Volume 8
Number 4

ISSN 0730-9295

Information Technology and Libraries

December 1989

CONTENTS

) 345–472 (1989); Inf Technol Libr

ISSN 0730-9295

Information Technology and Libraries

Volume 8, Number 4: December 1989

CONTENTS

ITLBDC 8(4) 345–472 (1989); 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:* Linda J. Knutson. The journal is issued quarterly in March, June, September, and December.

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 83 of the March 1989 issue.

Manuscripts of articles, communications, and news items should be addressed to the editor-designate: Thomas W. Leonhardt, *Information Technology and Libraries*, University of the Pacific, Stockton, CA 95211. Copies of books for review should be addressed to: Susan B. Harrison, *ITAL Book Reviews*, The New York Public Library, 455 Fifth Ave., New York, NY 10016. Copies of software for review should be addressed to: George S. Machovec, *ITAL Software Reviews*, Arizona State University, Tempe, AZ 85287. Advertising arrangements should be made with Kurt Murphy, Hayden Library, Arizona State University, Tempe, AZ 85287.

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology Association. Subscription price, $17.50, is included in membership dues. Nonmembers may subscribe for $35 per year in U.S., Canada, Mexico, Spain, and other PUAS countries; $40 in other foreign countries. Single copies, $9. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. *Postmaster:* Send address changes to *Information Technology and Libraries*, 50 E. Huron St., Chicago, IL 60611.

Circulation and Production: Central Production Unit (Chere Elliott, CPU Supervisor; Bruce Frausto, Debra M. Hall, and Colleen King), American Library Association, 50 E. Huron St., Chicago, IL 60611.

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 *CompuMath Citation Index, Computer Contents, Computer Literature Index, Current Contents/Health Services Administration, Current Contents/Social & Behavioral Sciences, Current Index to Journals in Education, Education, Library Literature, Magazine Index, NewSearch*, and *Social Sciences Citation Index*. Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

Inter-Indexer Consistency in Subject Cataloging

Lois Mai Chan

Better quality control over the subject access points in cataloging records is essential if we are to improve retrieval effectiveness or to facilitate cooperative cataloging. An important measure of catalog quality is how consistently individual records are indexed. Findings of earlier inter-indexer consistency studies indicate that there is a direct correlation between retrieval effectiveness and inter-indexer consistency; that is, high inter-indexer consistency in assignment of indexing terms appears to be associated with a high retrieval effectiveness of the documents indexed.[1]

Most MARC records are prepared according to widely accepted standard tools such as *Anglo-American Cataloguing Rules*, second edition, (AACR2), *Library of Congress Subject Headings* (LCSH), *Dewey Decimal Classification*, and *Library of Congress Classification*. When these tools are carefully followed, the result is what we accept as standard practice, which, for want of an absolute measure, we define as congruence with Library of Congress (LC) cataloging. Some discrepancy in record treatment is inevitable, given the variance in the nature of the material cataloged and the degree of subjective judgment required to use the tools. We find justifiable inter-cataloger differences even in descriptive cataloging, where circumstances are more clear-cut than in subject cataloging. Discrepancy is naturally considerably higher for subject cataloging, where guidelines are less distinct and where there is much more scope for differences in assessing the material. Still, it must be our goal to reduce inter-indexer inconsistency as far as possible. To do so, we need more information than we have on the inconsistencies we all know exist.

The problem has been addressed in numerous studies in the past.[2] Many of the earlier studies, however, suffer from the lack of reliable means of measurement because of the presence of uncontrolled variables. Some have relied on samples prepared specifically for the experiment, having the same documents indexed by two or more indexers for comparison, resulting in consciously or artificially prepared samples. Others have compared records extracted from similar databases but possibly indexed according to different policies. Clearly, a more valid way of measuring inter-indexer consistency is desirable. The work reported here details a new approach to measuring inter-indexer agreement on subject heading assignment.

The purpose of the current study has been twofold: first, to develop a valid methodology for studying indexing consistency in MARC records and, second, to study such consistency in subject cataloging practice between non-LC libraries and the Library of Congress.

METHOD OF STUDY

Data Collection

It appears that a policy in effect at OCLC does provide an effective (that is, valid and reliable) instrument for consistency measurements. OCLC prefers LC MARC records to records input by member libraries and routinely replaces member

Lois Mai Chan is a professor at the College of Library and Information Science, University of Kentucky, Lexington. Based on the results of a project sponsored by the OCLC Library School Research Equipment Support Program.

records when their LC MARC equivalents become available. The result is pairs of records, with one of each pair prepared at LC and the other not. The member records are presumably prepared in the absence of LC records for the same titles. Since LC catalogers do not have online access to the OCLC database, it is safe to assume that they prepare the LC MARC records without consulting the OCLC member records. By collecting such pairs of records (capturing the member records from the OCLC database before they were replaced by LC records and pairing the two) and comparing the subject headings assigned to the same titles by different catalogers who used the same subject headings list and presumably following the same policies, that is, LC practice, it is possible to measure indexing consistency in terms of degrees of matching.

For the present study, a pool of approximately 500 records prepared and entered into the OCLC Online Union Catalog during 1986 by non-LC libraries paired with records prepared subsequently by the Library of Congress for the same bibliographic items was supplied by the Research Office of OCLC. From this collection, 100 pairs were selected as the sample for this study. In selecting the records from the collection supplied by OCLC, the following categories were excluded:

1. Foreign language records were excluded in order to eliminate the variable of language proficiency on the part of the individual catalogers preparing the records.

2. Records for later editions or issues of a work were excluded because it is possible that the cataloging of the non-LC record for the later edition or issue might have been based on an LC record for an earlier edition or issue.

3. Records for works that contain CIP (Cataloging-In-Publication) information were excluded because the CIP information was originally provided by LC.

The first 100 pairs of records that survived these exclusionary criteria were selected for this study.

Data Analysis

Record Comparison

The non-LC and LC records for the same titles were paired and compared for the purpose of identifying variations among the subject headings assigned. Only LC subject headings were studied. No attention was paid to any additional subject headings derived from other sources, such as MeSH (*Medical Subject Headings*) and children's subject headings.

Pattern Identification

Patterns of variations were identified and sorted according to the component parts of LC subject headings—main headings, subdivisions, capitalization, punctuation, spelling, and field tags.

Data Tabulation

The number and percentage of each pattern of variation were recorded and tabulated.

Heading Verification

Headings found in non-LC records that are not exact matches of headings in corresponding LC records were analyzed for their validity in terms of LC policies. These headings were first checked in *Library of Congress Subject Headings* (10th edition, 1986).[3] Headings listed in LCSH were presumed to be valid. A list of the remaining Headings was sent to the Subject Cataloging division of the Library of Congress for verification.[4] The list was returned with the invalid headings identified and explained.

RESULTS AND ANALYSIS

The results are presented below in the order of decreasing degree in indexing consistency. Perfect matches are presented first, followed by partial matches and nonmatches. Differences in main heading are considered to contain greater conceptual variation than differences in subdivisions, since main headings usually represent the main subject of the document, while subdivisions reflect its aspects.

Distribution of Headings

The distribution of subject headings appearing on the 100 pairs of records is shown in table 1. A total of 190 headings appears in the 100 non-LC records, whereas the 100 LC records were assigned a total of 214 headings. The average number of headings assigned to non-LC records is 1.9 and that of LC records is 2.14, representing a difference of 0.24. In other words, each LC record received an average of 12.63% more headings than a non-LC record.

Compared with the results of an earlier study by O'Neill and Aluri which showed an average of 1.41 subject headings per record in the OCLC database, the average number of subject headings per non-LC record in the current study shows an increase of 0.49 heading per record.[5] The average number of subject headings per LC record shows an increase of 0.73.

The range of numbers of headings assigned to non-LC records is identical to that of LC records; both non-LC records and LC records received from 0 to 5 headings each. The mode is 2 headings per record, with 38 (38%) non-LC records and 33 (33%) LC records being in this category. The study by O'Neill and Aluri cited above identified the mode to be 1 heading per record, represented by 44.4% of the sample studied.[6]

The next most common number of headings per record is 1; both groups contain 28 records with 1 subject heading each. In the other categories, 8 non-LC records received 0 headings in comparison with 6 LC records with no subject heading. On the other hand, more LC records received greater numbers of headings than non-LC records: 13 LC records received 4 headings each, and 4 LC records received 5 headings each in comparison with six 4-heading records and only one 5-heading record in the non-LC group. Nineteen non-LC records and 16 LC records were assigned 3 headings each.

Table 1. Distribution of Headings

	Non-LC	LC
Total sample		
Number of records	100	100
Number of headings	190	214
Average number of headings per record	1.90	2.14
Range of numbers of headings per record	0–5	0–5
Number of records with		
0 heading	8	6
1 heading	28	28
2 headings	38	33
3 headings	19	16
4 headings	6	13
5 headings	1	4

Table 2. Profile of Matching

	No. of Records Non-LC	No. of Records LC
Study sample	100	100
Perfectly matched records	15	15
Partially matched records	80	80
Unmatched records	5	5

Profile of Matching

When the sample was analyzed record by record, it was found that out of the 100 pairs 15 were perfect matches; that is, with exactly identical subject cataloging (same number and same headings). (See table 2.) Of the remaining records, 80 (80%) pairs of records have some headings that matched completely or partially, and 5 (5%) pairs are totally unmatched.

The term "perfectly matched records" means a pair of records describing the same bibliographic item showing the same number of subject headings which match character for character. The term "completely matched headings" is defined as a pair of headings that are matched character for character, including field tags, subfield codes, punctuation, and capitalization. "Partially matched headings" refers to a pair of headings that contain identical main headings but different subdivisions, as well as a pair of headings that contain one or more, but not all, identical words in the main heading portion.

Perfectly Matched Records

A total of 12 headings appear in the 15 pairs of records that matched perfectly (see table 3). Out of these 15 pairs, 6 pairs received 0 heading, 7 pairs received 1 heading each, 1 pair received 2 headings each, and 1 pair received 3 headings each.

It is interesting to note that the perfect match in 6 of the 15 pairs is due to the fact that no subject heading is required for these titles. In actuality, only 9 pairs of records have subject headings that are perfectly matched.

Table 3. Perfectly Matched Records

	Records No. (in pairs)	Records %	Headings Non-LC No.	Headings Non-LC %	Headings LC No.	Headings LC %
Total sample	100	100	190	100	214	100
Perfectly matched records	15	15	12	6.32	12	5.61
with 0 heading each	6	6	0	0	0	0
with 1 heading each	7	7	7	3.68	7	3.27
with 2 headings each	1	1	2	1.05	2	0.93
with 3 headings each	1	1	3	1.58	3	1.40

Table 4. Imperfectly Matched Records

	No. of Pairs of Records	No. of Completely Matched Headings
Imperfectly matched records	80	
Records with 1 or more completely matched headings	*46*	58
Records with 1 completely matched heading	36	36
Records with 2 completely matched headings	8	16
Records with 3 completely matched headings	2	6
Records with only partially matched heading(s)	*34*	0

Table 5. Completely Matched Headings

	Records No. (in pairs)	Records %	Headings Non-LC No.	Headings Non-LC %	Headings LC No.	Headings LC %
Total sample	100	100	190	100	214	100
Records with at least 1 completely matched heading	55	55	70	36.84	70	32.71
Records with 1 completely matched heading	43	43	43	22.63	43	20.09
Records with 2 completely matched headings	9	9	18	9.47	18	8.41
Records with 3 completely matched headings	3	3	9	4.74	9	4.21

Table 6. Partially Matched Headings

	Headings Assigned by Non-LC Libraries No.	Headings Assigned by Non-LC Libraries %	Records No. of pairs	Records %
Headings not completely matched	120	63.16	85	85
Partially matched headings	*67*	35.26	52	52
Headings with exact match in main heading portion	45	23.68	38	38
Headings with partial match in main heading portion	22	11.58	19	19
Unmatched headings	53	27.90	38	38

Imperfectly Matched Records

Of the 80 pairs of records that are not perfectly matched, 46 (46%) pairs of records have one or more, but not all, completely matched headings (see table 4). Among these, 36 pairs contain one completely matched heading each (a total of 36 headings), 8 pairs contain 2 completely matched headings each (a total of 16 headings), and 2 pairs contain 3 completely matched headings each (a total of 6 headings).

Of the 80 pairs of imperfectly matched records, 34 pairs contain only partially matched headings, that is, identical main headings with different subdivisions or main headings that have one or more component words in common.

Completely Matched Headings

In the *total* sample of 100, 55 (55%) pairs of records have at least one completely matched heading (see table 5). Among these, 43 (43%) pairs of records have exactly one completely matched heading each, 9 (9%) pairs have 2 matched headings (for a total of 18 headings), and 3 (3%) pairs of records have 3 matched headings each (for a total of 9 headings).

In other words, out of the 190 headings assigned to the non-LC records, 70 (36.84%) are complete matches of headings on corresponding LC records.

Partially Matched Headings

Of the 120 (63.16%) headings assigned to non-LC records which do not have exact matches on corresponding LC records, 67 headings (appearing on 52 records) are partial matches (see table 6), that is, pairs of headings with the main headings being totally or partially identical.

Forty-five headings appearing in 38 non-LC records match precisely those in corresponding LC records in the main heading portion. The variations occur in the assignment of subdivisions, with some of the headings showing variations in more than one subdivision.

Twenty-two headings appearing in 19 pairs of records match in one or more, but not all, words in the main heading portion. That leaves 53 headings (appearing in 38 non-LC records) that have no matches on corresponding LC records.

The partially matched headings are discussed below in the order of decreasing degree of match.

Variations in Subdivisions

Of the 67 partially matched headings, 45 headings appearing in 38 pairs of records match precisely in the main heading portion (see table 7). Variations occur in the subdivisions. These headings contain a total of 49 variations, with some headings showing more than 1 variation.

Table 7. Headings with Matched Main Headings and Variations in Subdivisions

	No. of Variations in Subdivisions	No. of Headings	No. of Records
Exact matches in main heading	49	45	38
Variations in Form subdivisions	*17*		14
Absence of (Non-LC:10; LC:4)	14		11
Different	3		3
Variations in Chronological subdivisions	*2*		2
Absence of (Non-LC:1; LC:1)	2		2
Different	0		0
Variations in Geographic subdivisions	*7*		7
Absence of (Non-LC:4; LC:1)	5		5
Different	2		2
Variations in Topical subdivisions	*18*		17
Absence of (Non-LC:8; LC:6)	14		13
Different	4		4
Different types of subdivisions	*5*		5

Following are examples of variations in subdivisions:

Variations in form subdivisions:

Title: *RSANB, 1926–1958 Retrospective South African National Bibliography for the Period, 1926–1958*

Non-LC heading:	651	0	South Africa $x Imprints *$x Bibliography*.
LC heading:	651	0	South Africa $x Imprints.

Title: *Water Atlas of Saudi Arabia*

Non-LC heading:	650	0	Water resources development $z Saudi Arabia $x *Atlases*.
LC heading:	650	0	Water resources development $z Saudi Arabia $x *Maps*.

Variations in chronological subdivisions:

Title: *The Showa Anthology: Modern Japanese Short Stories*

Non-LC heading:	650	0	Short stories, Japanese *$y 20th Century* $x Translations into English.
LC heading:	650	0	Short stories, Japanese $x Translations into English.

Title: *A History of the United States since 1861*

Non-LC heading:	651	0	United States $x History.
LC heading:	651	0	United States $x History *$y Civil War, 1861–1865*.

Variations in geographical subdivisions:

Title: *The Volunteer Organization Handbook*

Non-LC heading:	650	0	Voluntarism $x Handbooks, manuals, etc.
LC heading:	650	0	Voluntarism *$z United States* $x Handbooks, manuals, etc.

Title: *Budgetary Issues Affecting the Great Lakes Basin*

Non-LC heading:	650	0	Federal aid to water quality management *$z Great Lakes*.
LC heading:	650	0	Federal aid to water quality management *$z Great Lakes Region*.

Variations in topical subdivisions:

Title: *The Idea of Nation: The Romanians of Transylvania*

Non-LC heading:	651	0	Transylvania (Romania) *$x Social conditions* $x History.
LC heading:	651	0	Transylvania (Romania) $x History.

Title: *Rheology of Wheat Products*

Non-LC heading:	650	0	Wheat $x *Quality* $x Congresses.
LC heading:	650	0	Wheat $x *Testing* $x Congresses.

Different types of subdivisions assigned under same main heading:

Title: *Topics in Employee Benefits for Association Executives, 1985*

Non-LC heading:	650		Employee fringe benefits *$x Congresses*.
LC heading:	650	0	Employee fringe benefits *$z United States*.

Title: *A Parent's Treasure Box of Ideas for Preschoolers*

Non-LC heading:	650	0	Education, Preschool *$y 1965-*
LC heading:	650	0	Education, Preschool *$z United States $x Parent participation*.

Variations in Main Headings and in Tagging

Main headings vary in several aspects: capitalization, punctuation, spelling, entry element, and wording. Variation in capitalization does not alter the meaning of the heading and normally does not affect retrieval. Punctuation and spelling, on the other hand, may affect the meaning and term matching in retrieval. Variations in the entry element or in the wording of the heading can definitely have an impact on index display and retrieval. In the study sample, most variations in entry element involve name headings.

A total of 24 instances of partial matches in the main heading portion were identified (see table 8).

Following are examples showing variations in main headings:

Title: *1985 Report of the Task Force to Examine the School Construction Program*

Non-LC heading:	650	0	School *building* $z Maryland $x Design and construction $x Finance.
LC heading:	650	0	School *buildings* $z Maryland $x Design

and construction $x Finance.

Title: *The Van Voorhees Family*
Non-LC heading: 600 30 *Van* Voorhees family.
LC heading: 600 30 *Voorhees* family.

Title: *Managing Today and Tomorrow with On-Line Information*
Non-LC heading: 650 0 Management *information systems.*
LC heading: 650 0 Management $x Data processing.

Headings in Totally Unmatched Records

Five pairs of records in the sample show total failure to match in any portion of the main headings in corresponding records. In 2 of the 5 cases, the failure to match is due to the fact that the non-LC records were not assigned any subject headings. Although the sample is too small to draw valid conclusions, it is interesting to note that 3 out of the 5 titles in this group fall in the subject area of business and/or management. Examples are shown below:

Title: *Christmas Collectibles: Tree Ornaments and Memorabilia*
Non-LC heading: 650 0 *Collectors and collection.*
LC heading: 650 0 *Christmas* $x Collectibles.

Table 8. Variations in Main Headings and in Tagging

	No. of Variations	No. of Records
Variations in main heading	24	22
Variations in capitalization	1	1
Variations in punctuation	2	2
Variations in spelling	1	1
Variations in entry element	3	3
Variations in wording	17	15
Variations in tagging	1	1

Table 9. Headings Assigned to Non-LC Records

	No. of Headings	%
Unmatched headings	120	100
Headings printed in LCSH (10th ed.)	42	35
Unprinted headings	78	65
Valid LC headings	54	45
Unauthorized headings	24	20

Title: *Involving Managers in Competitive Analysis*
Non-LC heading: 650 0 *Business intelligence.*
LC heading: 650 0 *Strategic planning.*

Title: *Capital Facilities Planning: A Tactical Approach*
Non-LC headings: 650 *Strategic planning.*
650 0 *Facility management.*
LC heading: 650 0 *Capital investments* $x Planning.

Analysis of Headings Assigned by Non-LC Libraries

The 120 headings found in non-LC records that do not match completely with headings on corresponding LC records were analyzed as to their validity in terms of LC policies (see table 9). These headings were checked against the 10th edition of *Library of Congress Subject Headings* (1986), and 42 were found listed. The remaining 78 headings were sent to the Subject Cataloging Division of the Library of Congress for verification.[6] Among these, 54 were identified as valid LC headings though not listed in LCSH, that is, unprinted headings that are nonetheless valid.

Examples of unauthorized headings assigned by non-LC libraries are shown below with LC's explanations:

Antitrust *legislation* $z European Economic Community Countries
(Heading does not exist. Use *Antitrust law*—Local subdivision)

Computer simulation $x *Analysis*
(Subdivision not authorized)

Coos County (N.H.) $x Genealogy $x *Sources*
(Subdivision redundant and not necessary)

Hanako, $d 1868–1945
(Ota, Hanako)

Unauthorized Main Headings

Among the 24 headings that are not valid LC headings listed above, 7 (29.17%) headings contain unauthorized elements in the main heading portion (see table 10). Of these, 4 (16.66%) are name headings (for example, HANAKO, 1868–1945 rather than OTA, HANAKO, 1868–1945), 1 (4.1%) involves the question of singular vs. plural noun (SCHOOL BUILDING rather than SCHOOL BUILDINGS), 1 (4.17%)

heading represents a reversed citation order (EMIGRATION AND IMMIGRATION $z UNITED STATES rather than UNITED STATES $x EMIGRATION AND IMMIGRATION), and only 1 (4.17%) heading involves a conceptual variation (ANTITRUST LEGISLATION rather than ANTITRUST LAW).

Unauthorized Subdivisions

The remaining 17 (70.83%) headings show the use of unauthorized subdivisions or subdivision combinations (see table 10). Among these, 8 instances involve the improper use of form subdivisions such as $x MAPS or $x BIBLIOGRAPHY, and 2 involve the use of $x ADDRESSES, ESSAYS, LECTURES, a subdivision recently abandoned by the Library of Congress. It cannot be determined whether the non-LC cataloger was not aware of the change or whether the assignment of this subdivision occurred before the change in policy was announced by LC.

Only 1 (4.17%) instance of improper use of subdivisions involves a chronological subdivision. In this case, the non-LC cataloger subdivided an inverted literary heading by period. It is LC's policy not to subdivide such headings chronologically.

There are 3 (12.50%) instances of unauthorized use of geographic subdivisions. In 1 heading, the non-LC cataloger subdivided geographically a main heading which has not been designated (*indirect*), that is, subdivisible by place. In another case, the non-LC assigned heading shows a geographic subdivision that has not been established by the Library of Congress. In the third instance, the geographic subdivision in the non-LC-assigned heading (WATER-SUPPLY $z GREAT LAKES) is technically correct, but not totally appropriate conceptually. LC, in commenting on the heading, suggests a more appropriate heading (WATER-SUPPLY $z GREAT LAKES *REGION*), probably because the concept of "water supply" relates to the region rather than the lake.

The remaining 5 (20.83%) unauthorized headings involve topical subdivisions. Of these, 4 are results of unauthorized combinations of main headings and subdivisions, and only 1 represents an incorrect form of the subdivision. This indicates that catalogers outside of LC are having greater difficulty in following LC's policy regarding synthesis of subject heading strings than in arriving at the correct form of subdivisions.

In 13 out of 17 cases of unauthorized subdivisions, the greatest difficulty occurs in the $x subfield, that is, form and topical subdivisions. Chronological subdivisions appear to cause the least problem.

CONCLUSION AND RECOMMENDATIONS

Methodology

As noted above, one of the purposes of this project has been to develop a research method for studying consistency in cataloging and indexing. The method used in this study relied on samples drawn from the real cataloging environment of a large database. The records were prepared by catalogers using the same tools and following the same policies. As a research instrument, this method should also be useful in comparing other access points in MARC records, such as name headings, Dewey Deci-

Table 10. Unauthorized Headings

	No. of headings	%
Unauthorized headings	24	100
Unauthorized main headings	7	29.17
Name headings	4	16.66
Grammatical forms	1	4.17
Citation order	1	4.17
Concept variation	1	4.17
Unauthorized subdivisions	17	70.83
Form subdivisions	8	33.33
Chronological subdivisions	1	4.17
Geographic subdivisions	3	12.50
Topical subdivisions	5	20.83

mal class numbers, and the Library of Congress call numbers.

Should a similar study on a larger sample be undertaken later, the investigator might consider setting up a separate analytical category for records with no subject headings assigned. Comparing pairs in this class with pairs in which only one member is given subject treatment could yield interesting results.

Inferences and Recommendations

In spite of the fact that the study's small sample size limits how far its numerical results may be generalized, the findings are strongly suggestive. They indicate that total consistency among indexers is difficult to achieve, even when they use the same controlled vocabulary and attempt to follow the same indexing policies. Only 15% of the sample shows total consistency, and out of this small group, 40% resulted from the fact that no subject headings are required for the documents in question.

Partial agreement among indexers dealing with the same document appears to be the norm. In this study, 80% of the records contain some common elements between corresponding records: identical heading strings, identical main headings, or at least some identical word(s) in the main headings. The ideal is to have total consistency. Where that is not achievable, the fact that the majority of the records on the same subject share certain common elements in the subject heading strings could at least bring partial success in recall in systems with keyword searching capabilities in the controlled vocabulary fields.

Only 5% of the records studied in this experiment show total inconsistency in the assignment of main headings. It is difficult to generalize on the causes of the inconsistency because of the small size of the sample.

This study has identified patterns of inconsistency in subject cataloging between LC and non-LC cataloging records. Some of these have few remedies, such as different interpretations of the subject content or focus of documents. Other variations, however, can be lessened by efforts to improve LCSH and by better communication of LC's policies to catalogers outside, particularly with regard to synthesis of subject heading strings.

One main area of variation occurs in the use of form and topical subdivisions. Since LCSH is a partially synthetic system, catalogers must often build subject headings strings that are not listed in LCSH, particularly when free-floating subdivisions or phrases are involved. The *Subject Cataloging Manual*, published since 1984, has served as an instrument for communicating LC's subject cataloging policies to professionals outside of the library. Nonetheless, many fine points appear to be lost. With regard to the structure of LCSH, consideration should be given to simplification of subdivisions, for example, having fewer subdivisions that overlap in meaning and are difficult to distinguish. Another suggestion is to move from a largely precoordinate system to one relying more heavily on post-coordination. For instance, many concepts represented by topical subdivisions may be treated as distinctive elements and established as separate headings. This could reduce inconsistency in heading combinations.

A tool which might also help to reduce the instances of invalid subject heading combinations appearing in non-LC records would be a file of authorized headings compiled from subject headings that have appeared in LC MARC records, including unprinted headings, particularly those with free-floating subdivisions. This would be a true subject authority file: it would be apart and different from *Library of Congress Subject Headings*, and it would not need to contain cross references. Catalogers outside of the Library of Congress could use this file to verify locally assigned headings. Such a tool would be useful to library personnel responsible for day-to-day catalog maintenance and to those engaged in retrospective conversion, as well as to catalogers seeking to verify unprinted headings.

Because of the small size of the sample used in this particular study, its results must be viewed with caution when cited as conclusive evidence of indexing consistency, or of inconsistency for that matter. Unfortunately, there are no similar inter-indexer consistency studies that can be used for comparison or validation of the results of this one. Earlier studies on consistency in subject cataloging used different methodology and focused on different aspects of subject indexing.[7,8] With the OCLC Online Union Catalog as a source for study samples, it is hoped that more studies on interindexer consistency will be pursued by interested investigators.

REFERENCES

1. L.E. Leonard, *Inter-Indexer Consistency Studies, 1954–1975: A Review of the Literature and Summary of Study Results* (Champaign, Ill.: University of Illinois, Graduate School of Library Science, 1977).
2. K. Markey, "Interindexer Consistency Tests: A Literature Review and Report of a Test of Consistency in Indexing Visual Materials," *Library & Information Science Research* 6:155–77 (1984).
3. Library of Congress, Subject Cataloging Division, *Library of Congress Subject Headings*. 10th ed. (Washington, D.C.: Library of Congress, 1986).
4. Library of Congress, Subject Cataloging Division. Comments on Non-LC Assigned Subject Headings (1987).
5. E.T. O'Neill and A. Rao, *Research Report on Subject Heading Patterns in OCLC Monographic Records* (Columbus, Ohio: OCLC, Inc., Research Department, Research and Development Division. Report No.: OCLC/RDD/RR-79/1, 1979).
6. Ibid.
7. O.L. Lilley, "Evaluation of the Subject Catalog," *American Documentation* 5:41–60 (1954).
8. J.T. MacMillan and I.D. Welt, "Study of Indexing Procedures in a Limited Area of the Medical Sciences," *American Documentation* 12:27–31 (1961).

DOCLINE: A National Automated Interlibrary Loan Request Routing and Referral System

Gale A. Dutcher

NLM made the DOCLINE interlibrary loan request routing and referral system available to biomedical libraries in the United States beginning in March 1985. This system, based upon online links to bibliographic, institutional, and serials holdings information, automatically routes interlibrary loan requests to potential lenders. The eight DOCLINE system modules permit users to create requests, monitor request status, receive requests, update requests with the action taken, receive messages, receive statistical reports, and support automatic routing and system monitoring. By October 1988 there were over 1,600 libraries using DOCLINE. During the latest full year of operation nearly 1.3 million requests were routed by the system, with an overall fill rate of 92 percent. The Regional Medical Library Network plays a significant role in the success of DOCLINE, since the DOCLINE routing patterns reflect the preexisting network interlending patterns. These interlibrary lending relationships have been established over the approximately twenty years of network activity.

One of the goals of the National Library of Medicine (NLM) is to provide health professionals and biomedical research scientists with timely access to information. To accomplish this goal, NLM coordinates a national biomedical library network, provides document delivery services to augment local resources, and develops and supports an automated document request routing and referral system called DOCLINE.

The Regional Medical Library (RML) Network, composed of libraries in the United States with collections in the health sciences, is intended to provide health professionals, researchers, administrators, and educators with timely and convenient access to biomedical information. The Network includes

1. Seven Regional Medical Libraries, large libraries under contract to NLM, are responsible for coordinating activities within their regions. These activities include document delivery, reference service, development and maintenance of regional bibliographic locator tools, and referral services for basic training and consultation.

2. One hundred thirty resource libraries with large health sciences collections are responsible for providing primary document delivery service in their regions and for other resource sharing activities.

3. Approximately 4,000 Basic Health

Gale A. Dutcher is Technical Information Specialist, Specialized Information Services, National Library of Medicine, Bethesda, Maryland.

Sciences Libraries (BHSL) normally represent the health professional's first point of access into the network (e.g., hospital libraries).

Even before the implementation of any automated interlibrary loan processing, the RML Network used traditional methods of routing and handling to provide document support to users for materials that were not available locally. An estimated two million interlibrary loan requests are handled annually within all levels of the RML network.

NLM began implementation of DOCLINE in 1985. This automated request routing and referral system is a significant improvement over the older, manual methods. The twofold aim of DOCLINE is (1) to expedite the routing and referral of interlibrary loan requests in order to improve access to the literature and (2) to facilitate the collection and analysis of interlibrary loan and collection management information. Although other interlibrary loan systems and networks exist, the approach used in DOCLINE is unique. The system includes the use of online linkages to bibliographic, institutional, and serials holdings information and the use of automated routing, as well as a strong resource sharing network.

HISTORY

As early as 1974 NLM began using the term DOCLINE when discussing a concept of an automated interlibrary loan system. The first generation of DOCLINE was a single database using the INQUIRE[1] software. The DOCLINE database was created in May 1974 and was used for a prototype referral system with the British Library Lending Division (BLLD; currently known as the British Library Document Service Center, BLDSC) at Boston Spa, England. Staff at the National Library of Medicine entered requests selected for referral (serials only) into the database, which staff in England could then access as they would any database on the NLM computer. The BLDSC staff also updated requests by noting the action taken (e.g., filled or not filled) and the date. The traffic was in one direction only.

The next step in DOCLINE development was the addition of the UCLA Biomedical Library to this limited system in 1975. UCLA used the system to refer requests for monographs and serials to NLM. Again, the traffic was unidirectional: UCLA to NLM. NLM could also refer UCLA's requests to BLLD.

NLM created two additional databases that were intended to support the development of the automated interlibrary loan system. In 1979 NLM began the design of the National Biomedical Serials Holdings database, now called SERHOLD.[2] This database contains library holdings information for biomedical serials. The current SERHOLD format is based upon the 1980 ANSI Standard at the Summary level (Z39.42-1980). SERHOLD was originally created with data from many existing regional databases. This database was intended to serve as one of the major supporting features of the new DOCLINE system.

The DOCUSER (Document Delivery User) database was developed in 1981. DOCUSER is a directory database containing information about libraries participating in interlibrary lending. Included in each record is the institutional name and address, as well as management information such as the type of institution (e.g., primary health care, academic health sciences) and the source of funding (e.g., federal, state, municipal, private). The database was created from addresses on interlibrary loan request forms received by NLM.

NLM made a commitment in 1980 to the RML directors to give their libraries access to an automated interlibrary loan request system. In response to that commitment, DOCLINE was expanded to support access by five of the RMLs. This interim system was built with the user language capabilities of INQUIRE and did not require any programming support. The system permitted creation of new requests and status checking. Entry of requests was facilitated through an online link to DOCUSER that contained the name and address information for the borrowing libraries. In addition, a link to the SERIALS database, NLM's serials processing database, permitted the retrieval of the full serial title, call number, and unique identifier into the ILL request (transaction) record. Again, the flow of requests was unidirectional: from the RMLs to NLM. This project provided NLM with valuable information about the format of the request, the database structure required for future development, and potential network problems.

In addition, the Region 4 RML (University of Nebraska Health Sciences Library), with funding from NLM, developed a local interlibrary loan system, OCTANET, for use within that region. The OCTANET[3] system pioneered several of the concepts subsequently incorporated into DOCLINE, notably, the use of a hierarchical routing table.

In 1982 NLM management decided to proceed with the development of a fully functional document request routing system. The development team, composed of librarians and programmers, worked for over two years to produce the DOCLINE system, which became operational in March 1985.

DESIGN

The DOCLINE system designers made a number of decisions and assumptions that influenced the course of the development:

1. To rely on communications capabilities of the value-added networks such as Telenet and TYMNET, rather than to use dedicated leased lines solely for a single system;
2. To accommodate the diversity of equipment already in use within the RML network, including low-speed printing terminals, CRTs, and various microcomputers with different telecommunications software. Since NLM did not wish to prescribe equipment, the system was designed to run on virtually any type of ASCII terminal;
3. To create a system that would be very easy to use and learn so that no training in its use would need to be provided by NLM or the RMLs; and
4. To use an interactive approach that would permit the most flexibility for users to create, route, and update requests.

