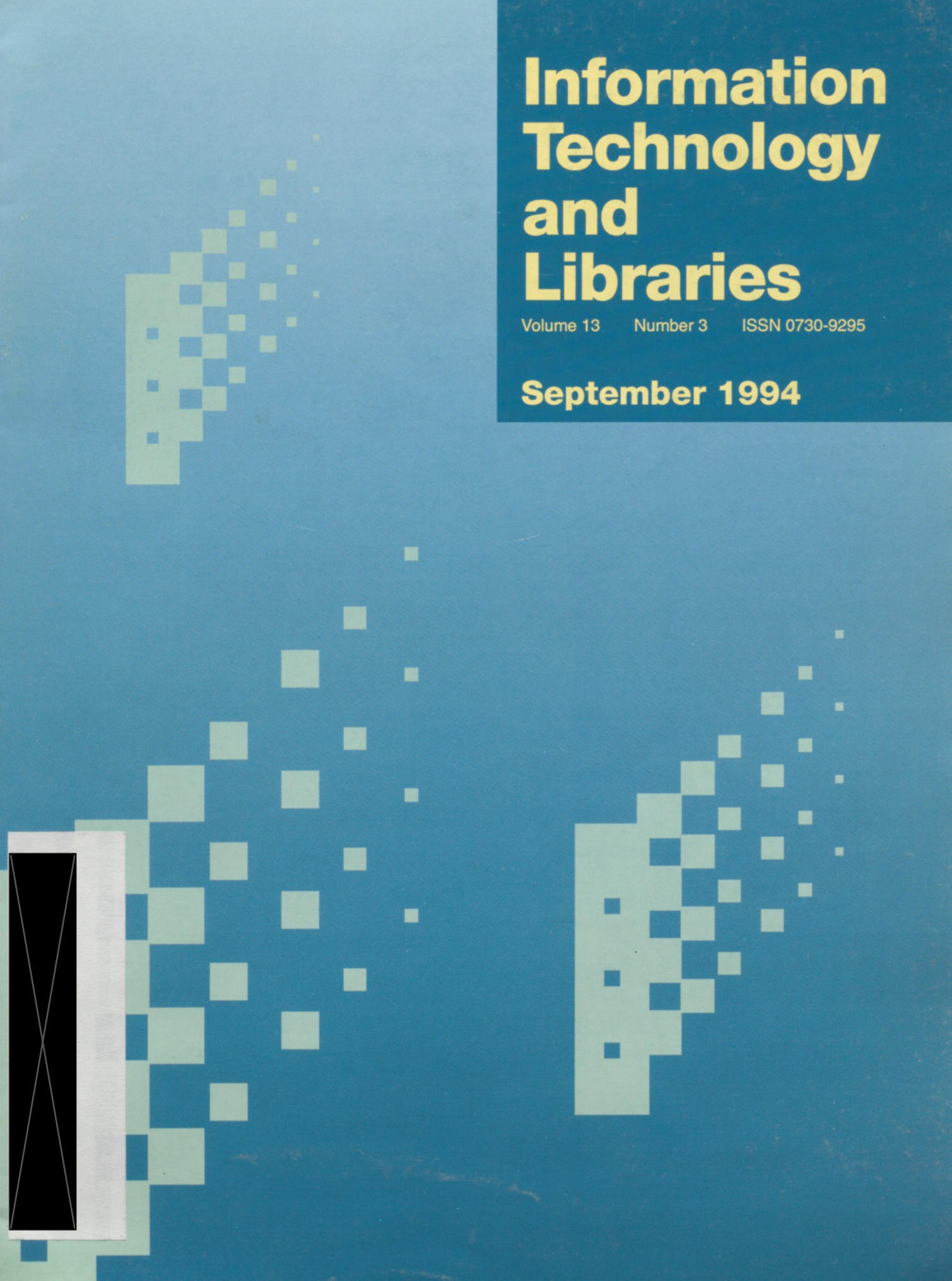
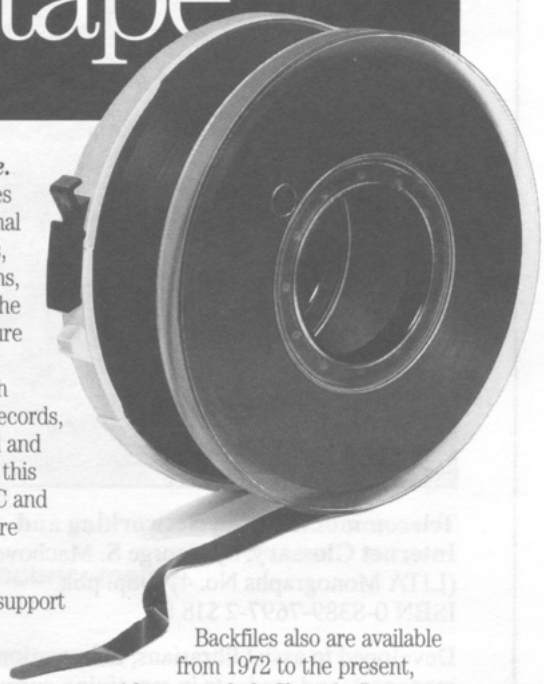
Information Technology and Libraries

Volume 13 Number 3 ISSN 0730-9295

September 1994



Now get unlimited, low-cost access to the PAIS Database via leased tape



For as low as $3,000 per year, you can have unlimited access to the PAIS International Database, the premier bibliographic index to public and social policy issues.

Economical, fixed-fee pricing. With leased magnetic tapes, there's no need to worry about usage and teleconnect fees. Your users can search PAIS as often as they like, for as long as they like, all for one annual rate.

Flexibility and ease of use. PAIS is compatible with most OPACs. That means your users don't have to learn different search software, and *you* don't have to acquire new hardware.

Institution-wide availability. With PAIS International on your network, users in dorm rooms and offices across an entire campus or organization can access the database 24 hours a day.

A total information resource. The PAIS magnetic tape includes references to over 380,000 journal articles, government documents, statistical directories, monographs, conference reports and more. The database also references literature published worldwide in English, French, German, Italian, Spanish and Portuguese. Bibliographic records, a Thesaurus File, and Periodical and Publisher Directories round out this unique resource. ISSN, ISBN, LC and Superintendent of Documents are included in the data.

Full telephone support. PAIS support is never more than a fast phone call away.

An annual subscription to PAIS on magnetic tape includes 12 monthly updates and complete documentation.

Backfiles also are available from 1972 to the present, either as a whole file or in five-year units. For more information or a free sample tape, use the coupon below or call toll-free, 800-288-PAIS.

☐ Please send me more information on PAIS Database tapes.

☐ Please call me about receiving a sample tape and tape documentation.

Name _____

Institution _____

Address _____

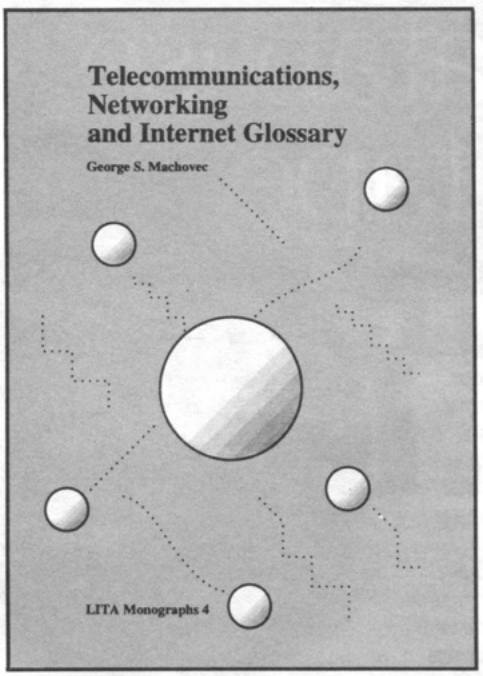
City/State/ZIP _____

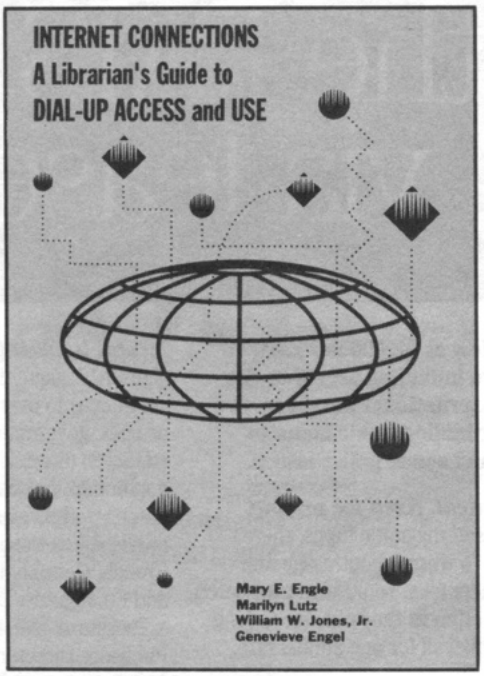
Phone _____ Fax _____

PAIS® **No one looks at the world like PAIS**

Public Affairs Information Service®, Inc.
521 West 43rd Street, New York, NY 10036-4396
800-288-PAIS, 212-736-6629 (in NYC)
Fax: 212-643-2848

KEEP CURRENT WITH TELECOMMUNICATIONS AND INTERNET





Telecommunications, Networking and Internet Glossary, by George S. Machovec. (LITA Monographs No. 4) 106p. pbk. ISBN 0-8389-7697-2 $18.00

Developed to assist librarians, information managers, and students in remaining current with the constantly increasing amount of new terminology in the high-growth communications field, this new glossary from CARL (Colorado Alliance of Research Libraries) technical coordinator George S. Machovec provides hundreds of definitions to telecommunications, networking and internet terms that frequently appear in library automation technology. A handy reference for today's librarian.

Internet Connections: A Librarian's Guide to Dial-Up Access and Use, by Mary E. Engle, et al. 166p. pbk. ISBN 0-8389-7677-8 $22.00

This guide is for librarians and library patrons who are not already members of the Internet family of users. It is a snapshot of the networking services available at the time of publication, a discussion of the concepts and terms in use, and a bibliography of the notable guides that document the search techniques, navigational tools, and information resources available. It is designed to help you get started right away.

To order, call 1-800-545-2433 and press 7. Call between 8:30 a.m.-4:30 p.m. CST.

LITA Library and Information Technology Association
a division of the
American Library Association

Information Technology and Libraries (ISSN 0730-9295) is the official publication of the Library and Information Technology Association, a division of the American Library Association, 50 E. Huron St., Chicago, IL 60611; *Executive Director:* Linda J. Knutson. The journal is issued quarterly in March, June, September, and December.

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 168 of the June 1994 issue. Manuscripts of articles, communications, and news items should be addressed to the editor: Thomas W. Leonhardt, *Information Technology and Libraries*, Bizzell Library, 401 W. Brooks, University of Oklahoma, Norman, OK 73019-0528; (405) 325-2611; Internet: leonhardt@aardvark.ucs.uoknor.edu. Copies of books for review should be addressed to: Susan B. Harrison, *ITAL Book Reviews*, The New York Public Library, 455 Fifth Ave., New York, NY 10016. Copies of software for review should be addressed to: George S. Machovec, *ITAL Software Reviews*, Colorado Association of Research Libraries, 3801 E. Florida Ave., Ste. 370, Denver, CO 80210-2542. ADVERTISING: contact Stuart Foster, Ad Sales Manager, and Dolores LaPointe, Ad Coordinator, c/o *Choice*, 100 Riverview Center, Middletown, CT 06457; (203) 347-6933.

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.

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology 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.

Circulation and Production: ALA Production Services (Eileen Mahoney; Dianne Rooney; Bruce Frausto, and Donavan Vicha), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by LITA or the ALA.

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 239.48-1984. ∞

Information Technology and Libraries

Volume 13, Number 3 September 1994 ISSN 0730-9295

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

OPACs in Twelve Canadian Academic Libraries: An Evaluation of Functional Capabilities and Interface Features

Joan M. Cherry, Nancy J. Williamson, Carol R. Jones-Simmons, and Xin Gu

OPACs (online public access catalogs) in twelve Canadian academic libraries were evaluated using a checklist. Particular emphasis was placed on evaluation of functional capabilities and interface features which have been suggested in the literature by researchers. The findings of the study suggest that, among the ten areas evaluated, "Screen Display" is the best developed area, whereas "Subject Search Aids" is the weakest. Possible directions for the future development of OPACs are identified and areas for future OPAC research are suggested.

Introduction

Numerous studies of OPACs have been conducted since their introduction in libraries in the early 1980s. Many studies have been devoted to the investigation of public acceptance, satisfaction with, and success in the use of OPAC systems. In other words, they have investigated the attitude and behavior of OPAC users. In this regard, the OPAC studies were fruitful. They revealed that the 1980s had witnessed the overwhelming success and growing popularity of OPACs in the minds of both librarians and the public, even though some technical problems remained unsolved.[1,2] On the other hand, as Peters observed, "The development of online catalogues in the [1970s and] 1980s was governed by bibliographic networks, vendors of automated systems, and technical services librarians, not by the needs and expressed wishes of the library patrons."[3] There were many studies of the capabilities and features of individual OPAC systems. The purpose of these studies was to enhance the capabilities of specific OPAC systems or to create special functions for special uses. Developed by different OPAC vendors or library institutions, various OPAC systems were produced, and are still undergoing change. Hildreth, in 1982, noted that "the systems [the OPACs] not only differed in the range and complexity of their functional features, they [the OPACs] used different terms to describe them and different commands for invoking them."[4]

Previous Research

Several researchers have been engaged in the analysis and evaluation of the functional and interface features of existing OPACs. Hildreth and his team investigated ten OPAC systems.[5] They collected data through first-hand use of each system, by sending questionnaires to system producers or owners, and by reviewing the system documentation in each case. Numerous checklists were used to analyze the ten systems. This was the first comparative OPAC study that involved a large number of OPACs. The methodology provided a model for describing and comparing OPACs in later research projects, including the present one. The major findings of Hildreth's study were:

- Little similarity existed in command language vocabulary or syntax among the ten OPACs studied;
- Display formats of the ten OPACs varied significantly;
- The number of available help features and displays varied considerably across the ten systems.

Salmon,[6] employing Hildreth's conceptual framework and features classification, observed and compared the functional characteristics of twenty OPACs in U.S. libraries. In his paper, screen displays of various systems were used to illustrate the functional features of these OPACs. He presented feature comparison tables that showed that the OPAC field had evolved rapidly. As a result of his study, Salmon presented a list of features he suggested OPAC systems should offer.

Fayen discussed, in general, the types of OPACs available in terms of their approaches to searching.[7] She briefly described features of twelve OPAC systems she thought were representative and suggested a list of factors that should be considered when planning to build or buy an OPAC. These included the size of the database, computer equipment available, costs, authority file capabilities, display options, and retrieval features such as Boolean operators, truncation and wildcard operators, browsing features, and numeric value searching.

Another comparative study of OPAC features was conducted by Boss and Harrison,[8] who developed a survey that was sent to all OPAC vendors. According to the survey results, Geac Advance, NOTIS, and DRA were the most complete products as of mid-1989. Each of these systems satisfied more than 80% of the composite requirements. Systems that met the lowest percentage of the requirements were the TBMS and OCLC Local systems. Since the main purpose of this study was to provide libraries planning to go online with a guide to

Joan M. Cherry and **Nancy J. Williamson** are faculty members at the Faculty of Information Studies, University of Toronto. **Carol R. Jones-Simmons** and **Xin Gu** are graduates of the Faculty of Information Studies, University of Toronto.

OPACs available and to help them make a choice, recommendations were made concerning OPAC systems procurement. In an effort to continue the investigation of the state-of-the-art features of present-day OPACs with respect to functional capabilities and interface characteristics, a research project for a master of information science degree was conducted at the University of Toronto. This paper discusses the major findings of that project.[9]

Methodology

Selection of OPACs

This study examined OPACs in twelve CARL (Canadian Academic Research Libraries) libraries. They represented about half of the total CARL members. The twelve institutions were the University of Calgary, Carleton University, Dalhousie University, the University of Manitoba, the University of New Brunswick, Queen's University, the University of Saskatchewan, the University of Toronto, the University of Waterloo, the University of Western Ontario, the University of Windsor, and York University. These institutions were selected because their OPACs were publicly accessible through the Internet. Among these twelve OPACs, there were seven different types of systems, including DOBIS, Geac, NOTIS, PALS, DRA, and two home-grown systems.

Although four institutions had Geac systems, the systems were very different from one another. Therefore, in this study they were treated as individual systems. Similarly, three institutions had NOTIS systems, which were also very different from one another. Thus, the twelve OPACs were regarded as different systems and consequently were examined individually.

The Evaluation Checklist for OPACs

In order to evaluate the twelve OPACs, a checklist of functional capabilities and interface features of an ideal OPAC system was designed and used. This checklist served as the evaluation standard or benchmark for evaluating the strengths and weaknesses of the OPACs. The format for the checklist was modeled on a checklist for evaluating the usability of human-computer interfaces developed by Ravden and Johnson.[10]

Much research had been done to determine what features should be included in an OPAC system. In addition to the basic functions that were commonly believed to be offered by most OPAC systems, many advanced features had been suggested by OPAC researchers. For the purpose of this study, the ideal OPAC was defined as one which possessed all of the features proposed in the literature. One hundred seventy features were included in the checklist. In order to permit easy review and facilitate comparative evaluation of the features of different systems, these 170 features were grouped into ten categories: "Database Characteristics," "Operational Control," "Searching," "Subject Search Aids," "Access Points," "Screen Display," "Output Control," "Commands," "User Assistance," and "OPAC Usability via Remote Access." This classification was developed by reviewing several similar classifications of OPAC characteristics. Among the studies consulted in drawing up the checklist were those of Hildreth,[11] Matthews, Lawrence and Ferguson,[12] Salmon,[13] Dwyer,[14] and Northover.[15] "Remote Access" was added to the traditional classifications because this feature had emerged in recent research as an important functional element in OPACs.[16] Each category appeared as a separate section in the checklist.

Each question on the checklist represented one OPAC feature. All questions required a yes or no response. Space was also available for comments about each feature. If the question was not relevant to a particular OPAC, the "Comments" column was marked "Not Applicable" or "NA." The checklist was used uniformly across all systems investigated.

In the analysis of the data, ten tables were produced. Each table corresponded to a category in the checklist. The checklist data were entered into tables as follows. If the checklist answer was yes (indicating that the system provided this feature), an *x* was placed into the corresponding cell in the table. If the checklist answer was no (indicating that the system did not have this feature), the corresponding cell in the table was left blank. If the answer was not applicable, then *NA* was entered into the corresponding cell.

Scores were calculated for each system. Each time a cell was checked, the overall score for that OPAC increased by one point; otherwise, the score did not change. In other words, the score for a system was the total number of cells checked for that system. Thus, if a system had many features, the score for that system was high. The highest possible score for a category, (i.e., the score of the ideal system) was the total number of features included in that category. The highest possible overall score was 170. It was recognized that some of the features in the checklist were not as important as other features for an OPAC system. However, for the sake of simplicity, all features were assigned the same weight (i.e. one point for each feature). With these scores, we

compared the twelve OPAC systems in each category and overall.

Data-Collection Method

Initially, data were collected for each of the twelve OPACs twice by one of the authors from June 1992 to November 1992. Some differences were found between the two sets of data. These differences were a result of changes in the systems or errors in evaluation. Between December 1992 and February 1993, a second evaluator (another author of this paper) collected data independently. In late February of 1993, the two sets of data were checked for disagreements. All differences were resolved by checking those features again online, and the resulting data reflect the status of these OPACs as of February 1993.

Results

Database Coverage and Bibliographic Access Points

Many researchers had given suggestions regarding the content and coverage of an OPAC database. Some believed an OPAC database should be a database of bibliographic citations for all materials in the library's collection.[17] For example, Hildreth thought that an OPAC should be a "full-collection access tool."[18] He identified one of the major functional improvements of third-generation OPACs to be "expanded coverage and scope." Others, like Matthews, had even higher expectations, believing that an OPAC should be "much more than an inward looking tool for patrons to gain access to a library's collection but rather it would, or should, become

Table 1
Coverage

OPAC	Univ.	Provides public access to:
DOBIS	A	1. Book, serials & special material catalog 2. Govt. publications catalog 3. Comm. media film catalog
Geac	B	Library catalogs of eight colleges/universities
	C	Library catalog
	D	Library catalog
	E	Library catalog
NOTIS	F	1. Library catalog 2. Other university library catalogs
	G	Library catalogs
	H	1. Library catalog 2. Circulation information 3. Other university library catalogs
PALS	I	1. Library catalog of the university 2. Library catalog of a local hospital 3. Library catalog of one other local university
DRA	J	1. Library union catalog 2. Journal article databases 3. Other university library catalogs
Home-grown	K	Library catalog
Home-grown	L	1. University info. system (calendar, news, phone/fax/e-mail, campus bulletin board, etc.) 2. Library catalog for users accessing through the INTERNET

a facilitating tool that would allow a patron to gather, evaluate and download appropriate articles, text of chapters of a book, musical sound records, audio recordings, motion pictures, slides, etc."[19] Table 1 gives a general overview of the coverage of the twelve systems.

None of the twelve OPACs provided access to the full collections of their libraries. Most did not provide access to journal articles and manuscripts. Thus this study reveals that there has been little progress made in providing access to all of the resources of a library through its OPAC.

Only five OPACs provided access to holdings of other libraries. OPACs at universities F, H, and J provided access to a number of other university OPACs (including some in the United States). The OPAC at university B provided access to the holdings of eight colleges and universities, including itself. The OPAC at university I provided access to its own holdings and the holdings of one other local university and hospital. The other seven OPACs provided access only to their own library holdings.

Table 2 presents the data collected on database characteristics. It shows that all twelve OPACs provided standard bibliographic information, such as call number, author, title, subject headings, edition, publication, volume, and name of the library where the item is

Table 2
Database Characteristics

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
Does the system provide the following information in either short or long record displays? (a) call number	x	x	x	x	x	x	x	x	x	x	x	x
(b) author	x	x	x	x	x	x	x	x	x	x	x	x
(c) title	x	x	x	x	x	x	x	x	x	x	x	x
(d) subject heading(s)	x	x	x	x	x	x	x	x	x	x	x	x
(e) edition information	x	x	x	x	x	x	x	x	x	x	x	x
(f) publication information	x	x	x	x	x	x	x	x	x	x	x	x
(g) copy information	x		x		x	x		x	x	x	x	
(h) volume information	x	x	x	x	x	x	x	x	x	x	x	x
(i) location for current issues of serials	x	x	x		x	x	x	x	x	x	x	x
(j) name of the library where the item is located	x	x	x	x	x	x	x	x	x	x	x	x
(k) loan status	x	x	x	x	x	x	x	x	x	x	x	
(l) document abstract												
(m) table of contents												
(n) citations within the text												
(o) keywords from the book index												
(p) book review												
Score: (maximum 16)	11	10	11	9	11	11	10	11	11	11	11	9

Table 3
Operational Control

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
1. (a) Is there an introduction to the system?	x	x	x	x	x	x	x	x	x	x		x
(b) Is there a set of instructions?	x	x	x	x	x	x	x	x	x	x	x	x
2. Does the system provide a choice of command-driven or menu-driven interface throughout?												x
3. If menu choices are by letters, are they mnemonic?		x	x	x	x	NA	NA	NA	NA	x	NA	NA
4. In the command mode: (a) Is a display of the commands available?	NA	x	x	x	x	x	x	x	x	x	x	x
(b) Are examples of how the commands are used available?	NA	x		x	x	x	x	x	x	x	x	
5. Does the system allow the registered user of this library to place holds/reserves on items at the terminal?			x						x		x	
6. Does the system allow the user of this library view the list of items charged out to him/her at the terminal?			x			x		x	x		x	
7. Does the system display system or database changes, system access notices or library news?	x	x	x	x	x	x	x	x	x	x	x	x

located, in either short or long record displays (table 2, questions a–k). None, however, provided expanded subject information through such elements as document abstracts, table of contents, citations within the text, keywords from the book index or book reviews (table 2, questions l–p), although many researchers had recommended augmenting MARC records to include more information.[20] Table 2 also shows that the highest score received was 11/16. Eight OPACs received this score. The lowest score, received by two OPACs, was 9/16, indicating a small gap between the highest and the lowest score. The data indicate that these twelve OPACs had reached approximately the same level in the development of their databases.

Operational Control

Table 3 presents the data collected on operational control features. Table 3 shows that only the OPAC at university L provided a choice of command-driven or menu-driven interfaces throughout (table 3, question 2). Most systems were hybrids of the two modes. That is, sometimes they offered menu choices; sometimes they required users to type commands; and sometimes the two modes were available at the same time. All but one of the twelve OPACs studied offered an introduction to the system (table 3, question 1 (a)). All OPACs offered a set of instructions (table 3, question 1(b)) and database or library news features (table 3, question 7).

Although early in 1982 Hildreth suggested that OPACs should allow users to place a hold or renewal on items at the terminal,[21] only three OPACs (at universities C, I, and K) offered this function (table 3, question 5). Hildreth also suggested that OPACs should support the downloading of bibliographic records to users' personal computers. Only two OPACs (those at universities J and K) supported this function (table 3, question 11). Table 3 also shows that only three OPACs (at universities A, J

Table 3 (continued)
Operational Control

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
8. Can the user select the length of the system messages? (e.g., long or short error messages, or help displays)									x			
9. Is there a ``break'' or ``interrupt'' key or command? (e.g., ``Esc'' key)	x									x		x
10. Can the user return to a previous screen and change a selection?			x	x	x					x		
11. Does the system support downloading bibliographic records to the local user's personal computer?										x	x	
12. Can the user communicate interactively with library staff through the OPAC whenever he/she has trouble finding material on a subject?												
13. Can the user transmit search results via e-mail?												
14. Does the system offer online mailboxes for user suggestions and comments?	x									x	x	x
15. Can the user send messages to a specific library staff member or department?										x		
Score: (maximum 17)	5	6	8	7	7	6	5	6	8	11	8	7

Note: "NA" means "not applicable"

and L) offered a "break" or "interrupt" key (table 3, question 9). For OPAC L, although the system indicated that it had an "interrupt" key, no equivalent keys were explained for remote users. The suggested features of interactive communication between users and library staff via e-mail also were lacking (table 3, questions 12, 13, 14, and 15). None of the twelve OPAC systems received high scores in table 3. The average score was seven, less than half of the features of the ideal OPAC. The OPAC at university J, which received the highest score (11/17 points), only had 65% of the features suggested by researchers. Indeed, two OPACs scored only 5 points, a mere 29% of the suggested features (table 9).

