documentation in print . . . . The situation is analogous to the early days of the automobile when the body was made to look like a buggy while demands for new technology specific to motoring had to emerge slowly with experience."[34]

In the online environment, there are several factors causing change in the presentation of bibliographic records and creating a need for change in how these records are created. The fragmentation of the record through the MARC format enables record elements to be rearranged and displayed selectively, thus calling into question rules that produce context-sensitive data, such as the uniform titles in MARC field 240 and abbreviated names of publishers named fully elsewhere in the record. The interchangeability of descriptive data elements and access points (since each can be made to serve both functions online) makes the separate creation of description and headings seem pointless and burdensome. Labeling of data elements (made possible through the mapping of terms to MARC fields) creates a need for simpler, less ambiguous bibliographic data definitions than are appropriate for the dense and context-rich narrative-style records catalogers continue to create. System support for linkages among records of different types—or of the same type—creates demands both for new types of data (e.g., tables of contents, book reviews, records for parts of an item) and for creative uses of existing data (e.g., links to authority records, multiple-version records, hypertext links).

At the same time, catalogers no longer determine the forms the records they create will assume. When names of editors appear beside the label "Author:" or title-page statements of responsibility are suppressed from display, that reflects policy and programming decisions generally made outside cataloging units, and

often (anecdotal evidence suggests) without prior input from technical services librarians. If outside access to catalogs through Z39.50 browsers becomes commonplace, control over certain aspects of record content will, in many cases, rest with either the end user or the remote accessing library, through the choice or programming of the browser software. It can be argued that the role of catalogers has become more restricted than in the past, that while they still collect the data elements and give them initial organization, final organization and interpretation are now performed elsewhere. But that final editing is essentially a process of machine translation, with all the potential for degradation of meaning that that phrase implies.

Together, these factors suggest that cataloging standards will need to be rewritten in order to provide the kind of data flexibility expected in online systems. The central requirement has already been identified by Wool et al.: "records flexible enough to be added to, subtracted from, and rearranged without loss and garbling of meaning. What is needed is a modular record structure, in which every segment of data can stand on its own with appropriate labeling and which can support all possible display lengths and combinations of data elements."[35] Besides a revamping of the cataloging rules (and perhaps of MARC data definitions), the path to this goal could well take in such proposed changes as integration of rules and data format, linkages between bibliographic and authority records, conscious application of the relational database model to cataloging standards, and use of SGML (standard generalized markup language) to generate and organize data. The requirements of database design, the preferences of catalog users, the needs of shared cataloging and international record exchange, and especially the undiminished importance of data integrity and record quality should all be taken into account.

Even so, the most logical approach may well be through the current effort to define a common metadata standard for all types of information providers.[36] After all, as one of its participants, Priscilla Caplan notes: "Metadata really is nothing more than data about data; a catalog record is metadata; so is a TEI header, or any other form of description."[37] The growing tendency to present the catalog record as a labeled list of attributes makes this approach especially attractive.

Research Suggestions

It seems safe to assume that in the next few years, OPAC record displays will less and less resemble card records, not only visually, but in structure and content as well. Lagging behind this development, but just as inevitable, will be changes in cataloging practice and ultimately in cataloging standards to reflect the modular nature of bibliographic records online. How these changes are managed will have a long-term impact on the integrity, and thus the usefulness, of catalog information (a major investment, unavoidably, on the part of libraries). The knowledge needed to make the transition successfully will come from several kinds of research:

- comparative studies of display design and effectiveness, including user studies
- continued monitoring of record-display practices
- studies of staff and public use of catalog information
- analysis of the underlying structure of bibliographic information, and of its various representations
- comparisons of cataloging practice (including copyediting) among various institutions subscribing to a standard
- evaluation of new ways to store, access, and present bibliographic data

More work is needed in all of these areas in order to inform the judgment of those who will determine (whether by action or reaction) the future of library catalogs.

Notes

1. Gregory J. Wool and others, "Cataloging Standards and Machine Translation: A Study of Reformatted ISBD Records in an Online Catalog," *Information Technology and Libraries* 12 (Dec. 1993): 383–403.

2. "The OCLC Online Union Catalog," in *Furthering Access to the World's Information: OCLC Annual Report 1993–94* (available on the World Wide Web at http://www.oclc.org/oclc/ar94/oluc.htm).

3. Walt Crawford, Lennie Stovel, and Kathleen Bales, *Bibliographic Displays in the Online Catalog* (White Plains, N.Y.: Knowledge Industry, 1986). A methodological introduction to the study can be found in Walt Crawford, "Testing Bibliographic Displays for Online Catalogs," *Information Technology and Libraries* 6 (Mar. 1987): 20–33.