These considerations were intended to make the system available to even the smallest library by keeping costs and the level of effort required low.

The design team used the Yourdan[4] technique of structured analysis to determine the fundamental components of the system, the flow of data among the various files, and the processes required. The system was subsequently programmed using highly structured code, with the programs written in the PL/1 language. Most of the system databases use the INQUIRE database management system. There are also additional files in other formats used only by the program.

DATABASES USED BY DOCLINE

TRANSACTION contains the active interlibrary loan requests, including the bibliographic data, Library Identifier (LIBID) of the borrowing library, and the routing history. There may be 40,000–60,000 records in this database at any time.

DOCUSER describes the libraries involved in interlibrary loan, including their LIBID, address, and telephone number. Information required by the DOCLINE system, such as the access codes, status, and stored Routing Table are also included in the DOCUSER records. In addition to the active DOCLINE users, there are approximately 10,000 libraries (both domestic and foreign) represented in DOCUSER.

SERHOLD contains serials holdings of biomedical libraries in the United States. As of October 1988, SERHOLD contained over 1,100,000 holdings statements for more than 35,000 serial titles held by approximately 2,600 libraries. This includes the holdings of the approximately 1,600 DOCLINE libraries.

SERLINE (Serials Online), MEDLINE (MEDLARS Online), CATLINE (Cataloging Online), AVLINE (Audiovisuals Online), and SERIALS (Serials Processing) are bibliographic databases in the MEDLARS system that are searched in order to provide correct information in the ILL request. MEDLINE consists of citations to journal articles; CATLINE and AVLINE contain cataloging information of printed materials and audiovisuals, respectively; SERLINE contains the records of serials held by NLM and other biomedical libraries; and SERIALS is an INQUIRE database used for NLM serials processing and to produce the SERLINE database.

ILLUSE is an archival database for completed, inactive interlibrary loan requests that have been deleted from the TRANSACTION database. Records are maintained in this database until the annual statistical reports are produced. This database contains well over one million records by the end of a year.

SYSTEM MODULES

The databases described above are used by the various modules of the DOCLINE system in a manner transparent to the user. That is, the user is not aware of which databases are being accessed or the structure of any file or record. There are eight main modules in the DOCLINE system. Five of these are online functions that may be selected by the user (BORROW, RECEIPT, LEND, STATUS, and MESSAGE); one produces offline reports: (STATISTICS); and two function without direct selection by the user (ROUTING and TIME-TRIGGERED ACTIONS.) (See figure 1.)

BORROW

The BORROW module permits a user to initiate an ILL request. In this module, the user may identify the institutional borrower, if it is different from the creator of the ILL request. This is accomplished using an automatic link to the DOCUSER database. The code used to log on to the system identifies the user to DOCLINE and sets the borrower's LIBID for the session. The acceptance of the system-supplied LIBID

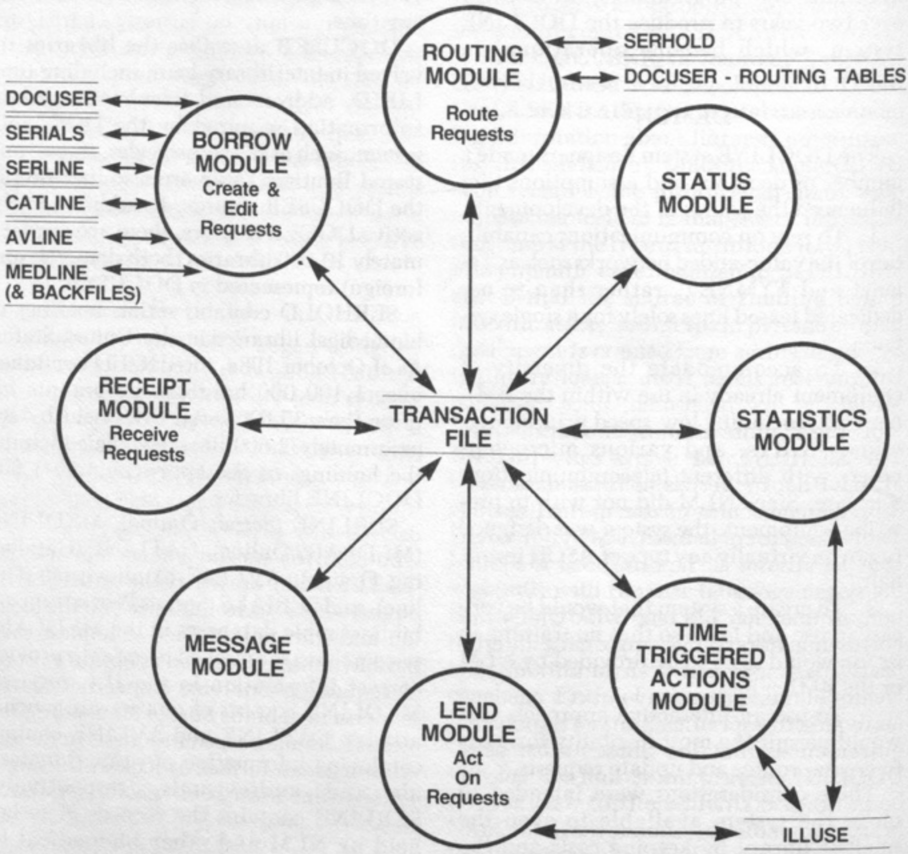


Figure 1. DOCLINE Overview.

or the entry of a different borrower's LIBID identifies the institution and causes the mailing address to be automatically supplied on the ILL request. This module is also linked to the bibliographic databases, so that the creator of the ILL request need only enter the MEDLINE, SERLINE, CATLINE, or AVLINE Unique Identifier (UI); the bibliographic data will then be generated into the ILL request. Alternatively, the user may key the full citation directly. The borrower is also prompted to add all the administrative information pertinent to an ILL request, including certification of compliance with copyright laws. The DOCLINE system will not accept an ILL request without this indication of copyright compliance. The request is stored in the TRANSACTION database while further action on it proceeds.

ROUTING

Once the request has been entered into the DOCLINE system, the ROUTING module is activated. ROUTING uses information in the ILL transaction, SERHOLD, and DOCUSER to send the ILL request to the most appropriate lender. Only ILL requests containing the Unique Identifier (UI) of the serial title from the SERLINE database will be automatically routed by the system. The SERLINE UI may be entered in the record either directly by the creator or by the system if the Unique Identifier from the citation in MEDLINE has been used, since this contains the identification of the serial title. The SERLINE UI is the only link to SERHOLD, the serials holdings database. A predefined routing pattern, the Routing Table, is stored in DOCUSER for each DOCLINE user at the time he or she joins the system. This Routing Table may contain up to eighty libraries with which the borrowing library has preexisting interlibrary loan agreements. In addition, the creator of the request may direct a particular request to a specific library. The only requirement is that the library must be a DOCLINE participant; it does not have to be in the user's Routing Table. If that library cannot fill the request, the automatic routing process begins. The logic of routing is discussed in greater detail below.

RECEIPT

The RECEIPT module permits a potential lender to receive incoming ILL requests. The system prints the requests in a 6-by-9-inch format, the size of a standard ILL form. The request is also formatted with the address of the borrower printed at the top of the form so that it may be used as a mailing label. The system provides for reprinting of all or selected requests in the event of printer failure. Users are expected to use this module to receive their incoming interlibrary loan requests on a daily basis.

LEND

In order to respond to the request, the potential lender must access the LEND module. In this module, the potential lender updates the TRANSACTION record with the action taken on the ILL request. That action may be either FILLED or NOT FILLED. If the request is not filled, the respondent is asked to supply a reason such as NOS (Not on shelf) or BDY (at bindery). DOCLINE uses the ANSI terminology and categories for the nonavailable reasons. Although a few reasons for not filling a request terminate the request (e.g., canceled at borrower's request or incomplete information), most of the reasons for not filling a request reactivate the ROUTING module to route the request to another potential lender using the same criteria as for the first route. The potential lender takes no action on the ILL request other than to update the TRANSACTION with the action code. The lender does not need to actively refer the request.

STATUS

Both borrowers and lenders may use the STATUS module to verify the status of ILL requests. Borrowers may check on the location and status of any active request that they have entered, as well as obtain a list of all requests older than a specified date. Lenders may obtain a list of all pending requests. Both may reprint requests in the interlibrary loan format or look at the complete routing history of a specific request.

MESSAGE

The MESSAGE module is still under development. Phase I, currently available, permits users to obtain an online list of all requests that have routed completely through DOCLINE but remain unfilled. This expedites the use of the STATUS module to verify the location of requests. Future development will permit users to send brief messages to any other user of the system. For example, NLM will use this ability to provide other possible locations for unfilled requests. It will also permit NLM and the RMLs to send messages to classes of users. For example, each RML will be able to send brief messages to all users in their region. The message capability is not intended to replace existing electronic mail systems but rather to permit users to send brief messages about specific ILL requests or questions or comments about the system itself.

STATISTICS

DOCLINE users receive management reports monthly and annually as part of the STATISTICS module. These reports are printed at NLM and distributed to users by mail. Reports include monthly counts of requests initiated as a borrower, requests received as a lender, requests filled and not filled as a borrower and as a lender, and annual frequency lists of requested serial titles (see figure 2). The DOCLINE coordinators at each RML receive lists of participating libraries not using the RECEIPT or LEND modules regularly.

TIME-TRIGGERED ACTIONS

In order to keep the requests moving through the system as rapidly as possible, regardless of the responsiveness of individual libraries, the TIME-TRIGGERED ACTIONS (TTA) module was developed. This module functions as part of the nightly update processing and works on ILL requests in three areas. If a potential lender has not used the RECEIPT function on an ILL request after one full working day, the TTA module routes it to the next potential lender. If a potential lender has RECEIPTed the ILL request but not updated it for four working days, the TTA module reroutes it to the next potential lender. In both of these instances the routing history of the ILL request is updated with this information. The TTA module also removes completed requests from the TRANSACTION database and prepares them for addition to the ILLUSE database for statistical reporting. Completed requests include those that are filled and those that have been routed to all potential lenders but remain unfilled.

ROUTING

The routing algorithm is based upon preexisting interlibrary loan patterns stored in Routing Tables in DOCUSER, serials holdings data in SERHOLD, and the specific citation in the ILL request.

A Routing Table is created by the user library at the time it joins the DOCLINE system. Each library establishes a Routing Table that reflects its interlibrary borrowing patterns, including the guidelines developed within its region. The library may make adjustments to its Routing Table if problems arise or as changes occur within its region.

The stored Routing Table is arranged in seven hierarchical groups, called cells, each of which may contain the LIBIDs of up to twenty libraries, with no more than eighty in the entire table. NLM is automatically listed as the only library in Cell 8. Although each participating library determines its own Routing Table, NLM recommends that each cell contain libraries that are relatively similar in terms of serials holdings.

The ANSI standard describes three reporting levels for serials holdings: LEVEL 1—library symbol and date of report; LEVEL 2—library symbol, date of report, completeness code, and acquisition status code; LEVEL 3—library symbol, date of report, completeness code, and enumeration and chronology data supplied in standard format. NLM has defined an additional level for use by libraries reporting to SERHOLD, LEVEL X, which is identical to the LEVEL 3 format except enumeration and chronology are in a nonstandard format. Only LEVELs 3 and X are used by SERHOLD participants. NLM uses a modification of the LEVEL 3 format (see figure 3), the only one that can be interpreted completely by the DOCLINE program. The modifications were made to provide clear delineation of the parts of the state-

NATIONAL LIBRARY OF MEDICINE
DOCLINE STATISTICS
RANKED LIST OF SERLINE TITLES REQUESTED
JANUARY - DECEMBER 1987

12345X HOSPITAL AND MEDICAL CENTER HEALTH SCIENCES LIBRARY/MIDDLETOWN MD

# OF TIMES REQUESTED	NUMBER FILLED	YEAR(#/YR)	TITLE (SERLINE UI)
3	3	1986 (1) 1985 (1) 1983 (1)	AMERICAN JOURNAL OF CLINICAL ONCOLOGY (A24010000)
2	2	1987 (1) 1985 (1)	AMERICAN JOURNAL OF GASTROENTEROLOGY (A24195000)
2	1	1955 (1)	BRITISH MEDICAL JOURNAL (B35880000)
2	2	1987 (2)	CANCER RESEARCH (C05900000)
2	2	1983 (2)	CANCER TREATMENT REVEIWS (C06120000)
2	2	1985 (2)	CLINICAL AND EXPERIMENTAL IMMUNOLOGY (C06120000)
2	1	1986 (1)	EMERGENCY MEDICINE CLINICS OF NORTH AMERICA (E06036000)
1	1	1981 (1)	ANNALS OF INTERNAL MEDICINE (A34995000)
1	1	1984 (1)	ARCHIVES OF GENERAL PSYCHIATRY (A60930000)
1	1	1985 (1)	NURSING ADMINISTRATION QUARTERLY (N34620000)
1	1	1987 (1)	AMERICAN JOURNAL OF KIDNEY DISEASES (A24335000)

Figure 2. Sample DOCLINE Report.

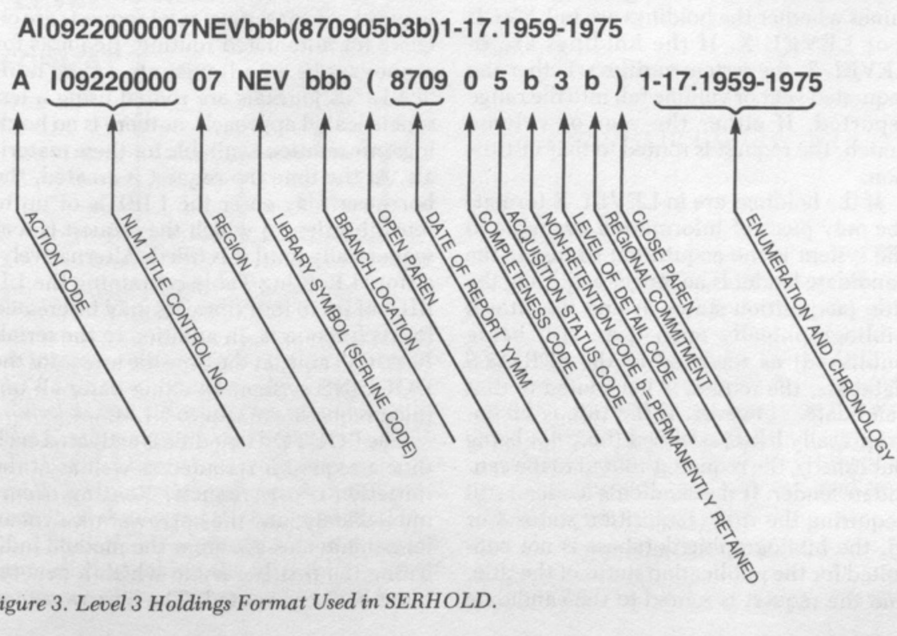


Figure 3. Level 3 Holdings Format Used in SERHOLD.

ment to facilitate automated routing. The essential information used by DOCLINE for automated routing is the library symbol and serial volumes and years owned.

When use of DOCLINE began in 1985, only 20 percent of SERHOLD was in the LEVEL 3 format. The remaining 80 percent was in a variety of LEVEL X formats. As of October 1988, approximately 40 percent of SERHOLD was in the LEVEL 3 format.

In routing (see figure 4), the system randomizes the LIBIDs in Cell 1 of the borrower's Routing Table and selects the first one on the list. Using the SERLINE UI (also known as the TCN, Title Control Number) from the ILL request, it checks SERHOLD to determine whether or not the candidate library has reported owning the title. If the title is not reported as owned by that library, the system selects the next LIBID on the list and checks again for ownership. The system will continue down the randomized list until a library reporting owning the title is located. If no libraries in Cell 1 report owning the title, the system moves to Cell 2. This can continue until the entire Routing Table has been exhausted. The request will then be sent to NLM. However, if a candidate library reports holding the requested title, the routing module determines whether the holdings are in LEVEL 3 or LEVEL X. If the holdings are in LEVEL 3, the system verifies whether the requested year or volume fall into the range reported. If either the year or volume match, the request is routed to that institution.

If the holdings are in LEVEL X format, the only piece of information available to the system is the acquisition status. If the candidate lender is no longer acquiring the title (acquisition status 5) and the title is bibliographically open (i.e., still being published) as reported in the SERIALS database, the request is not routed to that candidate. However, if the title is bibliographically listed as closed (i.e., not being published), the request is routed to the candidate lender. If the candidate lender is still acquiring the title (acquisition status 4 or 0), the bibliographic database is not consulted for the publication status of the title, and the request is routed to the candidate lender. Since the majority of interlibrary loan requests from health professionals are for material with recent publication dates, if a library no longer has a subscription to a title which is still being published, it is more advantageous to the borrower not to try that candidate lender.

Automated DOCLINE routing has been found to be relatively accurate. In a study conducted by NLM staff it was found that only 8 percent of all requests using the automatic routing algorithm were updated by the first potential lender with a code that would indicate an inaccurate routing, i.e., requested item not owned by the potential lender (NLM Routing Study, 1987). It was further found that an additional 25 percent of this group would have routed the same way even if the holdings statement had been reported in the LEVEL 3 format. Another 25 percent of this group would have routed correctly if the acquisition code had been reported accurately. Thus, only 4 percent of all requests undergoing automatic routing are at all likely to be routed inaccurately due to the routing algorithm devised for DOCLINE.

Automated routing is limited to requests for serials that contain the NLM title control number (SERLINE UI). Since well over 90 percent of ILL requests in the RML network are for serials, most requests are eligible for automated routing. Requests for monographs, audiovisuals, and non-SERLINE journals are routed using a less sophisticated approach, as there is no holdings information available for these materials. At the time the request is created, the borrower may enter the LIBIDs of up to four libraries, to which the request is sent sequentially until it is filled. Alternatively, a stored Routing Table containing the LIBIDs of up to four libraries may be created for such requests, in addition to the serials Routing Table, at the time the user joins the DOCLINE system. In either case, all unfilled requests are sent to NLM.

The ROUTING module is activated each time a request is rerouted as well as at the initiation of the request. Routing occurs immediately, and the borrower receives an immediate message from the module indicating the first library to which a new request has been routed. The borrower may

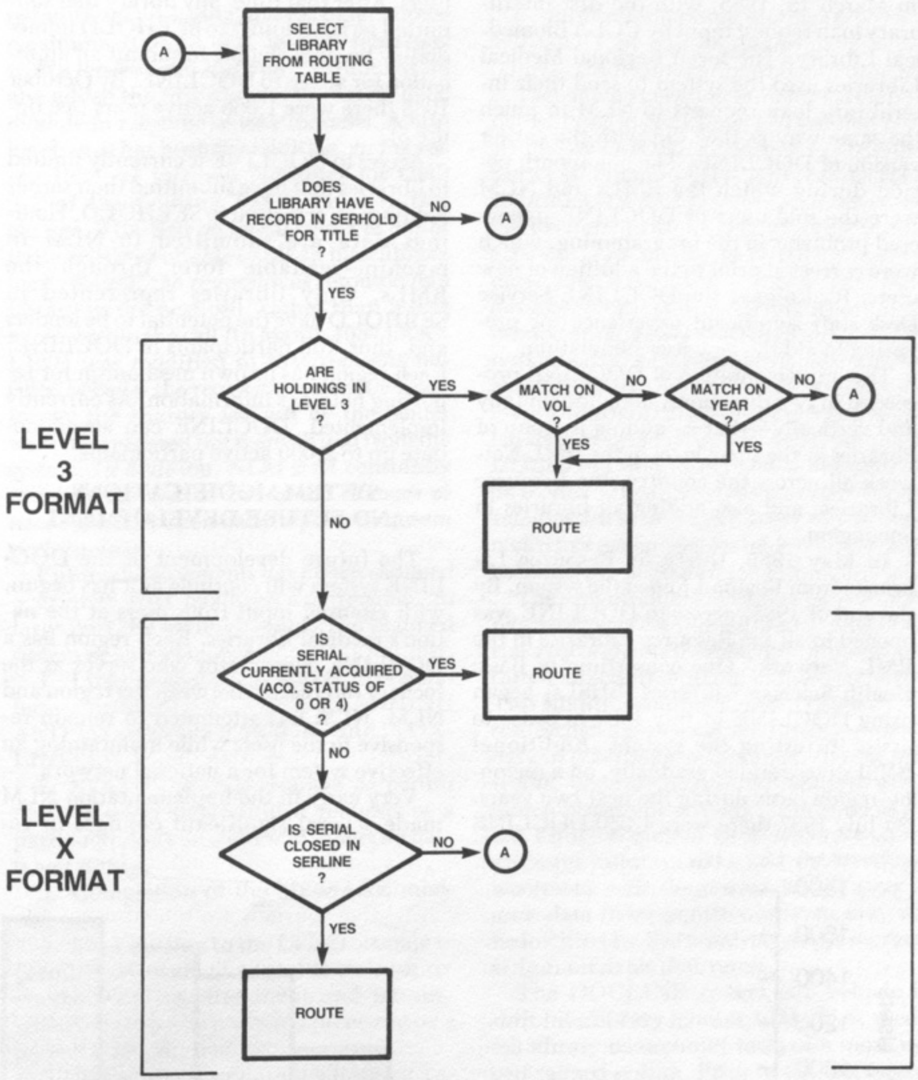


Figure 4. DOCLINE Routing Algorithm.

determine subsequent routing by using the STATUS module. If an ILL request is updated as Not Filled in the LEND module, the routing algorithm is activated. TIME-TRIGGERED ACTIONS also calls the ROUTING module whenever a request is rerouted due to nonreceipt or nonaction.

The ROUTING module is the main feature distinguishing the DOCLINE system from all other currently available interlibrary loan systems. In no other system is routing so automated with so little intervention required by a borrowing or referring library. The OCTANET system, which served as the forerunner for the routing table concept, used only ten lending libraries in the routing algorithm.

IMPLEMENTATION

The current DOCLINE began operation

on March 15, 1985, with the first interlibrary loan request input by UCLA Biomedical Library. The seven Regional Medical Libraries used the system to send their interlibrary loan requests to NLM in much the same way as they did with the earlier version of DOCLINE. The one-month period during which the RMLs and NLM were the sole users of DOCLINE uncovered problems in the programming, which were corrected prior to the addition of new users. It also gave the DOCLINE Service Desk staff significant experience for preparing to aid the new user population.

The implementation of DOCLINE proceeded in two directions, both horizontally and vertically—that is, adding a group of libraries at the same level in the RML Network all across the country, the Resource Libraries, and also adding all libraries in one region.

In May 1985, thirty-six Resource Libraries from Region 1 joined the system. By the end of 1985, access to DOCLINE was opened to all the Resource Libraries in the RML Network.[5] One consortium of Basic Health Sciences Libraries (BHSLs) began using DOCLINE in July 1985 in order to assist in testing the system. Additional BHSLs were added gradually, on a region-by-region basis during the next two years. By July 1987 there were 1,350 DOCLINE users. After that time, any library that submitted serial holdings to SERHOLD immediately became eligible to submit an application for access to DOCLINE. By October 1988 there were 1,668 active users (see figure 5).

Access to DOCLINE is currently limited to libraries that have submitted their serials holdings information to SERHOLD. Holdings data are submitted to NLM in machine-readable form through the RMLs. Only libraries represented in SERHOLD have the potential to be lenders and, thus, full participants in DOCLINE. Each region has its own mechanism for reporting holdings information. As currently implemented, DOCLINE can accommodate up to 2,000 active participants.

SYSTEM MODIFICATIONS AND FUTURE DEVELOPMENT

The future development of the DOCLINE system will continue as it has begun, with essential input from users at the nation's medical libraries. Each region has a DOCLINE coordinator who serves as the focus of interaction between the region and NLM. NLM has attempted to remain responsive to the users while maintaining an effective system for a national network.

Very early in the implementation NLM made several significant changes in re-

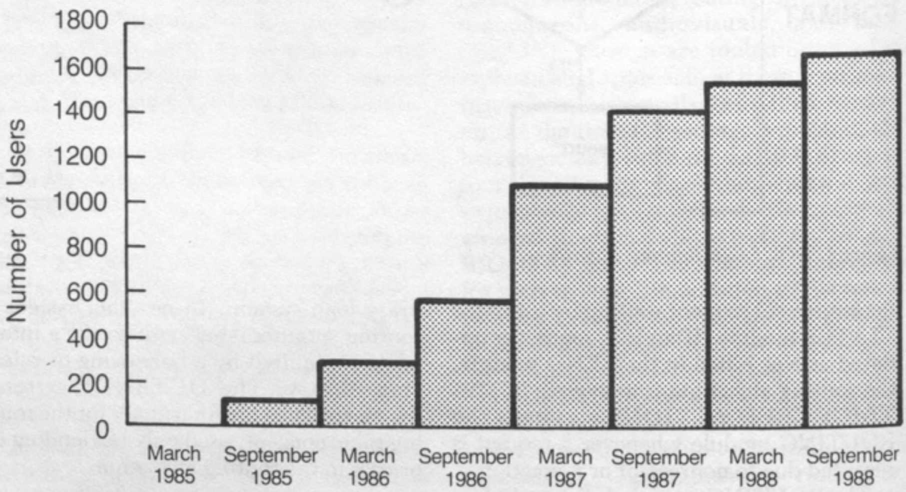


Figure 5. Growth in DOCLINE Users.

sponse to user requests, such as reducing the maximum time an ILL request remained with a lender from eight days to four. NLM also added the first phase of the MESSAGE module in response to user requests. A major change has been the addition, in February 1988, of intracell rerouting for Cells 3–7 of the Routing Table. Prior to this change, if the first potential lender in one of those cells was unable to fill an interlibrary loan request, the request was rerouted to a library in the next cell, rather than trying again within the cell. In March 1987 DOCLINE was modified so that requests unfilled because the requester was not willing to pay the charges assessed by the lender were not rerouted but were retired from the system. In addition, NLM staff constantly monitor the system and make changes as needed, based upon daily reports of system performance.

Future developments have been assigned a priority that has been determined, in large part, by network input. The planned system enhancements are

1. Links to additional MEDLARS databases, primarily HEALTH PLANNING AND ADMINISTRATION and TOXLINE;[6]
2. STATUS module changes to permit users to select only unfilled requests or a particular day's requests and to sort by date when listing;
3. Completion of the MESSAGE module;
4. Modification to the LEND module to permit borrowers to cancel their own requests any time after input and routing, when the request is not being processed by a potential lender; and
5. Increase in the size of the Routing Table from 80 libraries in seven cells to 120 libraries in ten cells.

Enhancements under consideration for the future include

1. Link to GRATEFUL MED, NLM's microcomputer software that assists health professionals in performing online searches;
2. Command stacking to move directly from one menu to another;
3. Microcomputer front end;
4. Gateways to other ILL systems; and
5. Expansion to other groups of users.

CONCLUSION

The design chosen for the DOCLINE system would not have been possible without the existing infrastructure of the RML Network. NLM and the system developers did not have to create any type of interlending network or structure, but merely used one that had already proven its effectiveness. The RML network and the RMLs have been essential in the design, testing, implementation, and management of this system at all stages. By using this existing network, NLM was able to implement the system without the problems inherent in creating a network at the same time. The DOCLINE developers could concentrate on technical system problems and were not called upon to deal with many of the organizational and policy issues associated with network development.

The use of DOCLINE has proven that relatively accurate automatic routing of serial requests can be achieved without detailed machine-readable holdings data. This finding has potential implications for the recording and use of serials holdings data in other types of resource sharing networks, even if they do not employ automated routing of requests. The level of local effort required to contribute more extensive holdings data and the overhead associated with the storage and display of such data in cooperative systems may not be justified by the relatively small increase in ILL referrals that result.

The DOCLINE system is a vehicle to shift interlibrary lending in the U.S. medical library community toward a more decentralized system. Prior to the full implementation of DOCLINE, a borrowing library would likely attempt to request material from perhaps two libraries within its region before requesting the material from NLM. The DOCLINE system facilitates more widespread request routing, with NLM serving only as the very last resort. Changes to DOCLINE will continue to encourage this decentralized approach. In FY 1988, nearly 1.3 million requests were entered into DOCLINE. Eighty-five percent of these requests were filled by network libraries. In addition, since the DOCLINE system is available to even the smallest medical library, materials are now more accessible to health practitioners throughout the United States.

REFERENCES AND NOTES

1. INQUIRE is a database management system licensed by NLM from the Infodata Corporation, Falls Church, VA.
2. William J. Willmering, Martha R. Fishel, and Dianne E. McCutcheon, "SERHOLD: Evolution of the National Biomedical Serials Holdings Database," *Serials Review* 14, no. 1–2: 7–13 (1988).
3. Millard F. Johnson, Jr., and Richard B. Pride, "OCTANET—An Electronic Library Network: I. Design and Development," *Bulletin of the Medical Library Association* 71, no. 2: 184–91 (Apr. 1983).
4. Edward Yourdan and Larry L. Constantine, *Structured Design: Fundamentals of a Discipline of Computer Program and Systems Design* (Englewood Cliffs, N.J.: Prentice Hall, 1979).
5. All the libraries in Region 4 joined DOCLINE simultaneously in May 1987, switching from their regional system, OCTANET, to the national system.
6. The link to the HEALTH PLANNING AND ADMINISTRATION database was completed in March 1989. ■■

The UC MELVYL MEDLINE System: A Pilot Project for Access to Journal Literature through an Online Catalog

Clifford A. Lynch and Michael G. Berger

The University of California has run a three-year pilot project to provide access to biomedical and health sciences literature through the integration of part of the National Library of Medicine's MEDLINE database into the University of California's MELVYL online union catalog.

This paper reviews the parallel but historically distinct approaches to providing access to monographic and journal literature and presents the MELVYL online catalog as the environment in which MELVYL MEDLINE operates. The pilot project itself is described, including its design, implementation, and evaluation to date. Finally, the paper introduces some of the current planning efforts to mount additional databases as part of the MELVYL catalog and discusses the key technical and policy issues that UC's experience with MELVYL MEDLINE has raised for this process.

For about a decade, libraries have been replacing traditional card catalogs with online catalogs to provide access to monographic holdings. The vastly improved capabilities of computerized information retrieval systems, coupled with the greatly increased access provided by computer networking technology, have opened libraries' book collections to library patrons. Libraries have done relatively little, however, to provide similar access to the journal literature, even though journal articles are perhaps more crucial than monographs in fields such as science and engineering and contribute a vital part of the literature in all fields of scholarship.

This situation is now changing. The increased patron expectations inspired by the new level of collection access offered by online catalogs have motivated libraries to attend to the journal literature. This paper describes a three-year pilot project at the University of California to provide access to biomedical and health sciences literature through the integration of part of the National Library of Medicine's MEDLINE databases into the University of California's MELVYL online catalog.[1]

A review of the parallel but historically distinct approaches to providing access to monographic and journal literature opens the paper, and a brief discussion of the MELVYL online catalog as the environment in which MELVYL MEDLINE operates follows. The planning and organization of the project and the design and implementation of the MELVYL MEDLINE user interface and database are then described. The

Clifford A. Lynch is Director and Michael G. Berger is Associate Director of the Division of Library Automation, University of California, Oakland, California.

MELVYL MEDLINE project includes a significant evaluative component, and the MELVYL catalog is instrumented to capture usage and performance data. Much has been learned about who uses the system, how it is used, and the successes and shortcomings of the design of MELVYL MEDLINE. Some of this data is reviewed, and the impact of MELVYL MEDLINE on the UC community is examined.

While the MELVYL MEDLINE project has been a major development effort in its own right and represents an important operational resource for the University of California, it is also a pilot project that is expected to pave the way for the mounting of databases to support other disciplines. The paper concludes with a discussion of some of the current planning efforts for such additional databases, key technical and policy issues at the center of this planning process, and the ways in which UC's experience with MELVYL MEDLINE has influenced this planning.

ACCESS TO THE JOURNAL LITERATURE

The explosion of knowledge in the twentieth century and the accompanying explosion in publication volume necessitated the differentiation between bibliography (for journal articles) and catalogs (for monographic material). Initially this separation of bibliography from the catalog was reflected in the use of card catalogs for books and of various printed reference tools for journal articles. The development and implications of this dichotomy and the prospects for repairing it through computer-based technology have been described by Michael Buckland, the assistant vice-president for library plans and policies at the University of California when the MELVYL MEDLINE project was conceived.[2-4] During the late 1960s and early 1970s, computers were providing access to the journal literature, while access to monographic holdings remained via 3-by-5-inch cards in virtually all libraries. The emphasis on the journal literature, primarily in science and engineering, was fueled by the funding supporting these disciplines during this period and by the interest from industry and well-supported parts of the research community in obtaining high-quality access to information in these disciplines. Commercial electronic database vendors and database utilities such as Dialog and BRS responded to these demands.

Automated library access to monographic holdings developed later. While the early 1970s saw the rapid rise of the bibliographic utilities and the wide adoption of the MARC cataloging formats, it was not until the late 1970s that libraries amassed machine-readable records for a significant part of their holdings, and the cost of computing technology dropped enough to enable the development of online public access catalogs. The late 1970s and the first few years of the 1980s saw the development of large online catalogs as research projects in major libraries. By the mid-1980s commercial (turn-key) versions of this technology were widely available, and online catalogs became commonplace in academic libraries and larger public libraries.

By the mid-1980s online catalogs had gained wide use in university settings, but the commercial database utilities had only minimally affected most university end users. There were several reasons for this. Online catalogs (at least the best of them) had set a comparatively high standard for ease of use by untrained searchers, while the commercial information retrieval services continued to be targeted for use by searchers trained in their complex and computer-oriented query languages and user interfaces. While access to online catalogs was made available at no charge to the patron (library operating funds supported the online catalog), searching a commercial information retrieval service represented a significant expense—as much as $50 to $100 per hour—which was directly recharged to the end user in many cases. (Some institutions did fund a limited amount of mediated searching on behalf of end users.) Although costs for the commercial services were occasionally factored into grant applications, very few academics had the funding to pay bills of this magnitude. Expanded access to journal literature for the academic community depended on three criteria that were not being met at the time by the commercial search services:

- Access to journal article databases would have to be available at no cost to the user, just as searching of the library catalog was free. As a corollary, given the massive unsatisfied demand for database access, meeting this criteria would have implied that the commercial utilities offer universities reasonable flat rates for essentially unlimited use.

