Volume 4
Number 2

ISSN 0730-9295

Information Technology and Libraries

June 1985

CONTENTS

ITLBDC 4(2) 89-196 (1985); Inf Technol Libr

Information Technology and Libraries

Volume 4, Number 2: June 1985

CONTENTS

ITLBDC 4(2) 89-196 (1985); Inf Technol Libr

Information Technology and Libraries

Information Technology and Libraries 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:* Donald P. Hammer. 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 88 of the March 1985 issue.

Manuscripts of articles, communications, and news items should be addressed to: William Gray Potter, Editor, *Information Technology and Libraries*, Hayden Library, Arizona State University, Tempe, Arizona 85281. Copies of materials for review should be addressed to: Karin A. Trainer, *ITAL Book Reviews*, Yale University Library, 120 High St., P.O. Box 1603A, Yale Station, New Haven, CT 06520. Advertising arrangements should be made with William Z. Schenck, University of Oregon Library, Eugene, OR 97403.

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology Association. Subscription price, $12.50, is included in membership dues. Nonmembers may subscribe for $25 per year. Single copies, $7.50.

Circulation and Production: American Library Association, 50 E. Huron St., Chicago, IL 60611. Please allow six weeks for change of address.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by the Library and Information Technology Association or the American Library Association.

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 *Computer Contents, Computer Literature Index, Current Contents, Current Index to Journals in Education, Education, Library Literature, Magazine Index, and NewSearch*. Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

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.*

Preparing for the Use of Classification in Online Cataloging Systems and in Online Catalogs

Pauline A. Cochrane and Karen Markey

INTRODUCTION

Not all aspects of the cataloging process have been aided by computer-based systems since their introduction in the middle 1960s. While recent computer-based system developments have accelerated the importance of files for bibliographic records and related files for name and subject authority records, there has been, unfortunately, little parallel effort to automate the file of classification schedules that controls the shelf arrangement of library material and that may help to improve subject access online. The Library of Congress Classification (LCC) and Dewey Decimal Classification (DDC) schedules can provide the arrangement of catalog records in shelflist order as well as an outline of knowledge for browsing that is more systematic and logical than an alphabetical list of subject headings.[1,2] If properly exhibited online, these systematic arrangements may become very useful displays for the information seeker.

For purposes of discussion, we assume that someday there will be machine-readable versions of the Library of Congress Classification (LCC) and Dewey Decimal Classification (DDC) available to librarians, which they will use for many purposes. We believe the presumption that this will not happen eventually is false. There are precedents for such a file being made available as the following discussion reveals.

The Universal Decimal Classification (UDC) was put into machine-readable form almost in its entirety in the mid-1960s. Some European libraries use abridgments of this file for printing their classified catalogs, and a research team used this file in one of the first online catalog/retrieval systems that predated DIALOG.[3-5]

In 1979, the nineteenth edition of the printed *Dewey Decimal Classification (DDC)* was produced by computerized photocomposition. The print tapes of this edition are being converted presently into a format that will facilitate editorial management of this edition and publication of future editions through an online editorial support system.

Selected portions of this machine-readable file of DDC Schedules and Relative Index are the object of a research project begun in 1984 by the OCLC Office of Research and supported by Forest Press (publisher of *DDC*), the Council on Library Resources, and OCLC. The Dewey Decimal Classification (DDC) Online Project team has taken a machine-readable file of records from selected portions of the DDC Schedules and Relative Index to create an online searcher's tool for subject access, browsing, and display in an online catalog with the bibliographic records of

Pauline A. Cochrane is professor, School of Information Studies, Syracuse University, and Karen Markey is research scientist, Office of Research, Online Computer Library Center (OCLC).

four different libraries. Even after this project demonstrates the utility of a classification system in an online catalog, there is still the question of how library classification schedule and index record files should be constructed to maximize their multiple use. Their use could eventually go far beyond photocomposition and shelf control to include more online displays.

LIBRARY CLASSIFICATION AND AUTOMATION, 1940–84

The application of automation to library classification schemes dates back to the late 1940s, when punched card equipment was used to express and sort Universal Decimal Classification numbers.[6] Freeman[7] and Rigby[8,9] summarized early experiments with computers and classification that predate *online* bibliographic retrieval systems. These early experiments were preoccupied with the difficulties of expressing and manipulating notations constructed from classification schemes with the available technology.[10]

By 1964, the information community envisioned the following applications of computers to classification:[11]

1. concept groupings for vocabulary control;
2. control and display of classification schedules in one or more languages;
3. systematic listings of titles and abstracts;
4. classified arrangements of bibliographies and indexes;
5. concept matching in systems for selective dissemination of information.

The first online bibliographic searching system in which a classification scheme was used for subject access and browsing was demonstrated in 1967–68 by Freeman and Atherton.[12,13] The researchers created a file of numbers for nuclear science and the English-language descriptions represented by these UDC numbers. Cross-references and scope notes from the schedules formed part of the record. They also loaded a bibliographic file where these UDC numbers appeared as part of the index record for bibliographic items. For searching and retrieval, an interactive, command-driven system called AUDACIOUS was used.[14] (It was based on early versions of NASA/Recon.) Users entered search arguments expressed as an English-language term or phrase or as a UDC number; Boolean operators were available for use. AUDACIOUS users could even display portions of UDC schedules. Figure 1 presents online displays illustrating the search scenario in AUDACIOUS.

The results of the AUDACIOUS experiments with the UDC led Freeman and Atherton to conclude that the UDC could be used as the indexing language in a mechanized system and that their success with the UDC might be generalized to other tools, such as the Dewey Decimal and Library of Congress Classifications.[15]

There are other experiments with computers and classification,[16,17] but they were narrower in scope than the AUDACIOUS/UDC project of Freeman and Atherton. To teach library school students about automation, Atherton and Tessier[18] implemented a computer-based retrieval system in which the index to the Z schedule of the Library of Congress was converted into a KWIC (keyword-in-context) index and made searchable in the retrieval system. Using computer-assisted techniques, Olson[19] produced and published a fifteen-volume set of alphabetical and classified indexes to the indexes of the Library of Congress Classification (LCC) schedules. Olson employed student assistants at Mankato State University to input indexes of the LCC schedules from which were produced an alphabetically arranged, geographical name index, a subject keyword-in-context index, a biographical subject index, an author index, and a classified index to persons. This file has not been used as an online searching tool, but it may have some use even though it is now somewhat out-of-date.

The *Computing Reviews* classification was organized into browsing displays by Fox and Palay[20] in the BROWSE system, which is used by library patrons to search the Carnegie-Mellon University Computer Science Department library collection. The BROWSE system is menu-based and utilizes the tree structure of the *Computing Reviews*' classification as the framework for the menu-based approach. BROWSE does not allow users to access the specific

Online Search			*Explanation*
> Radioactivity			*The user inputs* Radioactivity.
RADIOACTIVITY		539.16	*The system responds with lines*
RADIOACTIVITY IN THE ATMOSPHERE		551.510.7	*from the UDC schedule where*
RADIOACTIVITY IN THE EARTH		550.378	*the term* Radioactivity *appears.*
RADIOACTIVITY IN OCEANS AND SEAS		551.46:539.16	
RADIOACTIVITY IN RIVERS		551.48:539.16	
RADIOACTIVITY STANDARDS		539.16.081	
>551.510.7			*The user is interested in* Radioactivity in the Atmosphere *and inputs its number.*
551.51	STRUCTURE, MECHANICS, AND THERMODYNAMICS OF THE ATMOSPHERE		*The system response is a list of lines from the UDC schedules in the neighborhood of that UDC number, 551.510.7, thereby showing some of the hierarchy.*
551.510.3	PHYSICAL PROPERTIES OF THE ATMOSPHERE		
551.510.4	COMPOSITION OF THE ATMOSPHERE		
551.510.5	GENERAL STRUCTURE OF THE ATMOSPHERE		
551.510.7	RADIOACTIVITY IN THE ATMOSPHERE		
551.510.71	NATURAL RADIOACTIVITY		
551.510.72	ARTIFICIAL RADIOACTIVITY		*Confer (cf.) notes suggest alternate placement of this concept.*
551.510.721	RADIOACTIVE FALL-OUT Cf. 614.73, 628.511		

Source: Robert R. Freeman and Pauline Atherton, "File Organization and Search Strategy Using the Universal Decimal Classification in Mechanized Reference Retrieval Systems," in *Mechanized Information Storage, Retrieval, and Dissemination; Proceedings of the FIDIFIP Joint Conference*, ed. Kjell Samuelson (Amsterdam: North-Holland Publishing Co., 1968), p.146. Explanation added by Cochrane and Markey.

Fig. 1. UDC as an Online Browsing Tool for Subject Access.

part of the classification in which they are most interested. Rather, users have to travel from the most general parts of the classification to the most specific to arrive at the classification area of interest, treating hierarchy as a strict network path. Geller and Lesk's[21] experimental menu-based retrieval system at Bell Labs requires users to travel from the most general parts of a classification scheme to the most specific subjects instead of allowing users to go directly to the subjects of interest. Also, the Bell Labs' experimental menu-based system employs Library of Congress subject headings to represent captions from the DDC schedules.

Besides research efforts with experimental or prototype systems, classification in an online retrieval environment has been discussed in published articles. As early as 1968, Richmond enumerated three approaches to the application of computers to classification in libraries. The most interesting approach, in view of the proposed research, is the compilation of a thesaurus from classification schedules, indexes, and subject headings.[22] Writing six years later, Richmond[23] emphasized that LC classification *descriptions* and *indexes* would be valuable if converted to machine-readable form, augmented to MARC records, and indexed and/or compiled into a superthesaurus. The results of the Subject Access Project[24] also underlined the critical need to augment subject access to libraries' bibliographic records.

As libraries began to plan for online public access catalogs, interest in the findings of traditional library catalog use studies was renewed. When reviewing studies of traditional catalog use in relation to the online catalog, Atherton[25] demonstrated that findings of catalog use studies had gone unheeded, and she admonished librarians not to ignore these findings when designing online catalogs. Realizing the growing importance of subject access in online catalogs, she suggested in 1978 that "classification schemes used in libraries can provide subject access in online catalogs, and the words

in the classification schedule captions may provide the free or natural text needed for greater access to topics in the contextual areas of the schedule, thereby providing a browsing feature in online catalogs."[26]

After a hiatus of several years, interest was renewed in the application of computers to library classification in early 1981 because of the availability of online retrieval systems to library patrons. Many published articles appeared, and major conferences were devoted to classification in the online environment. Up to this time, access to online retrieval systems was chiefly the realm of library staff, who questioned patrons about their topics and performed the online searches for them.

Svenonius[27] suggested that classification could be used in an online environment to facilitate bookshelf browsing. Williamson[28] predicted that classification would assume a new and important role in the information environment of the future, especially as a browsing device in online catalogs. Mandel,[29] concerned about the need to improve subject access to libraries' book collections, suggested that LCC schedule captions be added to MARC records. Her suggestion is similar to the idea in Richmond's 1974 article.[30] Gorman[31] also noted that the computer will have great value for library classification schemes. Describing a subject retrieval function for OCLC's Online Union Catalog, Mischo[32] considered an enhancement that would add LC classification schedule terms to subject authority records by matching LC classification numbers (in 053 fields of subject authority records) to the LC classification schedules. Hildreth[33] recommended the provision of automatic linkages between subject headings and class numbers of retrieved items in online catalog searches to assist searchers in online catalog browsing activities. Hill[34] warned that classification access in online catalogs must consider past and present classification practices in libraries. She highlighted the problems of staff shortages and classification numbers, problems that could impede access in online catalogs. Council on Library Resources–sponsored research projects, position papers, and proceedings of invitational conferences have helped to stimulate interest in online access to classification schedules and index entries.

Conclusions from the CLR-sponsored Online Catalog Projects[35-37] underlined the critical need for improved subject access to libraries' online databases. Online catalog users surveyed in the projects ranked the "ability to view a list of related subjects" first among improvements to online catalogs.[38,39] Many systems designers have assumed that they could provide this search capability by allowing online catalog searchers to peruse the Library of Congress Subject Headings (LCSH) online, since this would enable searchers to check related, coordinate, or more specific terms. However, comprehensive review of LCSH criticism concludes that LCSH leads searchers *out* of the subject, not *into* it.[40] The weaknesses of LCSH's syndetic structure (the relationship between headings) are well documented. Some have suggested that the implementation of a classification schedule online might provide library users with the ability to view a list of subjects related to their search topic conceptually, since this is what library users have been doing all along when they browse at the bookshelf.[41]

The CLR-supported position paper *Subject Access in the Online Catalog* by Mandel[42] recommended development of a "program to add to MARC records the appropriate terms from the LC classification schedules whenever certain class numbers appeared in the record" to enhance subject access to libraries' bibliographic records. Response to Mandel's paper and preliminary results regarding subject access from the CLR-sponsored Online Catalog Projects inspired CLR to sponsor a June 1982 invitational meeting of library researchers, practitioners, and policy-makers to identify long- and short-term recommendations to improve subject access to bibliographic databases. One short-term issue/project was to "encourage Forest Press to authorize the design of a machine-readable format for the *Dewey Decimal Classification* . . . and to distribute, with periodical updates, for search and display only, the resulting database."[43] One long-term issue/project was to establish the utility of the *Library of Congress Classification* "in machine-readable form for users of on-

line public access catalogs and catalogers."[44]

The growth of online public access catalogs in libraries, their acceptance by library patrons, and the need for enhanced subject access to MARC records provided the impetus to rekindle interest in computerized library classification by 1984. The timing for a research effort into online classification was favorable in view of the interest and curiosity expressed by library researchers, theoreticians, and practitioners in very recent published literature. CLR, in restricting the first research it funded in this area to the Dewey Decimal Classification, was influenced by the availability of the DDC in machine-readable form.[45]

Knowledge that some existing online catalogs were providing classification number access and shelflist browsing capabilities pointed to some rudimentary developments in this area. The CLR research enhances and expands on that online capability.

CLASSIFICATION NUMBER ACCESS IN EXISTING ONLINE CATALOGS

Libraries' bibliographic records in online catalogs typically contain classification numbers like LCC and DDC. A number of online catalogs have incorporated class number searching as a mode of subject access. Although the actual search routines for class number searching differ from system to system, they all do something similar to searching the DDC number on BLAISE (the bibliographic retrieval system of the British Library): they search the inverted file of class numbers in the bibliographic database (the cataloging records) and retrieve all the books with the particular call or class number. The major drawback of this approach is that searchers are required to know the *exact class number* that represents their search topic. What they view as a result is a series of slightly related book titles, not a systematic outline of their subject. If they could search the text of classification schedules (as they could in the AUDACIOUS system of 1968), the text accompanying the class number would explain how broadly or narrowly the subject was covered by the class outline. Thus, the searcher could then match the topic with the terminology of the classification.

Figure 2 is an example of an online searching feature of LCS (Library Control System, the online catalog of the Ohio State University Libraries) that allows users to browse the shelflist. Users enter an LC class number that they have previously identified as relevant to their interest and retrieve a list of class number and titles in shelflist order matching or nearly matching the input call number. Users can browse forward and backward in this display.[46] While systems designers have implemented call number or shelflist searching in many ways, each requires the searcher to know the exact class number of interest.

The approach to classification number searching in LCS is one of five approaches summarized in table 1. The ways in which online catalog systems designers have implemented each of the five typical approaches to classification number searching identified in table 1 have serious flaws. In all five approaches, the display of retrieved items is seldom in shelflist order. Online catalog designers must make a special effort to ensure that display of items retrieved in class number searches are in shelflist order instead of the collating sequence of the online catalog's hardware. Library staff using the online catalog to classify books or to facilitate shelflisting tasks have openly expressed their displeasure with such displays, since they are not very helpful during the course of their daily classifying and shelflisting efforts.[47]

In the first three approaches to classification number searching, online catalog users do not know the subject matter of items retrieved by the entered class number until the records are displayed. The LCS approach to class number searching (see figure 2) overcomes this drawback. It lists call numbers in shelflist order and accompanies listed call numbers with title information (i.e., bibliographic information that often indicates the subject matter of the item).

A review of the features of forty-two existing online catalogs[48-52] reveals that almost three-quarters of these catalogs provided classification or call number access. However, this figure is somewhat inflated because call number searching in a number of online catalogs is associated with finding

Online Catalog Search				*Explanation*
SUB/RIOTS—UNITED STATES				
11 1 RIOTS—JAPAN—TOYAMA PREFECTURE—HISTORY				*The searcher enters a subject heading that describes his topic on racial riots in the U.S. In response, the online catalog shows that there are thirty-eight items assigned the user-entered subject heading (line 15).*
12 1 RIOTS—MASSACHUSETTS				
13 1 RIOTS—NEW YORK (CITY)—HISTORY				
14 1 RIOTS—OHIO				
*15 38 RIOTS—UNITED STATES				
16 8 RIOTS—UNITED STATES—ADDRESSES, ESSAYS, LECTURES				
•				
•				
•				
PAGE 2 OF 3 # FOR PRECEDING PAGE, ENTER PS1; FOR FOLLOWING PAGE, ENTER PS3; ENTER TBL/ AND LINE NO. FOR TITLES				
TBL/15				
RIOTS—UNITED STATES		(38 TITLES)		*The user selects the subject heading that matched the entered phrase. This results in a list of the first ten of the thirty-eight retrieved items. The searcher can scan the list, select a title, and obtain more detailed bibliographic or circulation information, or proceed to the second list of ten titles.*
01*Button, James W., 1942-	Black violence : political impact of	1978	FBR	
02 Lipsky, Michael.	Commission politics : the processin	1977	FBR	
03*Lipsky, Michael.	Commission politics : the processin	1977	FBR	
04*Lipsky, Michael.	Commission politics : the processin	1976	FBR	
05 Boskin, Joseph.	Urban racial violence in the twentie	1976	FBR	
06 Waskow, Arthur I	From race riot to sit-in, 1919 and t	1975	FBR	
•				
•				
•				

Fig. 2. Class Number Search in Ohio State's LCS Online Catalog.

Online Catalog Search	*Explanation*
PAGE 1 # FOR MORE TITLES ENTER PG2; FOR AVAILABILITY ENTER DSL/ AND LINE NO.	
TBL/5	*The searcher has chosen to display detailed bibliographic information for the fifth listed title.*
HV6477B61976 Boskin, Joseph. Urban racial violence in the twentieth century. 2d ed. Beverly Hills (Calif.) : Glencoe Press, 1976. xii, 188 p. ; 23 cm. Includes bibliographical footnotes. SUB: 1. RIOTS—UNITED STATES 2. UNITED STATES—RACE RELATIONS LC CARD #:75-14920 TITLE #:2054022 OCLC #:2341309 &ZZ770101	
SPS/HV6477	*The searcher then enters part of the call number (HV 6477) of the item using the shelf position search (SPS). The result is a list of titles in shelf-list order. The searcher can browse forward or backward to display titles of items in this call number area.*
11 HV6464G1962A / Ginzburg, / 100 years of lynchings/1969 12 HV6464R2 / RAPER, ART / THE TRAGEDY OF LYNCHING/1933 13 HV6464R21969B / Raper, Art / The tragedy of lynching/1969 14 HV6465F6M351982 / McGovern, / Anatomy of a lynching : the killin/1982 15 HV6477 16 HV6477A76 / ALLEN, HOD / VIOLENCE AND RIOTS IN URBAN AMERICA/1969 17 HV6477A831970 / Assembly o / The Kerner report revisited; final /1970 18 HV6477B58 / BOEHME, LI / CARTE BLANCHE FOR CHAOS/1970 19 HV6477B6 / Boskin, Jo / Urban racial violence in the twentieth c/1969 1A HV6477B61976 / Boskin, Jo / Urban racial violence in the twentie/1976 1B HV6477C51970 / CHIKOTA, R / RIOT IN THE CITIES/1970	
PAGE 2. FOR PRECEDING PAGE, ENTER PG1; FOR FOLLOWING PAGE, ENTER PG3	
PG3	
21 HV6477F44 / Feldberg, / The turbulent era : riot and disorder in /1980 22 HV6477G66MICROFILM / Goding, E / Urban race riots and social chan/1977 ● ● ●	

Fig. 2. Continued

Table 1. Five Approaches to Class Number Searching in Existing Online Catalogs

User Enters	System Response	User Option
1. Class Number	Number of items retrieved that exactly match user-entered class number	Display Bibliographic Records
2. Class Number	Number of items beginning with user-entered class number (implicit truncation)	Display Bibliographic Records
3. Class Number	List of corresponding class numbers in system usually in machine collating order (with or without number of retrievals)	Select line in list to get display of bibliographic records
4. Class Number	List of class numbers and brief bibliographic record	Browse forward and/or backward or select line in list to display bibliographic records
5. Call Number	Bibliographic record bearing call number entered and circulation information	(*Rarely*) message indicating how to decode listed circulation information

out the circulation status of an item, i.e., the fifth approach described in table 1.

Let us repeat the major drawback to call or classification number searching in all existing online catalogs: the searcher must know the subject matter represented by a particular call or class number to effectively use such number searching. The strategy employed by the LCS searcher in figure 2 is exemplary in this regard, for the searcher has first identified a class number relevant to his search query by entering the assigned subject heading "Riots—United States" and culling the class number from a displayed bibliographic record for subsequent entry into the online catalog. However, few online catalog users use or know how to use such a strategy. In fact, call or class number access points are the least frequent types of access points entered into online catalogs and amount to less than 10 percent of access points entered into online catalogs as determined by transaction log analysis.[53,54] Considering that subject access points (i.e., in the form of user-entered precoordinated subject headings and/or keywords) represented the majority of access points entered into online catalogs,[55] call and class number searching could hold a great potential for improving online catalog users' subject searching experiences and providing possibilities for subject search strategies not feasible through the alphabetical approach of subject headings and/or keywords presently supported by online catalogs. But the method needs improving if users are to adopt it, and one way to improve it is to permit access to the *meaning* of those numbers.

FUNCTIONS OF CLASSIFICATION RECORDS IN ONLINE CATALOGING SYSTEMS AND IN ONLINE CATALOGS

There are several functional areas of library work that can be served when classification information is available in machine-readable form. For the editor of a classification, the most obvious functions of a machine-readable classification file and online editing systems are to maintain, edit, update, and produce the classification in a variety of forms including print, micrographic, and online. The editor would be able to see all changes, either approved or waiting approval. There could be a timely distribution of these updates by the publishers of the classifications, so that printed, micrographic, and machine-readable versions of the classification could be distributed.

For library technical services staff performing classificatory tasks, the machine-readable files of their library classification could be accessed through a centralized cataloging subsystem of the type offered by OCLC, RLIN, or WLN. Library staff

would then have up-to-date schedules and the classification practice expressed by millions of catalog records. Shelflisting could also be facilitated.

When library systems staff or turnkey vendors integrate the library classification file into an online catalog, they could produce from the classification file a tailored version of the library classification that follows that library's practice. Libraries incorporating a classification scheme into the subject searching capabilities of their online catalog should consider using the information provided in relocation and discontinued notes for reclassifying their collection through a global update feature of the kind employed in online catalogs' local subject authority function.

In the DDC Online Project mentioned above, bibliographic records were enhanced with captions from the DDC Schedules and Relative Index entries. Captions were entered into a field entitled "Dewey Subject" in bibliographic records, displayed in full record displays in subject searches, and processed into various searchable subject indexes in the experimental online catalog. DDC Relative Index entries were not displayed in bibliographic records but were processed into various subject indexes. Fields from the DDC Schedules that contained subject-rich information (i.e., "class here" note, example note, scope note, and "including" note) were processed into subject indexes and displayed in screens of captions from the classification schedules. Fields that contained relocation information—i.e., relocation note, discontinued note, standard subdivisions exception note, and optional provisions note—were also displayed in screens of captions from the classification schedules and treated like *see* references to alert users to another (or other) class number(s) where they would find additional material. Other fields from the original DDC Schedules were discarded because they were not considered useful to subject searchers of a library's online catalog.

For the online catalog user, library classification becomes a tool for subject access, browsing, and display. In the DDC Online Project, the classification records provided the following benefits to subject searchers:

- Browsing capability to an exact location or multiple locations in the DDC Schedules (using captions) in response to:
  1. Subject keyword searches
  2. Class number searches
  3. Selection of an entry from the DDC Relative Index contained in online alphabetical lists of Relative Index entries.
- Enriched subject description in bibliographic records by the addition of classification captions that were processed into searchable subject keyword indexes to applicable bibliographic records.

The DDC Relative Index records provided the following benefits to subject searchers:

- Browsing capability to alphabetically arranged Relative Index entries in response to subject searches for directing users to bibliographic records or records in the classification file.
- Enhanced subject access to bibliographic records by directing users to bibliographic records when these records' class numbers and the class numbers corresponding to Relative Index entries matched.

The ways in which the CLR-funded project fashioned the DDC into an online users' tool for subject access, browsing, and display are indicative of some but not all of the possible functions of classification records. An experiment with the LC Classification online could demonstrate links between machine-readable LC Subject Authority records bearing LC class numbers, and the LC class numbers in bibliographic records. More than 30 percent of the records in the present Subject Authority file produced and distributed by LC have at least one class number noted. For example, the subject authority record for the LC subject heading "Electric Machinery, Synchronous" bears the class number TK 2731. The Subject Cataloging Division of the Library of Congress is quick to emphasize that the class numbers in the Subject Authority file may not be accurate and up-to-date, since there has been no concerted effort in recent years to maintain the integrity of this information. If the subject

heading was established in 1921 and never revised, then the class number listed with it in the Subject Authority file has not been revised, even though a better number may have been found to arrange the subject in a classified sequence.

The present LC Subject Authority file used for printing editions of *LCSH* may not have up-to-date information to support a correlation between a subject heading and a class number that would agree with the information in the bibliographic file. Some may question whether this information needs to be accurately correlated. The art of classifying and subject cataloging is a subjective process, and our files represent the practices of too many hands to ever expect more than an approximate uniformity. Others assert that the value of an amalgamation of these various subject access routes has a valuable use as a pointer—not as pinpointers.

When users select a subject heading from a retrieved bibliographic record or from a list of subject headings, the online catalog system can determine, as a computation, what classification areas are common to the greatest number of items bearing that chosen subject heading. The user would see the library classification displays with captions and could retrieve additional items pertinent to the user's topic. In recent years, two commercial publishers have tried correlations between LCC and LCSH in an attempt to provide libraries with new tools for subject analysis,[56,57] but no one has used these files online. The closest online approximation to such an operation of suggesting classification areas common to the majority of retrieved items has been implemented in the University of Illinois' Online Catalog, which suggests to the user subject headings common to the majority of items retrieved in a title keyword search.

Before classification schemes become converted into and available in machine-readable form, we should address several questions including the following:

1. What are the purposes for which we want to use such a file?
2. How should the record be structured to facilitate those uses?
3. Should "better numbers" be included in machine-readable subject authority records?
4. What notes in the machine-readable record need rewording to be more understandable to library patron users of online catalogs?
5. How can the editing process be facilitated and improved by automated techniques?
6. What if existing layout and typefont style for printed schedules could be improved? How would it be changed in online displays?
7. What parts of schedules and auxiliary table information are to be considered useful in online prompts for novice classifiers?
8. What notes or instructions in classification schedules are considered essential to classifiers? To library patrons?
9. How can we alter subfield coding of classification numbers in bibliographic records to facilitate the identification of the source(s) of synthesized class numbers?
10. Should the various subject access mechanisms used in MARC bibliographic records (subject headings, class numbers [DDC and LCC]) be linked in the various authority files and online catalog files? If so, how?

SPECIFICATIONS FOR RECORDS IN A MACHINE-READABLE LIBRARY CLASSIFICATION AUTHORITY FILE

For purposes of constructing a file of LCC and DDC number records, it will be necessary to take into account the relationship of a *class number* to the various operations, uses, users, and products associated with it. It will not be enough to think of it as a *line on a page* of a printed product entitled, for example, "Library of Congress Classification, Class T, TECHNOLOGY, Fifth Edition, 1971."

Even within the context of a single product, i.e., the printed classification schedules, a *class number record* has several uses. It will help form a line in the schedule; it may contain the lines in the alphabetic index related to it; it may have a purpose in

the *Outline* or *Synopsis* portions of the schedule; and it should be traced to any mention of it in a confer note (cf.) or *see* reference elsewhere in the file.

The class number has a caption in the classification schedule that presents a set of data elements with a self-organizing utility (though it is now only a rough version of what it could become in machine-readable form). If manipulated properly, class number records with several data elements could be used in an online mode for bibliographic file segmentation and search output sorting and as browsing aids for users. These records and data elements might also be useful in update and maintenance functions across several bibliographic files and indexes.

Table 2 lists a set of minimal data elements for records representing either the Library of Congress or Dewey Decimal Classification. In general, online displays of the classification for use by library staff or patrons could be composed of fields from authority class number records.

The number of different notes that appear in classification schedules and auxiliary tables poses a separate problem for presentation online. Table 3 lists notes occurring in the printed LCC and DDC Schedules and auxiliary tables. Most notes pertain to classifiers' activities, but other notes (such as scope note, example note, "including" note, and "class here" note) provide additional subject information about a particular class number or class number range that would be valuable in an information retrieval application. "Class elsewhere," cross-reference, *see also*, former heading, standard subdivisions exception, "discontinued," and relocation notes provide useful information, often of a historical nature, for identifying other area(s) of the classification where books on a topic are (or were) classified. Online classifiers will need the information for building classification numbers contained in the schedules or tables in "add" notes, "arrange alphabetically" notes, "successive Cutter" notes, "standard subdivision table" notes, "arrange alphabetically by Cutter as specified" notes, and "standard subdivision zeros" notes. Besides these notes, there are a variety of footnotes in schedules and tables that explain how to add to the base number from tables or schedules or how to exercise an option to successive Cutter numbers.

The back-of-the-schedule index to individual LCC schedules and the DDC Relative Index are specially composed index entries. They may represent a classification number or a range, and the index contains various cross-references specially prepared. There is some consulting of LCSH as the index entries are prepared, but no attempt is made to achieve compatibility between index entries and LCSH headings used when the class number is applied to classify a book. The integration of all these subject access vocabularies will be a formidable task, if ever attempted. Whether each index entry is to form a separate record or be incorporated as data elements of Schedule records is a question that needs further exploration. For the present, we suggest here a separate record format for classification index records. Table 4 lists data fields and subfields that would be needed in such index records.

NATURE OF THE DDC AND LCC RECORD AND FILE STRUCTURE

Unlike a bibliographic record file, a name-authority file, or a subject-authority file, a classification schedule file is a file of *nested* records and requires a special design. The analyses, in the mid-1960s, of a machine-readable file for the UDC[58,59] and later, in 1969, for the schedule of LCC[60] led to the conclusion that there was a need for linking or nesting individual lines or segments of the classification schedule. Although linear, the schedule has dimensions, ranges, and divisions that must be accommodated if it is to be manipulated by the computer.

Figure 3 is an example of this nesting or linkage for the DDC number 621.3133 if a *unit record* in the DDC is conceived of as *a classification number with a caption and notes*. Bars on the bar chart represent the amount of the DDC Schedules linked to that unit record because its classification number, 621.3133, is subsumed by other DDC records for various *ranges*

Table 2. Minimal Data Elements for Classification Schedule, Auxiliary Table, and Outline Records

Tag	Tag Name	Sub-fields	Description	DDC Example(s)	LCC Example(s)
cn	classification number	a	The classification number or number range that the rest of the record describes	621.312	TK 2731
cap	caption	a	The caption or heading of the meaning of the last significant digit(s) of the classification number in the cn tag	Generation, modification, and storage (for 621.312)	Types of alternating current machinery (for TK 2731)
chn	central heading number	a	The classification number range that introduces some general remarks about some cn's treated as a group	621.182–621.184	(Implied by LCC schedule at TK 2661–2699)
chc	central heading caption	a	The caption or heading of the meaning of the last significant digit(s) of the central heading in the ch tag	Generation (for 621.182–621.184)	Special direct-current machinery (for TK 2661–2699)
smh	summary heading	a	The field that identifies the record as a summary record	SUMMARY	
smn	summary heading number	a	Each of the classification numbers in a summary record	a 621.1	
		t	Textual portion of a summary record	t Steam engineering	
otl	outline heading	a	The field that identifies the record as DDC Summaries or LCC Outline	SUMMARIES	OUTLINE
otn	outline general number	a	General classification number in outline	a 600	a T
		t	General textual heading corresponding to general number	t Technology (Applied sciences)	t Technology (for T)
		i	Including note		i Including geography and description of individual regions and countries (for D)
		r	Cross-reference note		r For general philosophical treatises *see* B 10–41 (for B)
		s	Scope notes describing the scope, range, or subject areas classed in this number		s For works too general or comprehensive to be classed with any particular subject, however broad (for A)

Table 2. Continued.