This suggests that generally these OPACs were weak in operational control features. Specifically, the following features were missing in most of the twelve OPACs:

- both command-driven and menu-driven interfaces, and easy movement from one mode to the other;
- a facility for users to place holds on items at the terminal;
- a facility for users to view the list of items charged out at the terminal;
- an option for users to select the length of system messages (e.g., long or short error messages);
- a "break" or "interrupt" key or command;
- support for downloading bibliographic records to users' personal computers;
- online interactive communication with library staff through the OPAC;
- support for transmitting search results via e-mail;
- a facility for sending messages to a specific staff member or departments;

- online mailboxes for user suggestions and comments.

Searching

This section contained 22 questions (See appendix, section 3). Scores of the twelve OPACs varied a great deal. The best OPAC scored 33/46 (72%) and the worst scored 10/46 (22%). On average, the twelve OPACs provided more than half, 53%, of the desired features for searching (table 9).

NOTIS-based systems scored higher in search features than Geac-based systems. Among the four Geac-based systems, OPAC C scored much lower than the other three. For section 3, question 8, "Which Boolean operators are available?" and section 3, question 16, "Can searches be limited by publisher, type of materials, library location, etc.?" OPAC C indicated that the functions were available, but they could not be activated through remote access.

All systems allowed users to set default values for search type and display format (appendix, section 3, question 1 (a) and (c)), as well as allowing for default values to be reset during a session (appendix, section 3, question 2). All but two systems also allowed users to set default values for specific fields in keyword searches (appendix, section 3, question 1(b)). Approximately half the OPACs allowed users to start a search anywhere (appendix, section 3, question 3) or to continue or start a search directly from the help screen (appendix, section 3, question 4).

All systems permitted keyword searching of the controlled vocabulary (appendix, section 3, question 5(a)); all but one also supported keyword title searching (appendix, section 3, question 5(b)). Only half of the OPACs supported keyword searching anywhere in the record (appendix, section 3, question 6). All but one OPAC had stop words; however, only four of the OPACs had a list of stop words available for display, and only half indicated to users who searched a stop word that the word was not indexed (appendix, section 3, question 7).

Most OPACs (9) allowed for Boolean searches with operators AND, OR, NOT, as well as supporting Boolean searches using keywords in author, title, and subject fields and across two or more fields. Six of those nine also supported keyword searches not limited to any field(s) (appendix, section 3, question 9). Six systems supported word adjacency operators; only four supported word proximity operators. All three NOTIS systems supported both (appendix, section 3, question 11 and 12). All but two systems allowed users to use right-truncation; only three systems used wildcard characters, and three allowed users to indicate specific limits on truncation. None of the systems offered left-truncation or variable-length wildcard characters (appendix, section 3, question 14). More than half (7) of the twelve OPACs allowed users to limit searches; the most popular limitations being language, type of material, and date (appendix, section 3, question 16). None of the OPACs supported weighted term searches (appendix, section 3, question 15); none allowed search strategies to be saved for later use (appendix, section 3, question 20); and only two allowed users to save search results (appendix, section 3, question 21). More than half (7) of the OPACs made it easy for users to switch between one type of search and another (appendix, section 3, question 22).

Hildreth indicated that for many users "browse searching is the most useful and preferred approach when the search aim is not specific (regarding, for example, discipline or topic, type of publication, level of treatment, perspective, etc.), the desired results are not precisely known in advance, or the correct terms for representing users' queries (which may be vague) are not known at the outset."[22] The data (appendix, section 3, question 19) showed that all twelve OPACs offered browse searching using author, title, and subject searches; however, only five also included cross-references in the index.

Hildreth also suggested: "Related document browsing and discovery can be facilitated in OPACs through richer precoordination in the database of multiple subject/topic clues found in bibliographic records (e.g., linking title terms with subject headings with call numbers, etc.), and by providing more search navigation options between retrieved and unretrieved (but linked) records, that is, record to record 'jumping' at the discretion of the searcher (e.g., 'show me more books from this publisher,' 'What other titles are in this series?' 'What documents cite this work?'"[23] None of the twelve OPAC systems provided these dynamic navigation features (appendix, section 3, questions 17 and 18).

OPAC Systems' Responses to Unsuccessful Queries

Many previous studies of OPAC transaction logs suggest that increasing the result when too little is retrieved and reducing the result when too much is retrieved are two major problems frequently encountered by OPAC users.

Hildreth commented that one of the major reasons the OPACs of the 1980s failed to satisfy users' needs was "partially implemented search strategies and missed opportunities to retrieve relevant materials."[24] He suggested that searches resulting in large retrieval sets should be scanned or narrowed in size, and title key-

word searches should be followed by searches on the call numbers or subject headings of the retrieved records. Burke proposed the "no dead end" OPAC. She stated: "The late 1990s catalog will not bring the user to a dead stop with screens like 'No Hits.' Instead, it will suggest ways to modify or expand the search."[25] Markey has presented specific recommendations to help subject searchers who failed to retrieve any output at all or retrieved too much output.[26] Her recommendations are summarized below.

To increase output, Markey suggested that a system:

- provide online cross-references directing users from the input term to the correct term;
- display related word lists directing users to more general terminology—that is, when users enter subject headings that were not assigned to any library materials, a list of terms broader than the entered heading with an explanatory message should be displayed;
- automatically truncate users' input subject term(s);
- respond with a message telling users to try keyword subject searching, or automatically perform the keyword subject search and report results to users;
- augment bibliographic records with subject headings.

To reduce output, Markey suggested that a system:

- offer prompts telling users how to reduce the number of retrievals;
- display related word lists to help users find more specific terms or phrases, or find additional terms or phrases to enter into the ongoing search strategy through an implicit Boolean AND;
- show a list of subject headings alphabetically adjacent to the entered term, class number areas (in which the books are assigned to the displayed subject heading), and the number of books found in each class number area.

In addition to the checklist data (appendix, section 6, question 9), this study gathered descriptive data on how these twelve OPACs addressed problems of unsuccessful searches—either zero hits or too large retrievals—in an attempt to gauge whether the above-mentioned suggestions had been adopted in present OPACs or not. Although not all researchers' suggestions were implemented, some systems had made progress in this area. For example, OPACs at universities F, G, H, and I not only provided possible reasons for zero hits, but also provided users with specific techniques to increase the number of hits. The OPAC at university J provided alternate suggestions in author, title, and subject searches that resulted in zero hits. The messages presented by the OPAC at university I when searches resulted in sets with too many hits ("use AND command with additional WORDS or LIMITING command to reduce results") may also be helpful in modifying searches to reduce the number of hits.

OPACs at universities A, K, L, and the four Geac-based systems did not give any helpful suggestions when search requests resulted in zero hits or too-large retrieval sets. In author, title, and subject searches that resulted in zero hits, these OPACs simply displayed a list of headings alphabetically close to the search. This display of a list of index terms may help users to identify appropriate headings. When keyword searching retrieved no hits, the Geac-based systems only asked users to revise the search request but did not suggest how to accomplish this. When keyword searching retrieved a set that was too large, the Geac-based systems could display up to five hundred records.

Subject Search Aids

Subject searches here refer to LCSH subject searching, since the Library of Congress Subject Headings was the principal controlled vocabulary used for constructing online indexes for all twelve libraries.

Subject searches have been an enduring problem since traditional card catalogs were first used. Over the years many studies have been conducted to examine searching tactics employed by users performing subject searches and various methods for improving subject searching. With the introduction of online public access catalogs, this research activity has increased significantly. The hopes and expectations of researchers and librarians were that OPACs would support new features never before possible and would vastly improve subject searches.

Evaluating a system's performance on subject searches is a difficult process in that there are different objectives for subject searching. Different expectations of users result in varying definitions of what is "successful" subject searching. For example, should the search result be an exhaustive bibliography, or only the most relevant items on the topic? For use in this study, Mandel's definition of a successful subject search was revised.[27] Here, a successful subject search is one that:

- leads users from the topics they have in mind to the relevant vocabulary terms available in the catalog;
- provides users with the most relevant items in the catalog on the topic in question—the upper bound

of the number of items provided should be in the range of 150 (± 50)[28];
- provides users with enough information to decide whether to look for the item through a search in the stacks.

Why have researchers put so much effort into studying subject searching? There are two important reasons. First, many online catalog use studies snow that subject searches predominate over other types of OPAC searches, although users are likely to perform more "known-item" than subject searches as they increase their years of formal schooling.[29] Second, users encounter more problems searching by subject than with any other single aspect of their interaction with the online catalog.[30] For these two reasons, subject searches deserve more emphasis in terms of improvement and enhancement of online catalogs than other types of searches.

In view of various problems reported in previous subject search studies conducted by different researchers, the major question for investigation becomes, What is the source of subject search failures? The findings from several studies on user subject search behavior may throw some light on this question:

- Only about half of the terms used by readers in their first try of a subject search correspond to a subject heading or a reference found in the catalog.
- Users often select terms that are either too broad or too narrow.
- Not all users persist in subject searching until they are successful.[31,32]

Evidence suggests that users feel that the nature of subject searching is, in general, more difficult as compared with known-item searches, such as author and title searches. Karen Markey indicated that "subject searches, which are by nature more abstract, often required more mental effort to conjure up many phrases or single terms to express a topic or concept than did known-item searches."[33] Moreover, the cognitive process and search techniques are not the only causes of failure. An extensive literature review revealed that the shortcomings of LCSH were another factor in subject search failures. Inconsistencies in LCSH subject terms, lack of specificity and currency in LCSH subject terms, lack of *see* references, and awkward phrase construction are flaws frequently discussed by researchers. On the other hand, some researchers believed that "the frequently voiced complaints regarding lack of specificity in LCSH do not necessarily reveal that the actual vocabulary of the list is the cause of the search failure. For example, readers' requests may be too specific for the level of subject indexing applied in library catalogues. American library catalogues generally index whole books and do not contain subject entries for parts of books. This difference between the general practice in the subject cataloguing of monographs and readers' requests could account, at least in part, for the demonstrated superiority of a system that adds information derived from the indexes and table of contents of monographs."[34]

Aside from policy decisions, the indexers or subject catalogers are identified as sharing some of the responsibility for search failures. Edward Blume pointed out that subject headings could have been created as needed, but often catalogers chose not to do so. The result can be the use of an inappropriate, or less than satisfactory, subject heading, simply because it exists and its use avoids changes in authority files and bibliographic records. Research suggests that many search failures are not due to the limitations of the system as such, but rather are the result of extremely poor cataloging decisions.[35] Moreover, traditional subject cataloging provides provisional subject access at best. As further explained by Charles R. Hildreth, "effective retrieval of records resulting from a match of the LCSH search terms with an index record is further diminished by the fact that, on average, less than two subject headings are assigned by the Library of Congress to each item catalogued."[36]

Although there are many problems with LCSH, this system will continue to be used for political and (especially) economic reasons for the foreseeable future. Millions of subject card catalogs and machine-readable subject entries are already stored in libraries around the world. In particular, LCSH is still the closest approximation available to a common subject vocabulary authority system. Therefore, there is a prevalent opinion that it is more realistic to make an effort to find ways to improve LCSH subject searches rather than to replace the LCSH, as some subject catalogers have suggested.[37]

Various ways have been suggested for this to be accomplished. Among the suggestions are the use of keyword searching on various parts of the record, including LCSH headings; enhancement of the records by the addition of more natural language terms; the use of classification as an online browsing tool; the display of authority control data (subject headings and references) for users; and search simplification through the use of transparent references. Several researchers, including Cherry,[38] have suggested the use of keyword subject searches, title searches, and keyword title searches to bolster LCSH subject searches. Cherry's study showed that conversion of zero-hit subject search queries to keyword subject, keyword title, or title searches using the original queries from users' zero-hit subject searches were as fruitful, or more fruitful, than using new

searches constructed from cross-references provided by LCSH. For a small number of zero-hit subject searches, LCSH cross-references provided the fruitful approach. For many queries, however, the better choice was to change the original search to a keyword subject, keyword title, or title search.

Some librarians have argued that besides the words in the title of a book, and very occasionally a contents note, "there simply are not enough descriptive words in standard records for monographs to permit adequate subject retrieval."[39] To improve subject access, some researchers have suggested augmented MARC records, which include tables of contents and indexes of books. For example, Atherton created a database of MARC records enriched with descriptive terms from the tables of contents and the indexes of the books represented.[40] She reported that the enriched database was clearly superior for subject searching. None of the OPACs in this study, however, incorporated these recommendations regarding enhanced records (section 5, Access Points), probably due to the labor costs this would entail. Mandel pointed out that "academic libraries, faced with budget cuts, are not likely to adopt cataloging practices that require additional labor. In fact, the last decade has shown greater reliance than in the past on standard Library of Congress records in most libraries. Convincing libraries that they need to enrich their records for subject access will require a large body of research and a well argued plan to show that the benefits of such an enrichment outweigh the costs."[41]

One of the approaches of subject searchers is so-called "bookshelf-browsing." To facilitate the bookshelf browsers' search, Karen Markey proposed an alphabetical display of assigned subject headings with class numbers.[42] Such a display would direct the user to the most relevant class number areas, or at least link the user's subject with one area of the classification scheme. That is, once the user entered a subject, the system in response would display a list of the subject catalog's controlled vocabulary terms or phrases alphabetically near the input term(s). The user could scan forward and backward in it. She suggested that this alphabetical display include assigned subject headings, number of postings per assigned subject heading, class numbers common to the majority of the items assigned a particular subject heading, and line selection numbers (or letters). Results of the CLR survey also suggested that "the Dewey and LC Classification Tables, including the defining words and phrases tied to the class numbers, might be placed online and made available for searching and browsing."[43]

Later in 1986, Markey reintroduced this idea of putting classified lists of subject vocabulary online. She stated: "Online classified lists could be taken from the English-language captions or headings of classification schedules such as the Library of Congress, Dewey Decimal, or Universal Decimal classifications. This feature could help online searchers match their vocabulary with that of the online catalog, view classification in context, and identify area(s) of the classification that are pertinent to their topic of interest. Online classified lists would satisfy users' needs for lists of related terms and subject augmentation." Markey predicted that "since the Dewey Decimal Classification and parts of the Universal Decimal Classification are in machine-readable form, online classified lists of subject vocabulary may be implemented in online catalogs in the near future."[44]

As early as 1981, CONDOC (Consortium to Develop an Online Catalog at the University of Notre Dame Libraries) recommended that an authority file for subject headings with appropriate cross-references be a standard feature of an OPAC.[45] A study conducted by the OCLC Research Department,[46] which was intended to provide OPAC designers with features that would support and enhance the present subject search tactics of library users, revealed that users were enthusiastic about the idea of online displays of thesaurus terms to help them choose broader, narrower, and related terms; they also preferred that the system transparently (i.e., automatically) translate their input terms into the terms used in the catalog.

Table 4 shows the data for the subject search aids provided by the twelve OPACs evaluated in this study.

As shown in question 1 of table 4, none of the OPACs supported browsing of the display of classification outlines or schedules. Table 4 also shows that eleven of the twelve OPACs offered displays of subject headings that begin with the search term(s). When users entered a term or phrase, the system responded with a list of subject headings beginning with the search term or phrase, the number of titles assigned to each listed subject heading, and a list of consecutive numbers from which users could select to view brief bibliographic record displays of titles assigned to the listed subject heading. All of the systems allowed users to browse forward and backward in the list of subject headings. However, in this approach, the subject headings can only be retrieved when the search term or phrase exactly matches the subject headings, or at least the initial terms of the subject headings (table 4, question 2a).

The keyword-in-context subject heading display approach is more flexible. It produces a display of not only those subject headings in which the keyword is the initial term in the headings, but also those in which the keyword is embedded in the headings. The OPACs at universities A (DOBIS) and C, D, E (three of the four Geac systems) offered such keyword-in-context displays of subject headings (table 4, question 2b).

Although online cross-reference suggestions have

Table 4
Subject Search Code

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
1. Can the user browse a display of: (a) classification outlines?												
(b) classification schedules?												
2. Can the user view a group of subject headings: (a) which begin with the search term(s)		x	x	x	x	x	x	x	x	x	x	x
(b) which include the search term wherever imbedded in the subject headings	x		x	x	x							
3. Does the system display the following cross-references? (a) SEE/USE	x					x	x	x		x		
(b) SEE ALSO/BT/NT/RT	x					x	x	x		x		
4. Does the system have transparent SEE/USE references which automatically substitute the user's input term with the correct subject heading without informing the user?												
5. Does the system convert an original zero hit Subject search to Title, Keyword Title or Keyword Subject search?												
Score: (maximum 8)	3	1	2	2	2	3	3	3	1	3	1	1

been made for years, and these features have been used successfully to improve subject access in a variety of information databases, less than half of the OPACs studied here had this feature (table 4, question 3). Question 4 shows that no OPACs automatically translated users' search terms to the subject headings used in the catalog. Five of the twelve OPACs (including all three NOTIS systems) provided users with displays of cross-references (table 4, question 3). None of the twelve OPACs converted users' zero-hit subject queries to other types of searches, nor did they advise users to do so (table 4, question 5).

In short, it appears that much research has been done on ways to improve subject searches, and various techniques have been proposed by researchers. Unfortunately, data collected in this study revealed that none of the OPACs had satisfactorily incorporated the modifications suggested. More effort should be made to incorporate these features in the next generation of OPACs.

Access Points

All twelve systems could be accessed by author (including corporate author), title (including series title), subject, and LC call number. Eleven of the twelve systems offered keyword subject search (see appendix, section 5, question 10). Most OPACs in this study (except those at universities A and K) could also be searched by keyword(s) in title (appendix, section 5, question 9).

One of the findings of the CLR survey was that among the three highest user priorities for improvements of any type (systems, library services, databases) was the ability to search a book's table of contents, summary, or index.[47] As indicated in the previous section on subject search aids, none of the twelve OPACs incorporated researchers' recommendations regarding enhanced records. None of them provided access points to table of contents, citations within the text, or indexes of books (appendix, section 5, question 18–21).

On average the twelve OPACs offered approxi-

mately half of all the access points proposed by researchers, 56% (see table 9, page 189).

Screen Display

Screen display is an important aspect of the OPAC interface because it affects the ease with which users can absorb displayed information. Most of the OPACs' screen display scores were high. Five OPACs offered 90% of the features of an ideal system. Even the worst OPAC scored 50% (table 9). All the systems reported the number of retrieved records first, before displaying results, so that limiting techniques could be used to reduce large retrievals, or other techniques could be used to increase small retrievals (appendix, section 6, question 3). All systems provided at least two display formats (appendix, section 6, question 4), brief and full bibliographic record, as recommended by researchers. All systems numbered items successively (appendix, section 6, question 8), so that users always knew how many records they had viewed.

A study carried out by Saracevic, Mokros, and Su indicated that users did not want large outputs.[48] They found that the upper bound users were willing to view was in the range of 150 items (± 50). Users seldom examined specific items from sets larger than this. None of the OPACs evaluated in this study set the display limit as low as Saracevic et al. suggested (appendix, section 6, question 9). The OPACs at universities B, C, D, and E displayed up to 500 hits. Those at universities F, G, and H limited the display to 250 hits. The OPAC at university I displayed 20 hits each time and allowed users to choose to continue displaying the next 20 records or to revise the search query. The OPACs at universities A, J, K and L displayed the search result with no upper limit.

Output Control

Table 5 shows the data for output control characteristics of the twelve OPACs. In general, all twelve OPACs were very weak in the area of output control. They did not give users much control over the format or the order of records displayed. The average percentage of output control features provided by the twelve OPACs was approximately 31%. Even the best OPAC provided only 50% (table 9).

The display formats of these OPACs were not flexible. For example, none of the OPACs allowed users to select specific fields for display (table 5, question 1). All the OPACs allowed users to select a single record for display; however, only one OPAC could display several records out of sequence, and only three OPACs could display a range of records. Among the twelve OPACs, only the one at university I could display results in all three ways (table 5, question 2). Only three OPACs allowed the results of several searches to be merged for display (table 5, question 3). Only one-third of the OPACs (four) let users specify the sort order of search results (table 5, question 4). None supported ranked document display in decreasing order of probable relevance to the search query (table 5, question 5). All systems displayed results by paging (table 5, question 6).

Commands

Most OPACs did well in the area of command characteristics. They had, on average, 65% of the features that had been suggested to improve the usability of commands. The OPACs at universities B, D, E, and J each had a score of 10/13, or 77% of the recommended features.

All of the OPACs assigned each command a consistent role (appendix, section 8, question 1). Only four OPACs used function keys; those that did had consistent definitions for keys and used them to reduce the number of keystrokes for commonly used commands (appendix, section 8, question 2 and 3). Ten OPACs had a standardized syntax for commands (appendix, section 8, question 4), and ten minimized the use of punctuation, or at least used familiar punctuation (appendix, section 8, question 5 and 6). Eleven OPACs used mnemonic abbreviations for commands (appendix, section 8, question 7), which facilitated users' remembering commands. However, only the Geac systems supported stacking commands (appendix, section 8, question 8).

Section 8 also included several questions (numbers 9–12) about leading articles, punctuation, word order and upper/lowercase. Only six OPACs ignored leading articles, which allowed users either to enter them or not and not have the search results affected (appendix, section 8, question 9). Seven OPACs allowed users to omit dashes preceding LCSH subdivisions. Most systems ignored punctuation entered by users that was not required (appendix, section 8, question 10). All OPACs accepted author's names only in inverted order (appendix, section 8, question 11). All OPACs accepted both upper- and lowercase (appendix, section 8, question 12). In retrospect, we feel that these questions would have been more appropriately placed in section 3 on "Searching."

User Assistance

Table 6 shows data for user-assistance characteristics of the twelve OPACs. The twelve OPACs possessed, on average, only 53% of the user-assistance features suggested for an ideal OPAC (table 9). The scores of the best OPAC and the worst OPAC were very polarized, the best system scoring 82%, and the worst scoring 24% (table 9).

Table 5
Output Control

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
1. Can the user select specific field(s) for display?												
2. When multiple records are retrieved in a single search, can the user select: (a) any single record for display?	x	x	x	x	x	x	x	x	x	x	x	x
(b) several records not in sequence for display? (e.g., record #2, #5, etc.)									x			
(c) a range of records for display? (i.e., by specifying the first and the last records. e.g., from record #5 to #9)									x	x		x
3. Can the results of several searches be merged for display?									x	x		x
4. Can the user specify that search results be sorted by: (a) author	x	x		x	x							
(b) title	x	x		x	x							
(c) subject		x		x	x							
(d) call number	x											
(e) date of publication	x											
5. Does the system support ranked document display in decreasing order of probable relevance to the search query?												
6. Does the system display results by paging?	x	x	x	x	x	x	x	x	x	x	x	x
Score: (maximum 12)	6	5	2	5	5	2	2	2	5	4	2	4

Two-thirds (eight) of the OPACs provided a list of accessible databases (table 6, question 1). All provided a list of search types (table 6, question 2). Only two offered an online tutorial (table 6, question 3). Eleven OPACs offered general help messages with information on various aspects of search strategies, but the messages in over half of the systems were not very helpful (table 6, question 4). Nine OPACs offered contextual help messages, of which eight were considered helpful (table 6, question 5). Ten OPACs offered routine procedural prompts, of which eight were considered helpful (table 6, question 6).

Only two systems provided an explanation of what the system was doing during a long search (table 6, question 7). Although all OPACs provided error messages, those in more than half of the systems were judged not clear enough (table 6, question 8). Although most OPACs (ten) displayed a message saying that users could ask a librarian for further help, only two indicated where users could get offline assistance, such as a detailed instruction sheet or a user's manual (table 6, question 9). Only two OPACs indicated to users how to edit input for typing errors, etc. (table 6, question 10). None of the OPACs offered a "spell check" function (table 6,

Table 6
User Assistance

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
1. Does the system provide a list of accessible databases?	x		x			x	x	x	x	x		x
2. Does the system provide a list of search types?	x	x	x	x	x	x	x	x	x	x	x	x
3. Is there an online tutorial?										x		x
4. (a) Are there general help messages, providing information on various aspects of search strategies, which can be called up at any point?		x	x	x	x	x	x	x	x	x	x	x
(b) Are they helpful?					x	x	x	x	x	x	x	
5. (a) Are there contextual help messages, specific to the point in the search reached by the user?	x	x	x	x	x	x	x	x		x		
(b) Are they helpful?	x	x		x	x	x	x	x		x		
6. (a) Does the system routinely provide procedural prompts or guiding comments to indicate possible next steps during a search?	x	x	x	x	x	x	x	x	x	x		
(b) Are they helpful?	x			x	x	x	x	x	x	x		
7. Is an explanation of what the system is doing displayed when searching takes a long time?	x									x		
8. (a) Does the system provide error messages?	x	x	x	x	x	x	x	x	x	x	x	x
(b) Are they clear enough?						x	x	x	x	x		
9. (a) Does the system identify who to ask if the user needs help?	x	x	x	x	x	x	x	x		x		x
(b) Does the system identify where printed instructions are available?	x											x
10. Does the system make it clear how to edit input?										x		x
11. Is "spell check" software available to the user?												
12. Does the system show the elapsed session time?												
Score: (maximum 17)	10	7	7	8	9	11	11	11	8	14	4	8

Table 7
OPAC Usability via Remote Access

Questions	A	B	C	D	E	F	G	H	I	J	K	L
	DOBIS	Geac				NOTIS			PALS	DRA	Home	Home
1. Is there adequate logon instruction (i.e., explain which terminal types are supported)?	x						x		x		x	
2. Are the contents and coverage of the OPAC clearly explained?	x	x	x	x		x	x	x	x	x		x
3. Are the key equivalencies explained for remote user's keyboard?	x	NA	NA	NA	NA		NA	x	NA	x	NA	x
4. Is there adequate logoff instruction?		x	x	x	x	x	x	x	x	x	x	x
5. Is the screen display always clean? (i.e., no garbage characters)	x	x	x	x	x	x	x	x	x	x	x	x
6. (a) Is remote access unrestricted in terms of time of day?		x	x	x	x	x	x	x	x	x	x	x
(b) Does the system tell the user if there is a time limit to remote sessions?												x
(c) Does the system give a warning message of automatic logoff if there is no user input?				x	x					x	x	
7. Does the remote user have access to the same OPAC as those who use dedicated terminals in the library?	x	x		x	x	x	x	x	x	x	x	
8. Does the system indicate where the remote user can get additional help?	x			x						x		x
Score: (maximum 10)	6	5	4	7	5	5	6	6	6	8	6	7

Note: "NA" means "not applicable."

question 11) or showed the elapsed session time (table 6, question 12).