• Databases would have to be easy to use; the user interface should be no more complex than that provided by an online library catalog. Further, some commonality of user interface would be highly desirable so the user would not have to learn a different query language for each database or when passing from the online monograph catalog to a journal article database.

• Access to the databases would have to be accomplished easily through the computer networks that are developing to serve the academic community (i.e., the Internet) rather than through public packet-switched networks such as Telenet or dial-up access. Online library catalogs, it should be noted, have often failed significantly in this area, and only in the late 1980s is network access receiving its share of emphasis in library automation planning.

Libraries examining the integration of journal article databases with library catalogs faced any number of formidable problems:

• Journal article databases are typically proprietary and often carry sizable license fees. In many cases, they are leased rather than purchased. Identifying funds and developing acquisitions policies to lease such databases was a major problem for most libraries.

• Many database vendors wanted (and still want) to apply usage-sensitive pricing schemes, i.e., the more a database is used, the more the library pays. This is an unacceptable arrangement for most libraries (unlike the commercial information utilities, which merely pass on the costs to their customers), where there is no way to generate income in proportion to database use.

• Unlike bibliographic records, which are well standardized through the MARC communications format, there is no standardization for journal article abstracting and indexing record data elements. There are great variations in indexing practice, the use of thesauri and other classification schemes, and the use of controlled vocabulary. Coverage among different databases overlaps in untidy ways: Some databases cover some journals exhaustively, and some cover selected journals selectively. If one licenses multiple databases, one may well have a number of records for the same article. Another complication of the wide variation in format and content of these databases was that there was very limited software available commercially that could be used to mount abstracting and indexing databases onto a local library's computer. In addition, most turn-key online catalogs could not load journal article abstracting and indexing databases, so a new supplementary software system had to be incorporated.

• Compared to MARC-based bibliographic databases, journal abstracting and indexing databases are enormous, particularly those that cover a wide range of disciplines. Individual records in these databases may be much larger than typical MARC records because they include abstracts. There is an implied need for massive disk storage to house the databases and significant computational capacity to maintain them and to provide access. In many cases, the system support requirements for the new databases overwhelmed the capacity available on local systems supporting online catalog functions.

• The effective deployment of journal article databases in a university environment created demands for new support services or the massive expansion of existing services (such as document delivery). A higher level of integration between library computing and the overall computing environment is needed to accommodate the increased demand for downloading records to workstations and receiving current-awareness services through electronic mail. Availability of journal databases would fuel demand for such services.

These are daunting challenges for libraries, perhaps even more overwhelming than the prospects faced in the late 1970s and early 1980s, when online catalogs and the retrospective conversion of monographic collections to create databases for online catalogs were first considered. These challenges must be considered in a financial environment already strained by the requirements of other automation projects and by pressure from the acquisitions budget in the face of rapidly rising costs for books and particularly for journals. Nonetheless, journal literature must be electronically accessible. There may be increasingly little value in subscribing to any journal not indexed by a readily accessible database. Further, while all students, faculty, and researchers in the humanities and social sciences have benefited from online catalogs, these systems have had less impact on advanced students and researchers in the sciences and engineering, who make extensive use of the journal literature. These members of the university community must be supported as well.

THE UNIVERSITY OF CALIFORNIA MELVYL ONLINE UNION CATALOG

As a result of a planning process initiated in 1979, the University of California developed the MELVYL online union catalog, which consolidates the holdings of the nine UC campuses into a single database.[5,6] This system was deployed in prototype in 1981 and in production form in 1982, bringing together cataloging from over thirty cataloging units within UC created through OCLC, RLIN, and various local campus systems. In the mid-1980s the California State Library also became a MELVYL catalog participant. The system has grown into one of the largest online catalogs in the world, containing over ten million holdings (five million unique titles) and over a million periodical holdings (some 620,000 unique serials titles). The catalog also contains records for nonbook materials (e.g., sound recordings, maps, and visual materials). During the academic year, the system services over a million queries a month and displays some nine million records a month to its user community.[7-10] The MELVYL catalog is supported by a large-scale, TCP/IP-based, packet-switching network that is integrated with the UC campus local area networks and the national TCP/IP internet, providing convenient access to the catalog not only from in-library terminals but from computers and workstations located throughout the UC system and nationally.[11]

The MELVYL system was designed for easy use by the general public, without training, from any type of lowest-common-denominator, line-by-line ASCII terminal. It supports both a simple menu-based interface called LOOKUP mode that is intended for the casual user and a powerful command language–based user interface. Hundreds of help and tutorial screens are available. While the system is used successfully without training, it is supported by an extensive user education program conducted by the UC libraries; many users find that these training sessions enrich their use of the catalog.

The initial concept of the MELVYL catalog, as developed in the 1979 Plan for Development, dealt with the monographic literature only. By 1982 the Division of Library Automation (DLA), the universitywide unit that develops and runs the catalog, was already considering extending the catalog to cover the journal literature.[12] Funding and computing resource constraints, as well as the vast amount of work still needed on the online catalog proper, precluded any serious efforts in this area until the mid-1980s. At that time, funds were obtained through a Medical Library Resource Project Grant (NLM Grant #GO8 LM04466) to support a pilot project to mount the most recent three years of the National Library of Medicine's MEDLINE database as a part of the MELVYL catalog. MEDLINE seemed an ideal choice for a number of reasons: The biomedical and health sciences community was already one of the larger users of journal article databases (MEDLINE and others) through commercial search services, and MEDLINE had a well-established constituency. Further, library use in these disciplines leans heavily toward journal article literature; many biomedical libraries actually have more journals than books. Obviously, MEDLINE was an important resource for this community. Within UC many campuses had biomedical or health science libraries, providing a strong focus for the instruction and evaluation of a UC MEDLINE implementation.

ORGANIZATION AND PLANNING FOR MELVYL MEDLINE

From the very beginning, MELVYL MEDLINE was a cooperative effort between the health science libraries and the Division of Library Automation. The then director of DLA, Edwin Brownrigg, and the assistant university librarian at the UCLA Biomedical Library, Mary Horres, were coprincipal investigators for the grant. A Medical Library Task Force of health science librarians guided DLA during the development and implementation of MELVYL MEDLINE at campus sites. In addition to the task force, a new librarian training coordinator position provided support for implementing the project in UC libraries. The task force also convened a Users' Advisory Council of end-user faculty, staff, and students to review prototypes and development in progress to ensure that the MELVYL MEDLINE implementation met its users' needs.

The five medical libraries at UC have made significant staff and financial resources available to support the project. These medical libraries contribute half the funds for the database license; the other half comes from universitywide shared acquisitions funds (traditionally used to purchase microfiche sets and similar systemwide resources).

This collaborative approach reflected the lessons learned during the development of the original MELVYL online catalog. During the initial development of the MELVYL catalog, little consideration was given to the issues of user training and the impact the MELVYL system would have on library operations. Insufficient review of the system occurred prior to deployment, resulting in many system changes based on feedback from librarians at the campuses. (Some of the initial MELVYL design errors could have been identified and corrected prior to system release through a more extensive review process, although many could not. There was no base of experience to help evaluate systems such as the MELVYL catalog at that time. And much of what we learned through the deployment and evaluation of MELVYL MEDLINE could only be determined through experiment and experience.)

The MELVYL MEDLINE project was planned as three phases. The first phase involved analyzing MEDLINE functions and data structures, developing programs for loading and maintaining the MEDLINE database as part of the MELVYL system, and defining extensions to the existing MELVYL user interface and command language to support MEDLINE. At the conclusion of this first phase the university leased a three-year subset of the database from the National Library of Medicine. Because the university collectively holds about 90% of all journals indexed in MEDLINE, articles for all journals were included, thus committing UC to approach MEDLINE as a bibliography rather than an extension to traditional cataloging (i.e., a summary of UC's holdings alone). The size of the UC system and the elaborate UC interlibrary loan system actually blurs the distinction between catalog and bibliography; basically, one can obtain what was intended as a bibliography and present it to the user as both catalog and bibliography. In contrast, many institutions mounting MEDLINE have in fact treated it as an extension of their catalogs by including only locally held journals.

The second phase of the project made the MELVYL MEDLINE system available to the health and life science libraries as a prototype. The librarians at these facilities worked closely with DLA in introducing the system, testing training materials, and providing rapid and detailed feedback as end users began to use it. The third phase of the project, which started in the summer of 1988, made the MELVYL MEDLINE system available in all UC libraries and to the entire UC community through campus local area networks and direct dial-up.

Implementation was an iterative process that involved continuing refinement of the specifications and subsequently of the system itself. The design of the system for opening day (November 1, 1987) targeted a nearly full-featured system to avoid retraining users as additional major features were implemented. Nonetheless, the system remained dynamic as information about how MELVYL MEDLINE was actually being used was analyzed and translated into appropriate system enhancements.

REMOTE ACCESS TO MELVYL MEDLINE

Providing access to MELVYL MEDLINE from terminals and workstations outside the UC libraries required special planning considerations. The MELVYL online catalog is a true public access system: One can access the system through dial-up or across the internet without a user ID. The license agreement for the MEDLINE database required that access to the MEDLINE database be restricted to UC faculty, staff, and students. Yet we had no intention of completely converting to a user ID/password system, since it would have created an impossible administrative burden and an unnecessary barrier to easy MELVYL catalog access.[13] To resolve this dilemma, we created a password system for MELVYL MEDLINE and profiled each access point within the university system. If a UC campus is prepared to guarantee that access to a specific host or network (or subnet) within the UC internet is limited to members of the UC community, then for MEDLINE sessions originating from that location or set of locations no password is required. From other locations where such guarantees cannot be made, the user is prompted for a password before access is given to the MEDLINE database. Passwords are issued by the UC libraries and are valid for a fixed period of time (one to four years) or indefinitely (for tenured faculty, for example), depending on individual campus policy. The user must agree to keep the password confidential and to use it only for personal work.

THE DESIGN OF MELVYL MEDLINE

The vision for MELVYL MEDLINE was to extend an existing online catalog that was familiar to much of the user community within UC. Creating a completely new user interface would have required users to learn and master new search techniques and commands. The duplication of existing vendor-based interfaces would have been familiar to some part of the potential MEDLINE user community within UC but would have been completely foreign to the rest of the user community. In addition, we felt that existing user interfaces were not really easy to use. There were several basic design assumptions for adapting and extending the MELVYL catalog to support MEDLINE:

- MELVYL MEDLINE must provide all the essential MEDLINE features available in existing MEDLINE implementations. A full system can thus satisfy the needs of the segment of the university community already using sophisticated information retrieval features available in these implementations. (In hindsight, we may have overestimated the amount of use that these sophisticated features received and the importance of replicating them in our implementation. This important issue is discussed later.)
- MELVYL MEDLINE must use the existing MELVYL user interface structure and command language, adding new features or commands only to implement new capabilities that have no analogue in the existing MELVYL catalog.
- MELVYL MEDLINE must first meet the needs of end users in the health and biomedical sciences and secondly introduce MEDLINE to users in other disciplines. The interdisciplinary nature of both the MEDLINE database and the UC research and instructional programs requires a general approach, not a narrow focus on the life and health sciences.
- The system must complement and support library efforts to teach MEDLINE and other medical bibliographic topics. Use of MELVYL MEDLINE can be only an introduction to a comprehensive use of medical informatics.

MELVYL MEDLINE must be a separate database that users can select with the MELVYL catalog. Having a separate database facilitates development both for the MELVYL catalog and MELVYL MEDLINE, limits the impact of MEDLINE on existing MELVYL catalog users who are not interested in the MEDLINE database, and reduces system load. (This design assumption was debated at great length prior to adoption. We were deeply concerned about the huge retrievals generated by users of the MELVYL catalog on a five-million-record database and did not want to aggravate this problem by enlarging the result set with journal articles from MEDLINE unless the user wanted them. We also felt that it would be confusing if the user could obtain journal article records only from MEDLINE but not from any other field. Finally, the performance of two separate databases was judged to be better than one single merged database, as the indexes would be smaller and could be searched much more quickly. Ultimately, these largely pragmatic considerations came to dominate the various theoretical arguments that can be posed for a merged database. We are not entirely comfortable with this decision based on expediency, but the results work.)

- MELVYL MEDLINE would contain a limited subset rather than the entire MEDLINE database. Several factors were involved: the availability of the full database on commercial services; the performance problems and disk space costs inherent in mounting the full database; and the heavy emphasis on recent material in most user searching of MEDLINE. We chose to start each year with two full calendar years of the database and to grow a third year into the database, so that at the end of the year the database contained three years of coverage. We are now planning to extend this coverage to four years, building to five, because the user community desires access to more than three years of the database. We seem to have overemphasized the degree to which people want recent material. There is a sizable part of the user community that feels even five years will be inadequate and would prefer ten years. Some of the erroneous assumptions we made here may have arisen from an attempt to extrapolate experience with the use of MEDLINE by rather specific communities—such as clinical medicine—to the very large and diverse UC user base.
- Because the stability and performance of the MELVYL online catalog is so critical to the overall university community, MELVYL MEDLINE must have minimal impact on the performance and stability of the MELVYL catalog as a whole.

We did choose to violate one of these principles—maximum compatibility with the existing MELVYL user interface—in the design of the menu-based user interface for MEDLINE. The MELVYL system has two interfaces, a command-language mode and a very simple menu-based mode called LOOKUP. We had believed for some years that LOOKUP mode was too limited and needed extensive revision, and the MEDLINE project became a testing ground for many of the concepts we hoped to incorporate into a revised LOOKUP mode. This new mode was called ASSIST mode. We felt that the variations between the LOOKUP and ASSIST modes would not be a major problem for our user community, since both of these modes guide the user and do not rely heavily on prior knowledge of system use.

THE MELVYL MEDLINE USER INTERFACE

Working with the specifications from the MEDLINE Task Force and an analysis of functions in other MEDLINE implementations, DLA attempted to match required MEDLINE functions to the existing MELVYL command language. The MELVYL catalog uses two primary commands: FIND (for searching a number of indexes—including personal author, title, subject, and series, including both keyterm containment and exact matching—and supporting Boolean combinations) and DISPLAY (for displaying retrieved records, with user-specified levels of completeness, at the terminal). The catalog also offers a BROWSE command to examine authority-controlled headings for personal and corporate authors, series, and subjects and a SELECT command to use such headings as retrieval criteria. Other commands are available for such functions as requesting HELP and saving results. While the majority of these functions could be translated directly to the MEDLINE database, there were a number of MEDLINE functions that had no analogue in the MELVYL catalog command language. For example, *MeSH* tree number searching, subheading display and manipulation, and navigation through broader and narrower headings all required command language extensions. The ability to "explode" headings (i.e., to search a heading and all of its subordinate headings, a feature offered by other MEDLINE user interfaces) and the ability to search on the major subject of an article (as designated by NLM indexing) also required extensions to the command language.

The basic structure of the MELVYL interface was retained as well: A user chooses the MEDLINE database and has the option of directly issuing commands, choosing a brief online tutorial, or selecting a menu approach that guides the user through searching for articles, browsing *MeSH* subject headings, and displaying records. The MELVYL context-sensitive HELP facility was maintained, as was the EXPLAIN system, which lets a user receive an explanation of command language verbs and options or other terms at any time without disturbing the search context.

The major development challenge was providing access to *MeSH*, MEDLINE's controlled indexing vocabulary. *MeSH* is much more tightly controlled than the Library of Congress subject heading scheme, which underlies subject searching in the online catalog. *MeSH* headings are arranged in hierarchical tree structures of related headings, with broader headings at the top of the tree and narrower ones at the bottom. The majority of these headings are further qualified by subheadings that are standardized aspects of a subject, such as "complications," "diagnosis," and "etiology."

In the menu/system-guided dialogue mode, called ASSIST mode in the MEDLINE user interface, users can search for keywords taken from titles, subject, or abstracts or BROWSE *MeSH* headings to determine appropriate heading or heading/subheading combinations. The keyword search, which is very much like searching in the MELVYL catalog, is advertised as a quick approach to finding information by subject; it is helpful if a user does not know the precise subject heading for a topic or if NLM has not yet assigned a specific *MeSH* heading (e.g., "designer drugs" or "sick building syndrome"). When the keyword search option is selected in ASSIST mode, the user has the option of listing only title and *MeSH* headings for the retrieved articles; this parallels the library instruction technique of looking at subject headings of books to determine appropriate subjects for a further search. Users also have the option of selecting a "find more/find less" function to further control search results. This begins an attempt to provide a more intelligent user catalog interface and will ultimately move into the MELVYL catalog as well as MEDLINE. For small results (less than twenty citations), the user has the option of expanding the search term list (the system executes an OR function); for large results, the user can limit the search term list by supplying more keyterms (automatic AND), an author name or publication date, language, or "check tags" (i.e., a common MEDLINE-specific indexing element, such as "human," "male," or "review").

If the ASSIST mode user chooses the BROWSE approach rather than direct keyterm searching, supplied terms are matched against the *MeSH* vocabulary. The system introduces the BROWSE function and most common options and suggests the most productive way to search for *MeSH* terms. Users may first try a sample BROWSE sequence to see how BROWSE works prior to starting their own BROWSE. The result of a BROWSE command returns all matched headings; terms are searched against a keyterm pool derived from all subject headings and subheadings. These headings and subheadings are listed in alphabetical order, along with an indication of the number of articles associated with the heading or subheading. Once a list of headings is obtained, the user can employ the SELECT command to choose all articles matching one or more headings in the displayed list. From ASSIST mode, the user can then rejoin the interface logic paths for display and more/less result management common to both modes.

ASSIST mode functions are a subset of the functions available to users of the full command language mode. About 80–85% of the users recognize the advantages of learning a relatively simple command language. By design, users are encouraged to learn the command language, which is considerably more flexible and powerful than ASSIST mode searching. Library training programs and online and printed documentation at the terminals also emphasize the use of COMMAND mode. The ASSIST mode is designed not only to guide the user who does not know the COMMAND mode but also to introduce that user to the most common commands, thus easing the ultimate transition to COMMAND mode. An ASSIST mode user has the option of switching to COMMAND mode at any time. As with the MELVYL catalog, the listing of available modes in the introductory screens encourages the user to choose COMMAND mode, and as with the MELVYL catalog, users can view a brief online guide to commands upon entering COMMAND mode. Many users, particularly those with prior experience with computers, immediately select COMMAND mode, review the brief tutorial, and begin searching.

THE MELVYL MEDLINE DATABASE

Records received from the National Library of Medicine in its standard distribution format are restructured into DLA's MARC format, which involves creating a leader, a directory, and MARC-like variable fields containing tags, indicators, and subfields. However, this restructuring is merely one of communications format. It is not a semantic map from NLM's data elements to the MARC data elements. Early in the MELVYL MEDLINE project the decision was made to restructure records rather than to map MEDLINE data onto existing MARC-content designators directly and then process the resulting MARC records through existing input streams to the MELVYL catalog. Having data in the MARC format structure allows the use of many existing programs while maintaining the full specificity of NLM tagging. In addition, the use of NLM tags allows us to continue to rely on the NLM documentation in sharing information about the system with the UC community already familiar with the NLM database.

The database is recreated yearly to include new *MeSH* headings; after the *MeSH* file is reloaded, snapshots for the two preceding years are reloaded. Because only a three-year file is maintained, we felt it was cost-effective to use a design based on reloading the file annually rather than to develop all of the complex logic necessary in the load programs to perform annual maintenance as the *MeSH* headings change. Even with the larger four-year file currently under development, the reloading should not consume unreasonable computational resources, although we are finding we need more and more transient disk space to support the file regeneration process.

The MEDLINE data is stored as part of the MELVYL database. Loading and maintenance are performed by a complex set of PL/1 programs; Software AG's ADABAS database management system is used to manage the database. The full three-year MEDLINE database, including indexing, uses about five gigabytes of disk storage in our implementation.

AN EVALUATION OF THE MELVYL MEDLINE DESIGN

MELVYL MEDLINE became operational in November 1987 in the UC health and life sciences libraries. On campuses that did not have such a library, a specific location within some other library and a MELVYL MEDLINE coordinator were designated. In the summer of 1988 access was extended to all libraries in the UC system; in September 1988 access was further extended to include campus networks and direct dial-up by UC faculty, staff, and students. Use of MELVYL MEDLINE has been high both in the expected high-use locations (the medical, health science, and life science libraries) and in the other UC libraries. Remote use through campus networks and direct dial-up has been growing steadily. Table 1 shows the breakdown of use from different sources for January through March 1989.

Use of MELVYL MEDLINE generally parallels that of the online catalog. MELVYL MEDLINE users are more willing to use COMMAND mode (about 15% use ASSIST mode on MELVYL MEDLINE, while 27% use the LOOKUP mode on the MELVYL catalog). In both systems, users prefer the simple keyword approach with titles, subjects, and abstracts to the more elaborate BROWSE approach. Average result sizes both for the catalog and for the MEDLINE database are enormous—average result size is 140–160 for the MELVYL catalog and 400–500 for MELVYL MEDLINE. The large size of the MELVYL MEDLINE retrievals is particularly impressive because the MEDLINE database contains only about 660,000 citations, and the MELVYL catalog contains about five million citations. The huge retrievals may be due to the rich indexing of journal articles by NLM and the availability of generic keyword searching for the MEDLINE file which searches abstracts, subjects, and titles.

Few users exploit the more advanced features available in MELVYL MEDLINE; only 2.4% use the BROWSE command, for example. This is better than in the online catalog user community, where BROWSE is used in only about 1 out of 200 sessions. Less than 4% of the searches in COM-

MAND mode (the only mode in which the features are available) employ options such as EXPLODE or major subject headings. The reliance on the most basic features of the system has a number of possible causes:

- Users are unwilling to invest the time to learn the more advanced features, perhaps because they are unconvinced of the value of these features. Yet they are willing to wade through enormous result sets as an alternative. Currently about half of the users display less than 25 citations in a session; about 25% display 25–50 citations; about 15% display 50–100 citations; and, incredibly, about 15% display over 100 citations in one sitting.
- The average MELVYL MEDLINE user does not need to make a comprehensive search and is satisfied with only some results (which are easily produced by the keyword approach). It is possible that the user simply does not have to be either comprehensive or highly selective, or perhaps the user accomplishes the highly selective search by another means, such as scanning the bibliographies of consulted articles. Clearly, we need to better understand user goals as they approach such systems.
- The interface fails to advertise effectively the advantages of the more sophisticated searching features and to make those features easy to use. This is probably more a problem of making the user aware of the features than a problem with the commands themselves. Most MELVYL MEDLINE searchers do not view searching from the same perspective as a trained NLM researcher; complex Boolean requests, explodes, and check tags are foreign to them. In ASSIST mode, however, where users are prompted and guided to use *MeSH*; to use Boolean searching; to limit by check tag, date, and language; and to consult subject headings for leads to other relevant subject headings, users incorporate these features slightly more than their COMMAND mode counterparts. But even in ASSIST mode the numbers are small: BROWSE makes up 3.7% of the commands in ASSIST mode, compared with only 2.4% of the commands overall. Users in ASSIST mode perform 27% of the limit operations by date, language, and check tag, while accounting for only 15% of the total MEDLINE use. Interestingly, users in a novice mode are willing to try more advanced features, which clearly suggests that there is room for a great deal more development.

Evaluation of the MELVYL MEDLINE system was a major part of the overall project plan and included not only an analysis of statistics and command logs but also an investigation of how users viewed MELVYL MEDLINE in context. The full results of this evaluation are reported.[14] Data from the evaluations is available as part of the MELVYL MEDLINE final report, available from the Division of Library Automation. A few summary statistics give a feel for the high degree of acceptance, however. Based on 1,614 questionnaires distributed at eight medical libraries in November 1988, we found that 94.7% of the users were either "satisfied" or "very satisfied" with the system; 79.8% of the users claimed that it allowed them to find journal articles more quickly; 54% said they were able to find more relevant material; and 81% found and read more articles from outside their specialty. A full 94.9% of the users stated that MELVYL MEDLINE had changed the way they used the library. Of the users surveyed, 74% said MELVYL MEDLINE was the first online database they had used; 84% had used the system before; and 29.8% had used it more than fifteen times.

Graduate students are the largest group of users in the population polled by the questionnaire (38%), followed by house staff, residents, and postdoctoral students (16.9%). Undergraduates, surprisingly, were a full 16.3%. (A March 1989 questionnaire, for which we received 2,173 responses, showed undergraduate use at 23%.) Faculty or academic staff constituted 13.7% of the use.

While generally satisfied, users were also realistic about problems with MELVYL MEDLINE. In addition, there was a great demand for additional years of coverage (now being addressed), more instructional materials, and better printing facilities.

Users also revealed something about themselves in their answers. Over 50% of the respondents were aware of the more sophisticated features in MELVYL MEDLINE, but a much smaller percentage reported using the features. Of the users who offered suggestions for improvement of the system, 11% requested features that were in fact already available. A good part of the MELVYL MEDLINE user community uses the system regularly. In the March 1989 survey mentioned above, 39% of the users (and 51% of the users searching the system from terminals or workstations outside the library) said they had used the system at least fifteen times before. Only 17% (6% of the remote users) said this was their first time using the system. Forty-eight percent of the total users (and 70% of the remote users) said they expected to use the system at least once a week in the future. Another 40% said they expected to consult the catalog at least once or twice a month.

User comments about database coverage in the March 1989 evaluation were particularly interesting. Forty percent felt that more historical coverage of the database was needed. An earlier survey (in November 1988) attempted to probe just how extensive a back file of MEDLINE was desired by the UC community: About 60% of the users would be satisfied with five years of coverage, but over 20% of the users wanted ten years. In the March 1989 survey, while 30% felt that they were retrieving too many irrelevant citations, 19% said that they needed citations on topics not covered by the MEDLINE database. Fifteen percent wanted more recent citations. In response, we will be receiving tapes from the NLM biweekly instead of monthly in the coming year.

THE IMPACT OF MELVYL MEDLINE

As anticipated, use of the MEDLINE database by UC faculty, students, and staff has reached unprecedented levels. In February 1989 we processed over 150,000 MEDLINE searches and displayed 1.8 million journal article citations to the UC community. There has been a corresponding decrease in the use of Index Medicus and direct searching of MEDLINE on NLM and other utilities. The MELVYL MEDLINE system has also made it possible to consult MEDLINE at the reference desk and thus speed up the reference process.

The impact of MELVYL MEDLINE on the medical libraries is harder to assess. As part of the project, the librarian training coordinator provided printed materials and training for nonmedical library staff at each campus to ease the burden of educating users when the system became generally available. It is clear from the accompanying table that MELVYL MEDLINE is accessed widely from outside the traditional constituencies; anecdotal evidence also suggests that many people from related disciplines such as anthropology and even traditionally distinct areas like physics have discovered that MELVYL MEDLINE can be a valuable resource in their research. The relatively high use of the system by undergraduates also suggests that MEDLINE is finding a reception beyond its traditional user community. Faculty members are examining how resources such as MELVYL MEDLINE should be explicitly integrated into academic programs. It will be interesting to see how far MEDLINE extends its influence.

The implementation of MEDLINE has clearly placed additional demand on the libraries. Reference desk traffic, particularly in the biomedical and health science libraries, has increased substantially. Demand for terminals has increased, causing congestion at some locations until the number of terminals can be increased. Some libraries in the UC system have chosen to constrain access to MELVYL MEDLINE from in-library terminals since they feel their resources (terminals and/or staff) are inadequate. There is considerable concern about the increased load in training and user support that the system is generating.

PLANNING FOR FUTURE DATABASES

MELVYL MEDLINE has successfully made access to the journal literature in the medical and health disciplines widely and easily available to the UC community. UC is actively engaged in planning for the acquisition and implementation of additional journal article databases. Knowledge gained from the MELVYL MEDLINE project illuminates a number of issues that must be carefully considered:

- Coordination with the libraries is essential to the success of such efforts, as are instruction and training. Extensive library participation in user interface design is of central importance.
- There is an enormous unsatisfied demand for access to the journal article litera-

ture. This translates into demand for terminals and for reference desk support, as well as for computing resources.

• While electronic access to the journal literature is still limited, the availability of databases will cause uneven use of collections. Collections indexed by a database will be used more than those for which indexing is unavailable. Thus, there is considerable interaction between acquisitions policy and index database selection.

• When removed from all of the economic and technical distortions caused by the expense and difficulty of using the commercial online utilities, there is a good deal of evidence that databases traditionally considered as specialized resources of specific subject areas will be used in a highly multidisciplinary fashion and that mounting databases without these distortions will promote multidisciplinary use of the databases. We must be cautious in extrapolating our experience with small user communities on commercial services when making decisions about databases to mount and about coverage for selected databases, and we must continually reassess these decisions in light of empirical data.

• It is unclear whether one should pay the development cost to replicate all of the elaborate searching features used by trained searchers on the commercial utilities when implementing a database as part of an online catalog. A sizable portion of the implementation cost of MELVYL MEDLINE supported the development of *MeSH* heading BROWSE, explode, and similar capabilities that are not exploited. We must convince users that these features are valuable and successfully educate them, or we must decide whether implementation of such features is reasonable in a public access database implementation. One factor arguing for the inclusion of the sophisticated features is that their use approximates mediated or librarian use of the catalog. If these features are not provided, will the interface meet the needs of the established user community for a given database? (On the other hand, such a user community is tiny, if vocal—how important is it to meet the needs of this community?) This particularly difficult issue may be divisive in the future: Many users will opt for simpler implementations of more databases sooner; librarians may argue for a slower but more comprehensive implementation of each database.

• It is also unclear whether it makes sense, in a major research university such as UC, to assign a high priority to the acquisition of databases that are primarily oriented for undergraduate use. MEDLINE is intended primarily as a research-level database, yet a sizable percentage of the users seem to be undergraduates. In a period of limited resources and infinite demand, the value of diverting efforts into specifically undergraduate-oriented databases is not clear. Balancing this observation, however, we have no empirical data about use at UC of undergraduate-oriented databases.

CONCLUSIONS

The MELVYL MEDLINE system has been available for less than two years, and its level of use is impressive. The implementation has been successful, and the database is fulfilling an important need for the university.

There is still a great deal of work to be done with MELVYL MEDLINE, however; the system is far from mature, and the user interface will be upgraded based on usage information for many years. We have yet to explore the full promise of integrating journal article databases with an online catalog and a university computing environment, which will lead to functionality that commercial information retrieval systems and CD-ROM implementations cannot offer. In late 1989 or early 1990 we will link the MEDLINE database to the university's union list of serials (CALLS), so that users can determine where retrieved journal citations are held. Additionally, we hope in 1990 to build a current-awareness service that allows users to register interest profiles and delivers updates through the university's electronic mail system. Downloading from the database into personal computers or workstations is still crude; improvements to these facilities would be a major boon to the user community. Finally, at some point we shall have to link MELVYL MEDLINE and the university's developing document delivery systems.

Ultimately we will have to face some of the hard design questions that we have been able to sidestep with a pilot project. Imagine that fifty journal article databases, rather than just one, are available as part of an online catalog. Should we continue to present them to the user as fifty distinct databases, or should we attempt to merge them, either into a single journal article database or even into a single bibliographic database that also includes the current material from an online catalog? This will become a key issue in the late 1990s as public access journal article databases proliferate. Under the current approach of fragmenting databases, we can expect that users will have at least as much difficulty figuring out where to look for information as finding it once they have picked a database.

ACKNOWLEDGEMENTS

We would like to thank Vice-President Cornelius Hopper and Associate Vice-President Richard West; Edwin Brownrigg and Michael Buckland for their support throughout the grant; Frank Arana, Elizabeth Baer, Debora Boyd, Robert Brandriff, Carol Connolly, Karen Coyle, Jayne Dickson, Ellen England, Laine Farley, Linda Gallaher-Brown, Nancy Gusack, Karen Helm, Lisa Hirayama, Kris Hoye-Brown, Fathilah Kamaluddin, David Lindholm, David McCutcheon, Teresa Montgomery, Mary Jean Moore, Sean O'Hara, Sherrie Pierson, Bennett Price, Barbara Radke, Dennis Smith, Meriel Stewart, Michael Thwaites, Margery Tibbetts, Ann White, David Wills, and Deborah Wolfe for their work on the MELVYL MEDLINE project, and the UC Campus Task Force, especially Mary Horres and Susan Starr. Our thanks also to Nancy Gusack for her editorial assistance.

REFERENCES AND NOTES

1. MELVYL is the registered trademark of the regents of the University of California. MEDLINE is the registered trademark of the National Library of Medicine.
2. Michael K. Buckland, "Bibliography, Catalogs, and Library Records," *Library Planning Discussion Paper DP-86-5* (Berkeley, Calif.: Library Plans and Policies, Office of the President, University of California, July 16, 1986).
3. Michael K. Buckland, "The Vanishing Library Catalog," Samuel Lazerow Memorial Lecture, Florida State University, 1987.
4. Michael K. Buckland, "Bibliography, Library Records, and the Library Catalog," *Library Resources & Technical Services* 32, no.4:299–311 (1988).
5. *The University of California Libraries: A Plan for Development 1978–1988* (Berkeley, Calif.: Office of the Executive Director of Universitywide Library Planning, Office of the President, University of California, July 1977).
6. Stephen R. Salmon, "Planning for Library Automation: The Origins of MELVYL," *Information Technology and Libraries* 1:350–58 (Dec. 1982).
7. Division of Library Automation, *MELVYL Reference Manual* (Berkeley, Calif.: University of California, Apr. 1985).
8. Clifford A. Lynch and Robert K. Brandriff, "The Evolution of the User Interface in the MELVYL Online Catalog, 1980–1985," *Proceedings, Forty-Eighth ASIS Annual Meeting, Las Vegas, Nevada, October 20–24, 1985* (White Plains, N.Y.: Knowledge Industry, 1985), p.102–5.
9. Barbara S. Radke, Katharina Klemperer, and Michael G. Berger, "The User-Friendly Catalog: Patron Access to MELVYL," *Information Technology and Libraries* 1:358–71 (Dec. 1982).
10. Charles Hildreth, *Online Public Access Catalogs: The User Interface* (Dublin, Ohio: OCLC, 1982).
11. Clifford A. Lynch, "From Telecommunications to Networking: The MELVYL Online Catalog and the Development of Intercampus Networks at the University of California," *Library Hi Tech* 7, no.2:21–47 (Apr. 1989).
12. Edwin B. Brownrigg and Clifford A. Lynch, "Online Catalogs: through a Glass Darkly," *Information Technology and Libraries* 2:104–15 (Mar. 1983).
13. Clifford A. Lynch and Edwin B. Brownrigg, "Public Access Bibliographic Databases in the Multicampus University Environment," *Proceedings, 1987 International Conference on Databases in the Humanities and Social Sciences*.
14. Susan Starr, "Evaluating the MELVYL MEDLINE Project" (Paper presented at the Medical Library Association Annual Meeting, Boston, Massachusetts, May 1989). ■■

The Development and Impact of a Global Information System

Douglas A. Kranch

Three revolutionary technological changes will be required to bring about affordable individual access to global online information: efficient large-scale database construction and maintenance, high-speed digital transmission networks, and highly precise intelligent searchware. As these technological revolutions appear over the next several decades, they will result in a worldwide information system that will have a major impact on the entire information industry. A projected model illustrates the pricing and retrieval rate changes that may be expected over the next fifty years.