4. Wool and others, "Cataloging Standards," 385, 399.

5. Ibid., 385–86, 399–400.

6. Ibid., 386, 400.

7. Anthony G. Curwen, "International Standard Bibliographic Description," in *Standards for the International Exchange of Bibliographic Information: Papers Presented at a Course Held at the School of Library, Archive and Information Studies, University College London, 3–18 August 1990*, ed. I. C. McIlwaine (London: The Library Association, 1991), 73–81.

8. Curwen, "International Standard Bibliographic Description," 80.

9. Rodney M. Brunt, "The Code and the Catalogue: A Return to Compatiblity," *Library Review* 41, no. 3 (1992): 22–32.

10. Nancy Lee Shires and Lydia P. Olszak, "What Our Screens Should Look Like: An Introduction to Effective OPAC Screens," *RQ* 31 (Spring 1992): 357–69.

11. Ann O'Brien, "Online Catalogs: Enhancements and Developments," in *Annual Review of Information Science and Technology*, vol. 29, ed. Martha E. Williams (Medford, N.J.: Learned Information, 1994), 220–21.

12. Keith Pickens, "The Relationship of Bibliographic Database Design to the Structure of Information: A Case Study in Education," *Journal of Documentation* 50 (Mar. 1994): 36–44.

13. Ibid., 42.

14. Leslie Troutman, "The Online Public Access Catalog and Music Materials: Issues for System and Interface Design," in *Advances in Online Public Access Catalogs*, vol. 1 (Westport, Conn.: Meckler, 1992), 20–21.

15. E. Paige Weston, "An OPAC for Every Public: Customizing the Catalog for Individual Researchers' Needs," in *Academic Libraries: Achieving Excellence in Higher Education: Proceedings of the Sixth National Conference of the Association of College and Research Libraries*, ed. Thomas Kirk (Chicago: Assn. of College and Research Libraries, 1992), 377–83.

16. F. H. Ayres, "Duplicates and Other Manifestations: A New Approach to the Presentation of Bibliographic Information," *Journal of Librarianship* 22 (Oct. 1990): 236–51; Yasmin S. Warrick, "The Use of Abbreviations in Area V of AACR2," *Technicalities* 14 (Mar. 1994): 11–13; Frederick G. Kilgour, "Effectiveness of Surname-Title-Words Searches by Scholars," *Journal of the American Society for Information Science* 46 (Mar. 1995): 146–51; Barbara M. Wildemuth and Ann L. O'Neill, "The 'Known' in Known-item Searches: Empirical Support for User-Centered Design," *College & Research Libraries* 56 (May 1995): 265–81.

17. R. Conrad Winke, "Discarding the Main Entry in an Online Cataloging Environment," *Cataloging & Classification Quarterly* 16 (1993): 53–70; Esther Green Bierbaum, "A Modest Proposal: No More Main Entry," *American Libraries* 25 (Jan. 1994): 81–84.

18. Carol Mandel, "Cataloging for Access," *The Reference Librarian*, no. 34 (1991): 61–68.

19. Ibid., 62.

20. Alan Jeffreys, "AACR after 1978," in *AACR, DDC, MARC and Friends: The Role of CIG in Bibliographic Control*, ed. John Byford, Keith V. Trickey, and Susi Woodhouse (London: Library Association Publishing, 1993), 58; Ross Bourne, "MARC: Strait-Jacket or Opportunity," in *AACR, DDC*, 82.

21. Gregory H. Leazer, "An Examination of Data Elements for Bibliographic Description: Toward a Conceptual Schema for the USMARC Formats," *Library Resources & Technical Services* 36 (Apr. 1992): 189–208.

22. Michael Heaney, "Object-Oriented Cataloging," *Information Technology and Libraries* 14 (Sept. 1995): 135–53.

23. Charles R. Hildreth, *Online Public Access Catalogs: The User Interface* (Dublin, Ohio: OCLC, 1982); Stephen R. Salmon, "Characteristics of Online Public Catalogs," *Library Resources & Technical Services* 27 (Jan./Mar. 1983): 36–67.

24. Hildreth, *Online Public Access Catalogs*, 144–58.

25. Salmon, "Characteristics," 54.

26. Joan M. Cherry and others, "OPACs in Twelve Canadian Academic Libraries: An Evaluation of Functional Capabilities and Interface Features," *Information Technology and Libraries* 13 (Sept. 1994): 174–95.