Remote Access

Table 7 shows the data for remote access characteristics of the twelve OPACs. Although remote access was a fairly new function, these OPACs scored quite well in this area. They possessed, on average, 59% of the ideal OPAC's remote access features. Even the worst OPAC scored higher here than in most other categories except "Screen Display" and "Database Characteristics" (table 9).

Most OPACs (eleven) had clear log-off instructions (table 7, question 4). Ten OPACs explained the contents and coverage of their databases (table 7, question 2). All OPACs had clean screen displays (no garbage characters) (table 7, question 5). All OPACs but one offered unrestricted remote access in terms of time of day (table 7, question 6). Ten OPACs offered remote users the same content and coverage of their databases as those using terminals in the library (table 7, question 7).

Many OPACs can support several types of terminals. Before users can communicate with the OPAC

Table 8
Rankings of the Twelve OPACs

Rank	Institution	OPAC	Percentage of Features Desirable
1	J	DRA	67.1
2	I	PALS	59.4
3	H	NOTIS	58.2
4	F	NOTIS	57.1
5	E	Geac	56.5
6	G	NOTIS	55.3
7	D	Geac	54.7
8	L	Home-grown	54.1
9	B	Geac	53.5
10	C	Geac	44.1
11	A	DOBIS	41.2
12	K	Home-grown	38.8

properly, they must select the right type of terminal; however, only four OPACs provided adequate log-on instruction, including an explanation of what terminals were supported (table 7, question 1). Many OPACs limit a remote search session to a certain period of time (e.g., sixty minutes); however, only one OPAC in this study provided a message that there was a time limit on the remote search session (table 7, question 6(b)). Eight OPACs simply disconnected users when the time limit had expired (table 7, question 6(c)).

Since remote users may not use the same kind of terminal as those in the library, OPACs should explain key equivalencies to users, especially for function keys; however, only four did this (table 7, question 3). Seven did not appear to use function keys. Only four OPACs indicated where remote users could get help when in trouble (table 7, question 8).

Summary of Findings

The highest possible score (the ideal system) was 170. On average, the twelve OPACs scored 90.8 (53%); thus, on average, they possessed slightly more than half of the features of the ideal OPAC.

Table 8 shows the ranking of the twelve OPAC systems in order of the overall percentage of desirable features incorporated. The best system (that at university J) possessed 67% of an ideal system's features. Three systems (those at universities A, C, and K) possessed less

Table 9
Ranking of Categories

Rank	Section Name	Mean Percentage of Desirable Features	Minimum Percentage of Desirable Features	Maximum Percentage of Desirable Features
1	Screen display	79.2	50.0	90.0
2	Database characteristics	65.6	56.3	68.8
3	Commands	64.7	38.5	76.9
4	Remote access	59.2	40.0	80.0
5	Access points	56.4	38.1	71.4
6	Searching	53.1	21.7	71.7
7	User assistance	52.9	23.5	82.4
8	Operational control	41.2	29.4	64.7
9	Output control	30.6	16.7	50.0
10	Subject search aids	26.0	12.5	37.5

than half of the features of the ideal OPAC. The remaining systems possessed slightly more than half of an ideal system's features. Table 8 also indicates that, on average, NOTIS-based OPACs had more of the desired features than Geac-based OPACs. In general, the findings of this study show that although many suggestions had been made to improve OPAC functional capabilities and interface characteristics, the OPACs evaluated in this study still had major deficiencies in both areas.

Table 9 shows the ranking of the ten categories for the twelve OPACs in the order of their mean percentages of desirable features.

Generally speaking, all the systems had fairly high scores in the area of "Screen Display," while "Subject Search Aids" was the weakest category. "Remote Access" ranked fourth in development among the ten areas. In retrospect, we realized that some questions used would be more appropriate in other sections of the checklist, specifically question 9 in section 6 on "Screen Display" (limits on displays of hits), and that questions 9 to 12 in section 8 on "Commands" (leading articles, punctuation, word order for authors, and case) should be moved to section 3 on "Searching." There was also some overlap between the sections on "Database Characteristics" (section 1), "Subject Search Aids" (section 4), and "Access Points" (section 5). Examination of the data shows that these changes would not affect the overall order of ranking of the categories.

Conclusions and Future OPAC Research

This paper evaluated the functional capabilities and interface features of OPACs in twelve CARL libraries using a checklist. Data collected on the twelve OPACs were used as the basis for a comparison of the development of these OPACs. The ten areas evaluated were "Database Characteristics," "Operational Control," "Searching," "Subject Search Aids," "Access Points," "Screen Display," "Output Control," "Commands," "User Assistance," and "OPAC Usability via Remote Access."

While many improvements have been made in OPAC systems since they were first introduced, the results of this study show that there is still a wide gap between the systems evaluated in this study and the ideal OPAC system suggested by researchers. These OPACs, on average, had only approximately half of the features of the ideal OPAC. Among the ten categories evaluated, major shortcomings were found in the areas of "Subject Search Aids," "Output Control," and "Operational Control." Future research might focus on determining which of these major shortcomings users would rank as most important for successful searching.

As indicated earlier, the checklist itself could be revised to eliminate overlap and to move questions from the "Commands" and "Screen Display" sections to the "Searching" section, where they would be more appropriate evaluative measures. We point out that some questions were difficult to answer using the methodology of this study. For example, questions about "Database Characteristics," such as "Does the system include the following types of records? (a) monographs, (b) serials, etc.," and "Does the system display records for (a) items 'on order,' (b) items 'in process'" were originally included in the checklist. However, when actually answering these questions, we found that there was no convenient and reliable way to find the answers. The only possible method was to keep trying until relevant data were located (e.g., found a record of a map or an item "on order"); however, this would be too time-consuming. Originally also a number of questions related to printing were included; however, in many cases, since these systems were examined from remote access, it was difficult to tell whether the systems supported the functions or not. It was decided to eliminate these questions from the checklist.

In order to monitor the progress of OPAC development, studies such as the present one should be carried out every two or three years. By comparing data collected each time, it would be possible to monitor the progress made in each category, and by each system during a two- or three-year period. Based on the findings of the study at different times, patterns of OPAC development could be identified. By examining these patterns, trends might be predicted for the future development of OPAC features.

The major limitation of this study is the lack of weights assigned to the relative value of a particular feature. To address this shortcoming, future research could include a survey that would ask respondents to weight the relative importance of each feature. These weights could then be applied to this study's results to produce a ranking of catalogs which would not only reflect the number of features, but perhaps provide a truer picture of how well the catalog meets users' needs.[49]

References and Notes

1. Charles R. Hildreth, "Online Public Access Catalogues," in *Future of Online Catalogues: Essen Symposium, 30 Sept.–3 Oct. 1985*, ed. Ahmed H. Helal and Joachim W. Weiss (Essen: Essen University Library, 1985), 3–55.

2. Thomas A. Peters, *The Online Catalog: A Critical Examination of Public Use* (Jefferson, N.C.: McFarland, 1991), Introduction.

3. Ibid., 15.

4. Charles R. Hildreth, *Online Public Access Catalogs: The User Interface* (Dublin, Ohio: OCLC, 1982), 8.

5. Ibid., chapters 7–10.

6. Stephen R. Salmon, "Characteristics of Online Public Catalogs," *Library Resources & Technical Services* 27 (Jan./Mar. 1983): 36–67.

7. Emily Gallup Fayen, "The Online Public Access Catalog in 1984," *Library Technology Reports* 20 (Jan./Feb. 1984): 7–59.

8. Richard W. Boss and Susan B. Harrison, "The Online Patron Access Catalog: The Keystone in Library Automation," *Library Technology Reports* 25 (Sept./Oct. 1989): 635–726.

9. Xin Gu, *Checklist Evaluation of OPACs in Twelve CARL Libraries* (master of information science research project report, Univ. of Toronto, 1993).

10. Susannah Ravden and Graham Johnson, *Evaluating Usability of Human-Computer Interfaces* (New York: Wiley, 1989).

11. Hildreth, *Online Public Access Catalogs*.

12. Joseph R. Matthews, Gary S. Lawrence, and Douglas K. Ferguson, eds., *Using Online Catalogs: A Nationwide Survey* (New York: Neal-Schuman, 1983).

13. Salmon, "Characteristics of Online Public Catalogs."

14. Gill Dwyer, "Evaluating Online Public Access Catalogs: A Checklist," *State Librarian* 32 (Nov. 1984): 42–43.

15. Steven A. Northover, "A Checklist for Choice of an OPAC," *LASIE* 21 (Sept./Oct. 1990): 51–56.

16. Peters, *The Online Catalog: A Critical Examination of Public Use*, 3.

17. Boss and Harrison, "The Online Patron Access Catalog," 639–43.

18. Charles R. Hildreth, "Advancing Toward the E^3 OPAC: The Imperative and the Path," in *Think Tank on the Present and Future of the Online Catalog: Proceedings*, ed. Noelle Van Pulis (Chicago: American Library Assn., 1991), 37.

19. Joseph R. Matthews, "The Online Catalog: Time to Move Beyond the Boundary of a 'Catalog'!" in *Think Tank on the Present and Future of the Online Catalog: Proceedings*, ed. Noelle Van Pulis (Chicago: American Library Assn., 1991), 14.

20. Matthews, Lawrence, and Ferguson, *Using Online Catalogs*, 141.

21. Hildreth, *Online Public Access Catalogs*.

22. Hildreth, "Advancing Toward the E^3 OPAC," 26.

23. Ibid., 30.

24. Ibid., 29.

25. Jane Burke, "The Future OPAC: Content, Form and Boundaries," in *Think Tank on the Present and Future of the Online Catalog: Proceedings*, ed. Noelle Van Pulis (Chicago: American Library Assn., 1991), 46.

26. Karen Markey, "Users and the Online Catalog: Subject Access Problems," in *The Impact of Online Catalogs*, ed. Joseph R. Matthews (New York: Neal-Schuman, 1986), 61–64.

27. Carol A. Mandel, *Subject Access in the Online Catalog* (Washington, D.C.: Council on Library Resources, 1981), 5.

28. Tefko Saracevic, Hartmut Mokros, and Louise Su, "Nature of Interaction between Users and Intermediaries in Online Searching: A Qualitative Analysis," in *Proceedings of the 53d ASIS Annual Meeting* 27 (1990): 53.

29. Karen Markey, *Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs* (Dublin, Ohio: OCLC, 1984), 81.

30. Matthews, Lawrence, and Ferguson, *Using Online Catalogs*, 177.

31. Mandel, *Subject Access in the Online Catalog*, 7–8.

32. Markey, *Subject Searching in Library Catalogs*, chapters 4 and 5.

33. Ibid., 82.

34. Mandel, *Subject Access in the Online Catalog*, 9.

35. International PRECIS Workshop, "The PRECIS Index System: Principles, Applications, and Prospects," in *Proceedings of the International PRECIS Workshop*, ed. H. Wellisch (New York: Wilson, 1977), 197.

36. Charles R. Hildreth, "Online Browsing Capabilities," in *Proceedings of the 45th ASIS Annual Meeting* 19 (1982): 129.

37. Mandel, *Subject Access in the Online Catalog*, 13.

38. Joan M. Cherry, "Improving Subject Access in OPACs: An Exploratory Study of Conversion of Users' Queries," *Journal of Academic Librarianship* 18 (May 1992): 95–99.

39. Mandel, *Subject Access in the Online Catalog*, 15.

40. Pauline A. Atherton, *Books are for Use: Final Report of the Subject Access Project to the Council on Library Resources*, (Syracuse, N.Y.: Syracuse University School of Information Studies, 1978).

41. Mandel, *Subject Access in the Online Catalog*, 16.

42. Markey, *Subject Searching in Library Catalogs*, 91–93.

43. Matthews, Lawrence, and Ferguson, *Using Online Catalogs*, 140.

44. Karen Markey, "Users and the Online Catalog," 58–59.

45. Emily Gallup Fayen, *The Online Catalog: Improving Public Access to Library Materials* (White Plains, N.Y.: Knowledge Industry Publications, 1983), 7.

46. Neal K. Kaske and Nancy P. Sanders, "Evaluating the Effectiveness of Subject Access: The View of the Library Patron," *Proceedings of the 43d ASIS Annual Meeting* 17 (1980): 323–25.

47. Matthews, Lawrence, and Ferguson, *Using Online Catalogs*, 135.

48. Saracevic, Mokros, and Su, "Nature of Interaction between Users and Intermediaries in Online Searching," 53–54.

49. Recommendation of an anonymous referee.

Appendix: Evaluation Checklist for OPACs

The checklist consisted of ten sections, each of which focused on a different functional or user interface feature of an OPAC system. It was presented in tabular form with columns for yes/no answers and comments. The checklist questions are reproduced below in sections, each one corresponding to a separate table in the study. All questions were yes/no answers. If a particular question on the checklist was not relevant to the OPAC being evaluated (e.g., a question referring to the consistency of function key definitions for systems having no function keys), then evaluators were instructed to mark the question "Not Applicable." They were also encouraged to write comments about any question in the column titled "Comments."

SECTION 1
Database Characteristics

1. Does the system provide the following information in either short or long record displays?
   a. call number
   b. author
   c. title
   d. subject heading(s)
   e. edition information
   f. publication info.
   g. copy information
   h. volume information
   i. location for current issues of serials
   j. name of the library where the item is located
   k. loan status
   l. document abstract
   m. table of contents
   n. citations within the text
   o. keywords from the book index
   p. book review

SECTION 2
Operational Control

1. a. Is there an introduction to the system?
   b. Is there a set of instructions?
2. Does the system provide a choice of command-driven or menu-driven interface throughout?
3. If menu choices are by letters, are they mnemonic?
4. In the command mode:
   a. Is a display of the commands available?
   b. Are examples of how the commands are used available?
5. Does the system allow the registered user of this library to place holds/reserves on items at the terminal?
6. Does the system allow the user of this library view the list of items charged out to him/her at the terminal?
7. Does the system display system or database changes, system access notices or library news?
8. Can the user select the length of the system messages (e.g., long or short error messages, or help displays)?
9. Is there a "break" or "interrupt" key or command (e.g., Esc key)?
10. Can the user return to a previous screen and change a selection?
11. Does the system support downloading bibliographic records to the local user's personal computer?
12. Can the user communicate interactively with library staff through the OPAC whenever he/she has trouble finding material on a subject?
13. Can the user transmit search results via e-mail?
14. Does the system offer online mailboxes for user suggestions and comments?
15. Can the user send messages to a specific library staff member or department?

SECTION 3
Searching

1. Can the user set default values for:
   a. search type (e.g., A, T, S)?
   b. field(s) in which a keyword is to be searched?
   c. display format?
   d. dialogue mode (command or menu)?
2. Can the user reset the default values during a search session?
3. Can the user start a search anywhere?
4. Can the user continue or start a search directly from the HELP screen?
5. Does the system:
   a. permit keyword searching of the controlled vocabulary (i.e. keyword subject search & keyword author search)?
   b. support keyword title search?
6. Does the system support keyword searching anywhere in the record?

7. a. Does the system have stop words?
   b. Is there a list of stop words available for display?
   c. Does the system indicate that the word is not indexed, when the user tries to search a stop word?
8. Which of the following Boolean operators are available?
   a. AND
   b. OR
   c. NOT
9. When is Boolean searching supported?
   a. in keyword author search
   b. in keyword title search
   c. in keyword subject search
   d. in keyword search not limited to any fields
   e. in cross-fields searches (i.e. two or more fields)
10. Is there an unlimited number of Boolean operators which can be used in a single search?
11. Is a word adjacency operator available?
12. Is a word proximity operator available?
13. In multi-word searches, does the system explain that a space is the same as AND or ADJ?
14. Can the user specify:
   a. left-truncation (e.g., #ism)?
   b. right-truncation?
   c. wildcard characters (e.g., WOM#N)?
   d. variable length wildcard characters (e.g., BEHAVI#R gets both BEHAVIOR and BEHAVIOUR)?
   e. user specified limits on truncation (e.g., LIBRAR*3 to get LIBRARY and LIBRARIES but not LIBRARIANSHIP)?
15. Does the system support weighted term search by ranking the search terms by:
   a. the user?
   b. the system?
16. Can searches be limited by:
   a. publisher?
   b. type of material (e.g., serials, monograph, etc.) ?
   c. library location?
   d. date of publication?
   e. language of publication?
17. Does the system allow the user to indicate which of the retrieved records are relevant to the search question and use the feedback information to automatically generate searches based on some algorithm to locate other items in the collection that are similar to the relevant record?
18. Having retrieved a document record, can the user directly access another document cited in it?
19. Can the user "browse" up a list of index terms which are near the search term/phrase:
   a. in author search?
   b. in title search?
   c. in subject search?
   d. Do the indexes include cross-references?
20. Can the user save a search strategy to be used again later?
21. Can the user save search results in sets for later use?
22. Can the user easily switch from one type of search to another (e.g., author search to title search)?

SECTION 4
Subject Search Aids

1. Can the user browse a display of:
   a. classification outlines?
   b. classification schedules?
2. Can the user view a group of subject headings:
   a. which begin with the search term(s)?
   b. which include the search term wherever imbedded in the subject headings?
3. Does the system display the following cross-references:
   a. SEE/USE?
   b. SEE ALSO/BT/NT/RT?
4. Does the system have transparent SEE/USE references which automatically substitute the user's input term with the correct subject heading without informing the user?
5. Does the system convert an original zero hit subject search to title, keyword title or keyword subject search?

SECTION 5
Access Points

Does the system provide access via the following?

1. Personal author
2. Corporate author
3. Author/title
4. Title
5. Subject
6. Series
7. Notes
8. Author keyword
9. Title keyword
10. Subject Keyword
11. Notes Keyword
12. ISBN
13. ISSN

14. Government document number
15. LC card number
16. LC call number
17. Dewey Decimal Classification number
18. Circulation bar code number
19. Table of contents of books
20. Citations within the text
21. Indexes of books

SECTION 6
Screen Display

1. Does the display text use both UPPER and lower case?
2. Are the fields of the display labeled: (e.g., AUTHOR=)
   a. in brief bibliographic display?
   b. in long bibliographic display?
3. Is the number of hits retrieved reported before they are displayed, so that if more hits have been retrieved than are really wanted, one of the limiting devices can be used?
4. Does the system offer both brief bibliographic display and long bibliographic display?
5. Is the search request always displayed on the screen, so that the user can see what was typed while viewing the hits?
6. Is the circulation status of an item always shown on the same screen with its call number?
7. Is the total number of items to be displayed identified in the display of each item (e.g., item 1 of 100)?
8. Are items in a set numbered successively (e.g., 1 to 8, 9 to 18, etc.) when there are more items than can be displayed on one screen?
9. Are the limits to the number of hits which can be displayed equal to 150(± 50)?

SECTION 7
Output Control

1. Can the user select specific field(s) for display?
2. When multiple records are retrieved in a single search, can the user select:
   a. any single record for display?
   b. several records not in sequence for display (e.g., record #2, #5, etc.)?
   c. a range of records for display (i.e. by specifying the first and the last records. e.g., from record #5 to #9)?
3. Can the results of several searches be merged for display?
4. Can the user specify that search results be sorted by:
   a. author?
   b. title?
   c. subject?
   d. call number?
   e. date of publication?
5. Does the system support ranked document display in decreasing order of probable relevance to the search query?
6. Does the system display results by paging?

SECTION 8
Commands

1. Does each command have the same role in every context?
2. Are function key definitions consistent (e.g., F1 always invokes help)?
3. Can function keys be used to reduce the number of keystrokes required to enter commonly used commands?
4. Do all the commands have a standardized syntax?
5. Is the number of keystrokes kept to a minimum?
6. Is there minimal or familiar punctuation in the commands (i.e., 3)?
7. Are mnemonic abbreviations used for the commands (e.g., *A* for Author)?
8. Can commands be stacked (i.e. typing in several commands in a group and executing them all at once, e.g., to indicate that the search results are to be displayed in a brief format, arranged alphabetically by title)?
9. In title or title/author searches, are leading articles ignored by the system?
10. a. Can the user omit "—" for LCSH subdivisions?
    b. Does the system ignore punctuation entered by the user when they are not required?
11. Will the system accept an author's name in any order (e.g., Smith A or A Smith)?
12. Can searches be entered using a mix of UPper and lowER case?

SECTION 9
User Assistance

1. Does the system provide a list of accessible databases?
2. Does the system provide a list of search types?
3. Is there an online tutorial?

4. a. Are there general help messages, providing information on various aspects of search strategies, which can be called up at any point?
   b. Are they helpful?
5. a. Are there contextual help messages, specific to the point in the search reached by the user?
   b. Are they helpful?
6. a. Does the system routinely provide procedural prompts or guiding comments to indicate possible next steps during a search?
   b. Are they helpful?
7. Is an explanation of what the system is doing displayed when searching takes a long time?
8. a. Does the system provide error messages?
   b. Are they clear enough?
9. a. Does the system identify who to ask if the user needs help?
   b. Does the system identify where printed instructions are available?
10. Does the system make it clear how to edit input?
11. Is spell check software available to the user?
12. Does the system show the elapsed session time?

SECTION 10
Remote Access

1. Is there adequate logon instruction (i.e., explanation of which terminal types are supported)?
2. Are the contents and coverage of the OPAC clearly explained?
3. Are the key equivalencies explained for remote user's keyboard?
4. Is there adequate logoff instruction?
5. Is the screen display always clean (i.e., no garbage characters)?
6. a. Is remote access unrestricted in terms of time of day?
   b. Does the system tell the user if there is a time limit to remote sessions?
   c. Does the system give a warning message of automatic logoff if there is no user input?
7. Does the remote user have access to the same OPAC as those who use dedicated terminals in the library?
8. Does the system indicate where the remote user can get additional help?

Reserve On-line: Bringing Reserve into the Electronic Age

Halcyon R. Enssle

The virtual library, libraries without walls, the invisible user . . . these are some of the terms being used to describe the library of the future or the library of the twenty-first century. Much of the attention on features of these libraries has been centered on the OPAC, CD-ROM databases, and electronic publishing. Online reserve systems began to be utilized in the late 1980s, but even those librarians who have foreseen the incredible possibilities for libraries utilizing campus networks, the Internet, and networked CD-ROMs to provide new kinds of services to their users have focused on access to bibliographic and full-text databases, with expanded document delivery. Material placed on reserve has remained in traditional formats, and a search of the literature has not revealed any attempt to present reserve material in an alternate manner, with the exception of the San Diego State University project. Most efforts to automate the reserve operation have focused on processing.

In the summer of 1992, however, the Association of Research Libraries (ARL) proposed a Reserve Materials Publishing Project, which is designed to explore and develop new services to meet the instructional needs of students. New advances in capabilities for storage, retrieval, networking, and manipulation of images are providing exciting possibilities for their use in traditional library arenas such as reserve rooms. It was recognized that much of the traffic in reserve was for the purpose of checking an item out only to photocopy it and return it. Reserve processing is time-consuming, and problems of loss and damage are frequent. To see whether reserve materials could be provided in alternate ways, ARL proposed a project that included the following objectives:

> Explore, and measure faculty and student acceptance of, a changing paradigm for the provision of instructional support materials that traditionally have been made available through reserve book rooms . . . .
>
> Provide instructional support materials of the highest quality to students at the lowest possible cost . . . .
>
> Evaluate a variety of mechanisms for providing instructional support materials in a number of formats (paper-to-paper, paper-to-electronic, or electronic-to-electronic products) . . . .
>
> Explore copyright issues that affect the various participants . . .
>
> and
>
> Develop a prototype and vision for providing instructional support materials within an electronic environment.

Libraries were also encouraged to tailor projects according to the local practices and needs of their universities. Colorado State University has a well-developed electronic information network and is connected to the WAIS gopher. The University Libraries has its own LAN and utilizes the CARL online catalog, which will in the near future provide online access to reserve course lists. Currently faculty and students are able to submit interlibrary loan requests electronically through the campus network. Bibliographic information is pulled from the online catalog and transferred to an online form, which can then be forwarded to our ILL department. We wanted to develop a similar mechanism for the provision of reserve materials to students.

To achieve this, the author drafted a proposal that had the objective of providing instructional-support materials within an electronic environment. By scanning or downloading course reserve material that would normally be photocopied into a database that would be available on the campus network, students would be able to access reserve material through their own microcomputers or through microcomputers available in laboratories. They would be able to read, print, or download the information without paying photocopy charges. An ultimate goal of the project is to enable a student to access the reserve list on CARL, our online catalog; choose an item identified as being available electronically; pull the text up on their own PC or a CARL PC terminal; and either read, print, or download the text. Reserve material will be available twenty-four hours a day, from remote sites and at no cost to the student. Although we were aware of other projects making reserve material available through a computer, such as the project under way at San Diego State University, this project does not currently provide access over the campus network and charges students for copies of the reserve readings. Our goal was to provide wider access across campus and, if possible, at no cost to the student.