Libraries represent the effort by civilizations to store and retrieve knowledge about themselves and their discoveries. Yet despite the technological progress of this century, libraries remain passive information warehouses requiring the patron's memory and intelligence to find and interpret the world's knowledge. "Electronic" libraries are little more than data on paper converted to equally static data on screens. The mechanical button-pushing of online searching has been automated by personal computer software such as ProSearch and dial-up front-end systems like EasyNet, but the intellectual aspects of searching remain manual. Limited expert and natural language systems for information retrieval exist only in the laboratory or for very limited subject areas.[1] Getting information today from large databases requires skill, experience, and a little bit of luck.

Automated large-scale retrieval must be practical and worth its cost in order to be accepted. It requires a method for encoding massive amounts of information with little delay, affordable memory, and rapid, inexpensive access. To date, the closest thing the information industry has to such a "universal" system is the personal computer coupled to CD-ROM and dial-access sources. Because of incompleteness or time delays between the publishing of information and its availability in a database, many who need the latest accurate information turn to the "invisible college" (acquaintenances, preprints, and conferences). But changes within the next half century in data conversion processes and storage, data transmission, and searchware will bring about a revolution that will make global information retrieval practical.

DATA CONVERSION AND STORAGE

Anything remotely resembling a universal retrieval system must have a universal knowledge base as its foundation. Creating large-scale knowledge bases will be a massive undertaking. If an automatic database-building process is not developed, the database must be created manually, as several groups are attempting to do.[2]

These databases will require extremely dense storage media. The $6,000 1-gigabyte erasible optical disk drive available today could not meet future storage needs economically.[3] Both optical storage operating at much higher light frequencies

Douglas A. Kranch is Head, Instructional Services Department, General Libraries, Emory University.

and "nanotechnology," where information storage occurs on the molecular level, promise increased density and decreased storage costs. There is great market pressure for cheap, dense, reliable storage, assuring significant breakthroughs in the decades to come. I have projected the cost for a gigabyte of storage to drop from the current $6,000 to a mere $100 (1989 dollars) within fifty years. Considering storage costs trends over the past forty years, this estimate is conservative.

DATA TRANSMISSION

At current telecommunications costs, CD-ROM databases and in-house online systems such as BRS' OnSite can be attractive alternatives to remote services, but they have inherent limitations.[4] Today's market can support both in-house and remote systems because each has unique economic or access advantages.[5] Over the next several decades, optical fiber technology and competition among long-distance companies will reduce long-haul rates on high-density routes. The private or leased long-distance data networks that many online information services now use will prove impractical when faced with cost increases of 25 to 50 percent and more. The sheer economy of scale will make commercial long-lines the choice for future information systems. The choice becomes even more certain with the development of packet switching and all-digital communications systems with an estimated capacity several thousand times greater than the fastest current system. These Integrated Services Digital Networks (ISDN) promise the ability to send image, sound, and other data simultaneously over the same lines with fast, clear, error-free transmission and without using separate networks or modems.[6]

SEARCHWARE

Today's online catalogs depend on the user for accuracy. Boolean searching is usually provided along with some form of truncation and delimiting, but the burden of tactical assault on the online information fortress remains on the user. The complication of multiple-database access will require that the burden of search formulation be shifted to the system. Artificial intelligence (AI) and hypermedia links among machine-readable sources offer hope for the user. There are predictions that library catalogs may become the most massive AI applications in existence within ten years.[7]

New technology is leading toward the development of effortless links between local and external databases. Tying personal computers together into networks is becoming the norm. One estimate sees 58 percent of personal computers connected to a Local Area Network (LAN) by 1992.[8] New chip designs incorporating the Extended Industry Standard Architecture (EISA) and Micro Channel Architecture allow multiple processing at very high speeds. One computer can act as a communications server for the entire network, freeing up individual processors to concentrate on other computing tasks while the server processes and sorts database queries.[9] Connect time to databases will shorten as transmission rates increase, rendering inconsequential the higher switched access and local measured service costs resulting from the Modified Final Judgment (MFJ) breaking up the Bell System.[10] These LANs will become the power users of online information.

Developments such as the 486 processor and the OS/2 environment will enable networks of microcomputers to handle AI applications using "groupware."[11] The processing power available to the individual will be greatly multiplied by a network link. Networks can become "idea processors," automatically routing information requests or results (along with voice, audio, animation, and other supporting material) to the LANs most likely to have or need the answers.[12]

As computer chip prices fall, processing speeds increase, and knowledge and linguistic barriers are overcome, parallel processing will allow much larger bodies of knowledge to be handled by computer programs. It will be commonplace for systems to interact with users in human language. The areas of expertise will be broadened, the systems made "wiser" and less susceptible to breakdown. Networks will expand their databases by inputting and sorting information automatically. Screen presentation will be refined to help users clearly visualize information links. An AI interface between the user and the network will turn the personal computer into an active source of information, providing knowledge to fit specific situations.

The entire globe could then be considered a single information system with no single location to be pointed to as its home. Instead, there will be a "memex effect"[13] resulting from the seamless interfacing of innumerable databases by AI gateways communicating across digital networks. Some foresee the development of certain cities as "knowledge factories" operating with the new technologies.[14] The economics of information handling will steer the development of the worldwide network toward global nodes containing multiple billions of gigabytes of stored textual, aural, and visual information. To avoid expensive redundancy, the nodes will tend to have uniquely specialized information. These nodes will be interconnected to each other and to (possibly) two mega-databases in Japan and Europe. (See figure 1.)

Publishing will be drastically affected. Some items, perhaps textbook-type material, will remain in print, but journals reporting factual information will probably become fully electronic. The amount of "scholarly" electronic publishing will increase as the number of printed journals declines. College and university accreditation bodies must then formally acknowledge that information access is as important to academic standards as library book holdings. The academic system must also sanction the weighting of formally reviewed electronic publishing as equal to similar pa-

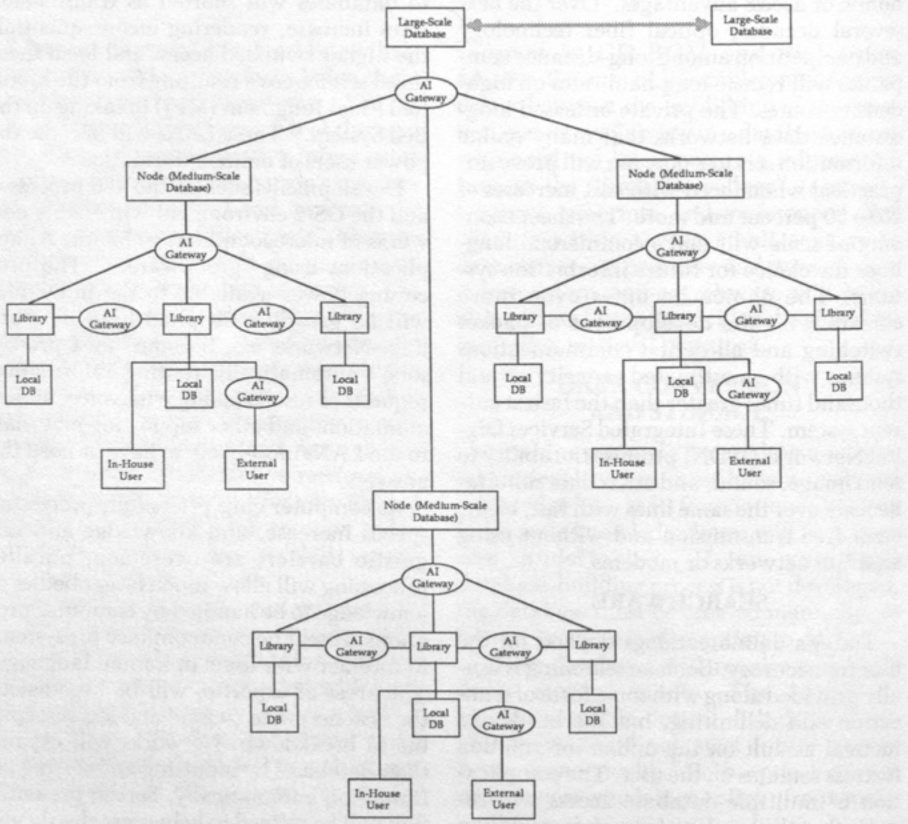


Figure 1. A Universal Retrieval Network Producing the "Memex" Effect.

per publishing for tenure review. When this happens, many authors will bypass text in favor of the immediacy of electronic knowledge systems. Maintaining the electronic knowledge base will gradually become part of the publishing industry.

THE LIBRARY AS INFORMATION ACCESS CENTER

During the next half-century, growing demand for information will pressure libraries to depend more on networking with shared databases across distributed systems. Local information files will be electronically linked to regional and national databases. High-demand data access by optical disc will be replaced by access through networks. Libraries will purchase major database access from vendors the same way they now purchase cataloging, indexing, or online search services.

Figure 1 is an example of how AI interfaces could link the components of the global information system together. A library patron, either in the library or at home, could query the library's searchware. Information relating to the query could be retrieved from the local library files, from regional database files, or from files half a world away. The AI interfaces would pass the query on from network to network and database to database. The only indication of any information's source would be a location code supplied by the searchware and any costs for access. The system could indicate experts who may have the needed information and leave query statements in their electronic mailboxes.

Information professionals, previously immune to automation-induced job loss, will discover that the advanced information searchware will have an impact on them, increasing the need for those with a facility for AI development and use and decreasing the need for information retrievers. Arbib notes how the importance of different mental skills changes with the availability of technology.

> The ability to write down our thoughts has long since relieved us of having to acquire the skills of rote memory that were common in preliterate societies. With computers we will develop yet different memory skills to learn about retrieval, and about how best to organize information within our computer network so that we can find what we want.[15]

As much as 80 percent of patron information needs will likely be satisfied by the intelligent interface. The remaining 20 percent, the difficult, unusual, hard-to-pin-down questions, will be referred to human intermediaries. But even then, the nature of the questions will not be so much "Where can I find?" as "How do I go about searching for?," for the worldwide information network will be one vast reference work with searchware as its index. Knowing in which book relevant information lies will no longer be the chief task of the reference librarian. Rather, librarians must become competent at formulating searches and be familiar with the quality of the thousands of databases to which access might be gained. Likewise librarians now creating the catalog of items in the library will then be supervising the adding of selected information to the local database, editing electronically developed indexing.

As artificial intelligence systems are developed and marketed, the library will gradually become a "knowledge server," able to retrieve and exhibit information electronically, summarize it, pursue relationships, and, by drawing analogies, produce its own new concepts. The library "user" may be another knowledge system, a library, or a person. The future library will be a network of collaborating people and machines, a group intelligence supplying individual information needs.

There is legitimate concern regarding electronic information access for the economically disadvantaged. Historically, information gaps have been filled by the public libraries. The public libraries will once more find their historic roles needed but with a technological twist. The motto of public libraries in the early twenty-first century could well be "Mass information for the masses." Computer literacy may become a priority equal to reading literacy. Funding for free or reduced-cost searching by patrons of the major databases will be required to equalize information access.

CONCLUSION

The distributed memex will arise when storage costs, communication speed, and search capability produce the revolutionary environment required for its birth. First, the information on which the system will feed must be encoded quickly, continuously, and economically. Second, an information network capable of high-speed data transmission in all formats must be set up worldwide. Third, the software interface to make searching such a global information system possible must be developed. It is in especially the third area that work must be done. Strategies are needed by which expert system interfaces can give the required 90%+ search precision. Local information databases must be produced and linked for searching. For hypertext and hypermedia to become the tools for building intelligent access, standards for applying links and indexing must be established. In other words, information specialists must develop the standardized information environment within which the dream of the universal retrieval system will come to pass.

Even larger issues will be raised by the development of global information processing. Some warn that the economic basis of the knowledge industry might extend the gap between the have and the have-not nations to include information.[16] However, it is also possible that the "memex effect" could make technologically developed nations dependent on foreign sources for major portions of their information needs. Information embargoes by international monopolies could be possible. Information workers might protest the cheap data processing occurring in developing nations. Cheap and easy access to information services may transform information delivery into a highly competitive service industry where knowledge of the location of pieces of information will be an irrelevant memory game. Information professionals must plan to adapt their skills, their profession, and the training of future professionals to the revolutionary economic and political changes these new technologies will produce.

REFERENCES AND NOTES

1. Donald T. Hawkins, "Applications of Artificial Intelligence (AI) and Expert Systems for Online Searching," *Online* 12, no.1:31–43 (Jan. 1988).
2. For example, the Japanese Fifth Generation Project has spun off the Japan Electronic Dictionary Research Institute, which is attempting to encode massive amounts of verbal information. In the United States, BRS is constructing full-text databases for end-users, largely in the medical field.
3. The 70,000 new titles reported in the United States alone would require more than 40 such disks costing together over $240,000, assuming 3 kilobytes per page and 200 pages per average title.
4. Large databases require multiple laserdisks with their attendant swapping, and networking them to other stations can be a complex technical and legal task. Technical and financial burdens limit site-based online systems to sites with the equipment and infrastructure to support them.
5. Mick O'Leary, "Dialog's New Era," *Online* 12, no.6:15–21 (Nov. 1988).
6. Larry L. Learn, "Networks: The Telecommunications Infrastructure and Impacts of Change," *Library Hi Tech*, 6, no.1 (consecutive issue no.21):13–31 (1988).
7. John Drabenstott, ed., "Beyond the Online Catalog: Great Potential and Profound Change," *Library Hi Tech* 6, no.1 (consecutive issue 21):101–111 (1988).
8. Christine Strehlo, "The Promised Lan," *Personal Computing* 13, no.1:93–98 (Jan. 1989).
9. Russ Lockwood, "Hot Technologies to Watch in 1989: Five Advances That Extend Computing Performance," *Personal Computing* 13, no.1:101–06 (Jan. 1989).
10. Learn, *Library Hi Tech*, no.21:13–31.
11. Strehlo, *Personal Computing*.
12. Lockwood, *Personal Computing*.
13. This term is based on the name coined by Vannevar Bush for his black box mechanical memory aid designed to hold "all of [a person's] books, records, and communications." See, "As We May Think," reprinted in Bush, Vannevar, *Endless Horizons* (Washington, D.C.: Public Affairs Press, 1946), p.32.
14. Edward Feigenbaum, et al., *The Rise of the Expert Company: How Visionary Companies Are Using Artificial Intelligence to Achieve Higher Productivity and Profits* (New York: Times Books, 1988), p.256.
15. Michael A. Arbib, *On Being Human in the*

Computer Age. In Trappl, Robert, ed., *Impacts of Artificial Intelligence* (North-Holland—Amsterdam: Elsevier, 1985), p.52–63.
16. Feigenbaum, et al., p.267.
17. Elizabeth Betz Parker, "The Library of Congress Non-Print Optical Disk Pilot Program," *Information Technology and Libraries* 4, 289–99 (1985).
18. If international standards for databases and their searchware are developed allowing them to link together without interfaces, these specialized subject databases could become modular components of custom-designed home and library information systems.
19. Du Pont, for example, has discovered that expert systems can do 80 percent or more of the decision work of experts. It may be similarly expected that expert systems could locate 80 percent or more of the information required by an interrogator. Fiegenbaum, p.151.

Global Information Retrieval: A Fifty-Year Projection

With the help of computer modeling, the economic forces affecting the coming information changes can be projected. The model with the pricing projections shown in table 1 assumes that costs for off-site retrieval will be adjusted to remain competitive with on-site database costs. It also assumes that the technological changes described in the article will all be in place about the year 2015.

In the years before 2015, the technical obstacles of system standards and compatibility should be overcome. As inter-system cooperation becomes technically feasible, several specialized mega-databases will be developed and marketed, each with its unique software interface. However, knowledge bases approaching the "universal" level will be so expensive that only government-sponsored long-term projects could be successful. Japan, already at work on such a database, will likely be the home of one; Europe, pooling the economic might of the several states, the home of a second. It is doubtful whether any other technologically capable countries, including the United States, will feel the need to commit resources to similar "universal" databases as long as information can be purchased elsewhere at reasonable cost.

Developing alongside the knowledge bases and the telecommunications networks will be the AI interfaces to manage the information flow. In the early stages of development, users will pay a fee based on the use of AI in the search. For complex searches, this fee could be quite substantial. Because of this, these interfaces will be found almost exclusively on databases where retrieval time and comprehensiveness outweigh the additional expense. As AI becomes more common, more generally applicable, and more precise, added fees will decrease markedly.

The copyright problems inherent in a distributed system will have to be resolved, problems which can become very complex, as the Library of Congress disk project demonstrated.[17] Charging royalties for downloading is one answer.

The database vendors will be expanding by using leased lines and building their own communications systems. Rising telecommunications rates, value-added search costs, online royalties, and the introduction of increasingly dense storage will make local systems more cost-effective than remote searching. These on-site systems will be relative small and slow, with search precision on the order of 50 percent. The information industry will be comparing the trade-offs between maintaining local databases and searching remote databases. Successfully marketed full-text remote databases will tend to be current reference material useful to professionals.

About the year 2015 (in my model), the three revolutionary technological changes will be in place, radically changing the global information picture. Affordable AI interfaces will be in place that increase search precision to beyond 90 percent. Information vendors will switch to Integrated Services Digital Networks (ISDN), forming a worldwide transportation system for the information industry. This network will multiply input/output and search speed, allowing the implementation

Table 1. Pricing Projections

	1990	1995	2000	2020	2040
Remote Search Costs					
Average Search Charge per Value-Added Search Parameter	$ 8.00	$ 5.00	$ 6.50	$ 3.00	$ 2.25
Average Value-Added Parameters per Search	1	3	4	10	20
Average Royalty per Downloaded Item	$ 15.00	$ 15.00	$ 7.50	$ 6.00	$ 3.00
Average Number of Downloaded Items per Search	1	2	3	8	12
Telecommunications Charges per Hour	$ 12.00	$ 15.00	$20.00	$25.00	$30.00
Average Number of Minutes per Search	10	7	5	2	1
Local Database Costs					
Cost of 1 Megabyte of Storage	$ 6.00	$ 2.00	$ 1.00	$ 0.33	$ 0.10
Average Number of Items Added per Year for 50,000 Item Database	1.5	2	100	725	1,750
Average Number of Items Added per Year for 25,000 Item Database	1.5	2	50	338	900
Average Number of Items Added per Year for 5,000 Item Database	1.5	2	10	70	178
Average Cost of Each Item Added	$300.00	$200.00	$50.00	$10.00	$ 5.00

Note: All costs are given in 1989 dollars.

of the AI interfaces to lower drastically telecommunications costs per retrieved item. The networks will already have the advantages of well-developed databases of current material, powerful search "groupware," and links to other systems. Market forces will adjust the search and copy royalties to make the overall cost of remote searching competitive with on-site systems.

The competing databases will expand between 2015 and 2025 while major government-funded suppliers of textual information come online. A decade of merging will follow, as well as the distribution of pre-packaged databases on specific subjects for home and library.[18] Royalties will continue to drop to keep remote databases competitive. Information users will discover that, for many information needs, buying information, even with the added costs of "smart searching," pays off in retrieval costs that are lower than the cost of on-site storage.[19] Computer storage costs should have dropped enough to make onsite databases economically feasible for information of purely local interest. Single-user CD-ROM stations will disappear, major worldwide information nodes will be developed, and the birth of the universal retrieval system will be upon us.

The graphs compare local acquisition and storage of electronic information with online searching and downloading of items from an external commercial database. Costs are compared for three sizes of microcomputer-based local databases: 5,000 items, 25,000 items, and 50,000 items. Annual acquisition rates are adjusted so that the final database size is reached in the year 2040. For simplicity, no weeding is assumed to take place.

Figure 1, "Cost (1989 Dollars): Local Acquisition vs. Off-Site Retrieval," compares the cost of annual purchases and storage of local items with costs for searching and downloading the same number of items from a commercial source. (For the purposes of this projection, an "item" is taken to be roughly one bibliographic unit, such as a monograph or one issue of a serial. Current remote retrieval of an "item" was calculated using six downloaded articles requiring twelve minutes search and download time from a database with a $75 per hour connect charge.) The initially

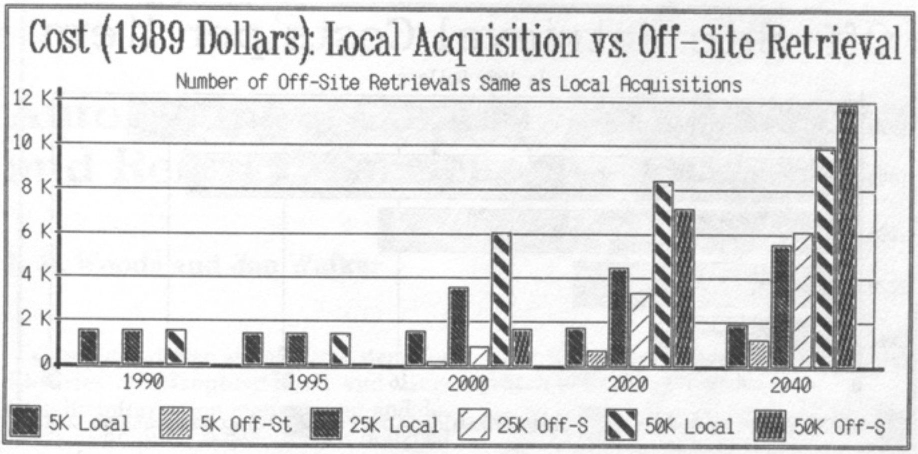


Figure 1. Local Acquisition vs. Off-Site Retrieval.

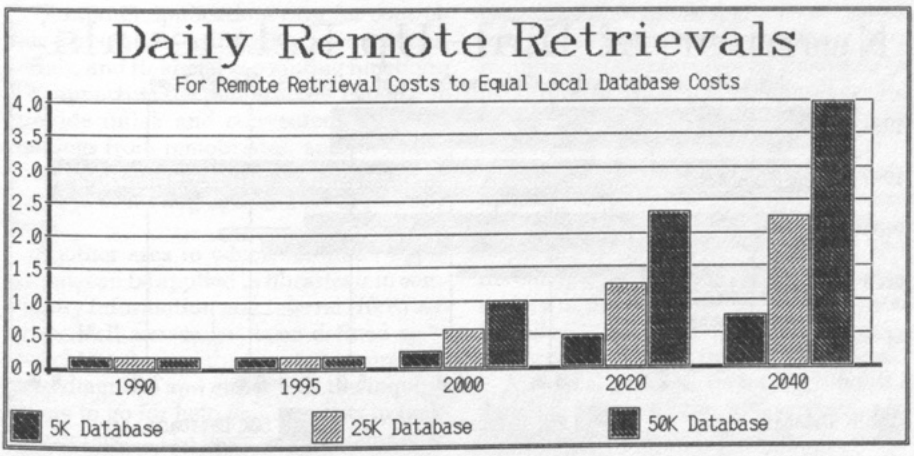


Figure 2. Daily Remote Retrievals.

high local acquisition cost compared with outside retrieval costs reflects the increased purchasing of local database items which have dropped in price from current levels but are still not cheap. Subsequent years show a rough parity between local and remote searching costs due to continually decreasing local item cost and an intentional attempt to match off-site retrieval costs to local database costs.

Figure 2, "Daily Remote Retrievals," shows the average number of commercial downloads that could be done each day for the same money spent on acquiring and maintaining the local databases. It is an indication of the minimum number of useful items that would have to be retrieved from the local database to make its maintenance cost-effective. Its gradual increase after the year 2000 reflects the assumption that AI interfaces and lower search and downloading costs will make commercial database searching more competitive.

Figure 3, "Off-Site Retrieval Cost Per Item," projects the costs for each item retrieved from a remote database, showing the portions that comprise telecommunications charges, search charges, and connect

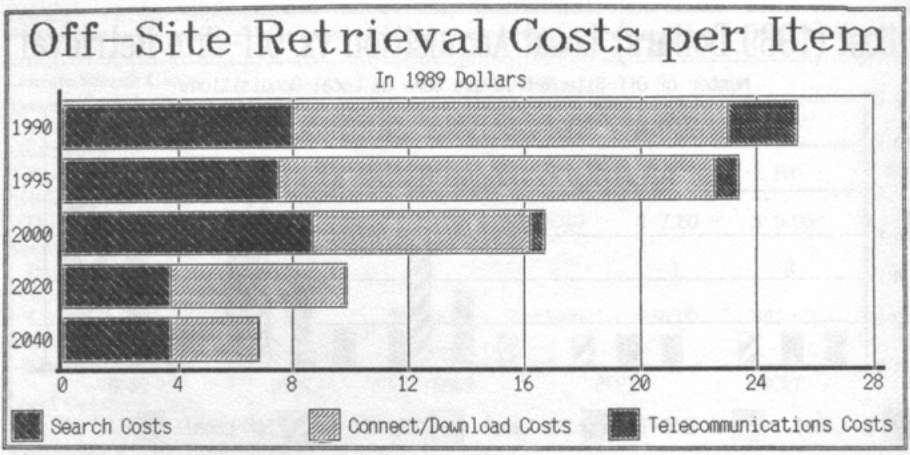


Figure 3. Off-Site Retrieval Costs per Item.

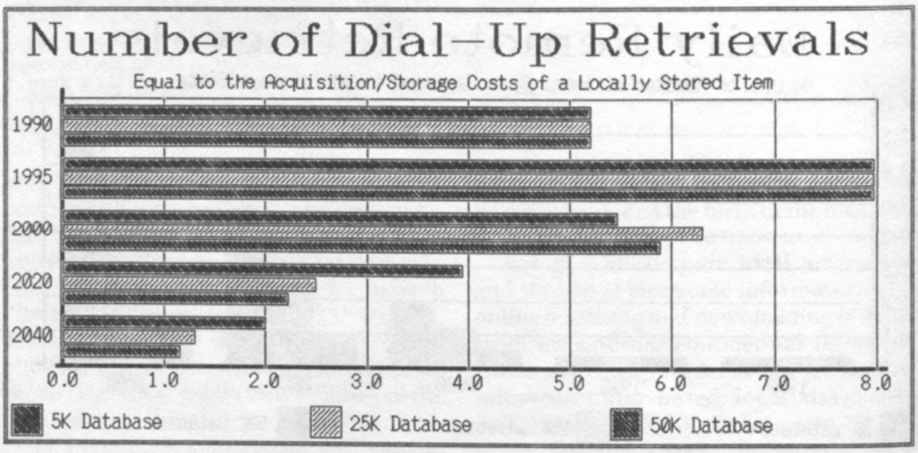


Figure 4. Number of Dial-Up Retrievals.

and downloading costs. Notice the increasingly small part telecommunications costs will play. This reduction is a direct result of AI's allowing faster, more precise searching of the databases. Search costs will first increase, then decrease to a value lower than today's costs, as the AI searchware moves out of the development stage and into the general production stage. Connect and downloading costs (which includes royalties) will be drastically lowered after the year 2000 in an attempt to remain competitive in the information industry.

Figure 4, "Number of Dial-Up Retrievals," shows how many remote retrievals each locally added item would buy. The closer to 1.0 the value is, the more equivalent the costs become. In this model, the number decreases greatly over the fifty years finally to approach 1.1 (for the 50K database), indicating that owning an item would be just a little more expensive than retrieving it from a remote source when needed.

As a model, specific years and numbers are speculation, but the overall trend seems clear. The result will be an information environment that can support mass information storage and retrieval on demand from either local or remote databases. ■■

Automation of Community Information and Referral Services

L. B. Woods and Jon Walker

Computer-driven automated systems for libraries offer sophisticated and user-friendly information management and delivery techniques. When properly designed and operated, automated systems store and retrieve information more efficiently and effectively than do manual systems.

Computer applications exist for controlling circulation, cataloging, acquisitions, serials, and financial accounting functions; for supporting the public access catalog; to provide quick and convenient access to holdings from remote sites; and for online access to indices, abstracts, and full-text displays of monograph and serial publications.

Another area to which computer applications can be applied in libraries is in community information and referral (I&R) services. I&R service has been defined as "a switching station that performs preliminary diagnosis and either tells the inquirer where to go for help or . . . assists in making contact."[1] The basic requirements of an I&R service include the organization of a file of information about community organizations, services, and programs; dissemination of this information; and continual updating of the file. Generally speaking, computers can simplify all of these functions.[2]

This paper will address the applications of computer technology to I&R services, basic planning considerations for automating an I&R service, and the development of an effective automated I&R system; give examples of computer-supported I&R services that are in operation; and describe the Tulsa City-County Library System's approach to I&R automation.

PLANNING THE AUTOMATED I&R

The collection of local community information is often difficult. Investigating sources, collecting and verifying data, organizing information, maintaining its currency, and disseminating the information requires skill and involves an enormous expenditure of time and energy. An automated system can improve the ability of I&R staff to do these functions.

If the I&R resource file is online, it can be updated readily, and both subject access and cross-referencing are simplified. This is important because information in an I&R file changes constantly. Unlike a printed directory or manual card file, an automated file allows staff to make changes quickly and easily, as soon as they are discovered.

A typical method for compiling I&R data is to fill out a form with a community organization's name, location, telephone number, hours of operation, purpose, details about services and fees, and appropriate subject access terms.[3] Each element of information is assigned to appropriate fields in the computer file that, in turn, comprise a record. Each record in the master file is analogous to the "agency card" of traditional manual files. A CRT is the standard tool for inputting information into the computer. The computer software is programmed to allow information in the master file to be retrieved by subject, organization name, and other appropriate fields. In more advanced applications, the computer

L. B. Woods is a library consultant based in Tyler, Texas; Jon Walker is Head of the Department of Community Information and Referral Services, Tulsa City–County Library, Tulsa, Oklahoma.

software permits searching by combinations of fields, i.e., by Boolean logic.[4]

Despite the advantages of automation, conversion from a manual file can be costly. It requires careful planning in order to devise and implement properly an automated system that improves I&R services. The first step is to conduct a needs assessment. The first component of this assessment is the determination of goals and objectives of I&R services. The second describes step-by-step the details of existing procedures. This requires a close examination of current tasks and a determination of which should be eliminated or continued in the automated environment.[5]

The third step involves weighing cost-benefit relationships. The approximate costs of both the hardware and the software are determined. Certain benefits are also easily quantifiable, e.g., potential reduction in personnel and support services and faster response times to patron requests for information. Other benefits are more difficult to quantify, e.g., effects on staff morale, quality of service, improved accuracy, and currency of information.[6]

The fourth step is to develop a Request for Proposal (RFP). The RFP is a statement to software and hardware vendors requesting a formal bid on the cost of providing a computerized system that meets the I&R service's stated software and hardware requirements.[7]

In planning for automation, it should be remembered that the computer is a tool, and albeit a sophisticated one, the computer exists to serve the goals and objectives of the organization.

AUTOMATING TO IMPROVE I&R SERVICES

An important goal of I&R automation is to acquire, process, and display information for the benefit of I&R staff members who provide information to the public and who make referrals. Another goal is to be able to assess the nature of information problems and to generate appropriate referral. Understanding how to make a referral is not the same thing as looking up information in a file, although there are many cases for which this may suffice. Making a referral is a complex process because the public is not generally well versed in the terminology used by I&R staff. They may not be able to follow through on referrals that I&R staff consider elementary.

A patron will often describe only a portion of his or her true information need. As a result, the I&R staff may be hard-pressed to determine quickly and accurately an appropriate referral. The difficulties faced by the I&R staff increase if a sufficient quantity of relevant information must be assembled from manual files. Individual staff member's decisions are based on incomplete descriptions, his or her own past experience, feedback from past patrons, and conversations he or she has held with other staff members. The system works but is inefficient compared to a more complete view of potentially available data.[8]

Access to a larger portion of data is within grasp, given two assumptions: (1) the existence of computers capable of handling complex files and (2) a system design that takes advantage of computer capabilities for storing and retrieving data. The development of the second assumption depends on at least three considerations: (1) applying computers less to areas of quick success and high payback and more to areas requiring increasingly sophisticated, difficult-to-describe, and slow payoff service situations; (2) developing better models of the I&R assessment process; and (3) better understanding by computer experts of I&R assessment problems.[9]

Automated systems in I&R services have a significant role to play both now and in the future. Despite some current difficulties, the prospects for increasing the quality of I&R services are too great to ignore.