Tag	Tag Name	Sub-fields	Description	DDC Example(s)	LCC Example(s)
		n	Other types of notes		
ots	outline specific number	a	Specific classification number in outline	a 621	a TK 2000–2891
		t	Specific textual heading corresponding to specific number in outline	t Applied physics	t Dynamoelectric machinery
		i	Including note		i Including geography and description of individual regions and countries (for D)
		r	Cross-reference notes		r For collections published under the auspices of learned bodies, *see* AS (for AC 1–195)
		s	Scope notes describing the scope, range, or subject areas classed in this number		s For works on the race as a social group and race relations in general (for HT 1501–1595)
		n	Other types of notes		n Arranged by special classes of persons as determined by age, defects, occupation, race, economic status, etc. (for HV 697–4959)
nXX	notes		Comments on the cn, chn, cap, or chc. Individual note fields are covered in table 3	(See table 3)	(See table 3)
sou	source	a	Source of classification records	a Index	a Schedule
		e	edition	e 19	e 5
		c	class		c T
rcs	record status	a	Old, or added, or changed. All initial input is "old"		
		d	Date of last activity		

Table 3. Notes in Classification Schedule and Auxiliary Table Records

Tag	Tag name	DDC Example	LCC Example
nad	add notes	Add to base number 621.322 the numbers following 72 in 725–728, e.g., lighting for libraries 621.32278 (for 621.3225–.3228)	
nce	class elsewhere note	Class engineering of dams for hydroelectric power in 627.8 (for 621.312134)	cf. TK 1141–1168, alternating current engineering (for TK 2711)
nch	class here note	Class here comprehensive works on electrical bridges (bridge circuits) (for 621.3742)	
ncr	cross-reference note	For rectifiers, see 621.3137 (for 621.3135)	Central station testing, *see* TK 1831; generators and motors, *see* TK 2433 . . . (for TK 2316)
nco	citation note	Unless other instructions are given, class complex subjects with aspects in two or more subdivisions of this schedule in the number coming first in the schedule, e.g., aged Chinese women 331.398 (not 331.4 or 331.6251) (for 331.3–331.6)	
ndn	discontinued note	Auxiliary generating stations. Numbers discontinued; class in 621.3121 (for 621.[31214])	
nex	example note	Examples: garden and patio lighting, advertising and display lighting (for 621.3229)	e.g., Conduits, manholes (for TK 3261)
nfa	add as instructed under footnote	Add as instructed under 792.1–792.8 (for 791.1)	For Table I, *see* p.263–265. Add country number in table to 0 (for TJ 21–126)
nfc	footnote for alternate successive Cutter e example subfield t title of alternate successive Cutter		The following subarrangement may be substituted: (for T 170) e A1-4 serial t Official publications
nfo	option footnote	Use is optional; prefer treatment described under 944.012 (for 944[.0121])	
nfs	other types of footnotes		Special countries (for T 20)
nfz	footnote for standard subdivisions zeros	Use extra 0's for standard subdivisions; see instructions at beginning of Table 1 (for 332.45609)	
nin	including note	Including thermionic converters (for 621.31243)	Including switchboard meters, switches, rheostats, etc. (for TK 2821)
nit	introduction to tables note	The following notations are never used alone, but may be used . . . (for Table 7, Persons)	The numbers in this list are intended to be used as a guide . . . (for Table II,T,5th edition)

Table 3. Continued.

Tag	Tag name	DDC Example	LCC Example
nlp	arrange alphabetically note	Arrange alphabetically by title of film (for 791.4372)	By region, A–Z
	e example subfield		e e.g., .N4 New England
nls	arrange alphabetically by Cutter as specified note		Special measures and appliances, A–Z (for T55.3)
	e example subfield		e .H3 Hazardous substances
	c class elsewhere subfield		c cf. HD 7263–7265.5, Industrial Hygiene
nop	optional provisions note	If preferred, give priority in notation to the currency of the jurisdiction requiring local emphasis, e.g., libraries in the United States. Class exchange rate between currencies of United States and United Kingdom in 332.4560973041 (for 332.45609)	Class preferably with special subject in TA-TT (for T 40)
not	other types of notes	Business taxes 336.207 (for 336.201–336.207)	General works only (for T 54)
nph	former heading note	Former heading: Obtaining employment (for 331.128)	
nrl	relocation note	Direct-energy-conversion-powered stations. Class in 621.3124 (for 621.312137–.312139)	
nsa	*see also* note	(New note for edition 20)	*See also* Engineering laboratories, TA416–417 (for TJ148)
nsc	successive Cutter note (with subfields identifying data in note)		Special countries (for T23–31)
	c Cutter number		c .A1A–Z General works
	n instruction subfield		n under T 23–31
	e example subfield		e e.g., T 26.A7 Austria
nsn	scope note	Firing, feeding feed-water treatment (for 621.194)	Combinations of single- and two-phase currents transmitted over single line and used as single and two-phase current at point of application (for TK 3159)
nst	standard subdivisions table note	Notations from Table 1 (for 621.01–.03)	Table 1[1] (for TK 5121–5226)
nsx	standard subdivisions exception note	Do not use; class treatment by governmental level in 336.01; general, historical and geographic treatment in 336.09 (for 336[.009])	
nsz	standard subdivision zeros note	Use 331.63001–331.63009 for standard subdivisions (for 331.63)	

Table 4. Data Elements for Classification Index Records

Tag	Tag Name	DDC Example	LCC Example
ixa	index entry (first indention level)	Technology	Dynamos:
ixb	index entry (second indention level)	machinery (when ixa is Direct-current)	Alternating current (when ixa is Dynamos:)
ixc	index entry (third indention)	arts (when ixa is Diplomas and ixb is prints)	
. . . to ixh (eighth indention level)			
ixn	class number reference	600 (for Technology) 621.3132 (for Direct-current machinery technology and manufacturing)	TK 2411–2491 (for Dynamos) TK 2661–2665 (for Direct-current Dynamos)
ixq	cross-reference from Index entry (without class number listed)	s.a. special applications, e.g., General electrical power (under Machinery, electrical, technology and manufacture)	*See also* Automobiles; Motor vehicles; Railroads and railroading, etc. (under Transportation engineering)
ixr	*see also* reference in Index entry (with class number listed)	s.a. psychology of other special groups, also special applications, e.g., Industrial fatigue studies 658.544 (for 658.544)	
ixs	*see* reference from Index entry (without class number listed)	Dynamos, direct-current, *see* Direct-current machinery	Dynamics, Structural, see Structural dynamics
xxx	class number in cross-reference from *see also* reference	658.544	

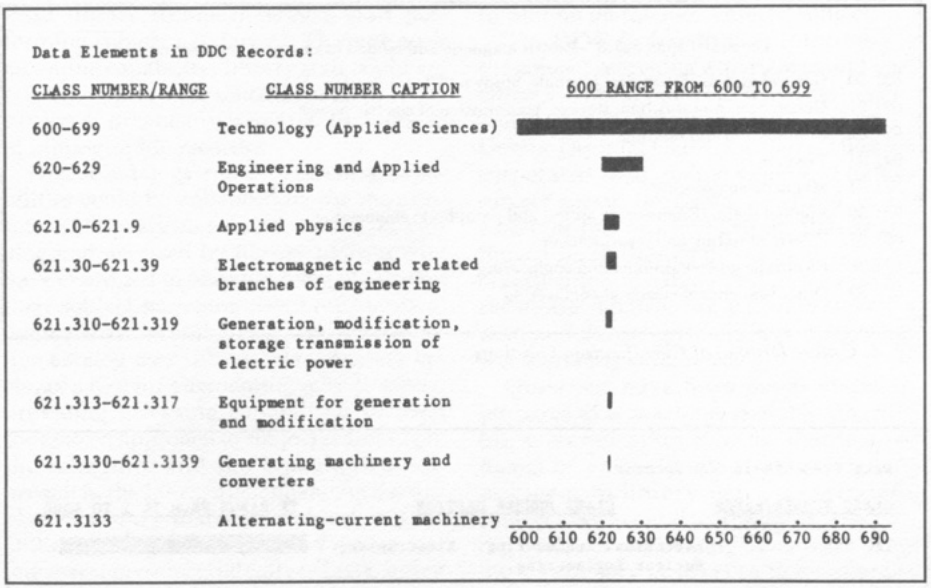

Fig. 3. Nesting of the Record for the DDC Number 621.3133.

of the schedule, e.g., 620–629 or 621.310–621.319.

Nesting of the Dewey Decimal Classification records can be done by automatic procedures by manipulating DDC unit records, i.e., class numbers with a caption and notes, in multiples of ten from zero to nine. In fact, this automatic manipulation has been demonstrated in the CLR-supported DDC Online Project when the experimental online catalog system responds to user-entered keywords or class numbers with a list of ten or fewer captions from the DDC Schedules. For example, a user-entered class number 621.31 results in the online display of classification captions for the range 621.30 to 621.39 shown in figure 4.

Automatic creation of such displays is only possible because the DDC notation usually expresses the broader, coordinate, or narrower relationships among class numbers. Wajenberg[61] suggests encoding hierarchical relationships among class numbers using indicator values in the 082-tagged field (i.e., DDC classification number field) of bibliographic records. Such encoding would only be redundant in bibliographic records and is more appropriate for library classification authority unit records. Although the availability of centered headings and the classification's hierarchical notation could streamline encoding hierarchical relationships, such encoding in the DDC would largely be a manual process.

Figure 5 is an example of nesting for the LCC number TK 2731 if a unit record in the LCC is thought of as *a classification number with a caption and notes* and the explicit ranges noted in the schedule are recorded *typographically* in a conversion of the schedule into machine-readable form. Bars on the bar chart represent the amount of the LCC Schedule linked to that unit record because its classification number, TK 2731, is subsumed by records for various *ranges* of the schedule, e.g., TK 2000–2891 or TK 2711–2745.

Nesting of the Library of Congress Classification records cannot be done by the same automatic procedures used to manipulate DDC unit records, since LCC notation does not express relationships between class numbers. Encoding the hierarchy of LCC unit records is a *requirement* in the conversion of printed *LCC* Schedules to machine-readable form. The presence of indentions in these Schedules, typing conventions such as underlining and uppercase characters, and centered headings help to

General Topic: 621.3 Electromagnetic and related branches of engineering

621.30 General aspects. Examples: Journals, Study and Teaching.
621.31 Generation, modification, storage, transmission of electric power
621.32 Illumination and lighting
621.33 Traction
621.34 Magnetic engineering
621.36 Applied optics (Engineering optics) and paraphotic engineering
621.37 Electrical testing and measurement
621.38 Electronic and communication engineering
621.39 Other branches of electrical engineering

Fig. 4. Online Display of Classification Captions.

Data Elements in LCC Records:

CLASS NUMBER/RANGE	CLASS NUMBER CAPTION	TK RANGE FROM TK 1 TO 4000
TK	Electrical Engineering. Electronics. Nuclear Engineering.	
TK 2000-2891	Dynamoelectric machinery and auxiliaries.	
TK 2181-2799	Dynamoelectric machinery.	
TK 2711-2799	Alternating-current machinery.	
TK 2711-2745	General works, testing, types, etc.	
TK 2731	Types of alternating-current machinery, e.g., synchronous, polyphase, etc.	

1 1000 2000 3000 4000

Fig. 5. Nesting of the LCC Record for the LC Class Number TK 2731.

interpret the ranges to the human classifier, but not to the machine. The hierarchy-encoding process in LCC has to be a manual task and must precede the production of machine-readable LCC unit records. This task must be completed when the LCC Schedules are entered into machine-readable form for the first time.

CONCLUSION

For purposes of constructing files of library classification records, it is necessary to consider the relationship of a class number to the various operations, uses, users, and products associated with it. This paper has enumerated data elements for library classification schedules, and auxiliary tables where a class number in a classification is conceived of as a nested record made up of a number related to other numbers, captions, and notes. These records are related to index information records.

Machine-readable files of library classifications can benefit the editors and publishers of the classification and libraries' staff and patrons. The editors and publishers of library classifications could maintain, edit, update, produce, and distribute the classification in a variety of forms including print, micrographic, and online on a timely basis. Library service organizations such as OCLC, RLG, and WLN could provide online searching and display of the latest versions of library classifications to

assist library technical services staff performing subject cataloging. Through such an online capability, library staff could access both up-to-date schedules and the classification practice represented by millions of bibliographic records.

Online catalogs' subject searching capabilities could be enhanced by the incorporation of library classifications into the catalog and accessed by library patrons who have expressed in online catalog use studies their subject searching needs for improvement of existing online catalogs. For the online catalog user, library classification becomes a tool for augmenting subject access, providing browsing capabilities through the classed approach to subject searching in the schedules and the alphabetical approach in the index, and enhancing the display of library materials' subject matter. Online catalog use studies have shown that few searchers select the call or class number option to search for subjects. The major drawback to the call or class number option of existing online catalogs is that searchers are required to know the exact class number that represents their subject of interest. The incorporation of a machine-readable library classification into an online catalog will relieve searchers' burden of knowing the exact class number and its corresponding subject and enable them to enter their subject in English-language words and phrases to find out the "meaning" of a classification number area and/or the number and shelf location of books containing information on the user-entered subject.

The DDC Online Project is focusing on the subject searching experiences and preferences of library patrons when performing searches in an online catalog in which the Dewey Decimal Classification has been incorporated as an online searcher's tool for subject access, browsing, and display. The results of this project have the potential to impact on the design philosophy of library classification schedule- and index-making, the future of patron use of library classification, and the design of online catalogs' subject searching capabilities.

Questions have been posed about the purposes of a machine-readable file of library classification records, the data elements in a library classification that are pertinent to library patrons performing subject searches and to library staff classifying books, or performing other tasks. Finding satisfactory answers to these questions early in the development of this new tool can ensure that the machine-readable version of library classifications will serve a variety of users and result in a multitude of useful products in various physical formats.

ACKNOWLEDGMENTS

John Finni of Inforonics and Julianne Beall, Winton E. Matthews, and Gregory R. New of the Decimal Classification Division of the Library of Congress are acknowledged for their groundwork in establishing field and subfield coding for a library classification.

REFERENCES

1. Jesse Shera and Margaret Elizabeth Egan, *The Classified Catalog: Basic Principles and Practice* (Chicago: American Library Assn., 1956).
2. James G. Williams; Martha L. Manheimer; and Jay E. Daily, *Classified Library of Congress Subject Headings*, 2d ed. (New York: Dekker, 1982).
3. Robert R. Freeman and Pauline Atherton, *Final Report of the Research Project for the Evaluation of the UDC as the Indexing Language for a Mechanized Reference Retrieval System* (New York: American Institute of Physics, Report AIP/UDC-9, 1968).
4. Robert R. Freeman and Pauline Atherton, "File Organization and Search Strategy Using the Universal Decimal Classification in Mechanized Reference Retrieval Systems," in *Mechanized Information Storage, Retrieval, and Dissemination: Proceedings of the FID/IFIP Joint Conference*, ed. Kjell Samuelson (Amsterdam: North-Holland Publishing Co., 1968), p.122–52.
5. Robert R. Freeman and Pauline Atherton, "AUDACIOUS—an Experiment with an On-line, Interactive Reference Retrieval System Using the Universal Decimal Classification as the Index Language in the Field of Nuclear Science," in *Information Transfer: Proceedings of the 31st ASIS Annual Meeting*, V.5. (New York: Greenwood Publishing Co., 1968), p.193–99.

6. Malcolm Rigby, *Automation and the UDC, 1948–1980*, 2d ed., FID Publication 565 (The Hague: International Federation for Documentation, 1981), p.2.
7. Robert R. Freeman, "Computers and Classification Systems," *Journal of Documentation* 20, no.3: 137–45 (1964).
8. Malcolm Rigby, *Computers and the UDC: A Decade of Progress 1963–73*, FID Publication 523 (The Hague: International Federation for Documentation, 1974).
9. Rigby, *Automation and the UDC*.
10. Ibid., p.3
11. Rigby, *Computers and the UDC*, p.6.
12. Freeman and Atherton, *Final Report*.
13. Freeman and Atherton, "File Organization."
14. Freeman and Atherton, "AUDACIOUS."
15. Freeman and Atherton, *Final Report*.
16. Rigby, *Computers and the UDC*.
17. Rigby, *Automation and the UDC*.
18. Pauline Atherton and Judith Tessier, "Teaching with MARC Tapes," *Journal of Library Automation* 3:24–35 (1970).
19. Nancy B. Olson, *The Combined Indexes to the Library of Congress Classification Schedules* (Washington, D.C.: United States Historical Documents Institute, 1974).
20. Mark S. Fox and Andrew J. Palay, "Machine-assisted Browsing for the Naive User," in *Public Access to Library Automation*, ed. J. L. Divilbiss, Clinic on Library Applications of Data Processing (Urbana, Ill.: Graduate School of Library and Information Science, Univ. of Illinois at Urbana-Champaign, 1981), p.77–97.
21. Valerie Geller and Michael Lesk, "An Online Library Catalog Offering Menu and Keyword User Interfaces," in *National Online Meeting Proceedings—1983*, comps. Martha E. Williams and Thomas H. Hogan (Medford, N.J.: Learned Information, 1983), p.159–65.
22. Phyllis A. Richmond, "General Advantages and Disadvantages of Using the Library of Congress Classification," in *The Use of the Library of Congress Classification; Proceedings of the Institute on the Use of the Library of Congress Classification*, eds. Richard H. Schimmelpfeng and C. Donald Cook (Chicago, Ill.: American Library Assn., 1968) p.219.
23. Phyllis A. Richmond, "A Reconsideration of Enumerative Classification for Current Information Needs," *Ciencia da Informacao* 3, no.1: 10 (1974).
24. Pauline Atherton, *Books Are for Use: Final Report of the Subject Access Project to the Council on Library Resources* (Syracuse, N.Y.: Syracuse Univ., School of Information Studies, 1978).
25. Ibid.
26. Pauline Atherton, "Catalog Users' Access from the Researcher's Point of View: Past and Present Research Which Could Affect Library Catalog Design," in *Closing the Catalog: Proceedings of the 1978 and 1979 Library and Information Technology Association Institutes*, eds. D. Kaye Gapen and Bonnie Juergens (Phoenix, Ariz.: Oryx Pr., 1980), p.109.
27. Elaine Svenonius, "Directions for Research in Indexing, Classification, and Cataloging," *Library Resources & Technical Services* 25: 88–103 (Jan./Mar. 1981).
28. Nancy J. Williamson, "Is There a Catalog in Your Future? Access to Information in the Year 2006," *Library Resources & Technical Services* 26:122–35 (Apr./June 1982).
29. Carol A. Mandel, *Subject Access in the Online Catalog* (Washington, D.C.: Council on Library Resources, 1981), p.16.
30. Richmond, "A Reconsideration."
31. Michael Gorman, "The Longer the Number, the Shorter the Spine; or, Up and Down with Melvil and Elsie," *American Libraries* 12:498–99 (1981).
32. William H. Mischo, *A Subject Retrieval Function for the Online Union Catalog*, OCLC Technical Report Series, no. OCLC/DD/TR-81/4 (Dublin, Ohio: OCLC, 1981). (ERIC: ED 212 263)
33. Charles R. Hildreth, "Online Browsing Support Capabilities," in *Information Interaction: Proceedings of the 45th ASIS Annual Meeting*, V.19, eds. Anthony E. Petrarca, Celianna I. Taylor, and Robert S. Kohn (White Plains, N.Y.: Knowledge Industry Publications, 1982), p.127–32.
34. Janet Swan Hill, "Online Classification Number Access: Some Practical Considerations," *Journal of Academic Librarianship* 10: 17–22 (1984).
35. University of California, Division of Library Automation and Library Research and Analysis Group, *Users Look at Online Catalogs: Results of a National Survey of Users and Non-users of Online Public Access Catalogs; Final Report to the Council on Library Resources* (Berkeley, Calif: The University, 1982).
36. Karen Markey, *Online Catalog Use: Results of Surveys and Focus Group Interviews in Several Libraries*, V.2 of the *Final Report to the Council on Library Resources*, OCLC Research Report Series, no. OCLC/OPR/RR-83/3 (Dublin, Ohio: OCLC, 1983).

37. Karen Markey, *Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs* (Dublin, Ohio: OCLC, 1984).
38. University of California, *Users Look at Online Catalogs*, p.107.
39. Markey, *Online Catalog Use*, p.141.
40. Pauline A. Cochrane and Monika Kirtland, *Critical Views of LCSH—the Library of Congress Subject Headings: A Bibliographic and Bibliometric Essay;* and *An Analysis of Vocabulary Control in the Library of Congress List of Subject Headings (LCSH)* (Syracuse, N.Y.: ERIC Clearinghouse on Information Resources, 1981).
41. Pauline A. Cochrane, "Classification as a User's Tool in Online Public Access Catalogs," in *Universal Classification: Subject Analysis and Ordering Systems*, proceedings of the 4th International Study Conference on Classification Research, 6th Annual Conference of Gesellschaft für Klassifikation, ed. Ingetraut Dahlberg (Frankfurt: Indeks Verlag, 1982), p.260–67.
42. Mandel, *Subject Access*, p.16.
43. Keith W. Russell, *Subject Access: Report of a Meeting Sponsored by the Council on Library Resources* (Washington, D.C.: Council on Library Resources, 1982), p.70.
44. Ibid., p.72.
45. Benjamin A. Custer, ed., *Dewey Decimal Classification and Relative Index*, V.1., 19th ed., devised by Melvil Dewey (Albany, N.Y.: Forest Pr., 1979).
46. Susan L. Miller, "The Library User Meets LCS," *Illinois Libraries* 64: 31–37 (1982).
47. Markey, *Online Catalog Use*, p.47.
48. Charles R. Hildreth, *Online Public Access Catalogs: The User Interface* (Dublin, Ohio: OCLC, 1982).
49. Joseph R. Matthews, *Public Access to Online Catalogs: A Planning Guide for Managers* (Weston, Conn.: Online, 1982).
50. Markey, *Subject Searching*.
51. Betsy Kelly and others, "Bibliographic Access & Control System," *Information Technology and Libraries* 1: 125–32 (June 1982).
52. Gary L. Horowitz and Howard L. Bleich, "PaperChase: A Computer Program to Search the Medical Literature," *New England Journal of Medicine* 305: 924–30 (1981).
53. Christine L. Borgman, *End User Behavior on the Ohio State University Libraries' Online Catalog: A Computer Monitoring Study* OCLC Research Report Series, no. OCLC/OPR/RR-83/7 (Dublin, Ohio: OCLC, 1983).
54. John Tolle, *Current Utilization of Online Catalogs: Transaction Log Analysis*, V.1 of the *Final Report to the Council on Library Resources*, OCLC Research Report Series, no. OCLC/OPR/RR-83/2 (Dublin, Ohio: OCLC, 1983).
55. Markey, *Subject Searching*, p.77–81.
56. *Subject Authorities: A Guide to Subject Cataloging* (New York: Bowker, 1981).
57. *Classified Library of Congress Subject Headings*, 2d ed., eds. James G. Williams, Martha L. Manheimer, and Jay E. Daily (New York: Dekker, 1982).
58. Freeman and Atherton, *Final Report*.
59. Freeman and Atherton, "File Organization."
60. Atherton and Tessier, "Teaching with MARC Tapes."
61. Arnold Wajenberg, "MARC Coding of DDC for Subject Retrieval," *Information Technology and Libraries* 2: 246–51 (Sept. 1983). ■■

NELINET: A Case Study of Regional Library Network Development

Betsy Kruger

The New England Library Information Network (NELINET) and OCLC began at the same point in time with almost identical objectives: to create a shared bibliographic database based on the use of the newly available machine-readable tapes from the Library of Congress and contributed cataloging from member libraries. OCLC's success should not eclipse the important contributions of other pioneers in library computer applications. This paper follows NELINET's development since its 1964 beginning and uses the events of its history as stepping stones from which to explore important issues in library network development: the early and continuing economic imperative for cooperative efforts in automation; issues of funding, governance, and organizational structure; and the evolving role of regional networks and their relationship to major nationwide networks such as OCLC.

In 1964, one year before Lawrence Buckland published his landmark report on the potential applications of the Library of Congress' MARC tapes,[1] representatives from a group of New England land grant universities met to discuss ways in which cooperative efforts could reduce the economic burden faced by the individual institutions during a period of unprecedented growth and expansion. To pursue this goal, the New England Library Information Network (NELINET) was formed. Among NELINET's earliest objectives was the creation of a regional computerized bibliographic database based on the use of LC MARC tapes and contributed cataloging from member libraries.

About 750 miles west of NELINET's Wellesley, Massachusetts, headquarters, a group of Ohio academic libraries was also exploring MARC tape applications. Today, Online Computerized Library Center, Inc., (OCLC), formerly the Ohio College Library Center, has computer terminals in more than three thousand libraries nationwide and abroad and NELINET serves mainly as a broker of OCLC services in New England. It is tempting to synopsize, as one writer already has,[2] that NELINET and OCLC began at the same point in time with almost identical objectives and "the rest is history," but to do so is to risk dismissing an illuminating example of regional library network development in the United States. OCLC's success should not eclipse the important contributions of other pioneers in library computer applications.

This paper will follow the evolution of NELINET from its beginnings in 1964 to the present day and use the events of its history as stepping stones from which to explore important issues in library network development. The "events" are interesting in themselves, as NELINET's history is certainly noteworthy: its pioneering regional automation efforts and subsequent decision to broker OCLC services in New England; its expansion of non-OCLC services to the region; its separation in 1978 from its parent organization; and its pursuit of a new

Betsy Kruger is a master's degree candidate at the Graduate Library School, University of Chicago.

technological role for itself in the early 1980s. Further, the hows and whys of these events illuminate important issues and trends in library network development: the early and continuing economic imperative for cooperative efforts in automation; issues of funding, governance, and organizational structure; and the evolving role of regional networks and their relationship to major nationwide networks such as OCLC. NELINET is of particular interest for exploring this last issue; unlike most of the regionals, which formed almost entirely to broker OCLC services, NELINET has developed side by side with OCLC—and more often in its shadow—for the last twenty years.

EARLY DEVELOPMENT AND EFFORTS IN AUTOMATION

Institutions of higher education in the United States experienced enormous growth and expansion during the 1960s, nurtured by a financial lifeline from state and federal government sources that is merely a fond memory for such institutions today. As New England state governments financed extensive educational programs in that region, their state universities grew in physical size, student population, and number of graduate programs, "each insisting on departmental excellence rather than sufficiency."[3] This sudden growth applied to the libraries serving these institutions as well: the University of Connecticut not only doubled its student enrollment between 1960 and 1970, it also tripled the holdings of its library.[4] It became evident in New England, as elsewhere, that traditional methods of library processing could not keep pace with the libraries' physical growth.

The early formal discussions among the six New England state universities (Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, and Vermont) on how to coordinate their development to best meet both individual institutional and regional needs coincided with the emerging application of computer technology to library functions. The Library of Congress had developed a machine-readable format for bibliographic data and began, in 1966, to record catalog data on magnetic tapes and, in 1969, to distribute these tapes to interested libraries. By 1966, NELINET had formed, and systems analysis began in each of the member libraries for the purpose of establishing a regional computerized catalog. One year later, a pilot project to create catalog cards by batch processing through a central computing facility was underway. In July of that same year, 1967, OCLC was incorporated as a not-for-profit organization and the design of a central computerized catalog for Ohio academic libraries commenced in earnest. It was no small revolution.

NELINET was originally established as a program of the New England Board of Higher Education (NEBHE), a public interstate agency. NEBHE has operated since 1955 under the terms of a legal document titled "The New England Higher Education Compact," the purpose of which is "to provide acceptable and efficient educational facilities to meet the needs of the New England Region."[5] Each of the six New England states contributes funds to NEBHE, which partially funded NELINET and oversaw all its administrative functions until 1978. The problems inherent in this administrative arrangement and its contributions to NELINET's problems and successes will be discussed later in this paper. Between 1967 and 1970 NELINET received all its funding for research and development, administration, and special projects from the Council on Library Resources (CLR) and the United States Office of Education (USOE).[6]

NELINET contracted with a commercial firm for the design of its proposed bibliographic database. This firm, Inforonics, Inc., (then of Maynard, Massachusetts) had received MARC tapes from LC during the MARC Projects and successfully used them to produce catalog cards and book and spine labels from a computer. Ron Miller, NELINET's director in the early 1970s, noted that as a by-product of these efforts Inforonics also demonstrated the feasibility of "building an ever-enlarging file of bibliographic records coupled with local holdings statements which could become a central union catalog in machine readable form. These records could be used for a seemingly limitless number of pur-

poses. It was a heady time indeed."[7]

As Norman Stevens points out in his essay, "An Historical Perspective on the Concept of Networks,"[8] the notion of using a centralized bibliographic database of MARC records as the basis for all automated library services can be directly traced to the early planning efforts of NELINET and OCLC. Technical processing (cataloging and acquisitions) of English language monographs received initial emphasis in the NELINET system, but the greater vision included shelflists; local author files; union catalogs; circulation and interlibrary loan control; the inclusion of foreign language monographs, serials, and government documents; and, eventually, communication with non-NELINET databases.[9] Fred Kilgour's 1968 vision of the OCLC database was identical to that of NELINET's creators: "The entire system, including shared catalog, bibliographic information retrieval, circulation control, serials control, and technical processing, will be based on one file, thereby achieving a truly comprehensive system."[10]

Librarians are just now beginning to challenge this earlier vision of a "totally integrated library system." Richard De Gennaro has cogently described the trend toward decentralization of automated library services.[11] Calling the goal of a total system a "moving target," he describes the new economic incentives and advances in computer technology that are now encouraging librarians to approach some automated services from the local level. In the 1980s, shared cataloging and ILL are likely to remain the province of the large utilities such as OCLC, while circulation, serials control, and acquisition functions are managed more efficiently on a local or regional level. But for NELINET and OCLC in the early days of library automation, a totally integrated system on a network level seemed the most viable alternative for managing the exorbitant costs of the new technology.

From the beginning, both NELINET and OCLC envisioned online access via remote terminals to the central database. NELINET's system never realized this goal fully, although it achieved moderate success as an offline batch processing system during its operational stage. After an enthusiastic response to its pilot project at the University of New Hampshire, NELINET went operational in April 1970. Member libraries teletyped bibliographic data for books requiring catalog cards and labels to Inforonics in Maynard, Massachusetts. The Inforonics staff ran these tapes against LC tapes and extracted cataloging data for found titles. Catalog cards and book and spine labels were produced and mailed to the library.[12] Titles that were not found remained in the computer's request file for ten weeks during which they were periodically checked against new MARC tapes.[13] (The early offline service offered by OCLC was patterned on a similar technique using punched cards.) Composed as it was solely of MARC records, the early NELINET database lacked the element of member-created bibliographic copy that makes modern networks true "shared cataloging" efforts. Nonetheless, the desirability of the services provided by NELINET at that time is illustrated by the increase in network membership from five to thirty during 1971. By year's end, NELINET had provided 36,000 catalog card sets to its members, and its master file contained more than 160,000 records.[14]

Despite this progress, the database and its services were limited and the network was operating with severe organizational and funding difficulties. By June 1970, only two months after its catalog products system began operating, NELINET's funds ran out and there were no grant proposals being prepared. Research and development halted. NEBHE and five member libraries contributed funds to keep the project going.[15] Ron Miller, NELINET's second director, was hired in 1971, replacing Samuel Goldstein.

Meanwhile, back in Ohio things were going quite well. In mid-1970, when NELINET faced its funding crisis, OCLC began offline production of catalog cards at the rate of nine thousand cards per week for its twenty-four member libraries.[16] Then, in the fall of 1971, sooner than NELINET could ever have accomplished the feat given the organizational and financial dif-

ficulties it faced, OCLC went online with a cost-effective shared cataloging system; fifty member libraries accessed the central database in Columbus via CRT terminals.[17] NELINET turned a hopeful gaze to the west.

De Gennaro describes the mid-1970s as library networking's "era of benevolent monopoly" when "OCLC could afford to tolerate and even encourage its weak and struggling sister networks and participate as a senior partner in studies and plans . . . ."[18] Putting contemporary networking controversies aside for the present discussion, it should be noted that OCLC early envisioned a system that could be duplicated outside of Ohio. "Design of a regional system that could be a node in a national network seems to be the obvious choice, and the decision at OCLC was to proceed along this avenue in hopes of producing a system whose hardware, programs, and data could be duplicated for operation in another region and with minimal development expense."[19]

NELINET began exploring the feasibility of replicating the OCLC system in New England. In 1971, CLR awarded two grants to NELINET: one to support a comprehensive technical and user audit of its current shared cataloging support subsystem and the other to study the transferability of the OCLC system to New England. Representatives from NELINET, OCLC, and the Washington State Library met with Henriette Avram of the Library of Congress to discuss coordinating their efforts in systems development. NELINET drafted a proposal for replicating the Ohio system in the two other regions.[20] (At this point, the Washington State Library withdrew from the discussions and " the rest is history. . . .")

A commercial firm called COMRESS, which had performed a simulation of early OCLC computer configurations, simulated the performance of the Ohio system when loaded with profiles of NELINET members. The simulation demonstrated that, with only minor adjustments, the OCLC design could work quite well in New England. It also showed that the use of the OCLC system by three or four times its present membership was technically feasible.[21]

Despite the technical feasibility of the transfer, the immediate availability of funding for such an undertaking was questionable, and it quickly became evident that OCLC would control the cost and performance of the transferred network through pricing and new release of its system software. For the interim, NELINET signed a three-year agreement for OCLC services. NELINET's director, Ron Miller, expressed the hope that "the library base (would) expand along with matching support to enable the region to support a standalone system after transfer costs are paid."[22] Up until 1974, OCLC restricted its membership to Ohio libraries although services were "unofficially" being provided not only to New England libraries, but to the fourteen members of the Pittsburgh Regional Library Center as well. In January 1974, the OCLC Board of Directors decided to extend services formally to libraries outside of Ohio. That September the NELINET Board, with the consent of its advisors, abandoned its decision to replicate the Ohio system in New England, and its members officially jumped onto the OCLC bandwagon.

NELINET AND OCLC: EARLY LESSONS IN NETWORKING

Both NELINET and OCLC, in considering the use of MARC tapes in the mid-1960s, stood in the forefront of a revolution whose time had come: the need was great and the necessary technology was rapidly evolving. A closer examination of the organization, funding, and technical approaches of NELINET and OCLC at that time should shed more light on their problems and successes in the early days of library automation.

Brett Butler, in describing various factors in library network development, notes that "it is a fine balancing act to judge the time needed to develop the idea of a new service and to balance against that the technology available to support the service when it reaches a growth stage" and while "OCLC stretched the state of the art at the time in the initial plans for online service,

gains in computer technology have in fact favored the OCLC online concept during the near-decade required for the marketing of the system concept."[23] Both OCLC and NELINET envisioned online access to their databases, yet at the time OCLC was able to offer such access, NELINET had not yet moved beyond a teletype tape-oriented system, partly because NELINET and Inforonics had underestimated the amount of time and money required to realize the system's final design fully. NELINET director Ron Miller described the frustrating limitations of a tape-oriented system: "Speaking anthropomorphically, the system has no voice and no eyes at the moment. Conversing with the system is fraught with delay, which any interactive system operating through the mails in a batch processing mode is subject to. It is a slow dream-like state which can cause frustration. It has no eyes because librarians cannot see the file from which their products emerge. Even manual LC proof slip files and their microform derivatives provide this needed reassurance and connection with the past."[23]

The technical and user audit that NELINET conducted in 1971 uncovered several other shortcomings of this early system. The catalog products support system was "woefully underutilized"; librarians felt that the cost of these products was "unacceptably high"; the system had enabled only one library to reduce its staff load, and this had only been by one person; the access points did not include main entry/title, which users felt was a severe limitation; and users found it cumbersome to re-enter the system to correct errors or update holdings.[24]

Both NELINET's technical inadequacies and OCLC's more successful application of the MARC tapes are attributable, at least partly, to the original system designer's level of sensitivity to the libraries' needs. OCLC hired its own technical staff, whereas NELINET contracted with an outside commercial firm for research, development, and operational support. NEBHE and the NELINET Board intended that NELINET would eventually hire its own technical staff to take the project over from Inforonics, but, for the time being, all technical, pricing, and production decisions were made by Inforonics. John P. McDonald of the University of Connecticut libraries suggested that while an outside contractor prevented any one institution's needs from dominating the system's design, it did at times seem "that the project staff and contractor tended to lose sight of local needs."[26] He also alluded to problems in "the area of proprietary rights to programs and other software" since Inforonics, as an outside contractor, also developed systems and services for other clients.