The proposal was accepted by ARL; however, funding was not available through this project, so the libraries administration agreed to fund the cost of the pilot project. It was thought that beyond the cost of the equipment (a 486/33 PC with 16 MB RAM and a 200 MB hard drive, HP Scanjet scanner, and Omnipage Professional software) the primary cost would be for programming. However, in discussions with my staff, staff member Tom Delaney volunteered to do the programming, as he had the needed UNIX expertise. It was agreed that release time from his normal duties would be provided and that he would work on programming for the project.

Work began in October 1992, first in designing the Reserve On-line program and then in actual programming. The programming was done in UNIX and was quite complex. Since the files were to be available on the

Halcyon R. Enssle is a planning librarian at Colorado State University's Morgan Library.

campus network through Lamar, the university mainframe, it was essential that access to files be limited to the actual readings themselves and that there was no possibility that users could access the programs. The program itself is responsible for creating and manipulating the directories and subdirectories that move the user through the program. Reserve On-line also is able to track the number of times each reading is accessed. Reserve On-line is currently allocated 20 MB of space on the university computer, and to date this has proved quite adequate. If use of the service expands, however, an additional 20 MB is available. Files will not remain online beyond the end of a semester. It is planned to download and store the files when the course is removed from reserve, making space available for new files.

Once the preliminary programming was done and testing had shown that the concept was feasible, work was done on refining the pathways and menus. We wanted the program to be as user friendly as possible and so worked very hard at designing clear, easy-to-understand menus. The programming was completed in January, and the actual pilot project began.

The process begins with the identification by the professor of the material to be put on reserve. In the pilot project this was restricted to material that was not under copyright, such as course notes, syllabi, and solutions. The material is scanned, using the HP Scanjet II, and then edited, if necessary, using Omnipage (an optical character recognition software able to display any ASCII text files that are submitted). Omnipage allows us to bring the document into WordPerfect and edit the document online before downloading it. Once the editing (which is usually minimal) is completed, the document is uploaded into the Reserve On-line program on Lamar and placed in the appropriate file for the course for which it is assigned. If the material is already on disk, the faculty member has the option of sending the file by e-mail or sending a copy on disk. The files will be uploaded directly into the Reserve On-line program.

Through Lamar and the CSU gopher, any student with access to a computer or a suitably equipped computer lab is able to view the Reserve On-line files. No computer account is needed, and the service is available to all students. To monitor effectiveness and to ensure access is open to all, any material put online is also available in paper form at the reserve desk.

If the student has access to Lamar, once at the Lamar prompt all that is necessary is to select the reserve prompt. A second option is provided through the CSU gopher. Reserve is a menu item on the CSU network menu (see figure 1) and through gopher. Students may go into the reserve course readings, where they are given a menu of courses available on Reserve On-line (see figure 2). Selection of the course number takes them to a menu showing the readings available (see figure 3). When they have selected the reading, they can then scroll through the entire reading. If they have the appropriate software, they can download the reading for later review. When through reading the chosen selection, students may return to the course menu to choose another reading or exit the program.

A second option also available in electronic reserve is a new service providing faculty and teaching instructors the option of sending reserve list requests electronically. Access to the electronic reserve request program is also through the university network. Faculty need only enter information on items to be placed on reserve and the length of the checkout period (see figure 4). The requests will then be forwarded by e-mail, downloaded, formatted, and processed for reserve.

The pilot project for Reserve On-line consisted of readings for five courses. At the end of the semester, evaluation forms were distributed to the professors and students in the classes involved, asking for feedback on the project. Students who used Reserve On-line were extremely enthusiastic and encouraged us to develop the project further. A surprising number did not use the service because they did not have a computer and were not aware of access to computers through the many labs on campus. The disappointing results were that many of the students in the classes reported that they did not use the service because they were not aware of it. Since this was a pilot project, and we were not sure of our ability to respond if we became overwhelmed with requests, we did not advertise the service widely in campus media. We did have publicity concerning the project at the reserve desk, but we relied primarily on the professors to advertise the service. This does not work. We realize that for the next round we want to send staff into the participating classrooms and make sure that all students are aware of the service and exactly how to access it. All of the professors who used the service and responded to the questionnaire were enthusiastic and will use the service again.

Comments received from professors inquiring about the service tell us that there is a great demand for the service as soon as copyrighted material can be placed online, so the second phase of the pilot project will be to explore inclusion of copyrighted materials.

Investigation to date shows that the issue of copyright of electronically reproduced material made available for educational use only is far from being resolved. Only one article, by Mary Brandt Jensen (1993), has directly analyzed the issue of electronic reserve and copyright. Jensen found that the use of scanned material stored on a central computer and available to many patrons falls under section 106 of the copyright law concerning exclusive rights reserved to the copyright

```
IMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
     :Colorado State University Wednesday, January 31, 1994. 10:21am
LMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
     ::   H - Host name lookup
     ::
     ::   P - Ping
     ::
     ::   I - ILL E-MAIL (login as: ill)
     ::
     :    T - Telnet:
     ::
     :    A - AIS:
     ::
     ::   F - Ftp
     ::
     ::   G - Gopher
     ::
     :    C - CARL:  R - Reserve Online -Login: reserve
     ::
     :    M - CSU Electronic Directory:
     ::
     :    L - CSU UCC PACX:

HMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
     H:\HENSSLE
     1 Go to Shell; 2 Clipboard; 3 Other Dir; 4 Setup; 5 Mem Map;
     6 Log: (F7 = Exit)
```

Figure 1
Network Menu with Reserve Option

holder and under section 107, fair use. She found particularly in section 107 that nothing in this section "or its legislative history indicates that Congress intended to limit fair use copying to facsimile form." Furthermore, she found that if the scanning is done to facilitate educational purposes, it also is defined under fair use. It will be important to monitor the character of the material scanned. The copying of journal articles for reserve traditionally has been considered to fall under fair use. However, if the material is consumable, such as workbooks, or is intended to be a compilation that would replace a textbook, then the fair use clause would probably not apply. Jensen's final conclusion is that sections 107 on fair use and 108 on reproduction by libraries must be relied on to "sanction the making of the electronic copy."

In line with making copyright material available under the guidelines of fair use, we are investigating limits to access, such as restricting access to students actually enrolled in the class. We are also looking into what may be involved in subsidizing the copyright costs for material put on Reserve On-line. We hope to expand the pilot project to a limited number of copyrighted articles in the fall. ARL has taken on this issue, and perhaps through this avenue some further guidelines will develop.

Still another line of service we are pursuing involves making available graphic material that cannot be accommodated by our current software. Graphics files (such as maps, drawings, and photographs) require equipment to scan, store, and display such material. We have acquired two Macintosh SV II's that we will link in a mini-network within the library. We will use the HP scanner to scan graphic material and the Macintosh software Deskscan II to edit and then make available this material at one workstation within the library. The university network does not currently completely support Macintosh in the networked environment, so we cannot connect this service to the network.

Progress on the pilot project has been rapid. The final step in developing a completely integrated system will be working with CARL Systems Inc. to develop necessary software for the second phase—accessing course reserve lists through the online catalog. Students would be able to search the online catalog, choose an item identified as being available on Reserve On-line, pull the text up on their own terminal or a CARL PC terminal, and either read, print, or download the text.

```
AG AGRONOMY

BK MARKETING

EE ELECTRICAL ENGINEERING

LI LIBRARY INFORMATION

HO HORTICULTURE

Enter the 2 LETTER ABBREVIATION for the department you want to read.
Or Press "X" to exit from Reserve.

                                                        lamar 10:24:51
```

Figure 2
Courses Available on Reserve On-Line

```
1) EE362

Select a NUMBER for the selection you want to read.
Or enter X to eXit, or B to go Back to MAIN MENU

                                                        lamar 10:25:45
********************************************************************************

1) Homework #9
2) Homework #10
3) Homework #11

Enter the NUMBER for your selection.
X to exit from RESERVE,
or P for PREVIOUS menu

                                                        lamar 10:25:55
```

Figure 3
Course Selected with Readings Available

```
You may place Library books on reserve.  In addition,
if you have personal articles or materials that we do not
own, you may place them on RESERVE as personal copy material.

In order to provide you with more efficient service,
we will retrive Library material from the stacks for your list,
and you may send us personal copy material via campus mail.

If you want to leave without entering data, press (CNTRL+C).
P) to place your personal copy material on reserve.
L) to place library books on reserve.


                                                        lamar 10:27:49

FTP Software PC/TCP tnvt Version 2.04 pl2
Copyright (c) 1986, 1990 by FTP Software, Inc.  All rights reserved.
Escape character is F10

*******************************************************************************


If you make an error, or leave a field blank
don't worry: you'll be prompted to verify all information
before we continue.

IF YOU WANT TO EXIT, PRESS (CNTRL+C)

Please enter your LAST nameEnssle
Enter your SSN or university id number123456789
Enter DepartmentLibrary
Enter office or department telephone number4911865
Enter course numberLI 300

                                                        lamar 10:29:06
```

Figure 4
Initial Menu and Personal Information Required for Placing Material on Reserve

This may take quite some time to develop, as CARL is currently developing a new version of the reserve software and we will not begin working on this phase until this software is completed and in release. Negotiations must also be made with CARL to support the project.

Response to the project has been very positive. It is an exciting next move toward the library of the twenty-first century.

References

1. Susan D. Gyeszly, "Reserve Departments and Automation: A Survey of ARL Libraries" *Information Technology and Libraries* 7 (Dec. 1988): 401.

2. Ray Metz, "Integrating Local Library Systems and Services into a Campus Network Environment," *Computers in Libraries* 10 (June 1990): 19.

3. Don L. Bosseau, "Photo Essay: Anatomy of a Small Step Forward: The Electronic Reserve Book Room at San Diego State University," *Journal of Academic Librarianship* 18, no. 6: 366–68 (Jan. 1993).

4. Mark Stover, "Using PC-File: Db for Reserves in an Academic Library," *Computers in Libraries* 9: 11–12 (Sept. 1989); James S. Aagaard and Elizabeth J. Furlong, "Automating Reserve Activities at Northwestern University," *College & Research Libraries News* 51 (Feb. 1990): 98–101; Gregory A. Crawford, "Using a DBMS to Maintain a Reserve Reading Room," *College & Research Libraries News* 51 (Oct. 1990): 860–62; James Self, "The Automation of Reserve Processing," *Information Technology and Libraries* 4 (Sept. 1985): 215–19.

5. Duane E. Webster, "Request for Participation," electronic mail message from the ARL/NACS Reserve Materials Publishing Project Meeting, August 21, 1992.

6. Mary Brandt Jensen, "Electronic Reserve and Copyright," *Computers In Libraries* 13: 40–45 (Mar. 1993).

7. Ibid., 42.

8. Ibid., 44.

Expert Systems in Technical Services and Collection Management

Carol Pitts Hawks

This manuscript examines the topic of knowledge-based systems, particularly expert systems in the areas of technical services and collection management. Background is provided on the goals and purposes of expert systems, their components, selection of appropriate domains, the process of knowledge acquisition, and development issues. The manuscript focuses on existing research projects in expert systems related to collection management, acquisitions, serials, cataloging, and preservation.

Technical services operations have traditionally evaluated their effectiveness in terms of quantity—number of orders placed, number of serial issues checked in, number of titles cataloged. Recent budget cuts and technological changes have, however, required libraries to redefine effectiveness. Many have begun to adopt current management thinking, which considers an organization's ability to adapt to its environment in addition to quantitative measurements. The ability to adapt may be increased through the empowerment of staff to take control of issues and resolve them without extensive hierarchical intervention or approval. This decentralization of decision making enables library staff at all levels to respond rapidly and innovatively.

Frank D'Andraia notes another compelling issue that will affect the future of technical services. He predicts a major staffing crisis in academic libraries in the 1990s, for technical services in particular. The focus is now on knowledge skills rather than clerical skills. Staff work very independently and must understand the larger processes in addition to their specific tasks. Traditional clerical work has been replaced with more interdependent and varied automated tasks.[1] What technical services manager has not seen the increasing need for acquisitions staff who understand OCLC, or copy catalogers who can interpret acquisitions records?

One of the ways to capture and build on existing expertise is through the use of knowledge-based systems. Although there are exceptions, these systems are not in wide use in libraries today. Certain aspects of acquisitions, serials control, collection management, cataloging, and preservation are ideal candidates for artificial intelligence development. However, as the following remarks indicate, little has been done outside the cataloging arena to develop knowledge-based systems in these areas.

Knowledge-Based Systems

Background

What exactly are knowledge-based systems? Knowledge-based systems are the broad category of systems that use some knowledge to perform their functions. They need not use either heuristics (rules of thumb) or artificial intelligence techniques in performing their tasks. Knowledge-based systems may be as simple as an online personnel manual that is easier to use because it can be searched more efficiently. Intelligent systems are a subset of knowledge-based systems. They display intelligent behavior, but not necessarily at the level of an human expert. The same online personnel manual may fit into the category of the intelligent system if it includes linkages from one policy to another, for example, between salary schedules and pay classifications. Expert systems, a more specific category, use heuristics to perform tasks previously done by human experts.[2] The "expert" version of the personnel manual would incorporate the personnel librarian's knowledge of classification to allow the novice supervisor to select the appropriate classification after answering a few questions asked by the system. In essence, a well-developed expert system should provide the same answers that an expert would give when approached with a particular problem.

Use

Expert systems have five primary uses:

1. To make existing expertise more readily available, particularly in multiple locations;
2. To reach new levels of expertise by accelerating complex problem solving;
3. To free the expert to handle difficult cases while the system handles the more routine ones;
4. To preserve expertise that might be lost through retirement or resignation; and
5. To enhance training through observation and analysis of the reasoning used by the system to reach its decisions.[3]

In more colloquial terms, "the computer can answer queries when the expert gets tired or takes a vacation; it doesn't forget key components when under

Carol Pitts Hawks is Head, Acquisition Department, Ohio State University Libraries, Columbus. An earlier version of this presentation was given at the Ninth Texas Conference on Library Automation, April 2–3, 1993, Houston, Texas.

pressure. A well-written and well-formulated expert system will perform consistently and 'mindlessly.'"[4]

Expert systems can be developed to function in a variety of roles within an organization. As a consultant, the system provides consistent advice to specialists or serves as an advisor to less-experienced staff who need access to particular expertise. As a checklist, the system reminds the user of factors to be considered and prompts with questions related to the problem. As a trainer, the system can provide initial training to the less-experienced or improve the expertise of the experienced user. As a communicator, the system serves as a ready reference to information that can be easily located and cross-referenced.[5]

Appropriate Domains

How does a designer determine what would be an appropriate domain or problem for an expert system to address? The literature is fairly consistent in its definition of appropriate domains for expert systems:

1. The scope and domain should be quite finite and bounded.
2. The problem and its solution should have a logical structure, i.e., the solution should not depend upon the use of common knowledge or everyday know-how.
3. The problem should be repetitive, yet sufficiently complicated to warrant the effort of creating the program. One conventional standard is to tackle problems that would take an expert more than five minutes but less than thirty minutes to resolve. Make sure there is not another, more cost-effective way to offer the information. Very complicated problems should be left to the now-liberated expert.
4. The solution should be clear-cut and not involve opinion.[6]

One additional factor is uncertainty. If there are not a number of possible answers to a given question, then the problem is too simple for an expert system. A non-expert would see that a group of possibilities were plausible solutions to the problem; the expert would be the person most likely to select the most promising solution.[7]

Components

The basic structure of an expert system includes a knowledge base, an inference engine, a user interface, and some form of explanatory capability to demonstrate the reasoning behind the conclusions. The essential components of the system are the knowledge base and the inference engine. The knowledge base must include all the information and strategies needed to solve the problem. The inference engine is the logic or reasoning portion of the system that enables it to use the knowledge base to reach conclusions.[8]

Knowledge Acquisition

Knowledge acquisition or knowledge engineering is the process of obtaining and organizing expert knowledge to build the knowledge base. This process is widely recognized as the bottleneck in the development of expert systems. Alberico and Micco suggest that reference librarians and catalogers have ideal skills to pursue the two portions of this process. Obtaining information through tact and interpersonal skills is well-suited to the skills of reference librarians, and catalogers are expert at organizing and classifying knowledge.[9]

Interviewing is the most common technique used to gather information from an expert. Observation of the expert at work can be used to supplement the interview process, but observation alone is rarely sufficient. The single greatest factor that handicaps the interview process is the expert's inability to explain or understand the thought processes used in the decision-making process. "Often the expert claims to 'just know' how to do certain things; and it's extremely difficult to reduce intuitive logic like that to a decision rule, or to relate it to other decision rules."[10]

Expertise is also a moving target—it changes on a regular basis. Just as the human expert has to keep up with changes in his field of expertise, the knowledge engineer's job is not complete once the first knowledge base is constructed. Knowledge bases must be continuously revised and maintained.

Methods

The inference engine was mentioned earlier as the logic or reasoning portion of an expert system. Ford identifies nine structures used in developing these inference engines: rule-based systems, semantic networks, frames, object-oriented systems, grammar rules, reasoning with uncertainty, blackboards, machine learning via automatic rule induction, and machine learning via neural networks.[11] Rule-based systems, which use rules as the main form of knowledge representation, are the simplest type of intelligent systems to develop and the most common inference engine used in expert systems for libraries. A rule-based system might be designed to do no more than indicate a library's opening hours, taking into account weekends, holidays, and end-of-term periods. The IF-THEN rules would be very clear-cut. For example:

If today is a bank holiday or today is a Sunday, then the library is closed.

If today is vacation and today is a weekday, then the library is open 9–5.[12]

Development Issues

There are additional development issues beyond selection of an appropriate domain and inference engine. Cost can be a deterrent. Staff time is the greatest expense entailed in developing an expert system, although the decision to hire external expertise or train staff internally will also increase the costs. However, more sophisticated software and more powerful hardware are bringing these costs down.

The selection of tools to be used in developing the system can significantly affect cost. Tools for expert systems development come in two varieties: programming languages and shells. To use a programming language, the designer needs to know how to write programs. Shells, by contrast, are often designed with the nonprogrammer in mind and require no programming knowledge. (A shell may be thought of as a knowledge-based system without the domain knowledge. The framework already includes the inference engine and the user interface. All the developer has to do is add the knowledge base and define the problem to be solved.)[13] A relatively simple system should not require an expensive shell. "Conversely, to attempt to use a cheap, rule-based shell to create a system requiring more sophisticated knowledge representations and reasoning mechanisms is likely to be very expensive in terms of staff time and frustration, and unlikely to result in a useful product."[14]

Most importantly, using a shell frees the developer to focus on the construction of the knowledge base. Carrington concludes that "using an expert system shell . . . is about as difficult as learning a word processing program such as WordPerfect or Microsoft Word."[15] Both Quinn and Alberico and Micco[16] provide excellent guidance on evaluating expert system shells.

Applications

The primary reason for developing expert systems for technical services (acquisitions, serials, cataloging, preservation) and collection management is to bring the improvements that technology can provide to bear on existing tasks. As with the introduction of new management tools, such as Total Quality Management, if it helps the library provide better service or work smarter, it is worth exploring.

Cataloging

Excellent summaries of the first known expert systems in cataloging exist throughout the literature.[17] Jeng has provided much new information on understanding human interpretation of bibliographic data by analyzing two hundred title pages for visual and linguistic cues.[18] Researchers have also identified the deficiencies of various expert system methods, such as classification with simple rule-based systems.[19] Cataloging has long been considered a bottleneck in the process of attempting to get newly acquired material to the patron in a timely fashion. Thus, cataloging would seem to be fertile ground for the development of expert systems. Cataloging involves three basic activities: describing the item and choosing access points for names and titles, assigning classification numbers, and assigning subject headings. Most of the research to date has focused on the first activity, descriptive cataloging, because it uses a rule-based cataloging code, *AACR2*. However, as the following discussion of existing prototypes will confirm, *AACR2* is now considered too unmanageable to be used in its entirety for a single expert system.

OCLC Automated Title Page Project

One of the most significant projects in cataloging is the OCLC Automated Title Page Project. OCLC's study examines the viability of scanning title page information into an automated cataloging system. The system would evaluate this data and produce a first-level bibliographic description as defined by *AACR2*. Three of the seven elements needed for the first-level description can usually be found on the front of the title page. The development and evaluation of this project resulted in only a moderate success rate, with 73 percent of the bibliographic elements being identified correctly by the system. Even so, OCLC was encouraged by the results and believes that it is feasible to automate much of the initial effort involved in descriptive cataloging.[20]

Anticipating that the currently limited sophistication and reliability of optical character recognition (OCR) will be improved, OCLC has continued to build upon its work with title pages. Stuart Weibel has developed a more complex system to automatically catalog and index the structural components of documents by using OCR technology to read text. The system identifies bibliographic components of the documents, such as author and title, as well as structural components, such as abstracts and indexes, and encodes the components in accordance with the Standard Generalized Markup Language (SGML). Weibel expects that the system will be able to produce descriptive catalog entries, to convert existing catalog cards to MARC records

automatically, and to index documents for full-text retrieval systems.[21]

MAPPER

MAPPER is an experimental cataloging advisor developed by Zorana Ercegovac as her doctoral dissertation at UCLA. This prototype uses the domain of descriptive cataloging for maps, an expensive and complex portion of the cataloging process. MAPPER was developed for two purposes: to make expert advice available to novice map catalogers and to improve conventional instruction in map cataloging. The system is essentially an interactive tutor that asks specific questions to elicit information, advises, explains its decisions, and provides the MARC tags and subfields. After the session, it displays a completed catalog entry and asks the user to verify each element.[22] It integrates cataloging rules related to maps from *Anglo-American Cataloguing Rules, Cartographic Materials: A Manual of Interpretation for AACR2,* and *Library of Congress Rule Interpretations,* as well as expert knowledge from map catalogers at LC.

The system was tested by library school students who had completed the introductory cataloging class but had no other experience in map cataloging. Three cataloging tasks were tested: determining the person or corporate body responsible for the map, determining the title and statements of responsibility, and deciding on the elements to be included in the publication area. As expected, students assisted by MAPPER provided significantly better cataloging answers than those not assisted by MAPPER.[23]

Jeng's Title Proper Project

For her doctoral dissertation, Ling Hwey Jeng investigated the linguistic and typographical properties of two hundred title pages in order to construct an expert system that could correctly identify the title proper. Information on the title page was blocked into nine sections: author, author affiliation, edition, other title information, place, publisher, series, title proper, and year.[24]

ShelfPro

ShelfPro, developed by Karen Markey Drabenstott and others, addresses another aspect of the cataloging process—shelflisting.[25] Shelflisting is concerned with assigning a book number, as opposed to the class mark portion of the call number, to an item. The manual shelflisting process involves (1) examining the book itself, (2) answering questions about its relationship to other books in the collection, and then (3) deciding what device to give the book so that it will appear in its proper place on the shelf. The system asks questions concerning steps 1 and 2, then, based on the answers provided, suggests appropriate shelflisting devices for the book (step 3).

In actual practice, the cataloger must assign the class mark of the book in hand first. If the book falls within the classification section used for this prototype (computer science, QA75 to QA76.95), consultation on ShelfPro begins. To initiate the cuttering, the system prompts the user for the first letter of the principal author's surname. Once the system suggests the appropriate author Cutter number, the user is advised via an information window to check the shelflist to make sure the number is unique. The system also prompts for other characteristics of the book that would affect its shelflisting—whether it is a translation, a conference or congress proceeding, etc.

ShelfPro would save time and improve consistency if incorporated into a library's automated library system, particularly its online shelflist. Consistency of shelflisting would improve because everyone involved in the process would be answering the same sequence of questions. Practices traditionally passed on in internal memorandums or by oral transmission would be documented. The online shelflist would be available to every workstation, eliminating the need for a manual file. Finally, ShelfPro would automate the practice of creating author and topical Cutters from the printed Cutter-Sanborn tables.

CATALYST

CATALYST is a simpler, less sophisticated system developed at Strathclyde University that serves as a consultant to the cataloger on the choice and form of access points. The user moves through a sequence of menus that ask questions related to access points, such as whether the authorship is mixed, unknown, or shared. Based on the user's selections, the system responds with advice as to how the main entry should be made and whether additional entries are needed. CATALYST does not produce a catalog entry but rather helps the user rapidly find and display the appropriate rules from *AACR2* to generate the entry. The system is designed for both novice and expert catalogers and so must balance the amount of explanatory information given to provide enough for the novice without boring the expert.[26]

CATALYST is a good example of a system that is knowledge-based but cannot be considered an expert system. "It is more of a 'smart manual' than an expert system. Its knowledge is basically algorithmic rather than heuristic, and it could easily have been programmed using traditional, not knowledge-based, programming techniques."[27]

CatTutor

CatTutor, developed at the National Agricultural Library (NAL), is another hybrid system, better described as a hypertext training tool than as an expert system. Its

primary goal is to educate novice catalogers in creating bibliographic records for computer files. The system incorporates portions of *AACR2*, second edition, 1988 revision; the MARC format for computer files; a glossary; sample bibliographic records; quizzes; and a mastery test. Specifically,

> CatTutor aims to familiarize the user with the specialized vocabulary of descriptive cataloging and computer files; to link related concepts in standard reference tools to enhance access to these tools, thereby facilitating cataloging; to simulate an actual cataloging environment by providing the users with a graphic representation of the item being cataloged; and to lead the trainee through the creation of cataloging records.[28]

The system was tested and evaluated at NAL and other institutions across the country. Sarah Thomas provides a detailed assessment of the effort and costs associated with the development of this prototype. Knowledge engineers and programmers invested a great deal of time in the project. A minimum of $125,000 was spent on the development and testing of the prototype alone. "To improve the prototype to the point that it could be used as an effective tool in libraries or library schools would require many additional hours. To create a tutorial that would cover the spectrum of cataloging training, including descriptive cataloging of formats other than computer files, subject analysis, classification, and authority work, would require a major commitment."[29]

The developers also concluded that few libraries have the technological environment to support the system as envisioned. It was imagined as a permanent reference tool that would be a component of each cataloger's workstation. In addition, there were persistent questions about whether the tool was intended for experienced or novice catalogers. As with CATALYST, experienced catalogers wanted much of the basic material removed. Novices still thought that the system was not clear enough and that it required more detail.