27. Walt Crawford, *The Online Catalog Book: Essays and Examples* (New York: G. K. Hall; Toronto: Maxwell Macmillan Canada, 1992).

28. AUTOCAT: Library Cataloging and Authorities Discussion Group [online]. Available e-mail: listserv@ubvm.cc.buffalo.edu.

29. Karl E. Johnson, "MARC Meets the OPAC and Loses," *Technicalities* 13 (Mar. 1993): 11–12.

30. "AACR2000: Toward the Future of the Descriptive Cataloging Rules," preconference institute sponsored by the Association for Library Collections and Technical Services, June 22, 1995, Chicago, Ill. Presenters were Clifford Lynch, David Levy, Crystal Graham, Steven Hensen, Martha Yee, Barbara Tillett, Michael Gorman, and Lynne Howarth.

31. Despite the widespread assumption that libraries add LC records to their catalogs with little or no editing, the records in this study contain numerous variations from the MARC records found on OCLC. Two hundred eleven MARC variable fields (out of a total of 1,996, or 10.6 percent) were either displayed in altered form in the OPACs studied, or were omitted altogether in ways not readily attributable to the display software. (The local MARC records themselves were not examined.) Ninety-seven of 165 records (from the 34 non-LC catalogs, excluding five spoiled printouts) contained at least one apparent editing change. It appears likely that two of the LC records originated as CIP copy, and that a number of the libraries in the study took in the CIP version and never updated it. Since the possible editing of LC copy was not being examined, however, further data pertaining to it were not gathered. The lack of such data affects the present study in minor ways; for instance, label sequence in some catalogs, especially among the notes fields, could not always be fully documented.

32. HYTELNET gopher version 2.0.14, Northern Lights, Saskatoon, Saskatchewan, Canada. Developed by Peter Scott. URL http://library.usask.ca/hytelnet/

33. The layout of the LC records differs from the two-column model in that most labels appear above, rather than beside, the data they refer to. However, since the labels are left-justified, and the data are indented slightly and run (in most cases) well to the right of where the labels end, that distinction may not be readily apparent to most people.

34. Steven R. Howe and Robert J. Graham, "Metadata and User Interfaces: Promises and Problems," *IASSIST Quarterly* 17 (Spring/Summer 1993): 4.

35. Wool and others, "Cataloging Standards," 397–98.

36. Priscilla Caplan, "You Call It Corn, We Call It Syntax-independent Metadata for Document-Like Objects," *The Public-Access Computer Systems Review* 6, no. 4 (1995). URL http://info.lib.uh.edu/pr/v6/n4/capl6n4.html.

37. Caplan, "You Call It Corn," paragraph 3.

Appendix A
Composite Bibliographic Display Formats Compiled from Records for Five Monographs (figures 1–5) (August 1994)

** indicates blank line above label
& indicates label is repeated for multiple fields
# indicates numbering of labeled fields (1, 2, etc.)
MARC tags in brackets are displayed paragraph style

Table A-1
Auburn University (NOTIS)

Label	MARC Tags
Author:	100, 111
**Contributors:	700, 710
**Title:	245 (no $c)
**Publication Information:	260
**Description:	300
**Subjects:	600, 650
**Notes:	500, 504
[or:] Contents:	505
**Series:	440 (no $x), 800, 830
**ISBN:	020

Table A-2
Brigham Young University (NOTIS)

Label	MARC Tags
Author:	100, 111
**Title:	245
**Published:	260
Description:	300
**Subjects:	600, 650
**Other authors:	700, 710
**Notes:	500, 504
**Contents:	505
**Series:	440 (no $x), 800, 830

Table A-3
Indiana State University (NOTIS)

Label	MARC Tags
Author:	100, 111
Title:	245
Published:	260
Description:	300
Other contributors:	700, 710
Series:	440 (no $x), 490, 800
Uniform series:	830
Notes:	500, 504
**Contents:	505
Subject headings:	600, 650
ISBN:	020 (no $c)

Table A-4
Indiana University at Bloomington (NOTIS)

Label	MARC Tags
Author:	100, 111
Title:	245
Published:	260
Description:	300
Other contributors:	700, 710
Series:	440 (no $x), 490, 800
Uniform series:	830
Notes:	500, 504
**Contents:	505
Subject headings:	600, 650
ISBN:	020 (no $c)

Table A-5
Northwestern University (NOTIS)

Label	MARC Tags
Author, etc.:	100, 111
Title:	245
Published:	260
Description:	300, 490
"Series:	440 (no $x), 800 (altered, no $f), 830
"Notes:	500, 504
"Contents:	505
"Library of Congress Subject Headings:	600, 650
"Other authors, etc.:	700, 710
"ISBN:	020 (no $c)