Unlike OCLC, which formed originally to serve libraries in a single state, NELINET was from its beginnings a multistate library network. During its early years, this regionalism was both an asset and a hindrance. On the one hand, it provided a valuable link among the vast educational resources of the several small New England states and could also provide a voice for certain shared geopolitical interests. On the other hand, NELINET was subject to the vicissitudes of diverse geopolitical interests, particularly as it searched for a stable financial base. Ron Miller described the notion of starting a regional network serving libraries in six states as "a curious and paradoxical process in practically every respect"[27] and even expressed doubts as to whether "a commitment to building an interstate regionally administered network was worth the effort."[28] NELINET's regionalism and its parent-child relationship to the New England Board of Higher Education also presented difficulties for its early automation efforts. While OCLC originally had strong ties to the Ohio College Association (OCA), apparently this relationship was overwhelmingly beneficial. OCA was, in 1967, a stable hundred-year-old organization. Kilgour suggests that "perhaps the major factor behind the establishment of OCLC is this long and quiet tradition of effective cooperation among Ohio academic institutions."[29] Furthermore, OCLC's organization as a nonprofit corporate entity no doubt gave it a certain independence not enjoyed by NELINET.

As a child of NEBHE, NELINET lacked a certain autonomy in the management of its own affairs. Eventually this would lead to a formal split between the two groups in

1978, but until that time NEBHE signed all NELINET's contracts, grants, and agreements. It also administered all the network's business activities including the receipt, deposit, and authorization for disbursement of funds; auditing; and the hiring and firing of personnel. NELINET's staff were the board's employees and its director served in a second capacity as NEBHE's director of Library Programs.[30] This dual role divided the director's attention between NELINET and the board's wide range of library programs for all 249 institutions of higher education in New England. While NELINET's executive committee and membership council recommended policy, NEBHE held final veto power. It is likely that these administrative arrangements posed certain problems for NELINET in pursuing a hitherto uncharted course in library automation. A certain lack of management expertise within NELINET during those years has also been suggested. Originally, NELINET was a membership organization whose members did not pay dues and that found itself, two months after its computerized database went operational, without funds and without financial prospects. This suggests that NELINET had not managed to kindle the direct financial support of its members until late in the game.

NELINET's funding difficulties also had roots in the network's regionalism. Because the organization's headquarters were in Massachusetts, financial support from individual states was viewed by some as money being sent out of state. This is an attitude with which several regional networks since have had to contend despite the obvious reality that few, if any, states or individual institutions could shoulder such an economic burden alone. Another funding difficulty NELINET specifically faced was that state funding for library programs in New England had historically favored public libraries; all of NELINET's original members were higher education institutions.[31]

NELINET's early automation efforts fell short of its goal of a regional online shared cataloging system. Political, administrative, and fiscal factors influenced this outcome. With a more autonomous organizational structure and a more stable financial base, perhaps NELINET could have continued its own research and development efforts. However, OCLC's system was already successful, cost-effective, and capable of further expansion. The NELINET Board recognized, with foresight, that its goals could be achieved sooner and at less cost by obtaining OCLC services than by continuing parallel development of a separate system.

THE LAST TEN YEARS

Perhaps because NELINET and OCLC were both pioneers in the early days of library networking, NELINET's development over the last ten years has been marked by recurrent efforts to establish an independent role for itself beyond that of a promoter of its successful sister network's services.

During the first few years after contracting for OCLC services, NELINET devoted most of its time to profiling member libraries for OCLC services, purchasing and installing terminals, and training NELINET and member library staff on the use of the OCLC system through subregional workshops. Public and special libraries began to join the network. As a result of the OCLC connection, NELINET's membership increased steadily: 48 members by 1973,[32] 73 members by 1975,[33] and 119 members by 1980.[34]

During these years, NELINET was also involved in regional non-OCLC activities: the Task Group on Government Documents worked towards publishing an expanded title index to government publications, and the Task Group on Serials compiled an annotated list of union lists of serials in New England.[35] In collaboration with the National Science Foundation, NELINET became involved with the Northeast Academic Science Information Center, a project to promote the use of online bibliographic databases in New England academic libraries. A joint project undertaken with the National Agricultural Library involved the large scale microfilming of agricultural publications held by the six New England state universities.[36] All these projects illustrate NELINET's longstanding commitment to regionalism.

Yet at the same time that NELINET's executive committee decided to abandon its plans to replicate the OCLC system in New England, Ron Miller also announced the network's intention to acquire "a teleprocessor computer for the immediate purpose of enhancing efficiency and stabilizing the cost of the communications network and the added potential for long range benefit by facilitating the establishment of additional network services."[37] The executive committee officially approved this as a major development area for NELINET in December 1974.

NELINET acquired a PDP-11 minicomputer, and in mid-1977 reached a long sought agreement with OCLC to channel all telecommunications between OCLC and its member libraries in New England through the computer at NELINET's Wellesley headquarters. This was a definite first for a regional network. NELINET itself had developed the software, which it felt could eventually help other regionals. The teleprocessor gave NELINET a direct telecommunications link to its member libraries, and it provided the potential for communications between libraries and for access to non-OCLC agencies such as LC and BALLOTS (later RLIN) and online bibliographic retrieval services such as Lockheed, SDC, and BRS.[38] NELINET's William Matthews, in an interview with *Library Journal*, described the new arrangement as "a filtering out of services to their proper levels—national, regional, local. A sifting process is going on and people are beginning to understand it."[39]

The following year, in another move toward independence, NELINET's membership unanimously voted to separate from the New England Board of Higher Education. This action was heralded as another step in NELINET's "evolution . . . toward financial stability through member support and toward the assumption of major responsibilities in the network future of the region and the nation."[40] NELINET was not the first regional to take such action: SOLINET and AMIGOS had already split from their parent organizations.

Matthews' view that a "sifting process" was beginning illustrates the growing tension between secondary networks and OCLC as the former began to assert a measure of independence and the latter began to recognize competition within its ranks. The question of the proper roles for secondary and primary networks persists to the present day. Viewpoints over the years have ranged from that which sees secondary networks as proliferating services to keep themeselves in business, to that which insists that OCLC has become an unreceptive monopoly that must realize it cannot be all things to all libraries.[41]

For the most part, regional networks recognize their dependence on OCLC cataloging services and deny that they are attempting to become competitors. Rather, they insist that some activities are best performed on a regional or statewide basis, or among a group of libraries. NELINET's neighbor to the south, SOLINET, had long been fighting for the right to reuse records that its members contributed to the OCLC database. When SOLINET failed to get foundation funding for its Regional Support System (RSS), a stand-alone regional catalog utilizing WLN software, OCLC withdrew a long-promised loan of $500,000 for the project. OCLC's withdrawal sufficiently squelched the project, and twenty-three SOLINET employees were laid off.[42] OCLC's critics saw this move as an effort by OCLC to stifle what it had come to see as a threatening competitive situation. Rowland Brown of OCLC insisted that without other funding, the RSS project involved more risks than rewards; SOLINET's director, Lee Handley, countered that the RSS "would have improved OCLC and not taken away from it. . . . The regionals have a rightful role in developing projects and services."[43]

NELINET was apparently beset by more political and funding difficulties that delayed its efforts to expand computer services to its members. But in the spring of 1980, the NELINET membership once again approved the board's proposal for a new technological strategy. This statement had been drafted through input from a membership questionnaire. NELINET envisioned its new technological role as assisting its members to meet their own automation goals through the provision of online catalogs and interlibrary loan and circula-

tion control systems. To this goal, it would investigate the purchase of various software systems and stated that it would "work closely with OCLC to develop the appropriate communication system to tie everything together."[44] John Linford, the man hired to direct NELINET during this new technological phase, emphasized that "NELINET is not interested in competing with OCLC . . . the real object is not to create strong networks, but to create strong libraries.[45]

However, less than two years later, NELINET was once again forced to abandon its technological aspirations. Ironically, NELINET was now facing competition similar to that which it and other regional networks posed to OCLC: a decentralization of technology evidenced by the growing number of message-switching, microcomputing, and electronic mail services that were allowing libraries to communicate with each other without the assistance of NELINET as a middleman. Linford said that "NELINET had a strong entrepreneurial style but this spirit was snuffed out by economics," but "we will never again have a consensus among libraries for such OCLC services as interlibrary loan and acquisitions. These services appeal to a smaller universe of libraries than did cataloging."[46] As noted earlier, while control of automated acquisitions is moving toward the local level, interlibrary loan along with shared cataloging will probably remain major services of primary networks for some years to come.

During the past two years, NELINET's activities have centered around regional programs and the brokering of OCLC services in New England. Many of the issues raised on the library networking scene over the last ten years of NELINET's development persist, many more will arise, and with it all, NELINET's role will continue to evolve. Along with other regional networks, NELINET now faces the challenge of reexamining its role to determine how it can best meet the changing needs of its member libraries. The network's future is likely to be as interesting and problematic as its past. NELINET is not merely the network that did not become OCLC; it is one of the many potent forces and organizations that have shaped the contemporary face of library networking in the United States.

REFERENCES

1. Lawrence F. Buckland, *Recording of Library of Congress Bibliographic Data in Machine Form.* (Washington, D.C.: Council on Library Resources, 1965).
2. Norman Stevens, "An Historical Perspective on the Concept of Networks: Some Preliminary Considerations," in *Networks for Networkers: Critical Issues in Cooperative Library Development* (New York: Neal-Schuman, 1980), p.35.
3. William R. Nugent, "NELINET: The New England Library Information Network," in *Proceedings of the International Federation for Information Processing* (Edinburgh, Scotland: International Federation for Information Processing, 1968), p.1315.
4. John P. McDonald, "NELINET: One Approach to Library Automation," in *Collaborative Library Systems Development* (Cambridge, Mass.: MIT Pr., 1971, p.102.
5. "New England Library Information Network," in *Bowker Annual of Library and Book Trade Information, 1975* (New York: Bowker, 1975), p.100.
6. Ron Miller, "The New England Library Information Network: Progress and Prospects," in *Proceedings of the Indiana Seminar on Information Networks* (West Layfayette, Ind.: Purdue University Libraries, 1972), p.20.
7. *Ibid.*
8. Stevens, "An Historical Perspective on the Concept of Networks," p.39.
9. Inforonics, Inc., *Development of a Computerized Processing Center for the New England State University Libraries* (Maynard, Mass.: Inforonics, 1967).
10. Fred Kilgour, "Initial System Design for the Ohio College Library Center: A Case History," in *Proceedings of the Clinic on Library Applications of Data Processing* (Urbana: University of Illinois Graduate School of Library Science, 1969), p.86–87.
11. Richard De Gennaro, "Library Automation and Networking: Perspectives on Three Decades," *Library Journal* 108: 629–35 (Apr. 1, 1983).
12. Russel Shank, "Networks," in *Bowker An-*

nual of Library and Book Trade Information, 1970 (New York: Bowker, 1970), p.293.
13. New England Board of Higher Education, *A Technical and User Audit of the Shared Cataloging Support Subsystem of the New England Library Information Network* (Washington, D.C.: Council on Library Resources, 1973), p.I-8.
14. Marvin C. Gechman, "Machine Readable Bibliographic Databases," in *Annual Review of Information Science and Technology* (White Plains, N.Y.: Knowledge Industry, 1972), p.35.
15. Miller, "The New England Library Information Network," p.20.
16. Dan Clemmer and Russell Shank, "Interlibrary and Information Networks," in *Bowker Annual of Library and Book Trade Information, 1976* (New York: Bowker, 1976), p.301.
17. Richard De Gennaro, "Computers and Regional Networks," in *New Opportunities for Research Libraries* (Association of Research Libraries, 1972), p.63.
18. "From Monopoly to Competition: The Changing Library Network Scene," *Library Journal* 104: 1216 (June 1, 1979).
19. Kilgour, "Initial System Design for the Ohio College Library Center," p.84.
20. Ron Miller, "NELINET: A Regional Network." in *New Opportunities for Research Libraries* (Association of Research Libraries, 1972), p.63.
21. *Ibid.*, p.65.
22. *Ibid.*, p.67.
23. Brett Butler, "The State of the Nation in Networking," in *Bowker Annual of Library and Book Trade Information, 1976* (New York: Bowker, 1976), p.108.
24. Miller, "NELINET: A Regional Library," p.63.
25. NEBHE, "A Technical and User Audit," p.I-7–I-10.
26. McDonald, "NELINET," p.104.
27. Miller, "NELINET: A Regional Network," p.62.
28. Miller, "New England Library Information Network: Progress and Prospects," p.25.
29. Kilgour, "Initial System Design for the Ohio College Library Center," p.79.
30. "New England Library Information Network," in *Bowker* (1975), p.100.
31. Miller, "New England Library Information Network: Progress and Prospects," p.21.
32. "Membership Categories Redefined by NELINET," *Library Journal* 98: 1076 (Apr. 1973).
33. "New England Library Information Network," in *Bowker* (1975), p.102.
34. Susan K. Martin, *Library Networks 1981–1982* (White Plains N.Y.: Knowledge Industries, 1981), p.52.
35. New England Board of Higher Education, *NELINET Director's Annual Report* (1972–1973), (Wellesley, Mass.: NEBHE, 1972–73), p.3–4.
36. "New England Library Information Network." in *Bowker* (1975), p.101.
37. "OCLC Reports More Growth; NELINET to Try New Systems." *Library Journal* 99: 2028 (Sept. 1, 1974).
38. "NELINET Gets Some OCLC Responsibilities," *Library Journal* 102: 1444 (July 1977).
39. *Ibid.*
40. "NELINET Splits from NEBHE," *Library Journal* 103: 1456 (Aug. 1978).
41. Richard De Gennaro, "Will Success Spoil OCLC? *Library Journal* 109: 626 (Apr. 1, 1984).
42. "$$ Shortage Forces SOLINET to Abandon RSS Effort," *Library Journal* 106:1973 (Oct. 15, 1981).
43. *Ibid.* p.1974.
44. "Online Catalogs and Circulation New Step for NELINET," *Library Journal* 115:2370 (Nov. 15, 1980).
45. "NELINET Sees Technology as Next Network Step," *Library Journal* 105:118 (May 15, 1980).
46. "NELINET Abandons Plans to Play Technological Role," *Library Journal* 107:126 (Jan. 15, 1982).

FURTHER READING

Breedin, Brent. "The Council on Library Resources." In *Bowker Annual of Library and Book Trade Information, 1973.* New York: Bowker, 1973.

Clemmer, Dan, and Russel Shank. "Library and Information Networks." In *Bowker Annual of Library and Book Trade Information, 1971.* New York: Bowker, 1971.

"CLR Grant to NELINET." *Wilson Library Bulletin* 43:825 (May 1969).

Curran, Ann T. "Library Networks: Cataloging and Bibliographic Aspects." In *Clinic on Library Applications of Data Processing.* Proceedings of the 1969 Meeting. Urbana: University of Illinois Graduate School of Library Science, 1970.

Ford, Joseph. "Network Service Centers and Their Expanding Role." *Information Technology and Libraries* 1:28–36 (Mar. 1982).

Grove, Lee E. "The Council on Library Resources, Inc." In *Bowker Annual of Library and*

Book Trade Information, 1972. New York: Bowker, 1972.

Inforonics, Inc. *New England Library Information Network: Catalog Data File Creation for the New England Regional Library Technical Processing Center*. Submitted to the New England Board of Higher Education. Final Report. June, 1968. (ED 026 077).

Kilgour, Fred. "Computerized Cooperation." In *New Opportunities for Research Libraries*. Association of Research Libraries, Minutes of the 80th Meeting. Atlanta, 1972.

"NELINET Split Final." *Library Journal* 104:448 (Feb. 15, 1979).

"NELINET Studies Automation for R.I. Library Network." *Library Journal* 105:2462 (Dec. 1, 1980).

"NELINET to Furnish Online Catalogs." *Wilson Library Bulletin* 55:331 (Jan. 1982).

New England Board of Higher Education. *Annual Report*, 1966.

"N.E. Processing Center Gets Project Director." *Library Journal* 93:4088 (Nov. 1, 1968).

"Technology as Priority Endorsed by NELINET." *Library Journal* 105:1352 (June 15, 1980).

"Two Library Consortia to Go Online with OCLC." *Library Journal* 97:3528 (Nov. 1, 1972).

Vincent, Donald E. "New England State University Libraries' Regional Processing Center." *ALA Bulletin* 61:672–73 (June 1967). ■■

Information Technology, Library Management, and OCLC

Mary Jo Lynch

An exploratory study of OCLC's role in providing management information to libraries was conducted by the author as OCLC Visiting Distinguished Scholar in 1983–84. Based on a series of interviews with OCLC staff, network directors, and librarians in member libraries and a review of related literature, it was recommended that OCLC take several steps to improve the statistics it currently makes available and to seriously consider a leadership role in helping libraries use microcomputer technology to provide statistics for library management.

INTRODUCTION

In August 1982, the author submitted to OCLC's Office of Research a document entitled "Management Information Systems for Libraries through OCLC: A Proposal for Research and Development." The author was subsequently invited to OCLC as a visiting scholar and conducted a study during three nonconsecutive months between May 1983 and May 1984. This article discusses why the project was initiated, how the focus changed, and what was finally recommended to OCLC. It includes a review of the small body of literature available on the use of computer technology to assist library management.

The proposal presented to OCLC in 1982 was an outgrowth of the increasing concern, in several parts of the library community, for the "statistics problems" that can be described briefly as follows: data about libraries on the national level are collected by a number of different agencies; some needed data are not collected at all; the data collected are often incompatible and are frequently hard to find or unavailable in time to be of use to managers. In "Towards the Development of a Library Management Information System," Robert Runyon proposed a partial solution to the problem: one of the networks should collect and disseminate statistics that would assist the library manager.[1] Runyon's idea served as the basis for the author's exploratory study at OCLC.

PROJECT GOALS AND METHODOLOGY

Although often unrecognized, there are major differences between the statistics needed for planning and monitoring library service at the state and national level and the statistics needed for local library management. Such factors as what data should be collected, how often they should be gathered, and how they should be organized change as one moves from a consideration of statistics needed for external reporting and statistics useful for internal management. There are problems to be solved regarding both kinds of statistics in libraries, but it appeared that OCLC could most appropriately help libraries with statistics for internal management.

The original goal of the project was to design a model management information system that OCLC could offer as a service

Mary Jo Lynch is director, ALA Office for Research.

to its members and then to test the system in one type of library. Data on a selected number of management topics would be collected by participants, organized by OCLC, and sent back to the libraries, thereby enabling them to compare themselves with similar institutions.

After spending several weeks talking to OCLC staff and members, however, that idea seemed premature. For one thing, OCLC already made some management information available, but no one person knew exactly what was distributed or how it was used by networks and member libraries; those matters needed investigation. Secondly, there was little evidence that provision of additional management information should be a priority for OCLC in view of the many other matters needing attention. For those two reasons, the focus of the project shifted from development of a management information service to exploration of the information currently available and how it was used.

Everyone on the OCLC staff who had any responsibility for disseminating statistics was interviewed in order to produce a list of what was currently distributed. Almost all network directors were interviewed about how they used the OCLC data, which data they sent to member libraries, and how data were used by members. Librarians in over two dozen libraries were interviewed following recommendations by network directors. The final recommendations to OCLC were based on information gained through these interviews and on a review of relevant literature.

LITERATURE REVIEW

In order to enrich the OCLC project and link it to previous work in the field, a literature search was conducted to discover materials about management information systems in general and recent writings describing the use of such systems in libraries. "Management information system(s)" is the phrase frequently used to describe the collection and manipulation of statistics by means of computer technology for the purpose of managerial decision making. A related phrase, "decision support systems," is also used, sometimes almost as a synonym.

There is a large and growing body of literature about management information systems in various public and private enterprises. Although the phrase is used to describe a variety of things, it usually means that an automated system is developed for the purpose of supplying management with information for making decisions. Judging from the literature, librarians seem to be much more interested in using computers to solve operational problems. The generation of management information is almost always a secondary consideration. This may be entirely appropriate given the nature of the library enterprise, but it does raise questions about the meaning of the phrase "management information system" in the library context.

F. W. Lancaster avoided it in the title and introduction to the proceedings of the 1982 Annual Clinic on Library Applications of Data Processing published late in 1983 under the title *Library Automation as a Source of Management Information*. Lancaster notes in the introduction to this work that

> in the last twenty years, two of the most significant developments in the library field have been the increase in the adoption of automated procedures and growing interest in the measurement and evaluation of library services. Yet, the marriage of these two trends—that is, the use of automated systems as sources of data to permit improved management and decision-making—has not been a major factor of professional interest.[2]

A review of the literature in "Management Information for Library Decision Making," written in 1982 by Charles McClure for volume 13 of *Advances in Librarianship*, supports Lancaster's statement.[3] In addition, Michael Bommer and Ronald Chorba's 1982 publication *Decision Making for Library Management*,[4] which explores the feasibility of a data-based decision support system using computer technology, cites almost no previous or current work to make such a system operational. This is not surprising since the basic textbooks on library management rarely mention using statistics in that activity. The reason for this omission may be that until the advent of computers, collecting and organizing data were more trouble than they were worth. Then, once com-

puters became available, librarians were busy using them for other functions. Information for library management was not perceived as a vital need.

In a very early article on the use of automation to produce management information for libraries, Maidment expresses surprise that so little attention has been given to questions of management and policy-making information in library applications of automatic data processing in contrast to the attention paid to these matters in the commercial world. Maidment explains this by noting that "it is probably rare for the librarian's problem of access to information via different characteristics to appear in an acute form . . . major policy decisions are scarcely daily occurrences and, in any case, require analysis of longer term trends."[5] Morris Hamburg notes a related characteristic of libraries when he explains why it was difficult to develop a model management information system for them. "In the private sector of society, the profit-making objective is quite clearly defined. There a natural framework exists for a management information system . . . no similar natural framework exists for libraries."[6]

Despite the scarcity of operational management information systems or decision support systems in libraries, there *is* occasional use of those terms in books and articles about library management. Literature using those phrases seems to fall into three somewhat overlapping categories: material that talks about management information as a by-product of the automation of library processes, theoretical material that discusses decisions made in libraries and how data might be collected and used to support the decision maker, and material that talks about the collection and use of statistics in library management. There are also many articles in the last two categories that do not mention management information systems or decision support systems or computers. Those articles will not be reviewed here because the focus is on literature that discusses management information and computers.

The first major mention of management information systems of libraries appears in the 1974 monograph *Library Planning and Decision Making Systems*, by Morris Hamburg and others. This work begins with chapters on library objectives, a framework for library decision making, and the description of existing models of library processes before describing what should be included in a management information system for libraries. Hamburg predicted that statistics would become more important in library management as a result of computerization: "One more significant trend for the future development of libraries is the increasing awareness of the relevance and usefulness of library statistics for an effective management information system. To a large extent these statistics will be offshoots of the computerization of library functions."[7] Hamburg then described how management decisions could be made on the basis of data collected from particular automated processes.

Also published in 1974 was an article by Robert Burns presenting "An Empirical Rationale for the Accumulation of Statistical Information."[8] Burns was involved in the Colorado State University Library's change from a manual circulation system to an automated one and needed to specify what management information should be provided by the new system. That necessity caused him to articulate a number of concepts and guiding principles regarding the development of what he called management information systems in libraries.

In 1976 the paper "A Management Information System in a Library Environment"[9] was presented by Sutton and Black at the Seventh Ontario Universities Computing Conference. This paper describes a system planned at the University of Guelph to collect data both manually and through computerized systems for use in the decision-making process. The system was never fully implemented because other events intervened, but the staff at Guelph is still working along these lines.[10]

Although Burns and Sutton and Black seemed to have a good grasp of the potential for management information statistics in libraries, evidently such thinking was not widespread. Dranov's 1977–78 study of automated circulation systems notes that "in theory, statistics garnered through the system can be used to guide libraries in acquisitions, in allocating resources among

the various branches, and in other budgeting considerations. In practice, however,. . . the reports and statistics capabilities of automated systems have yet to have an impact on these library operations."[11] (It seems likely that a survey in 1985 would show more activity in this area but the literature does not yet reflect a major change.)

Despite the reality described by Dranov, library leaders continued to advocate the use of data generated by automation in library management. Teague wrote an excellent overview of the possibilities in the *Canadian Library Journal* for October, 1979.[12] D. J. Milne, another Canadian, wrote on library management information systems from an entirely different perspective in 1980. Instead of starting with automation and showing how it can produce management statistics, Milne started with the statistics needed in managing libraries and described how automation can help provide them.[13] Also in 1980, Adelaide A. Del Frate described a study done at the NASA Goddard Space Flight Center using two automated databases—cataloging and circulation. The two were "mutually profiled to reveal the changing degree of coincidence between holdings and use." The analysis proved so interesting that "it accelerated the decision to store use history in the machine-readable shelflist record as a management information system data element."[14]

Runyon's 1981 article in *College & Research Libraries* has already been mentioned. It suggested that libraries could and should work together to assemble management information. Runyon described a basic set of data needed by library administrators and proposed that a network or bibliographic utility offer a service that would help libraries gather and organize these data. His proposal was based on "standardized terminology, machine-aided data collection, and customized computer processing and reporting as well as systematic training and documentation."[15] The only previous suggestion that networks could or should be involved in providing management information was made by Lemke in 1978.[16] Lemke noted in particular OCLC's "TOTARC" (which provided monthly reports on catalog production, holding updates, original input, and reclassification projects) and the use of OCLC/MARC tapes in the collection development project conducted by Glyn Evans at SUNY/OCLC.

A different approach to the subject was presented in Michael Bommer and Ronald Chorba's 1982 book, *Decision Making for Library Management*.[17] Bommer had been associated with Morris Hamburg in the work cited earlier. The 1982 monograph describes the principles of a decision support system and then explains how such a system might work in an academic or special library. This system would use data generated by automated processes in the library as well as data available from other automated sources in the parent institution. Bommer and Chorba cite only one piece of previous work integrating library records with records from the parent institution—work done by Glyn Evans in the SUNY/OCLC network. Among other things, Evans used machine-readable data on course loads by subject in connection with library acquisition records to examine the extent to which a library was purchasing heavily in areas where courses were being offered.[18]

The most recent literature on this topic is the volume of papers presented at the annual data processing clinic of the Graduate School of Library and Information Science at the University of Illinois held in April 1982.[19] Included are theoretical articles on management information systems, descriptions of how particular libraries use data resulting from automation, and warnings about how the use of such data may be resisted or how it can be misused.

Three articles stand out as being of particular practical value in that they describe operational systems. "Beyond the Numbers—a Decision Support System" by Kenneth Dowlin and Lynn McGrath describes the system currently in place at Pike's Peak Library District in Colorado.[20] "Computer-Derived Management Information in a Special Library" by Robert A. Kennedy describes what is done in the library system of Bell Laboratories.[21] A paper by Glyn Evans and Albert Beilby describing "A Library Management Information System in a Multi-campus En-

vironment"[22] provides information hitherto available primarily in government reports about the information SUNY/OCLC can provide to requesting member libraries regarding their collections in relationship to the academic program of the campus.

Finally, there is work going on in this area that has not yet been described in the literature. There are two libraries where management information systems are being developed in the classic way, i.e., analysis of decisions made, determination of where and how to gather useful data, design of automated systems to collect data.

At Cornell, J. Gormly Miller, formerly director of the Cornell University Library, is working with Shirley F. Harper, library director at the School of Industrial and Labor Relations, to develop a statistical reporting system for that library that could eventually be used for all Cornell libraries. They are now collecting and analyzing data on a microcomputer using Lotus 1-2-3; in the future they expect to use dBase III also.

Miller and Harper began by studying data reported by each unit of the Industrial and Labor Relations Library and used the *Library Data Collection Handbook*[23] as a framework for the system and as a source for definitions. As of March 1985 they had developed a "Resource Distribution Database" with three subfiles (circulation, interlibrary loan, copy service), a "Reference and Information Services Database," and a "Cataloging Services Database." Structure and formulas for an "Acquisitions and Collection Development Database" were almost complete. Data are collected manually for the system, which provides basic statistical reports (e.g., sums and averages). Harper and Miller would like to add financial information to the system and to develop cost models. They recently sought grant funds to accomplish those tasks but have not yet been successful in attracting support. Meanwhile work continues on the system. Eventually the developers hope to be able to market it to other libraries.

Jack Benford, head of the Office of Planning, Evaluation, and Development at the Free Library of Philadelphia, is directing a project established in March of 1983 to develop a management information system for that library. LSCA funds were used for the first fifteen months of the project, which builds upon the planning report prepared for the library by Lowell Martin Benford and his team, which includes Dr. Anthony Di Primio, an expert on management information systems from the faculty at Temple University, began with a careful examination of decision-making activities in the library. They studied organization charts and questionnaires returned by unit heads reporting on what information they needed and how they obtained it. Interviews were then conducted with these managers to extend and clarify responses to the questionnaire. Next the team created a list of kinds of information needed, where it is found, and how it is used. Then they designed the subsystems using DBase II and Lotus 1-2-3 on a personal computer.

As of March 1985 they had developed and implemented a "Library Services and Resources" subsystem for use by the forty-eight branches of the Free Library. Basic demographic data from the U.S. Census tapes are on the system; branches are encouraged to use it in connection with library data on such topics as registration, circulation, turnover rates, and other measures of library output. Benford and his team have suggested methods of analysis for all data.

Also under development at the Free Library are a "Processing Operation" subsystem, which will be based on information available from the CLSI circulation system; and "Administrative Services" subsystem, which will organize data on shipping and supply operations and on building and vehicle maintenance; and a "Human Resources" subsystem, which will utilize information from the city personnel system and a "Financial Control" subsystem. Since Philadelphia is currently developing a citywide automated payroll system, this part of the library project will not be complete until the more comprehensive work is done.

Whether they use the phrase "management information system" or not, it seems likely that librarians will use statistical information resulting from automation in the management of their institutions. In addition to the literature, several trends point in that direction. Library educators fre-

quently write and talk about the need to help students become more numerate, and the number of brochures this author has received about workshops or seminars on statistics has increased considerably in recent years. Automation is being used in more and more libraries and for more and more purposes. Leaders in the field are taking advantage of automation to generate information for management and writing about their experiences.[24] Finally, librarians are learning how to use microcomputers to organize data so that it becomes information useful for decision making. New items in the library press often allude to such activity, and conversations at conferences frequently concern it.

This probability of increased use of statistics produced through automation for the purposes of library management was recognized by at least two of the eight librarians who described the future in the June 15, 1984, issue of *Library Journal*, which focused on information technology. Richard Boss noted that "acquisition decisions will be based on extensive management information about library users and patterns of collection use."[25] Maurice Freedman also observed that the "computer's capacity to supply management data is phenomenal, and that will provide better support for and thus enhance the professional's judgment."[26]

RECOMMENDATIONS

Both the literature search just described and the interviews noted earlier shaped this visiting scholar's recommendations to OCLC. The final report made suggestions about three different matters. Regarding statistics already provided to networks, the report suggested a change in the way these data are perceived. Instead of viewing them as statistics to be used primarily for managing OCLC, the statistics should be viewed from the perspective of network management and individual library management. If currently produced statistics were viewed in this way, several things should happen that would benefit all concerned. OCLC might take action to make the statistics more usable by networks and libraries (e.g., by providing more meaningful labels for columns and rows, by supplying clearer explanations of how to interpret the tables). Also, OCLC might promote the use of its statistics by networks and encourage networks to do the same with member libraries. To achieve those ends, it was recommended that statistics currently produced by OCLC be carefully reviewed from this new perspective by a task force of members and staff. The author's interview notes, included as an attachment to the report, could be made available to such a task force to stimulate thinking.

A second suggestion related to existing statistics concerns educational programs. OCLC has already had considerable success in mounting programs to help people use the systems it offers. A similar effort should be made for management statistics. Both network staff and member libraries could profit from written materials and interactive workshops designed to teach them how to use OCLC statistics in making management decisions. The author's interview notes describe how OCLC data are currently used in some networks and libraries. Those notes, supplemented by additional contact with some of the persons interviewed, could be used to design educational programs.

In addition to improving statistics currently available, the author recommended that OCLC consider offering other products based on data currently collected. Two kinds of products were suggested: special reports based on existing files (e.g., Marketing Trends Reports data arranged by state within a multistate network) and software to enable a library to analyze its own tapes.

Finally, suggestions were made about other statistics OCLC might collect and disseminate. Although a few people have expressed great interest in having OCLC collect and distribute statistics in addition to those generated through use of the OCLC subsystems, there is not enough demand for this service to ensure that it would be a viable new product at this time. However, several factors exist that make it likely that this situation could change in the near future: libraries are becoming more complex institutions to manage, librarians are becoming more knowledgeable about how to use statistics in management, and microcomputers are making it possible for statis-

tics to be collected and manipulated with relative ease. Therefore, it was recommended that OCLC take a leadership role in helping libraries use microcomputer technology to provide statistics for library management.

That last suggestion was supported by one of the attachments to the final report entitled "Statistical Data for Library Management: Proposal for an OCLC Pilot Project." This outline, developed by a small advisory group convened at OCLC to discuss possibilities in this area,* proposes the development of a package based on microcomputer technology that will assist librarians in capturing, manipulating, and using statistical data about library operations and the testing of this package in one segment of the library market for the purpose of determining the feasibility of OCLC's offering such a service.