Developers of expert systems for cataloging have invested great effort in making *AACR2* manageable or reducing it to manageable subsets. (Davies argues for extensive revisions of the cataloging codes to remove rules that rely on human judgment.) Even if the capital costs were too high for most libraries, installing such systems in the national libraries could dramatically improve the volume of material cataloged at those libraries, an effect that would ultimately trickle down, eliminating backlogs at the local level. Davies concludes that systems that serve an advisory role, as a sort of partner with the cataloger, are the simplest and most promising in the short run. However, where expertise is rare (as in the case of cataloging maps, the problem MAPPER addresses) systems which incorporate real expertise and not just the relevant AACR2 rules are the most useful.[30]

Davies draws a further conclusion that all in technical services would be wise to heed: The performance of automatic cataloging should be assessed primarily on the nature of the mistakes made. Do mistakes occur primarily in description, or do they affect information retrieval? And how easy are the mistakes to detect and put right?[31] Hjerppe and Olander speculate that only 20 percent of the cataloging done today poses problems for human or computer catalogers. "In such cases, great demands are placed on the cataloger in terms of wide cataloging experience, high level of education and knowledge in general, and profound knowledge of the cataloging code as well as local cataloging practice and tradition. The element of interpretation is increasingly important to 'fringe' cases, and this kind of expertise is also increasingly difficult to formalize to the extent required for computer manipulation."[32] They also conclude that a system to manage the standard cases of cataloging is within reach today, but would probably not be cost-effective for most libraries. However, such a system would free catalogers to address the problematic 20 percent of the titles and to extend the role of cataloging to include other pieces of the record, such as tables of contents, or to include cataloging for other resources, such as databases found on the Internet.[33]

Acquisitions/Serials Control

In their 1989 survey of artificial intelligence and expert systems in libraries, Hsieh and Hall identified twelve articles addressing applications in technical services. They note that the majority concern cataloging, speculating that this is due to the ease with which *AACR2* rules can be manipulated. They rightly acknowledge that this is not the case in acquisitions, "where there are no set rules to guide the creation of expert systems."[34] Although acquisitions librarians would likely argue that there are *some* set rules, the basic assumption is valid. Since their study, at least two expert systems in acquisitions have been developed and reported in the literature.

Monographic Acquisitions Consultant

Pam Zager Rebarcak's Monographic Acquisitions Consultant was designed to eliminate the discretionary component in monographic vendor selection, replacing it with a more quantitative decision-making model. The system was also developed to support the library's philosophy of using multiple vendors for monographic ordering. Elements of the process addressed include selecting vendors or suppliers, determining what types of orders to send to particular vendors, and determining what special policies or procedures apply.

The expertise of the head of the firm order unit, a Library Assistant IV in the monographic acquisitions section, was captured through interviews and observation. During the knowledge acquisition process, a number of decision factors emerged that were subsequently built into the decision tree for the inference engine:

Is the publisher direct only?
Do we have a blanket order with publisher?
Do we have a standing order with publisher?
Is this a membership?
Is the publisher foreign?
Is the publisher a university press?
Is the publisher on our exchange list?
Is this an approval vendor?
Is the publisher scientific?
Is the publisher known as predominately trade?[35]

In addition to these factors, the knowledge base includes supplier addresses, messages and instructions, publishers, and performance variables derived from the *ALA Guide to Performance Evaluation of Library Materials Vendors*.[36] These variables include service, delivery time, accuracy, discounts, shipping and handling, and additional charges. The vendor's performance on each of these variables was weighted, resulting in a composite score reflecting how well the vendor had performed in the library's experience. In the selection process, the vendor with the highest score who can supply a given type of material is recommended. Once a certain number of orders has been sent to that vendor in a given time period, the vendor with the next highest rating will be selected instead, supporting the library's goal of using multiple vendors.

The system was validated in a small test of twenty orders that had previously been assigned to vendors by an expert. Fifteen of the vendors selected by the system matched the decisions made previously by the expert. The remaining five orders were for publishers that were not part of the system's knowledge base. A conversation with Pam Rebarcak and the study's functional expert, Jerie Schwartz, confirmed many of the problems articulated theoretically in the literature. Based on the cost-benefit analysis done regularly by the library's assistant director for technical services, Dilys Morris, Schwartz was spending 20 percent of her time selecting vendors. Releasing Schwartz from this activity would thus free her to tackle larger problems. Rebarcak also encountered the classic problem of the expert not always being able to articulate her reasons for making a selection or the factors considered. Hardware problems and the constant need to maintain the knowledge base have precluded the system from being used in production. However, the system has replaced a huge manual file of vendor arrangements.[37]

Pennsylvania State University System

The expert system developed at Pennsylvania State University by Lynne Branche Brown determines whether a title requested for order would be received on any of the extensive approval plans maintained by the library. The receipt of books on approval plans is determined by a set of rules called the plan profile, which could be incorporated into an expert system. In addition, staff use a variety of sources of information which could be incorporated as knowledge bases into the system: (1) a list of publishers whose titles are received on approval, (2) a spreadsheet which identifies subject inclusions organized by LC class, (3) a list of contemporary authors whose works are received on approval, (4) a list of geographic coverage of the approval plan program, and (5) a vendor database which includes the vendor's treatment of each title (including nonsubject parameters such as academic level and publication type). Incorporating these sources into an expert system would make this data accessible to a wider audience of acquisitions personnel and subject selectors.

As with the Rebarcak system, the prototype was tested against twenty orders that had already been evaluated by the approval plan staff. Most of the titles were identified correctly by the system. Those that were not were traced to errors in logic in the rule base. For example, if the system reached an early conclusion, it did not test the title against later rules; changing the order of the rules corrected the error. Once again in this system, the need for continuous maintenance was evident. The system must be updated as changes are made in each profile, as, for example, when publishers are added or deleted.[38]

In summary, the work in acquisitions and serials control has focused on monographic vendor selection and approval plan receipts. Another possible application that warrants exploration is the extension of the existing predictive serials check-in systems. "The next step for these systems is for them to 'learn' about publication patterns of individual titles, and adjust claiming cycles, based on actual receipt dates, rather than a mathematically-derived length of time between issues."[39] Much effort is also being expended in some library acquisitions departments to monitor and approve license agreements for many of the new electronic products, such as CD-ROMs. Are there elements of this process that could incorporate the expertise of the legal services department at the institution? In the hypertext arena, libraries have developed information finders for the public. Could these be extended to the technical services area, incorporating, for example, "who to contact for what" documents? Such a system would bear some resemblance to the supermarket systems which

enable the shopper to locate the aisle for a particular product.

Collection Management and Development

With the continual increase in number of publications and reductions in materials funding, it is more important than ever to select the best and most relevant material for the library's patrons. Johnston and Weckert provide two additional arguments for the capture of collection development expertise in expert systems. First, this expertise could be put to use in smaller libraries that could never afford the services of a full-time human expert. Second, larger libraries could use the system as a second opinion to improve consistency in the decision-making process. Collection development is also an appropriate domain because perfect results are not required, nor is it clear what perfect results would be in this area.

Selection Advisor

Selection Advisor, the system developed by Johnston and Weckert, uses six categories of selection criteria (in declining order of importance): subject, intellectual content, potential use, relation to collection, bibliographic considerations, and language. Issues within these categories are grouped into first, second, and third priorities. The system interacts with the user through a series of thirty questions for each book or journal being considered for purchase, for example: Is this a major critical study? Is it likely to be of research interest? Is it a popular treatment? Using the Prolog programming language, the system evaluates responses to these questions and recommends either purchase or rejection of the title.[40]

In practice, answering thirty questions per item would be too time-consuming. Practicing librarians confirmed for the authors that most decisions were made fairly quickly when scanning lists of titles from publishers. The system would provide no time savings for the selectors over the current method. "A more promising approach would seem to be that of tapping the information in electronic databases of publications, including synopses of content. It is here that the power of such a system scanning thousands of entries would become apparent. Even if such a system yielded results no more quickly, each individual result could reflect the sum of considerable assessment."[41]

Monograph Selection Advisor

Monograph Selection Advisor was developed by Steven Sowell at Indiana University. Sowell selected a narrow subject field—classical Latin literature—because its scope was primarily limited to the works of a few dozen writers and secondary works about those writers and their works. Knowledge acquisition came via interviews with bibliographers in this field, which included questions about basic information on the subject, research and teaching needs of faculty and students, selection sources, and budgetary constraints. Based on the interviews, each factor was given a weight to reflect its importance in the selection process. For example, secondary materials in particular languages were more important than works in other languages. A series of questions was developed based on these factors; based on the user's responses to these questions the system would make one of five recommendations: must be bought, should be bought, can be bought, should not be bought, or more information is needed.

The testing/refinement stage of this project was well designed and implemented. Examples were evaluated by the system and the expert. Where there were differences, the expert was interviewed and the system's weighting structure refined as appropriate. This cycle of testing and refinement continued until the developer was satisfied with the program's performance. Since the knowledge base in this area is relatively stable, little revision will be required on an ongoing basis except to reflect changes in the teaching and research interests of faculty and students. A subject area experiencing rapid change will require significantly more updating.

Like Johnston and Weckert, Sowell observed that the expert often took in the information at a glance, arriving at a selection decision much more quickly than the system could. He suggests that the system might be more effective using frames as opposed to the current rule-based scheme. Sowell reaches a conclusion similar to that of the developers of the Selection Advisor: that an expert system could effectively review large quantities of machine-readable bibliographic information on newly published materials and make selection recommendations based on a library's specific knowledge base. Although similar to an approval plan profile, the system could be programmed and refined to exclude unwanted items. In addition, he sees a potential use for the system in the training of subject bibliographers, especially through full use of its ability to explain itself. Finally, he speculates that an expert system might be the most effective way of presenting collection development policies.[42]

Journal Expert Selector

Journal Expert Selector was developed by Roy Rada, editor of *Index Medicus*, to capture the expertise of human journal selectors at the National Library of Medicine who were making decisions as to which journals should be indexed in *Index Medicus*. The main criteria of

the JES included (1) composition of the journal, (2) producers of the journal, (3) information in articles, and (4) authors of articles.[43] The JES contains thirty rules that interact with particular journal attributes. To operate the JES, an expert responds to each journal attribute, indicating the importance they assigned that attribute. Rada's evaluation of the first prototype can be generalized to other expert systems where those with expert knowledge may be threatened by the implications that a expert system might replace them.

> A set of rules was devised and tested, and the expert system performed reasonably well as a first prototype. The experts themselves were, however, generally reluctant to accept the possibility of any formal codification of their knowledge. The view is typical in situations where the experts have long practiced an art without needing to justify in detail the decisions rendered during that practice.[44]

The Bibliographer's Workstation

The Bibliographer's Workstation, developed by John Meador and Lynn Cline at Southwest Missouri State University, again represents the use of a hypertext tool rather than an expert system. The system models the four-step collection decision process: identification of material, evaluation, selection (or rejection), and acquisition. Each stage relies on different sets of data. The data in the Bibliographer's Workstation are organized into four groups: (1) bibliographic data, such as the library's local OPAC; (2) critical and contextual data, such as collection development policies and accreditation standards; (3) financial data, such as the library's materials budget allocations; and (4) commercial data, such as BIP+ or other vendor databases. The user can approach the system by working through the process of selection, by going directly to the needed data set, or by discipline. The current system is essentially a collection of data and databases linked through hypertext connections to databases accessible via the Internet. However, the developers envision the system's evolving into a tool for filtering bibliographic data.[45]

Collection development systems to date have focused primarily on providing information to enhance the selection process. As an acquisitions librarian, I can envision cases where these selection decisions could be captured to enhance the acquiring process. Not only could available bibliographic data be captured to avoid rekeying, but the thinking that resulted in the decision to buy could also be captured. For example, how much effort should be expended in locating a piece that is out of print? What is most important: getting the best price, getting an item quickly for use in a class, or getting a specific or first edition? But collection development and management includes much more than the selection process. Systems could be explored relating to journal cancellations, weeding or storage decisions, evaluation of the collection, and usage studies.

Preservation

Preservation activities are almost entirely unexplored territory in the expert systems area. The only evidence I can find of development in this area is a system known as CALIPR, marketed by the California State Library Foundation. This needs-assessment instrument provides some of the expertise of a preservation consultant "to help assess, quantify, and prioritize the preservation needs of your collection."[46] Collection development and management, along with acquisitions, serials, and preservation, have likewise received little attention in the expert systems arena; there is much new ground to be explored.

Conclusion

Charles Bailey has articulated a number of the general barriers to the development of expert systems, such as cost, the tedium and difficulty of knowledge acquisition, the difficulty of natural language processing, problems in scaling up prototypes to operational systems, and the limited pool of artificial intelligence expertise in the library world. In addition, he discusses an issue he labels as "risk aversion." "When library administrators invest scarce resources in innovative projects, they usually expect success, preferably rapid success. Unfortunately the closer to the cutting edge a project is, the greater the chance that it will fail to produce a fully functional system."[47]

Librarians must also adjust their thinking about expert systems to match current thinking in the artificial intelligence world. According to Philip J. Smith of the Cognitive Sciences Laboratory at Ohio State University, the field has shifted away from the use of the term "expert systems," which are considered a failure, to the term "cooperative systems." Developers of cooperative systems acknowledge that it is very difficult to replicate with a machine what the expert actually goes through. Instead, developers take a task that an expert does, identify a piece of that task that the machine can assist with (typically 20 percent of the task), and develop a cooperative system around that aspect.[48]

Despite the number of articles written about expert systems, development in the library world has been

limited, with almost no operational systems in widespread use. Bailey compares the development of expert systems to the development of home-grown automated systems and those marketed by vendors. When did everyone start buying automated systems? The answer is, once a few entrepreneurial libraries and vendors developed viable ones.

> Today, few libraries develop their own integrated library system; most buy a turnkey system from a vendor. This is the major reason why integrated systems are so prevalent today—each library does not have to build its own system. As long as we are in an era of hand-crafted intelligent systems, libraries will make limited use of these systems. We need turnkey intelligent systems, which can be modified for local use. As in the past, the source of these systems may be mixed, with both vendors and a few exceptional libraries producing systems that vendors can successfully market. To accomplish this goal, vendors and a small number of progressive libraries will need to create powerful, transportable, and marketable intelligent library systems, based on the continuing advances made in the commercial AI [artificial intelligence] marketplace.[49]

With downsizing, "rightsizing," and the emergence of contracting out technical services, particularly cataloging services, to outside vendors and utilities, it is unlikely that libraries will devote much attention to developing expert systems for technical services, or that they will have the technological environment and the technological and technical services expertise for such projects. Instead, it will become incumbent on these new service providers to use developments in artificial intelligence to enhance their provision of these contract services. For example, a vendor may develop and provide an expert system application to the libraries using its services. The system would enhance the provision of information to the contractor to expedite the cataloging or acquisitions activities that they have assumed for the library. Ideally, these systems will be components of the library's integrated library system, so that information can flow seamlessly from one area to another. Thus, partnerships between providers of automated library systems and providers of technical services are essential to efficient use of expert systems in technical services.

It is important not to lose sight of the obvious: that artificial intelligence is nothing more than a tool having unique strengths and weaknesses. "Our true goal is not to create systems based on artificial intelligence technologies—it is to create the most powerful, flexible, and easy-to-use systems possible for ourselves and our patrons. AI is one tool in the toolbox, which should be employed when the characteristics of the task at hand indicate that an AI solution is called for."[50]

References and Notes

1. Frank D'Andraia, "Knowledge Workers in Technical Services," in *Foundations in Library and Information Science*, ed. Thomas W. Leonhardt (Greenwich, Conn.: JAI Pr., 1992), 25: 241–48.
2. Nigel Ford, *Expert Systems and Artificial Intelligence: An Information Manager's Guide* (London: Lib. Assn. Pub., 1991), 8–9.
3. Ibid., 13–14.
4. Bessie M. Carrington, "Expert Systems: Power to the Experts," *Database* 13 (Apr. 1990): 47.
5. David N. Wilson, "Expert Systems: Project Management Implications," in *Libraries and Expert Systems*, ed. Craig McDonald and John Weckert (London: Taylor Graham, 1991), 153.
6. Carrington, "Expert Systems," 47; Charles Fenley, *Expert Systems: Concepts and Applications*, Advances in Library Information Technology, no. 1 (Washington, D.C.: Cataloging Distribution Serv., Library of Congress, 1988), 24–25. Fenley provides an extensive list of essential and highly desirable characteristics for a suitable expert system domain.
7. Ralph Alberico and Mary Micco, *Expert Systems for Reference and Information Retrieval* (Westport, Conn.: Meckler, 1990), 258.
8. Hilary Drenth, Anne Morris, and Gwyneth Tseng, "Expert Systems as Information Intermediaries," in *Annual Review of Information Science and Technology*, ed. Martha E. Williams, 26 (1991): 114–15.
9. Alberico and Micco, *Expert Systems for Reference and Information Retrieval*, 107.
10. R. E. Podhorn, "In Pursuit of Expert-Ease: A Methodology," *Library Software Review* 9 (Jan./Feb. 1990): 8.
11. Ford, *Expert Systems and Artificial Intelligence*. Ford discusses each of these methods in individual chapters.
12. Ibid., 225–26.
13. Ibid., 22.
14. Ibid., 233.
15. Carrington, "Expert Systems," 47.
16. Kenneth Quinn, "Expert System Shells: What To Look For," *Reference Services Review* 18 (Spring 1990): 83–84; Alberico and Micco, *Expert Systems for Reference and Information Retrieval*, 59–63.
17. Roy Davies, "Expert Systems and Cataloguing," in *The Application of Expert Systems in Libraries and Information Centres*, ed. Anne Morris (London: Bowker-Saur, 1991), 133–66; Barbara Anderson, "Expert Systems for Cataloging: Will They Accomplish Tomorrow the Cataloging of Today?" *Cataloging & Classification Quarterly* 11 (1990): 33–48.
18. Ling Hwey Jeng, "An Expert System for Determining Title Proper in Descriptive Cataloging: A Conceptual Model," *Cataloging & Classification Quarterly* 7 (1986): 61; Ling Hwey Jeng, "The Knowledge Representation of the Visual Image of a Title Page," *Journal of the American Society for Information Science* 42 (1991): 99–109.
19. S. J. Cosgrove and J. M. Weimann, "Expert System Technology Applied to Item Classification," *Library Hi Tech* 10 (1992): 34.

20. Diane Vizine-Goetz, Stuart Weibel, and William Oskins, "Automating Descriptive Cataloging," in *Expert Systems in Libraries*, ed. Rao Aluri and Donald E. Riggs (Norwood, N.J.: Ablex, 1990), 123, 125–26, 134.

21. Ford, *Expert Systems and Artificial Intelligence*, 133–34.

22. Harold Borko, "Getting Started in Library Expert Systems Research," *Information Processing and Management* 23 (1987): 87.

23. Zorana Ercegovac, "A Multiple-Observation Approach in Knowledge Acquisition for Expert Systems: A Case Study," *Journal of the American Society for Information Science* 43 (1992): 507, 513.

24. Anderson, "Expert Systems for Cataloging," 41; Jeng, "An Expert System for Determining Title Proper," 61.

25. Karen Markey Drabenstott, Leslie C. Riester, and Bonnie A. Dede, "Shelflisting Using Expert Systems," in *Classification Research for Knowledge Representation and Organization*, ed. Nancy J. Williamson and Michele Hudon (Amsterdam: Elsevier, 1992), 199–207.

26. Forbes Gibb and Carolyn Sharif, "CATALYST: An Expert Assistant for Cataloging," *Program* 22 (January 1988): 62.

27. Ford, *Expert Systems and Artificial Intelligence*, 132.

28. Sarah E. Thomas, "CatTutor: A Prototypical Hypertext Tutorial for Catalogers," *Library Resources & Technical Services* 36 (Oct. 1992): 508.

29. Ibid., 513–14.

30. Davies, "Expert Systems and Cataloguing," 157, 159.

31. Ibid., 160.

32. Roland Hjerppe and Birgitta Olander, "Cataloging and Expert Systems: AACR2 as a Knowledge Base," *Journal of the American Society for Information Science* 40 (Jan. 1989): 28.

33. Ibid., 38.

34. Cynthia C. Hsieh and Wendy Hall, "Survey of Artificial Intelligence and Expert Systems in Library and Information Science Literature," *Information Technology and Libraries* 8 (June 1989): 212–13.

35. Pam Zager and Omar Smadi, "A Knowledge-Based Expert Systems Application in Library Acquisitions: Monographs," *Library Acquisitions: Practice & Theory* 16 (1992): 148.

36. Guide to Performance Evaluation of Library Materials Vendors, Acquisition Guideline no. 5 (Chicago: American Library Assn., 1988).

37. Conversation with Pam Zager Rebarcak and Jerie Schwartz, Iowa State University, 27 October 1992.

38. Lynne C. Branche Brown, "An Expert System for Predicting Approval Plan Receipts," *Library Acquisitions: Practice & Theory* 17 (1993): 156–57, 160–62.

39. Ibid., 155.

40. Mark Johnston and John Weckert, "Expert Assistance for Collection Development," in *Libraries and Expert Systems*, ed. Craig McDonald and John Weckert (London: Taylor Graham, 1991), 70, 72, 75–77.

41. Ibid., 83.

42. Steven L. Sowell, "Expanding Horizons in Collection Development with Expert Systems: Development and Testing of a Demonstration Prototype," *Special Libraries* 80 (Winter 1989): 47–49.

43. Roy Rada and others, "Computerized Guides to Journal Selection," *Information Technology and Libraries* 6 (1987): 180.

44. Ibid., 183.

45. John M. Meador, Jr., and Lynn Cline, "Displaying and Utilizing Selection Tools in a User-Friendly Electronic Environment," *Library Acquisitions: Practice & Theory* 16 (1992): 290–94.

46. Marketing flyer for the CALIPR system from the California State Library Foundation.

47. Charles W. Bailey, Jr., "Intelligent Library Systems: Artificial Intelligence Technology and Library Automation Systems," in *Advances in Library Automation and Networking* (Greenwich, Conn.: Jai Pr., 1991), 4: 5–10, 11–12.

48. Conversation with Philip J. Smith, Cognitive Sciences Laboratory, Ohio State University, 30 July 1992.

49. Bailey, "Intelligent Library Systems," 12, 18–19.

50. Ibid., 13.

To Merge and Not To Merge: Israel's Union List of Monographs in the Context of Merging Algorithms

Susan S. Lazinger

ALEPH, Israel's research library network, was implemented as a highly decentralized network consisting of nearly thirty separate files and no union catalog. To solve the problem of searching these separate files, the Union List of Monographs was implemented in 1991. In reality neither a union list nor confined to monographs, the ULM is, in effect, a union index for locating bibliographic items by author or title, in order to ascertain where they can be found without searching each library's file separately. A survey of the literature of merging files and records is presented, followed by the history and development of the algorithm which produces this unique union list.

ALEPH (Automated Library Expandable Program Hebrew University), Israel's research library network, has always been a system that reflected the needs and limitations of the country it was designed to serve: "Like Israel itself, a country always long on problems and short on financial resources, both the designers of ALEPH and the decision-makers governing the network structure have been influenced every step of the way by the constraints of stark pragmatism. . . . The approach was, necessarily, better to get it done than to get it perfect."[1]

This highly pragmatic approach, which led to a decentralized network consisting of nearly thirty separate files of records for individual university libraries and special projects, has been described in detail in an earlier article on ALEPH.[2] The current article presents the equally pragmatic approach that led to a solution, albeit partial and imperfect, of the problem of searching in these separate files. Like ALEPH itself, the solution is unique: a union catalog that was never intended to replace the individual records it merges, but rather to serve as a supplementary search tool. ALEPH's Union List of Monographs (ULM) is, in fact, a whimsical misnomer, neither a union list nor limited to monographs. What it is, in effect, is a union index for locating bibliographic items by author or title, in order to ascertain where they can be found without searching each library's file separately or choosing libraries by guesswork. Once a particular bibliographic item is located by one of these fields, the ULM retrieves a short cataloging record, with a list of libraries holding the item and pointers to the full cataloging record in each library's file. The user can then retrieve the full record by inputting the pointer associated with a particular library, in the form of that library's code and the system number of the record.

The title of the present paper, therefore, derives from the curious situation that engendered the ULM. In seeking to make the decentralized system, which lacked the union catalog most American bibliographic networks offer, more useful, it was decided to attempt, in the words of Monty Python, "something completely different." An algorithm was sought that would merge matching records into a short record listing the library code and system number of each record merged—cheaply, automatically, and reasonably accurately—without replacing the original records. What emerged from this merging-while-not-merging was a location device that, at least as far as this author was able to determine from the literature, is unique.

The Literature of Merging

The professional literature on merging online records explores four issues: (1) the goals of merging, (2) problems and solutions in merging records from diverse cataloging systems using disparate software, (3) selection and testing of merging algorithms, and (4) problems involved in matching and merging records.