Table A-6
Southern Methodist University (SMU) (NOTIS)

Label	MARC Tags
TITLE:	245
AUTHOR:	100, 111
OTHER AUTHORS:	700, 710
"PUBLISHER:	260
DESCRIPTION:	300
ISBN:	020 (no $c)
SERIES:	440 (no $x), 490, 830
"CONTENTS:	505
NOTES:	500, 504
"SUBJECTS:	600, 650

Table A-7
State University of New York (SUNY) at Buffalo (NOTIS)

Label	MARC Tags
TITLE:	245
"OTHER TITLES:	240
"AUTHOR:	100, 111
"CONTRIBUTORS:	700, 710
"PUBLISHED:	260
"DESCRIPTION:	300
"SUBJECTS (SL=):	600, 650
"SERIES:	440 (no $x), 800, 830
"NOTES:	500, 504
"CONTENTS:	505

Table A-8
University of Alabama (NOTIS)

Label	MARC Tags
Author:	100, 111
"Title:	245
"Publisher:	260
"Description:	300
"Series:	440 (no $x), 800, 830
"Notes:	500, 504
[or:] Contents:	505
"Subject headings:	600, 650
"Other contributors:	700, 710
"ISBN:	020 (no $c)

Table A-9
University of California at Santa Barbara (NOTIS)

Label	MARC Tags
Author:	100, 111
**Title:	245 (no $c)
**Published:	260
Description:	300
**Subjects:	600, 650
**Other authors:	700, 710
**Notes:	500, 504
**Contents:	505
**Series:	440 (no $x), 800, 830

Table A-10
University of Toledo (NOTIS)

Label	MARC Tags
Author:	100, 111
**Title:	245 (no $c)
**Published:	260
Description:	300
**Subjects:	600, 650
**Other authors:	700, 710
**Notes:	500, 504
**Contents:	505
**Series:	440 (no $x), 800, 830

Table A-11
University of Virginia (NOTIS)

Label	MARC Tags
AUTHOR	100, 111, 700, 710
**TITLE	245
**PUBLISHED	260
**DESCRIPTION	300, 500, 504
**CONTENTS	505
**SUBJECT TERMS	600, 650
**SERIES	440 (no $x), 800, 830
ISBN	020 (no $c)

Table A-12
Vanderbilt University (NOTIS)

Label	MARC Tags
Author:	100, 111
Title:	245
Publisher:	260
Description:	300
**Series searchable as:	440 (no $x), 800, 830
**Series:	490
**Other authors:	700, 710
**Notes:	500, 504
[or:] Contents:	505
**LC subjects:	600, 650

Table A-13
Yale University (NOTIS)

Label	MARC Tags
Author:	100, 111
Title:	245
Published:	260
Description:	300
Series:	440 (no $x), 800, 830
Notes:	500, 504
Contents:	505
Subjects (Library of Congress):	600, 650
Also listed under:	700, 710

Table A-14
York University (NOTIS)

Label	MARC Tags
Author:	100, 111
**Title:	245
**Published:	260
Description:	300
**Series:	440 (no $x), 490, 800 (altered, no $f), 830
**Notes:	500, 504
**Contents note:	505
**ISBN:	020 (no $c)
**Other access points:	700, 710
**Subject headings: [SU=]	600, 650

Table A-15
Data Research Associates/Library of Congress (DRA)

Label	MARC Tags
Type of material:	fixed field code
**LC Call Number:	050
**Author:	100, 111
**Generic title:	240
Title:	245
**Publication Info:	260
Phys. Description:	300
Series Name:	440, 490
**&Notes:	500, 504, 505
**&Subjects:	600, 650
**&Other Names:	700, 710
Other Series Names:	800, 830
**LC Card Number:	010
ISBN:	020

Table A-16
Duke University (DRA Atlas) (does not include Davis record)

Label	MARC Tags
MATERIAL:	fixed field code
AUTHOR:	100, 111
**GENERIC TITLE:	240
TITLE:	245
**PUBLICATION:	260
DESCRIPTION:	300
SERIES:	440, 490
**&NOTES:	500, 504, 505
**&SUBJECT:	600, 650
**&ADDED ENTRY:	700, 710
SERIES ADDED ENTRY:	800 (no $f), 830

Table A-17
University of North Carolina (DRA Atlas)

Label	MARC Tags
MATERIAL:	fixed field code
AUTHOR:	100, 111
**TITLE:	245
**PUBLICATION:	260
DESCRIPTION:	300
SERIES:	440, 490
**&NOTES:	500, 504, 505
**&SUBJECT:	600, 650
**&ADDED ENTRY:	700, 710
SERIES ADDED ENTRY:	800, 830