*Involved in developing that outline were several interested experts on the OCLC staff and the following advisers from outside OCLC: Philip Clark, Willis Hubbard, Robert Mason, Nolan Pope, and Robert Runyon.

CONCLUSIONS

One of the original goals of OCLC was to "provide management information." Reasons why this has not yet been a priority for OCLC were well explained in a paper presented at the 1982 Illinois data processing clinic by Jacob and Kaske:

> At OCLC much of our early work focused on acquiring only those pieces of data which we actually needed to run and support the operation—in particular, the information required for billing. The design of our systems from a management support perspective leaves much to be desired.[27]

OCLC's interest in this author's study is an indication that OCLC may pay more attention to management information in the future. The several suggestions made in the final report have been reviewed by the OCLC managers responsible for making changes in the appropriate areas. Some changes have been made already; some will be made as resources become available; others will have to wait until more pressing matters are settled.

REFERENCES

1. Robert S. Runyon, "Towards the Development of a Library Management Information System," *College & Research Libraries* 42:539–48 (Nov. 1981).
2. F. Wilfrid Lancaster, ed., Introduction to *Library Automation as a Source of Management Information;* papers presented at the 1982 Clinic on Library Applications of Data Processing, April 25–28, 1982 (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983), p.1.
3. Charles McClure, "Management Information for Library Decision Making," in *Advances in Librarianship, Volume 13*, Wesley Simon, ed. (New York: Academic), p.1–47.
4. Michael R. W. Bommer and Ronald W. Chorba, *Decision Making for Library Management* (White Plains, N.Y.: Knowledge Industry Publications, 1982).
5. William R. Maidment, "Management Information from Housekeeping Routines," *Journal of Documentation* 27:37 (Mar. 1971).
6. Morris Hamburg, "Statistical Methods for Library Management," in *Quantitive Measurement and Dynamic Library Service*, Ching-chih Chen, ed. (Phoenix: Oryx Pr., 1978), p.32.
7. Morris Hamburg and others, *Library Planning and Decision-Making System* (Cambridge, Mass.: MIT Pr., 1974), p.211.
8. Robert W. Burns, Jr., "An Empirical Rationale for the Accumulation of Statistical Information," *Library Resources & Technical Services* 18:253–58 (Summer 1974).
9. Michael J. Sutton and John J. Black, "A Management Information System in a Library Environment," paper presented at the Ontario Universities Computing Conference (Waterloo, Ontario, Canada, June 2–4, 1976) ED 127 915.
10. Telephone conversation with John Black, associate librarian, University of Guelph, August 1, 1984.
11. Paula Dranov, *Automated Library Circulation Systems, 1977–78* (White Plains, N.Y.: Knowledge Industry Publications, 1977), p.75.
12. Jean Teague, "Computer Potential for Management Information," *Canadian Library Journal* 36:268–70 (Oct. 1979).
13. Dorothy J. Milne, "Automating the Management Information Systems of Libraries," *Canadian Journal of Information Science* 5:159–67 (1980).
14. Adelaide A. Del Frate, "Use Statistics: A

Planetary View," *Library Acquisitions: Practice and Theory* 4:248 (1980).
15. Runyon, "Towards the Development of a Library Management Information System," p.539.
16. Darrell H. Lemke, "Library Networks and the Provision of Management Information," *Catholic Library World* 49:376–79 (Apr. 1978).
17. Bommer and Chorba, *Decision Making for Library Management.*
18. Glyn T. Evans and Albert Beilby, "A Library Management Information System in a Multi-campus Environment," in *Library Automation as a Source of Management Information*, F. Wilfrid Lancaster, ed. (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983) p.164–96.
19. F. Wilfrid Lancaster, ed., *Library Automation as a Source of Management Information;* papers presented at the 1982 Clinic on Library Applications of Data Processing, April 25–28, 1982 (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983).
20. Ken Dowlin and Lynn Magrath, "Beyond the Numbers—a Decision Support System," in *Library Automation as a Source of Management Information* F. Wilfrid Lancaster, ed. (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983), p.27–58.
21. Robert A. Kennedy, "Computer-Derived Management Information in a Special Library," in *Library Automation as a Source of Management Information*, F. Wilfrid Lancaster, ed. (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983), p.128–47.
22. Glyn T. Evans and Albert Beilby, "A Library Management Information System in a Multi-campus Environment."
23. Mary Jo Lynch, ed., *Library Data Collection Handbook* (Chicago: American Library Assn., 1982).
24. Lancaster, *Library Automation.*
25. Richard W. Boss, "Technology and the Modern Library," *Library Journal* 109:1189 (June 1984).
26. Maurice J. Freedman, "Automation and the Future of Technical Services," *Library Journal* 109:1202 (June 1984).
27. Mary Ellen Jacob and Neal K. Kaske, "Management Information Systems in a Network Environment," in *Library Automation as a Source of Management Information*, F. Wilfrid Lancaster, ed. (Champaign-Urbana, Ill.: Univ. of Illinois, Graduate School of Library and Information Science, 1983), p.115. ■■

Browsing with Sound: Sound-Based Codes and Automated Authority Control

Karen G. Roughton and David A. Tyckoson

With the development of online systems, libraries have the opportunity to implement new techniques for linking and retrieving bibliographic records. This paper investigates the use of sound-based codes for authority control and record retrieval. Three problems need to be overcome to utilize sound-based codes: code generation, file size and structure, and code retrieval. The SOUNDEX code and the Davidson Consonant Code are explained, and their uses with small, medium, and large files of personal names are demonstrated. Both codes would be relatively simple to implement and would provide useful additional access points in an online authority control file. It is further theorized that these codes could provide more precise retrieval for title searching in an online catalog and could be the basis for alternative catalog access points for the visually impaired and handicapped.

Catalogers take great pains to ensure that library users can locate the materials they need. To meet this goal, authority control attempts to bring together all forms of an author's name, no matter what form is used in the work itself. However, as Malinconico states, "consistent and uniform records describing items in a library collection are only of value *after* they are located by a catalog user. The manner in which records are organized determines whether they will be located."[1] The cataloger provides variants of an author's name so the user need not know the authoritative form when searching. Most variants supplied by librarians presuppose knowledge of an author's pseudonym or a more complete or incomplete form of a name. Enhanced traditional search capabilities would combine authority files with vocabulary files to provide a positive response to Miller's question, "Will an online catalog provide keyword and phonetic searches that will yield access without reliance on authority hierarchies?"[2] Sound-based codes provide mechanical variants and bring together unlimited forms of an author's name, no matter what form is used by the searcher. Sound-based codes make possible the alternative of browsing by sound.

SOUNDEX is a system of coding names according to their phonetic sounds. This technique has been used since the turn of the twentieth century to retrieve names based on pronunciation rather than spelling. It was originally developed to search for names in the 1890 census files. SOUNDEX indexes to the U.S. Census were compiled in the 1930s because widespread misspellings caused difficulties for the Social Security Administration in matching names of persons applying for old age benefits who had no birth or other proof of age records. The Davidson's Con-

Karen G. Roughton is head of catalog management and David A. Tyckoson is information services librarian at Iowa State University Library, Ames, Iowa. This paper is based on a presentation given before the LITA ISAS/RTSD Discussion Group on Authority Control in the Online Environment at ALA Midwinter Meeting, January 7, 1985.

sonant Code is another sound-based system that was designed for and is currently used in airline passenger reservation systems.[3] Both are phonetic coding systems that produce the same code for different words and also have the tendency to create the same code for homonyms and near homonyms.

Davidson's Consonant Code

1. The first character is always the first letter of the surname.
2. The fifth character is always the initial letter of the first given name, if available. Otherwise the fifth character may be a special character (such as the period, .), which is programmed to match any and all initials, or the fifth character may be a dummy initial of infrequent occurrence (say X).
3. The second, third, and fourth characters of the coded name are set up first as space characters and will remain as such unless replaced in step (4).
4. The second, third, and fourth characters of the coded names are then, in turn, replaced by consonants obtained from the surname only, by the following procedures:

   Strike out all vowels and all appearances of H, W, or Y (except for the first letter of the surname).

   Squeeze the remaining letters, including the first letter.

   Each time that any given letter is repeated contiguously, after this squeezing process, eliminate all but one of that letter.

   The first three remaining letters (or whatever number, if any remain, if less than three) which are immediately adjacent to the preserved first letter, then are placed in the second, third, and fourth characters of the coded names, respectively.
5. Each coded name always consists of five characters. If the result of step (4) provides fewer than three additional consonants, the coded name will contain one or more of the space characters inserted at step (3).
6. The first four characters of the coded name will be called the coded surname.

The SOUNDEX Algorithm

1. Retain the first letter of the name, and drop all occurrences of A, E, H, I, O, U, W, Y in other positions.
2. Assign the following numbers to the remaining letters after the first:

B,F,P,V	1
C,G,J,K,Q,S,X,Z	2
D,T	3
L	4
M,N	5
R	6

3. If two or more letters with the same code were adjacent in the original name (before step 1), omit all but the first.
4. Convert to the form letter, digit, digit, digit by adding trailing zeros (if there are less than three digits), or by dropping rightmost digits (if there are more than three).

Sound-based codes may be sufficient to find entries in a small file such as a personal telephone directory, but they would only form part of a search algorithm for a larger index. This system of consonant compression, or number/letter code, reduces many forms of an author's name to a standard form or code and presents the user with an invisible search strategy. Thus, homonymic names such as *Matteson* (M325) and *Madison* (M325) are retrieved together using the SOUNDEX algorithm. Basic problems with the algorithms include a misspelling of the first letter of the surname, e.g., *Craft* (C613) and *Kraft* (K613) and surnames with two concurrent consonants, e.g., *Rodgers* (R326) and *Rogers* (R265). The following examples illustrate sound-based codes and some of their possibilities and peculiarities.

	SOUNDEX		
Johnsen	JNSN	J525	J52
Johnson	JNSN	J525	J52
Johnston	JNSTN	J5235	J52
Johnstone	JNSTN	J5235	J52

		Davidson	SOUNDEX
Henry	Madison	MDSNH	M325
	Matteson	MTSNH	M325
David	Tyckoson	TCKSD	T222
	Ticosen	TCSND	T225
	Tikasin	TKSND	T225
Karen	Roughton	RGTNK	R235
	Rowten	RTNbK	R350
	Rufften	RFTNK	R135
Barbara	Tillett	TLTbB	T430
	Tilit	TLTbB	T430
Jean	Rhys	RSbbJ	R200
	Rees	RSbbJ	R200
	Reece	RCbbJ	R200

Two 1984 articles illustrate the need for computerized reduction of all forms of an author's name to a standard form. Beecher Wiggin's "How Important Are Unique Access Points?"[4] makes two points. First, that user convenience should remain paramount in librarians' efforts to make a bibliographic item accessible, and second, that at the heart of authority control is the concept of uniqueness among headings. Unique headings and user convenience are not always compatible concepts. Sound-based coding focuses on the similarities found among the many forms of a name that a person might use when searching a database and brings them together for a searcher's perusal. Robert Burger, in "Artificial Intelligence and Authority Control"[5] argues that "it will be necessary to apply the knowledge gained in other related disciplines to solve the problems of automated library systems." Online systems must be designed to include consideration of how best to build the user-computer interface so that poorly constructed queries can be converted to well-structured forms that the computer subsystem can handle. Sound-based codes can provide a linking mechanism so that the variants provided by the librarian can be augmented by the extremely literal variants provided by a computer program to offer the library user augmented search strategies. "Augmentation occurs where there is a feedback mechanism in the information retrieval system that presents the user with choices that determine the machine's further actions. With delegation, once the search process is begun, an algorithm, or previous system design, proceeds without intervention by the human user."[6] Here is the subtle shift from choosing a proper heading, with applicable cross-references, to also identifying the various manifestations of a name under which a user would search.

TECHNICAL ASPECTS

There are three primary technical difficulties to overcome in adding sound-based codes to an online authority file—code generation, file space and structure, and code retrieval. The system must be able to create these codes without creating an additional burden on the cataloging personnel or on the online catalog. Codes must be generated entirely by the system, the file must be structured so that searching time and file space are kept at a minimum, and the search codes must be simple to use yet retrieve a valid output.

Code Generation

In order to generate sound-based codes without creating more work for the cataloger inputting records, an online authority file with sound-based capabilities must contain an algorithm to generate and index each sound code. Fortunately, sound-based codes are extremely simple to generate. The original SOUNDEX algorithm was written in 1918 (cf. U.S. Patents 1261167 [1918], 1435663 [1922]), and this algorithm has already been enhanced to run on a wide variety of computer equipment ranging from micros to mainframes. Two very simple BASIC algorithms have been published in *Byte*, one of only twenty-six lines for the Commodore PET[7] and one of thirty-six lines for the Apple.[8] A more sophisticated program for the IBM PC appeared in *Practical Computing*[9], and even a small database manager such as dBase II has had a SOUNDEX program written for it.[10] A good, usable system of sound coding should have some improvements on the original SOUNDEX. However, for each system of sound indexing, a straightforward method of coding the data is used that should easily be translated into an algorithm allowing the online system to generate sound-based codes without any extra effort on the part of the cataloging/inputting personnel. The systems that currently use these codes do not appear to have any difficulties with the code generation aspects of the system.[11-13]

File Size and Structure

An authority file incorporating sound-based codes is obviously larger and requires more searching capabilities than the same file without sound access. However, when measuring the additional space requirements against that of the file as a whole, it is clear that the inclusion of this coding does not substantially affect the size of the file. The most commonly available online authority file currently in use is the one on OCLC. A typical MARC authority file rec-

ord consists of a large fixed field and four to five variable fields. The inclusion of sound-based coding would require one additional variable field consisting of as many characters as the code requires (four characters for SOUNDEX, five characters for Davidson codes or SOUNDEX plus first initial, etc.). This increases the record by 20 percent in terms of fields, but only by a very minute amount (1 percent–2 percent) in terms of characters. In the total OCLC catalog of more than ten million records, the authority file accounts for an additional one million records. Because this file accounts for less than 10 percent of the complete catalog and the addition of sound-based coding to the authority file causes a 2 percent increase in the size of the file, the effect on the entire online catalog is a mere 0.2 percent. In view of the storage requirements for an online catalog, sound-based codes have an insignificant effect.

The area where these codes will have a much greater effect on the structure of the information stored in the computer is in the index files to the authority file and online catalog. At present, the OCLC online authority file uses an eight-character search key composed of four letters from the last name, three from the first name, and the middle initial. Every record in the authority file is indexed by a search key of this structure. A search on any eight-character key will retrieve every record that is indexed under that key and only those records that are indexed under that key. The inclusion of sound-code access would require an additional search key composed of the four (or more) character code generated by the appropriate algorithm. This would necessitate a second index file to the authority control records based on the chosen sound-based code. For the simple SOUNDEX codes, this requires an additional key of four characters (one letter, three numbers) for every record in the database. The number of access keys would double and the amount of storage required for file indexes would increase by at least 75 percent. OCLC currently has access by search keys based on author, title, author/title, ISBN, ISSN, LC Card number, SuDocs Number, corporate author, and CODEN. An additional index file of authority/sound-based codes will have a negligible effect on the overall disk requirements.

Searching SOUNDEX Codes

Even though sound-based coding is simple to generate, requires little file structure change, and uses minimal additional file space, it would be worthless if it did not retrieve useful information from the file. These codes can only be useful if they provide the searcher with additional and different access points from the search keys already in use on the system. Because sound-based codes are designed to be searched only after an unsuccessful search using traditional search keys, the retrieval of a substantially different set of records is essential. A number of studies on files of various sizes have indicated that sound-based codes fulfill this need.

One of the first applications of this type of coding in an automated system was for airline reservations. Because these systems handle a large number of personal names and the possibility of losing a reservation through spelling errors is high, a sound-based code has been implemented to increase accuracy in searching the file. In a study by Davidson in 1962,[14] the Davidson code was matched against searching by exact surname. These codes were found to be highly successful when the traditional search method (exact match of the surname) did not retrieve the desired record. By varying the specificity of the match of the Davidson code, the system was able to retrieve exact matches on progressively varying forms of the name. This system was so successful that it is still in use by airlines and travel agencies today as the standard method of record retrieval.

In another study of a file of 25,382 names,[15] a comparison was made between the retrieval capabilities of SOUNDEX, Davidson codes, exact name searching (the standard search key), and a search on secondary characteristics such as birth date, age, and sex. For total retrieval of similar names, the exact name search found 227 records, 113 of which were successful hits, the SOUNDEX codes found 425 items, of which 213 were successful, and the Davidson codes was successful on 235 of 436 retrievals. Although all three search methods

retrieved approximately the same percentage of "correct" records, both the SOUNDEX and Davidson codes retrieved almost double the number of records found by the name alone. A searcher who performs an unsuccessful search by name or a searcher desiring more records than were retrieved by the exact name search can double the number of hits by using either of these sound-based codes.

Two studies by the Sampling and Survey Research Staff of the Dominion Bureau of Statistics of Canada[16] have also indicated that SOUNDEX enhances retrieval from a large file. In a system of more than eight hundred thousand names from the Canadian census files, SOUNDEX was compared to three other schemes to link together groups of like records with certain characteristics. In comparison to search keys consisting of the name, the first five letters of the name, and the second five letters of the name, SOUNDEX was the most efficient at blocking like names. However, it performed much less adequately when searching non-Western names. For most cases, SOUNDEX was recommended as the method for linking records consisting of personal names.

All three of these studies indicate that some type of sound-based code could be extremely useful in an authority file. SOUNDEX and the Davidson code retrieved 50 percent more items while still maintaining the same rate of accuracy. This provides the searcher with a much broader base from which to narrow the search. The nature of searching the authority file also lends itself to sound-based searching. Because these codes will generally be searched following an unsuccessful exact name search, the searcher will be looking for a greater number of hits with sound-based codes than from the name search.

SOUND-BASED CODES IN A PUBLIC ACCESS CATALOG

Another application for these codes that may have an even wider impact than their use in authority control is the use of sound-based coding as an additional access point in a public access catalog. As in the authority file, the technical aspects of introducing sound-based codes are minimal. The algorithms already exist for generating these codes and, if sound-based authority control were implemented, would already be part of the system. The file structure would require only slight modification. The MARC format can accommodate sound codes by the addition of a single field. A sound-code field would be treated in much the same way as an ISBN or LC card number field. It would be displayed on the screen and could be searchable by the simple four-or five-character sound code. Just as in ISBN, CODEN, ISSN, or LC card number, the search key would be composed of the code itself. The only modification to the index file structure is the addition of the index file based on the sound-based codes.

Searching the catalog by these codes would have many advantages for public access with very few disadvantages. If the patron is unsuccessful using the standard search keys of the system, it provides another alternative for searching the catalog. Rather than respond "Tyck,Dav,A is not in Author index" or "Record not found" or some other negative message, the system could respond by giving the patron a message such as, "This name was not in the file. Would you like to find all of the names that sound like Tyckoson?" This would not only allow the user another chance to find the information, but give the psychological impression that the system is doing all that it can to assist the patron in retrieving the desired materials.

Just as there are individuals who continually search by author when a title search might be more appropriate, there will always be users who misuse the sound-based code access by using it as the initial search query instead of using it after a failed initial search. This could cause frustration on the part of the user when a much larger set of records is retrieved than is expected. Educating the user about effective search techniques is an important part of making any system work.

If sound-based coding was expanded further to code titles as well as authors, the user would have even greater access to the catalog. When the exact spelling of the title is not known, these codes can provide the same additional access for titles that it does for authors. However, even if the spelling of a title is known exactly, there will still be some advantages in title searching. Because

sound codes eliminate vowels and like consonants, its possibilities for title searching may be even more far-reaching than for author searching. Since most standard search keys are constructed on the basis of the first *n* letters of the first word(s) of the title, they often retrieve many more records than necessary due to the fact that many words begin with the same word stem. Sound coding can actually eliminate some of these false drops because it searches solely on the consonants. For example, the word *computer* generates the search key "comp" on many systems. This same search key retrieves the words *complete, comprehensive, comparable,* and *compassionate,* not to mention the many other possible word endings. However, the SOUNDEX code for *computer,* C513, will only pick up those words belonging to the same linguistic class as the original. Although this will retrieve some additional terms, such as *competent* or *camped,* the number of terms may be smaller than for the standard key. If the Davidson code for *computer* is used, CMPTR, the retrieval is even more exact. Soundex codes have the advantage of not only providing a broader search for authors, but also a more specific search by title.

THE TALKING CARD CATALOG

Taking the sound-based concept to its extreme, it is possible to construct a system in which both input and output are based entirely on sound. The user would input a request through a microphone, the computer could translate the voice input into SOUNDEX or other codes, search the catalog, and respond to the user with computer-synthesized speech output. Although this may sound like a 1950s science fiction story, it is technically feasible with equipment readily available on the market today. A number of companies are currently marketing speech synthesizers for computers. Originally designed to make computerized information available to the visually handicapped, these devices are in everyday use in a number of industries, including the telephone, automotive, and electronics industries.

Voice input is also possible through the translation of sound into some type of SOUNDEX code by the computer. Although voice input is currently in use in some military and commercial applications, this part of the system is less well developed at this time. Problems occur due to varying speech patterns, inflections, and accents used by different persons. Despite the fact that voice recognition programs have been developed for many different languages (including Spanish, French, German, Japanese, and Chinese), each program is incompatible with any other. Variations in phonetics between languages require that all voice input into a system be in one standard language. However, as voice input technology becomes more developed and voice output technology becomes more refined, the day when we can talk to the catalog becomes more of a reality. The "talking catalog" would greatly assist some of the users who may have difficulty using the standard online catalog, such as the visually handicapped, elderly, or the functionally illiterate. Ideally, the system would also have a set of sound-based codes for each of the languages of the library user, allowing a non-English-speaking patron to input the request in his/her native language. All that is required is a little more "sound" technology.

REFERENCES

1. S. Michael Malinconico, "Bibliographic Data Base Organization and Authority File Control," in *Authority Control: The Key to Tommorrow's Catalog,* Mary W. Ghikas, ed. (Phoenix, Ariz.: Oryx Pr., 1982), p.3.
2. R. Bruce Miller, "Authority Control in the Network Environment," in *Authority Control: The Key to Tommorrow's Catalog,* Mary W. Ghikas, ed. (Phoenix, Ariz.: Oryx Pr., 1982), p.42.
3. Leon Davidson, "Retrieval of Misspelled Names in an Airline Passenger Record System," *Communications of the ACM* 5:169–71 (1962).
4. Beecher Wiggins, "How Important Are Unique Access Points?" *RTSD Newsletter,* 9:88 (1984).
5. Robert H. Burger, "Artificial Intelligence and Authority Control," *Library Resources & Technical Services* 28:337 (Oct./Dec. 1984).
6. Ibid., p.341.

7. Tom Munnecke, "Give Your Computer an Ear for Names," *Byte* 5:196–200 (May 1980).
8. Jacob R. Jacobs, "Finding Words That Sound Alike—the Soundex Algorithm," *Byte* 7:473–74 (Mar. 1982).
9. David Clarke, "Sounds Familiar," *Practical Computing*, 7:90–93 (Feb. 1984).
10. John Gillen, "The Soundex Routine," *Data Based Adviser* 2:46–47 (Aug. 1984).
11. Emil H. Levine and others, *Pathfinder I—a Minicomputer DBMS. Low Cost High Powered Computing*, Proceedings of the 40th ASIS Annual Meeting, Chicago, Ill., September 26–October 1, 1977, ed. by B. M. Fry (White Plains, N.Y.: ASIS, 1977), p.46–47.
12. Eryatar Shayevitz, "LIS—Library Information Systems for Mini-Micro Computers," in Carl Keren and Linda Perlmutter, eds., *The Application of Mini- and Microcomputers in Information Documentation and Libraries* (Amsterdam: North Holland, 1983), p.161–67.
13. Davidson, "Retrieval of Misspelled Names."
14. Ibid.
15. R. H. Greenfield, "An Experiment to Measure the Performance of Phonetic Key Compression Retrieval Schemes," *Methods of Information in Medicine* 16:230–33 (1977).
16. A. B. Sunter, *A Statistical Approach to Record Linkage* (Ottawa: Dominion Bureau of Statistics, 1967) (Available through NTIS. PB179786.)
17. A. B. Sunter and I. P. Fellegi, *An Optimal Theory of Record Linkage* (Ottawa: Dominion Bureau of Statistics, 1967). (Available through NTIS. PB179787.) ■■

Special Section: Optical Disks

The Information Workstation: A Confluence of Technologies Including the CD-ROM

John C. Gale

The microcomputer, the optical disc, and—in particular—the CD-ROM (compact disc—read only memory) have enabled a new opportunity for distributing electronic information products directly to microcomputer users. This opportunity has become a practical reality due to the confluence of six interrelated technologies.

CONFLUENCE OF TECHNOLOGIES

The Altair computer kit first sold by MITS in 1974 was the first true microcomputer. The next major stride occurred in 1977 when Apple, Tandy, and Commodore all announced products. In August 1981 IBM "legitimized" the personal computer by introducing the IBM-PC. Microcomputer processor cycle times, memory capacities, and access times have steadily improved, especially during the last two years.

Some view the advance of the microcomputer and improvements in the capacity and error rates of removable optical mass storage media as being the only requirements for successful information workstations. The achievements in digitally encoded videodisc (the digital data storage offshoot of the analog videodisc for home movies provided by Reference Technology and Laserdata) and in CD-ROM (the digital data offshoot of the compact audio disc for home stereos provided by Hitachi, Sony, Phillips, and others) are certainly impressive.

However, the substantial changes in the power of microcomputer operating systems and other software engineering achievements, such as those in database retrieval and workstation environments, have also been essential.

Each of the above technologies has achieved substantial progress in the last few years. As a related achievement, sales of computer-based workstations were higher than $4,500,000,000 in 1983. In addition, there is a strong demand for electronically distributed information—total 1983 sales were $1,200,000,000. As a result, information workstations are a soundly based concept whose time has come.

Next, successful vendors will offer user interfaces tailored to specific vertical markets. Local area networks for connecting these systems and telecommunications for communicating with other systems will also be an integral part of future systems.

THE INFORMATION WORKSTATION

The purpose of an information workstation is to provide reference information, information content, or work product support. MARVLS, the MARC and REMARC Videodisc Library System, is an example of a reference workstation product. DISCAT and DISCON are examples of work product support systems for current cataloging and retrospective conversion.

Our present view of an information workstation includes a library of information products on removable media, workstation and retrieval software, media drives, and microcomputer hardware. The International Thomson Information/Carrollton Press approach to designing these systems has been to develop a series of prototypes using both digitally encoded videodisc and CD-ROM.

The databases involved have been pro-

John C. Gale is vice-president and chief technical officer of International Thomson Information, Arlington, Virginia.

vided by sister companies including Carrollton Press, Inacom, COADE, and COMPress. Several of the efforts have involved different applications of the REMARC database. For example there have been several generations of the MARVLS full-text Boolean retrieval system for use in business and library environments. These provide free text searching of the 6,500,000 bibliographic records in the MARC, REMARC, and LawMARC databases, which represent the complete Library of Congress shelflist.

In addition, there have been prototypes of library automation systems for current cataloging (DISCAT) and retrospective conversion (DISCON) systems using the MARC and REMARC databases. These systems have less sophisticated search capabilities. Their purpose is to support a certain work task, so they provide only the help, search, edit, display, store, print, and transmit facilities needed for the job at hand.

The Inacom MicroIndex is an index to 800,000 pages of electronic engineering data sheets. The index provides access by both product descriptors and company name, using a controlled vocabulary. The software assists users if they do not know the correct spelling of a subject term or company name.

Demo-Graphics: World Populations and Projections was created by Paul Handler, University of Illinois. The database includes data on population, fertility, and mortality for 146 countries. Users may modify the data to reflect recent changes or estimates of present and future fertility and mortality values. Various plots, based on user-controlled assumptions are available. Data can be used to forecast changes in population subsets for two countries of the user's choice, with the results plotted in color.

IR and NMR Spectroscopy are included in a segment of the *Introduction to Organic Chemistry* by Stanley G. Smith, University of Illinois. Infrared and proton nuclear magnetic resonance spectroscopy are introduced, and problems provide practice in identifying unknowns from their IR and NMR spectra. Spectroscopy is used throughout the rest of the programs in the development of functional group chemistry.

Several of these efforts have included cooperative ventures with outside organizations such as Cuadra Associates, the University of California, and the Library Corporation.

These ventures have been most helpful in exploring the cost, performance and user interface tradeoffs for developing appropriate information products. For example, Cuadra Associates' STAR software was used for one generation of the MARVLS reference system. This generation included 1,000,001 bibliographic records that, with indexing, required a total of 540,000,000 characters of optical "real estate." In all, six different information retrieval systems have been utilized for four different databases.

For information workstations to achieve substantial penetration into professional communities, user-friendly and user-transparent operating procedures will be needed. These easily understood user interfaces will consider the "effort-to-learn" and "effort-to-use" factors in their design. International Thomson Information, Carrollton Press, Inacom, COADE, and COMPress have made significant progress in this area by creating seven information product prototypes and exhibiting them in various arenas in order to gauge and respond to user reaction.

FUTURE IMPLICATIONS FOR SOCIETY

At a recent Information Industry Association conference a Phillips executive described the CD-ROM as the most significant advance in publishing technology in the last four hundred years.

It is now possible to distribute more information to more professionals in a more accessible environment with more attractive pricing than ever before.

IMPLICATIONS FOR USERS

It is obvious that the availability of information will improve. Additionally, the fixed-cost environment will clearly encourage unlimited browsing by users who will have the opportunity to "learn" a database

as never before. Information downloading and interfaces with word processing packages will become prevalent. Professionals will thus have better information available to provide documented support for decisions they must make. Researchers will have more time to make their exhaustive searches even more thorough.

The work restructuring that will result from the use of information workstations may have impacts on user work environments that cannot be foreseen. For example, the interplay of rational user information needs and user productivity issues with user organizational political power issues is far from clear at this time.

Security of discs will be an important issue, as will updating both information and retrieval software. User assistance and customer support will also be changed from previous patterns. And the user will have easier access to more information.

FUTURE TRENDS

Products that offer real value to their users will be most successful in the end. This seems to be a trivial observation, but many organizations fail to consider whether or not they offer true "value-added" in their products.

Over time, retrieval systems will evolve into advisory systems and then into true expert systems. During this process application generators and software engineering will have a profound impact on system design. Databases of digitized images and cadastral information will be accessible from networked, 32-bit microcomputers. In the foreseeable future there will be distribution of substantial information products to the home.

SUMMARY

A confluence of technologies and market demand has made the information workstation possible. It can support reference searches, information content perusal, and work product efforts. International Thomson Information, Carrollton Press, Inacom, COADE, and COMPress are well postured to assist their customers with information products and services based on their efforts with these state-of-the-art technologies.

BIBLIOGRAPHY

Gale, J. "CD-ROM Standards: Do the Users Really Benefit?" *CD Data Report*, March 1985.

Gale, J., "Use of Optical Disks for Information Storage and Retrieval." *Information Technology and Libraries* 3: 379–82 (Dec. 1984).

Paisley, W. "Design Considerations in the Shift from Online to Stand-Alone Microcomputer Retrieval Systems." Institute for Communication Research, Stanford University, 1984. ■■

Evaluating Laser Videodisc Technology for the Dissemination of Agricultural Information

Pamela Q. J. André

The National Agricultural Library is evaluating the use of laser videodisc technology as a means for disseminating agricultural information. The study is being funded as part of a strategic planning effort by the United States Department of Agriculture (USDA) to assess long-range opportunities in the utilization of technology as it relates to agriculture and agricultural information.

The focus of the evaluation is the potential of laser videodisc technology used in conjunction with microcomputers for storage and dissemination of agriculturally related full-text databases. In September 1984 the National Agricultural Library contracted with LaserData of Cambridge, Massachusetts, to develop a system utilizing these two technologies. The system will include all hardware and software necessary to store and retrieve a full-text file together with a videodisc as the storage me-

Pamela Q. J. André is chief, Information Systems Division, National Agricultural Library, Beltsville, Maryland.

dium for the database selected.

Since end-user experience is to be a critical part of this evaluation, the system will be installed in three locations: the National Agricultural Library, Purdue University Library, and the National Pork Producers Council. The evaluation of the system will be coordinated by the Agricultural Information Service at Purdue University with participation by pork producers, researchers, and various USDA personnel.

Laser technology has the potential to change significantly the means by which publications and databases are stored, distributed, and used. The results of this evaluation will provide data on the user acceptance of microcomputer/laser videodisc systems as a means of access to textual agricultural information.

THE DATA

One of the most exciting aspects of laser videodisc technology is the very high storage capacity, up to eight hundred million characters of data on one disc. The capability to store both digital and analog data makes use of the videodisc extremely flexible since both text and graphics can be stored on the same medium. These two characteristics were the primary focus in our database selection decision.

The database we selected for this evaluation project is the *Pork Industry Handbook (PIH)*. The *Handbook* in its printed form is a major teaching and resource tool for the pork industry. It is published jointly by Purdue University and the USDA Extension Service. The *Handbook* consists of ninety-six chapters in loose-leaf form. Each chapter describes an aspect of the care and feeding of pigs. This publication was selected for its variety of material, i.e., text, pictures, and charts, as well as its high use within a specific segment of the agricultural community. A page of the *Pork Industry Handbook* is shown in figure 1.

One drawback of the publication is its relatively small size, approximately two million characters. With a disc capacity of eight hundred million characters, the *Pork Industry Handbook*, even with full text, graphics, and indexing, does not come close to the disc capacity. So one of our early problems was a way to test the storage capacity of the disc.

AGRICOLA gave us the answer. *AGRICOLA* (*AGRIC*ultural *O*n-*L*ine *A*ccess) is the bibliographic database created and maintained by NAL. It contains more than two million bibliographic citations for monographs, serials, and journal articles and covers a wide variety of agricultural topics. Two hundred thousand current *AGRICOLA* records were selected for inclusion on the videodisc. These bibliographic records will be accessed separately from the *Handbook*. However, by using bibliographic records as part of this project, we will be able to experiment with the use of microcomputer technology to support a stand-alone "online" catalog. While the *AGRICOLA* records will be defined as a separate file for access, they can be used in conjunction with the *Handbook*.