The Goals of Merging

Stephen Toney, in a 1992 article on deduplication of an international bibliographic database, enumerates the three basic options available for merging records determined to be duplicates:

1. One record is chosen as the master record and the others are deleted.
2. All records are kept but clustered with a master record.
3. One record is chosen as the master record and variant fields from the duplicates are added to the master.[3]

Toney stresses that different techniques work for different databases and that the need for testing in designing a merging project cannot be overemphasized. Furthermore, testing must be an iterative process, with

Susan S. Lazinger is a full-time member of the faculty at the School of Library, Archive and Information Studies, Hebrew University of Jerusalem. The author wishes to thank Judith Levi of ALEPH-Yissum and Elhanan Adler of Haifa University for their technical advice.

tests made on samples of records. The method is modified based on the results of each test, and the cycle repeated until an acceptable level of success is achieved. The sensible rule he suggests in testing fields for inclusion in the match keys is to try the simplest approach until it proves not to work.[4]

Dorothy McPherson, Karen Coyle, and Teresa Montgomery—Coyle is the author of several articles on merging policies at the University of California Division of Library Automation (DLA)—describe a design similar to the RLIN clustered database concept, in which a composite database record is built from incoming records. The user, however, views a selected version of the merged record that has been designated as the "base record." The DLA design is depicted as a "composite" system in itself, contrasting with "some systems where only one version of a record is allowed, and with other systems where each contributor's records are kept in separate files."[5] The advantage of the composite bibliographic record is that it "minimizes storage while allowing for the possibility of reconstructing each component record by extracting the fields flagged for a particular cataloging unit," while masking the merged record on the screen.[6] Whether network design maintains a single database record for each distinguishable bibliographic work, as in OCLC and WLN, or keeps multiple copies of records, as in RLIN and DLA, determines the stringency of the record-matching algorithms, as Coyle explains in a later article on record matching:

> Both OCLC and WLN strive to maintain a single database record for each distinguishable bibliographic work. It is also important that this single database record be of a high quality. Both of these systems have record-matching algorithms that are very strict, as one of the matching records will be removed from the database. . . . Both RLIN and DLA keep multiple copies of records. . . . In these systems, the matching algorithm can be freer. Though mismatching can mask the presence of a record, data are not lost from the database and can be recovered when the error is found.[7]

Merging Disparate Formats

Even when all records accepted for inclusion in a union catalog are fully-coded MARC-format monograph records, as in the DLA system, which includes records from OCLC, RLIN, LC, and local UC files, differences in source records require translating transaction records into a standard record format. There are a number of factors that account for differences in source records in different databases. McPherson, Coyle, and Montgomery cite differences in the representation of local data and profile conventions.[8] Dale Flecker, describing the creation of the Harvard Union Catalog, notes that even merging machine-readable records from individual units of the same university creates problems:

> First, there is a problem of cataloging standards and practices . . . made more difficult when the various cataloging units do not have available the same basic reference sources. . . . Most important certainly is that two catalogers doing original cataloging of the same title will frequently create at least slightly variant records.[9]

When nonstandardized databases using widely varied formats and software are combined, the problem of converting all records to a standard format is naturally even more acute and complex. Yannakoudakis et al. describe a procedure for converting records from nonstandard databases by reducing all incoming citations to a short, uniform record comprised solely of candidate elements for the USBC-generation rules, and by adopting different conversion rules for each of the files.[10] Royan calls the millions of National Agency records, plus those of numerous libraries with automated catalogs with substantial databases, each in a different format, which were loaded into a bibliographic network for Singapore, a "veritable Babel of formats."[11] The strategy for dealing with this bewildering variety of formats was "to convert each file into one common exchange format and then to merge it into the SILAS [Singapore Integrated Library Automation Service] database in a single standard way."[12] In the case of Singapore, USMARC was chosen for pragmatic reasons as the exchange format, based on the likelihood that future records would be available in that format.

Developing the Merging Algorithm

Discussions of the merging algorithm in the literature address two issues: (1) the theoretical underpinnings of merging algorithms, i.e., what merging algorithms must accomplish and general principles of merging algorithm development, and (2) the development, testing, and application of specific algorithms for particular union catalogs.

Theory of Merging Algorithms

Toney lists eight topics that must be taken into account in planning any deduplication project:

1. The goal of duplicate checking
2. The question of whether a computer or an editor determines which records are duplicates
3. The processing algorithm
4. The selection of fields to use for duplicate checking
5. The preparation of match keys
6. The determination of what is considered a duplicate

7. The question of what to do with duplicates once identified
8. The need for testing.[13]

Basing his theories on previous literature, Toney goes on to analyze the two fundamental approaches to duplicate checking:

> There are two basic approaches to duplicate checking depending on the project goals. One approach, which I call the loose method, seeks to match records for human review; since the reviewers will be making the final determination, more rather than fewer matches are desired, and thus records with a lower degree of similarity will be matched. The tight method is for projects in which human review is impossible; records must match very closely because they will be matched automatically.[14]

Finally, Toney proposes the criteria for selecting the fields that should be used in a merging, or deduplication, algorithm:

1. Frequency. There is no point in using a field that only a few records contain....
2. Consistency. If the data in the field are inconsistent, they provide a less certain means for matching....
3. Simplicity of programming and processing. Simplicity of programming argued for having all records use the same fields.[15]

Toney notes, however, that while using fields that require little processing aids in the simplicity of merging records, author and title fields, which are not such fields, are nonetheless mandatory for deduplication. Although some of the specific algorithms did not use the author field as a match key, at least for the first pass of a two-stage algorithm, preferring, for example, a title-date key,[16] the title field was cited throughout the literature as the one match key field essential to any algorithm, tight or loose. As Coyle mentions, it is "the chief identifier of the bibliographic work."[17] MacLaury, writing on the development of fixed-length keys for automatic merging of monographic databases, also cites the title field as the primary source field for this type of key:

> In developing a fixed-length key . . . the title field (title and subtitle) was chosen as a primary source for this key because it is generally a good discriminator between records; it is one of the fields least subject to variation in cataloging practice, as the title is recorded as found on the title page and it appears in almost all records.[18]

Coyle and Gallaher-Brown relate to the specific differences in cataloging records for the same bibliographic item that must be overcome by a merging algorithm: (1) differences in the fullness of the record, (2) variations in cataloging practice, and (3) differences due to errors.[19]

Specific Merging Algorithms: Case Studies

Merging algorithms, as mentioned above, can be categorized as either tight (requiring a very close match because no human intervention is involved) or loose (requiring a lower degree of similarity for matching and leaving the final decision on whether to merge borderline matches to the staff). An example of the former, tight method is the University of Illinois algorithm, which because of strict matching conditions, makes very few incorrect matches.[20] The University of California, while using a weighted scheme (revised from its earlier algorithm) that "improves the chances that records will match without increasing the number of incorrect matches appreciably" can nevertheless be classified among the loose algorithms, since human intervention is involved: "Pool records having weights that are near the threshold are designed for review by DLA staff."[21]

Aside from these differences, and in spite of the addition or subtraction of certain nonprimary fields in constructing the match keys, most of the algorithms had certain features in common. First, data in the textual fields of the records was usually normalized—i.e., punctuation, spacing, special characteristics, and diacritics were removed and all letters were converted to uppercase.[22] This allowed the computer to ignore minor differences and determine that the fields matched. Next, most algorithms used a two-phase matching process.

In the "cleanup and deduplication" algorithm described by Toney, the first pass inspected the entire database and grouped the records into "pools." The second pass examined each pool with a more precise algorithm to find duplicates. The advantage to this two-pass method is that it saves the computer's having to examine each record minutely against all others.[23] The University of California's algorithm, as described by Coyle, first matched the records on a minimum set of highly reliable data elements—LCCN, date, edition, and title—tolerating very little variation, in order to bring together records derived from a standard LC cataloging in a minimum of processing time. A second phase processed records that didn't match on the minimum set of data elements through the remainder of the algorithm, comparing the cumulated weight at key points to defined upper and lower cutoff points.[24] The University of Illinois algorithm consisted of a first step that generated a key to partition the database into many small sets of possible duplicates based on the last two digits of the date and a sampling of eight characters from the first three and last three words in the title, i.e., a "title-date" key. The second step compared names, titles, and pagination of records. Name matching used the first five

characters in both main and added entries, while title matching used "Harrison keys" (bit strings derived from character strings) hashed into the keys. These title keys were then compared using the Hamming distance—the number of bits in one key that did not match those in a second key—as the criterion for matching.[25] Hickey and Rypka, who describe this complex algorithm with its very tight matching criteria, state that "the . . . program committed very few errors in matching different records but missed many duplicates."[26] Thus, in one sentence they pinpoint the major problem which must be faced in choosing a merging algorithm—the balance between an algorithm so strict it misses duplicates and an algorithm so loose it merges nonduplicates.

Problems in Matching and Merging Methods

There are two basic problems to deal with in any attempt to merge online bibliographic records: (1) the need to develop an algorithm neither so tight that genuine duplicates fail to merge, clogging the system, nor so loose that records that are not genuine duplicates merge, hiding or destroying bibliographic information, and (2) the need to decide whether all merging will be done automatically or whether, after running the records through the computer algorithm, the final decision will be left to human editors.

With regard to the first problem, agreement is unanimous throughout the literature that merging nonduplicates is a far more serious error than failing to merge duplicates. Missed duplicates do indeed cause a problem, as Hickey and Rypka note:

> Large numbers of duplicate records impede searching and increase the number of records that users must inspect and compare in detail. Manual comparison of suspected duplicates is difficult at best without two terminals side by side or without a hard-copy printout of the records. Duplication also increases the size of the data base that must be stored.[27]

Nonetheless, all agree that mismatches create a far worse problem:

> The adequacy of any scheme for creating a union catalog can be measured in terms of the two cardinal errors that may occur. Records that do not relate to the same item might be erroneously identified as duplicates—these are mismatches—and records that truly relate to the same item might not be identified as duplicates—these are missed matches. Of the two types of errors, mismatches are more serious because they result in information being permanently lost.[28]
>
> Over-identification is less serious than under-identification. False drops only tend to clog the line. Lost reference points, on the other hand, lead to lost information.[29]
>
> In addition to our desire to merge "like" records, an over-riding goal was to not merge records which did not represent the same item. The penalty for mis-merging of records is that one of the records is essentially hidden from the user, as only one version of a merged record is displayed.[30]

On the second issue—whether a computer algorithm can be developed that is adequate for matching and merging records without human intervention, or whether human editors must be used for the final step in any merging project—there is far less agreement. Toney believes that only human beings are able to perceive the nuances necessary to remove all duplicates without making mismatches:

> It is impossible to design a deduplication algorithm that will remove all the duplicates without removing some legitimate records. Since human beings are far better at perceiving patterns and nuances than computers, and computers are better at comparing large numbers of records to suggest possible duplicates, a strategy should be used that builds on these different abilities.[31]

Coyle holds that the reduced information in the online record, as opposed to the book in hand, coupled with the computer's lesser ability to infer, makes fully computerized record matching problematic:

> A cataloger can look at two books and decide if they represent the same edition of the same bibliographic work. The same cataloger, book in hand, may find it less easy to decide which of a number of plausible online records is the right one for the work. An algorithm comparing two machine-readable records is clearly working with a reduced set of clues with which to judge "sameness" and has much less inferential ability when faced with ambiguous or conflicting data.[32]

In two other articles on merging, however, Coyle, writing with others, emphasizes that even manual editors can never produce an algorithm that is 100 percent effective:

> Design of a merging algorithm that is 100 percent effective is probably not feasible given the almost unlimited possibilities for variations in the representation of bibliographic data. Even manual review of two cataloging records may not be sufficient to resolve the question.[33]
>
> We knew from our experience with bibliographic databases that no record matching algorithm could be perfect. Any large database has record pairs that even a human expert cannot make a positive decision on.[34]

Finally, Hickey and Rypka claim, not only that human intervention is unnecessary, but that testing indicated that human matching can actually be inferior to machine matching:

> In testing the algorithm, we found that human matching of records is inconsistent at best, even when librarians can agree on the definition of a duplicate. For example,

the algorithm refused to match three records of the duplicate pairs sample that two professional catalogers had identified as duplicates. Both catalogers missed the difference between the records—the fixed reproduction code. For such detailed comparison the algorithm can be more accurate than all but the closest human scrutiny.[35]

ALEPH's Union List of Monographs

In attempting to provide a search tool that would eliminate the need to jump from file to file to ascertain where in Israel's university system a specific book could be found, ALEPH's designers faced, initially, the problems universal in merging projects. First, bibliographic records for the system are produced by nearly thirty different cataloging units. In spite of the fact that all of these units use the same software—an advantage over many of the systems described above—there remain variations in cataloging practices among the contributing libraries and slight variations in field codes, since ALEPH allows total flexibility of record design. The Library of Congress call number, for example, is coded "LC" in some files and "CN" in others. Thus, the first requirement for merging records in ALEPH was to select the fields whose codes would be normalized throughout the system in order to run them through the merging algorithm. Four fields were selected: the main entry author field, title field, year of publication field, and language of publication field. Each library, therefore, has a table that automatically translates whatever codes it uses for these fields to AU = author, TL = title, YR = year, and LN = language—the codes used by most libraries in the network—when it prepares its files for input into the ULM.

Second, at the time of this writing the total number of records in the system stood at approximately 1.5 million, with thousands of records being added monthly. The size of the database and the lack of money to undertake a merging project that would require a staff of human editors made it imperative to develop a "tight algorithm"—one that would function adequately without human intervention for bibliographic decisions. What remained to be determined, in addition to the algorithm itself, was the working definition of "adequately."

On the other hand, the designers of ALEPH's merging algorithm had one less problem than the designers of the algorithms discussed above. Because the union list was to stand separate from and in addition to the individual files, no records were ever in danger of being discarded or hidden by a mismatch. The worst consequence of a mismatch would be that users searching in the union list would fail to locate a record in *this* file. The original record would remain in the individual file, unmerged and undamaged, and could be retrieved in a search of the individual library's online catalog.

A few years after the 1984 decision to decentralize Israel's research library network, allowing each library to maintain a separate file and make its own authority decisions, the system's designers embarked on their first attempt to provide a union list to supplement searching in the individual catalogs. At the time this author's 1991 article on ALEPH was written, the problems intrinsic in such a decentralized network had brought the project to a temporary halt:

> Because of the decentralized authority control in ALEPH, a program matching identical main entry and title fields (essentially a clustering program, similar to RLIN's) produced so many duplicate entries that the project was temporarily abandoned. The union catalog simply grew to unwieldy proportions too quickly.[36]

The problem, thus, was missed matches: the matching program was too tight for the rather chaotic state of ALEPH's authority control. The program failed to match so many records that represented the same bibliographic item but showed minor differences that the database became cluttered and ineffective.

As with all components of the ALEPH network, the ULM project was both hampered and facilitated by the pragmatic Israeli approach of "better to get it done than to get it perfect." It was this philosophy that enabled this resource-poor country to implement a nationwide research library network within a few short years by allowing autonomy of authority control to each library. The goal of the ULM project was always to provide a "quick and dirty" file, searchable by author and title only. Because no records would actually be lost even if mismatches occurred, since the individual files would remain untouched and accessible, it was decided from the first to make the matching algorithm rather loose (as tight algorithms go!), matching fewer fields than the algorithms used in most other projects, even though there was to be no human intervention. Once again, the objective was to get it done and functioning, even if imperfectly, as quickly as possible and without costly human editing. The first attempt at producing the ULM proved that ALEPH's authority control was so loose that even a loose algorithm—requiring a match only on author, title, and filing year—was too tight to prevent an unacceptable number of missed matches.

In mid-1991 a second algorithm was completed and tested on the network, and by late 1991 the ULM was implemented. The minutes of a meeting of Israel's Inter-University Cataloging Committee on January 20,

1992, describe the nature of the algorithm and the file it produces.[37]

The unification algorithm is based on a comparison of the nonfrequent letters in the title (seven characters) and author's name (three characters), in combination with the publication year. This algorithm has been found to function well enough to prevent the database's being clogged with duplicates (i.e., missed matches), at least with regard to monographs. Since the algorithm does not use the "MT," or material type, field in its matching, however, the ULM it produces includes cataloging material of all types, including journals and audiovisual material. Therefore, considerable overlap is created between it and the ULS (Union List of Serials, one of the individual files in the network), although the ULS is bibliographically controlled and includes nonuniversity libraries that are not part of the ALEPH network. When the catalogs are unified, the first cataloging record in the database remains the only cataloging record, with no attempt made to ascertain which of the cataloging records is the fullest, most precise, or most authoritative. There is also no means of correcting incorrect cataloging. A corrected entry in an update tape opens a new record (or combines with a previous record, possibly with the incorrect record). From time to time, therefore, it will be necessary to erase the entire ULM and rebuild it from the beginning; only thus can mistakes be corrected.[38]

During the discussion at this meeting a suggestion was raised for improving the algorithm with regard to journals. It was suggested that the automatic program for comparing records be modified so that when comparing titles it would ignore (1) the definitions "JOURNAL," "J," etc. when found in any form of parentheses, (2) the word "VOL." or "VOLUME" and what follows it, and (3) the letters "THE" at the beginning of a title. This improvement would require identification of the series by means of the field MT (material type) and identical content in this field for all libraries. A table of standardized input codes for this field was presented along with the minutes of the meeting, but as of this writing network-wide standardization of this field has not occurred and the algorithm has not been changed to include the MT field in its matching procedure. Another possibility suggested but not yet implemented was matching periodicals by means of the ISSN or the ULS number.

Nonetheless, the ULM's rather minimal matching algorithm, based on only four fields, has alleviated the problem of missed matches by sophisticating the whole-field approach of the first algorithm and matching only a standard number of frequent letters from author and title fields, plus publishing year and language. All letters are normalized to lowercase. During the summer of 1992, the ULM was rebuilt, implementing an additional rule for normalization, which further decreased the number of missed matches. There was, until this new rule was applied, sometimes a problem in matching the author fields of identical items, because some of the libraries cataloged at a somewhat substandard level—again choosing to "get it online rather than to get it perfect"—identifying authors by last name and first initial only. The new normalization procedure utilized only the word before the first comma in the author field and the first letter after this comma in merging. If the author field contained no comma, the algorithm made use of the text of the entire author field. With this improvement, it was felt that the algorithm was tight enough to prevent a significant number of mismatches and loose enough to prevent so many duplicates that the file would become useless. Quick and dirty it is, but it functions and points most users in most cases to a library in which the bibliographic item they are trying to locate can be found.

Another problem—updating the ULM—is about to be solved. Until the autumn of 1992 the ULM could be updated only by rebuilding the entire file. The basic file, built in the summer of 1991, was rebuilt and updated in the summer of 1992, at the time the second normalization procedure was implemented. This involved closing down the file for a period of about ten days. In autumn 1992 a new updating procedure was implemented, under which each library will be asked to send quarterly updates of its records to the ULM utilizing a batch utility developed for this purpose, which can be accessed through its regular batch file options. The updates will then be copied through DECNET OR TCP/IP to the ULM node, where another utility will update the ULM file. Libraries will be assigned staggered dates for submitting their updates, so that central updating will constantly be in process and the file will never be more than three months behind with regard to the holdings of any given library.

A search in the ULM by author or title, assuming it produces a hit, retrieves a short cataloging record, consisting of author, title, filing year, and language fields, followed by a list of holdings fields culled from the original records merged into this composite record. Each holdings field is made up of the library code (three characters, or three characters plus a period plus three more characters) and the system number of the original, full record in the holding library's online catalog. For example, a search for Abraham Cohen's book *Everyman's Talmud* in the ULM produces the record shown in figure 1.

A request for help is answered with the explanation that inputting the library code plus the system number (e.g., "BGU.BGU/1001490") will produce that library's full record for the item. Another menu available

```
SYSNO     0708452
Title     EVERYMAN'S TALMUD
Author    COHEN, ABRAHAM, 1887-1957
          COHEN, BOAZ
File-YR   1949
Lang      ENG
Holding   JMS JMS 1025993
BGU       BGU 1001490
TEC       TEC 2062153

  Type HELP or press HELP key for
            instructions.
```

Figure 1

in the Help mode defines all the library codes (e.g., BGU is Ben-Gurion University of the Negev in Beersheva, and TEC is the Technion in Haifa). Thus, a user who knows the title or author's name for any item can determine with a single search which libraries in the university library system have the book, and with one additional command can examine the chosen library's complete record. Although primitive, perhaps, by WLN standards, Israel's ULM has proven to be a reasonably effective, relatively inexpensive way to provide a union file for searching a highly decentralized network. Like the network itself, it was designed and implemented on a shoestring and, in typical Israeli fashion, improvised and improved along the way, using limited available resources to the utmost to produce a tool that, though not perfect, works. For Israel, as for other small countries, pragmatism is the default method of technological development and functionality the standard by which a system is judged. By this standard, the ULM, with all its anomalies and limitations, is a success.

References

1. Susan S. Lazinger, "ALEPH: Israel's Research Library Network: Background, Evolution, and Implications for Networking in a Small Country," *Information Technology and Libraries* 10 (December 1991): 289.
2. Ibid., 275–91.
3. Stephen R. Toney, "Cleanup and Deduplication of an International Bibliographic Database," *Information Technology and Libraries* 11 (March 1992): 23.
4. Ibid.
5. Dorothy S. McPherson, Karen E. Coyle, and Teresa L. Montgomery, "Building a Merged Bibliographic Database: The University of California Experience," *Information Technology and Libraries* 1 (December 1982): 375.
6. Ibid.
7. Karen Coyle, "Record Matching: A Discussion," *Information Technology and Libraries* 4 (March 1985): 58.
8. McPherson, Coyle, and Montgomery, "Building a Merged Bibliographic Database," 371.
9. Dale Flecker, "The Harvard Union Catalog: Creation and Maintenance of a Catalog Using Multiple Cataloging Systems," in *Sixth International Online Information Meeting*, London, 7–9 December 1982 (Oxford: Learned Information, 1982), 179.
10. E. J. Yannakoudakis and others, "Matching of Citations Between Non-Standardized Databases," *Journal of the American Society for Information Science* 41, no. 4 (1990): 601.
11. Bruce Royan, "Building a Bibliographic Network for Singapore," *Electronic Library* 5 (December 1987): 348.
12. Ibid., 349.
13. Toney, "Cleanup and Deduplication," 21.
14. Ibid., 23.
15. Ibid., 25.
16. Thomas B. Hickey and David J. Rypka, "Automatic Detection of Duplicate Monograph Records," *Journal of Library Automation* 12 (June 1979): 131; Martha E. Williams and Keith D. MacLaury, "Automatic Merging of Monographic Data Bases—Identification of Duplicate Records in Multiple Files: The IUCS Scheme," *Journal of Library Automation* 12 (June 1979): 159–60.
17. Coyle, "Record Matching," 58–59.
18. Keith D. MacLaury, "Automatic Merging of Fixed-Length Keys Derived from Title-Strings," *Journal of Library Automation* 12 (June 1979): 144.
19. Karen Coyle and Linda Gallaher-Brown, "Record-Matching: An Expert Algorithm," *ASIS '85: Proceedings of the 48th ASIS Annual Meeting* (White Plains, N.Y.: Knowledge Industry, 1985), 22: 77.
20. Hickey and Rypka, "Automatic Detection," 130.
21. McPherson, Coyle, and Montgomery, "Building a Merged Bibliographic Database," 377
22. Coyle and Gallaher-Brown, "Record-Matcning," 78; McPherson, Coyle, and Montgomery, "Building a Merged Bibliographic Database," 378; Toney, "Cleanup and Deduplication," 22.
23. Toney, "Cleanup and Deduplication," 21.
24. Coyle and Gallaher-Brown, "Record-Matching," 78.
25. Hickey and Rypka, "Automatic Detection," 131.
26. Ibid., 130.
27. Ibid., 139.
28. Williams and MacLaury, "Automatic Merging of Monographic Data Bases," 166.
29. James L. Dolby, "An Algorithm for Variable-Length Proper Name Compression," *Journal of Library Automation* 3 (December 1970): 260.
30. Coyle and Gallaher-Brown, "Record-Matching," 77.
31. Toney, "Cleanup and Deduplication," 21.
32. Coyle, "Record Matching," 57.
33. McPherson, Coyle, and Montgomery, "Building a Merged Bibliographic Database," 377.
34. Coyle and Gallaher-Brown, "Record-Matching," 77.
35. Hickey and Rypka, "Automatic Detection," 141.
36. Lazinger, "ALEPH," 282.
37. "The Union List of Monographs," translation from Hebrew of "Summary of Meeting No. 30, which Took Place on January 20, 1992."
38. Ibid.

Book Reviews

Susan B. Harrison, Editor

All about Internet FTP: Learning and Teaching to Transfer Files on the Internet

By David F. W. Robison. Internet Workshop Series, no. 2. Berkeley, Calif.: Library Solutions, 1994. 85p. paper (includes diskettes), $45 (ISBN 1-882208-06-4); paper (without diskettes), $30 (ISBN 1-882208-04-8).

One of the most frequently asked questions about the Internet is how to transfer files successfully using FTP (file transfer protocol). Robison gives a clear introduction to the process in this companion workbook to *Crossing the Internet Threshold* (Library Solutions Press, 1993). Both trainers and users will find much assistance in this volume. Chapters include an explanation of FTP and how it works, a definition of anonymous FTP, what to expect during an FTP session, and the process of downloading. Particularly useful is the chapter on locating files, which explains in sufficient detail how to browse an archive site. Various Archie sites are mentioned, with a road map for wending one's way through the resulting directories and subdirectories. The author also describes how a user can keep up with the increasing explosion of FTP sites and software by subscribing to electronic discussion groups such as the Net-Resources list.

The lecture chapters include notes to the instructor and a numbered icon of a slide projector to indicate which overhead goes with each section of text. The twenty-eight slides were designed with Powerpoint presentation software and are very professional in appearance. Also included are suggested exercises, such as how to FTP a Mac font, locating and getting graphics files, and how to read a test file without transferring it. The author includes a lesson plan with the suggested time needed to complete each module, as well as practical training tips. As with other volumes in this series, the spiral binding is designed for personal use and will not stand up to heavy circulation. Highly recommended.—*Linda Friend, Penn State University Libraries*

CD-ROM In Libraries: Management Issues

Ed. T. A. Hanson and J. M. Day. London: Bowker-Saur. 1994. 296p. $60 (ISBN 1-85739-086-5).