Table A-18
University of Toronto (DRA Information Gateway)

Label	MARC Tags
SHELF NUMBER:	050, local info.
**AUTHOR:	100, 111
**TITLE:	245
PUBLISHED:	260
PAGING:	300
SERIES:	440, 490
**NOTES:	500, 504
**CONTENTS:	505
**SUBJECTS:	600, 650
OTHER ENTRY:	700, 710
LCCN:	010
OCLC #:	[from fixed field]
ISBN:	020

Table A-19
Marquette University (INNOPAC)

Label	MARC Tags
AUTHOR	100, 111
TITLE	245
PUB/DATE	260
DESCRIPT.	300
SERIES	440 (no $x), 490, 800, 830
NOTE	500, 504
CONTENTS	505
SUBJECT	600, 650 (order varies)
ADDED AUTH	700, 710

Table A-20
University of Nebraska at Lincoln (INNOPAC)

Label	MARC Tags
TITLE	240, 245
AUTHOR	100, 111, 700, 710
PUBLISHER	260
DESCRIPT.	300
NOTE	504, 500
[or:] CONTENTS	505
SERIES	440, 490, 800, 830
SUBJECT	600, 650

Table A-21
Washington & Lee University (INNOPAC)

Label	MARC Tags
AUTHOR	100, 111
TITLE	245
PUBLISHER	260
DESCRIPT	300
SERIES	440~, 490
ALT SERIES	440~, 800, 830
BIBLIOG.	504
NOTE	500
[or:] CONTENTS	505
SUBJECT	600, 650
ADD AUTHOR	700, 710

~Displayed beside both SERIES and ALT SERIES on the Hoffman record, beside SERIES only on the Bieberbach record

Table A-22
Arizona State University (CARL)

Label	MARC Tags
AUTHOR(s):	100, 111
TITLE(s):	245
**[unlabeled area]	260, 300, 440, 490, 500, 504, local field
CONTENTS:	505
**OTHER ENTRIES:	600, 650, 700, 710, 800, 830

Table A-23
University of Hawaii at Manoa (CARL)

Label	MARC Tags
AUTHOR(s):	100, 111
TITLE(s):	245
**[unlabeled area]	260, 300, 440, 490, 500, 504
Contents:	505
**OTHER ENTRIES:	600, 650, 700, 710, 800, 830 (altered, no $f)

Table A-24
CLIC—Cooperating Libraries in Consortium (Minnesota) (DYNIX)

Label	MARC Tags
#AUTHOR	100, 111
**TITLE	245
TITLE VAR	800 $t
**PUBLISHER	260
**DESCRIPT	300
**#SERIES	440 (no $x), 490, 800 ($a and $d), 830
#NOTES	500
**ISBN	020 (no $c)
**#SUBJECT(S)	600, 650
**#ADD AUTHOR	700, 710
**DYNIX/LCCN	local info., 010

Table A-25
Wake Forest University (DYNIX)

Label	MARC Tags
#AUTHOR	100, 111
**TITLE	245 (no $c)
**PUBLISHER	260 (no $a)
**COLLATION	300
**NOTES	500
**#SERIES	440, 490, 800 (no $f)
**#SUBJECTS	600, 650
**#ADDED AUTH	700, 710

Table A-26
Rutgers University (Geac)

Label	MARC Tags
AUTHOR	100
TITLE	245
PUBLISHER	260
DESCRIP	300
SERIES	{440, 490, 800, 830}
NOTES	{500, 504}
CONTENTS	505
OTHER AUT	{700, 710}
SUBJECTS	{600, 650}
LC CARD	010
ISBN	020 (no $c)
RLIN ID	[from fixed field]

Table A-27
University of Rochester (Geac)

Label	MARC Tags
AUTHOR:	100
[or:] CONFERENCE AUTHOR:	111
TITLE:	245
IMPRINT:	260
PHYSICAL DESCRIPTION:	300
SERIES:	{440 (no $x), 490, 800, 830}
NOTES:	{504, 500}
CONTENTS:	505
OTHER AUTHORS:	{700, 710}
SUBJECTS:	{600, 650}
ISBN:	020 (no $c)
OCLC NUMBER:	[from fixed field]

Table A-28
Minnesota State University System (PALS)

Label	MARC Tags
LOCATION:	{local info., 050, 092}
LIBRARIES:	local info.
AUTHOR:	100, 111
TITLE:	245
PUBLISHER:	260
DESCRIPTN:	300
SERIES:	440
BIBLIOG:	504
CONTENTS:	505
&SUBJECT:	600, 650
SERIES:	800, 830