EXAMPLES OF AUTOMATED I&R SERVICES

A survey of the literature revealed descriptions of fifteen automated I&R sites. Seven of these are in libraries. In general, the others operate as independent agencies receiving substantial financial backing from the United Way. Below is a description of each:

1. The Fresno County Free Library's Information and Referral Network uses word processing equipment to maintain its I&R file. The file is not online; instead, data for

a printed directory is maintained on the computer. This makes it easier to print updates and revisions. The file is stored on disk, updated quarterly or whenever desired, and then printed for use by the I&R staff. This is among the least sophisticated of the fifteen services surveyed.[10]

2. The Pikes Peak Library District's I&R files are among the most sophisticated. The file is in database format and operates on the same computer as the library's circulation, inventory control, and acquisitions functions. The library has four separate community information files: CALL, which lists social service agencies; CLUBS, which lists clubs and social organizations; COURSES, which includes educational and recreational offerings in the community; and CALENDAR, a listing of local cultural events. The four files function as an integrated system, allowing staff to do I&R work for patrons from all files simultaneously.[11]

3. The Chicago Public Library's Native American Information and Referral Center has created an automated file to serve the Native American population in Chicago's Uptown area. While implementing the system, the staff of the Chicago Public Library (CPL) encountered some problems in designing and starting up the machine-readable I&R file. The problems lay with the I&R staff's difficulty in communicating the system's requirements to the project programmer. The result was an initial software design that was inadequate for doing the specified job. The CPL did not go online with their service until after scrapping the initial programs and carrying out a more detailed process of system and software design. This included improving communications between library staff and the programmer.

4. The Peninsula Library System's CIP is an online I&R file maintained on the SPIRES database at Stanford University's Center for Information Technology. Using distributed access processing, the online file is searched by CIP staff, librarians at seventeen libraries, and staff of the San Mateo County Social Services Department. It provides information to people working in social service agencies and answers patrons' requests for information about local human services. A printed directory is also produced from the computerized file.[12]

5. The Whitley County Opportunity Center is a social service agency that provides day treatment programs to the developmentally disabled. They have implemented a computerized information system to help clients as they progress through habilitation programs.

The I&R staff has drawn several conclusions from their work at Whitley:

(1) The delivery of human services must be interfaced with the automated system and receive feedback from it but should never be compromised just to accommodate it.
(2) The staff will support and utilize fully the automated system only if they perceive that it facilitates their work. They must feel that they have a vested interest in the system.
(3) The system must produce reports and information that are usable and assist the staff's ability to serve clients. To accomplish this requires greater investment in system design and software development and modification. But when one considers the ongoing impact of the resulting database on an agency's functioning, investment at this stage is very effective over time.
(4) Most people seeking to computerize an I&R service understand either human services or computers but seldom both. Each discipline has its own language and tends to solve problems differently. There is a strong need for the two to work together effectively.

6. The Broward County Community Service Council started a computerized information and referral system that not only helps I&R specialists to make proper referrals but also offers a case management system. It allows human service agencies to know via a distributed-access online system which agencies a particular client has already contacted and his or her current status. In this way, social workers can keep better track of and assist needy individuals.[13]

7. At the Community Chest I&R Center in Cincinnati, Ohio, the online resource file enhances and simplifies the data collection and retrieval that is necessary for operating

an effective I&R service. Human service agencies listed in the file are accessible by name or category of service provided. The system gives users the option of truncating search terms. In addition, it generates printouts of individual agencies and the services they offer, and these printouts can be given to individual clients. The automated system also produces monthly statistical reports that include profiles of the number of referrals made to each agency in the community.[14]

8. As with the CIP system described above, the AID service of Edmonton, Alberta, uses SPIRES. The software runs on a mainframe located at the University of Alberta's computer center. The experience in developing the AID service bears out the conjecture that a successful automation project requires good communication between programmer and I&R staff. Moreover, the Edmonton experience teaches that "technocratic neophytes" are often too optimistic in their projections regarding the time required for developing and implementing an automated system.[15]

9. The VD National Hotline in Palo Alto, California, provides I&R services to callers nationwide. It uses an automated online system that allows file updates to be available immediately to front-line staff. In addition, it simplifies data storage techniques and the creation of mailing lists and enhances the staff's ability to access in-depth information.[16]

10. CRIB in Los Angeles County represents another successful automation experience. The I&R system maintains on computer a comprehensive master file of 5,000 public and private community resources. Information is updated monthly and retrieved, reproduced, and distributed semiannually in a directory called "CRIB Books." CRIB is an example of a file, like the one at Fresno County Free Library (see 1 above), where information is input into a computer but reproduced and accessed in printed form rather than online.[17]

11. A system similar to CRIB is CRIS, the Community Resource Information Services, which is based in Philadelphia. CRIS is used to produce a 5,000-page COM catalog of services that are available in southeastern Pennsylvania.[18]

12. The Service Agency Inventory System is a project located in New York City that lists over 2,000 social service agencies. It is online on an IBM mainframe and is used to print a directory both in hard copy and on microfiche.[19]

13. LINC, a library-based I&R service in Memphis, uses a microcomputer-based software package developed by Storey/Ross/Barker, Inc. LINC was the initial test site for the software package, and the staff at LINC experienced both the excitement and the frustrations of serving in this role. LINC uses the Storey/Ross/Barker software to produce on demand, brief, topical printed directories of community services.[20]

14. ACCESS is a minicomputer-based I&R service developed in 1986 and housed in the Spokane Public Library. Information in ACCESS is organized both alphabetically by organization name and topically by service type. Two broad conclusions were reached while developing this file: (1) automated systems often take longer to develop than anticipated and (2) staff must often spend much more time than anticipated learning about computer technology.[21]

15. TIP is an I&R service sponsored by the Detroit Public Library. This program is generally recognized to be the first such service that was located in a library, having been established there in June 1973. TIP automated their services in 1987 using PC-INFO, a microcomputer-based system. The use of microcomputer technology makes TIP's automated files not only "information efficient" but highly cost-efficient as well.

There are four reasons TIP was automated: (1) information in an automated file is more easily updated than in a manual file, (2) computerized data allows more sophisticated search capabilities than does manually maintained information, (3) specialized directories can be readily printed, and (4) information can be shared electronically with other agencies.[22]

The automated systems summarized in this section range in sophistication from simple word processing files to more complex database systems with online distributed access. The systems also range from the use of large mainframe computers to small personal computers; each provides distinct advantages over manual systems. In the next section we will examine in greater detail the workings of one automated I&R system.

THE TULSA CITY-COUNTY LIBRARY MODEL FOR I&R

In September 1974, as part of a push at the federal and state levels to increase services to older Americans, the Tulsa City-County Library (TCCL) undertook the Senior Citizens Information and Referral Service (SCIS). A card file of the agencies in the Tulsa area that provide services to the elderly was compiled by the staff.[23]

The purpose of SCIS was expanded to provide a total community I&R service—age-integrated, relevant, and accessible to all citizens of Tulsa County. The name was changed to the Community Information Service (CIS), and card files were expanded.[24] Information fell into four basic categories: (1) government and human services, (2) clubs and organizations, (3) speakers and program presenters, and (4) continuing education opportunities.

Until the mid-1980s these card files were maintained manually. The system consisted of listings of essential service information on a main entry card, and the use of added-entry and subject-entry cards for cross-referencing.[25]

In late 1983 TCCL expanded its automated bibliographic capabilities by implementing an LC-MARC-based, IBM mainframe driven software package called NOTIS (Northwestern's Online Technical Information System). Soon thereafter, CIS files were automated by inputting them into the existing NOTIS database.

Information about community organizations was formatted for convenience of use and in order to take advantage of the organizational strengths of the MARC database. In other words, TCCL supplied a bibliographic processing system (NOTIS) to a specific file of nonbibliographic items. Moreover, in order to complement the library's online catalog, LC Subject Headings were used to provide the online I&R files with subject access. An abbreviated set of the available MARC tags was selected:

040: Used to show if a name authority record is established for the agency or organization. AACR2 and OCLC are used as aids in establishing names of entries in the fields. This helps to maintain consistency of format between the library's book catalog and the I&R file.

245: Contains the official name of the organization. This allows the user to search for organizations by name, using standard NOTIS protocol for title searching.

260: Denotes mailing or street address of the organization.

265: This tag is suppressed in public access displays in NOTIS. It is used to note information intended only for staff use.

500: Any number of notes can be used to show office hours; director, officers, or other contacts within the organization; purpose of the organization; services offered; and speakers or audiovisual materials available to the public. Such notes would also show when the organization was last contacted for updated information and whether a vertical file on that organization is maintained in the library.[26]

650: Contains Library of Congress Subject Headings (LCSH) assigned to the organization. This allows users to search for type(s) of service(s) using standard NOTIS procedures for subject searching.

LCSH is not often used to maintain the taxonomy of an I&R file.[27] In Tulsa, it enables staff and patrons who are accustomed to searching the library's book catalog to switch to the I&R files without having to rethink search terminology. LCSH can be unwieldy, but it provides users with the widest access to the total information resources available in the library.[28]

740: Includes acronyms, initialisms, and alternative names of the organization.[29]

At present there are nearly 4,000 listings in the CIS online file. A staff of 5.25 FTEs maintain the currency of the information and use the online file to respond to an average of 1,200 reference queries per month.

The strength of this system is that information in the bibliographic and nonbibliographic (I&R) catalogs is consistent in format and accessible via the same terminals. Clients using the online catalog to locate materials on a particular topic can readily access relevant community organizations.

The integration of information is a significant step forward in providing more complete information services.

CONCLUSION

As automated systems continue to evolve and improve, it is reasonable to assume that computers will be used for new and innovative tasks. The ability of libraries to make information accessible will be magnified. Traditionally, libraries have stored information in book and magazine format; more recently, on microform, audio- and videocassettes, and phono- and laser-discs. In order to provide convenient access to the information stored in these mediums, libraries have developed sophisticated classification and cataloging schemes that are machine readable. With the application to I&R of sophisticated file organization techniques such as LC-MARC and LCSH, the prospects for making I&R fully responsive to human information needs have never looked brighter.

REFERENCES AND NOTES

1. Emelie Shroeder, "Online I&R in Your Library: The State of the Art," *RQ*, no.21:128–35 (Winter 1981).
2. Ibid.
3. Ibid.
4. Ibid.
5. Robert Gunderson, "Is Computerization Right for My Agency?" *Information and Referral*, no.5:1–12 (Summer 1983).
6. Gunderson, p.10.
7. Gunderson, p.11.
8. Paul Licker, "Deciding on Decision Support: Building a Decision Support System for Social Service Professionals," *Information and Referral*, no.5:25–36 (Summer 1983).
9. Licker, p.28–29.
10. John Jewell, "Automating I&R Files: Word Processing, A First Step (Fresno County Information & Referral Network)," *RQ*, no.21:135–40 (Winter 1981).
11. Cynthia Slater, "Automated Information & Referral at Pikes Peak Library," *RQ*, no.21:140–43 (Winter 1981).
12. Jane Irby Light and Carol Yamamoto, "An Online Information and Referral File: Peninsula Library System's Community Information Program," *RQ*, no.21:146–55 (Winter 1981).
13. Kenneth B. Goldsmith, "Computerization of Information and Referral: The Broward County Experience," *Information and Referral*, no.5:49–54 (Summer 1983).
14. Evelyn Laux, "Community Chest Computerized Resource File," *Information and Referral*, no.5:55–60 (Summer 1983).
15. Barbara Ann Olsen, "A Decade of Information Systems," *Information and Referral*, no.5:61–66 (Summer 1983).
16. Bea Mandel, "The VD National Hotline: 'High Tech' Information and Referral," *Information and Referral*, no.5:67–76 (Summer 1983).
17. Shroeder, p.129.
18. Shroeder, p.129.
19. Shroeder, p.129.
20. Linda N. Ross, "PC INFO: The Information and Referral Software for Microcomputers," *RSR*, no.15:21–34 (Winter 1987).
21. Jim Wavada, "ACCESS: A Community Information Service," *RSR*, no.15:21–27 (Fall 1987).
22. Ross, p.33–34.
23. Phyllis Jobe, *A Journal of an Information and Referral Service: Tulsa City-County Library* (Tulsa City-County Library, 1977).
24. Ibid.
25. Ibid.
26. Jon Walker, Letter to Deidre Norton, Chairman of the Board, Galveston Public Library, 25 July 1988.
27. Georgia Sales, "Developing a Human Services Taxonomy: A Case Study," *RSR*, no.15:35–44 (Winter 1987).
28. Walker, p.1–2.
29. Walker, p.2.

BIBLIOGRAPHY

Cauffman, Joy Garrison. "SEARCH: A Link to Services." *Information and Referral*, no.3:89–105 (Summer 1981).

Chapman, Ronald W. "Success Criteria for Constructing an Automated I&R Agency." *Information and Referral*, no.5:13–22 (Summer 1983).

D'Urso, Lawrence A. "The Application of Microcomputers to New I&R Files: A Beginner's Experience." *RQ*, no.21:143–46 (Winter 1981).

Koppel, Ted, and Ward Shaw. "What's That Doing Here? Non-Bibliographic Data in a Bibliographic Environment." *RSR*, no.15:15–19 (Winter 1987).

Lekis, Lisa. "Using I and R Data for Planning: The Process, Problems, and Potentials." *Information and Referral*, no.2:15–29 (Fall 1980).

Mershon, Loretta K. "A Model Automated Resource File for an Information and Referral Center." *Special Libraries*, no.71:335–44 (Aug. 1980).

Shively, James A., and Gary Mitchell. "Program Accountability Made Practical Through an Information Management System: A Case Study." *Information and Referral*, no.5:37–48 (Summer 1983).

Statement of Ownership and Management

Information Technology and Libraries is published quarterly by the American Library Association, 50 E. Huron St., Chicago, IL 60611-2795. Annual subscription price, $17.50. American Library Association, owner; William Gray Potter, editor. Second-class postage paid at Chicago, Illinois. Printed in U.S.A. As a nonprofit organization authorized to mail at special rates (Section 423.12, *Domestic Mail Manual*), the purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes have not changed during the preceding twelve months.

Extent and Nature of Circulation

("Average" figures denote the average number of copies printed each issue during the preceding twelve months; "Actual" figures denote actual number of copies of single issue published nearest to filing date—the June 1989 issue.) *Total number of copies printed:* Average, 7,455; Actual, 7,528. *Paid and/or requested circulation:* not applicable (i.e., no sales through dealers and carriers, street vendors and counter sales). *Mail subscriptions:* Average, 6,794; Actual, 6,934. *Total paid and/or requested circulation:* Average, 6,794; Actual, 6,934. *Free distribution by mail, carrier or other means, samples, complimentary, and other free copies:* Average, 174; Actual, 200. *Total distribution:* Average, 6,968; Actual, 7,134. *Copies not distributed: office use, left over, unaccounted, spoiled after printing:* Average, 487; Actual, 394. *Returns from news agents:* not applicable. *Total (sum previous three entries):* Average, 7,455; Actual, 7,528.

Statement of Ownership, Management and Circulation (PS form 3526, Feb. 1989) for 1989 filed with the United States Post Office Postmaster in Chicago, September 29, 1989.

The Use of SQL and Second Generation Database Management Systems for Data Processing and Information Retrieval in Libraries

William Leigh and Noemi Paz

SQL, or "Structured Query Language," is the result of an American National Standards Institute effort to standardize the language used for querying computer databases. This standard is the common element in a host of new commercial database management system products, a second generation of database management systems. The great scale, homogeneity, competitiveness, and resulting synergy of this market is fostering significant advances in the capability of smaller computers to manage large masses of data. Libraries should take advantage of these advances. This article discusses SQL, its implementations, associated products, and techniques for its use in online catalogs, circulation systems, and other library applications.

The Structured Query Language (SQL) is a database querying language codified as a standard by the American National Standards Institute.[1,2] The SQL standard has supplied a long-awaited basis for interconnection of database products and applications. This redefinition of the market for database management products has prompted the entry into the marketplace of a large set of second generation products that embrace the new standard and introduce other significant innovations as well.

The group that developed SQL is composed of database users and vendors who are primarily concerned with business applications of database management systems. Many of the computer applications in the library, such as circulation control, are of a transaction-processing nature that may be supported readily by SQL and the application development features of SQL-based, second generation database management systems. These application development features include tools for user programming of reports and data entry forms. These tools use SQL internally to communicate with the database processing engine.

This use of standard database management system components will reduce the cost of library information systems. The use of SQL could facilitate this conversion to standard components.

However, SQL is not oriented directly to the problems of information retrieval in libraries or in the bibliographic database industry. SQL is too difficult for library patrons to use. Online catalog and other library information retrieval applications require the services of a professional programmer to develop a suitable interface. This interface can then communicate with the database internally using SQL, reliev-

William Leigh is a professor in the Department of Management, College of Business Administration, and Noemi Paz is in the Industrial Engineering Department, University of Central Florida, Orlando.

ing the programmer of the time-consuming tasks of developing low-level database access routines.[3] As the power of SQL and second generation database management systems become better known, it can be expected that such library-oriented interfaces to SQL will appear commercially.

Equally as significant as the language features of SQL are the architectural improvements occasioned and inspired by the SQL revolution. All of the new, second generation database management system implementations employ a client/server architecture on local area networks. The database accessing component, or database engine, executes on dedicated, powerful, data server computers. Many other client computers can be members of the network and direct their database requests to the server machines. The hardware and software standardization present in modern networks removes the problems of protocol conversion that complicated and frustrated the development of systems such as this only two or three years ago.

Thus, client machines deal only with user interaction and the generation of SQL statements. The server machines accomplish the actual database manipulation according to the SQL statements submitted by the clients. Only the desired data is returned to the client machines. This arrangement makes possible mainframe-scale, multiple user applications on personal computers. The cost savings inherent in this architecture must compel library and other bibliographic information providers to investigate its use.

THE RELATIONAL MODEL AND SQL

SQL is based on the relational modeling of data. An introduction to this technique is a prerequisite to a comprehension of SQL.

From the standpoint of the user, the relational layout of a database is not far removed from a conventional file model. Data is arranged in two dimensional tables with rows and columns. The significance of the relational model lies in the undergirding logical theory. This theory, which is explained in database textbooks,[4] makes the efficient manipulation of this intuitive and convenient tabular approach possible.

For example, the primary elements of a circulation system include books, checkouts, and patrons. Occurrences of these entities may be organized and represented relationally (see figure 1).

SQL is a language for querying, that is, for retrieving subsets of the database according to search criteria. Selection is the basic operation supported by SQL. For example, the application of selection to the Checkouts table to obtain a list of books due before 10/01/89 would return the table:

Call No.	Card No.	Due Date
256	12	09/19/89
351	45	09/22/89

Books:

Call Number	Title	Author
132	ORGANISING KNOWLEDGE	Rowley
256	DICTIONARIES	Landau
351	INFORMATION RETRIEVAL	van Rijsbergen
811	INFORMATION ANALYSIS	Kent

Checkouts:

Call Number	Card Number	Due Date
256	12	09/19/89
351	45	09/22/89
811	12	10/21/89

Patrons:

Card Number	Name	Address
12	Sally Jones	1256 Elm St.
39	Kim Smith	13 E. Garden Ct.
45	Jack Wagner	42 Rose Trail

Figure 1. Primary Elements of a Circulation System.

The particular SQL statement for accomplishing the selection is

```
SELECT CALL_NUMBER, CARD_NUMBER,
  DUE_DATE
FROM CHECKOUTS
WHERE DUE_DATE < 10/01/89
```

Selection in SQL is powerful enough to involve more than a single table. A list of overdue books would be more useful with the patron name and address appended. This requires a joining of the Checkouts and Patrons tables and is accomplished with the SQL selection statement:

```
SELECT CALL_NUMBER, CARD_NUMBER,
  DUE_DATE, NAME, ADDRESS
FROM CHECKOUTS, PATRONS
  WHERE DUE_DATE < 10/01/89 AND CHECK-
  OUTS.CARD_NUMBER = PATRONS.CARD_
  NUMBER
```

The data resulting from this query:

Call No.	Card No.	Due Date
256	12	09/19/89
351	45	09/22/89

Name	Address
Sally Jones	1256 Elm St.
Jack Wagner	42 Rose Trail

Finally, an overdue listing with patron information as well as book information may be obtained with SQL:

```
SELECT CALL_NUMBER, CARD_NUMBER,
  DUE_DATE, NAME, ADDRESS, TITLE, AU-
  THOR
FROM CHECKOUTS, PATRONS, BOOKS
WHERE DUE_DATE < 10/01/89 AND CHECK-
  OUTS.CARD_NUMBER = PATRONS.
  CARD_NUMBER AND CHECKOUTS.CALL_
  NUMBER = BOOKS.CALL_NUMBER
```

It should be clear that SQL can serve as a powerful report generator for data which resides in a data processing system such as circulation control.

INFORMATION RETRIEVAL IN SQL

Relational data modeling[5] and SQL can be applied directly to information retrieval, but the result is cumbersome. Consider an online catalog system with the data shown in figure 2.

Obtaining a list of holdings that are indexed under the subject headings "Libraries" AND "Thesauri" requires the following query in SQL:

```
SELECT CALL_NUMBER, TITLE
FROM BOOKS
WHERE CALL_NUMBER IN
  SELECT CALL_NUMBER
  FROM BOOKS_SUBJECTS FIRST, BOOKS
  _SUBJECTS SECOND WHERE FIRST.
  CALL_NUMBER = SECOND.CALL_NUMBER
    AND FIRST.SUBJECT = 'Libraries'
    AND SECOND.SUBJECT = 'Thesauri'
```

The awkwardness of SQL as a direct interface for patron catalog searching is self-evident from this example. However, it would be unfair not to point out that the preceding example, though it is typical of information retrieval use, does not show SQL in its best light. For example, the OR version of the query above, that is, obtaining a list of holdings that are indexed under

Books:

Call Number	Title	Author
132	ORGANISING KNOWLEDGE	Rowley
256	DICTIONARIES	Landau
356	INFORMATION RETRIEVAL	van Rijsbergen

Books_Subjects:

Call Number	Subject
132	Libraries
132	Indexing
132	Thesauri
256	Dictionaries
351	Libraries
351	Indexing

Figure 2. Data for an Online Catalog System.

the subject headings "Libraries" OR "Thesauri," can be accomplished with the following, still awkward but not as bad, query in SQL:

```
SELECT CALL_NUMBER, TITLE
FROM BOOKS, BOOKS_SUBJECTS
WHERE BOOKS.CALL_NUMBER = BOOKS
  _SUBJECTS.CALL_NUMBER AND SUBJECT
  IN ('Libraries', 'Thesauri')
```

Free text and partial string search are a possibility with SQL, which you might not expect. Finding all books with "Dictionar" in the title can be accomplished with the query:

```
SELECT CALL_NUMBER,TITLE
FROM BOOKS
WHERE TITLE LIKE '%Dictionar%'
```

This partial string search capability can accomplish keyword searching but not without drawbacks. A column can be added to the BOOKS table, above, to contain "Subjects" with data entries such as "Libraries,Thesauri." Then the problem of obtaining a list of holdings that are indexed under the subject headings "Libraries" AND "Thesauri" requires the following query in SQL:

```
SELECT CALL_NUMBER,TITLE
FROM BOOKS
WHERE SUBJECTS LIKE '%Libraries%'
  AND SUBJECTS LIKE '%Thesauri%'
```

This is a briefer, more easily understood solution to this problem. The difficulty here is that in order to process this version of the query the system will have to search every record in the database individually. There is no way in a conventional database management system to prepare an auxiliary data structure, called an index, for all of the appropriate substrings of a field such as "Subjects" in this example. The use of an auxiliary data structure, which is easily possible for the BOOKS_SUBJECTS table, can speed up query processing by a factor of several thousand in a large database.

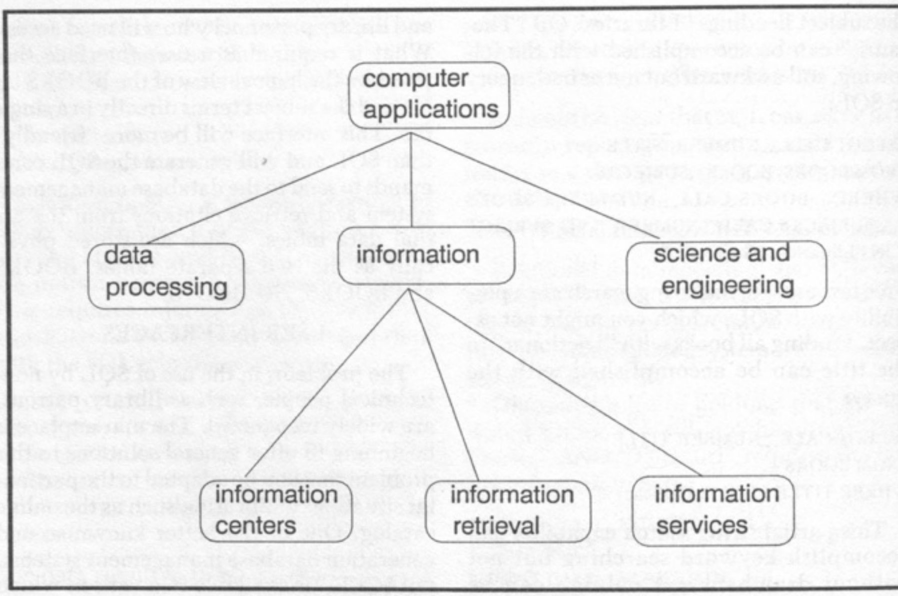
If SQL is to be used for a bibliographic database, then the use of the separate BOOKS_SUBJECTS table is the appropriate physical design. Under this arrangement, querying directly with SQL will be inconvenient for the majority of the patrons and library personnel who will need access. What is required is a user interface that provides the logical view of the BOOKS table and the subject terms directly in a single file. This interface will be more "friendly" than SQL and will generate the SQL commands to send to the database management system and retrieve citations from the actual data tables, which are stored physically as the two separate tables, BOOKS and BOOKS_SUBJECTS.

USER INTERFACES

The problems in the use of SQL by nontechnical people, such as library patrons, are widely recognized. The marketplace is beginning to offer general solutions to this problem that can be adapted to the particular situations of libraries, such as the online catalog. One of the better known second generation database management systems, ORACLE,[6] is available in a version which supports interfaces constructed in Apple Computer Corporation's popular HyperCard.[7,8] HyperCard is known for its capability to support the development of "user friendly" interfaces.

HyperCard can support the programming of a conventional online catalog or information retrieval interface to the database management system. Simplified versions of information retrieval languages have been devised[9,10] and can be used as models for building an online catalog or bibliographic database retrieval system using HyperCard and ORACLE. In this arrangement, existing files of bibliographic information and subject indexes convert into the second generation database management system directly, without recataloging.

HyperCard can also serve as a basis for a graphic-oriented retrieval interface. Examples of such graphical interfaces to an information retrieval system have been developed and explored.[11,12] A simple version presents graphical trees depicting the structure of the information stored in the system. Figure 3 is an image of a display showing a tree index to holdings on computer applications. A catalog user selects a subject with the cursor and presses the ENTER key or mouse button to perform the retrieval. An interface to a SQL-based data-



This is an example of a type of user interface that can be provided to a second generation database management system through the use of high-level, commercial tools such as Apple's HyperCard.

Figure 3. Graphical Trees Can Display the Subject Index in an Online Catalog.

base management system would perform the retrieval by preparing a suitable SQL statement and submitting it. This display and the underlying processing to generate the required SQL can be programmed readily in the HyperCard interface to ORACLE.

Interface-building tools that employ other new information technologies are becoming commercially available for second generation database management systems. Expert systems are a promising possibility.[13,14]

An expert system processes rules such as:

IF USER INDICATES "querying"
  THEN SET TERM "information retrieval"
IF USER INDICATES "data processing"
  THEN SET TERM "computer applications"
IF USER INDICATES "information centers"
  THEN SET TERM "information science"

The thesaurus and controlled vocabulary information of the system are embedded in these rules. The expert system carries on a dialog with the user determining the truth of the left-hand side (to the left of the "THEN") of the rules and carrying out the operation on the right-hand side as indicated. Graphical elements such as the index tree shown previously can accelerate this dialog process by determining the values of several rule left-hand side's in one user selection operation.

GURU[15] is one product that supports the development of expert system interfaces to a second-generation database management system. Nexpert*SQL[16] is a type of software package that allows linking of an expert system to an SQL-based database management system.

Natural language understanding is still a largely unfulfilled promise of artificial intelligence technology. GURU, in addition to supporting expert systems, has capabilities for storing and manipulating rules of a sort that make possible the understanding of simple natural language queries. This has obvious utility for an application such as an online catalog.

More mundanely, these language translation capabilities of GURU can be employed to understand existing, conventional retrieval languages such as that of the Dialog bibliographic database system. This would allow the direct use of second generation database management technology as a replacement for an existing bibliographic information retrieval system.

Obviously, the development of rules, graphical indexes, or natural language understanding capability can require a large investment of labor. At this time, the development of these nonconventional interfaces may be justifiable only in the case of tutorial systems or for small collections used by large numbers of people. The potential of HyperCard for inexpensively grafting a patron-oriented interface on a SQL-based catalog system or for GURU to translate an already in-use querying language to SQL are more within the realm of applicability for general library use.

THE CLIENT/ SERVER ARCHITECTURE

Query language standardization has well-understood benefits, but by itself it is not a compelling reason for libraries to consider converting to SQL-oriented database management systems. Probably the most attractive feature of the new systems is the cost-effective architecture which they employ rather than their use of SQL. (Remember, though, that it is the SQL standardization revolution that has fostered the market climate conducive to the development of so many radically innovative database products; this is termed the "second generation" in this article.)

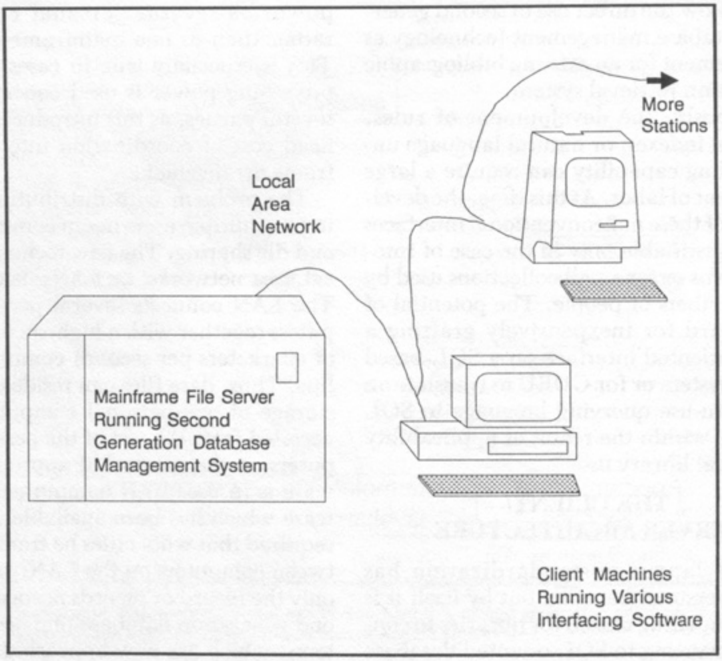
First generation database management systems execute on large, expensive mainframe computers. The central computer needs to be big and fast so that its services can be parcelled out to multiple terminals and their users. The terminals contribute no processing power but only serve for communication and display.

The economics of computers is such that computing power is cheaper in smaller packages. Computing power may be measured in "millions of instructions per second," or mips. A modern personal computer may possess processing power of approximately one mips and cost less than $3,000. A modern mainframe computer may have only five mips of processing power but may cost over $100,000. Thus, there is considerable economic motivation for acquiring multiple mips of processing power as several personal computers, rather than as one mainframe computer. This is especially true in cases where the processing power is used concurrently by several parties, as this introduces the overhead cost of coordination into the mainframe arrangement.

The problem with distributing processing to multiple users lies in communication and file sharing. The new technology of local area networks, or LANs, is a solution. The LAN connects several personal computers together with a high-speed (millions of characters per second) communications line. Thus, data files can reside on the disk storage of one personal computer and be accessed from the rest of the personal computers on the network if appropriate software is in use. LAN communication software which has been available in the past required that whole files be transferred between computers on the LAN, rather than only the record or records needed. The second generation database management systems, which are now appearing, comprise software that allows the transmission over the LAN of only the records of interest.

To attain these economies, the second generation database management systems employ a "client/server architecture."[17] Personal computers with special capabilities such as higher processing speeds and higher speed disk storage can be dedicated to the tasks of storing and retrieving data. These are the servers. Other personal computers, the clients, with color, high resolution displays but with slower processing elements and little or no mass storage capabilities, execute the user interface software that generates requests for data and transmits them over the LAN to the servers. This specialization of computing function exploits the economies inherent in purchasing computer power in smaller units but yields the system performance (for information retrieval and data processing) of mainframe computing systems.

The use of SQL as a common language in this client/server architecture makes it possible to mix the software products of multiple vendors in the same network system. A system builder can choose server software, which is the "engine" performing the actual database manipulation and access, from one vendor and choose client com-



Second-generation database management systems and commercial local area network systems provide protocol support and standardization for lower levels of hardware and software.

Figure 4. SQL Provides Application-Level Compatibility in a Distributed Environment of Heterogeneous Equipment.

puter, interfacing software, such as HyperCard, from another vendor. There is assurance that the components will work together if they all use SQL.

Some of the second generation database products run on a large range of different computing equipment. ORACLE, as an example, can be used in mainframe, minicomputer, or personal computer environments. This allows the database server to be any size computer, so that an installation can grow by installing a larger server, up to the very largest mainframe, or by installing multiple smaller servers, even on different brands of computer. Client computers from different manufacturers, using different brands and natures of interface software for different application tasks, can also be combined on the same network. Figure 4 is a diagram of a possible heterogeneous LAN-based library system.

In figure 4, a mainframe computer is shown filling the role of file server, as might be the case if an existing mainframe-based library system were extended with a LAN and improved interface software running on personal computers. This arrangement could improve service and extend the useful life of the mainframe at the same time. However, in the case of a new installation, or when complete conversion from an old system is called for, one of the more powerful personal computers would likely be an economical alternative to a mainframe file server.

CONCLUSIONS

SQL is the standard in database querying. It is directly applicable to many library data processing applications, such as circulation control. It is not directly applicable to patron-user computing, such as the online catalog. However, there are new SQL-interfacing products and methods available in the marketplace that can solve this problem.