"In" analytic or indexing records describing each of the ninety-six chapters of the *Pork Industry Handbook* are included in the *AGRICOLA* subset on videodisc. This means that end users interested in items on pork production but not aware of the *Handbook* can approach the bibliographic file and, through subject-oriented searches, retrieve citations to the *Handbook*. At that point, because both the *AGRICOLA* citations and the full text of the *Handbook* are on the same videodisc, the user can easily retrieve the appropriate full-text paragraphs for viewing on the same terminal.

More sophisticated users who know of the *Handbook* can approach the stand-alone system and access portions of the *Handbook* directly without having to go through the bibliographic file. Access will be provided to each word in the text of the *Handbook*, excluding a limited list of stop words, as well as through the author and title for each chapter. In addition, chapter subheadings, paragraph headings, and captions for tables and other graphics will provide access.

THE PREPARATIONS

Once the *Pork Industry Handbook* had

NUTRITION PIH-3

pork industry handbook

COOPERATIVE EXTENSION SERVICE • PURDUE UNIVERSITY • WEST LAFAYETTE, INDIANA

Energy for Swine

Authors
Palmer Holden, Iowa State University
Lowell Frobish, USDA-ARS, Washington, D.C.

Reviewers
Bill G. Diggs, Mississippi State University
Herbert M. Barnes, Ohio State University

The value of a feedstuff is based on several factors: acceptability (how well the material will be consumed by an animal), energy availability and as a source of other nutrients (protein, vitamins, minerals). Should a swine producer buy corn or wheat or oats as a feed ingredient? This will depend primarily on the cost of these ingredients and their value as a source of energy for the animal.

The pig requires energy to maintain normal body processes, to grow and to reproduce. Energy is the major component of all swine diets, and the intake of many other nutrients is related to the energy content of the diet. Carbohydrates from cereal grains are the most abundant energy source in swine rations. Fats and oils contain more energy than carbohydrates but are used to a lesser extent. Protein may serve as an energy source only if included in the ration in excess of the animal's requirement for protein.

Definition of Energy

To make sound decisions in selecting feed ingredients it is desirable to have an understanding of the system by which feedstuffs are rated for their energy content and the value of these ratings toward the pig's growth and production. The gross energy (GE) of a feed ingredient is defined as the heat produced when a substance is burned and is expressed as calories per unit weight. A calorie is the amount of heat required to raise the temperature of 1 gram of water from 14.5°C. to 15.5°C. A kilocalorie is 1,000 calories, and a megacalorie is 1,000 kilocalories. Since not all of the feed consumed is digested, some energy is lost in the fecal material (Fig. 1). GE is a poor estimate of energy for the pig.

The amount of energy remaining after subtracting the fecal energy loss from total energy intake is designated as digestible energy (DE). The difference between GE and DE may be quite large. The greater the digestible energy value (DE/GE) the greater its value as a source of energy to the animal. This is a much more meaningful measure for livestock producers. In the production and excretion of urine additional energy is lost. Digestible energy minus the urinary energy loss is termed metabolizable energy (ME).

In most cases, metabolizable energy represents approximately 95% of the digestible energy content, so the conversion from DE to ME can be made very easily. Metabolizable energy is the "usable" energy of a feed for

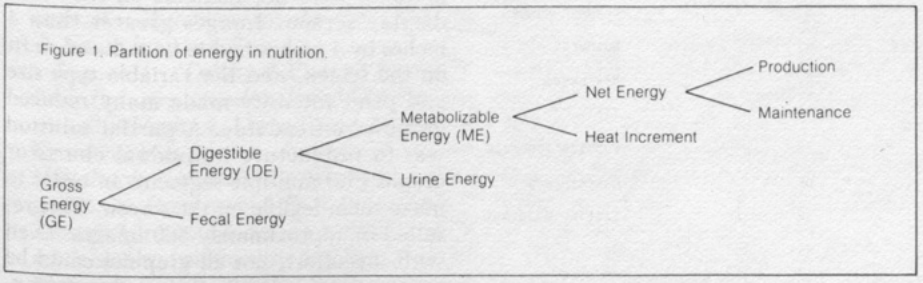

Fig. 1. The Title Page from the Pork Industry Handbook, *Chapter 3.*

been selected for the evaluation project, it had to be put into digital form in a format that could be processed by the retrieval software. There were two choices at this point. Since the *Handbook* was a printed publication created on a word processor, one alternative was to reformat the original keyed data. The other alternative was to key the full text based on the printed copy. Because of the difficulty of identifying the elements of the text from the word processing file, the decision was made to key the full text into a format that could be readily used for indexing and retrieval. A contract was let to Bibliographic Retrieval Services (BRS) to design the full-text format and key the data.

A major consideration in the structuring of the full-text file was to have in-depth access that would include indexing at the word level and display of specific paragraphs. In addition, access via author, title, chapter and paragraph headings was desired. The final structure of the data file is shown in figure 2.

In addition to the full text, BRS was able to key some portion of the charts in the *Handbook*. They also keyed the numbers and captions for the pictures and graphs. The full definition and keying activity took more than a year to complete. However, in February 1985 the full-text file, complete

```
Database Name:  PORK

Banner:  PORK INDUSTRY HANDBOOK ONLINE

Paragraph Labels and Names:

        SHORTNAME                LONGNAME

            NO                   Chapter number
            AU                   Authors
            RV                   Reviewers
            TI                   Title
            PD                   Publication date
            DE                   Descriptors
            SH                   Section Headings
            TX                   Text
            SHRT  equiv NO,AU,TI,DE
```

Fig. 2. Pork Industry Handbook *Data Structure.*

with field identifiers, was received at NAL. After a quality review process, the tape was forwarded to LaserData to be indexed and stored on the laser videodisc.

The *AGRICOLA* records exist in MARC format as part of NAL's bibliographic master file. Unfortunately, not all of the chapters in the *Handbook* had been indexed, so the first task was to create indexing records for each of the ninety-six chapters. Once completed, these records were merged with approximately two hundred thousand *AGRICOLA* records from 1984–85. The result was fourteen 1600 BPI tapes for LaserData. Conversion work began in early February 1985. Eventually the fourteen tapes will be replaced by a single videodisc. The structure of the *AGRICOLA* file is shown in figure 3.

The pictures, charts, and graphs in the *Handbook* presented the biggest challenge. There are 541 graphics in the ninety-six chapters of the *Handbook*. One hundred fifty-nine of those were keyed by BRS as part of the full-text file, leaving 382 to be transferred to videodisc using analog techniques. Originally, the plan was to get the originals of each graphic for transfer to the videodisc. However, after checking the quality of the printed copy, it was determined that the originals were not necessary and the printed copy could be used to produce high-quality display images.

It looked like a simple transfer from paper to disc was possible. This turned out to be far too simple. Working with samples from the *Handbook*, LaserData did some preliminary video images for review, and it was readily apparent that many of the graphics were not readable on the video display screen. Images greater than 4 inches by 4 inches had to be reduced to fit on the screen, and the variable type size and print intensity made many reduced graphics unreadable. A partial solution was to restructure individual charts or graphs into multiple segments in order to make them legible on the screen. This resulted in approximately 560 images. Even with this effort, not all graphics could be restructured and still retain any meaning. However, because this is an evaluation of new technology, some imperfections are to

Database Name: AGRI

Banner: AGRICOLA 1982-1984

Paragraph Labels and Names:

SHORTNAME	LONGNAME
IDNO	ID Number
AUTH	Author
TITL	Title
TIFL	Title Foreign Language
IMPR	Imprint
COLL	Collation
NOTE	Notes
SUBJ	Subject Terms
GCOD	Geographic Codes
SCAT	Subject Category Codes
LANG	Language of Text
YRPB	Year of Publication
STAB	Serial Title Abbrev.
HCIT	Host Item Citation
ISSN	ISSN
CODN	Coden
ABST	Abstract
NALN	NAL Call Number
LOCN	Local Call Number
DCTP	Document Type
PSCD	Publication Source Code
SFIL	Subfile Code

Fig. 3. AGRICOLA *Data Structure.*

be expected. We chose to include all graphics whether they could be restructured to be readable or not. The result will show the limits of current technology and will also give an indication of how publication patterns will have to change if this technology is to be effectively utilized. Examples of an original picture from the *Handbook* and its restructured version for the videodisc are shown in figures 4–7.

Figure 4, "Original Picture from the *Pork Industry Handbook* Publication," is actually "Figure 4, Three-Breed Rotational Cross System" from chapter 39 of the *Handbook*, entitled "Crossbreeding Programs for Commercial Port Production." This picture could not be reproduced in readable form on the system video monitor. Figures 5, 6, and 7 present the restructured version showing the original figure number and the restructured part number.

THE SYSTEM

Most of the hardware and software components of the microcomputer videodisc system are "off-the-shelf" technology, the exception being the LaserData controller. Given this fact, there were a number of alternatives in terms of the development of the system. However, because the heart of the system is the use of videodisc as a storage medium for digital data and the encoding and interface for utilizing that data is a specialized process, we elected to have a single vendor take responsibility for all components in the system. We could be assured of a working system without the aggravation of doing the technical interfaces ourselves.

Videodisc technology is based on laser optics, which uses light beams to read and write data. The result is a data storage technique that offers very high density, high-quality video, random access, and a mixed storage format utilizing both digital and analog data. The system implementation used for this evaluation study will utilize each of these characteristics.

The system contains the following components. The hardware includes an IBM PC-XT with 512 kilobytes of memory: a Pioneer LDV-1000 vidoedisc player; a LaserData TR10 110 controller with interface card; and a Panasonic TR124 monochrome video display monitor. The software consists of the PCIX operating system and a sophisticated retrieval package, BRS/SEARCH.

The key to the system is a unique blend of two very different technologies, laser videodiscs and database management systems. Data management software is required to preprocess and retrieve information. Managing hundreds of millions of characters of information is a complicated task, especially when the user wants specific information in a matter of seconds. Tying together the multimedia capabilities (video and data) complicates the matter even further. LaserData has developed the software tools necessary to manage these types and amounts of data.

In addition to its data encoding techniques and data management software, LaserData manufactures the interface controller that connects a standard laser vid-

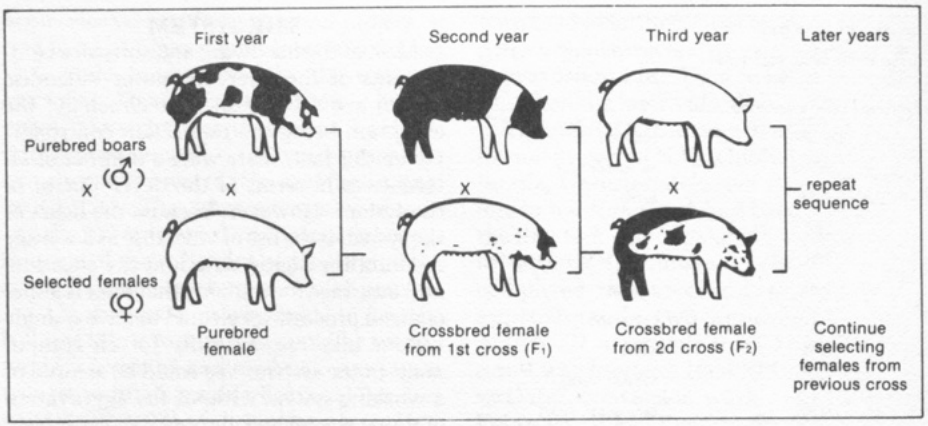

Fig. 4. Original Picture from the Pork Industry Handbook *Publication.*

eodisc player to the personal computer, allowing retrieval of the multimedia information that is encoded on the disc. The unit controls the videodisc player, retrieves digital information from the analog videodisc, and re-creates the digitally encoded material.

The retrieval software used to index, retrieve, and display both the *Pork Industry Handbook* and the *AGRICOLA* bibliographic records is BRS/Search. The package will be modified by LaserData to allow access to graphics. The retrieval system has a menu approach to subject searching that is powerful and easy to use. This is especially critical since we anticipate that most of our users will be novices in online searching. The capability to access any word in the full text and combine words using Boolean operators means more precise searching. Finally, the convenient display, browsing, and print options will help to provide maximum flexibility in using the *Handbook*.

Figures 8 and 9 show displays for the *Handbook* and *AGRICOLA*.

The most critical aspect of the system is the data conversion process and the mastering of the videodisc itself. The digital texts of the *Handbook* and the *AGRICOLA* file will first be indexed for access by BRS/Search. Next the graphics for the *Handbook* will be transferred to videotape using a standard videodisc camera. The data-

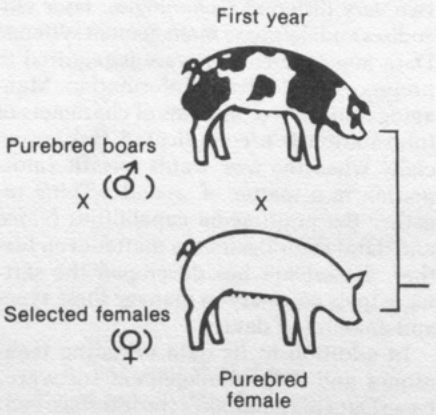

Fig. 5. Restructured Version of Figure 4, One of Three Images.

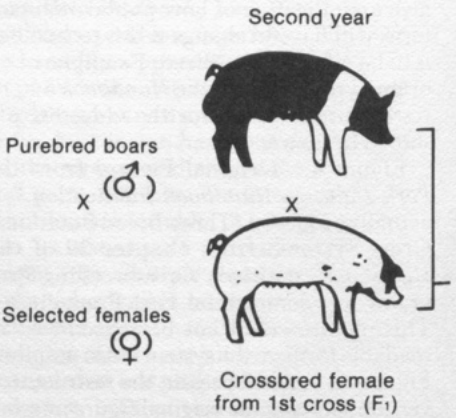

Fig. 6. Restructured Version of Figure 4, Two of Three Images.

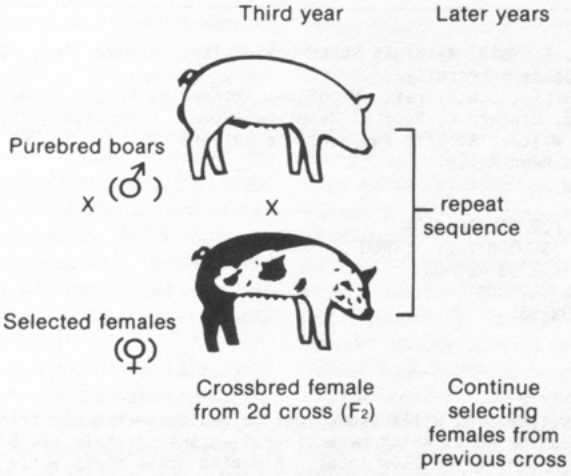

Fig. 7. Restructured Version of Figure 4, Three of Three Images.

bases and indexes will then be premastered to videotape using LaserData's proprietary encoding technique. This process will encode the digital information within the analog signal appropriate for storage on the videodisc. The videodiscs will then be created at a mastering facility. While the data are being mastered, the hardware will be integrated and tested. Final integration and acceptance testing will take place prior to the initial installation.

THE EVALUATION

The three videodisc/microcomputer systems are scheduled to be installed at NAL, Purdue University, and the National Pork Producers Council in early April, followed by a brief training period. The evaluation period, with data collection activities at the three sites, will extend from April through September. The data will be analyzed in October and a final report completed by November.

The evaluation aspect of the laser videodisc project will be directed by the Agricultural Information Service at Purdue University. Since the system will be installed at multiple locations, there will be a coordinator at each of the three sites to ensure that training, use, and data collection activities are monitored.

Staff at Purdue are developing an interactive training package on a microcomputer that will introduce users to the system. A program will guide each user through the rudiments of utilizing BRS/Search, including a simulated interaction with both the *Pork Industry Handbook* and *AGRICOLA* files. Once users complete the training, they will be free to use the two databases on the videodisc system.

Data collection is a critical aspect of the project. Unfortunately, plans for the data collection are still in the design stage, making it difficult to cite specifics. However, some general comments about the kinds of information we plan to collect may be helpful. A major concern is the usefulness of the file on microcomputer versus the printed version of the *Handbook*, so users will be queried specifically about their preferences. The in-depth access, i.e., word level searching and paragraph level display, is a major aspect of the retrieval system for full text that we plan to explore. We are also interested in the availability of the related pictures, charts, and graphs. We anticipate that the uncertain quality of the graphics will be a problem. Another issue to be explored is the need for a printed copy of selected materials. The system will not have an attached printer, although a printer hookup is technically feasible. The ease of use of the retrieval system for novice users will be explored. The question of cost will be broached as a practical consideration for future planning. How much would a user be willing to pay for his own machine ver-

```
NO  PIH.94
AU  Purkhiser, E. Dale; Michigan State University.  Bloome, Peter D.;
    Oklahoma State University.
RV  Bailey, Curtis; Schoolcraft, Michigan.  McKenzie, Bruce; Purdue
    University. Renbarger, Bob; El Reno, Oklahoma.  Singleton, Art; Logansport,
    Indiana.  Wilcox, Robert; Kansas State University.  TI  TITLE:  Calibrating
    Meter-Type Feed Mills.
PD  840300 new
DE  Nutrition
SH  PREPARING FOR CALIBRATION
    THE TIMED (SEQUENTIAL) METHOD
    THE SIMULTANEOUS METHOD
    ADDITIONAL METHODS
    QUALITY CONTROL
    SUMMARY

TEXT
TX  1 of 7
        Meter-type feed mills (sometimes called automatic electric mills) are
    feed processing units which meter ingredients separately and blend them
    together in controlled flow rates.  Blending takes place while the grain is
    ground and/or during conveying of the ration.  Finer ingredients may bypass
    the grinding chamber.  Each ingredient is supplied by a separate metering
    device.
        Different brands of mills feature different types and locations of
    metering devices (Figures 1 and 2).  Meter types include:
    *variable speed proportioning augers,
    *adjustable gravity flowspouts above belts or rotating tables,
    *fluted wheels with adjustable flow gates,
    *vibrators with adjustable flow gates.

        *Figure 1. A mill is shown with all meters above the
     grinding chamber.* (See Video Screen)

        *Figure 2. This mill has minor ingredient meters down-
     stream from the grinding chamber.* (See Video Screen)

        Ingredient meters measure material by volume while ration formulations
    are specified by weight.  Calibration is the process of determining the
    meter settings that will produce the desired ratio of ingredient weights in
    the ration.  Periodic calibration checks are used to determine if the
```

Fig. 8. Pork Industry Handbook *Text Display.*

```
IDNO  ADL84133159
AUTH  Laycock, G.
TITL  Hogs in the hills [European wild boars, eastern Tennessee, North
      Carolina]
SUBJ  Tennessee
      North Carolina
GCOD  G474
      G465
SCAT  L300
LANG  eng
YRPB  1984
STAB  Audubon
HCIT  Audubon; New York, N.Y.:  National Audubon Society.; Sept. 1984.
      v. 86 (5); p. 32, 34-35. ill.
ISSN  0097-7136
CODN  AUDUAD
NALN  S900.A8
DCTP  Journal Article
PSCD  6
```

Fig. 9. AGRICOLA *Bibliographic Display.*

sion of the *Handbook*, or any other agricultural database?

A primary reason for the three separate installations is to make the system available to the user community, that is the pork producers themselves. A major aspect of the evaluation project will be to identify potential users and make them aware of the availability of the system. At Purdue, one approach will be to coordinate with animal science faculty and introduce the system to the students as part of their class assignments.

At the National Agricultural Library in Beltsville, members of the agricultural research community will be invited to make use of the system. At the National Pork Producers Council in Des Moines, members of the pork producers community will be included in the evaluation project. A wide variety of interest and response is expected as a result of the geographic and professional differences in participants.

RELATED PROJECTS

The laser videodisc/microcomputer project is one of three projects that NAL has undertaken to evaluate various aspects of computer technology for the delivery of agricultural information. A related project actually predated the videodisc project. The earlier project, entitled "Full Text Information Online," focused on enabling USDA to gain experience in creating full-text machine databases; to test the quality of indexing and retrieval; and to evaluate the speed and reliability of full-text delivery. The *Pork Industry Handbook* was also used for this project, which resulted in the *Handbook* being made available as a private file on BRS. This online mainframe approach to database dissemination is being evaluated in conjunction with the laser videodisc/microcomputer file and the print publication.

The result of these joint evaluations will be a report comparing and contrasting various aspects of each, including cost, retrieval effectiveness, and user acceptance. The final report is scheduled for completion in fall 1985.

A second project is an outgrowth of these two. This project, entitled "Creating Full Text Online Databases from Photocomposition Tapes," has as its purpose the exploration of the possibilities for creating full-text online databases from preexisting photocomposition tapes, rather than having to key each database. The United States Department of Agriculture generates more than sixteen hundred printed publications annually. Many of these are created using photocomposition, which means that the data are keyed into machine-readable form as part of the publication process. These databases could be made available to a broader audience with more effective retrieval techniques through online access. This project will look at the various formats used by USDA in its publication process to determine the most effective approach to converting these files to online databases as another method of information dissemination.

There is a growing awareness within the agricultural community of the great potential of computer technology for information access and delivery. The videodisc/microcomputer evaluation in conjunction with the two related projects noted above will provide significant information on the potential uses of these technologies. The question that remains is What will be the next step? ■■

Encoding Databases on Optical Disks

Bryan G. Lewis

Conventional videodisc players can be used to store and retrieve any combination of digital data, still pictures, video sequences, or still-frame audio using the encoding/decoding schemes pioneered by LaserData, Inc.

LaserData has devised schemes for error correction and identification of digital information melded with the analog recording media and playback facilities of the vid-

Bryan G. Lewis is vice president for sales, LaserData, Inc., Cambridge, Massachusetts.

eodisc. The encoding and decoding system developed by LaserData uses a proprietary data framing algorithm to guarantee that the data can be found in the video signal. The system is completely insensitive to variances in the horizontal and vertical synchronizing signals of NTSC video.

This previously unavailable combination of video, digital, and now audio capabilities will further fuel the explosive growth that videodiscs are experiencing in fields like interactive training and education, point of sale, and database search and retrieval.

The frequent need to access very large databases can be performed by microcomputers with local storage. In the database area alone, those benefits include the economies of having local access without telecommunication costs; the speed of videodisc access compared to searching through dozens, or even hundreds, of floppy diskettes; and the convenience and time savings of a single, multilevel directory over microfiche or computer printouts.

LASERDATA RESEARCH

The research leading to the LaserData products addressed the performance characteristics of video in the optical videodisc replication process. LaserData research yielded a recording methodology having significant data densities and extremely high reliability.

The system's technical capabilities include:

- Error rates that are equal to or better than standard magnetic media used to store data (one error in one trillion bits).
- Capacities that far surpass even future magnetic media capacities with significant cost advantages.
- Convenience and ease-of-use features not reliable with any other data-storage and retrieval media.

LaserData has developed its technology with two basic assumptions in mind:

1. Standard, off-the-shelf laserdisc players are used as the "disc drive" in a LaserData system.
2. The standard laser videodisc mastering process is used to create digitally augmented videodiscs.

Both these approaches allow LaserData to take advantage of the many millions of dollars that have been spent on developing laserdisc technology. In addition, videodisc mastering companies are continually improving the mastering process, and videodisc player companies are continually improving player access speed and reliability.

The standard laser videodisc carries information on a fifty-four thousand revolution spiral that comprises a continuous sequence of pits read by laser. These pits represent an FM (frequency modulated) carrier, which when demodulated corresponds to the 4.2 MHz NTSC video signal. The LaserData digitally augmented videodisc is capable of storing up to eight hundred megabytes of digital information, encoded within the 4.2 MHz video bandwidth. The digital information is encoded during the active picture portion of the signal, with multilevel samples clocked at the video subcarrier frequency, yielding 15,288 bytes of data per frame including 64 header bytes reserved for the system. The LaserData encoding and decoding system uses a proprietary data framing algorithm to guarantee that the data can be found in the video signal, and is completely insensitive to variances in the horizontal and vertical synchronizing signals of NTSC video. (Other digital data encoding schemes rely on these signals and therefore have to maintain very tight tolerances during the mastering process.) LaserData is also capable of putting digital data only on specific lines of the video frame, allowing mixing of the data and video on a line-by-line basis.

The manufacturing process for making videodiscs is a multistep process and surface impurities are introduced on the face of the disc. These impurities cause dropouts on the FM carrier of the disc that translate into dropouts in the disc's information content. For analog pictures this is not very noticeable, especially since the players have what is called a "dropout compensator" that detects the errors and covers them up with information from adjacent video lines. For highly reliable digital information a different methodology must be employed to assure that "computer grade" information (i.e., error rate less than 10**-12) is retriev-

able from the disc. LaserData has developed a proprietary error correction scheme that optimizes the trade-off between overhead (the amount of extra information required on the disc to support the error correction process) and processor power (the amount of effort the computer's CPU has to expend to correct errors). This scheme, which is a combination hardware and software algorithm, assures the user of retrieving information with the same or better reliability than winchester disc drives.

To create a disc, the user need only submit magnetic tapes with data to LaserData for premastering onto one-inch videotape. If index preprocessing is required, LaserData will do the preprocessing for either LaserKey (for multiple-key ISAM lookups) or BRS MicroSEARCH (for full-text searching). There are several steps to the subsequent premastering process. First the data files are processed to include the error correction overhead (the overhead is used by the system when correcting errors). Then they are formatted into a digital representation of the final videodisc structure. Finally all the files are DMA transferred at approximately one megabyte/second through proprietary Unibus based D-to-A encoding hardware onto the final one-inch tape. The tape is then sent to the mastering facility where a videodisc master is created from which any number of copies can be replicated.

SUMMARY

Today significant barriers still exist between individuals and the information required for effective action and sophisticated decisions. The information exists. Yet with the current set of delivery systems, such as microfilm, online or specially trained personnel, it is inconvenient for users to quickly, simply, and cost-effectively retrieve the information.

An obvious solution is a system that puts entire sources of knowledge in a self-contained package completely under the user's control—a personal database. LaserData offers such a system, combining sophisticated access software with its unique, digitally augumented videodisc.

LaserData products are currently being sold in several segments of the electronic publishing market: for the replacement of online and microfilm applications and as electronic catalog and point-of-sale systems. The singular multimedia capacity of LaserData's system also has tremendous implications in the videodisc training and education markets, into which LaserData has begun selling via OEMs and systems integrators. ■■

Information Access Company's InfoTrac

Richard Carney

Capacity . . . cost . . . complexity.

These were the constraints that spawned a new product development project to fill a perceived void in the Information Access Company product line. Our original library product, computer output microfilm, provides a reference system that requires no user training and offers more than two million citations. The capacity of our microfilm products, however, is limited to the images that can be contained on a seven-inch reel of 42x reduction film. This system also restricts the user to linear searching by single topic.

At the other extreme, IAC offers its databases online, which breaks through the capacity restraints and supports highly sophisticated searching techniques. Some of IAC's online databases contain retrospective material twenty years old, and the searcher can use multiple-term subjects to search by controlled vocabulary or free-text methods. The cost and complexity of access to online databases, however, generally restrict their use to trained information professionals.

By April 1984, Information Access Company was meeting periodical reference needs with a Reference Network that provides a range of search options from simple,

Richard Carney is vice-president—product development, Information Access Company, Belmont, California.

single-subject citations to full-text retrieval in microform or online.

Further, because of the design of IAC's COM indexes and the simultaneous mounting of the online databases, a product line of discrete periodical indexes evolved: *Magazine Index*, *National Newspaper Index*, *Business Index*, and *Legal Resource Index*. The contents of the indexes are identical in both COM and online products, with the exception that the online contained much more retrospective material in most cases. Therefore, whether searches were of IAC's microfilm or online products, the user would have to search each index sequentially if the subject was likely to be covered in two or more of the databases. When IAC mounted its first full-text microfilm and online databases—*Magazine Collection* and *Business Collection*, and *Magazine ASAP* and *Trade & Industry ASAP*, respectively—they, too, directly complemented existing IAC online indexes for ease of database selection and searching. However, there was an identified need for a database that integrated selected material from all of IAC's existing indexes plus the addition of new citations to other periodicals to more clearly reflect the collections of major libraries.

Based on market studies and the recognized constraints of existing products, objectives for a new reference system had evolved by mid-1982;

- It was to be a self-contained system for in-library use.
- It was to demand little or no user training.
- It was to offer a more extensive database than any of the IAC's existing in-library products.
- It would, to be cost-effective, have to combine storage capacity, timely material, speedy access, search options, and simultaneous multiuser capability.

Our product objectives pointed unwaveringly to the use of computer access to an in-library data storage system that would be updated monthly. The evident benefits of computer access, such as speed, the ability to go instantly to *see also* references without the need to crank through all or most of a microfilm reel, and the capability of printing out the results of a search, were offset by the economics of such an electronic system.

IAC was no stranger to developing computer programs to access databases. In fact, the company's Search Helper front-end software, introduced in 1982, was the first major effort to address the cost and complexity issues of online searching. Search Helper offers coaching menus to enable users to establish their search strategies offline. The program then automatically dials the vendor computer, conducts the search, provides the option of reading the results on the screen and/or printing out up to twenty citations for further reference, then automatically signs off. Each search costs the user or library only $2.50 or $3.50 depending on which Search Helper package was purchased. Refinements in the Search Helper programming have been ongoing, and subscribers are provided with new generation software as it becomes available.

One of the things IAC learned in the development of Search Helper was the necessity of following a natural thought process to coach the user in structuring a search and of using natural-language instructions, devoid of jargon or ambiguities. This sounds like a simple lesson; however, when a programmer is struggling within the tight parameters of computer logic while trying to encompass the wide range of human logic, the result can be virtually incomprehensible to the user. Empirical testing, we discovered, was the only way to judge whether a computer program was people-proof. There is no way a program developer can design truly user-friendly software in an engineering lab. Real-world testing is vital, and refinements should be anticipated before a product is market-ready.

So, we entered the first phase of development of a self-contained, computer-accessible reference system, knowing that software design was to be our greatest challenge. A close and critical second consideration was cost. At the time Search Helper was developed, random access memory (RAM) was so expensive, the program had to be crammed within the confines of 16 kilobytes of RAM in order to be offered to libraries at an affordable level.

Advances in semiconductor design and

manufacture and increasing competition had, by late 1983, made possible personal computers with more than fifteen times the random access memory at the cost of their 16K predecessors of two years before. This lifted many of the limitations we had confronted in early search software.

At the time, however, 5¼-inch floppy magnetic disks were the accepted storage media for personal computers. Large-capacity hard disks that could contain up to 10 megabytes of information remained economically prohibitive for most library applications and, in fact, could not provide the storage capacity we demanded in the reference system we envisioned. Videodisc was just emerging as a viable medium for mass storage, and the necessary industry support, such as premastering studios and dependable mastering services, had developed to the point that we could be confident in evaluating the medium for our use.

In videodiscs we found the storage capacity we needed: 800 megabytes on a single disc or the equivalent of 530,000 pages of print material. This would enable IAC to deliver selected contents from most of its index databases in a single product. Videodiscs also provided additional appeal as we tested them for our application. First, they are virtually indestructable in routine use, far more resistant to abuse than traditional audio records or magnetic disks. Next, they're easily replaced. As each updated videodisc arrives from IAC monthly, the librarian simply slides out the disc in use and slips in the new version. Finally, as a read-only memory medium, videodiscs are comfortable in the publishing business, as all of our products are distributed in read-only form: books, magazines, records, microfilm, and even online databases.

The only drawback we saw in videodiscs was their slower access speed when compared to magnetic media. While virtually imperceptible to the user, we felt the necessity to overcome this constraint and were able to provide compensatory performance in the software.

In developing the videodisc, IAC provides the premastering studio, in our case LaserData of Cambridge, Massachusetts, with magnetic tapes containing the full database information including front-end search programming. LaserData processes the tapes through proprietary software that provides for error correction when data are read in the duplicated discs. Error control is vital because of the nature of the videodisc. Information will be "burned" onto the surface of the disc by a laser beam into fifty-four thousand frames, each containing 16,000 bytes of information (approximately sixteen thousand characters). Working at this microscopic level provides little tolerance for error.

LaserData then converts the digital data into a video signal and records it on videotape. This videotape master is then delivered to the 3M Corporation where the videodisc master is created. Pressings are made from this disc master to create the number of copies IAC orders each month.

Hardware and storage media decisions were largely resolved by early 1984, software development was under way, and the design of a multistation controller to act as the interface between up to four computers and the videodisc player was complete. Economics dictated that a single controller be used to direct the entire system. Economy was also the driving force in our four-system design. To make the product, which was eventually to be named InfoTrac, highly attractive in practical application, we believed we would have to provide the benefits of at least four COM readers, in addition to the search options possible with electronic access. InfoTrac was not designed as a replacement product for IAC COM indexes, but as a system to address an unserved market niche—the larger public or academic libraries that regularly have multiple COM indexes in use and are still not meeting patron demand either for periodical reference aid generally or for access to broad business topics specifically. We see InfoTrac as an addition to the product line that is complementary to our existing information delivery systems. COM, videodisc, and online delivery each address specific search needs and levels of searcher training.

Software development for InfoTrac continued well after we had placed the first systems in libraries for patron testing. The initial database offered on videodisc contained material from *Magazine Index*, *Business Index*, and *Legal Resource Index*,

with coverage, in some cases, from 1979 to the present. Much as an online searcher must, initial users of the InfoTrac system had to search each of the index databases sequentially rather than collectively. We purposely put the database and the software in use at test sites before either were completed in order to remain receptive to recommendations from users and librarians regarding the final product.

The test software did, however, contain the features we felt were most vital to the programming: the ability to "browse" a thesaurus to locate the most precise search terms, immediate display of citations under the selected subject, instant display of chosen *see also* references, simultaneous access by up to four users, the completion of a search in less than ten seconds, and immediate printout of relevant citations.

The initial InfoTrac system includes up to four IBM Personal Computers, up to four H-P Thinkjet printers (selected because of their speed and almost silent operation), a Pioneer disc player, a controller and, of course, the IAC-developed software. InfoTrac is not dependent on any specific hardware product and when (and if) more practical or cost-effective computers, printers, or storage media become available, they can be substituted for existing hardware relatively easily.