Table A-29
Carleton University (in-house) (does not include Davis record)

Label	MARC Tags
010@@$a	010
020@@$a	020 (no $c)
0901@$a	090
1@@@@$@	100
24500$a	
[or:] 24510$a	
[or:] 24504$a	245 (no $c)
260@@$a	
[or:] 260@@$@	260 (no $c)
300@@$@	300 $a
5@@@@$@	500
6@@@@$@	600, 650 (order varies)
7@@@@$@	700, 710, 111
84000$t	440
85001$n	local field

Table A-30
Harvard University (in-house)

Label	MARC Tags
AUTHOR:	100, 111
TITLE:	240, 245
PUB. INFO:	260
DESCRIPTION:	300
SERIES:	440, 490
CONTENTS:	505
NOTES:	500, 504
ISBN:	020 (no $c)
**#SUBJECTS:	600, 650
#AUTHORS:	100, 111, 700, 710, 800
OTHER TITLES:	440 (no $x), 830
**LOCATION:	local info., 050

Table A-31
ILLINET [state of Illinois union catalog] (in-house)

Label	MARC Tags
CALL-NO:	local info.
AUTHOR:	100, 111
TITLE:	245
PUBL.:	260 (no $c)
FORMAT:	300
DATE:	fixed field data
OTHER NAME:	700, 710
SERIES:	440 (no $x), 830
**SUBJECT:	600, 650
**ISBN:	020 (no $c)
LANGUAGE:	fixed field code
CONTENTS:	504, 505
NOTES:	500
RID #:	local and fixed-field info.

Table A-32
Library of Congress (in-house)

Label	MARC Tags
[unlabeled area]	100, 111, 240, {245, 260, 300}
**LC CALL NUMBER:	050
**SUBJECTS:	600, 650
**ADDED ENTRIES:	700, 710
**SERIES TITLES (Indexed under SERI option):	440, 490, 800 $t, 830
SERIES (Indexed under AUTH option):	800
**DEWEY DEC:	082
**CONTENTS:	505
**NOTES:	500, 504
**ISBN:	020
GEOG. AREA CODE:	043
LCCN:	010

Table A-33
MELVYL [University of California union catalog] (in-house)

Label	MARC Tags
Author:	100, 111
Uniform title:	240
Title:	{245, 260}
Description:	300
**Series:	440, 490
**Notes:	500, 504, 505
**Subjects:	600, 650
**Other entries:	700, 710, 800, 830
**Call numbers:	local info., 050, 090

Table A-34
Ohio State University (in-house)

Label	MARC Tags
[unlabeled area]	050, 100, 111, {245, 260, 300}, {440, 490, 500, 504}
#SUB:	{600, 650}
#AE:	{700, 800 (no $f)}
	[missing data]
LC CARD #:	010
TITLE #:	local data
OCLC #:	fixed field data

Note: LC CARD #:, TITLE #:, and OCLC #: display sequentially on a single line

Table A-35
University of Kansas (in-house)

Label	MARC Tags
AUTHOR:	100, 111
TITLE:	245
PUBLISHER:	260
SERIES:	440, 800, 830
SUBJECT:	
[or:] #SUBJECTS:	600, 650
OTHER ENTRIES:	700, 710
DESCRIPTION:	300
NOTES:	490*, 500, 504, 505

*Preceded by print constant: "Series statement on item:"

Table A-36
University of New Brunswick (in-house)

Label	MARC Tags
LOCATION:	local info.
CALL #:	050
AUTHOR:	100, 111
TITLE:	245
IMPRINT:	260
COLLATION:	300
SERIES:	440, 490
&NOTES:	500, 504
CONTENTS:	505
&SUBJECTS:	600, 650
&OTHER:	700, 710
SERIES:	800, 830

Index to Advertisers

News and Announcements

Tom Leonhardt Assumes LITA Presidency

At the conclusion of the ALA Annual Conference in New York City in July, Thomas W. Leonhardt, director of technical services at the University of Oklahoma Library, began his term as president of LITA. Leonhardt had served the previous year as vice-president/president-elect of LITA and has held several other positions within the organization, including editor of *ITAL* from 1989 to 1995. He takes over the position from Michelle Newberry, who now assumes the role of immediate past president.