SQL standardization has sparked a second generation of database management systems. These new systems are oriented to distributed application on local area networks. This client/server architecture makes available mainframe computer performance at personal computer prices, which is a development with many implications for library computing.

REFERENCES

1. American National Standards Institute, *Database Language—SQL*, Document ANSI X3.135-1986 (American National Standards Institute, Inc., New York, 1986).
2. C. J. Date, *A Guide To The SQL Standard* (Addison-Wesley Publishing Co., Inc., 1986).
3. Jitender S. Deogun and Vijay V. Raghavan, "Integration of Information Retrieval and Database Management Systems," *Information Processing and Management* 24, no.3, p.303–13 (1988).
4. Gio Wiederhold, *Database Design* (McGraw-Hill, New York, 1983).
5. Robert G. Crawford, "The Relational Model in Information Retrieval," *Journal of American Society of Information Science*, p.51–64 (Jan. 1981).
6. Oracle Corporation, ORACLE (20 Davis Drive, Belmont, Calif.).
7. Apple Computer, Inc., HyperCard (Cupertino, Calif.).
8. Steve Roti, "ORACLE Meets HyperCard," *DBMS* 2, no.1, p.52–62 (Jan. 1989).
9. R. G. Crawford and H. S. Becker, "A Novice User's Interface to Information Retrieval Systems," *Information Processing and Management* 22, no.4, p.287–98 (1986).
10. R. G. Crawford and Mary-Ellen Edwards, "A Prototype Mouse-Based Interface to Drug Related Information," *Online Review* 9, no.6, p.471–87 (1985).
11. Clifford Burgess and Kathleen Swigger, "A Graphical Database Interface for Casual, Naive Users," *Information Processing and Management* 22, no.6, p.511–21 (1986).
12. H. P. Frei and J. F. Jauslin, "Graphical Presentation of Information and Services: A User-Oriented Interface," *Information Technology: Research and Development* 2, p.23–42 (1983).
13. Peretz Shoval, "Principles, Procedures and Rules in an Expert System for Information Retrieval," *Information Processing and Management* 21, no.6, p.475–87 (1985).
14. Cynthia A. Kehoe, "Interfaces and Expert Systems for Online Retrieval," *Online Review* 9, no.6, p.489–505.
15. mdbs, Inc., GURU (P.O. Box 248, Lafayette, IN 47902).
16. Neuron Data, Inc., NEXPERT*SQL (444 High St., Palo Alto, CA).
17. Karen Watterson, "Server Technology: Divide to Conquer," *Data Based Advisor* 7, no.3, p.92–100 (March 1989). ■■

Geological/Geographical Name Subject Access

Joan E. Binder, Nancy M. Gustafson, and Meredith Merritt

Catalogers, geo-science faculty members, and graduate students from several Rocky Mountain Region institutions were surveyed. The catalogers were asked about their provision of subject access through geologic or geographic names, while faculty and students were consulted about their use of such access points. Strategies for providing optimal access are explored, balancing the needs of users for vernacular or variant forms of names and the sometimes conflicting needs of the cataloging community for adherence to protocols governing the construction and tagging of "correct" forms.

In the December 31, 1909, issue of *Science*, the secretary of the Geological Society of Washington summed up the dismal state of stratigraphic nomenclature with the declaration, "The present incubus of names is something to be shaken off at the earliest moment."[1] While this problem may be less onerous for stratigraphers these days, it continues to be a concern for contemporary librarians and patrons working with geological and geographic name subject terms in library catalogs. While wrestling with the incubus on the home front, we realized we needed more details on what problems exist for catalogers and patrons, how the problems arise, and what we can do about them.

There is much evidence that, in general, subject searching in library catalogs leaves much to be desired and yet is one of the most frequent, if not the most common, methods of gaining access to library collections. In their review of technical services research during 1987, Geraldene Walker and Judith Hudson find that "early studies of card catalog use led to the theory that subject access was of relatively little value to users, but more recent studies of online catalogs revealed a radical change."[2] With online catalogs, retrieval of specific information is made even more complicated due to peculiarities of individual systems and the vastly increased number of access points. Research has shown that nearly half of online catalog subject searches retrieve nothing—Judith Adams condemns this as "an appalling state of affairs."[3] Lois Mai Chan acknowledges that "online catalog use studies bear witness to the fact that our subject access arrangements fall far short of fulfilling the potential of the online catalog. Improvement is needed—urgently needed—in the light of how fast the MARC database is growing."[4]

That subject access is increasingly of concern is borne out in recent studies. Chan explores the possibilities of using Library of Congress classification as an alternative approach to subject access by pointing out its unique features, implications, and future work needed to make this a realistic alternative.[5] In a comprehensive and notable examination of this topic, Karen Markey consolidates many of the findings of earlier

Joan E. Binder is the Technical Services Librarian at the College of Law, University of Wyoming. Nancy M. Gustafson is the Principal Serials Cataloger at the University of Wyoming, and Meredith Merritt is a Monographic Cataloger at the University of North Carolina at Charlotte.

studies sponsored by OCLC and the Council on Library Resources.[6]

Micco, Smith, Hsiao, and Intaravitak, in attempting to design a prototype database for an integrated system, maintain that the MARC bibliographic format as a tool for "bibliographic record keeping" has distinct limitations for subject access: "It is not practical for periodical literature, and is virtually useless for reference books. Library of Congress subject headings were never designed or intended for postcoordinate searching, nor do they provide systematic subject access as demonstrated in study after study."[7]

In a recent article, Carolyn Frost and Bonnie Dede approach the notion of matching catalog headings with terms used in LCSH.[8] Markey proposes in another paper that patron access to Library of Congress Subject Headings in machine-readable form (LCSH-mr) may be a potential breakthrough in subject searching if an interface can be provided to overcome such obstacles as errors, outdated headings, free-floating subdivisions, subdivisions applied on the basis of "model" or "pattern" headings, and the addition of geographic subdivisions to headings, for which free-floating lists are not available.[9]

James Ross, as one of the few investigators interested primarily in geographic name access, echoes some of these same obstacles as barriers to geographic name access, lamenting "the changes AACR2 has brought to heading structures of all kinds, the existence of pre-AACR2 headings that have not been revised to AACR2 forms . . . the ambiguous and inconsistently applied instructions sometimes found on the LC list . . . and the difficulty of verifying headings in online cataloging or in the online authority file."[10] Names for geographic entities are also examined by William Studwell, although he focuses on the differing treatment by LC of corporate jurisdictions and subject place names. He postulates, "Perhaps this dual function of geographic name has been an element of confusion in the development of the system for geographic names as subjects."[11] He goes on to suggest that LC adopt the same pattern as has been used for personal names—that is, to use only one form in subject access.

Aside from these studies by Ross and Studwell, the peculiar problems inherent in geographic and geological name access in subject searching have not been specifically addressed in much of the literature to date. This topic merits attention, however, since having an area name will theoretically help to narrow a search, especially in an online environment. People do have trouble refining topical subject searches due to such factors as controlled vocabularies and system designs. When geographic or geologic names, with all the attendant problems of discerning the correct form, are added to search queries, the situation is doubly complicated. Furthermore, a good deal of effort on the part of catalogers may be expended when newly establishing a geographic or geologic name heading according to standards and procedures adopted by LC and AACR2.

In an earlier study on faculty subject searching, Carolyn Frost points out that "the focus of research on subject searching in library catalogs has been primarily on catalog use by students. . . . But faculty use of the subject search is also of concern."[12]

With that in mind, and presented with the vexing problem of establishing geographic and geologic names ("geo-names") in a remote area not much explored by LC's subject catalogers, we began gathering information both from catalogers and users (faculty and graduate students). Three surveys were designed on the basis of our experience, similar to the idea recently expressed by the Association of Research Libraries, Systems and Procedures Exchange Center: "Short, specifically targeted surveys with a limited objective can be very effective means of obtaining information to improve existing services, determine use patterns and/or frequencies, and plan new services, especially regarding use of new technologies." The Center also observes: "In many cases, strictly random samples may not be necessary for libraries to achieve a high confidence level in survey results."[13] Surveys developed included: one for catalogers or heads of cataloging departments, one for geo-science faculty, and one for graduate assistants. This final one was restricted to University of Wyoming students, but the first 2 questioned people at selected institutions in the entire Rocky Mountain Region. Response rates for the 3 surveys were 100, 33, and 24 percent. Although the response rates for the last 2 are not overwhelming, the results do provide valuable qualitative information.

CATALOGER SURVEY

We sent surveys to 13 cataloging departments at 12 universities chosen from the University of Wyoming's administrative listing of comparable institutions in the Rocky Mountain Region and to 5 state library cataloging departments in Colorado, Idaho, Montana, Utah, and Wyoming. The mailing was preceded by a telephone call to each institution ascertaining willingness to participate; since this was positive in every case, we determined exactly who was responsible for cataloging operations and would receive the survey. All surveys were returned, either by department heads or principal catalogers involved with name authority procedures. Ten of the libraries belong to OCLC, 6 participate in WLN, 3 in RLIN, and 1 in NOTIS. (One university belongs to 3 networks.)

The questions sent to catalogers were based on problems we have encountered in establishing geo-name subject headings at the University of Wyoming. The foremost one is the problem of discrepancies between forms of names used in searching and forms established in library catalog authority systems. Do these discrepancies adversely affect user success rates in gaining access to pertinent information sources?

Several factors contribute to the problematic nature of finding or setting up such names in libraries where Library of Congress Subject Headings (LCSH) are used. In order to assign MARC subject heading tags 650 or 651 with the second indicator of zero a cataloger is supposed to follow Library of Congress examples or adhere to LC's published guidelines for formulating geographic/geologic names. If the heading has not been established by or for the Library of Congress in the authority file or in LCSH, the next step is to look for an occurrence of the name as a subdivision in LC bibliographic records. Since subheadings are generally not indexed in many catalogs, finding a particular one can be a matter of intuition, diligence, and luck.

In the cases where LC has not used the name in its cataloging, or when a discrepancy is discovered in actual LC practice, the cataloger might establish the heading locally using the same resources LC would use: personal contact with BGN (United States Board on Geographic Names), *Columbia Lippincott Gazetteer*, *Webster's New Geographical Dictionary*, *National Gazetteer of the United States of America*, and other standard sources as available. These reference works would be used in conjunction with the specific Library of Congress guidelines for establishing such names in *Subject Cataloging Manual*, (H690 rev. 11/25/86 and H810 rev. 8/14/85),[14] together with policy decisions published in issues of *Cataloging Service Bulletin*.[15] The Board on Geographic Names issues *Decisions on Geographic Names in the United States* quarterly[16] and also responds to catalogers' requests to help resolve a problem. BGN staff are extremely helpful in verifying locations, resolving conflicts and providing information recently added to their files. Furthermore, the Library of Congress will respond to catalogers needing a clarification of policy or advice on establishing a local heading. The major problem with the two last-mentioned resources is that initiating correspondence adds another step to cataloging and requires more time than many understaffed libraries can spare.

Whatever the cataloger determines the official form to be and however carefully LC's guidelines are applied in formulating the geo-name for catalog entry, there is always the risk that LC will construct the entry differently in the future. Although LC usually accepts the BGN form of name, it retains the option of rejection.

Another significant problem encountered in formulating geographic or geologic names is that the form of entry prescribed by LC rules may differ markedly from the form appearing on the chief source of information for cataloging or the form in everyday use.

According to Carolyn Frost,

> It is widely acknowledged that a major source of difficulty in subject searching is the user's failure

to match a search term with the term used in the catalog as an access point. It has also been pointed out often that users are either unaware of the existence of the catalog's source of subject terms—i.e., the Library of Congress Subject Headings List (LCSH)—or find it difficult or inconvenient to use.[17]

An example is the feature in Wyoming and Montana known as the Powder River Basin. Although often considered a geologic basin, LC rules dictate that it be established as the Powder River Watershed (Wyo. and Mont.), not Powder River Basin. LC uses "watershed" to designate river basins and drainage basins. While the first form is "correct" according to Section H690 of the *Subject Cataloging Manual* and has been used in AACR2 LC bibliographic records, it would rarely occur to a patron to construct such a search strategy. To confound the matter, the name could appear on pre-AACR2 cataloging records as "Powder River Basin"; "Powder River Basin, Wyoming"; "Powder River Basin, Wyo."; "Powder River Basin, Wyo. and Mont."; or simply, "Powder River."

CATALOGERS' SURVEY RESULTS

Determining the official name as described above theoretically solves the problem of standardizing geo-name access at the network level but leaves the dilemma of providing access to the variety of names often given to a single place or feature. Bearing in mind the difficulty facing a searcher who knows perhaps only one of the many possible variants for a given heading, and the time-consuming nature of geo-name authority work, we designed a survey to determine:

- how much time catalogers spend on geo-name authority work;
- what kinds of materials need the most authority work;
- what kinds of features are the most troublesome to formulate;
- how often conflicts between local usage versus the Library of Congress' controlled vocabulary are found and how they are resolved; and,
- how catalogers assign subject subdivisions to provide access to the area in question.

This survey reveals that 8 of the 18 departments surveyed spend as much time on geo-name authority work as on other types of authority work, while 5 spend more time, and only 4 spend less time. One provides no comment. Original cataloging requires more geo-name authority work than does comparable adaptive cataloging for 12 departments.

Different types of publications being cataloged demand differing amounts of geo-name authority work, with state publications needing the most overall and theses following as a close second. Commercial publications generally take less than either of these. While several catalogers say maps generate a good bit of such work, it is misleading to include them in the ranking, since not all responding departments catalog maps as completely as other materials.

Bodies of water are the type of feature most often needing authority work, but basins, valleys, mountains, and ranges, along with national parks and recreation areas also cause concern. Other features generating authority work for some respondents include quadrangles, dams, watersheds, geologic formations, and archaeological sites.

One of the survey questions concerns the frequency of discrepancy between form of name used on the work being cataloged and form of name used by LC or that resulting from application of LC's rules. Only one department feels there would always be a conflict, apparently assuming that a qualifier would need to be added in any case. Another department finds conflicts very frequently, 6 somewhat frequently, and 10 occasionally. As a solution to this situation, 13 departments employ cross-references, and 4 of the 8 with online catalogs provide keyword access as an alternative. Only 2 lack provisions for access to variant forms of geo-names.

Seven departments refer conflicts and inquiries to LC, 4 to regional or state agencies, 4 to their network (WLN), and only 2 to BGN. Some respondents refer to more than one agency, and others solve conflicts internally.

When asked to choose among subject headings they would use in their library for an item on Lizard Gap Basin located in Wyoming (a made-up name), 6 catalogers choose a county-level subdivision (e.g., Paleontology—Wyoming—X County), 11 subdivide by LC form of feature name (e.g., Paleontology—Wyoming—Lizard Gap River Watershed), and 5 subdivide by the feature's common name (e.g., Paleontology—Wyoming—Lizard Gap Basin). Some respondents have marked more than one option, and others have marked no option at all.

In summary, it appears that most catalogers are aware of the need for geo-name access and make some provisions for these needs.

FACULTY SURVEY

In an effort to determine if the frequency of such prescribed terminology has an impact on information access, a survey on geo-name searching practice was sent to faculty members in geology and geography departments at the 12 universities whose library cataloging departments responded to our first survey. We estimated from directories and college catalogs that there were approximately 300 geo-science faculty members at these schools and sent the appropriate number of surveys to each department head for distribution.

The survey was designed to elicit the following information:

- how faculty gain access to the collections for research and teaching (e.g., card catalogs, online catalogs, browsing, database searches, etc.) and which methods are preferred;
- how they use geo-names in subject searching and to what extent;
- how successful they are when looking for materials using a geo-name;
- how often they encounter different forms of headings than they are expecting to find;
- if geo-name access is more important in some specializations or subdisciplines than others; and
- what modifications to current subject access by geo-name is desired.

Of the responses we received, 100 faculty respondents (33 percent of the surveyed group) show considerable interest in geo-name access points.

Of these respondents, 94 say they use the card catalog, and 44 search online (an encouragingly high figure, considering that only 8 of their libraries have online catalogs). Overlap exists here, since several respondents use both card and online catalogs. For the record, faculty also show interest in browsing, along with some use of database searches and assistance from reference librarians. Many "other" methods are listed by faculty respondents—the most usual are periodical indexes and other relevant bibliographies, citations in published literature, and discussions with colleagues.

When conducting a subject search in the card catalog, 3.2 percent of the faculty members always use a geo-name as their sole search element, while 18 percent do so very frequently, 7.4 percent somewhat frequently, 47.7 percent occasionally, and 19 percent never. In the online catalog, 6.8 percent always apply this strategy, 15.9 percent use it very frequently, 20.4 percent somewhat frequently, 43.1 percent occasionally, and 27.2 percent never. (See figure 1.)

Among faculty card catalog searchers, 6.4 percent always include a geo-name in combination with another search element, 31.8 percent add one very frequently, 14.8 percent somewhat frequently, 29.7 percent occasionally, and 10.6 percent never. Of the online searchers, 11.4 percent always adopt this approach, 31.8 percent adopt it very frequently, 27 percent somewhat frequently, 29.5 percent occasionally, and 15.9 percent never. (See figure 2.)

Our next question asked how often faculty searchers find what they are looking for when they search using a geo-name. Only 2.1 percent of the card catalog users say they always find what they are looking for, but 17 percent say their searches are very frequently successful, with 32.9 percent succeeding somewhat frequently, 33.9 percent occasionally, and 2.1 percent never. Of the online respondents, none succeed always in finding what they are looking for, 22.7 percent succeed very frequently with geo-name searches, 36.3 percent somewhat frequently, 31.8 percent occasionally, and 6.8 percent never. (See figure 3.)

The answers to our question about how often catalog entries for geo-names appear in different forms from the users' search terms are hard to interpret. The figures

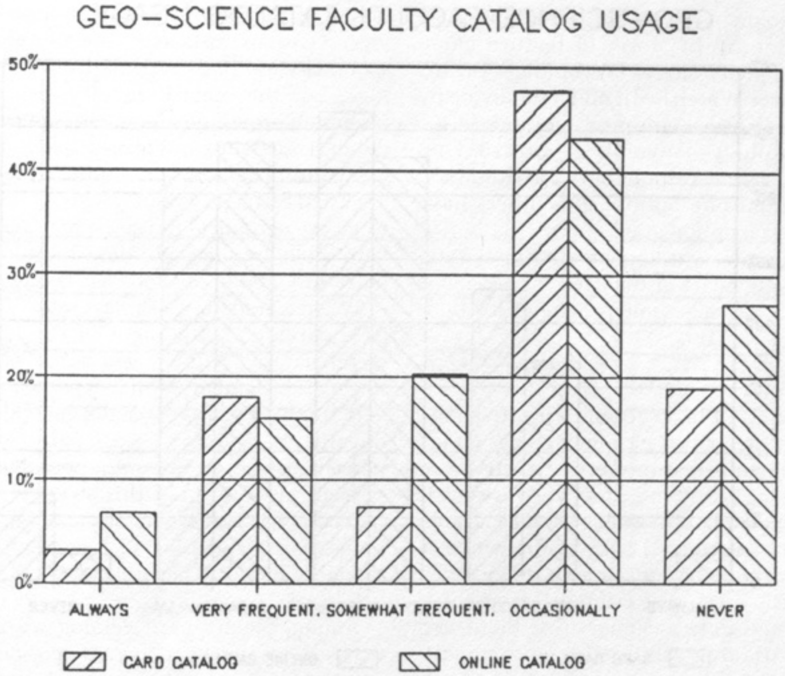


Figure 1. How Often Faculty Searchers use a Geo-name as Their Sole Search Element.

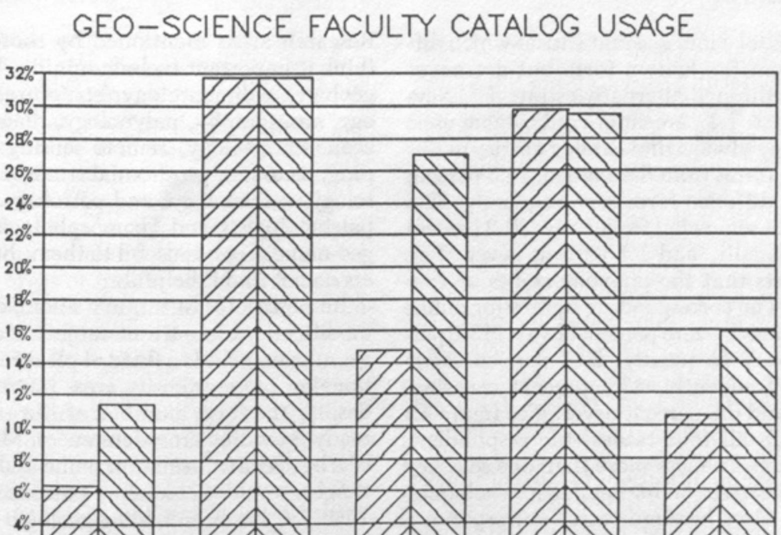


Figure 2. How Often Faculty Searchers Include a Geo-name in Combination with Another Search Element.

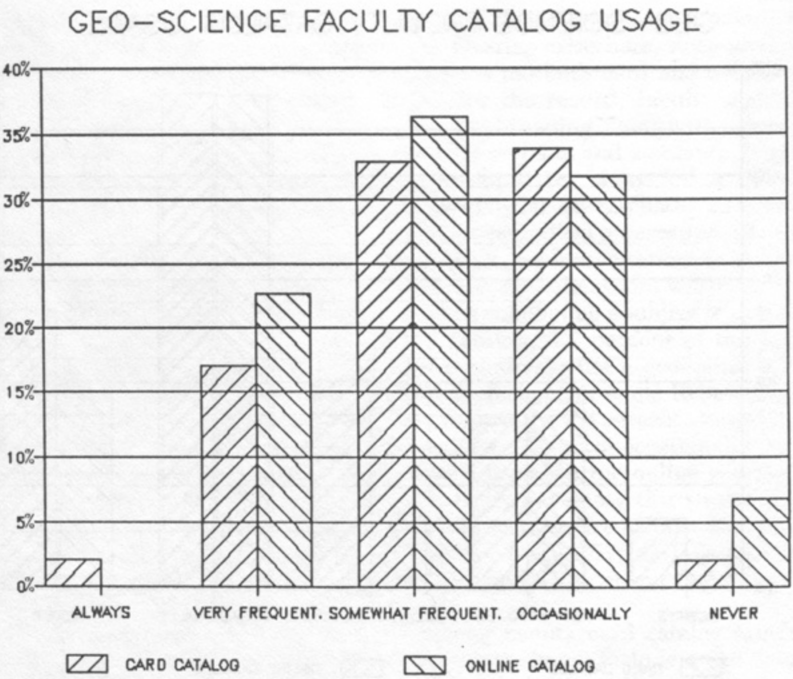


Figure 3. How Often Faculty Searchers Find What They Are Looking for When They Search Using a Geo-name.

cannot take into account entries which differ from the known form but are never found through alternative strategies. Nevertheless, 1.1 percent of card catalog users say they always find catalog entries in different forms from their usage, 14.8 percent find a different form very frequently, 22.3 percent somewhat frequently, 39.2 percent occasionally, and 1.1 percent never. This suggests that the phenomenon is of concern. The corresponding figures for online searchers are zero percent always, 13.6 percent very frequently, 20.4 percent somewhat frequently, 43.1 percent occasionally, and 6.8 percent never. (See figure 4.)

With all four tables, the respondents sometimes choose more than one selection under card or online catalog for each question. Also, in some cases, no response is given.

Over half of the geo-science faculty surveyed (55) feel access by geographic or geologic name is definitely important in their subdisciplines. Fifteen find it helpful sometimes. Only 10 feel such access is rarely of interest, and 11 see it as definitely unimportant. Nine did not answer this question. Research areas mentioned by those who think it important include mining, hydrogeology, sedimentology, structural geology, stratigraphy, palynology, diagenesis, economic geology, remote sensing, mapping, tectonics, geothermal studies, climatology, volcanology, and petrology. Some paleontologists and geomorphologists say geo-name access is useful to them, but others do not find it helpful.

Responses to an inquiry about desired modifications to current subject access by geo-names include a flood of pleas for additional or more extensive cross-referencing, despite the large number of libraries already providing cross-references. More access by specific formation name and quadrangles would please some patrons. (Since 1986, LC has been adding geologic formation names to subject heading lists, a policy announced in *Cataloging Service Bulletin* 34.)[18] Overall, our respondents would simply like to see more geo-names in their catalogs, in the forms that are familiar to them.

GRADUATE SURVEY

A third survey was designed to gather

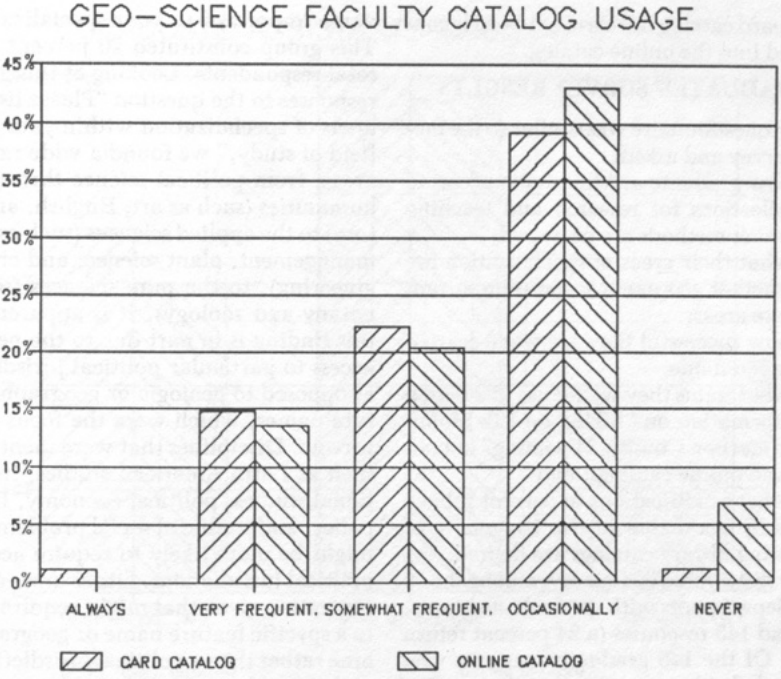


Figure 4. How Often Faculty Searchers Find Geo-name Catalog Entries in Different Forms from Their Search Terms.

data from other users of the catalogs—specifically, graduate assistants. We chose this population for several reasons. On the premise that some graduate students provide assistance with faculty research, it became apparent that their search methods also needed to be assessed. Additionally, we were curious to see what disciplines rely on geo-name searching at the graduate level, where areas of specialization and research might be more established than those of undergraduate students. We wanted to survey within the Rocky Mountain Region, in order to make a comparison to the previous surveys, a population of students likely to have an interest in the area's geological and physiological features. In order to limit the population size and distance, we elected to survey the University of Wyoming's graduate assistants.

At the time of the survey the University of Wyoming's online catalog contained about 80 percent of the collection, including all post-1978 monographs and approximately 85 percent of the periodicals and serials. It did not represent government document holdings, which were accessible only through standard printed indexes at that time. Searching of the online catalog could be done by author, title, series, subject, ISBN/ISSN, and keyword.[19] Keyword searching accessed most fields except imprint, notes, and some serials linking fields and could be enhanced through the use of Boolean operators and truncation. Unconverted holdings were found only in the card catalog, which was separated into author/title and subject entries and arranged by the 1968 *ALA Filing Rules*. At the time of the survey, many of the geological materials were not in machine-readable format, either lacking LC copy or still not converted from Dewey; therefore, we might safely assume that researchers interested in geo-names were familiar with the card catalog as well as the online catalog.

Headings in both catalogs have been generally assigned according to LC practice, either taken directly from LCSH or established locally on the basis of LC's guidelines (or lack thereof, since their proliferation has occurred comparatively recently in *Cataloging Service Bulletins* and in LC's subject manual). Cross-references are used in the card catalog but have not been incorporated into the online catalog.

GRADUATES' SURVEY RESULTS

Our questionnaire was similar to the faculty survey and asked:

- how graduate assistants gain access to the collections for research and teaching and which methods are preferred;
- what their areas of specialization are and whether geo-name access is important for these areas;
- how successful they are when searching by geo-name;
- what terms they would use in a search for information on "hiking on Elk Mountain in Carbon County, Wyoming" in both card and online catalogs; and
- what modifications to current subject access by geo-name in the University of Wyoming library catalogs are desired.

We distributed 604 surveys among the 19 UW departments with graduate assistants. We had 145 responses (a 24 percent return rate). Of the 145 graduate assistants who responded, 49 percent say that they use the card catalog to locate library materials for use in their research and teaching, and 97 percent search via the online catalog. Respondents indicate that they use both the online and card catalogs.

Regarding other methods for locating library materials for use in research and teaching, 78 percent say they use printed abstracts, indexes, and bibliographies. Twenty-six percent use computerized abstracts, indexes, and bibliographies, such as databases on Dialog and BRS. (This seems appropriate since the University of Wyoming charges for computerized searches.) Fifty percent browse the book and periodical stacks. Thirty-seven percent query the reference librarians, while 9 percent use other sources, such as referrals from other students and faculty, citations found in the literature, University of Wyoming's Document Delivery (a journal article delivery service), *Current Contents*, and the "new periodicals shelf." Once again, the respondents could mark all that applied.

Since an aim of our study was to investigate the importance of geo-names in searching, we centered our attention on the group giving a positive response to the question "Is locating materials by geo-name important in your specialization?" This group constituted 20 percent of the total respondents. Looking at this group's responses to the question "Please list your areas of specialization within your major field of study," we found a wide range of areas, from political science through the humanities (such as art, English, and history) to the applied sciences (such as range management, plant science, and civil engineering), to the pure sciences (such as botany and zoology). It is apparent that this finding is in part due to the need for access to particular political jurisdictions as opposed to geologic or geographic feature names, which were the focus of our surveys. Disciplines that were mentioned, such as Latin American studies, international studies, political economy, foreign policy, and studies of social problems, also might be more likely to require access to political jurisdictions. Other areas of concentration given that might require access to a specific feature name or geographical area rather than a political jurisdiction are archaeozoology, faunal studies, engineering geology, soil and rock mechanics, stratigraphy, sedimentology, veterinary parasitology, exercise physiology, and a host of specializations relating to water management and ecology. E. J. Coates in "Significance and Term Relationships in Compound Headings" cites the topics "Geography, Geology, History, Social Sciences, Linguistics, Literature, Ecology, and names of groups of natural organisms" as "significantly conditioned by locality"; therefore, they merit entry under place-name.[20]

Of the 20 percent group mentioned above, 68 percent answered on the usage of the card catalog. Of this 68 percent, 35.3 percent find most of the time what they are looking for when they search using a geo-name, 50 percent some of the time, and 14.7 percent not very often. Regarding this same 20 percent group, 94 percent answered on the usage of the online catalog. Of this 94 percent, 29.8 percent find most of the time what they are looking for when they search using a geo-name, 47.8 percent some of the time, and 22.4 percent not very often. (See figure 5.)

Of the respondents from this 20 percent who list other methods for locating re-

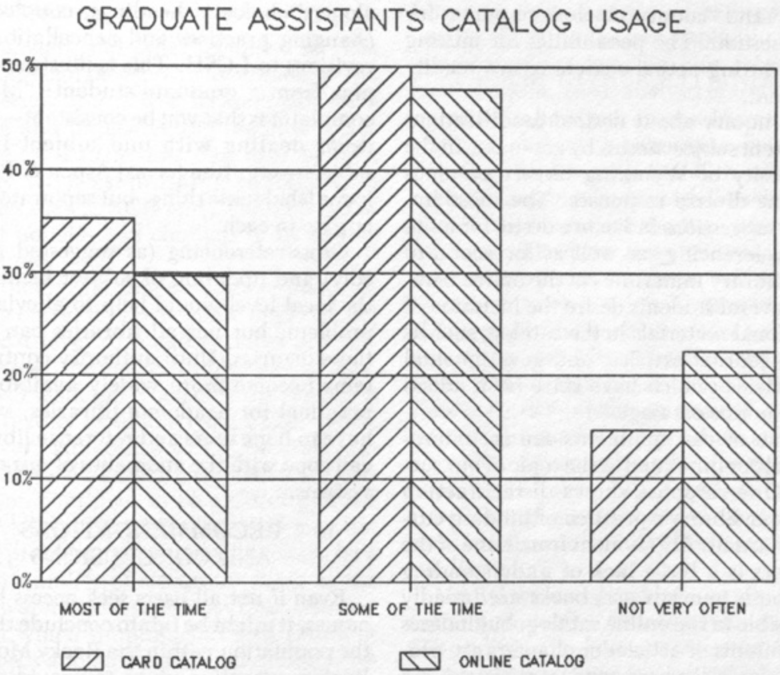


Figure 5. How Often Graduate Assistants (Who Felt Finding Materials by Geo-Name Is Important) Find What They Are Looking for When They Search Using a Geo-name.

search or teaching materials, 76 percent use printed abstracts, indexes, and bibliographies, while 28 percent use computerized abstracts, indexes, and bibliographies including databases on Dialog and BRS. Forty-eight percent browse book and periodical stacks, 31 percent query reference librarians, and 10 percent list other methods. This last category includes referrals from other students and faculty and citations from published papers. Many use more than one method.

We were interested in what subject terms this same 20 percent group (those who feel that locating materials by geo-name is important in their area of specialization) would use to search in the card and online catalogs for an item about "hiking Elk Mountain in Carbon County, Wyoming." A cataloger using LCSH and LC's subject manual could assign the following subject headings for the catalogs: (1) Hiking—Wyoming—Elk Mountain (Carbon County : Mountain); (2) Elk Mountain (Carbon County, Wyo. : Mountain); and (3) Hiking—Wyoming—Carbon County. As expected, the feature "Elk Mountain" is the most frequent search strategy for both the card and online catalogs. The second most frequent response for the card catalog is "Carbon County." For online searchers, it is a tie between "Hiking" and "Carbon County." The third most frequent response for the card catalog is "Wyoming," whereas for the online catalog it is "Recreation."