I believe it is important to emphasize that IAC's main business is not developing hardware systems. The company's prime focus is the development of information databases, which then demand supportive delivery systems that meet users' needs. InfoTrac is an example of the company's objective. A need was identified in IAC's existing markets that was not being met by current products, technology was examined to define how to meet that need most effectively and economically, and a delivery system was created with the involvement, each step along the way, of the library community and end users.

When InfoTrac was first shown at the ALA Conference in Washington in January 1985, it incorporated many of the recommendations gathered in empirical testing. The first systems, to be delivered to subscribers in March, will contain further refinements. For instance, the InfoTrac database will be completely integrated with material from *Magazine Index*, *National Newspaper Index*, *Business Index*, *Legal Resource Index*, and a substantial amount of business coverage never previously contained in an IAC database—more than one thousand publications in all with coverage from January 1982.

The computers in the InfoTrac system can be placed up to one hundred feet from the videodisc player, allowing access from different areas in a library reference section.

WHERE DO WE GO FROM HERE?

The questions we received from librarians in our first public demonstrations of InfoTrac inevitably centered on the possibilities of adapting the system to some specific need they had within their own library. For instance:

"Could a database made up of available information from IAC and other sources be created to meet our exact needs?"

Answer: Very definitely. However, the cost of such a custom database would be prohibitive for most libraries. If market research proves that the need is shared by a sufficient number of other libraries or companies, IAC will create and offer such an InfoTrac database.

"Can the InfoTrac system be expanded to contain more extensive or added databases?"

Answer: Yes, simply by adding additional disc players at a cost of approximately $1,000 for the player and supportive hardware the system can be expanded to contain up to 3.2 gigabytes of information (four disc players).

"Will it ever be possible to have remote access to InfoTrac over telephone lines so that all branches could use a single, central library system?"

Answer: This use is possible right now with the addition of another microcomputer to act as a file server. However, until we can validate that there is a market for remote access, we won't make the software modifications necessary to support remote access.

"Will you eventually put full-text files,

including graphics, on InfoTrac?"

Answer: At the speed at which technology moves, this may become an economic possibility; however, though it is now technically possible to record the full images of magazine pages on videodisc in black and white, the process and site hardware are too costly for most library applications. *Magazine Collection* and *Business Collection* are more practical, cost-effective methods for delivering full images of magazine pages.

"Microfilm, online, now videodisc—what can we expect next?"

Answer: Virtually whatever you want that can be delivered within your budget. Information Access Company's product plans grow out of the needs expressed by librarians and other information professionals. InfoTrac will continue to be refined to meet your expanding needs for information delivery, and other products will be developed as niches appear that are unfilled in our total Reference Network concept. ■■

Communications

OSU Libraries' Use of Library of Congress Subject Authorities File

Lorene E. Ludy

The Ohio State University Libraries (OSUL) applied the Library of Congress Subject Authorities File (SAF) to its online public access catalog (the Library Control System, LCS) in September 1984. This is thought to be one of the first broad-scale uses of machine-readable subject authority data for an online public access catalog.

THE LIBRARY OF CONGRESS SUBJECT AUTHORITIES FILE

The Library of Congress Subject Authorities File contains machine-readable records (in the MARC Authorities format, preliminary edition) for LC subject headings and associated references. There are records for name as well as topical headings. The Library of Congress' analysis of the 1982 SAF[1] shows the following types of headings (MARC field tag in parentheses) and what percentage of the file is accounted for by each:

Topical heading (150)	84.12%
Geographic Name (151)	10.22%
Personal Name (100)	4.03%
Corporate Name (110)	1.18%
Uniform Title (130)	.42%
Meeting Name (111)	.04%

References are recorded in 4xx and 5xx fields. There are *see-from* references (4xx fields) associated with 39 percent of the headings and *see-also-from* references (5xx fields) among 57 percent of the headings. Records that contain references average 1.8 references per record.

The SAF contains the same information as the cumulated microfiche editions of *Library of Congress Subject Headings* (LCSH): namely, established headings, references, scope notes, class numbers, explanatory references, and example under notes. The authority record for a heading does not, however, look like the entry in LCSH in that the redundancy of the printed list is eliminated. The reciprocal relations, *see/see-from* and *see-also/see-also-from*, are printed in both directions in LCSH (e.g., "LIBRARY AUTOMATION *see* LIBRARIES—AUTOMATION" and "LIBRARIES—AUTOMATION*x* LIBRARY AUTOMATION"). These relationships are recorded only once in the SAF. Figure 1 compares the printed LCSH entry for a heading with its machine-readable counterpart.

Some relationships between or among headings do not fit this structure. LCSH includes notes that explain a general relationship with headings in a certain pattern, e.g., "*sa subdivision* STUDY And TEACHING *under special subjects*. . . ." Explanatory *see-also* notes appear in the 360 field of the authority record. Explanatory *see* notes, however, are associated with *see* references that are recorded as 4xx fields in the SAF. To accommodate these notes, there are separate authority records for these nonauthoritative headings; these records comprise one-half of one percent (.52 percent) of the records in the SAF. (Figure 2 illustrates a general *see* reference and its SAF record.)

A new SAF is produced each year in conjunction with the production of the cumulated LCSH microfiche. The OSU Libraries used the 1981 SAF, which contained the headings from the ninth edi-

Lorene E. Ludy is authorities librarian, Ohio State Unviersity Libraries, Columbus.

Library education *(Indirect) (Z668-9)*
Here are entered works on the education of librarians. Works dealing with the instruction of readers in library use are entered under the heading Libraries and readers.
sa Children's librarians—In-service training
Interns (Library science)
Librarians—In-service training
Library employees—In-service training
Library orientation
Library schools
Library science libraries
subdivision Study and teaching *under special subjects, e.g.* Book selection—Study and teaching
x Education for librarianship
Librarians, Education of
Librarians, Training of
Library school education
Library science—Study and teaching
xx Education
Library schools
Professional education

```
053  Z668-9
150  Library Education
360  $i subdivision $a Study and teaching $i under special subjects,
     e.g. $a Book selection--Study and teaching
450  Education for librarianship
450  Librarians, Education of
450  Librarians, Training of
450  Library school education
450  Library science--Study and teaching
550  Education
550  Library schools
550  Professional education
680  Here are entered works on the education of librarians.  Works
     dealing with the instruction of readers in library use are
     entered under the heading Libraries and readers.
```

Fig. 1. Comparison of Printed LCSH Entry for Heading (Left) with Its Machine-Readable Counterpart (Right).

Buildings, Public
***See* Public buildings**
subdivision* Public buildings *under names of countries, cities, etc.

Fig. 2. A General See *Reference in LCSH (Top) and Its SAF Record (Bottom).*

```
SAF AUTHORITY RECORD

    130 Public buildings

    430 Buildings, Public


SAF REFERENCE RECORD  (identified by "b" in byte 9 of 008 field)

    130 Public buildings

    260 $i subdivision  $a Public buildings  $i under names of

      countries, cities, etc.
```

tion of LCSH and the supplements for 1979, 1980, and 1981. This was the most current SAF available in summer 1984 (The 1982 SAF was distributed in October 1984). The 1981 SAF contained 127,133 heading records (637 for nonauthoritative headings).

SUBJECT HEADINGS IN OSU LIBRARIES' ONLINE CATALOG (LCS)

Subject headings in LCS are stored in a headings file that was created by extracting all headings (author, subject, series, and

uniform title) from the machine-readable cataloging records produced by OSUL through OCLC from 1973 and author headings from all pre-1973 brief records.[2] Each heading has a record that stores information about the heading: i.e., whether the heading is authoritative (a correct heading or a reference), how frequently it appears in bibliographic records, its relationship to other headings (*see* and *see-also* references), and notes about the heading. These records are not in the MARC format, but contain much of the information found in MARC authority record. Unlike the MARC authority record, however, headings records on LCS display both *see* and *see-from* references (figure 3). In August 1984, LCS included more than 410,000 unique subject headings. The headings file serves as OSUL's authority file by showing which headings have been compared to and match an authority, such as LCSH. These authoritative headings are flagged "verified" in the headings file. Prior to applying the SAF, 21 percent of the subject headings in LCS had been verified, i.e., they had been found on the LC Name Authority Tape (names used as subjects) or had been used by the Library of Congress in current (AACR2) cataloging.

APPLYING THE SUBJECT AUTHORITY FILE

The Ohio State University Libraries used the SAF to add references and notes to LCS and to control subject headings on LCS by verifying those that are established Library of Congress headings. The SAF was applied using a matching program. This program identified those headings in the SAF that

```
Library education   (GEOG)  (1981)
02    11555   SUB: 43
03 (053) Z668-9
04 (680) Here are entered works on the education of librarians. Works dealing
           with the instruction of readers in library use are entered under
           the heading Libraries and readers.
07 SEE FROM: 5
08  1880017  Education for librarianship
09  1880019  Librarians, Education of
10  1880018  Librarians, Training of
PG1.  ENTER PG2 FOR NEXT PAGE.

pg2

Library education       11555
12 SEE FROM:
13  1880020  Library school education
14  1880021  Library science--Study and teaching
15 SEE ALSO: 5
16   220192  Interns (Library science) (1)
17    54673  Librarians--In-service training (10)
18  1352822  Library employees--In-service training (1)
19    48231  Library orientation (23)
20    11554  Library schools (1)
PG2.  ENTER PG1 FOR PRECEDING PAGE; ENTER PG3 FOR NEXT PAGE.

pg3

Library education       11555
22 (360) subdivision Study and teaching under special subjects, e.g. Book
           selection--Study and teaching
24 SEE ALSO FROM: 3
25     8484  Education (266)
26    11554  Library schools (1)
27    47904  Professional education (20)
PG3 END.
```

Fig. 3. Heading Record on LCS (Staff Display).

were also present in LCS, and only the authority records for those headings were added to LCS. There were 41,214 exact matches on topical subject headings (field 150).

There were 70,576 *see-also* links established between headings and 48,737 *see-from* links were established. OSUL's Main Library card catalog had contained *see* and *see-also* references among headings, but when LCS became the official catalog (July 1982), patrons had to consult LCSH to identify both established and related headings. This information has therefore once again been integrated into the catalog (figure 4). In addition, the scope notes and class numbers associated with subject headings on LCS are supplied, information that was never placed in the card catalog. (LCS has a call number "shelf position search" to allow its use as a classified catalog.)

LCS includes only subject headings (and their associated references) that have been actually used for bibliographic records in the online public access catalog. Because the headings file serves, in effect, as an index to the bibliographic records, it was not considered desirable to include unused headings, with two exceptions: (1) Unsubdivided headings were added when LCS included the heading with further subdivisions. This allowed the reference structure for the heading to be added to the catalog. This has created some blind references (i.e., references that refer to a heading with no associated bibliographic records). However, a search on the unused heading will cause the further subdivided heading to display. (2) *See-also-from* headings (5xx fields) in an SAF record whose 1xx field matched an LCS heading were added to LCS. In either of these cases, a patron searching the heading is given pertinent related information even though the exact heading does not

```
 01       Library duplicates
 02          SEARCH UNDER:  Duplicates in libraries
 03       Library economy
 04          SEARCH UNDER:  Library science
>05    43 Library education       *(SEE BELOW)
 06     4 Library education--Addresses, essays, lectures
 07     1 Library education--ASIA
 08     1 Library education--AUDIO-VISUAL AIDS
 09     1 Library education--AUDIO-VISUAL AIDS--BIBLIOGRAPHY
 10     1 Library education--CANADA
ENTER TBL/line no. FOR TITLES.  *ENTER SAL/line no. FOR MORE INFORMATION.
ENTER PS- FOR PRECEDING PAGE; ENTER PS+ FOR NEXT PAGE.

sal/5



Library education
   POSSIBLE BROWSING NUMBER(S):  Z668-9  (SEARCH WITH SPS/)
   Here are entered works on the education of librarians. Works dealing with
      the instruction of readers in library use are entered under the heading
      Libraries and readers.
   SEARCH ALSO UNDER:
      Interns (Library science) (1 TITLE)
      Librarians--In-service training (10 TITLES)
      Library employees--In-service training (1 TITLE)
      Library orientation (23 TITLES)
PG1.  ENTER PG2 FOR NEXT PAGE.
ENTER PS1 TO RETURN TO LIST OF SUBJECTS

pg2


Library education
   SEARCH ALSO UNDER:
      Library schools (1 TITLE)
   SEARCH ALSO UNDER:  subdivision Study and teaching under special subjects,
      e.g. Book selection--Study and teaching
PG2 END.
ENTER PS1 TO RETURN TO LIST OF SUBJECTS
```

Fig. 4. See *and* See-also *References on LCS (Public Display).*

appear in a bibliographic record on LCS.

OSUL has also applied the LC Name Authority Master and update tapes to LCS.[3] These tapes control matching name headings in LCS, including names used as subjects. Therefore, names in the SAF (100, 110, 111, 130, and 151 fields) were treated separately from topical headings, i.e., references were added but the headings were not verified. This prevented pre-AACR2 names in the SAF from replacing valid AACR2 names already present in LCS. All names from LCSH (such as United States, Shakespeare, etc., with all their subdivisions) will be verified individually by OSUL staff. At the time the Name Authority Tapes were applied to LCS, the subject heading fields on those tapes (*see-also-from* references—field 550—and names further subdivided by subject subdivisions—$x, $y, $z) were not added. These were added to LCS when the SAF was applied, providing *see-also* links between subjects and names (for example, MUSEUMS—GERMANY, WEST *see also* MARTIN-VON-WAGNER-MUSEUM).

SUBDIVISIONS

Subject heading control is further complicated by the use of subdivisions. Although LCSH is a list of established authorized headings, the majority of the subject headings in any large library's catalog consist of established headings with further subdivisions, the combination of which cannot be explicitly represented in LCSH. For example, the SAF contained 127,000 headings, whereas there were 410,000 subject headings used in LCS. There are distinct authority records in the SAF for headings with subdivisions when they so appear in LSCH; and nearly one-third of the headings in the SAF do include further subdivisions. However, there are more than 500 authorized free-floating subdivisions, each of which can be combined with many established headings. In addition, geographic names may be used as subdivisions with 70 percent of headings, making the number of unique headings highly expandable.

In order to verify subdivided headings, in addition to those found in LCSH, a list of seventeen common subdivisions was devised. Any LCS heading that consisted of a main heading on the SAF and a listed subdivision was verified as the SAF was applied to LCS. This doubled the number of headings that were automatically verified. However, as a consequence, some invalid headings may be verified on LCS; for example, United States was on the list, so any LCS heading that matched an SAF heading subdivided by United States was verified, even if that heading was not one that may properly be subdivided geographically. It was judged that the benefits of machine matching and the large number of headings so verified far outweighed this disadvantage.

The overlap between topical subject headings on the SAF and LCS indicates the extent to which subdivisions are used: one-third of the headings in the SAF exactly matched one-tenth of the headings in LCS. Another 10 percent of the headings in LCS consisted of the established SAF/LCS heading with one of the seventeen subdivisions; about 44 percent additional headings consisted of the established SAF/LCS heading with subdivisions other than the seventeen chosen. Thus, even though only 10 percent of the topical subject headings on LCS exactly matched established LCSH headings, those established headings were the basis of more than 50 percent of the subject headings found on LCS. It has been found, since fall 1984, that 75 percent of subject headings new to LCS are, in fact, previously used headings with new subdivisions. These can be routinely verified by staff using the online authority data on LCS.

OUTDATED DATA

There were some problems because the data in the SAF were necessarily not as current as some subject headings then present in LCS.

Before applying the SAF, the 1982 printed *Supplement* to LCSH was reviewed and all headings that had been cancelled were deleted from LCS or established as *see* references to the new heading. For example, ANIMALS, FOOD HABITS OF was replaced by ANIMALS—FOOD, and the superceded heading removed from

the list. In contrast, DIVORCEES was replaced by DIVORCED WOMEN, but the old heading remained in the file as a *see* reference to the new heading. When the SAF was applied, many of the removed headings were re-added to LCS because they were part of the reference structure of other headings. ANIMALS, FOOD HABITS OF reappeared as a *see-also* reference to CARNIVORA and FOLIVORES, even though all the titles on this subject had the correct heading ANIMALS—FOOD. Local programming specifications prevented any changes to existing *see* references in LCS. Thus DIVORCEES remained a *see* reference even though there was an SAF authority record for this heading. The 1982 *Supplement* is being reviewed manually once again to incorporate such additions and changes into LCS. It is clear that some subject headings in LCS will necessarily be more current than any cumulated SAF used to control them. This presents a challenge to provide ongoing control of current headings.

One of the benefits of using the Name Authority Tape was the global change that was made possible when an LCS heading matched a *see* reference. This also happened with the application of the Subject Authorities File: 1,039 subject headings in LCS (.25 percent) matched 4xx fields in SAF authority records. The 4,234 bibliographic records showing these headings were automatically changed ("flipped") to show the established heading (1xx field). A printed report documented all such "flips" for staff review. It was found that approximately 3 percent of the flips were erroneous; for example, headings that had been *see* references in 1981 and had since been established; or references referring to more than one heading, requiring a distinction among individual bibliographic records. Such errors had been anticipated and were corrected manually. In addition, approximately one thousand headings in LCS consisted of a heading that matched a 4xx field with a further subdivision. A new heading consisting of the established heading with the subdivision was created. The bibliographic records associated with these headings were not flipped. All were reviewed manually to determine the appropriateness of the subdivision to the new heading and flipped individually.

SUMMARY OF THE LC SUBJECT AUTHORITY FILE (SAF) LOAD ONTO LCS

1. Heading records on the 1981 SAF: 127,133
2. Subject headings on LC at time of load: ca. 410,000
3. Exact matches (topical headings only) between SAF and LCS: 41,214 (33% of SAF headings, 10% of LCS headings)
4. *See-also* reference links established: 70,576
5. *See* reference added to LCS and links created: 48,737
6. Headings added to LCS (*see-also* references, "grandfathers"): 56,531
7. Headings verified (exact matches and exact matches with a listed subdivision): ca. 81,300
8. Headings "flipped"
   automatically (during SAF application): 1,039
   manually (after SAF application): 1,089

Applying the Library of Congress Subject Authority File has brought much improved control to subject headings in OSUL's online public access catalog. The number of verified subject headings in LCS has more than doubled, and the data are recorded to evaluate more easily and control headings on an ongoing basis. Most importantly, the addition of more than 100,000 reference links, scope notes, and suggested call number searches will assist catalog users in their subject searches.

REFERENCES

1. James L. Godwin, "Descriptive Tabulation: Library of Congress Internal MARC Format Display: 1982 Masterfile Subjects" Internal report of Automated Systems Office, Library of Congress, Nov. 1984.
2. Lorene E. Ludy and Susan J. Logan, "Integrating Authority Control in an Online Catalog," *Proceedings of the ASIS Annual Meeting* 19:176–78 (1982).
3. Lorene E. Ludy, "LC Name Authority Tapes Used by Ohio State University Libraries," *Information Technology and Libraries* 3:69–71 (Mar. 1984). ■■

The Evolution of an Online Union Catalog: Impact of User Feedback

Kathryn A. Hammell and Kay Goldberg

INTRODUCTION

In 1982, the Midwest Health Science Library Network (MHSLN) began operating a regional online catalog and interlibrary loan system with funding from the National Library of Medicine.[1] The creation of this system gave critical support to MHSLN members who are responsible for providing health science practitioners with timely, convenient access to biomedical information sources.[2,3]

Because of the vital nature of this mission, it was important to develop a regional system that could be easily understood and that functioned in a manner acceptable to network users. Several authors, among them Penniman and Dominick,[4] Arnovick and Gee,[5] and Cochrane,[6] discuss the need to include evaluation of the system by its users in guidelines for system design. In an attempt to assess user needs, the project coordinators monitored system data as well as solicited user feedback during a six-month trial period. This article focuses on the extent to which user feedback impacted on system modifications for the Online Catalog and ILL System.

SYSTEM FEATURES

Bibliographic Retrieval Services (BRS) created the Online Catalog (CMHS) from previously existing machine-readable records in MARC formats contributed by MHSLN Resource and Basic Health Science Libraries. Any duplicate records were eliminated by running each institutional file against the others. BRS reformatted the data within each record into fields known as paragraphs.

All data in a CMHS record is fully searchable and will display in full when a print command is entered. The level of retrieval is dependent upon the operator that is entered. The standard operators of "AND," "OR," and "NOT" specify retrieval at the document or word level and the positional operators of "SAME," "WITH," and "ADJACENT" specify retrieval at the paragraph or sentence level in addition to word level. A sample record is displayed in figure 1.

The Online Catalog ILL System provides creation, transmission, and referral of requests with a single entry of user data. The process begins when the user has located the desired title in the CMHS database. A command is entered using the search statement and document number of the title that was retrieved. The system automatically captures, but does not display, data from the accession number, location, call number, main entry, title, and imprint paragraphs. The system then prompts the user to supply type of institution entering the request, bibliographic type of the title requested, copyright compliance, patron name, patron type, authorizing name for the request, and the date by which material is needed. The user may enter comments and supply a regional or national institutional identification code.

The system routes the request to the two libraries that are nearest and smallest, first within the state and then within the remaining states in the region. It routes to two regional locations only based on regional protocols. If the second library cannot fill the request, the system will refer it to the National Library of Medicine. Collection size was used to determine library size. The system will follow the same routing process by default if the lender does not act upon a request within seven calendar days.

The daily transaction file shows the status of requests that the library has entered. It must be updated each time a transaction is completed. The borrowing library updates its file to indicate receipt and return

Kathryn A. Hammell is head, Acquisitions Department, Library of the Health Sciences, University of Illinois at Chicago, and Kay Goldberg is program administrator for communications applications, Medical Information Network, American Medical Association, Chicago, Illinois.

```
LO M1.  M2.  M3.  M7.  MA.
CN IAXC: WG-590-M626-1978.  B: B: 617-481-In86-1974.  IAVM:
   WG-590-M626-1978.  GZHA: WL-355-I604m-1974.  HSL: RD594-2-M52.
HO IAXC IAXP.  B.  IAVM.  GZHA.  HSL.
TI Microvascular anastomoses for cerebral ischemia / Jack M. Fein and
   O. Howard Reichman, editors.
IM New York: Springer-Verlag, c1978.
CO xv, 324 p.: ill.; 28 cm.
NT Includes bibliographies and index.
AE Fein, Jack M.  Reichman, Owen Howard, 1932-.
SU LC: Cerebral-ischemia; Surgery, Congresses.  LC: Cerebral-arteries;
   Surgery, Congresses.  LC: Microsurgery; Congresses.  LC:
   Cerebrovascular-disease; Surgery, Congresses.  NLM:
   Cerebral-ischemia-Transient; Congresses.  NLM:
   Arteriovenous-anastomosis; Congresses.  NLM: Microsurgery;
   Congresses.
NO LC: 78001977.  ISBN: 0387902406.
NU DDC: 617-481.  NLM: W3-IN919M-2d-1974m.  LC: RD594-2-M52.
FF YR: 1978.  CP: nyu.  LG: eng.
CS MNU-B.
AN 333219.
```

Fig. 1. Sample Online Catalog Record.

of the loaned item and the lending library updates its file to note the date the returned item is received.

USER SELECTION AND TRAINING

Participants for the trial period of use were selected from among interested Medline Centers in Illinois, Indiana, Iowa, Minnesota, North Dakota, and Wisconsin. Fifty libraries were chosen and representatives from each were invited to attend a one-day training session. BRS Basic System training and specialized training on the operation of the ILL System were offered. The *Database Guide and Interlibrary Loan Manual* was distributed at each training session. The *BRS Basic System Training Syllabus*, a handout on creating a dictionary file, and the CMHS Search Key were also distributed. Participants were encouraged to communicate questions and suggestions to the project coordinator throughout the trial period.

USER FEEDBACK

Actively soliciting user feedback on system features and functions is a rational approach to designing a system that will satisfy user needs. This is underscored by Oliver[7] who suggests that within six months of operation, users are identifying new functions for a system. In addition, testing a system on a trial basis has the advantages of allowing modifications and enhancements to the system before it becomes available to a broader market.

The project coordinator planned to have users of the Online Catalog and ILL System complete a monthly log sheet during the trial period. The log sheet requested information on the number and type of searches performed and allowed for general comments (figure 2). A formative evaluation was to be administered near the end of the six-month period as a way to obtain feedback on all system features and functions. Scriven[8] indicates that a formative evaluation is an important tool for assessing the need for modifications of a trial system. The evaluation of the Online Catalog and ILL System would prove to be significant in determining the final design and use of the system.

The authors analyzed the Online Catalog, ILL System, and training program for the factors that were essential to the evaluation. The evaluation objectives were (1) to determine user satisfaction with the Online Catalog, the ILL System, and the training

MIDWEST HEALTH SCIENCE LIBRARY NETWORK
REGIONAL ONLINE UNION CATALOG OF MONOGRAPHS AND AUDIOVISUALS

LOG SHEET FOR THE MONTH OF: _____

Institution: _____

_____

_____

Total Number of CMHS Searches performed _____ Problems Encountered (please use reverse)

Number of Author Searches: _____

Number of Title Searches: _____

Number of Call Number Searches: _____

Number of Subject Searches: _____

Number of Free Text Searches: _____

Other Searches (Please Specify): _____

_____

_____

_____

Number of Items Located in CMHS: _____ Problems Encountered (please use reverse)

Number of ILL's Requested: _____

Days Required to Fill These Requests:

0	_____	7	_____
1	_____	8	_____
2	_____	9	_____
3	_____	10	_____
4	_____	11-14	_____
5	_____	15-21	_____
6	_____	22+	_____

Number of Requests Not Filled (After Referral) _____

Fig. 2. Online Catalog Monthly Log Sheet.

program and (2) to identify additional modifications and enhancements to the system and the training that had not yet been suggested by users. A six-page questionnaire was mailed to the 103 system users along with a cover letter explaining the purpose of the survey. A self-addressed envelope, coded to monitor user responses, was also included. A follow-up mailing to nonrespondents was done after the first deadline. After the second mailing, a total of 76 usable questionnaires had been returned for a response rate of 74 percent.

Questions covered each of the three components being evaluated: training, the CMHS database, and the ILL System. Users were asked to evaluate their satisfaction with features on a five-point scale or with yes or no answers. Suggestions for modifications came from the open-ended component of each question.

In the section on training, users were asked to evaluate the training program and the instructional materials that were distributed. Although the majority rated BRS Basic System training and ILL System training as useful or somewhat useful (table 1), over half of the respondents urged that the training program be modified. Hands-on experience or at least an online demon-

stration of each command, transaction, strategy, etc., were overwhelming concerns. The respondents urged that separate instruction be conducted for inexperienced searchers. In the area of training materials, respondents requested that more examples of transactions at each stage in the ILL process be prepared and distributed. Respondents were also interested in obtaining a brief command chart, additional examples of search techniques, detailed instructions on BRS commands, and sign-on and sign-off procedures.

Users were asked to evaluate the format and content of the bibliographic record, searching and printing capabilities, and online time in the section on the Online Catalog (table 2). The majority were very or somewhat satisfied with the bibliographic content of the record. Including the circulation status of each title in the record, moving the location (LO) and call number (CN) paragraphs to the end of the record, and replacing the director's name with the name of the department responsible for interlibrary loan were suggested. An overwhelming majority of respondents were very or somewhat satisfied with the time it takes to verify and locate a title on CMHS, and slightly more than half were convinced that full free text search capability of all data in the record had improved their retrieval.

In the ILL section, users were asked to evaluate the work form and commands, ILL processing, online time, and turnaround time (table 3). The majority were very or somewhat satisfied with the bibliographic citation, the requester data, and the lender data. Many respondents commented negatively on the length of the ILL request form. Fewer than half were very or somewhat satisfied with the format of the transaction file. Approximately two-thirds of the respondents were very or somewhat satisfied with the ILL commands, but some respondents felt that they were too long and contributed to user error. The majority were very or somewhat satisfied with the online time necessary to initiate a request, acknowledge a received loan, and indicate when the requested item is returned. Lending library respondents expressed similar satisfaction with the online time to fill a request, refer a request, and acknowledge that the returned loan had been received (table 4). Rather than remarking upon a reduction in turnaround time, respondents noted that using the Online Catalog increased the number of titles that could be verified and located and allowed them to inform a patron quickly if a title was available within the region (table 5). The majority of respondents were very or somewhat satisfied with routing within the state and then within the remaining states in the region (table 6). The convention of routing smallest to largest generated the least support. A number of respondents urged that the user be able to select the library to which the request would be routed or at least be able to override system routing.

SYSTEM MODIFICATIONS

Many of the modifications suggested by users on the log sheets were adopted prior to the administration of the evaluation questionnaire. For example, the notation for "Date Requested" was changed to "Date Needed by" in the ILL request form to reflect more accurately the nature of the data. The requester and lender portions of the request form were single spaced to reduce the size of the printed form. The renewal function was modified to permit the user to enter "No" at the renewal prompt instead of using the carriage return as the default. Additional action codes for title not owned (TNO) and part not owned (PNO) were introduced to allow the National Library of Medicine to reply accurately to a request. The reason for referral was also added to the information displayed when a request was referred. Other suggestions were addressed but not implemented. For example, the ILL System accepts only the current date during a transaction update to avoid introducing errors in the statistical compilations. The suggestion to print only new requests instead of all outstanding requests could not be accommodated due to the high cost of reprogramming. Users also suggested that the bibliographic portion of the request be compressed, but this could not be implemented because of the variable length of titles and statements of authorship.

Respondents to the evaluation question-

Table 1. Usefulness of Training Session and Instructional Materials (N = 76)

	Definitely Useful		Somewhat Useful		Not Very Useful		Definitely Not Useful		Did Not Attend		No Response		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
BRS Basic System training	26	34.2	19	25	6	7.9	8	10.5	12	15.8	5	6.6	76	100
CMHS ILL System training	30	39.5	21	27.6	7	9.2	1	1.3	12	15.8	5	6.6	76	100
CMHS Database Guide	47	61.8	15	19.7	4	5.3	0	0	4	5.3	6	7.9	76	100
CMHS ILL Manual	52	68.4	13	17.1	4	5.3	0	0	3	3.9	4	5.3	76	100
Monthly Updates	26	34.2	20	26.3	7	9.2	0	0	14	18.4	9	11.8	76	100
BRS Syllabus	25	32.9	26	34.2	5	6.6	2	2.6	8	10.5	10	13.2	76	100

Table 2. Satisfaction with the Online Catalog Bibliographic Record (N = 76)

	Satisfied		Not Satisfied		No Response		Total	
	No.	%	No.	%	No.	%	No.	%
Display of bibliographic data	65	85.5	2	2.6	9	11.8	76	100
Arrangement of data	65	85.5	3	3.9	8	10.5	76	100
Improvement in retrieval	45	59.2	12	15.8	19	25	76	100
Time to verify and locate title	62	81.6	5	6.6	9	11.8	76	100

Table 3. Satisfaction with ILL System Formats and ILL Commands (N = 76)

	Very Satisfied		Somewhat Satisfied		Somewhat Dissatisfied		Very Dissatisfied		Do Not Know		No Response		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
ILL Request														
Bibliographic citation	39	51.3	17	22.4	0	0	1	1.3	5	6.6	14	18.4	76	100
Requester data	23	30.3	23	30.3	6	7.4	1	1.3	10	13.2	13	17.1	76	100
Lender data	24	31.6	17	22.4	9	11.8	1	1.3	11	14.5	14	18.4	76	100
ILL Transaction File	20	26.3	17	22.4	8	10.5	2	2.6	14	18.4	15	19.7	76	100

Table 4. Satisfaction with Online Time

	Very Satisfied		Somewhat Satisfied		Somewhat Dissatisfied		Very Dissatisfied		Do Not Know		No Response		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Borrowing Library (N = 76)														
Initiating a request	38	50	16	21.1	1	1.3	0	0	8	10.5	13	17.1	76	100
Acknowledging receipt of requested item	26	34.2	15	19.7	4	5.3	2	2.6	16	21.1	13	17.1	76	100
Indicating return of requested item	25	32.9	15	19.7	5	6.6	2	2.6	16	21.1	13	17.1	76	100
Lending Library (N = 20)														
Filling a request	11	55	3	15	1	5	0	0	2	10	3	15	20	100
Referring a request	8	40	3	15	2	10	0	0	3	15	4	20	20	100
Acknowledging receipt of returned item	9	45	4	20	0	0	1	5	2	10	4	20	20	100

Table 5. Effect of Online Catalog ILL System Use on ILL Processing

	Significant Increase		Some Increase		No Difference		Some Decrease		Significant Decrease		Do Not Know		No Response		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Borrowing Library (N = 76)																
Rate of Fulfillment	8	10.5	9	11.8	38	50	2	2.6	0	0	7	9.2	12	15.8	76	100
Turnaround time (Fill speed)	14	18.4	14	18.4	20	26.3	6	7.9	1	1.3	9	11.8	12	15.8	76	100
Lending Library (N = 20)																
Rate of Fulfillment	3	15	5	25	6	30	0	0	0	0	2	10	4	20	20	100
Turnaround time (Fill speed)	3	15	3	15	8	40	0	0	0	0	3	15	3	15	20	100

Table 6. Satisfaction with ILL System Processing (N = 76)

	Very Satisfied		Somewhat Satisfied		Somewhat Dissatisfied		Very Dissatisfied		Do Not Know		No Response		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Routing smallest to largest	24	31.6	16	21.1	7	9.2	4	5.3	10	13.2	15	19.7	76	100
Routing within state first	34	44.7	14	18.4	0	0	3	3.9	11	14.5	14	18.4	76	100
System referral after 6 days	24	31.6	19	25.0	7	9.2	0	0	12	15.8	14	18.4	76	100
Two referrals maximum within region	22	28.9	21	27.6	7	9.2	1	1.3	11	14.5	14	18.4	76	100

naire suggested that changes be made in the organization of the training program, the development of additional instructional materials, and the single method of routing possible within the system. The *Database Guide and Interlibrary Loan Manual* was revised to include computer-generated illustrations of each search technique and the use of each ILL transaction command. An appendix of internal holdings codes was compiled and added to the *Database Guide* to indicate circulation status of materials. The training program was completely revised to include exercises, demonstrations, and hands-on use in consecutive half-day sessions on the Online Catalog and the ILL System. Finally, the routing algorithm was modified by BRS to permit users to override system routing and to select the two regional locations to which the request would be routed.