In the election held in the spring, Linda D. Miller, senior automation planning specialist at the Library of Congress, was voted vice-president/president-elect. Miller will assume the LITA presidency at the end of the 1997 ALA Annual Conference in San Francisco. Also elected to the LITA board as directors-at-large were Kate Wakefield, Internet Services coordinator at WLN, and Flo Wilson, associate director at Vanderbilt University. Wakefield and Wilson will serve three-year terms on the LITA board.

LITA Executive Director Named

Jacqueline Mundell has accepted the position of executive director of LITA. She will start officially in October but will be attending September division leadership meetings.

Jacqueline Mundell

Prior to this announcement, Mundell was executive director of the Cleveland Area Metropolitan Library System (CAMLS) located in Cleveland. CAMLS is a consortium of

seventy-nine multitype libraries serving three million patrons in ten Ohio counties. Before working for CAMLS, Mundell was director of NEBASE OCLC Network, in Lincoln, Nebraska.

Mundell holds a bachelor's degree from the University of Colorado-Boulder, an MPA from the University of Nebraska at Omaha, and an MLS from Emporia State University, where she was a member of Beta Phi Mu, an honorary society.

National Digital Library Competition

Ameritech Corporation and its Ameritech Library Services subsidiary have teamed with the Library of Congress to establish a grant program through which selected libraries in the United States can digitize their unique Americana collections for incorporation into the Library's National Digital Library (NDL) program. The Ameritech Foundation will make a $2 million gift to establish the Library of Congress/Ameritech National Digital Library Competition. Complete details on the competition were to be announced in midsummer.

CAUSE '96

"Broadening Our Horizons: Information, Services, Technology" is the theme of the twenty-fifth CAUSE conference being held in San Francisco on December 3–6, 1996. Up-to-date conference information and complete seminar descriptions can be found on the World Wide Web at http://cause-www.colorado.edu/.

Also from CAUSE—Professional Paper no. 14 explores the potential for technology to not only enhance but actually transform the delivery of higher education in the information age. In *Distributed Learning*, authors Diana G. Oblinger and Mark K. Maruyama explain the impetus for changing to a learner-centered paradigm and offer real-world examples of such a change. Funding by IBM enabled publication and distribution of complimentary copies of the paper to more than 3,500 CAUSE member representatives. Copies are

available to anyone on a CAUSE-member campus for $16 per copy or to nonmembers for $32. To order, call (303) 939-0310, or send e-mail to orders@cause.colorado.edu.

Preserving Digital Information

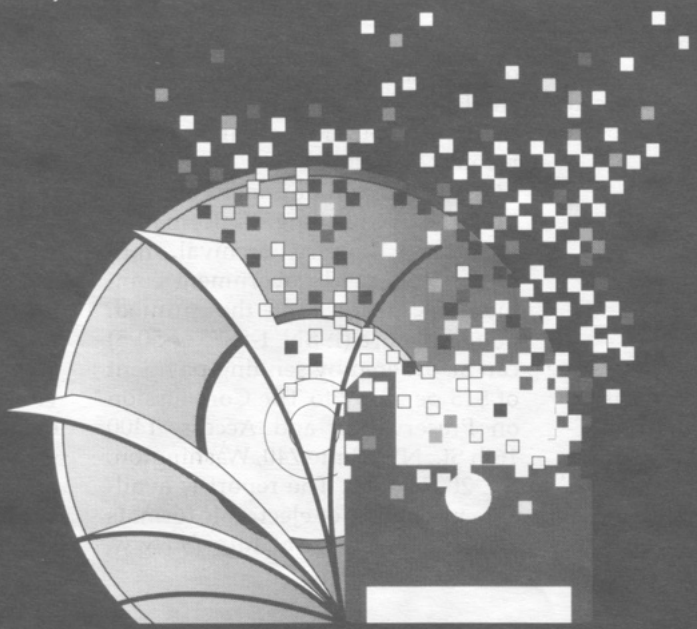
The Commission on Preservation and Access and the Research Libraries Group have issued *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* by Donald Waters and John Garrett, cochairs of the task force. This sixty-four-page report of the twenty-one-member task force outlines the challenges of archiving digital information and provides a summary and recommendation based on a year-long exploration of issues confronting the publishing, information technology, library, archival, museum, legal, and government communities. Copies of the printed, bound report (ISBN 1-887334-50-5) can be ordered by sending payment of $15 per copy to The Commission on Preservation and Access, 1400 16th St., NW, Suite 740, Washington, DC 20036-2217. The report is available in a choice of electronic formats at RLGs home page at: http://www.rlg.org.

NISO on WWW

Earlier this year the National Information Standards Organization (NISO) mounted a home page on the World Wide Web. You can now find NISO at http://www.niso.org. Through the NISO home page you can get information on NISO, contact NISO officers, learn the status of new standards in development, view the NISO Press catalog, and obtain forms for suggesting new standards projects and joining NISO.