In some cases, respondents use "Wyoming" in combination with "Elk Mountain" and "Carbon County." The success rate for finding items using these combinations in the card catalog is greatly affected by the patron's familiarity with LC's policies of direct as opposed to indirect subdivisions and of qualifying non-unique names. Assuming a patron understands that both headings and subdivisions are indexed in our online catalog, the user might also need to be familiar with LCSH terminology, the practice of abbreviating names in headings, and perhaps the truncation feature in the online catalog. In fact, some graduate assistants abbreviate "mountain," "Wyoming," and "county" in their responses for this question. The possibilities for missing items during actual searching are readily apparent.

An inquiry about desired modifications to current subject access by geo-name in the University of Wyoming library catalogs brought diverse responses. The most frequent suggestions indicate a desire for more cross-referencing, as well as for access to more library materials via the online catalog. Several students desire the inclusion of additional materials in the catalog, such as maps, journal articles, and governmental periodicals (which have since been added to the online catalog).

Some of the comments are incomprehensible or unrelated to the topic of our survey. One response shows dissatisfaction with our library's practice of binding current journals. Also evident from some of the answers is a basic lack of understanding that both journals and books are broadly accessible in the online catalog, but indexes for contents of articles or chapters are provided for elsewhere, and explicit volume holdings information is found in yet another source (probably the "microfiche" referred to in one of the responses). One comment suggests the consolidation of our card catalog with the online catalog, and another one reads "spend more money on books, periodicals and the like instead of spending it on new cataloging systems" indicates a certain naïveté concerning sources and expenditures of library funding, which one might expect. At the same time, however, it is asking a great deal of users to hope they can understand the searching idiosyncracies of two dissimilar systems.

User instruction in the intricacies of searching an automated system is even more important than ever. Not only is each individual system a little different in how things are indexed, but also it is crucial that the user understand the consequences of incorrect data entry (e.g., incorrect spelling or punctuation, entering data in an inappropriate search field, or attempting to search fields not indexed at all).

The difficulties of subject retrieval in machine searching are due in part to the discrepancies of different systems with variant indexing designs.[21] Other problems with subject retrieval are influenced by catalogers' choice of headings, coupled with changing practices and cancellations and revisions to LCSH. This is illustrated by a plea from a graduate student: "My only complaint is that you be consistent—put all items dealing with one subject in that subject—e.g. [the terms] Aspen & Populus index [the] same thing, but sep[arate] items pop up in each."

Cross-referencing (as suggested repeatedly) and updating of subject headings at the local level would help to alleviate this problem, but not all libraries can afford these luxuries. Until authority control systems become more widely available and practical for academic libraries, we will have to hope users and reference librarians can cope with the anomalies of our current systems.

RECOMMENDATIONS AND CONCLUSIONS

Even if not all users seek access by geonames, it might be fair to conclude that, for the population within the Rocky Mountain Region, attention given to providing such access by catalogers is well worth the effort. Ours is a region of considerable interest to geologists and geographers. There is substantial literature on the region, and in many cases if catalogers do not provide the access there may very well be none, especially for materials in regional publications. Libraries in other regions of interest to the earth sciences may find themselves in a similar position.

A multifaceted approach taken by librarians, researchers, and designers of information technology would ameliorate the difficulty of providing increased bibliographic access to geologic and geographic names. Catalogers need to provide appropriate and sufficient access points and, together with public services librarians, foster effective searching skills in users. Since access is also dependent on the design of the system itself, librarians should further direct their attention to systems and system-suppliers. Increasing demand for and manufacture of such features as user-friendly interfaces, online thesauri and authority control, and flexible searching capabilities may eventually reduce users' dependence on knowledge of standardized forms.

Cataloging

Catalogers and other technical services librarians can develop an awareness of system design and newer technologies, promoting and implementing the most effective ones. Librarians can also benefit greatly from establishing contact with agencies such as BGN and LC for clarification of policies and conflicts in setting up geo-names.

Further, regarding the form of name for catalog entry, it may be more important at the local level to make variants or commonly used forms accessible than to agonize over what the "correct" form is when it is not readily available from LC or BGN. However, adherence to LC or other national library authority standards for formulation continues to be a priority at the network level. In many cases this may not pose a problem, but there will always be a certain number of situations, such as the basin/watershed discrepancy discussed earlier, in which simultaneous satisfaction of user needs and national standards may create direct conflict. In such cases, each library may wish to arrive at a compromise to connect the needs of users with the available system. The keyword approach to online searching, bypassing the use of controlled vocabulary terms, might be espoused as one solution to this dilemma. Lois Mai Chan remarks in *Library of Congress Subject Headings: Principles and Applications:*

> The power of keyword searching on words in titles alleviates to a certain extent problems with lack of currency or failure to reflect common usage. Many subject indexes rely entirely on keyword display based on article titles or abstracts. Titles of documents that are expressive of content increase the effectiveness of keyword title access. Furthermore, this benefit is not available only to online catalog users; generous use of partial-title-added entries, to bring potentially useful access terms into filing position, can have much the same effect in the manual environment. Helpful as keyword searching can be, however, few theorists advocate abandoning the use of controlled vocabulary in library catalogs. Keyword access to terms in titles can be an adjunct to systematic indexing, but no more.[22]

Automated authority control systems and patron-accessible authority files may be a partial solution in the not-too-distant future for larger academic or research libraries. LCSH does not explicitly list all possible unique heading plus subdivision combinations. In an analysis of topical and geographic headings in the University of Michigan Library's catalog to determine degree of match with the 10th edition of LCSH, Carolyn Frost and Bonnie Dede find a higher than expected agreement (88.4 percent) of headings without subdivisions; however, a comparison of subdivisions alone reveals only a 31 percent match of those in LCSH, which might have been higher if a file of free-floating subdivisions were also matched. They also show that "less than 12% of the geographic subdivisions were found in LCSH" and acknowledge that "geographic subdivisions, however, will pose a problem unless free-floating lists of some kind can be developed for these as they have been for topical subdivisions."[23]

MARC formats provide some alternative tagging strategies. For example, in fields 650 and 651 (for books and serials) OCLC allows second indicator 7 with subfield 2 to define authorities for subject terms other than the national authorities. Locally formulated subject terms are valid on OCLC in fields 690 and 691, with the stipulation that they are "not part of the standard LC-MARC formats" for original records.

Despite these provisions, LC subject authority resulting in 650 and 651 fields with the second indicator 0 will be the most attractive option for most libraries. Use of the tags for local subject terms may entail unnecessary complexity in setting up profiles for automated systems, and this can contribute to problems in standardization for shared access when several libraries plan to merge data for union catalog purposes. Use of tags for local subjects also affects the level of programming required. In these situations, the mandate to accept LC forms of subject terms and to formulate original terms by LC rules saves programming effort and implementation headaches.

User Education

Developing user awareness may be the most immediate and potentially most effective means of guiding users through the intricacies of bibliographic systems and library materials arrangement. Such an effort should involve both technical and public services in bibliographic instruction in conjunction with site-specific user surveys to learn more about those interested in geo-name access and how best to serve their needs. It may also be worthwhile for cataloging departments to develop an awareness of which departments at their institutions are most likely to use geo-name search strategies and to direct bibliographic instruction appropriately.

Indeed, enhanced bibliographic instruction would go a long way toward getting users to realize the potential of current library indexing systems, regardless of their particular fields of study. Libraries still having divided systems should be sufficiently staffed to make sure people are using the tools appropriately. Complete conversion needs to remain a top priority. If users' first interactions with the library are successful, there is a greater likelihood they will make subsequent visits.

Technology

Improved technology applied to sophisticated, integrated library systems can complement quality cataloging and bibliographic instruction. Authority control and online thesauri should not be overlooked as beneficial tools for users and librarians alike, possibly with user-friendly interfaces that connect these to the databases. Current experiments with incorporating table-of-contents data into online systems containing keyword access, if successful, may eventually obviate the cataloger's concern with the pitfalls of LCSH.

A certain degree of caution must be observed, however, when expecting users to undertake searching such a file without appropriate training and understanding of the indexing protocols. For example, a keyword query "Bismarck" might retrieve the following array of results, including author, title and subject entries: Bismarck, Otto, Fürst von, 1815–1898; Bismarck, Herbert, Fürst von, 1849–1904; Bismarck (N.D.); Bismarck Archipelago (Papua New Guinea); Bismarck Range (Papua New Guinea); Grain Transportation Forum, Bismarck, N.D.; Ōkubo Toshimichi, the Bismarck of Japan; Bismarck-Museum; Bismarck National Bank (Bismarck, N.D.); Bismarck Bank (Bismarck, N.D.); Bismarck family; Bismarck (Ship); and, Bismarck Sea, Battle of, 1943.

Obviously, the main problem in searching without a controlled vocabulary is the retrieval of items unrelated to the desired topic. One way of reducing false drops is a flexible system that enables users to apply Boolean logic, and qualifiers, providing they understand how these strategies work and what the situations demand.

Further Research

Issues such as those mentioned above necessitate more survey research and analyses to find out more about the users themselves—the strategies they employ and why they use them; to what extent they perceive bibliographic instruction to be helpful; what happens in cases of failed searches; how well they understand bibliographic systems and the application of truncation, Boolean logic, and qualifiers; how they react to table-of-contents and other keyword indexes; and whether they are equally adaptable to both controlled and uncontrolled vocabularies.

The topic of user instruction raises several questions, such as whether bibliographic instruction is effective and how it might be more so; what user groups at what levels it can benefit the most; whether automated authority files and user-friendly systems lessen the need for instruction; how to ascertain searchers' abilities and thus direct instruction appropriately; how technical services librarians might contribute jointly towards this effort; and what other means are available to promote user awareness about library systems. Can we guarantee that users find what they need by providing them with enough skills to effectively use library catalogs rather than relying on the sheer persistence of a small percentage of users?

Questions such as these are beyond the scope of this paper but are fertile ground for continued investigation. We hope the potential to better accommodate users can be realized by designing improved systems and ensuring that the confidence in being able to use them is ultimately transferred as automatically as the data itself.

ACKNOWLEDGMENT

This paper represents the fruits of a study

conducted over the last two years. A very brief report of the first two surveys appeared in a regional newsletter, *Action for Libraries* (February and March 1987 issues). The surveys and resulting study were made possible by a grant from the Office of Research, University of Wyoming and in part by the authors' affiliated institutions.

The cooperation extended by the catalogers, geo-science faculty, and graduate assistants who responded to the surveys should be recognized, as should the patience of many library staff members in allowing the authors time to prepare the analysis.

REFERENCES AND NOTES

1. Arthur Keith, "The Status of Geologic Names," in Philip S. Smith, "The Geological Society of Washington," report of a paper presented at the 224th Meeting of the Geological Society of Washington, December 8, 1909, *Science* n.s. 30, no.783:975 (Dec. 31, 1909).
2. Geraldene Walker and Judith Hudson, "Research Methodology in Technical Services: The Case of 1987," *Library Resources & Technical Services* 32, no.4:360 (Oct. 1988).
3. Judith A. Adams, "The Computer Catalog: A Democratic or Authoritarian Technology?" *Library Journal* 113, no.2:33 (Feb. 1, 1988).
4. Loi Mai Chan, "Library of Congress Classification as an Online Retrieval Tool: Potentials and Limitations," *Information Technology and Libraries* 5, no.3:191 (Sept. 1986).
5. Chan, "LC Classification as Retrieval Tool," p.181–92.
6. Karen Markey, *Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs* (Dublin, Ohio: OCLC, 1984).
7. Mary Micco, et al., "Knowledge Representation: Subject Analysis," *Library Software Review* 6, no.2:84 (March–April 1987) citing George Martin Sinkankas, et al., *A Study of the Syndetic Structure of the Library of Congress List of Subject Headings* (Pittsburgh: University of Pittsburgh, Graduate School of Library and Information Science, 1968).
8. Carolyn O. Frost and Bonnie A. Dede, "Subject Heading Compatibility Between LSCH and Catalog Files of a Large Research Library: A Suggested Model for Analysis," *Information Technology and Libraries* 7, no.3:288–99 (Sept. 1988).
9. Karen Markey, "Integrating the Machine-Readable LCSH into Online Catalogs," "Special Section: Online Subject Access," *Information Technology and Libraries* 7, no.3:299–312 (Sept. 1988).
10. James Ross, "Geographic Headings Online," *Cataloging & Classification Quarterly* 5, no.2:27 (Winter 1984).
11. William E. Studwell, "Subject Suggestions 1: Some Concerns Relating to Geographic Areas," *Cataloging & Classification Quarterly* 8, no.2:93 (1987/88).
12. Carolyn O. Frost, "Faculty Use of Subject Searching in Card and Online Catalogs," *Journal of Academic Librarianship* 13, no.2:86 (May 1987).
13. Association of Research Libraries, Office of Management Services, Systems and Procedures Exchange Center, "User Surveys," *SPEC Flyer* 148 (Washington, D.C.: Association of Research Libraries, 1988).
14. Library of Congress, Subject Cataloging Division, *Subject Cataloging Manual: Subject Headings*, rev. ed. (Washington, D.C.: Library of Congress, 1985).
15. Library of Congress, Processing Services, *Cataloging Service Bulletin* 11–12, 15–16, 21–22, 28, 30, 34, 38–39 (Winter 1981–Winter 1988).
16. United States Board on Geographic Names, *Decisions on Geographic Names in the United States* (Washington, D.C.: Dept. of the Interior, 1963).
17. Frost, "Faculty Subject Searching," p.90.
18. Library of Congress, Processing Services, *Cataloging Service Bulletin* 34:46 (Fall 1986).
19. The University of Wyoming online catalog is a locally designed public catalog and database management system developed with IBM software licensed by Brodart. It has been available to the public since 1984.
20. E.J. Coates, "Significance and Term Relationships in Compound Headings," in *Theory of Subject Analysis: A Sourcebook*, ed. by Lois Mai Chan, Phyllis A. Richmond, and Elaine Svenonius (Littleton, Colo.: Libraries Unlimited, 1985), p.193–94.
21. Carol A. Mandel includes some graphic examples showing how system design affects subject retrieval in *Multiple Thesauri in Online Library Bibliographic Systems: A Report Prepared for Library of Congress Processing Services* (Washington, D.C.: Cataloging Distribution Service, Library of Congress, 1987), p.74.
22. Lois Mai Chan, *Library of Congress Subject Headings: Principles and Application*, 2d ed. (Littleton, Colo.: Libraries Unlimited, 1986), p.352.
23. Frost and Dede, "Subject Heading Compatibility," p.297. ■■

Small College Experiences in Retrospective Conversion of Periodicals

John C. Stachacz

A retrospective conversion project for a 216-year-old small college library periodical collection requires cooperation and planning from all staff members. This paper traces the history of such a project, highlighting the cooperative efforts of several departments in its planning and implementation. Of primary focus is the development of a Serials Cataloging Manual detailing rules utilized in editing OCLC/MARC records. Examples of edited MARC records are included.

This paper will outline the steps taken to develop and implement a periodicals retrospective conversion project in a small college library that is organized collegially and committed to the notion that all librarians bear responsibility for all library functions and services. Therefore, the paper will highlight the cooperative arrangements worked out by the cataloging and serials departments for this project.

Dickinson College, founded in 1773, is a coeducational liberal arts college of approximately 1,900 students. The library has approximately 400,000 volumes accessible through an online catalog, AutoCat, created and developed at Dickinson through the cooperation of the library and the computer center.[1] AutoCat is developed along the lines of online systems such as Dialog or BRS in that author, title, and subject keyword and exact match searches are performed utilizing Boolean operators, truncation, and nesting capabilities. AutoCat, located on the college's main Digital VAX 1186 computer in the computer center, can be accessed through one of fourteen dedicated terminals in the library or through personal VAX accounts on any of 200 terminals and microcomputers distributed across the campus. On the first day of operation, in April of 1987, 99 percent of the monographs, scores, and sound recordings were accessible due to the foresight of the staff in initiating a retrospective conversion of the collection when the library began using the OCLC cataloging system in 1975. Unfortunately, periodicals and government documents were not in machine-readable form when AutoCat became available to the public.

The library subscribes to approximately 1,300 current periodical titles. In all, the collection consists of over 3,600 titles spanning the 216-year history of the college. At no time in the library's history was this collection cataloged or converted into machine-readable form. Access to the collection had been through the use of Kardex cards in the reference area. In the late 1970s an in-house computer-generated periodi-

---

John C. Stachacz is a librarian at Dickinson College, Carlisle, Pennsylvania.

The author would like to express his sincere thanks to his colleague Joan Bechtel for criticism of a draft of this article.

cals holdings list was developed and is still utilized. This list, continually updated by student workers, is produced on an irregular basis throughout the year. This list is more flexible than the Kardex system, since it can be mass-produced and distributed throughout the main library and five small science libraries located across the campus. Unfortunately, it is out-of-date the minute it is produced and lacks subject access. Of prime concern to the library staff was the conversion of all periodicals into MARC form via the OCLC cataloging subsystem. The computer staff could then load and index them in AutoCat along with records for new monographic titles that appear on the OCLC/MARC biweekly archive tapes.

The serials department purchases periodicals and monographic series, with the latter cataloged by the cataloging department. Not only did the library wish to have periodicals entered into the AutoCat database but also wished to tighten the rules and principles of cataloging monographic series. To this end, the cataloging and serials staffs began discussions to formulate a plan of action and a manual.

Most of Dickinson's periodicals are shelved alphabetically by title, although a few were classified and cataloged. As rules and staff changed over the years so did the shelving practices. In most cases titles were shelved, contrary to current cataloging rules, by latest entry rather than successive entry. Rule changes and the long history of the periodicals collection mandated that in order for the retrospective conversion to take place, a complete inventory had to occur.

The library staff is collegially organized, rotating the position of chairperson every three years from among the nine librarians. Furthermore, it is desired that all librarians have a working knowledge of several positions in order to enhance each person's personal expertise and to facilitate the rotation of all professional responsibilities when the chairperson changes. In this arrangement, each librarian is involved with reference and cataloging to some degree, with the serials librarian having some duties in three areas. In addition, each librarian acts as a liaison to four or five academic departments, responsible for collection development and bibliographic instruction.

SERIALS DEPARTMENT PREPARATIONS FOR THE PROJECT

Several preliminary projects were planned and completed before the retrospective conversion project began. First and foremost was a complete inventory of the periodical collection. The computer-generated holdings list, although current, contains only general holdings statements and is in a format dissimilar to that of the OCLC/MARC record. Students had to double-check the entries to ensure that the correct information was recorded.

Another major project was the determination of the correct OCLC record to be used. This task was made relatively easy due to the library's previous participation in the Association of College Libraries of Central Pennsylvania Union List of Serials (ACLCP Union List) located on the OCLC union list subsystem. Begun in 1986, this union list project targeted, to within 95 percent accuracy, the correct OCLC records for Dickinson's periodicals. Although the union list project did attach the Dickinson symbol to serials records, it did not add the records to the MARC tapes of the individual libraries. Therefore, Dickinson's retrospective conversion project was required both to add the records to the OCLC archive tapes and to enhance the data for loading into AutoCat.

The paper ACLCP Union List of serials records for Dickinson was purchased by the library, and under the direction of the serials librarian, student workers checked each record for its appropriateness for the particular serial and for AutoCat. The OCLC record number was then listed next to each title on a specially designated Dickinson-generated periodicals holdings list entitled, "The Master List of Converted Periodicals." In several cases, OCLC records did not exist. Students then referred to either *Ulrich's Periodicals Directory*, the *Standard Periodical Directory*, or *New Serials Titles* to determine ISSN numbers. The ISSN number of these records was added to the Master List. The ISSN number was the last resort for finding OCLC records. If a good record was not found, the title was set aside for original cataloging.

A third project was the compilation of a Key Title Index for those periodical indexes held by Dickinson. This listing consisted of the proper form of the index title and ISSN number and was used to update the 510 fields on the MARC record.

The serials and cataloging librarians together worked on the creation of a Master List of Location Codes. The goal was to standardize the form of location codes, OCLC stamps generated by $a in the 049 field, location statements currently used on the periodicals printout, and punctuation and capitalization. The resulting list indicates the form of each location and governs its use in the computer-generated holdings list, in subsequent 049 subfield a's as well as in subfield $o's in the 049 field. In all, 43 different location and format codes were developed for titles held in the main collection, branch libraries, special collections, microformat, vertical files, and storage. (See figure 1 for selected examples.) Upon completion of all four projects, the conversion of all periodicals began.

CATALOGING DEPARTMENT PREPARATIONS FOR THE PROJECT

Discussions began in 1986 between the catalogers and the serials librarian for the development of a new manual to tighten the cataloging practices of all serial publications ordered by the library. These discussions stressed the need to train the serials librarian in serials cataloging while enhancing the catalogers' ability to deal with monographic series. The manual utilized current AACR2 cataloging rules and is profusely illustrated with examples of series titles and ways in which to catalog them. Interaction and communication within the group ensured the compilation of an excellent manual.

The following four rules were established in order to ensure that the correct OCLC record was utilized:

1. The title in the subfield a of the 245 field must exactly match that of the title in question.
2. The S/L ent element in the fixed field must be 0 to ensure that the MARC record was for a successive entry.
3. The dates in the fixed field span Dickinson's holdings.
4. No X occurs in the 042 field.

Although the fixed fields are not altered, the following fields are deleted to conserve space in the AutoCat database: 012, 069, 032, 035, 050, 082, 090, 850, and 901. Only a single 092 field is permitted. The following fields are edited to reflect the correct holdings and location of Dickinson's periodical titles. The examples are primarily gleaned from four records: *Drug Enforcement* (OCLC #2243913), *Bioscience* (OCLC #1536472), *The Magazine Antiques* (OCLC #9273632 and *Black Perspective in Music* (OCLC #6801117). Examples of the fully edited MARC records are displayed in appendix A.

086—Dickinson College is a partial depository for United States Government documents. Although a retrospective con-

Master List of Location Codes

Location	*049, $a*	*Subsequent $a's and $o's*	*AutoCat Display Term*
1. Spahr Basement Periodicals	—	—	Main coll
2. Storage	—	—	Storage
3. Current Periodicals	—	—	Main coll
4. Chemical Library	DKCH	[Chem]	—
5. Microfilm or Microfiche	DKCC	[Micro]	Microfilm or Microfiche

Figure 1. Selections from the Master List of Location Codes.

version project is also being undertaken to add government documents to the library's MARC tapes, all government document periodicals are centrally controlled through the periodicals department. As such, government document periodicals are shelved alphabetically within the periodical collection. The Superintendent of Documents number in this field highlights government document titles.

Figure 2. Drug Enforcement.
>>086 $a J 24.3/2:

092—Since classification numbers are not assigned to most periodicals, an X is added to this field to conform to OCLC requirements. In AutoCat the X is translated as the location statement, "Periodicals" and appears in the Call Number location in all AutoCat displays.

Figure 3.
>>092 $a X
or, if assigned a call number:
>>092 $a 050/D537.2

049—Location and general holdings statements are indicated in this field. Codes from the Master List of Location Codes are entered along with the corresponding span of years for each individual location. Each different location and format is indicated. In all cases, holdings are rounded off to the nearest complete year. Specific holdings are indicated in the 949 field. The span of years is not added for titles located in only one location.

Figure 4.
For titles that occur in only one location. *Drug Enforcement.*
>>049 $a DKCC $o Main coll
For titles occurring in multiple locations. *Bioscience.*
>>049 $a DKCC $o Main coll $o (1964–1982); $o AV, Microfilm $o (1983–); $o Bio $o Current

260—Neither the $a or $b sections are edited. The inclusive dates for titles no longer received or a beginning date for current subscriptions are entered in subfield c.

Figure 5. Drug Enforcement.
>>260 01 $a [Washington, D.C.]: $b Drug Enforcement Administration, U.S. Dept. of Justice. $c 1974–

300—The number of volumes for titles no longer received is entered in subfield a of the 300 field. In the case of current subscriptions, the volume number is blank. The cataloging staff routinely adds and then corrects the volume number in the 300 field each time a new volume of a monographic series is received. However, the serials staff thought it impossible to continually update this field for periodicals given staff and time constraints.

Figure 6.
For open entry titles. *Drug Enforcement.*
>>300 $a v. : $b ill. (some col.) ; $c 28 cm.
For closed entries. *Magazine Antiques.*
>>300 $a 50 v. : $b ill. ; $c 31 cm.

500—The 500 fields are used for notes to clear up uncertainties for the patron. Occasionally, a title change may occur in the middle of a volume with the entire volume bound under one title. A note in this field would indicate both titles.

Figure 7. Magazine Antiques.
>>500 $a Bound as: Antiques Magazine.

510—The Key Title List is used to update the 510 fields to correspond to the index holdings of Dickinson library. Those indexes not subscribed to or unavailable through Dialog, currently the only database vendor available at Dickinson, are deleted. Databases listed in the Dialog catalog are noted (available online), signifying that these indexes can be accessed through a database search. The online service department aggressively extends this service to all faculty, administrators, and students.

Figure 8. Drug Enforcement.
>>510 1 $a Legal resource index (Available online) $a 1980–
>>510 2 $a Excerpta Medica (Available online)
>>510 0 $a Index to U.S. government periodicals $x 0098-4604

533—Information concerning titles exclusively, or primarily, in nontraditional formats is indicated in this field. Otherwise, it is deleted.

Figure 9. Black Perspective in Music.
>>533 $a Microfilm. $b Ann Arbor, Mich., $c University Microfilms International. $e reels. 35 mm.

In accordance with rules devised for the 300 field, the exact number of microforms is not indicated for open entry titles. Closed entry titles indicate the exact number of volumes in the 300 field and the exact number of pieces in the 533 field.

570—Unless the publisher has some connection to Dickinson, or is very famous, this field is deleted.

780—This field is left alone unless it is necessary to record a preceding entry that Dickinson holds and is not already listed.

785—The same is true for the 785 field, which lists succeeding entries.

949—Exact holdings, including volume, number, and year are noted in the 949 field regardless of format. The form of the entries in this field follows that detailed in the OCLC serials format for the 362 field. In AutoCat the 949 field displays immediately following the 300 field in order to provide users exact, detailed holdings in immediate proximity to the space in which they are accustomed to finding the extent of the volume or volumes associated with a particular title.

Figure 10.

For open entry. *Drug Enforcement.*
>>949 0 $ a Vol. 1, no. 3 (1974)-
For closed entry.
>>949 0 $ a Vol. 9 (1926)-v. 12 (1927)
For a closed, broken run entry.
>>949 0 $ a Vol. 5 (1882)-v. 9 (1884) ; v. 11 (1886)-v. 14 (1889) ; v. 16 (1891) ; v. 18 (1892)-v. 28 (1902)

When holdings information exceeds ten lines on the computer-generated holdings list, a note refers the user to a printed holdings list located at the reference desk. As in the case of unnecessary fields, storage capacity in AutoCat dictated the formation of this rule. Fortunately, there are relatively few of these long holdings statements.

Figure 11. Magazine Antiques.
>>949 0 $a Vol. 13/14 (1928)-v. 60/61 (1962), incomplete. **For details inquire at a Reference desk.

910—This field is created to indicate the initials of the person doing the conversion and the date it was done. In order to ensure that all information is correct, each record is scrutinized by a second person. The record is saved, and the date and file number of the record is registered on the Master List of Converted Periodicals. The second person calls up the record, records his/her initials, and presses the update key, thereby entering the record onto the OCLC/MARC archive tape; no cards are ordered. If corrections are warranted, a note is placed on the Master List noting the corrections. The original inputter then makes the changes and updates the record. Because this is a conversion process, the library has secured a retrospective conversion OCLC access number that allows the lower first-time use charge.

Figure 12.
>>910 $a 3/10/88/JCS/JB

Originally, the serials librarian input all data, with the head of cataloging double-checking the records. The serials librarian was to go on sabbatical in the year subsequent to the start of the project. To ensure a smooth rotation of department heads, a second librarian was trained to be the head of serials and to continue the project in the absence of the project initiator. Upon return of the serials librarian from sabbatical, both trained serials librarians will complete the project.

START OF THE PROJECT

The retrospective conversion project began when all preliminary planning and projects were completed. Due to the varied responsibilities of the serials librarian, and the need for the serials assistant to monitor the flow of work and student activities of the department, the progress was slower than would have occurred under a team effort involving the assistant. Nevertheless, more than 1,200 records, approximately one-third of the collection, were converted onto the MARC tapes during the first nine months of the project. Due to the continual loading of these tapes into the AutoCat database, these records are already available online.

It was also decided to start the retrospective conversion project with the titles beginning with A proceeding through the alphabet, rather than starting in the latter half of the alphabet where fewer successive entry problems occur. Although time-consuming, this method proved invaluable as many problems occurred early in the project, forcing the librarians to solve unanticipated issues immediately. The result was that the latter stages of the project moved along quite rapidly. Only the turnover in serials librarians slowed the project while the new department head learned the duties and functions of the department.

FOLLOW-UP PROJECTS

Technically, one of the finest features of AutoCat is a program called the PreLoad editor. Originally designed for the elimination of duplicates, this program, designed by the AutoCat team, allows for the quick correction of records before they are loaded and indexed into AutoCat. After loading, the AutoCat editor provides full screen editing and re-indexing. Editing in Preload or AutoCat itself can be done from VAX terminals or microcomputers all across campus, or even from equipment in remote locations. When a cessation or title change occurs, the record can be retrieved, edited, and reloaded very quickly. Currently, the serials librarian handles all editing of AutoCat records. Student workers will be trained to update the records when the project reaches the halfway point. The serials librarian will continue to add new titles to the OCLC/MARC records.

Another task is the education of faculty and students to the existence of the periodical records in AutoCat. Although AutoCat can be self-taught, incoming freshmen and new students are required to attend an established series of workshops on the use of AutoCat. Upper-class students and faculty can either attend one of a multitude of workshops or arrange a tutorial with a librarian.[2] The availability of periodical titles is mentioned, and examples shown of how to access them. Periodicals are relatively easy to access, as the procedures are identical to those of finding monographs. The main difference is the interpretation of the record.

While the retrospective conversion project is being undertaken, students and faculty are taught to refer to AutoCat and the regular holdings list for citations. Once the project is completed, the paper-generated list will be discontinued.

The last remaining project will be the shifting of the periodical collection to match the latest title entry of AutoCat. The librarians decided to wait until the completion of the project to shift the collection in order to avoid multiple shifts. Notes to patrons are placed in the 500 fields of records to warn them about such problems. These records are noted in a file and will be deleted when the shift occurs. The shift will take place during the summer when use of the library is at a minimum and plenty of student hours are available. Microforms are spliced and relabeled when titles are converted, since they are few in number and easy to handle.

CONCLUSION

The retrospective conversion project accomplished several goals. First, all periodicals records were converted into machine-readable form via the OCLC/MARC tapes. Second, the retrospective conversion was accomplished through the cooperation and enrichment of several departments. Third, the serials staff have become trained in serials cataloging, while catalogers have improved their expertise in monographic series and periodical cataloging. Fourth, the periodicals collection is now accessible though the college's online catalog by author, title, and subject. Students searching for materials on a topic now routinely find—along with books—periodicals that carry articles on the subject. Fifth, a much-needed, complete inventory of the periodicals collection is being accomplished. And, finally, the new technologies have enabled both enhanced cooperative efforts in the library and the capability for accurate and instantaneous updating of periodical holdings.

REFERENCES AND NOTES

1. Joan Bechtel, "Libraries and Computing Centers Working Together: A Success Story," *Libraries & Computing Centers*, N.S., 7 (March 1988) located in *The Journal of Academic Librarianship*, 14, no.1 (March 1988).
2. Joan Bechtel, "Developing and Using the On-Line Catalog to Teach Critical Thinking," *Information Technology and Libraries* 7, no.1: 30–40 (1988). This article contains a detailed description of the use of AutoCat as a tool for teaching critical teaching. Examples are included.

APPENDIX A.