SUMMARY

Although users may at times be overlooked in focusing on the technical aspects of system design, they are the reason for the existence of an information system, as recent studies have emphasized.[9]

Lancaster notes that feedback from user evaluations oftentimes focuses on ease of use and accessibility.[10] Evaluation responses from MHSLN members yielded similar results. As a result of suggestions from the log sheets and the formative evaluation, a significant number of modifications were made to improve the functioning of the overall system.

REFERENCES AND NOTES

1. The Online Catalog is supported by National Institutes of Health contract number NIH-N01-LM-3502. Irwin H. Pizer, RML director and university librarian for the Health Sciences, University of Illinois at Chicago, is project director. Kathryn A. Hammell was formerly project coordinator for the Online Catalog and Kay Goldberg was formerly program coordinator: Evaluation and Design for the Greater Midwest Regional Medical Library Network. When MHSLN was reconfigured to GMRMLN (Greater Midwest Regional Medical Library Network) to include the states of Kentucky, Michigan, Ohio, and South Dakota in January 1983, the original contract was extended and additional funds awarded to augment the Online Catalog with new holdings and subsidize new users.
2. Irwin H. Pizer, "An Online Regional Catalog of Monographs and Audiovisuals: A Demonstration Project," *INSPEL* 15:69–84 (1981).
3. Kathryn A. Hammell, "Regional Online Union Catalog of the Greater Midwest Regional Medical Library Network: Development and Operation," *Bulletin of the Medical Library Association* 72:155–61 (Apr. 1984).
4. W. D. Penniman and W. D. Dominick, "Monitoring and Evaluation of Online System Usage," *Information Processing and Management* 16:17–35 (1980).
5. George N. Arnovick and Larry G. Gee, "Design and Evaluation of Information Systems," *Information Processing and Management* 14:369–80 (1978).
6. Pauline Cochrane, "Improving the Quality of Information Retrieval—Online to a Library Catalog or Other Access Service . . . or . . . Where Do We Go from Here?" *Online* 5:30–42 (July 1981).
7. Dennis Oliver, "Electronic Mail Services in the Library and Information Center Community," paper presented at the Twenty-first Annual Clinic on Library Applications of Data Processing, Urbana, Ill., April 1984.
8. Michael Scriven, *The Methodology of Evaluation* (Lafayette, Ind.: Social Science Education Consortium, 1966).
9. Paul M. Anderson and Ellen G. Miller discuss user participation in automated bibliographic records at the University of Cincinnati in "Participative Planning for Library Automation: The Role of the User Opinion Survey," *College & Research Libraries* 44:245–54 (July 1983). Their concern with determining and satisfying user needs is echoed by Ray R. Larson and Vicki Graham, who report on a user survey administered online in "Monitoring and Evaluating MELVYL," *Information Technology and Libraries* 2:93–104 (Mar. 1983). The Council on Library Resources sponsored surveys of online catalog users in twenty-nine libraries, as reported on by Robert N. Broadus in, "Online Catalogs and Their Users," *College & Research Libraries* 44:458–67 (Nov. 1983). In "Surveying Users of Automated Turnkey Systems," *Information Technology and Libraries* 2:290–95 (Sept. 1983), Philip Schwarz counsels prospective users to seek the opinions of their peers before selecting a turnkey system.
10. F. Wilfrid Lancaster, *Information Retrieval Systems: Characteristics, Testing, Evaluation*, 2d ed. (New York: Wiley, 1979). ■■

We've taken book ordering out of the Dark Ages.

BaTaSYSTEMS™ is Baker & Taylor's newest generation of electronic book ordering services. It's especially designed to work with existing computer hardware, with built in flexibility that allows you to match the level of service to your library's unique needs.

Whichever service level you choose, you'll save time, reduce paperwork and speed book acquisitions—all at a lower cost. For example:

ORDER allows you to order books through your personal computer, using a modem and regular telephone lines. Just enter the ISBNs and the following day you'll receive electronic confirmation from which you can print order slips. All calls are toll free. You also save the cost and delay of postal delivery.

Or you can choose SEARCH AND ORDER. In addition to electronic ordering, this service gives you quick access to Baker & Taylor's diverse and comprehensive database of over 800,000 title records. It's your single source for virtually all the titles published or distributed in the United States. And you eliminate manual searching and purchase order typing.

Finally, BaTaSYSTEMS ACQUISITIONS offers on-line access to our database and electronic ordering plus a complete software package with fund accounting and full reporting functions.

These advanced service technologies are typical of how Baker & Taylor stays in step with the times, building on our experience to bring you the latest in library services.

BaTaSYSTEMS. It's nothing less than a renaissance in book acquisitions. Write or phone today for more information.

EXPERIENCE YOU CAN DEPEND ON
BAKER & TAYLOR
a GRACE company

Eastern Division, 50 Kirby Avenue, Somerville, NJ 08876 (201) 722-8000
Southern Division, Mt. Olive Road, Commerce, GA 30599 (404) 335-5000
Midwestern Division, 501 S. Gladiolus Street, Momence, IL 60954 (815) 472-2444
Western Division, 380 Edison Way, Reno, NV 89564 (702) 786-6700

Reports and Working Papers

1984 Automated Authority Control Opinion Poll: A Preliminary Analysis

Barbara B. Tillett

This paper was presented to the LITA ISAS/RTSD CCS Discussion Group on Authority Control in the Online Environment at the ALA Midwinter Meeting in Washington, D.C., on January 7, 1984.

I am reporting as the chair of the LITA ISAS/RTSD CCS Discussion Group on Authority Control in the Online Environment, which had its first meetings at ALA in Dallas last summer. Prior to our initial meetings I distributed the automated authority control opinion poll, which I will refer to as the LITA poll or the 1984 opinion poll. The objective of the opinion poll was to determine current opinion of U.S. librarians on the issue of future computerized systems for authority control to generate topics for discussion.

I started to write this talk as a lesson in how not to do a study, but instead will just give you some insight as to how the opinion poll came about before I present the results. I have had several opportunities to present the results from this poll, the first being a brief apology at the Dallas meeting for being overwhelmed with the responses and in the midst of processing the data. The second being a presentation of the preliminary results at the Illinois Library Association Conference of the Resources and Technical Services Section in October. And a third opportunity was the presentation of this paper at ALA Midwinter 1985 in Washington, D.C. I have prepared the results of the LITA poll in comparison with the Council on Library Resources' 1982 study, "The Name Authority Cooperative/Name Authority File Service," which I will refer to as the CLR survey.

When I began thinking about collecting opinions on authority control, I had contemplated doing a Delphi study, contacting the experts on the topic and then reviewing their responses for a second response, to come to some consensus. Another thought was to prepare a questionnaire for a perfectly selected sample of librarians. Still another method was an opinion poll or a more open-ended approach to give people a chance to express their hopes and desires and vent frustrations with existing systems. I rather liked the latter approach and prepared a draft instrument to use for a small survey of a few librarians.

After a brief review of the poll with several colleagues, I contacted LITA to see if I could arrange to have the poll distributed to a sample of their membership to test out the survey, and then I would use the results as an opener for the LITA/RTSD Discussion Group.

The 1982 CLR survey was sent to 263 libraries so I was anticipating a similar size sample for the 1984 LITA poll. Much to my surprise, LITA was more than generous and sent the poll to 1,776 members of LITA and RTSD's Cataloging and Classification Section. The poll even got some very unexpected publicity in the *LJ Hotline*, which caused another 75 people to write to me to request to be part of the survey. Mercifully, not all 1,851 people replied. However, as far as a return rate for a survey goes, I had only a little over 18 percent returned (the CLR survey had 69 percent returned), which would make the results some-

Barbara B. Tillett is head, Technical Services, Scripps Institution of Oceanography Library, University of California–San Diego, La Jolla, California.

what suspect from any statistical analysis point of view. And yet the responses totalled more than the original sample size I would have preferred and represented a broad range of libraries with nearly every size and type of library, representing forty-two states, the District of Columbia, and one from a U.S. Army library in Germany. It may be fairly representative of the population—and I can probably salvage the statistics, although I have not yet had time to check that out. Indeed the analysis is further complicated by the fact that the original sample that LITA used was sent out on mailing labels and appears to be impossible to replicate. I cannot get back to the nonrespondents to improve the return rate. Both LITA and I have learned from this experience!

In any case, what I have to present today is an analysis of responses from those people who chose to state their opinions. And, although at this point I cannot claim the results represent the population of the U.S. libraries, I believe the results are worth sharing.

The poll included two primary sections with a total of fifteen questions. For comparison, the CLR survey included thirty-six questions, both multiple choice and open-ended questions. The first section of the LITA poll gathered some general background on the libraries in which the respondents work, their practices in authority work, the use of authority files, the presence of an online catalog, and the level of staffing and type of authority work done. The second part of the poll was for opinion—ranging from comments on presently used online resource authority files to dreams for the ultimate authority control system. There was also a question about the possibilities for an international authority file and another on thoughts about requiring a single authorized form for future computer systems. I will focus on the background information and then share with you the responses on the questions relating to use of present online resource authority files and ideas about the ultimate authority control system. The responses to these and the other questions will be found in the final report.

Three hundred and thirty-four usable polls were returned with six others voided due to duplication or nothing written on the poll. Of those 334 respondents, the type of library worked out very closely to what one would expect based on membership in bibliographic utilities and membership in ALA:

- 67.6 percent academic (including community colleges, colleges, universities, research libraries, and the "special" libraries within universities) (190 acad., 8 med., 11 law, 8 other branch)
- 22.4 percent public (including secondary school libraries) (72 pub., 2 sec. schools)
- 5.95 percent government [federal (13–4.05 percent), state government (1—.3 percent), and state libraries (5—1.6 percent)]
- 4.05 percent corporate (13)

OCLC's membership runs about 66 percent academic, 21 percent special, and 13 percent public; the CLR survey was more heavily weighted toward academic and research libraries, with 86.3 percent in academic and research libraries and 13.7 percent in public libraries.

The use of bibliographic utilities by respondents to the poll also seems representative of the nation's libraries:

- 76 percent OCLC (253)
- 8.7 percent RLIN (29)
- 3 percent WLN (10)
- 5.7 percent other utility or vendor service
- 9.6 percent none (32)
- NOTE: Some of the libraries used more than one utility.

CLR's survey looked at the use of LC's Name Authority File on bibliographic utilities as one question and found 76 percent of the respondents used OCLC, the same percentage I found (18 percent no response, 2 percent UTLAS, 1 percent WLN, 1 percent multiple).

Looking at the presence of online catalogs, 18 percent of the libraries in the poll have online catalogs, that is 68 of the 334 respondents say they have an online catalog in operation. The majority of those libraries not yet online indicated they intend to have an online catalog within the next five years.

I have not yet figured the correlation between type of library and online catalogs, but for Illinois it turned out to be thirteen libraries (59 percent) not online, half academic and half public with one state library; and nine libraries (41 percent) online with six academic and three public. For Washington, D.C., it turns out to be one-third online, two-thirds not yet online—six libraries not online: three academic, one public, one special library, and one government library. The five respondents with online catalogs represent three libraries: two government and one special library with the SCORPIO/MUMS catalog, GEAC's SBIS catalog and the museum library with ILS.

Even if this 18 percent with online catalogs is an accurate reflection of the nation's libraries, it shows a tremendous increase over two years ago when the CLR-funded studies on online catalogs were conducted. Those studies looked at fifteen online catalog systems that were in use in twenty-nine libraries. Many more systems have come on the market in the last two years, particularly as derivatives of automated circulation systems. Some of the online catalogs reported on the LITA polls were in fact identified as derivatives of a circulation system. Typically the online catalogs lack files representing the library's full holdings since many types of library materials included in card catalogs are not yet included in the online catalogs, and nearly all online catalogs on the market today lack the full syndetic structure of a catalog. Several vendors of online catalogs claim to provide authority control, but just what they mean by authority control varies widely. Arlene Taylor, professor at the library school at the University of Chicago has been keeping track of the authority control features claimed by various vendors for their systems and intends to publish, if she hasn't already published it, a review article on the variety of systems now available, which she reported on at the Dallas meeting. You might want to keep an eye out for it.

The responses to the poll on the matter of authority files were very interesting. Jeannette Mosey, in her 1980 dissertation on name authorities in OCLC libraries, found about 75 percent of the OCLC libraries she studied had authority files. The CLR study showed 72 percent of the responding libraries maintained their own authority file. The LITA poll shows:

- 73.6 percent have authority files in card form separate from the card catalog (245)
- 13.2 percent have no authority files (44)
- 7.2 percent claim to have online authority files, but some indicate that was the authority file available in their bibliographic utility (24)
- 3.9 percent have microfilm authority files (13)

Of the remainder, eight libraries use their card catalogs for an authority file and six utilize other tools produced outside their library as an authority. It is interesting to realize that from 13 to 28 percent of the libraries in the United States do not have authority files, depending on which survey results you use. Reliance on the records from other libraries has made this option attractive to libraries, but the result is an incomplete catalog often lacking cross-references with maintenance of headings becoming a hit-or-miss activity since there is no record of which headings were used. I don't wish to debate whether or not authority files are necessary, or for that matter whether authority control is necessary, and would like to paraphrase Henriette Avram, that it is time to stop questioning the need for authority control and time to start doing something about it.

It is apparent that most libraries (at least 72 percent, using CLR's lowest figure of the three studies mentioned) do maintain their own in-house authority files, and we also know this is a great expense—R. Bruce Miller, now at Indiana University, in his presentation at the 1979 LITA Authority Control Institute estimated over five million dollars per year was being spent by academic libraries for authority control. The NACO project is of course a giant step forward in sharing the costs of authority work with the product being the shared national resource authority file. Having the resource authority file widely available through the various bibliographic utilities and on microfiche has been a great help for catalogers.

Both the 1982 CLR survey and the 1984 LITA opinion poll found a high use for the online resource authority files. On the LITA poll, 287 respondents (86.2 percent) indi-

cated they use an online resource authority file (12 percent no—40; 1.8 percent no answer—6).

This compares to the 1982 CLR survey, which showed 82 percent (150) used the LC Name Authority file as offered on any online system [14 percent (27) did not, with 2 percent (5) having no answer to that question]. You should note the CLR study only took into account the LC name authority file, while the LITA study looked at any available online resource authority file and so included more than the NACO file, in particular the WLN file and the various files on RLIN.

The comments from the poll on using the online resource authority files were very positive, as they were on the CLR survey. The overwhelming response on the LITA poll was elation at having such a resource and praise for a tool that was found to be extremely valuable and useful. Many people noted that it saves a lot of time and effort. In fact, one respondent noted that it was "such an improvement over any tool we've had in the past, that one hesitates to complain." But complain they did. Since most of the respondents used a resource authority file on some bibliographic utility, the comments included not only comments on the MARC authority records but also on the features for accessing and using the authority file on the utility.

OCLC USERS

The LC Name Authority file as loaded on OCLC was generally perceived as useful, but with the following frustrating problems (in the order of frequency from the polls):

- incomplete file
- duplicate and conflicting records
- OCLC searching problems
- maintenance problems
- format problems
- pre-AACR2 forms
- lack of products

Let me explain each of these categories with excerpts from the polls.

Incomplete File

"Not extensive enough," not big enough—the respondents wanted the file to contain everything. They wanted both older and really current names included—wanted it to include more authors in general. Several libraries indicated that they dealt with older materials, rare books, gifts, etc., and were not finding most of the names on the resource authority file; the same complaint came from libraries doing really current material, particularly for local and regional corporate bodies. They wanted a way to be able to add those government headings to the resource AF to save the time and effort of others in their region that would need to establish the same heading. They saw NACO as a tremendous help in the right direction, but felt it should be expanded—the quote is "NACO is a good start, but too limited and too slow for local/regional headings."

Poor international coverage for personal names (especially for Spanish, Puerto Rican, and Russian names) was noted.

With several libraries that cataloged music, one library complained there were too few jazz composers.

They would like more geographic names.

They want subject authorities (this was a very popular response!).

"LCNAF microfiche record number does not have a corresponding record online"—this response meant an empty or blank record results when searching online using a record number found on the microfiche.

The online file lacks the authority records for referenced *see also* names, and it lacks authority records for parent bodies when the subordinate bodies have been established.

The file is reportedly not current enough. The respondents wanted more frequent updates, preferably online with constant updating with the NAFS/LSP.

Duplicate and Conflicting Records

"Get rid of duplicate records"—this is not the OCLC duplicate record complaint but rather a complaint about the LC authority file. There were several complaints about multiple records for the same entity. There were reportedly conflicts between LC's bibliographic records and the authority records.

The respondents hoped these duplicates and conflicts would be alleviated with loading the new LC authority tape in the summer 1984, but my staff are still finding this a problem with the LC Name Authority File. For example, we had a problem with two personal names, both having two official, conflicting authority records on the LC Name Authority File, and this is becoming more and more noticeable with a recent case of LC and a NACO library entering conflicting records for the same corporate body.

OCLC Searching Problems

Many respondents like the ease with which one can get to the authority file and back to the bibliographic record being worked on, but they wanted more.

It was apparent that the search key approach on OCLC is too limited, particularly for corporate hierarchies, common personal names, and author/title entry uniform titles. The respondents indicated that they must augment the online search with a search of LC's NAF on microfiche, especially for corporate bodies with subordinate bodies and for author/title entries.

OCLC's limit of 256 records for a search is inadequate, particularly for corporate bodies. Examples most often cited were federal departments, primarily the Environmental Protection Agency; institutions, such as Akademiîa nauk SSSR; and for musical composers, J. S. Bach was the prime target of most respondents.

Suggestions made for improving searching included keyword search on OCLC with boolean capabilities, a better way to qualify a search, including limiting the results of a search to omit or include cross-references as desired.

They would like a browsable file, especially for author/title entry uniform titles.

There were conflicting responses with respect to the stoplist. Some respondents said it often makes searching difficult, while others wanted to increase the stopword list.

The respondents would like links from the authority file record to the bibliographic files that incorporate the heading in order to verify the relationship and insure proper assignment of the heading. Also the respondents would like an easy method or commands to continue a search for related authority records for complex heading searches, such as hierarchies, earlier/later forms, etc. (This response did my heart good, since this is the area I am examining for my Ph.D. dissertation.) The present system requires a cumbersome research for each separate heading.

Maintenance Problems

"The most annoying problem is that LC changes its mind so often, necessitating changes upon changes in authority records and catalog cards." Respondents want to be able to identify changes and have the bibliographic utility notify people who have used the form that is later changed.

Many librarians responded that they want to be able to indicate when they've used an authority record, e.g., by attaching their holding library code. Two respondents suggested that with the holding code on authority records, they could let the OCLC file be their online authority file.

They want an easy/fast method to update and resolve conflicts.

Format Problems

The MARC authorities format is difficult for several respondents to read and they would like an easier display in straightforward English.

There were lots of problems reported in deciphering the w subfield information. This

was before the change from thirteen characters to four. In general, the respondents would like a more explanatory format rather than so much coded information in w subfields.

Pre-AACR2 Forms

The respondents would like all authority records to reflect AACR2 headings and references with an added reference for the pre-AACR2 form (several libraries made this same remark). In other words, they want LC's file to be all AACR2 plus an added *see* reference for the pre-AACR2 form. They want the AACR2 and non-AACR2 headings and references to be clearly identified as such and not buried in a coded subfield.

There was some frustration with unauthenticated references, which mean more cleanup for catalogers.

Lack of Products

They want to be able to produce cross-reference cards as they now are able to produce catalog cards, and they want the capability to load authority records onto their own archival tape for local use.

RLIN USERS

The comments from the users of the resource authority file on RLIN primarily complained about how very out-of-date the file was, which was updated at the end of May, after most people responded to the poll. As a consequence there were complaints that now would differ for RLIN users.

There were complaints about the incompleteness of the file and complaints about headings not in AACR2 form, as there were for OCLC users. One respondent complained that RLIN's file lacks a full syndetic apparatus, but I am not sure whether that meant no access was then available to the cross-references in the authority file or exactly what the problem was.

Users with experience on OCLC would have preferred OCLC's ability to easily move back and forth between bibliographic and authority records—RLIN requires a cumbersome log off and log on procedure to search the authority file.

Frequent errors in the cross-references were mentioned, and the RLIN users felt it was difficult to interpret the coded record. They would like a more word-based display, a comment also made by OCLC users.

WLN USERS

The WLN users loved WLN's authority files for the most part, but a few complaints were made about the problems with multiple entries, especially for series in both AACR2 and pre-AACR2 forms. They liked the inclusion of the subject authority records—one person even commented that the subject authorities were the best part of WLN's files. One user had problems with unclear coding for references.

Several users complained that WLN lacked many of LC's name authorities but expected the inconvenience to be temporary with progess on the Linked Systems Project. Those of you who attended the NACO meeting during ALA Midwinter heard WLN should be operational in LSP in the fall of 1985.

LC'S AUTHORITY RECORDS—GENERAL COMMENTS

Other general comments came through about LC's authority practices and some of the limitations of AACR2.

Respondents want more 4xx's (*see* references) than LC's rule interpretations call for, particularly for references for transliterated names and pre-AACR2 forms for the heading. There was a recent study conducted by LC on the reduction of *see* references to eliminate some of the variations for abbreviated forms and other selected variations that would be normalized by a computer comparison. Those normalized types of variations were not mentioned in this request for more *see* references.

The move away from including history notes was seen as disastrous, with one dismayed respondent exclaiming, "They zapped our history notes!"

They want scope notes.

They want more dates, including birth and death dates for persons as well as dates and history information for complicated corporate name changes to simplify the resolution of conflicts.

They want earlier/later names preserved rather than just made into *see also* or *see* references. They often find the earlier name is not AACR2. They want the earlier/later relationship stated explicitly not just embedded in the w subfield, perhaps using a separate tag.

Several disheartened comments were made about disagreement with AACR2 decisions on the part of LC and then LC changes to an even more disagreeable decision for the same heading.

They want more information in the 6xx fields to record more background on the authority decision.

And problems were mentioned with LC's decisions on the predominant form to use, with CIP information conflicting with title page information, and both being in conflict with the LC authority record for the person.

Talk about frustrations. . . .

And yet, let's not lose sight of the fact that the majority of respondents were delighted that the file was available.

One point, which Janet Swan Hill observed in her response to the opinion poll is that there is a difference between what is needed in a shared *resource* authority file and what might be included in an authority file controlling a single catalog or a set of catalogs. Those characteristics for the local authority file will vary depending on the capabilities of the catalog or catalogs it controls. To expand on that, the local system or local online catalog capabilities will influence the characteristics of the local authority file and should reflect any "front end" capabilities for normalizing or otherwise matching search terms. Local catalogs and authority files should be capable of customizing the local catalog for local needs, while the national or international *resource* authority file would preserve the decisions made by national libraries and be shared by all. Obviously I am among those who do not believe that every library must follow the same authorized form for headings, and I feel that computer systems help make local variations economical as well as possible, while enabling more libraries to utilize and even contribute to the national resource through use of standards. A change to the MARC authorities format would facilitate local variation while preserving the national decisions and links to the local decision.

So when we look at the responses to the question on stating one's ideal or ultimate authority control system, there was considerable range of opinion depending on the point of view: whether a national resource system or a locally established control.

The most frequent responses were that authority control should be on a complete, integrated online system which includes all types of authority records (names, series, subjects, uniform titles, etc.) all linked to bibliographic records and automatically updatable (global update of headings in the bibliographic records). Display of the correct form regardless of the form entered into the system with an explanation or choices for the user when different headings were retrieved were also desired. This would be accomplished by including complete reference structures and links to bibliographic records to resolve conflicts when the same search term represented more than one choice of name, series, title, or subject.

The LITA/RTSD Discussion Group on Authority Control in the Online Environment will be looking into several of these issues with the goal of improving the resource authority files as well as offering suggestions for design of future local authority control systems. We currently have our Steering Committee working in three areas:

1. issues in bibliographic and authority record links;
2. general issues related to subject authority control; and
3. looking at alerting mechanisms for maintenance of NACO authority records to no-

tify users of changes to authority records (this includes general maintenance suggestions for online authority files).

We met during ALA Midwinter for small group discussions on these three areas, and will have further position papers for our discussions at the summer ALA Conference in Chicago.

It was clear from the Dallas ALA Conference that there is a resurgent interest in authority control as more libraries actually begin their own online catalogs, and we hope to provide a forum for identifying directions for the future of authority control.

Those are the results I have gleaned from the responses to date, but I expect to work on the statistical analysis of the results to see what interesting correlations there may be and to further analyze the sample in preparation of a more complete final report.

BIBLIOGRAPHY

Mosey, Jeannette Gail. "Name Authority Work in OCLC Libraries: A Survey of Practices and Expectations." Ph.D. diss., University of Southern California, 1980.

Name Authority Cooperative/Name Authority File Service, Task Force on a Name Authority File Service, Bibliographic Service Development Program. Washington, D.C.: Council on Library Resources, 1984. ■■

News and Announcements

"Talking" Terminal for the MELVYL Catalog

The California State Library has awarded fifty-two-thousand dollars to the University of California Division of Library Automation to develop a "talking" terminal for the visually impaired. The terminal will interface with the university's MELVYL online catalog.

The funds are allocated under the federal government's Library Services and Construction Act (LSCA) program.

The MELVYL catalog combines library holdings of all nine UC campuses into one source, accessible to UC's library patrons by computer terminals located in libraries.

The project, which includes a state-of-the-art survey and the design and implementation of the talking terminal, will be conducted between January and September 1985.

Marion Bourke, principal librarian of the California State Library's Braille and Talking Book Library, has been designated primary consultant for the project.

The Division of Library Automation, which develops and maintains the MELVYL catalog, is a unit of the Office of the Assistant Vice President—Library Plans and Policies, Office of the President, University of California.

For further information, contact Carol Connolly, Division of Library Automation, 186 University Hall, University of California, Berkeley, CA 94720; (415) 642-9485. ■■

OCLC Publishes Microcomputer Magazine

OCLC is publishing a magazine for users of its M300 Workstation and the IBM PC. Each issue of *OCLC Micro* features articles, tips from OCLC and network staff, users' comments, small programs, solutions to problems, hardware and software reviews, and article abstracts. Some issues will have diskettes containing helpful programs that users can take advantage of immediately.

OCLC Micro was created to share information among M300 Workstation and IBM PC users, but some articles will be geared toward other popular microcomputers, such as Apple II and the MacIntosh.

Beginning with the premiere issue in March, *OCLC Micro* will be published every other month, with monthly issues during November and December. The subscription cost is thirty dollars per year.

Libraries interested in contributing articles to or subscribing to *OCLC Micro* should contact their network offices or Ginni Voedisch at OCLC. ■■

WLN Introduces Micro-Recon

The Washington Library Network has introduced Micro-Recon, a new software program for Apple and IBM PC microcomputers. Micro-Recon helps member libraries convert their holdings to machine-readable form. It allows libraries to easily create search keys, which are then matched against the WLN Bibliographic Database. This new method of retrospective conversion supplements a similar online facility that has been available to WLN members for six years and offers libraries greater flexibility and economy in their conversion efforts.

Micro-Recon provides a formatted screen and prompts the terminal operator for each type of data to be keyed (e.g., title, LCCN, call number, etc.). Micro-Recon also has an editor to detect certain errors. Diskettes containing search keys are mailed to WLN where they are periodically matched against the WLN Bibliographic Database. When a match occurs, the li-

brary's holdings are added. The holdings become available to WLN users online and may be included in COM catalogs and the WLN Resource Directory.

Micro-Recon was developed by the Washington Library Network with the help of a development grant from the Alaska State Library. It is the first in a series of microcomputers software packages to assist libraries in effective use of microcomputer for library applications. The recon software is available only to WLN member libraries. For more information about other library software applications please contact David Andresen, Microcomputer Consultant, Washington Library Network, AJ-11, Olympia, WA 98504-0111; (206) 459-6589. ■■

University of Toronto Sells the UTLAS Library Information System to International Thomson

The University of Toronto, UTLAS Inc., and International Thomson have concluded an agreement under which International Thomson has acquired UTLAS Inc. from the University of Toronto.

UTLAS Inc. is a bibliographic and information utility that has grown out of research and development at U of T and has been supplying online database services and related products in English and French to Canadian libraries since 1973. In recent years, it has also added a number of U.S. and Japanese libraries to its user network. Over three hundred institutions, members of consortia, and government agencies maintain individual databases through the utility's facilities, and more than two thousand individual libraries of all types and sizes receive products and services from this system.

International Thomson, a large Canadian-based multinational company, has significant interests in publishing in a wide variety of professional, educational, and library information services. The company has placed special emphasis in the last few years on the electronic distribution of information and the building of databases. International Thomson views the UTLAS acquisition as a significant expansion of its commitment to provide a wide variety of library services.

There is significant compatibility between UTLAS' long-term plans and those of International Thomson. Activities of International Thomson that are of interest to librarians and information professionals are provided through companies such as Research Publications Inc., Carrollton Press, Inc. and its REMARC project, Richard De Boo Publishers, Thomas Nelson International, Derwent Publications, Van Nostrand Reinhold Company Inc., and Wadsworth Inc.

As part of the transaction, the university has entered into a long-term agreement with UTLAS, whereby UTLAS will continue to be the exclusive provider of current cataloging services as well as being given an opportunity to develop new circulation systems for the university's libraries. Also, the university will cooperate with UTLAS in developing alternative means of providing automated services. ■■

AMIGOS Announces Microcomputer Program

AMIGOS Bibliographic Council has announced plans to offer training and technical support for selected microcomputer software, beginning in July 1985. The program will emphasize library and popular business application software, and will be available to AMIGOS members as part of a subscription package that features training customized to library needs and access to full user support and troubleshooting on software packages supported by the program.

Members will also have access to discounts on purchases of microcomputer software and printers, modems, and other peripheral equipment. An AMIGOS staff member will be available for consulting with libraries in planning local microcomputer services.

Initially the program will feature software written for the IBM Personal Computer and compatible microcomputers. Preliminary plans were approved at a meeting of the AMIGOS Board of Trustees held February 7 and 8 in Dallas. The program is an outgrowth of a consultant study conducted in the summer and fall of 1984, which identified member needs for these services. ■■

Initial Funding Completed for RLG Recon Program

The William and Flora Hewlett Foundation has awarded $1,050,000 to the Research Libraries Group over a three-year period to support the RLG cooperative retrospective conversion program. This grant completes funding for the first phase of the program, which has also received a $750,000 grant from the Andrew W. Mellon Foundation and a $350,000 contract with the J. Paul Getty Trust. The funds will be used to subsidize retrospective conversion projects approved by the RLG Board of Governors.

The cooperative program is part of RLG's effort to support RLIN retrospective conversion—the adding of records from manual files or from other systems into the Research Libraries Information Network (RLIN).

During the first year, subsidized projects will cover the following subject areas:

- Science (Q-QR)
- History: Western Europe
- History: U.S.
- Literature and Languages (P)
- Music
- Art and Architecture

To participate, member libraries submit project proposals for approval by RLG's Board of Governors. A proposed project is then evaluated to see whether it will:

contribute to substantial coverage of a subject area listed in the project guidelines;

involve conversion of a collection identified as strong by the RLG Conspectus;

result in the addition of many titles, especially titles that are original to the database.

If the project is approved, RLG and the member library agree on a series of "milestones" to ensure steady progress. As the milestones are met, RLG reimburses the library in full for RLIN processing costs and provides funding to cover other costs based on the number of records added.

For the program's first phase, which extends from September 1, 1984, through August 31, 1985, thirteen recon projects were approved at the following member institutions: the American Antiquarian Society, the University of California at Berkeley, Brown, Columbia, Cornell, Iowa, Johns Hopkins, Michigan, Minnesota, New York University, Stanford, Temple, and Yale. ■■

CLSI and Blackwell Join Forces

CL Systems, Inc. (CLSI) and Blackwell Library Systems, Inc., the automation subsidiary of B. H. Blackwell, Ltd., have announced their agreement to establish a long-term relationship to develop, market, and sell Blackwell's computerized Perline 100 Serials Control and Bookline 100 Acquisitions Systems. Under the agreement, the Perline and Bookline Systems will be offered as stand-alone units as well as integrated modules with CLSI's LIBS 100 System.

Operating cooperatively, CLSI and Blackwell will be able to offer libraries an integrated system with acquisitions, serials control, public access catalog, and circulation control. ■■

RLG Study on Distributed Processing

The Research Libraries Group, Inc. (RLG) has published *Processing and Data Distribution within the Research Libraries Information Network*, the final report of a study undertaken with financial assistance from the Carnegie Corporation of New York.

The report describes RLG's plans for a distributed architecture, in which much of the cataloging, acquisitions, and local searching now done on the RLIN network computer will shift to systems at member institutions. The technical and financial issues involved in this shift are discussed.

The distributed architecture proposed in the report is a network that links RLIN to local systems. The central database will contain the information needed to carry on RLG programs such as interlibrary loan, shared resources, collection management, and cooperative preservation. This "data resource node" will accept data and referred searches from local systems at member institutions. The local systems will support local online catalogs and such activities as circulation, serials control, acquisitions, and cataloging. The architecture of the distributed system will allow a

variety of local choices while still ensuring that RLG's programmatic goals will not be compromised in any way.

The report also assesses the hardware and software needed for such a distributed system, surveying current technology, looking at past trends, and trying to predict future developments and costs.

Included are the results of a survey that asked RLG members about their requirements and plans for local systems. The responses helped to define a set of necessary or desirable local applications. Seven local systems currently available were evaluated based on their support for these applications and on their ability to communicate with the distributed network. The systems were Avatar, BLIS, Carlyle, CLSI, GEAC, Innovative Interfaces, NOTIS, and RLIN. Of these, BLIS, GEAC, and NOTIS were recommended, though with some reservations. Members were also asked how soon they expected to implement local systems. The information gathered was used to project changes in the amount of cataloging and acquisitions processing done on the central RLG system.

The final section of the report deals with the financial impact that distributed processing will have on RLG. It considers four "scenarios" based on assumptions about activity levels, expenses, and income at the central site over the next eight years. This part of the report also develops a rough model of the local costs that the "average" RLG member library will face over the same period.