Pioneer users of CD-ROM technology in libraries present unique management perspectives in this collection of overviews and case studies. The chapters were commissioned from authors representing diverse library environments. The overviews present general theoretical treatments of the topic and identify issues, trends, and developments. They provide the reader a perspective through which to evaluate each case study. The book covers special, public, and academic libraries, and has a definite United Kingdom orientation. The editors also suggest that the intense use of CD-ROM technology in the research community results in an academic emphasis in the work.

The editors and contributors have library practitioner as well as scholarly publishing expertise. The presentation of the information reflects their hands-on knowledge. The structure and editing are excellent. Chapters that discuss strategic management issues are followed by one or more representative case studies. Significant planning and implementation issues, including material selection, setup, and marketing, are well described. Chapters on the use of CD-ROM for localized purposes and full-text information storage are unexpected features. In addition to chapter citations, there is an extensive bibliography of recent (1988–1993) publications. The bibliography is well organized and international in scope.

A management focus is the contributors' attempt to avoid the "out-of-date-before-it's-published" problem with books about technology. Nonetheless, the reader often wishes for more recent information. Surveys of CD-ROM products and discussions of pricing and hardware issues exemplify the difficulties inherent in compiling a monograph about technology. In this rapidly changing electronic environment, books may not be the most effective format for even broad-scope technology management issues. This work offers more value for its documentation of CD-ROM pioneering in libraries than for assistance with strategic planning and decision making.—*Connie V. Dowell, Charles E. Shain Library, Connecticut College*

Document Imaging: A Management Guide

By Bob Wiggins. Westport, Conn.: Meckler. 1994. 185p. paper, $45 (ISBN 0-88736-869-7).

Electronic imaging is the latest hot technology touted as helping us arrive at the nirvana of the "paperless office" as well as allowing one to view photograph or other collections with greater ease and image clarity than with micrographics. This book is intended to be a comprehensive overview of this growing field. For the purposes of this book, *document imaging* is defined as including written records, like correspondence and reports, and also graphics, like photographs, illustrations, and charts, that may be part of other written documents. A more recent development is multimedia systems, which add video and sound. The target audience is anyone in an organization who might be affected by implementation of an imaging

system—business managers, systems designers, librarians, and records managers—and who may not have a grasp of computer technology.

The book is divided into four sections, which can be read independently of one another as needed. Part 1 discusses the importance of information in its various forms to an organization and how to approach the problem of managing the "information life cycle." It explains how to determine objectives in retaining records and how to identify what is necessary and what can be destroyed. As a librarian considering image scanning, I found this section the most useful, especially the list of questions to ask to determine how information is used in your organization and the physical characteristics of this information: dimensions, mono or color printing, halftones, etc. When one is considering a new technology, it can be difficult to know the right questions to ask to ensure one gets what is needed from a system.

Part 2 of the book is an overview of relevant computer technology and trends, explaining types of networks, transmission modes, software, and storage systems. Because the technology changes so quickly, this takes the form of a general discussion rather than review of specific products. Part 3 compares older micrographic systems like microfilming with electronic imaging, then proceeds to a fuller discussion of imaging features, including storage media, formats, and image quality and permanence. Part 4 discusses how to initiate and follow through on an imaging project and possible future developments. Throughout the book the author emphasizes the importance of good indexing or classification of the imaged materials, since without clearly understood guidelines for indexing, digitized images can be lost forever (unlike badly organized file drawers, where there is at least the possibility of finding an item).

My main criticism of the book is that technical terms are not always explained on their first introduction; the reader may be directed to another section of the book for a definition. Also, the glossary provided is very selective. Other terms are listed in the index and the reader is directed to the appropriate page, which means one must look these terms up twice. Otherwise, the book is a good starting place for anyone considering an imaging project. It will guide the reader through the important points to consider before startup to ensure a good final result.—*Diane Powers, the New York Public Library, MM Picture Collection*

Internet: Getting Started

Ed. April Marine and others. (SRI Internet Information Series). Englewood Cliffs, N.J.: Prentice-Hall. 1994. 390p. $28 (ISBN 0-13-289596-X).

Bookstore shelves are liberally stocked—the word "glutted" comes to mind—with works promising an (almost) painless transformation of the gentle reader into a cybernaut (or cybrarian). Such knowledge does not come cheaply: $30 is the starting point for a paperbound volume; more if a diskette is included.

Internet: Getting Started is an attractive offering from this genre. Ephemeral enough not to require acid-free paper, the book is thoughtfully laid out using a large, legible typeface, generous interline spacing, and effective use of blue type for section numbers and heading keywords. The table of contents is exceptionally well organized, giving a very clear structural overview via detailed section and subsection numbering and headings. The publisher would do well to reproduce it as a free-standing guide or bookmark: the reader could use this to stay oriented while browsing the book's immensely rich contents. Each chapter begins with a paragraph that presents what will be discussed and how it connects to material that precedes that chapter and immediately follows it.

The authors wisely provide an introductory chapter. Titled *Overview: What is this book about? Who should read this book?* the chapter contains two major topical divisions: (1) What you need to do to gain Internet access (chapters 1–6) and (2) the Internet: its "concepts," applications, organizations, and resources (chapters 7–12). There are nine appendixes and a comprehensive index. Because Stanford Research Institute (SRI) has been heavily involved in networking for three decades, it is no surprise that this book contains a wealth of information dealing with connectivity, much of it international in scope. Contact lists of all sorts abound; they are combined with succinct explanations of key concepts involved in using the Internet. The authors employ an interesting technique to explain new concepts and terminology: they point the reader to actual information sources out on the Net, and they "interview" experts on different applications and tools (e.g., Farhad Xerxes Anklesaria at the University of Minnesota on Gopher). Relatively complex topics are usually presented with great clarity and succinctness: e.g., subsections 4 and 5 of chapter 8 deal with Internet addressing and the domain name system—topics that lend themselves to obfuscation but are here set forth so that almost anyone can grasp them without undue difficulty.

There are some notable omissions in the first part of this book: while PPP (point-to-point protocol) is mentioned, the reader never learns where to obtain such software. For readers who want to use dial-in access, there is a surprising lack of discussion of DOS, Windows,

or Mac telecommunications programs, how to set them up, and how to use them for Internet access.

The second half of the book maintains the momentum of the first, but it suffers from a tendency to treat important topics too briefly. FTP (file transfer protocol) is covered in less than two pages, and while an example of an "anonymous FTP" session is reprinted, the all-important "why" behind the various commands is omitted. Setting the "FTP type" (binary vs. ASCII) is not even mentioned, nor is the "hash on" command, and "mget" does not appear. Gopher is given short shrift indeed—the discussion is too short to be really useful, with nothing on basic Gophering command keys. In fact, the space allotted to Gopher is less than that given many Internet organizations! A discussion of what can be retrieved under Gopher versus anonymous FTP would have been useful. There is no mention whatsoever of Mosaic, and WWW (World-Wide Web) is covered in less than half a page. Section 9.5.4 on Wide-Area Information Server (WAIS) is equally superficial, although again a contact person is given.

One problem with a work composed of many lists is the speed at which such lists become obsolete. SRI might want to contemplate setting up a Gopher to maintain access to continuously updated versions of the listings in this book. Similarly, the question of user interfaces deserves far more coverage: Gopher and WWW as complementary tools and the many GUI (graphical user interface) versions available for UNIX, Mac, and Windows clients.

In summary, this is a rich reference source, nicely planned and organized, with special strength in the area of how to select and obtain connectivity. But the reader who is literally "getting started" will need more in-depth coverage of what to do once connection has been established than is available here. This can be found in such works as Ed Krol's magisterial *The Whole Internet User's Guide & Catalog* (O'Reilly & Associates, now in its second edition), Paul Gistler's *The Internet Navigator* (Wiley) and the delightfully informative, *multum-in-parvo The Instant Internet Guide*, by Brent Heslop and David Angell (Addison-Wesley), to name just a few.—*Gerald M. Furi, Farmington Community Library, Farmington Hills, Michigan*

The Internet Guide for New Users

By Daniel P. Dern. New York: McGraw-Hill, 1994. 570p. $40 (ISBN 0-07-016510-6); paper, $27.95 (ISBN 0-07-016511-4).

With the publication of more and more books aimed at the new Internet user, each author strives to make his or her contribution unique. The strength of Dern's work is in presenting Internet access, facilities, and culture within the framework and history of networking and its creators. Although ostensibly designed to be useful even to a complete novice, the book is so comprehensive that a beginner may be overwhelmed by the detail, but experienced users will enjoy the treatment of the Internet community, "netiquette," and security, along with stories of notable developers and events in the Net's evolution.

The Internet Guide for New Users covers all the specifics of Internet access and involvement for every type of computer equipment and platform, particularly DOS, Windows, Apple, and UNIX, and with any connection that can attach to a modem or a network. It evaluates the pros and cons of various types of access, describes how to obtain names of current Internet providers (but doesn't actually include a list), and clarifies the advantages of different types of accounts (e.g., TCP/IP, SLIP, (Serial Line Internet Protocol) shell, etc.). Continuing into electronic mail, Usenet and navigational tools such as Gopher, FTP (file transfer protocol) and WAIS (wide-area information server), the author often gives instructions for various versions, rather than using just one as a representative sample. In every instance, background comments, examples, and interviews with some of the "giants of the Internet" provide information in context to smooth the transition into cyberspace.

Because the book is designed to be read as needed, by chapter or section, much of the information is repeated or reworded from section to section, with minor but sometimes significant variations. The index, however, is skimpy, making it difficult to retrieve these nuggets of information. For instance, a useful chart on types of files and an accompanying paragraph on how to decompress files is not cited in the index under *files, compression,* or *decompression.* Dern includes a bibliography, glossary, and other appendixes. Since the glossary is intended to define terms and acronyms not covered in the book, it is less useful as a quick reference than as a supplement to the book.

The Internet Guide for New Users is similar to Ed Krol's *Whole Internet User's Guide and Catalog,* but Dern covers more options and possibilities at every step of the searching process. However, if your first goal is to get up and running as quickly and painlessly as possible, then a more straightforward approach may serve you better. The book is chatty (or wordy, depending on your point of view), and there is no systematic organization of information that allows you to jump right into the "how-to" sections of each chapter. On first glance the graphics are attractive, but finding the practical information needed when you're in the middle of a transmission can be frustrating. Levine and Baroudi's *In-*

ternet for Dummies (IDG Books, 1993) is a better choice as a "how-to" guide.

Written with a sense of humor, Dern's book contains detailed explanations with a minimal use of jargon and often includes a comment about the way things "used to be" ten or twenty years ago. Anyone who wants to immerse him or herself in the culture of networking will find *The Internet Guide for New Users* both an extensive introduction and a continuing resource as knowledge and understanding of the Net grow. This book will not replace *The Whole Internet User's Guide and Catalog* but complements it well for those who want an in-depth familiarity with the Internet, especially if you plan to teach it to librarians or other intensive users.—*Dimi Berkner, Developmental Studies Center, Oakland, California*

Internet's World on Internet 94: An International Guide to Electronic Journals, Newsletters, Texts, Discussion Lists, and Other Resources on the Internet

Ed. Daniel P. Dern. Westport, Conn: Meckler, 1994. 451p. paper, $45 (ISBN 0-88736-929-4).

Meckler's *World On Internet 94* is the first in a projected annual series attempting to provide an overview of useful information sources and services available on the Internet. This first edition contains more than six thousand entries describing discussion groups, campuswide information systems, freenets, electronic text archives, WAIS (wide-area information server) databases, and commercial services, among other things.

The entries were compiled by the editors between May and November of 1993, and the process appears to be largely self-selection: an e-mail questionnaire was sent to more than 3,500 sites on the Net, and entries have been composed largely by the respondents themselves. Despite the potential problem of useful sources not appearing because of survey nonreturns, a cursory glance through the listings indicates that the usual suspects (i.e., those sources already familiar to a large number of librarians) do indeed appear. Editor Tony Abbott notes that despite some trimming and editorial polishing, the data are published as submitted. Whether the data were tested for accuracy is not mentioned.

The range of coverage is impressively broad. Entries are clustered in one of eight major categories: discussion groups, electronic journals and newsletters, text archives and FTP (file transfer protocol) sites, freenets, campuswide information systems, commercial services, Usenet groups, and WAIS databases. Access beyond these categories is provided by a subject index. Unfortunately, the index only covers five of the eight categories, and in one striking omission, Usenet news groups, about as subject-specific a category as one could hope for, are not included. The subject index covers forty pages, averaging between ten and twenty subject categories and about fifty individual entries per page; this serves to confirm that subject access, while useful, is limited in the present directory.

Coverage within categories varies in depth. There are more than 230 pages of (primarily listserv) discussion groups, but a mere fifty-five entries on five pages exist for campuswide information systems in the United States (at least one of which, Yale, seems to have changed its address). Of course, the editors deserve credit for even attempting to classify these resources in categories, but the ever-shifting Internet makes such an effort difficult at best: the campuswide information system concept has, because of the ubiquitous nature of Gopher software, broadened far beyond its original concept. In the book's most striking omission, the rapidly evolving World-Wide Web and its many clients such as Mosaic and Cello do not appear to be addressed, despite the broad range of Web servers now in operation.

Ultimately, of course, directories of this nature are bound to fail: they provide us with snapshots of a whirlwind. Some test of the decay rate of this data would be useful; my guess (based only, I admit, on spot checks) is better than 50 percent annually. In fairness to the editors it should be noted that this is largely due to the lack of adoption of some resource identification system on the Internet (such a system of universal resource identifiers, or URIs, has been proposed by the Internet Engineering Task Force). Files are moved, machines renamed, passwords mutate, and another source is lost. Until this situation is resolved, print directories such as this will serve a limited function: the equivalent of the over-worked reference librarian who, in response to a patron's request for information, gestures vaguely toward the stacks and says, "Try over there."—*Patrick Flannery, Texas Medical Center Library, Houston*

Introducing the Internet: A Trainer's Workshop

By Lee David Jaffe. Internet Workshop Series, no. 1. Berkeley, Calif.: Library Solutions Press. 1994. 90p. paper (includes diskettes), $45 (ISBN 1-882208-05-6); paper (without diskettes), $30 (ISBN 1-882208-03-X).

This volume includes a fully conceived lecture/lesson plan on the Internet and has everything a trainer

would need to produce a high-quality presentation. Topics include definitions, getting connected, Internet services, and eligibility for a network account, among others. In the *Introducing the Internet Plus* version, software on 3 1/2-inch disks is included for the Macintosh and DOS environments; the lecture, Powerpoint-produced slides, and the viewer are all on the disk. Training tips for each section appear as marginalia, along with numbered icons of a slide projector so that an instructor can easily tell which slide goes with which topic. The author also gives suggestions on how to integrate optional online demonstrations into a session.

As with its "parent" publication, *Crossing the Internet Threshold* (Library Solutions Press, 1993), the focus is on practical, introductory information with plenty of examples. It includes all the basics, such as an explanation of Internet addresses, as well as a thorough discussion of the standard three services—electronic mail, remote login, and file transfer. "Extended services" are also covered, including Gopher, WAIS (wide area information server), World-Wide Web, Archie, Veronica, and Mosaic. A particularly useful and practical chapter is "Instructions for Trainers," in which Jaffe, an experienced trainer and Internet authority, stresses instructional objectives and shares suggestions for making sessions successful ("beware of jargon") The index to the volume is very complete and a glossary is also included. Some chapters have lists of further reading. This supplemental volume also includes as an appendix an additional set of overheads that illustrate how easily the author's slides were adapted by another individual for a presentation on "What is the Internet, anyway?"

This publication would serve as an excellent resource for anyone responsible for basic Internet training. One caution—the spiral-bound format will not stand up long as a circulating volume and permanent binding will be necessary.—*Linda Friend, Penn State University Libraries*

Pocket Guides to the Internet. Vols. 1–6

By Mark D. Veljkov and George Hartnell. Westport, Conn.: Mecklermedia. 1994.

Vol. 1: *Telnetting*. 42p. paper, $7 (ISBN 0-88736-943-X).

Vol. 2: *File Transfer Protocol*. 54p. paper, $7 (ISBN 0-88736-944-8).

Vol. 3: *Using and Navigating Usenet*. 68p. paper, $7 (ISBN 0-88736-945-6).

Vol. 4: *The Internet E-Mail System*. 58p. paper, $7 (ISBN 0-88736-946-4).

Vol. 5: *Basic Internet Utilities*. 68p. paper, $7 (ISBN 0-88736-947-2).

Vol. 6: *Terminal Connections*. 48p. paper, $7 (ISBN 0-88736-948-0).

This six-volume set of guidebooks provides a useful and easy-to-follow introduction to the basic features and functions necessary for navigating the Internet. The material in each of these slim volumes is presented in the same format: a nine-page overview of the Internet, four to six chapters on the primary subject of the volume, followed by a glossary, sample addresses, a brief bibliography "For Further Reading," and an index to the topics covered in the volume. Each pocket guide focuses specifically on the named topic, giving concise, step-by-step instructions for mastering a particular skill.

One of the most useful features of this series of guides is the presentation of alternative methods for accessing Internet resources. For example, in volume 5, "Basic Internet Utilities," the authors describe three primary ways to use Archie and Archie servers: using a local Archie client installed on your PC or local Internet host, by Telnet, or by e-mail. Similarly, volume 6, "Terminal Connections," explores and explains a variety of hardware and software resources capable of providing different levels of access to the Internet. Individual chapters are devoted to the three basic types of Internet connections: direct or dedicated connections using Ethernet, serial connections using serial line internet protocol (SLIP) or point-to-point protocol (PPP), and dial-up connections using a modem. This type of broad approach to Internet instruction makes these guides useful to a wide range of people with a variety of computer configurations and capabilities.

Overall, these volumes should appeal primarily to new users of the Internet, offering clear definitions and instructions in language that is largely jargon-free. The more experienced Internet user should find use for individual volumes as new skills are desired. Typographical errors occur within some volumes and are annoying, but they do not detract from the instructional value of these pocket guides.—*Linda Gunter, Honnold/Mudd Library, Claremont Colleges*

Other Recent Receipts

A.V. in Public and School Libraries. Ed. Margaret J. Hughes and Bill Katz. New York: Haworth, 1994. 110p. $39.95 (ISBN 1-56024-461-5).

Bielefield, Arlene, and Lawrence Cheeseman. *Maintaining the Privacy of Library Records: A Handbook and Guide*. New York: Neal-Schuman, 1994. 212p. paper, $39.95 (ISBN 1-555570-066-7).

Bonura, Larry S. *The Art of Indexing*. Somerset, N.J.: Wiley, 1994. 232p. $26.95 (ISBN 0-471-01449-4).

Catalysts for Change: Managing Libraries in the 1990s. Ed. Gisela M. von Dran and Jennifer Cargill. New York: Haworth, 1994. 198p. $29.95 (ISBN 1-56024-516-6).

CD-ROMs In Print 1994: An International Guide to CD-ROM Multimedia, and Electronic Book Products. Ed.

Matthew Finlay and Regina Rega. Westport, Conn.: Meckler, 1994. 1136p. paper, $99.95 (ISBN 0-88736-953-7).

Cetron, Marvin J., and Owen Davies. *Mastering Information in the New Century*. Washington, D.C.: Special Libraries Association, 1994. 93p. paper $25 (ISBN 0-87111-431-3).

Current Practice in Health Sciences Librarianship. Vol. 1: Reference and Information Services in Health Sciences Libraries. Ed. M. Sandra Wood. Metuchen, N.J.: Scarecrow, 1994. 340p. $39.50 (ISBN 0-8108-2765-4).

Davis, Trisha, and James Huesmann. *Essential Guide to the Library IBM PC*. Vol. 12: Serials Control Systems for Libraries. Westport, Conn.: Mecklermedia, 1994. 169p. paper, $34.95 (ISBN 0-88736-186-2).

Discovery in the Archives of Spain and Portugal: Quincentenary Essays, 1492–1992. Ed. Lawrence J. McCrank. New York: Haworth, 1994. 590p. $79.95 (ISBN 1-56024-643-X).

Drabenstott, Karen M. *Analytical Review of the Library of the Future*. Washington, D.C.: Council on Library Resources, 1994. 200p. paper, $15 (no ISBN).

Drabenstott, Karen Markey, and Diane Vizine-Goetz. *Using Subject Headings for Online Retrieval: Theory, Practice, and Potential*. New York: Academic, 1994. 360p. $55 (ISBN 0-12-221570-2).

Eisenhart, Douglas M. *Publishing in the Information Age: A New Management Framework for the Digital Era*. Westport, Conn.: Greenwood, 1994. 312p. $55 (ISBN 0-89930-847-3).

Emerging Communities: Integrating Networked Information into Library Services. Ed. Ann P. Bishop. Champaign, Ill: Publications Office, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 1994. 304p. $30 plus shipping and handling (ISBN 0-87845-094-7).

51 Reasons: How We Use the Internet and What it Says about the Information SuperHighway. Ed. Martha Stone-Martin and Laura Breeden. Lexington, Mass.: FARNET (114 Waltham St., Lexington, MA 02173) 1994. 124p. paper, spiral-bound, $22.95 (no ISBN).

Gasaway, Laura N., and Sarah K. Wiant. *Libraries and Copyright: A Guide to Copyright Law in the 1990's*. Washington, D.C.: Special Libraries Assn., 1994. 270p. paper, $59 (ISBN 0-87111-407-0).

Gertzog, Alice, and Edwin Beckerman. *Administration of the Public Library*. Metuchen, N.J.: Scarecrow, 1994. 576p. $59.50 (ISBN 0-8108-2857-X).

Gilmer, Lois. *Interlibrary Loan: Theory and Management*. Englewood, Colo.: Libraries Unlimited, 1993. 230p. $30 (ISBN 0-87287-947-X).

Government CD-ROMs: A Practical Guide to Searching Electronic Document Databases. Ed. John Maxymuk. Westport, Conn.: Mecklermedia, 1994. 324p. paper, $47.50 (ISBN 0-88736-887-5).

Guidelines for Bibliographic Description of Interactive Multimedia. Ed. Laurel Jizba. Chicago, Ill.: American Library Assn., 1994. 43p. paper, $12, $10.80 ALA members (ISBN 0-8389-3445-5).

Gustafson, Kent L., and Jane Bandy Smith. *Research for School Library Media Specialists*. Norwood, N.J.: Ablex, 1994. 220p. paper, $24.50 (ISBN 1-56750-087-0).

Hernon, Peter. *Statistics: A Component of the Research Process*. Norwood, N.J.: Ablex, 1994. 225p. paper, $28.50 (ISBN 1-56750-093-5).

Information for Management: A Handbook. Ed. James M. Matarazzo and Miriam A. Drake. Washington, D.C.: Special Libraries Association, 1994. 181p. paper, $32 (ISBN 0-87111-427-5).

Information Vision: Professional Papers from the 85th Annual Conference of the Special Libraries Association, June 11–16, 1994, Atlanta, Georgia. Washington, D.C.: Special Libraries Assn., 1994. 67p. $25 (ISBN 0-87111-424-0).

Instruction for Information Access in Sci-Tech Libraries. Ed. Cynthia Steinke. New York: Haworth, 1994. 137p. $21.95 (ISBN 1-56024-666-9).

The Internet Library: Case Studies of Library Internet Management and Use. Ed. Julie Still. Westport, Conn.: Mecklermedia, 1994. 180p. paper, $37.50 (ISBN 0-88736-965-0).

Kearsley, Greg. *Public Access Systems: Bringing Computer Power to the People*. Norwood, N.J.: Ablex, 1994. 208p. paper, $22.50 (ISBN 0-89391-948-9).

LaGuardia, Cheryl. *The CD-ROM Primer: The ABCs of CD-ROM*. New York: Neal-Schuman, 1994. 130p. $39.95 (ISBN 1-55570-167-1).

Latin America: The Emerging Information Power: Papers Presented at the State-of-the-Art Institute, November 8–9, 1993. Washington, D.C.: Special Libraries Assn., 1993. 165p. paper, $25 (ISBN 0-87111-425-9).

Libraries as User-Centered Organizations: Imperatives for Organizational Change. Ed. Meredith A. Butler. New York: Haworth, 1994. 256p. $34.95 (ISBN 1-56024-616-2).

Library Technology Consortia: Case Studies in Design and Cooperation. Ed. Jerry Kuntz. Westport, Conn.: Mecklermedia, 1994. 112p. paper, $42.50 (ISBN 0-88736-886-7).

McGuire, Carmel, Edward J. Kazlauskas, and Anthony D. Weir. *Information Services for Innovative Organizations*. New York: Academic, 1994. 317p. $55 (ISBN 0-12-465030-9).

Meloche, Joseph. *Introductory CD-ROM Searching: The Key to Effective Ondisc Searching*. New York: Haworth, 1994. 205p. $39.95 (ISBN 1-56024-412-7).

Notess, Greg R. *Internet Access Providers: An International Resource Directory*. Westport, Conn.: Mecklermedia, 1994. 309p. paper, $30 (ISBN 0-88736-933-2).

Pagell, Ruth A., and Michael Halperin. *International Business Information: How to Find it, How to Use it*. Phoenix, Ariz.: Oryx, 1994. 384p. $74.95 (ISBN 0-89774-736-4).

Prytherch, Ray. *Information Management and Library Science: A Guide to the Literature*. Brookfield, Vt.: Gower, 1994. 323p. $59.95 (ISBN 0-566-07467-2).

Reference and Information Services in Health Sciences Libraries: Current Practice in Health Sciences Librarianship. Ed. M. Sandra Wood. Metuchen, N.J.: Scarecrow, 1994. 394p. $39.50 (ISBN 0-8108-2765-4).

Relyea, Harold C. *Silencing Science: National Security Controls and Scientific Communication*. Norwood, N.J.: Ablex, 1994. 240p. $42.50 (ISBN 1-56750-096-X).