Imagine a world where the complete spectrum of information technology, library management and leadership is available to everyone ...in this world all can realize their full potential.

LITA/LAMA National Conference



TRANSFORMING LIBRARIES

OCTOBER 13-16, 1996

PITTSBURGH

A national conference on leadership and technology in the information age.

Brought to you by ***the American Library Association (ALA)*** *and its divisions.*

Come–meet the world

Library Information and Technology Association
Providing opportunities for professional growth and performance in areas of information technology.

Library Administration and Management Association
Providing a framework for career development and improving administrative and managerial capabilities.

ALA Office for Information Technology Policy
Promoting the development and utilization of electronic access to information as a means to ensure the public's right to a free and open information society.

For more information contact:

LITA/LAMA National Conference
American Library Association
50 E. Huron St. • Chicago, IL 60611
(312) 280-4268 / voice • (312) 280-3257 / fax
u09112@uicvm.uic.edu

"In 1985, there were only 300,000 registered electronic mail users...Today, the number of electronic mail users is estimated to be more than 27 million."
From *"Connecting the Nation: Classrooms, Libraries & Health Care Organizations in the Information Age,"* a report from NTIA, the U.S. Department of Commerce.

His world

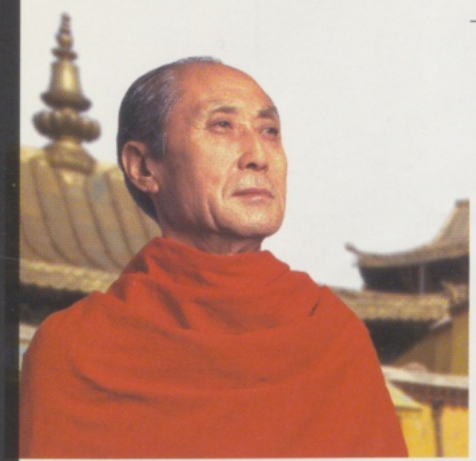
revolves around

silent meditation, yoga

and a library in

www.sirsi.com

pittsburgh.



unicorn™ It is possible, right now, for you to provide your clients and patrons with the absolute finest in library services and support—and extend those services around the world. It's an incredible capability that comes standard with Unicorn, the world's most powerful library automation system. Unicorn was created in 1982 by SIRSI. From the beginning, the right decisions were made to ensure that Unicorn users would always be able to meet change head-on. UNIX™ would be the operating platform. Client/server would provide the architecture. Unicorn would be the perfect blend of power, versatility and adaptability. These early decisions are the reasons why Unicorn is today's continuous source for technological breakthroughs. SIRSI is the only vendor offering complete public and staff functionality for MS-Windows™ and Macintosh™ clients. SIRSI's award-winning WebCat™ is the first product to tap the Web's potential for delivering library catalogs and other Z39.50 databases to patrons around the world. Today, strategic alliances with customers such as Carnegie Mellon and other industry partners ensure that Unicorn and other SIRSI products continue to lead the industry. Tomorrow, there will be new breakthroughs. And Unicorn users will be the first to reap the rewards.

SIRSI®

You need to reach the world. It needs to reach you.™

Worldwide Headquarters:

689 Discovery Drive • Huntsville, AL • Tel: 205-922-9820 • Fax: 205-922-9818

Total Networking SOLUTIONS

- Plug & Play CD-ROM Networking
- Compatible with All Major Operating Systems
- Choose 4X, 6X, or 8X CD-ROM Drives
- Configurations from 4 to 392 Drive Arrays
- Remote Access Systems
- Intranet Solutions
- CD Writers/Duplicators/Media
- Partners with Major Manufacturers
- Available on GSA Contract
- Toll Free Technical Support
- One-Year Warranty

New OPTI-NET Direct™
This hardware-based solution provides shared CD-ROM access over your Ethernet network. Plug & Play and web server features.

TODD ENTERPRISES INC.
800-445-TODD 516-487-3976 *FAX-516-466-6774*

yes It runs with NetWare

adaptec NSM PLEXTOR Stac SMART STORAGE

Microsoft OPTI-NET PIONEER TOSHIBA YAMAHA

Networkable CD-ROM Titles

Only Todd offers the Helping Hand... A FULL YEAR of FREE TECHNICAL SUPPORT

Call today for a FREE copy of our latest catalog!

TODD ENTERPRISES INC.
800-445-TODD 516-487-3976 *FAX-516-466-6774*