Processing and Data Distribution within the Research Libraries Information Network is available for $8 (plus state and local sales tax for purchasers within California) from: The Research Libraries Group, Inc., Attn.: Ms. Christina Schmehl, Jordan Quadrangle—Oak, Stanford, CA 94305. ■■

Online Catalog at the University of Illinois at Urbana-Champaign

One of the largest online catalogs in the country has been operational at the Library of the University of Illinois at Urbana-Champaign (UIUC) since late August 1984. The Online Catalog at the UIUC consists of two components. The first is the statewide LCS system, which contains brief bibliographic and detailed holdings-/circulation records for UIUC as well as twenty-five other academic libraries in Illinois. The second component is based on the software of the Washington Library Network WLN and its implementation is the result of a joint project with the River Bend Library System of Coal Valley, Illinois, funded through the Illinois State Library. This second component, known locally as the Full Bibliographic Record or FBR system, now contains over 850,000 full MARC records and affords complete author, title, and subject access through a network of over 250 terminals. At UIUC, the system is used as a public access catalog while its use at the River Bend Library System is intended for staff use.

Both the FBR component and the LCS component are available through the same terminal network and operate on an IBM 3081 run by the University of Illinois in Chicago. At any terminal, a patron can type a given LCS or FBR command and the combined system will determine which component to consult. After locating a full catalog record, through a subject search for example, the patron can retrieve the LCS record through a simple command to determine the location and circulation status of an item.

The combined LCS/FBR system provides a very fast system for known item searching (LCS) and a sophisticated bibliographic retrieval system (FBR) with such advanced features as keyword searching, subject retrieval, and authority control.

To assist patrons in searching, 58 of the 100 public terminals are actually IBM Personal Computers equipped with an interface program. This program, written by Professor C. C. Cheng of the Linguistics Department at UIUC, intercedes between the user and the Online Catalog interpreting one to the other.

In addition to terminals at UIUC, each of the eighteen regional library systems that serve library users throughout the state has a terminal tied into the Online Catalog and each of the twenty-five other academic libraries that use LCS have a terminal as well. ■■

University of Michigan
Library Acquires Microcomputers

The University of Michigan University Library has acquired seventy Zenith-150 microcomputers for staff use from Zenith Data Systems. The Michigan-manufactured microcomputers will be deployed in library units across campus; in addition each 320K, dual disk machine will include an internal modem, printer, and an integrated software package. The selection of Zenith equipment was based on a review of available hardware by library staff and available pricing agreements for computer equipment made with the university. In addition to the library's initial purchase of microcomputers, Zenith has donated to the university fifty Zenith-150 machines, which will be placed in the Undergraduate Library's Microcomputer Center jointly operated by the library and the Computing Center.

One aspect of the microcomputer plan includes the use of electronic mail and the creation of electronic conferences for staff to discuss issues and assist each other in using the machines effectively. Both communication media will use the campus network as a "switching center."

A unique aspect of the proposed communication network includes the School of Library Science, which has simultaneously acquired Zenith equipment for each of its faculty.

Training efforts for the microcomputers have been coordinated through the Library's Systems Office with the joint efforts of many individuals throughout the library. The multiple training sessions cover four functional areas and are intended to provide a cadre of staff with basic skills. Staff who have been trained are expected to train others in their units. In total, training will be provided for several hundred staff members.

While the library expects that innovative applications will evolve as staff become familiar with the capabilities of microcomputers, there are several general categories where activity is expected. Communication capabilities should help tie together a geographically distributed staff. The availability of word processing and editing capabilities will provide critical support for writing and the maintenance of files of written material. Microcomputers will also allow branch libraries to be able to access the library's large central database to determine items on order or check circulation files. A number of software programs could strengthen the planning functions of the library through the use of various software packages designed for management functions.

The library hopes to evaluate the impact of the deployment of microcomputers among library staff. Such a study might identify specific characteristics of the changes taking place in the library environment, focusing on those that might be generalizable to other similar environments. Data will also be collected that highlight issues relevant to implementation of networks such as this. ■■

LITA Conference at INFOMART

Hold the dates August 1–3, 1985, for a fabulous conference to be held by LITA and four other educational associations at INFOMART in Dallas. Program information and a registration brochure will be mailed to every member of ALA in June. INFOMART is the new (opened in January 1985) technology mart that is a national center for people-exposure to the information technologies. When fully occupied, it will consist of 225 permanent exhibits and many temporary exhibits. It has thirty-seven meeting rooms, an electronic resource center (the library), an auditorium, a gigantic atrium, restaurants, and complete telecommunications and teleconferencing facilities. The conference will zero in on technology use and its impact in libraries and on all levels of education. Hold the dates—it will be something to remember!

Recent Publications

Reviews

Genaway, David C. *Integrated Online Library Systems: Principles, Planning and Implementation.* White Plains, N.Y.; London: Knowledge Industry; 1984. 151p. Professional Librarian Series. ISBN: 0-86729-092-7, hardcover, $36.50; 0-86729-091-0, softcover, $28.50.

Interest in "integrated online library systems" has grown quite suddenly to dislodge "online public access catalogs" from its place as the central issue and organizing principle of contemporary library systems planning. This development reflects the experience of an ever greater number of libraries of all types with a broad variety of computer-assisted tools and expresses many hard lessons learned about using one system for catalog management, another for interlibrary loan, still another for acquisitions, a fourth for circulation, and so on. Interest in integrated online library systems does not draw its inspiration from the heady future of "electronic libraries." It is motivated instead by enduring concerns for reducing costs, increasing productivity, and shortening delays in material processing activities. It directs attention toward the structural aspects of library systems planning and away from mechanisms for access. It recommends a return to the "whole systems" conceptualizations of early automation writings and efforts with an urgency and dedication appropriate to rising clientele and constituency expectations and increasing institutional and competitive pressures.

The author of this short book states that his purpose is "to provide an overview of integrated online library systems and to outline some of the planning procedures, evaluation and selection criteria useful in acquiring and implementing these systems." For "integrated online library system" he offers a working definition of "a system that accesses a common machine-readable database and has two or more subsystems operational and online" and cites "circulation, acquisitions, online public access catalog, reference, documents, serials, cataloging, etc." as the subsystems he has in mind. Most of the book is devoted to general background information, key concepts, and miscellaneous advice regarded by the author as important for planning, evaluating, selecting, and implementing these systems. Three surveys also are presented and discussed: one was used at Youngstown State University (where the author is university librarian) to determine staff and patron attitudes toward administrative issues connected with system selection; the second studied the degree of satisfaction of seventy system users (with usable responses from forty-five of them) in May 1984; and the third inventoried the subsystems available from thirty-two vendors in the spring of 1984. Nearly one third of the main body of the book is taken up with the brief summaries of the individual responses obtained in the vendor survey and two appendixes supplement this information and provide a directory of suppliers. A final chapter provides some insight into how microcomputers are beginning to fit into the overall picture.

The author defines his audience as "all those involved in the selection and acquisition of such a system as well as all others interested in integrated online library systems." Even though regular readers of this journal are members of Genaway's target audience, it seems certain that most will find very little of value in this book, especially given its cost. By providing an overview and outlining some useful planning procedures and evaluation and selection criteria, the author has attempted to serve the interests common to the members of his vast audience, but the result is uneven, shallow, and anecdotal. The flaw of the author's approach is not so much in his definition of audience as in the relatively little at-

tention he pays to the unique requirements and challenges of integrated online library systems. Most of what the author has to say can be said about all library systems, be they integrated or not, online or not. This misdirection of attention is most clearly evidenced by the "profiles of integrated online library systems" that occupy so much of the book. Here can be found the familiar litany of vendors, give or take one or two, accompanied by more or less the same information provided elsewhere, regardless of whether "vendors of online public access catalogs," "circulation system vendors," or something else is being addressed. The book oversimplifies a complex and promising conceptual development in library automation planning. It is recommended only for those readers with minimal preparation who desire a quick introduction prior to pursuing a rigorous reading program.

There are a number of significant trends in library systems development that the word "integration" has been introduced to organize. The theme common to all of these is the need for and possibility of progressively closer cooperation of systems with dissimilar identities, functionalities, or localities. Providing circulation status information through online public catalog searches is a common example of how circulation and catalog management systems should be integrated. Automatic referral of a search from an online public catalog system to a cataloging utility or to the online public catalog system of another library is an example of how systems that manage local bibliographic records should be integrated with those that manage regional or national ones. Electronic transmittal of inquiries, orders, claims, and payments between library acquisition systems and the inventory, fulfillment, and accounts receivable systems of publishers and jobbers is an example of how library processing control systems should be integrated with the systems of suppliers. These examples illustrate different types of integration (functional, horizontal, and vertical, respectively) with different types of benefits (cost and time savings, information enrichments, etc.). The technical architecture of most contemporary systems features a single database supporting multiple "user views" that vary according to function and, therefore, command and display requirements. It is fair to say, though, that few (and perhaps no) systems merit designation as truly "integrated." Library system planners must still make choices regarding which systems are most important given clearly understood local circumstances while making judgments about hospitality not only to growth and change but to cooperation with other systems as well. This book provides nothing that is distinctive or new to readers who face such responsibilities.—*Paul Evan Peters, Columbia University Libraries, New York, New York.* ■■

Godden, Irene P., ed. ***Library Technical Services: Operations and Management.*** Library and Information Science Series. Orlando, Fla.: Academic, 1984. 272p. ISBN: 0-12-287040-9, hardcover, $32.

Editor Irene Godden states in the preface to *Library Technical Services: Operations and Management* that this work is "intended to give an overview of current operations and techniques associated with the acquisition, organization for access, and physical processing and maintenance of library materials." The overview is accomplished by chapters on technical services administration (Leslie Manning) and automation (Karen Horney) that serve as the framework for the chapters on acquisitions (Marion Reed), bibliographic control (Betty Bengston), preservation and materials processing (A. Dean Larson), and circulation (Leslie Manning). In addition to a bibliography and/or list of references, four of the chapters also include a section on "keeping up." While that idea is very attractive, particularly for new practitioners, comments are often redundant and a bit obvious, e.g., go to ALA, talk with your colleagues, etc.

The anticipated audience for the book is library administrators, librarians, and library school students. However, the descriptions of operations are primarily those of academic and research libraries.

The preface refers to Tauber's 1954 classic *Technical Services in Libraries* and suggests that while Tauber provides a compre-

hensive treatment of the topic for a generation ago, there is a great need to discuss current technical services realities. Indeed, this book does not replace Tauber, whose detailed descriptions are still useful, but rather supplements it. Comparing the indexes of Godden and Tauber serves as an instructive inventory of how we've changed. Tauber did not discuss access, AACR2, approval plans, archive tapes, authority control, automation, bibliographic utilities, Boolean searches, copy cataloging, deacidification, desuperimposition, disasters, ISBD, ISBN, ISSN, MARC, networks, online cataloging, OCR, retrospective conversion, standards, or theft.

The introduction by Godden provides a very good survey of technical services structure and the recent changes occurring through fiscal constraints and automation. She presents an interesting model of information systems as CPOD systems, involving the *C*ollection, *P*reservation, *O*rganization and *D*issemination of records. Godden contends that

> whatever the face of the future, it does seem safe to predict that, just as the recent past in technical services has been dominated by change, the continuing diversification of the information environment in the future will bring more and more accelerated changes in how the unique and central functions now called technical services will be performed.

The administration chapter by Leslie Manning covers basic management, personnel management, and measurement and evaluation procedures. While Tauber covered some similar topics, Manning provides in-depth coverage of current management practices and tools. Automation is not specifically covered in the chapter, but the author does note how current trends in automation might affect organizational patterns and personnel management.

Karen Horney's chapter on "Automation: The Context of the Potential" provides a clear, comprehensive overview of integration, linkage, standards, networks and cooperation, and the automation of specific technical services functions. This chapter, perhaps more than any other, could have been obsolete upon publication; it is to Horney's credit that it concentrates on concepts and trends rather than upon the details that are so rapidly changing.

The chapter on acquisitions by Marion Reed covers the traditional acquisition functions of vendors, records and files, ordering procedures, and special materials and problems. The chapter is a clearly written review of procedures. Even though the automation chapter discusses automated acquisitions, this reviewer would have preferred further delineation and discussion of the effect of automation on traditional acquisitions functions.

In the chapter on bibliographic control, Betty Bengston reviews organization and management, descriptive cataloging, subject cataloging and classification, original and copy cataloging, serials, authority control, filing and catalog maintenance, and retrospective conversion. Bengston's article successfully serves as an overview of current operations and trends as well as a review of major studies and sources to consult for further information.

"Preservation and Materials Processing," by A. Dean Larson, deals with the physical environment, processing and preservation procedures, deaccessioning, and related cooperative efforts. Preservation as a topic is probably second in popularity only to automation, and Larson provides a fine presentation on the current status and concerns of preserving our collections.

The final chapter on circulation is by Leslie Manning, who also contributed the chapter on technical services administration. It includes the traditional aspects of circulation and notes some of the effects of automation on this process. Many of the previous chapters noted that automated circulation is increasingly being assigned to technical services departments; Tauber's work also included circulation as a technical service function. It would have been helpful for this chapter to concentrate more on the ramifications of this changing organizational structure.

Compilations, unfortunately, usually lack some of the consistency of a work written by one individual and the present volume is no exception. Yet on the whole, *Library Technical Services* is a fine work, clearly written and surprisingly current. It does update Tauber and makes an important contribution to the field of technical

services. The introduction and the chapter on automation should attract a wide readership, and the book as a whole will be particularly well suited to the needs of library science students and technical services novices.—*Berna L. Heyman, College of William and Mary, Williamsburg, Virginia.* ■■

Gorman, Michael, ed. ***Crossroads: Proceedings of the First National Conference of the Library and Information Technology Association, September 17-21, 1983, Baltimore, Maryland.*** Library and Information Technology Series, no.1. Chicago, Ill.: American Library Assn., 1984. 261p. ISBN: 0-8389-3307-6, softcover, $40.

The forty-one papers in this book represent the keynote addresses, contributed papers, state-of-the-art papers, and panel presentations made at the first LITA national conference. The papers focus on the past, present, and future of library technology, and the range of subjects is best illustrated by this listing of the nine chapter headings: "Generalities" (i.e., keynote addresses), "States of the Arts," "User Response to Online Catalogs" (papers from a panel), "The Automation of Authority Control" (status reports from various institutions), "Networking and Technology," "The Administration of Technological Change," "Electronic Publishing and the Library," "Advances in the Automated Preparation and Storage of Library Information," and "Interactive AV: Videodiscs and Computer-Assisted Instruction."

Speakers included library administrators; public services, technical services, and systems librarians; library school faculty; management consultants; system vendors; scientists; publishers; university administrators; and one U.S. senator. Many of them are nationally known for their expertise on the topics they covered. For example, keynote speakers and the titles of their addresses were: Howard L. Resnikoff, "Libraries and the Magnification of Mental Power"; Ithiel de Sola Pool, "Looking Down the Road of Technological Change"; Hans H. Wellisch, "*Aere perennius*? Information, Technology, and the Durability of Records"; John Wicklein, "Will the New Technologies Kill the Public Library"; and Bruce A. Miller, "A Brief History of Telecommunications in the United States." Other examples of expertise, to choose but a few, include presentations by Gary Lawrence on online catalogs, Dick Boss on document delivery, and Joe Matthews on office automation.

This volume is similar to other conference proceedings. The papers present a snapshot of thinking on library automation and related topics at one point in time. Some of the material is timeless, while some of it was out of date before it was presented. Papers range from thought-provoking ones on general topics that are of interest to a large audience, to very specific, sometimes technical ones that are of interest to a very small audience. The papers vary in length (up to fifteen pages), and many include references (as many as thirty-one in one paper), figures, and tables. There are no abstracts.

In some ways this work is better than other conference proceedings. The quality of the printing and the aesthetics of the design are quite good and a refreshing contrast to the cut-and-paste, camera-ready-copy appearance of many proceedings. But in other ways this work is a disappointment. It lacks indexes to either subjects or authors, for example, and the institutional affiliations and positions of authors are listed in the back of the book, not on the first page of each paper where they belong.

A conference proceedings should do more than just deliver the formal presentations given at a conference; it should convey as much as possible the flavor and ambiance of the conference. Not many proceedings live up to this expectation, nor does this one. For example, a summary of the discussions that follow the formal presentations can be very educational, since much of the value of any conference lies in the exchange of information triggered by the presentations. Some published proceedings attempt to capture that information, but this one does not.

This was a significant conference, well planned and well executed by the many LITA members (credited in the volume) who worked hard to insure coverage of ap-

propriate topics, to assure quality in the presentations, and to provide the best setting possible for teaching and learning. In his interesting preface, editor Michael Gorman reflects on the conference and its importance. Nowhere in the work, however, is there any indication of the excitement (and even humor) shared by conference participants, or any mention of the unique aspects of this conference, such as the hands-on laboratory for learning about special technologies and applications, and the experimental use of an electronic mail messaging system for communication between conference participants. Although this is the official record of the conference for all posterity, no mention is made of such important things as the number and types of conference participants, the format of the conference, and other items that may be of interest to current and future readers.

The work is nice, but it could have been better. Nevertheless, it is the best record we have of this important conference, and despite its high price (for a paperback) it should certainly be in the collection of every library and every individual concerned with and interested in technology in libraries and related issues.—*Keith W. Russell, National Agricultural Library, Beltsville, Maryland.* ■■

Markey, Karen. ***Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs.*** OCLC Library, Information, and Computer Science Series, 4. Dublin, Ohio: OCLC, 1984. 174p. ISBN: 0-933418-54-X, softcover, $21.

In 1979, OCLC began the Subject Access Research Project to survey past catalog use studies regarding subject access and to examine in detail the process of subject searching in traditional library card catalogs. In 1981, the Bibliographic Service Development Program of the Council on Library Resources (CLR) funded five organizations—OCLC, RLIN, the Library of Congress, University of California Division of Library Automation, and J. Matthews and Associates—to conduct coordinated studies of online catalogs in thirty-one libraries. Some of the most important findings of the CLR evaluative project center around the role of subject retrieval in online catalogs. The CLR online catalog studies found (1) subject searches predominate over known-item catalog searches; (2) catalog users report more problems and dissatisfaction with subject searching; (3) users approach online catalogs with the expectation that they will find enhanced subject access to a broader field of materials (including periodicals) than in the card catalog; and (4) users recommend changes in online catalog features that would expand subject entry vocabulary and provide enhanced access to library materials. *Subject Searching in Library Catalogs* by Karen Markey of the OCLC Office of Research integrates the findings of the OCLC Subject Access Research Project and the federated CLR studies to offer a detailed model of user behavior in the subject retrieval process and presents a prescription for improved subject searching in online catalog environments. Markey does not attempt an abstract theory of information retrieval but rather uses data gathered through questionnaire administration, transaction log analysis of user-entered search arguments and corresponding system responses, and focused group interviews to examine the present capabilities and projected needs of subject retrieval functions for online catalogs. This book provides a very timely and useful summary of access features of present online catalogs and is required reading for designers of online catalogs and individuals involved in the selection of an online library catalog.

By summarizing the findings of the aggregate CLR studies (in chapters 1, 5, and 7) and focusing on the important element of subject searching in online catalogs, Markey provides a convenient overview of results released largely in separate reports by the individual agencies. The book complements the detailed summary of the findings of four of the project participants found in *Using Online Catalogs: A Nationwide Survey* by Joseph Matthews, Gary Lawrence, and Douglas Ferguson (Neal-Schuman, New York, 1983).

Chapter 3 describes the subject searching capabilities, including access points and

command language protocols, for seven representative operational online catalogs. Relevant features of other online library catalogs are mentioned throughout the book. One is struck by the wide variation in present online catalog search features. These catalogs are represented, at one end, by systems featuring sophisticated information retrieval capabilities such as keyword and Boolean searching, online authority control, and search limiting capabilities, and at the other end, by online catalogs that require user-entered exact matches (including proper punctuation) with LCSH and offer no cross-references.

By examining user entry vocabulary and analyzing subject retrieval failures, Markey is able to present a great many specific insights into the subject searching behavior of library users. She calls into question previous studies on percentages of entry vocabulary matches in subject searching by noting a phenomenon first demonstrated by Marcia Bates—that users who approach the catalog at an incorrect level of specificity will locate some materials but not the most relevant materials. She also presents a great deal of anecdotal evidence for the usefulness of both title word and subject component word searching.

Chapters 4 and 5 report on data derived through an analysis of the transaction logs generated from the Syracuse University online catalog. Among the findings: in two thirds of the searches, there was at least one access point that resulted in no retrievals; only 29 percent of searchers' input matched or closely matched LCSH; of the LCSH matches, 29 percent retrieved 100 or more citations; and in 61 percent of the searches, users entered terminology described as "whatever popped into their minds." The concurrent problems of too few or no hits and too many hits in subject searching receive a great deal of attention in Markey's discussion of online catalog enhancements.

In Chapter 6, Markey discusses the importance of online subject authority control and comments on the dissimilarity between the data elements in LC/MARC subject authority records and cross-references contained in the LCSH "Red Books." However, the missing "see also to" cross-references can be generated and collected from the machine-readable records by computer processing. Attempts at imposing further structure into LCSH have not met with general success. Several studies—most recently by Toni Peterson (*Journal of Academic Librarianship*, 1983, p.207)—have described the difficulties associated with restructuring LCSH into a modern BT, NT, RT thesaurus-style format by automatic or algorithmic methods.

Chapter 8 emphasizes the importance of the user-oriented interface in communicating with the online catalog. It is interesting to note the increasing use of the microcomputer as an information workstation, linking online catalogs with remote bibliographic databases, local information resources, and other online catalogs and circulation systems.

Markey offers specific suggestions and techniques for improving subject retrieval, including online vocabulary assistance and syndetic structure, augmentation of bibliographic records, refinement capabilities such as Boolean searching and limiting operations, tracking of searches with suggestive prompting, automatic truncation and spelling correction, and utilization of classification scheme information. The implementation of these capabilities requires a level of software sophistication and computing power that will challenge the present generation of online catalogs, particularly those systems that are minicomputer based. But, Markey notes, as online catalogs move out of their infancy, the important area of enhanced access needs to be addressed. This work will serve as the basis for all future investigations into improved subject retrieval.—*William H. Mischo, University of Illinois at Urbana-Champaign.* ■■

Mount, Ellis, ed. ***Fee-Based Services in Sci-Tech Libraries.*** New York: Haworth, 1984. 105p. ISBN: 0-86656-326-1, hardcover, $19.95. "Also . . . published as *Science and Technology Libraries*, Volume 5, Number 2."

In the sixties, libraries of all types enjoyed the benefits of lavish funding from

the federal government. Since the retrenchment of federal support that came in the seventies, cost recovery measures have come into vogue. All libraries today are familiar with cost recovery measures. In the eighties, fees for photocopying, bibliographic and online searching, and interlibrary loans are commonplace. The continued and probably permanent absence of federal funding has given new meaning to the adage "there's no such thing as a free lunch"—somebody has to pay. For libraries, "lunch" is the cost of doing business.

This latest Haworth Press publication is a collection of articles that seek to address issues involved with fee-based services in science and technology libraries. The first article, "Cost Allocation and Cost Recovery Considerations in a Special Academic Library: Georgia Institute of Technology," deals with cost recovery methods in a state school. It is interesting to note that by GT's own estimates, almost one fifth of the services rendered are for off-campus users. Needless to say, the library has an increased need to recover costs from non-student/faculty users. Another issue raised is the assessment of a direct library fee for students, much the same way an athletic fee is levied on undergraduate students. Campus research programs that receive outside funding are also scrutinized. Should these programs be charged for library services? If so, how and how much? Two existing programs for corporation support and an on-campus document delivery program for faculty and staff are also presented. Although the authors recognize that no library could function on a totally transactional basis, their article does suggest that there are ways of spreading the overhead costs to both inside and outside users more equitably.

A second article, by Shirley K. Baker of the Massachusetts Institute of Technology, amplifies the discussion by relating the experience at M.I.T. In addition to a well-documented history of the evolution of fee-based services at M.I.T., the author talks about the future. Such possibilities as offering access to the online catalog to outsiders with the logical corollary of increased document delivery requests are discussed. Another project involves the use of the five major subject libraries and branch libraries as nodes in a projected network of three thousand IBM and DEC microcomputer workstations for students and faculty. The libraries would, in addition to offering their catalogs online to users, respond to requests for reference services and document delivery via electronic mail. This service could be extended to outside users, thus generating revenues for the library.

The third article presents cost recovery efforts in the health-science library setting. The author, Lynn Kasner Morgan of the Mount Sinai Medical Center, presents an excellent overview of the AIMS project and what it hopes to achieve for medical libraries. The article also documents the development of interlibrary loan fee systems and the emergence of consortium development as a remedy to recover cost.

A fourth article, "Sci-Tech Libraries and Serials Agents: The Unused Leverage," deals with fees charged to libraries by journal subscription agents. Although an excellent and well-written article, it is somewhat tangential to the theme of the book.

There are several faults with this book. First, it is a reprint of the winter 1984 issue of *Science and Technology Libraries*. Second, out of 105 pages, only 32 deal directly with the topic of the book's title. The balance of the book, or should I say journal issue, consists of the following: the aforementioned article on serial agents' fees, a paper discussing evolving roles of end users and librarians in computer searching, a bibliography on artificial organs, a discussion of new reference works, online news, and reviews of journal articles. The bibliography and review section absorbs over half the book.

While it does begin with four well-presented essays, this book is not recommended for purchase. Overall, it is a well-rounded journal issue but a dubious excuse for a book. Redundant publishing, no matter what the publisher's intention, should not be rewarded.—*Tom Smith, Paul Himmelfarb Health Sciences Library, George Washington University Medical Center, Washington, D.C.* ■■

Other Recent Receipts

Listed here are books and other publications of potential interest to members of LITA, received for review. Some of these materials may be reviewed in later issues of ITAL.

Castonguay, Russell. *A Comparative Guide to Classification Schemes for Local Government Documents Collections.* Westport, Conn., and London: Greenwood, 1984. 144p. ISBN: 0-313-24208-9, hardcover, $35.

CLASS. *Directory of Microcomputer Applications in Libraries.* San Jose, Calif.: Cooperative Library Agency for Systems and Services, 1984. 128p. ISBN: 0-938098-08-X, spiralbound.

Fike, John L., and Friend, George E. *Understanding Telephone Electronics.* 2d ed. Understanding Series. Dallas, Tex.: Texas Instruments, 1984. 284p. ISBN: 0-89512-159-X, softcover, $14.95.

Friend, George E. and others. *Understanding Data Communications.* Understanding Series. Dallas, Tex.: Texas Instruments, 1984. 268p. ISBN: 0-89512-158-1, softcover, $14.95.

Gregor, Dorothy, ed. *Retrospective Conversion of Music Materials: Report of a Meeting Sponsored by the Council on Library Resources, July 18–19, 1984, Wayzata, Minnesota.* Washington, D.C.: Bibliographic Service Development Program, Council on Library Resources, 1984. 118p. Softcover, $6 prepaid.

Helal, Ahmed H., and Weiss, Joachim W. eds. *Local Library Systems: Essen Symposium, 24 September–27 September, 1984.* Veröffentlichungen der Gesamthochschulbibliothek Essen, 7. Essen: Gesamthochschulbibliothek Essen, 1984. 337p. ISBN: 3-922602-0808, softcover. "Festschrift in honour of Frederick G. Kilgour to his 70th birthday."

Information Systems Consultants Inc. *Videodisc and Optical Digital Disk Technologies and Their Applications in Libraries: A Report to the Council on Library Resources.* Washington, D.C.: Council on Library Resources, 1985. 191p. Softcover; $6 prepaid.

Jones, Maxine Holmes. *See, Hear, Interact: Beginning Developments in Two-Way Television.* Metuchen, N.J., and London: Scarecrow, 1985. 155p. ISBN: 0-8108-1720-9, hardcover, $15.

Li, Tze-chung. *An Introduction to Online Searching.* Contributions in Librarianship and Information Science, no.50. Westport, Conn., and London: Greenwood, 1985. 289p. ISBN: 0-313-24274-7, hardcover, $27.95.

Lipetz, Ben-Ami, and Paulson, Peter J. *A Study of Impact of Technological Change in Library Service Facilities: Changes in Use of the Public Catalog at the New York State Library Associated with the Introduction of Online Subject Searching Capability in the Public Catalog.* Albany, N.Y.: New York State Library, 1984. 31p. Softcover, $4. "Summary Report to the Council on Library Resources on Cooperative Research Grant CLR 30007-E."

Lovecy, Ian. *Automating Library Procedures: A Survivor's Handbook.* London: Library Assn., 1984. 247p. ISBN: 0-85365-516-2, hardcover, $28 from Oryx Press, Phoenix, Ariz.

Mason, Robert M., ed. *The Micro Consumer: Library Software: A Guide to Selection.* Atlanta, Ga.: Metrics Research, 1984. 110p. ISBN: 0-932393-01-2, softcover, $25.

Matthews, Joseph R. *Public Access to Online Catalogs.* 2d ed. Library Automation Planning Guides Series, no.1. New York and London: Neal-Schuman, 1985. 497p. ISBN: 0-918212-89-8, softcover, $35.

Peters, Paul Evan, ed. *Command Language and Screen Displays for Public Online Systems: Report of a Meeting Sponsored by the Council on Library Resources, March 29–30, 1984, Dublin, Ohio.* Washington, D.C.: Bibliographic Service Development Program, Council on Library Resources, 1985. 93p. Softcover, $4 prepaid.

Provan, Jill E., and Hunter, Joy W., eds. *Health Media Review Index: A Guide to Reviews and Descriptions of Commercially-Available Nonprint Material for the Medical, Mental, Allied Health, Human Service and Related Counselling Professions.* Metuchen, N.J., and London: Scarecrow, 1985. 844p. ISBN: 0-8108-1739-X, hardcover, $59.50.

Saffady, William. *Micrographics.* 2d ed. Library Science Text Series. Littleton, Colo.: Libraries Unlimited, 1985. 254p. ISBN: 0-87287-453-2, hardcover, $28 in U.S. and $33.50 elsewhere.

Schmitt, Neil M., and Farwell, Robert F. *Understanding Automation Systems.* 2d ed. Understanding Series. Dallas, Tex.: Texas Instruments, 1984. 276p. ISBN: 0-89512-164-6, softcover, $14.95.

Thomason, Nevada Wallis, ed. *Microcomputer Information for School Media Centers.* Metuchen, N.J., and London: Scarecrow, 1985. 316p. ISBN: 0-8108-1769-1, hardcover, $25. ■■

Letters

To the Editor:

A few typographical errors crept into pages 346 and 347 of *VDT Checklist: Another Look at Terminals* (December 1984). In the flurry of common and metric measurements under "Screen Size," centimeters and millimeters got confused, and some numbers got mislaid.

A 12-inch diagonal screen (305mm) is typically about 190mm high and 250mm wide, using extreme measurements. A "9-inch" screen would be about 225mm diagonal measure, and a 14-inch screen would be about 350mm.

The discrepancy between my recommended minimum character height (4mm) and R. Bruce Miller's (3mm) was based on a different definition of "character height." We discussed this at ALA Midwinter 1985, and came to the conclusion that he was specifying "x-height," the height of a typical lowercase letter (the *x*). I was specifying cell height, the distance from the bottom of one line to the bottom of the next.

A 3mm x-height is identical to a 6mm cell height for a typical character set. Typical good-quality character sets use 7 × 9 characters in a 9 × 12 cell. Of the twelve vertical elements in a cell, one is reserved for spacing (or underscores), two are used for descenders, and three are used for ascending elements. That leaves six elements in "x": exactly one-half the cell height. So, his "minimum acceptable" height requires a full 12-inch screen; for extended use, I would agree.

Looking back at the manuscript, the "seven-by-nine character matrix within an eight-by-ten cell" is marginal: it precludes well-formed descenders. Seven-by-nine within nine-by-twelve is probably the crudest character set acceptable for extended use. Who knows? By 1989 that may seem unacceptably crude for well-designed displays.—*Walt Crawford, Manager, Product Batch, The Research Libraries Group, Inc., Stanford, California.*

A Microcomputer Based
Serials Control System

Call Us.

800-225-6055 *(toll free)*
617-329-3350 *(collect)*

fa:on

Does your HELP key need a helping hand?

Directory of Consultants in Computer Systems from Research Publications helps you locate the consultant trained to meet your specific computer needs. We've listed and indexed over 7,000 highly qualified computer software consultants, consulting firms, and programmers throughout the United States to provide you with immediate access to the consultant who has just the expertise and credentials to help your business. These professionals will customize or modify your existing software package, or design a new one to fit your company's needs.

Directory of Consultants in Computer Systems is the first consultant directory designed with the computer owner/manager in mind. This easy-to-use directory includes consultants in every aspect of computer software, from micros to mainframes; from accounting to inventory. We require references for every consultant we list, so you can be assured of finding an active and qualified professional. Detailed listings and indexing with multiple cross-references will save you hours of blind phone calls and search time.

Directory of Consultants in Computer Systems — the comprehensive source that gets you to the professionals. Now available from Research Publications for only $75.00.

To order, complete and return the coupon.

Please send me _____ copy(ies) of **Directory of Consultants in Computer Systems** ($75.00 each; plus $4.00 S/H — price applies to U.S. and Canada only)
Connecticut residents add 7.5% sales tax.

_____ My check is enclosed for $ _____
(Please make checks payable to Research Publications, Inc.)

_____ Please charge to my credit card $ _____
___ AM EXP ___ VISA ___ M/C
Acct. # _____
Exp. Date _____

Name _____
Company _____
Address _____
City _____ State _____ Zip _____

Research Publications
12 Lunar Drive/Drawer AB
Woodbridge, CT 06525
(203) 397-2600
TWX: 710-465-6345
FAX: 203-397-3893

Outside North and South America
P.O. Box 45
Reading, RG1 8HF England
TEL: 0734-583247
TELEX: 848336NADL G

research publications®

Research Publications
12 Lunar Drive/Drawer AB
Woodbridge, CT 06525
(203) 397-2600
TWX: 710-465-6345
FAX: 203-397-3893

Outside North and South America
P.O. Box 45
Reading, RG1 8HF England
TEL: 0734-583247
TELEX: 848336NADL G

research publications®