Riechel, Rosemarie. *Public Library Services to Business*. New York: Neal-Schuman, 1994. 148p. paper, $39.95 (ISBN 1-55570-168-X).

Rose, Daniel E. *A Symbolic and Connectionist Approach to Legal Information Retrieval*. Hillsdale, N.J.: Lawrence Erlbaum Associates (365 Broadway, Hillsdale, NJ 07642-1487), 1994. 314p. $59.95 (ISBN 0-8058-1388-8).

Russel, Charlie, and Sharon Crawford. *Voodoo Unix: Mastery Tips and Masterful Tricks*. Chapel Hill, N.C.: Ventana Press, 1994. 300p. paper, $27.95 (ISBN 1-56604-067-1).

Stanton, Greta W. *Children of Separation: An Annotated Bibliography for Professionals*. Metuchen, N.J.: Scarecrow, 1994. 358p. $42.50 (ISBN 0-8108-2695).

Vickery, Brian C., and Alina Vickery. *Information Science in Theory and Practice*. New Providence, N.J.: Saur, 1994. 400p. $95 (ISBN 1-85739-017-2).

The Virtual Library: An SLA Information Kit. Washington, D.C.: Special Libraries Assn., 1994. 245p. paper, $15 (ISBN 0-87111-428-3).

Virtual Reality World's Virtual Reality Market Place 1994: Your Complete Sourcebook to the Virtual Reality Industry—Hundreds of Companies Offering Thousands of VR Products and Services. Ed. Sandra K. Heisel and Jeffrey Jacobson. 116p. paper, $39.95 (ISBN 0-88736-926-X).

Woodsworth, Anne, and others. *The Future of Education for Librarianship: Looking Forward from the Past*. Washington, D.C.: Council on Library Resources, 1994. 95p. paper, $15 (no ISBN).

Wren, Christopher, and Jill Robinson Wren. *Using Computers in Legal Research: A Guide to Lexis and Westlaw*. Madison, Wisc.: Adams & Ambrose (Dept. P, 1220 South Park St., P.O. Box 9684, Madison, WI 53715-0684) 1994. 771p. paper, $19.95 plus shipping (ISBN 0-916951-21-9).

Index to Advertisers

News and Announcements

Gaylord Bros. Announces New President and CEO

Gaylord Bros. has announced the appointment of Ronald Beckman as its president and chief executive officer. Beckman comes to Gaylord with thirty years of general financial and organizational management experience at General Electric. He has also had experience in marketing and sales, identification of customer needs, and the development of products and services for consumer and business markets.

In making the announcement, the board of directors expressed confidence that Beckman will provide the leadership and vision necessary to continue Gaylord's partnership with the library community. Gaylord Bros. has traditionally provided a variety of products and services to libraries, including the rapidly growing Gaylord Information Systems Division, and will celebrate its hundred-year anniversary in September 1996.

Hugh C. Atkinson Memorial Award Recipient Named

Dorothy Gregor, university librarian at the University of California, Berkeley, is the 1994 recipient of the Hugh C. Atkinson Memorial Award, cosponsored by four ALA divisions—ALCTS, ACRL, LAMA, and LITA. The award, $2,000 and a citation, recognizes outstanding contributions (including risk taking) to improve library services, or to library development or research, by an academic librarian in the areas of library automation or library management. The Faxon Company also awards a gold giraffe pin, designed by Tiffany's, to recognize the "sticking your neck out" aspect of the award.

"Choosing Dorothy Gregor to receive this prestigious award reflects her visions of automated services in libraries," said Willis M. Hubbard, chair of the Atkinson award committee. "She is recognized for her work in cataloging simplification, advocacy of bibliographic standards and her early work with machine-readable cataloging. Her pioneering efforts to provide information to all users of a library through innovative uses of technology are consistent with the leadership of Hugh Atkinson."

Gregor's twenty-six-year career includes positions as university librarian at the University of California, San Diego, chief of the shared cataloging division of processing services at the Library of Congress, associate university librarian for technical services and head of the serials department at the University of California, Berkeley, and reference librarian at the University of Hawaii. She has been a member of the OCLC board of directors since 1988 and the Center for Research Libraries board of directors since 1990.

The award was presented at the ACRL awards reception on Monday, June 27, during the 1994 ALA Annual Conference in Miami Beach.

Former ALA President Robert Vosper Dies at Age 80

Robert Gordon Vosper, who served as president of the American Library Association for the 1965–1966 term, died May 14 at his home in Los Angeles. He was eighty years old.

"Robert Vosper devoted his professional life of more than a half century to the advancement of librarianship, particularly through the administration of research libraries and service to the profession by his active participation in professional associations throughout the world," said ALA president Hardy R. Franklin. "By creating an environment in which people grow and serving as a mentor, he was responsible for the development of many of the profession's leaders."

Vosper received a bachelor's degree in classics from the University of Oregon and a master's degree in classics and a certificate in librarianship from the University of California, Berkeley. He began his career as a librarian in the accessions and reference departments of the university library at Berkeley in 1940. In 1942, he moved to Stanford to work in the reference department. He came to the University of California, Los Angeles, in 1944 as head of the department of acquisitions; he also served as the university's first assistant librarian and first associate librarian. Vosper retired as university librarian at UCLA in 1973, but continued to teach in the Graduate School of Library and Information Science and to serve as director of the off-campus Clark Memorial Library. He fully retired from university service in 1983.

Vosper was president of the Association of College and Research Libraries (1955–1956) and of the International Federation of Library Associations and Institutions (IFLA) from 1971 to 1977. He served as a member of the U.S. Commission to UNESCO from 1968 to 1973, as a member of the Government Advisory Committee on Overseas Book and Library Programs of the U.S. Department of State from 1970 to 1975, and as the leader of the Association of Research Libraries Farmington Plan Survey from 1957 to 1959.

Vosper received the ALA J. W.

Lippincott Award in 1985 and the UCLA Medal in 1988. An Honorary Fellow of IFLA, he was named a Guggenheim Fellow in 1969 and honored as "UCLA Man of the Year" in 1968 by the Franklin D. Murphy Associates, an organization composed of some of the university's most distinguished alumni and friends. Vosper was decorated as an officer of the Order of the Crown of Belgium in 1977.

Vosper is survived by his wife, Loraine; his children, Ingrid McCarroll, Kathyrn Katz, Elinor Charles, and Stephen Vosper; and eight grandchildren. Contributions in Robert Gordon Vosper's memory may be made to UCLA Research Library Special Collections or to the William Andrews Clark Memorial Library, Los Angeles, CA 90024.

RLG Reports Membership Surge

In 1993, The Research Libraries Group, Inc., attracted more new members than in any other recent year. RLG membership now stands at 143.

Eleven new members are outside the United States: the American Academy in Rome; McGill University in Canada; and the nine members of the Consortium of University Research Libraries (see preceding item). Four major U.S. universities have also joined in the past few months: Boston College, University of Maryland at College Park, University of Texas at Austin, and Wayne State University. Institutions of this size bring substantial new resources to RLG. University of Texas' General Libraries unit, for example, is the fifth-largest academic library in the United States.

Not all of the new members are large universities. Perhaps the most intriguing are the smaller, specialized research institutions. The American Academy in Rome is a case in point. Although a community of only seventy-five residents, it houses a research library of 112,000 volumes. The academy includes on its roster many of the leading artists and scholars of the past century, including architects John Russell Pope and Michael Graves, authors John Hersey and Nadine Gordimer, artists Frank Stella and Roy Lichtenstein, and art historian A. W. Janson.

Another new RLG member—Lawrence Berkeley Laboratory—has produced nine Nobel Prize winners. Its archives document the work of the many scientists associated with the laboratory during its fifty years.

The other new members are American University's Washington College of Law Library, the Center for Research Libraries, the Chicago Historical Society, the Institute for Advanced Study, Linda Hall Library, New York Botanical Garden, North Carolina State Archives, and the United States Holocaust Memorial Museum.

Titles in RLIN Bibliographic Files

The Research Libraries Group has just completed its semiannual RLIN database profiles. One of the most common questions we're asked is how many titles (not records) are in the RLIN bibliographic files. Table 1, provided by Karen Bendorf of RLG, shows the title and record count for each of the RLIN bibliographic files, and also indicates the files with the highest percentages of unique records.

RLG Adds Overseas Members

The Consortium of University Research Libraries has joined the Research Libraries Group. CURL's nine members—the libraries of the universities of Cambridge, Edinburgh,

Table 1
BIB Files by Category

FORMAT	RECORDS	TITLES	REC/CLUSTER
Clustered files			
BKS	55,929,501	18,208,431	3.07
MAP	265,250	233,007	1.14
REC	1,450,901	1,021,413	1.42
SCO	1,168,584	638,925	1.83
SER	3,899,581	2,237,050	1.74
VIM	226,779	200,813	1.13
Total clustered files	62,940,596	22,539,639	
Nonclustered files			
AMC	393,809	393,809	1.00
MDF	35,626	35,626	1.00
Total files	63,370,031	22,969,074	

Total BIB titles added March 1993–March 1994: 1,978,612

Source: RLIN profiling runs completed March 1994

Glasgow, Leeds, London, Manchester, and Oxford, and of Trinity College Dublin and University College London—are among the largest, oldest, and most prestigious libraries in the United Kingdom and the Republic of Ireland. The majority were in existence before 1600. One, the Bodleian Library, Oxford, was founded in 1320.

Although CURL member institutions may be ancient, their methods are up-to-date. Since 1987, their libraries have contributed to a bibliographic database of holdings, currently containing over seven million records. CURL's online bibliographic database, operated out of the University of Manchester, includes not only the cataloging input of all its members, but British National Bibliography files and Library of Congress files as well.

CURL plans to participate in a variety of RLG programs. The entire consortium will use RLIN for library support, and a number of individual member libraries have elected to participate in one or both of RLG's special membership activities: ShaRes and PRESERV.

ShaRes, RLG's Shared Resources Service, helps RLG members reduce costs and improve service by allowing them priority access to materials at other institutions within the RLG community. PRESERV, RLG's Preservation Service, helps members preserve and improve access to endangered research materials through services such as cooperative preservation microfilm projects, proposal preparation, grant administration, and exploration of new technologies.

ALA Gopher Debuts

The American Library Association has established its gopher on the Internet, in a move to expand the audience for association policies, publications, and conferences. (Gopher is software that allows document retrieval using menus.) Millions of Internet users around the world will be able to subscribe electronically to ALA journals such as *Booklist*, join ALA and several of its eleven divisions and seventeen roundtables, and register for some institutes and conferences.

Although the gopher is still under construction, much information about ALA is already available, including contact information for ALA staff (telephone numbers, fax numbers, and e-mail addresses), the *ALA Constitution, Bylaws, and Policy Manual* and all interpretations of the Library Bill of Rights, a complete listing of ALA awards, and a listing of Midwinter Meeting and Annual Conference dates and locations through the year 2004.

To access the ALA gopher, users can point their favorite gopher client to gopher.uic.edu (port 70). Once users access the gopher at the University of Illinois at Chicago, the ALA gopher can be reached by selecting the "Library" option on the main menu and then the "American Library Association" option on the submenu. Users with gopher clients that support bookmarks can then set them to make subsequent access to the ALA gopher quick and easy. Free gopher software can be secured for users without a gopher client by anonymous FTP to boombox.micro.umn.edu. The software can be found in the directory /pub/gopher. This directory contains a number of gopher clients that operate on several different platforms.

Questions or comments about the ALA gopher should be addressed to Charles Harmon, ALA headquarters librarian (charles.harmon@ala.org). Questions on using gopher software or accessing the University of Illinois at Chicago's gopher are best handled by Internet users' local system administrators.

LITA Announces New Electronic Journal: *Telecommunications Electronic Reviews*

A new electronic journal, *Telecommunications Electronic Reviews (TER)*, has been created by the Library and Information Technology Association.

TER provides reviews of and pointers to print and electronic resources for professionals dealing with networking and telecommunications, including books, articles, serials, discussion lists, software, training materials, and other items primarily related to libraries and information centers. Topics include telecommunications and networking technologies; hardware and software; network operation systems and applications; technical management issues; training and personnel issues; organizational, regional, and national networking; library perspectives; and research and development.

TER reviews are substantive and evaluative and include audience ratings and references to other similar materials. The journal will occasionally include multiple reviews of the same item and will invite readers and authors to comment on published reviews. Commentary on issues related to telecommunications and networking will also be published.

The new journal is being published irregularly using the following technologies: listserv (to be determined), gopher (info.lib.uh.edu), and World-Wide Web (to be determined). There is no subscription fee. Listserv and World-Wide Web information will be announced later on the ALA gopher (LITA section) and elsewhere.

TER seeks contributors with expertise in subject areas relevant to telecommunications and networking. Submissions should be sent as ASCII text e-mail to Thomas C.

Wilson, editor in chief, at lib4@jetson.uh.edu. For further information, contact Linda Knutson, Executive Director, LITA, 50 E. Huron St., Chicago, IL 60611; 1-800-545-2433, ext. 4270; e-mail linda.knutson@ala.org.

RLIN Database of Early European Printing Announced

At a meeting last week in Lisbon, the twenty-eight members of the Consortium of European Research Libraries selected the Research Libraries Group to supply database support for their first major project, the *Hand Press Book Database (HPB)*, which will cover European printing of the hand-press period (ca. 1455–1830).

According to the consortium's secretary, Lotte Hellinga of the British Library, "*HPB* can be expected to have a major impact on a variety of professions and disciplines in the library and academic world. By providing uniform access to records of the entire printed heritage of Europe, it will not only be a major boost for all types of historical research, but will also be a catalyst for library activities such as preservation, interlending, and substitution programs."

In the consortium's cautious estimate, the database will eventually grow to 4.5 million entries. The inclusion of this large pool of data in RLG's family of databases will further enrich RLIN's vast bibliographic resources.

The consortium emerged from meetings held in 1991 and 1992 in Munich, convened and cosponsored by the British Library and the Bavarian State Library to discuss a unified European approach to bibliographic records for early printed materials. The organization, whose members represent national and university libraries in eighteen countries, hopes to merge data from a number of national cataloging projects as well as individual libraries' files. Since member libraries have agreed to make available their already substantial machine-readable files, the project should progress rapidly.

The Research Libraries Group was selected to provide a database host, in Hellinga's words, "after three years of feasibility study, design, planning, and a tender exercise for procuring a database supplier." The project expands upon a long-standing interest at RLG, where an Early Printed Books (EPB) file on RLIN already incorporates data from some consortium members.

For information on the Consortium of European Research Libraries, contact Lotte Hellinga (phone 44-71-323-7581; fax 44-71-323-7736). For press information only, contact Richard Kohn (e-mail bl.rjk@rlg.stanford.edu).

Commission on Preservation and Access Appoints Interim President

Billy E. Frye, chairman of the Commission on Preservation and Access, Washington, D.C., has announced the appointment of M. Stuart Lynn, vice president for information technologies at Cornell University, as interim president of the commission. Lynn assumed the position upon the July 1, 1994, retirement of Patricia Battin, commission president since 1987.

"We are extremely fortunate to be able to call upon Stuart at this time to maintain the full and varied programs of the commission," Frye noted at the announcement. "His broad interests, visible leadership, and commitment to the commission's mission promise strong support to our constituencies." Lynn has served on the commission's Technology Assessment Advisory Committee (TAAC) since its founding in 1988 and has been a leader in Cornell's membership in the commission's Digital Preservation Consortium.

At Cornell, Lynn is responsible for policy, strategic planning, and coordination and development of information technologies, including library systems, network services, and associated support services. He is the author of a pivotal TAAC report, *Preservation and Access Technology: A Structured Glossary of Technical Terms* (1990), that forms the basis for a series of commission papers on the challenges of preservation and access in a digital environment. In 1992, he coauthored a report with the commission's international program officer Hans Rutimann, *Computerization Project of the Archivo General de Indias, Seville, Spain*, which assessed the technical and operational aspects of a large-scale image-scanning project.

Prior to his Cornell appointment in 1988, Lynn served for six years as director of computing affairs and professor of electrical engineering and computer science at the University of California, Berkeley. A graduate of Oxford University, Lynn received his M.A. and Ph.D. in mathematics from the University of California, Los Angeles.

Lynn has served as principal or coprincipal investigator for commission-sponsored research and development projects exploring the digital preservation of brittle documents, the application of digital technologies to the preservation of brittle books, and the application of photo-CD technology to the digital capture of image materials.

Frye indicated that all current commission initiatives, described in

the February 1994 *Working Paper on the Future* (available from the commission), will continue as planned while the board continues its search for a president. Commission initiatives are active in the areas of technology, science research, international programs, communications, central preservation microfilm collection, and education.

OCLC Gateway Software Now Available

The field test of OCLC Gateway Software came to a close on March 1, with all three test sites deciding to keep the new software interface because it significantly increased their productivity.

The UNIX-based OCLC Gateway Software connects non-OCLC terminals and workstations on a campus or local network to OCLC and other services. Using an easy, menu-driven interface, a library staff member can log on to two or more systems and toggle among them during the workday. Field testers reported that having access to several services from one workstation allowed them to be more productive. They also liked being able to transfer data between systems.

The final report from staff at the University of Pittsburgh said that the Gateway Software and the TLP telecommunications line "provide faster access to OCLC for all functions along with enormously enhanced functionality and flexibility for the same or a slightly reduced/increased net cost to the libraries." The two other test sites, the University of California, San Diego, and the University of Cincinnati, also reported excellent results.

With OCLC Gateway Software, it is very easy, for example, for reference librarians to look up a subject in the FirstSearch Catalog, check the library's online catalog to see if the item is available, and, if it's not, place an ILL request. It also allows catalogers to log on and work between the library's local system and the OCLC PRISM service, as well as to check their e-mail system conveniently.

The Gateway Software also helped the institutions extend the life of some near-obsolete equipment, such as IBM 3164 terminals, by giving them online access to multiple services.

OCLC Cataloging Symposium Proceedings Available

The proceedings of a February 4, 1994, symposium held at the ALA Midwinter Meeting in Los Angeles, titled "The Future is NOW: The Changing Face of Technical Services," are now available.

The symposium covered such topics as the role of outsourcing, extending local information resources, future electronic communications, and efficient centralized cataloging. Speakers included Michael Gorman, dean of library services, California State University, Fresno; Tia Gozzi, director of technical services, Stanford University; Arnold Hirshon, university librarian, Wright State University; Glen Holt, director, Saint Louis Public Library; Colleen Hyslop, assistant director of systems and access services and head of technical services, Michigan State University; and James Rush, executive director, PALINET. Martin Dillon, director of OCLC's library resources management division, moderated.

Copies of the proceedings are available on the Internet, in hard copy, and on videotape. The print version has been distributed to OCLC cataloging members and OCLC-affiliated regional networks. Additional copies of the print version are available on request, at no charge, from the OCLC Documentation Department, Mail Code 123, 6565 Frantz Road, Dublin, OH 43017-3395; include a complete mailing address and specify product code MAN2340 when requesting additional copies.

The 2-hour 38-minute video is available for 21-day loan through OCLC's PRISM interlibrary loan system. ILL requests, listing OCLC control number 29785444 and the video title, "The Future is NOW: The Changing Face of Technical Services," may be sent to the OCLC Information Center (holding symbol OCC). Libraries that do not participate in PRISM ILL may send requests on the ALA interlibrary loan request form to OCLC, P.O. Box 7777, Dublin, OH 43017-0702. Borrowed tapes may be duplicated for the borrowing organization's use.

For instructions on accessing the proceedings via the Internet, send the e-mail message "get ala_symposia/cataloging1994 instructions," in the body of the message, not in the subject line, to listproc@oclc.org.

WLN's *LaserCat* CD-ROM Database Now Available for Windows

LaserCat, WLN's CD-ROM database of more than four million USMARC cataloging records, is now available with a Windows user interface, with a Macintosh interface soon to follow.

New features of *LaserCat* for Windows include an integrated MARC record editor and an original cataloging module, as well as improved

keyword and Boolean searches. It is flexible and simple to use, requiring little or no training for those already familiar with Windows. *LaserCat* also includes browse and other powerful search capabilities, authority cross-references and notes, USMARC record downloading, interlibrary loan forms, MARC records for all types of materials, and over one million unique WLN-contributed records.

LaserCat is used by libraries throughout North America as a MARC record source for local systems and cataloging and as a reference tool. Some smaller libraries have had hit rates of 95–97 percent on *LaserCat* for retrospective conversion and current cataloging. A public access version, *Laser-Pac*, containing records for a single library or group of libraries, is also available from WLN.

LaserCat for Windows is offered at the same price as the DOS version. All current customers will continue to receive DOS *LaserCat* updates. For additional information and a free demonstration diskette of either version, contact WLN at 1-800-342-5956; fax (206) 923-4009; or e-mail info@wln.com. (Please specify desired version and diskette size, 3.5-inch or 5.25-inch.)

LC and OCLC Cooperate to Broaden Information Access

The Library of Congress (LC) and OCLC report significant progress in three cooperative projects to broaden access to bibliographic information. Using OCLC's PRISM service, LC has been able to reduce the books portion of its cataloging arrearage by almost one-half. OCLC is processing LC's PREMARC tapes and will be returning upgraded records for LC's older titles. And the Program for Cooperative Cataloging, formerly the National Coordinated Cataloging Program, which began using the PRISM service in mid-1993, is making information available to all OCLC member libraries more quickly than before.

For the fourth year, LC is using the OCLC PRISM service for its current cataloging and to copy-catalog titles from its arrearage. OCLC is helping in this effort by providing $75,000 worth of online cataloging credits to LC. In addition to cataloging monographic titles on PRISM, LC began in November to catalog its backlog of sound recordings by inputting records directly in the OCLC Online Union Catalog.

"We've made outstanding progress in reducing our books arrearage by 47.7 percent (from 898,030 to 466,720) since 1989 through a number of initiatives, including the use of copy cataloging," said Sarah Thomas, LC director for cataloging. "Our goal is to eliminate the entire books arrearage by Dec. 31, 2000, while remaining current with the cataloging of incoming receipts."

OCLC is also processing tapes of selected LC PREMARC records, items cataloged before 1968, the year LC began to produce cataloging in the MARC (MAchine-Readable Cataloging) communication format. LC cataloging created before 1968 was converted to machine-readable form by the Carrollton Press in the 1970s, but those five million records were never merged into the LC's current catalog. Because of subsequent updates to name and subject headings, the records are out of date and some of the information is incompatible with current cataloging standards.

OCLC member libraries have entered nearly two million pre-1968 LC records in OCLC's Online Union Catalog (OLUC). These records have been upgraded through numerous manual and automated corrections to the OCLC database, including AACR2 correction, the Duplicate Detection and Resolution program, three phases of the Name and Subject Headings Correction Project, and efforts by the Online Data Quality Control staff at OCLC. Library staff members at OCLC libraries have also edited, enhanced, and upgraded many OLUC records through cooperative database enrichment programs. Over the next three years, OCLC expects to process selected LC PREMARC records, matching them against existing OCLC records, and replacing approximately 1.5 million PREMARC records with upgraded OCLC records.

"The PREMARC replacement process enables us, in effect, to transform retrospective records into ones that can fit in our current catalog—a process that is beneficial to LC staff and users, with additional advantages in terms of timeliness and per-record cost," said Kay Guiles, LC senior cataloging policy specialist. "Were we to process the 1.5 million records at LC using the procedures originally planned, we estimate it would take 285 staff years."

Four of the seven libraries that participate in the Program for Cooperative Cataloging (PCC) are now cataloging records on the OCLC PRISM service rather than solely on the LC MUMS system. This change has speeded up the availability of records and reduced costs for PCC members.

The Library of Congress and OCLC work together on other programs and activities that benefit the library community and its users. The CONSER program, the United States Newspaper Program, the NACO Linked Systems Project, and the Fiction Project are examples of cooperation among LC, OCLC, and volunteer libraries.

TODD
NETWOR

from 4 to

CD-ROM
in ONE CA

Todd massive data retrieval Tota
dockable double-speed drives and
speed drives have a three-year w
feature of Todd Total Access Syste
Todd services include hardware
customized systems, and network in
Todd CD-ROM hardware will oper
environment with all administrating
256 CD-ROM drives with the Todd e.
445-TODD *or (718) 343-1040 for more*

Showing a
console and a
tower. Consoles available in five different sizes. Both consoles and tower with internal CPU units and dockable drives. Network towers, multidrives and consoles with your existing workstations.

Todd dockable drives slip in and out without tools or wiring. It is easy to expand your system.

since 1971
over 20 years
OF EXCELLENCE
TODD
Enterprises

224-49 67th Avenue · Bayside, NY 11364
718 343-1040 · FAX 718 343-9180 · BBS 516 829-0212

800 445-TODD

available on GSA contract

Control

the key to the effectiveness of any
ARS)™ compares your headings to
cal subject and/or WLN authority
authorized forms; and supplies a
authority records including notes
ized service throughout your project,
sive bibliographic record and author-
gramming based on your specifica-
rates, and three levels of manual
lored to your needs at a price you
member, invisible yet indispensable.

aration . Customized Programming . Smart
eview . CD-ROM Public Access Catalogs .
FTP.

thority control and how the
***RS)* can help my library.**

____ Title _____

_____

_____

_____

Telephone _____ Best time to be reached _____ a.m. _____ p.m.

Type of project: ❑ Authority control ❑ Holdings/item conversion ❑ Smart barcodes ❑ Other _____

If authority control: ❑ LC Name/Subject Authority File ❑ LC Children's ❑ NLM MeSH
❑ WLN Authority File ❑ Updating Service ❑ Other_____

Size of database _____ Local system vendor _____ Source of MARC records _____

Projected start date _____ Completion date _____

Comments _____

Signature _____

WLN

TURNING INFORMATION INTO KNOWLEDGE

WLN . P.O. Box 3888 . Lacey, WA 98503-0888 . 1-800-DIALWLN . (206) 923-4000 . FAX (206) 923-4009