I. FULLY EDITED MARC RECORD FOR THE JOURNAL *DRUG ENFORCEMENT*

```
1. Status: Approved
»»
Control: 2243913    Rec stat: c     Entrd: 760127     Used: 890227
Type: a             Bib lvl: s      Govt pub: f       Lang: eng
Source:             S/L ent: 0      Repr:             Enc lvl:
Conf pub: 0         Ctry: dcu       Ser tp: p         Alphabt: a
Indx: u             Mod rec:        Phys med:         Cont:
Frequn:             Pub st: c       Desc: a           Cum ind: u
Titl pag: u         Isds: 1         Regulr: x         Dates: 1973,9999

»010      ¤a 75647777 //r81
»040      ¤a DLC ¤c DLC ¤d NSD ¤d DLC ¤d m.c. ¤d GPO ¤d NST ¤d GPO ¤d RCS ¤d GPO
¤d HUL ¤d AIP ¤d OCL ¤d NST ¤d AIP ¤d DLC ¤d MUL ¤d NSD ¤d OCL ¤d m/c ¤d GPO
¤d DKC
»019      ¤a 1639247 ¤a 1233335 ¤a 1077343
»022 0    ¤a 0098-3470
»042      ¤a lc ¤a nsdp
»043      ¤a n-us---
»060 0    ¤a W1 DR521D
»086      ¤a J 24.3/2:
»092      ¤a X
»049      ¤a DKCC ¤o Main coll
»210 0    ¤a Drug enforc.
»222 00   ¤a Drug enforcement
»245 00   ¤a Drug enforcement.
»260 01   ¤a [Washington, D.C.] : ¤b Drug Enforcement Administration, U.S. Dept.
of Justice, ¤c 1974-
»265      ¤a Supt. of Docs., U.S. Govt. Print. Off., Washington, D.C. 20402
»300      ¤a v. : ¤b ill. (some col.) ; ¤c 28 cm.
»310      ¤a Irregular
»362 0    ¤a Vol. 1, no. 1 (fall 1973)-
»500      ¤a Title from cover.
»510 1    ¤a Legal resource index (Available online) ¤b 1980-
»510 2    ¤a Excerpta medica (Available online)
»510 0    ¤a Index to U.S. government periodicals ¤x 0098-4604
»515      ¤a Suspended after vol. 12, no. 1 (summer 1985).
»650  0   ¤a Narcotics, Control of ¤z United States ¤x Periodicals
»650  0   ¤a Narcotics, Control of ¤x International cooperation ¤x Periodicals
»650  0   ¤a Drugs ¤x Periodicals
»710 10   ¤a United States. ¤b Drug Enforcement Administration
»780 00   ¤t BNDD bulletin ¤x 0049-5468 ¤w (DLC)***75647641
»949 0    ¤a Vol. 1, no. 3 (1974)-
»910      ¤a 3/10/88/jcs/jb
««
```

II. FULLY EDITED MARC RECORD FOR THE JOURNAL *BIOSCIENCE*

```
1. Status: Approved
»»
Control: 1536472    Rec stat: c      Entrd: 750812     Used: 890228
Type: a             Bib lvl: s       Govt pub:         Lang: eng
Source: d           S/L ent: 0       Repr:             Enc lvl:
Conf pub: 0         Ctry: dcu        Ser tp: p         Alphabt: a
Indx: u             Mod rec:         Phys med:         Cont:
Frequn: m           Pub st: c        Desc: a           Cum ind: u
Titl pag: u         Isds: 1          Regulr: r         Dates: 1964,9999

»010     ǂa 82643645 ǂz sn 78005766
»040     ǂa MUL ǂc MUL ǂd NSD ǂd OCL ǂd UCU ǂd COO ǂd NSD ǂd YUS ǂd DLC ǂd IUL
ǂd HUL ǂd m.c. ǂd NSD ǂd DLC ǂd NST ǂd DLC ǂd SER ǂd RCS ǂd AIP ǂd NSD ǂd AIP
ǂd OCL ǂd m/c ǂd DKC
»019     ǂa 2083162
»022 0   ǂa 0006-3568
»030     ǂa BISNAS
»042     ǂa lc ǂa nsdp
»060 0   ǂa W1 BI913
»092     ǂa X
»049     ǂa DKCC ǂo Main coll ǂo (1964-1982) ; ǂo AV, Microfilm ǂo (1983-) ;
ǂo Bio ǂo Current
»210 0   ǂa Bioscience
»222 00  ǂa Bioscience
»245 00  ǂa BioScience.
»260 01  ǂa [Washington, D.C.] : ǂb American Institute of Biological Sciences,
ǂc 1964-
»265     ǂa American Institute of Biological Sciences, 1401 Wilson Blvd.,
Arlington, VA 22209
»300     ǂa v. : ǂb ill. ; ǂc 28 cm.
»310     ǂa Monthly
»350     ǂa $43.00 (institution, domestic) ǂa $47.00 (institution, foreign)
ǂa $32.50 (individual, membership)
»362 0   ǂa [Vol. 14, no. 1] (Jan. 1964)-
»500     ǂa Title from cover.
»510 1   ǂa General science index ǂx 0162-1963
»510 1   ǂa Magazine index (Available online) ǂb 1964-
»510 1   ǂa Readers' guide to periodical literature ǂx 0034-0464
»510 2   ǂa Biological abstracts (Available online) ǂx 0006-3169
»510 2   ǂa Chemical abstracts ǂx 0009-2258
»510 2   ǂa Current index to journals in education ǂx 0011-3565
»510 2   ǂa Energy information abstracts (Available online) ǂx 0147-6521
»510 2   ǂa Energy research abstracts ǂx 0160-3604
»510 2   ǂa Environment abstracts (Available online) ǂx 0093-3287
»510 2   ǂa Excerpta medica (Available online)
»510 2   ǂa GeoRef (Available online) ǂx 0197-7482
»510 2   ǂa Life sciences collection (Available online)
»550 0   ǂa "Official publication of the American Institute of Biological
Sciences."
»650   0 ǂa Biology ǂx Periodicals
»650   0 ǂa Biology ǂx Periodicals
»710 20  ǂa American Institute of Biological Sciences
»780 00  ǂt A.I.B.S. bulletin ǂx 0096-7645 ǂw (OCoLC)1460437
»936     ǂa Unknown ǂa Mar. 1982
»949 0   ǂa Vol. 14 (1964)-
»910     ǂa 2/3/88/jcs/jb
««
```

III. FULLY EDITED MARC RECORD FOR THE JOURNAL *THE MAGAZINE ANTIQUES*

```
1. Status: Approved
»»
Control: 9273632   Rec stat: c     Entrd: 830303    Used: 871220
Type: a            Bib lvl: s      Govt pub:        Lang: eng
Source: d          S/L ent: 0      Repr:            Enc lvl:
Conf pub: 0        Ctry: nyu       Ser tp: p        Alphabt: a
Indx: u            Mod rec:        Phys med:        Cont:
Frequn: m          Pub st: d       Desc: a          Cum ind: 1
Titl pag: u        Isds:           Regulr: r        Dates: 1928,1952

»010      ǂa sn 86025617
»040      ǂa AZS ǂc AZS ǂd NSD ǂd IUL ǂd DKC
»042      ǂa lcd
»092      ǂa X
»049      ǂa DKCC ǂo 2nd Floor Art ǂo (1928-1937), incomplete ; ǂo Room 107
ǂo (1935-1947), incomplete ; ǂo 2nd Floor Art ǂo (1948-1952)
»130 00 ǂa Magazine antiques (New York, N.Y. : 1928)
»245 04 ǂa The magazine antiques.
»260 00 ǂa [New York : ǂb Editorial Publications, ǂc 1928-52.
»300      ǂa 50 v. : ǂb ill. ; ǂc 31 cm.
»310      ǂa Monthly
»362 0  ǂa -v. 62, no. 1 (July 1952).
»362 1  ǂa Began with v. 13, no. 1 in Jan. 1928.
»500      ǂa Bound as: Antiques magazine.
»500      ǂa Description based on: Vol. 52, no. 1 (July 1942); title from cover.
»500      ǂa Imprint varies.
»555      ǂa Vols. 1 (1922)-60 (1951). (Includes also index to former title) 1 v.;
v. 61 (1952)-70 (1956). (Includes also index to later title) 1 v.
»650   0 ǂa Antiques ǂx Periodicals
»650   0 ǂa Collectors and collecting ǂx Periodicals
»650   0 ǂa Art ǂx Periodicals
»780 00 ǂt Antiques (Boston, Mass.)
»785 00 ǂt Antiques (New York, N.Y. : 1952) ǂx 0003-5939 ǂw (OCoLC)1481615
»936      ǂa July 1952
»949 0  ǂa Vol. 13/14 (1928)-v. 60/61 (1962), incomplete.**For details inquire
at a Reference desk.
»910      ǂa 11/13/87/jcs/jb
««
```

IV. FULLY EDITED MARC RECORD FOR THE JOURNAL *BLACK PERSPECTIVE IN MUSIC*

```
1. Status: Approved
»»
Control: o801117   Rec stat: c     Entrd: 801008    Used: 880205
Type: a            Bib lvl: s      Govt pub:        Lang: eng
Source: d          S/L ent: 0      Repr: a          Enc lvl:
Conf pub: 0        Ctry: miu       Ser tp:          Alphabt: a
Indx: u            Mod rec:        Phys med:        Cont:
Frequn: u          Pub st: c       Desc:            Cum ind: u
Titl pag: u        Isds:           Regulr: u        Dates: 1973,9999

»010      ǂa sf 86091875
»040      ǂa KEU ǂc KEU ǂd DLC ǂd NSD ǂd DLC ǂd DKC
»022      ǂa 0090-7790
»042      ǂa msc
»092      ǂa X
»049      ǂa DKCC ǂo AV, Microfilm ǂo (1973- ) ; ǂo Main coll ǂo Current year
»222 00 ǂa Black perspective in music
»245 04 ǂa The Black perspective in music.
»260 01 ǂa [Cambria Heights, N.Y., ǂb Foundation for Research in the
Afro-American Creative Arts] ǂc 1973-
»300      ǂa v. ǂb illus. ǂc 24 cm.
»362 0  ǂa v. 1-***spring 1973-
»510 0  ǂa RILM abstracts ǂx 033-6955
»510 0  ǂa Music article guide ǂx 0027-4240
»510 0  ǂa Music index
»533      ǂa Microfilm. ǂb Ann Arbor, Mich., ǂc University Microfilms
International. ǂe reels.  35 mm.
»650   0 ǂa Afro-Americans music ǂx History and criticism
»650   0 ǂa Music ǂz Africa ǂx History and criticism
»710 20 ǂa Foundation for Research in the Afro-American Creative Arts
»890      ǂa Black perspective in music.
»949 0  ǂa Vol. 1 (1973)-
»910      ǂa 2/3/88/jcs/jb
««
```

Communications

Beyond Find: Boolean Searching with HyperCard

Paul A. Carnahan

Boolean searching. The phrase warms the hearts of librarians everywhere. With the advent of commercial online databases, inhouse online catalogs, and now CD-ROM databases, librarians have come to expect powerful Boolean search capabilities from all automated databases. It is no surprise then that one of the first questions most librarians ask about HyperCard, Apple's hot new "information management system" for the Macintosh, is: "Does it do Boolean?" The answer to this question is a qualified yes. A little creative programming with HyperTalk, the programming language built into HyperCard, can give the program respectable Boolean capabilities.

HyperCard is relatively easy to use and hence is widely accessible to librarians who have an interest in making computers perform specialized tasks but who lack formal training in computer programming. The time spent learning to use HyperCard and HyperTalk will be well worth the investment because HyperCard offers two distinct advantages over traditional database programs: the user interface is terrific, and the program can be easily altered to meet in-house requirements. All of this, plus the purchase price (free with the purchase of a Macintosh or $42 when purchased separately), makes HyperCard an ideal tool for small databases designed to be used by untrained staff members or patrons.

This article will explain how to create a Boolean search routine using HyperCard. By following these step-by-step instructions you will be able to create a shell for a database that even the most computer-shy patron can operate successfully. I will explain in detail the HyperTalk scripts necessary for Boolean searching, but I will assume that the reader is familiar with basic HyperCard operations such as how to create a button, field, background, and simple script. There are numerous articles in library journals, computer magazines, and books that are available to readers who need an introduction to HyperCard.[1] Although there have been articles published on searching with HyperCard, none has shown how to use these techniques to create a database for use in a library setting.[2]

Paul A. Carnahan is Electronic Reference Services Librarian at Brandeis University, Waltham, Massachusetts.

THE FIND COMMAND

The Find command is the heart of HyperCard's searching capability. Find, which Apple has generously provided as a pulldown menu command, works much like the Find feature found on most word processors: HyperCard goes to the first occurrence of the search term and highlights it with a box. With the press of the return key, HyperCard goes to the next occurrence of the term, either on the same card or on another card. This process continues in an endless circle as long as the user keeps pressing the return key. Find can also be invoked from a HyperTalk script. Find in a script behaves just like Find from the menu, except that the user cannot go to the next occurrence of the term by pressing return.

The standard Find command includes an implied Boolean AND. If the user includes two words in the search string, HyperCard will go to the first card which contains both words, in any order, in any field or fields. HyperCard puts a box around the first found word from the search string, but it does not indicate where the other word is located. The "automatic AND feature" (my term, not Apple's) can be overridden with the "find whole" command introduced in version 1.2 of HyperCard. This command enables the user to search for phrases that include spaces.

The basic Find command searches for character strings at the beginning of words; it does not find characters embedded in words and does not limit itself to whole words. This has the effect of "automatic truncation": type in the root of a word, and HyperCard will find all variations. If truncation is not desired, additional commands can be added to the Find command: "find characters" (or the abbreviation "chars") will search for character strings anywhere in a word; "find word" will limit the search to whole words. A search can be restricted to one field by including the phrase "in field"; i.e., find [target] in field [field name].

Although the Find command can be invoked from the menu, most end-user HyperCard stacks do not use the standard menu bar but instead rely on developer-programmed buttons. These buttons allow the developer to limit a user's options to those that are relevant at a given time. They also allow the end user to perform complex operations with the simple click of a button. The remainder of this article is devoted to using the Find command in complex scripts that perform Boolean searches.

THE SEARCH CARD

My approach to Boolean searching with HyperCard requires the developer to create a search card on which the end user will formulate his/her search statement. This card must be in a separate stack from the database to be searched, preventing HyperCard from finding the term on the search card itself and permitting the end user to choose among several databases from one search card if the developer so desires.

Therefore, the first step in creating a HyperCard Boolean search stack is to design the search card. I have used a background design that suggests a stack of cards (see figure 1), but the developer can choose any appropriate design. The cards in the database stack can also be of any design. They do not have to be of the same design as the search card, but an integrated appearance is desirable.

After the background of the search card

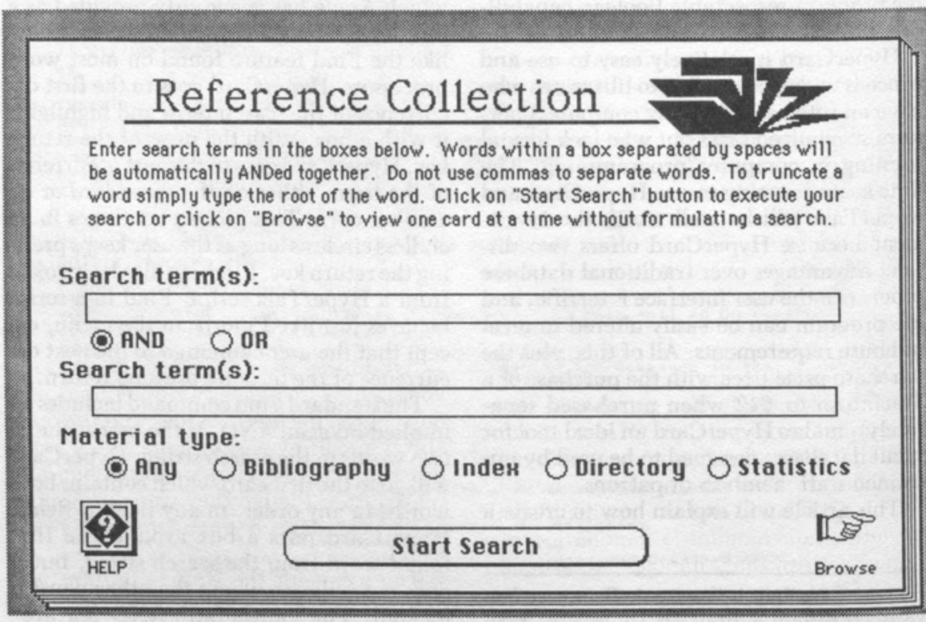


Figure 1. Search Card.

has been loosely designed (it can always be changed later), create six fields on the card. The first two fields should be long, narrow fields like the ones in figure 1; name them firstTerms and secondTerms, and be sure that they are card fields one and two, respectively. The next two card fields should be small and near the bottom of the card. Name them group1 and group2. These are going to be containers for information the user does not need to see, so hide the first by typing in the message box "Hide card field group1" and pressing return; then do the same for card field group2. The final card field should be a large field across the top of the card that will contain instructions to the end user. The text in this field should be "locked" via the "About Field" dialog box after the instructions have been entered; in this way they cannot be changed by users. If more detailed instructions are needed than will fit in this field, a help button can be created that shows a bigger field or that goes to cards that explain the operation of the system in detail.

RADIO BUTTONS

The next step is to create several sets of buttons. Create two radio buttons, name them AND and OR, and position them as in the illustration. Radio buttons in the Macintosh environment are generally used to present a user with mutually exclusive choices, such as in the standard print dialog box. Clicking on a radio button should turn it on and turn off all of the other radio buttons in its group. Unfortunately, the standard HyperCard radio button is a "dumb" button just like any other HyperCard button; clicking on one will not turn off the others. Therefore, the developer has to do a little programming to make HyperCard radio buttons behave like true Macintosh radio buttons. There are almost as many solutions to the radio button problem as there are HyperCard developers. I have used one here that I like (it is not original with me), but there are others that will work as well.[3]

The AND and OR buttons should both have the same handler (a handler is a section of a script that "handles" a particular HyperCard action such as mouseUp, closeCard, openStack, etc.):

```
on mouseUp
  doGroup1
end mouseUp
```

The line "doGroup1" calls a handler that you should enter into the card script:

```
on doGroup1
  get card field group1
  set the hilite of card button it to false
  put the short name of the target
    into card field group1
  set the hilite of the target to true
end doGroup1
```

Once you have done this and *before* you click on your new radio buttons, type in the message box: put "and" into card field group1; then press return. Now click on the AND button; it will turn on. Click on the OR button, and voila, it will turn on and the AND button will turn off.

You can also create a second set of radio buttons to help the end-user limit a search to a specific group of terms. In this example I have created a set of five buttons relating to reference book material types. Each button contains the handler

```
on mouseUp
  doGroup2
end mouseUp
```

If you use a second set of radio buttons, you should create a handler in the card script for doGroup2 just like the one for doGroup1 except with the names changed. Be sure to enter the name of one of your group2 buttons into the hidden field group2 before you click on one of these buttons, or you will get an error message. As you will discover later, it is important that the names of these buttons be the exact terms that you use in a field in your database stack.

BROWSE BUTTON

In addition to two sets of radio buttons, you should also create two other non-radio buttons. Name one "Search" and the other "Browse." Browse is for users who want to flip through the database cards one at a time, just as though they were browsing a card file; Search performs a Boolean search.

The script for the Browse button takes the user to the first card of the database stack, called "Reference Books" in this example:

```
on mouseUp
  global hitList
  set the cursor to 4
  lock screen
  push card
  put empty into hitList
  go to stack "Reference Books"
  show background button "next"
  show background button "previous"
  unlock screen with iris open
end mouseUp
```

In this script I have used HyperCard's "lock screen" command to freeze the screen action so that two hidden buttons can be displayed without the user seeing them flash on the screen. Your database stack will, of course, have to have two buttons called "Next" and "Previous" (see the section "Database Stack"). I have also used the command "push card." This sets aside the current card so that it can be easily returned to later with the "pop card" command. This script contains a global variable called "hitList." This variable is an essential part of the Boolean search routine, so when the user is just browsing the variable should be empty. Be sure to use the name of your stack in this script. If you want to get fancy, you can give the user a choice of stacks to search. You should be able to figure out how to do this with radio buttons and a hidden field after you have seen the script for the "Search" button below.

BOOLEAN SEARCH BUTTON

The script for the button named Search gives this card its Boolean search capabilities. As a result the script is rather complex (see figure 2). A large portion of the script is devoted to manipulating the information in the search request, massaging the search result, and navigating to the first found item. The Boolean portion of the script is contained in the "doSearch" handler, which appears at the end of the Search button's script.

The first three lines of the script take care of housekeeping details: identifying global variables (or containers) and setting the cursor to the stopwatch icon so that the end user knows that the computer is still working while this script is running. The first section of the script checks to see that the user has actually entered a search term in the first search field, firstTerms. If it is empty then the Macintosh displays a dialog box with a polite error message. If there are characters in the first field, the screen is locked and the script proceeds.

The second section of the script simply makes sure that certain global variables are "cleaned out" and do not contain any information left over from a previous search.

The third section sets the stage for a Boolean AND or OR search. First, searchCounter is given the value 2 if the OR button is on or the value 1 if it is not on. Then the information from the two search fields is put in the HyperTalk container (variable) searchHolder. If the search is an OR search, the search string is placed on two lines in the container; if it is an AND search, it is placed on only one line. Note that this script works with only two terms or phrases ORed together.

The third section of the script also takes the information from the hidden field group2 (which in this example contains material type information) and places it in the search string. HyperCard will be searching for the occurrence of the names of the group2 buttons in the database, so it is important that the names of these buttons correspond to information in the database. Notice that if the user has clicked on the "any" button, nothing is put in the search string.

Once the search string has been placed in searchHolder, the script pushes the current card and goes to the stack to be searched, in this case a stack called "Reference Books." Once in this stack the script performs a search using the handler doSearch. This handler could be integrated into the mouseUp handler (without the labels "on doSearch" and "end doSearch"), but I wrote it in a separate handler to emphasize its importance and to make it "transportable." This handler could reside at any HyperCard level so that it could be called by more than this one button script.

Essentially, the doSearch handler passes through the database stack and assembles a list of the ID numbers of all of the cards that contain the search string; I call the list hitList. DoSearch passes through the stack once if the user has requested an AND

```
on mouseUp
  global hitList,counter,totalMatches,viewCard,searchString
  global modifySearch,searchHolder,searchCounter
  set the cursor to 4

  -- Check to see that user has entered a search --
  if card field firstTerms is empty then
    beep
    answer "Please enter a search term"
    send tabKey
    exit mouseUp
  end if

  lock screen

  -- Initialize some variables --
  put empty into hitList
  put empty into searchHolder
  put empty into counter
  put empty into totalMatches
  put empty into viewCard

  -- Setup OR or AND search --
  if the hilite of card button "OR" is true then
    put 2 into searchCounter
    repeat with x = 1 to 2
      put card field x & space into line x of searchHolder
      if the hilite of card button "any" is false then
        put card field group2 after line x of searchHolder
      end if
    end repeat
  else
    put 1 into searchCounter
    put card field 1 && card field 2 & space into searchHolder
    if the hilite of card button "any" is false then
      put card field group2 after line x of searchHolder
    end if
  end if

  push card
  go to stack "Reference Books"

  -- "doSearch" is a handler in this button's script (see below)--
  doSearch

  -- Determine if there are any records found --
  put the number of words in hitList into totalMatches
  if totalMatches = 0 then
    put true into modifySearch
    pop card
    unlock screen
    beep
    answer "Search term(s) not found."
    exit mouseUp
  end if
```

Figure 2. Script for Button Start Search.

```
    -- "nextMatch" is a handler in the database stack "SciRef" --
    nextMatch

    -- show the buttons necessary for naviagation & show the first card --
    show background button "next match"
    show background button "previous match"
    show background button "browse"
    unlock screen with iris open
    set the cursor to 1
  end mouseUp

  on doSearch
    global hitList,searchHolder,searchCounter,tempList
    repeat with x = 1 to searchCounter
      put empty into tempList
      put line x of searchHolder into searchString
      repeat
        find searchString
        if the result is empty then
          put the short id of this card into thisCard
          if thisCard is in tempList then
            exit repeat
          else
            put thisCard & space after tempList
            if hitList contains thisCard is false then
              put thisCard & space after hitList
            end if
            go next
          end if
        else
          exit repeat
        end if
      end repeat
    end repeat
  end doSearch
```

Figure 2. Cont.

search (remember it is easy for HyperCard to perform AND searches) and twice if it is an OR search. If no matches are found, the doSearch handler exits back to the mouseUp handler. If at least one match is found, the handler proceeds. (The line "if the result is empty then" does *not* contain a typographical error: if HyperCard does not find what it is looking for, it creates a message to that effect, and "the result" is *not* empty; if HyperCard finds what it is looking for, it goes to that card, and the result is empty.) HyperCard stacks are circular, so doSearch has been written to end a given pass when the ID of the first card is found for a second time. The container tempList is used for this purpose. TempList and hitList will contain the same values during an AND search, but during the second pass of an OR search, tempList will contain only the IDs of the cards that were found during that pass.

After doSearch has conducted its search, control moves back to the mouseUp handler. If the search is unsuccessful, the Macintosh is instructed to display a message telling the user that the search terms were not found in the database stack. If the search criteria were met, then the computer goes to the first card identified in hitList. It does this through the handler nextMatch, which is in the *database stack*. (Remember that the computer is now in the database stack, even though the user still sees the search card "frozen" on the screen.) NextMatch will be described in the database section following. Before the script ends, it displays the buttons the user will need to navigate through the list of matches and then unlocks the screen with a visual effect.

Once the Search button is complete there remain only a few more simple scripts to enter into the search card. The card script needs a handler to (1) set the userLevel so that users can enter information into the search fields without being able to alter anything else in the stack and (2) to hide the standard HyperCard menu bar:

```
on openCard
  set the userLevel to 2
  hide menubar
end openCard
```

The card script should also have a handler to disable the keyboard's arrow keys:

```
on arrowKey
  beep
end arrowKey
```

This script "captures" the use of the arrow keys and does nothing except beep. This will prevent users from using the arrow keys to move from one card to another in your stack, circumventing the execution of your scripts and causing general confusion.

Finally, you may want to include handlers in the card script that "capture" the return and enter keys and make them behave like the "start search" button:

```
on returnKey
  click at the loc of card button "start search"
end returnKey
on enterKey
  click at the loc of card button "start search"
end enterKey
```

This may be desirable if you find users using these two keys like an enter key to initiate action. However, if you decide to use these handlers, *do not* create them until you have completed your stacks and are ready to put them out for public use, because if these keys are captured, the use of the message box is affected, which in turn severely limits your ability to change the userLevel so that you can make changes to your stack. If you need to edit the script of a card that you have made inaccessible in part by capturing the enter and return keys, you can use the handy HyperTalk "edit" command. Go to a different card or stack and set the userLevel to 5. Then type into the message box: edit the script of [object x]. The edit box for that card, stack, or background will appear, and you can make the necessary changes.

DATABASE STACK

The database stack is conceptually simple compared with the search card, but it contains a large number of navigational scripts. The background of the database stack, just like that for the search card, can be of any appropriate design (see figures 3 and 4). It can also have as many fields as are needed; they can be of any length and style, but they must be *background fields.*

In addition to at least one background field, the database stack must contain certain buttons, fields, and scripts in order to work properly with the search card I described above. There must be one background field called "matches monitor." This field will display the statement "Viewing match [x] of [y]," so it should be placed in a prominent position; I have placed it in the center of the card near the bottom (see figure 4). This field will be empty when the user is browsing through the stack, so its border should be transparent.

Create four background buttons. Name them "Next," "Previous," "Next Match," and "Previous Match." These buttons should have icons that are arrows or pointing hands to communicate their use to the user (see figures 3 and 4). Only one pair of these buttons will be visible at any one time, so they can be placed over each other in the background.

The scripts for the pair that is visible when the user is browsing through the stack (Next and Previous) are simple, so create them first. The script for Next is

```
on mouseUp
  visual effect scroll left
  go to next card
end mouseUp
```

Similarly, the script for Previous is

```
on mouseUp
  visual effect scroll right
  go to previous card
end mouseUp
```

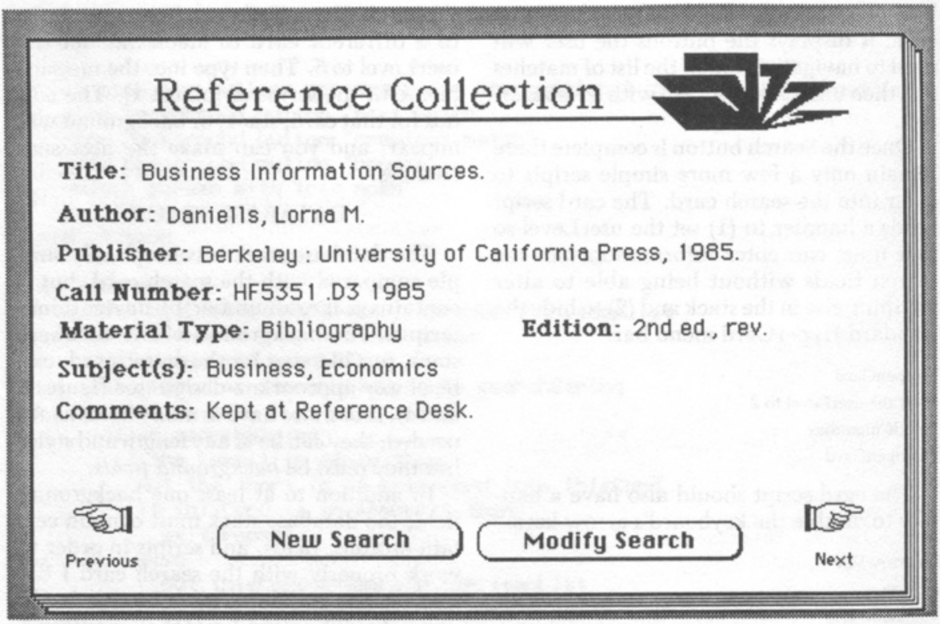


Figure 3. Database Stack in Browse Mode.

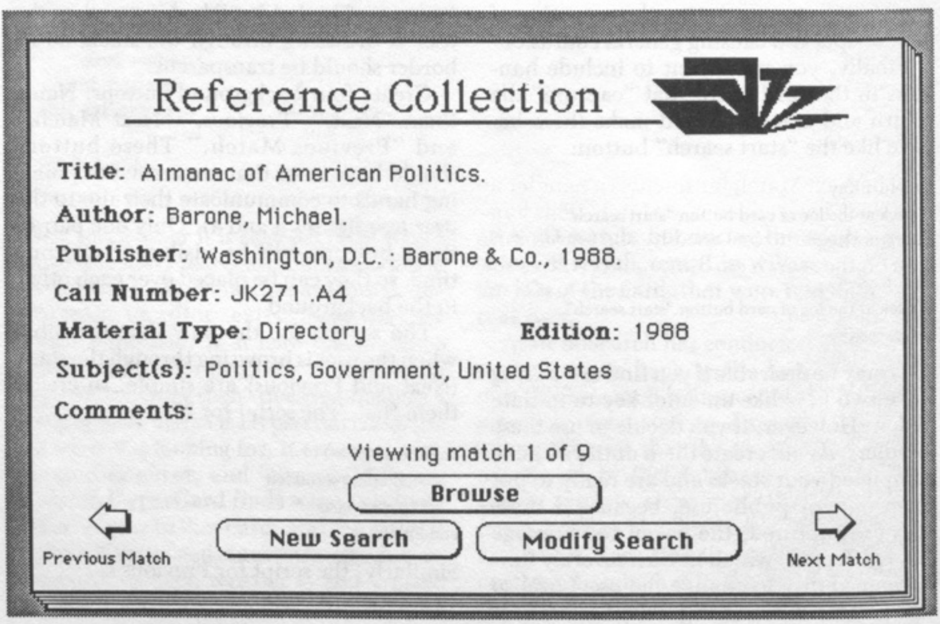


Figure 4. Database Stack after a Boolean Search.

The scripts for Next Match and Previous Match are a little more complex. They use the global containers hitList, counter, and totalMatches that were filled during the Boolean search. The script for the button Next Match should be as follows:

```
on mouseUp
  global counter, hitList, totalMatches, viewCard
  if counter < totalMatches then
    nextMatch—this is a handler at the stack level
  else
    beep
    answer "No more matches"
  end if
end mouseUp
```

The script for Previous Match should look like this:

```
on mouseUp
  global counter, hitList, totalMatches, viewCard
  if counter > 1 then
    subtract 1 from counter
    put word counter of hitList into viewCard
    put "Viewing match" && counter && "of" &&
    totalMatches into background field "matches
    monitor" of card id viewCard
    visual effect scroll right
    go to card id viewCard
  else
    beep
    answer "No previous matches"
  end if
end mouseUp
```

The Next Match button uses a handler at the stack level called nextMatch. This handler is the same one used by the search button on the search card when it retrieves the first match. Enter the following script in the stack script:

```
on nextMatch
  global counter, hitList, totalMatches, viewCard
  add 1 to counter
  put word counter of hitList into viewCard
  put "Viewing match" && counter && "of"
  && totalMatches into background field
    "matches monitor" of card id viewCard
  visual effect scroll left
  go to card id viewCard
end nextMatch
```

Basically what these three handlers do is use the container "counter" to keep track of where the viewer is in the list of matches, which is kept in another container called "hitList." The variable "totalMatches" is set by the search button ("put the number of words in hitList into totalMatches") and remains constant. It keeps track of the last word in the list of matches (a "word" in HyperTalk is anything set off by spaces).

Create two more buttons in the background of the database stack. Name one "Modify Search" and the other "New Search." Both of these buttons will take the user back to the search card; the former will leave the previous search in place, while the latter will discard it before the search card is shown. The script for Modify Search should be as follows:

```
on mouseUp
  global hitList, modifySearch
  —the following line makes the button behave like
  "new search" if there really is no old search to
  modify if hitList is empty then put false into
  modifySearch else put true into modifySearch
  goSearch—a handler at the stack level
end mouseUp
```

The script for New Search should look like this:

```
on mouseUp
  global modifySearch
  put false into modifySearch
  goSearch—a handler at the stack level
end mouseUp
```

Both of these buttons use a handler in the stack script:

```
on goSearch
  set the cursor to 4
  global modifySearch
  lock screen
  pop card
  if modifySearch is false then
    click at loc of card button "and"
    click at loc of card button "any"
    put empty into card field firstTerms
    put empty into card field secondTerms
  end if
  unlock screen with visual effect iris close
  send tabKey
end goSearch
```

This handler "pops" the card that was set aside by the "push" command as the user left the search card using either the Browse or Search button. If the user wants to start a new search, this handler clears the two search fields and sets up the search card by clicking on the "and" and "any" card buttons to establish two default values. The goSearch handler also "presses" the tab key once so that the cursor is placed in the first search field (if the user wants to modify the search, the contents of the first search field are highlighted).

Create one more pair of background buttons; name them "Browse" and "Show Matches." They can be placed on top of each other because they will not be displayed simultaneously (see figures 3 and 4). Browse will permit the end user to browse the stack card-by-card after conducting a search, while Show Matches will permit the user to return to viewing the list of matches. Neither of these two buttons will be visible when the user enters the database stack with the Browse button.

The script for the Browse button in the database stack follows (this is *not* the same as the script for the Browse button on the search card):

```
on mouseUp
  Answer "Browse starting with" with
  "First Card" or "This Card" or "Cancel"
  if it is "first card" then
    lock screen
    go to card 1
    browse
  end if
  if it is "this Card" then
    lock screen
    browse
  else
    beep
    exit mouseUp
  end if
end mouseUp
on browse
  global hitList
  set the cursor to 4
  hide background button "next match"
  hide background button "previous match"
  hide background button "browse"
  show background button "next"
  show background button "previous"
  show background button "show matches"
  put empty into bkgnd field "matches monitor"
  unlock screen with iris open to gray
end browse
```

This button displays a dialog box that gives the user the option of beginning to browse with the first card in the stack, or the current card, or of cancelling the action. The handler "browse" simply shows the proper navigational buttons for browsing through the stack along with replacing the Browse button with the Show Matches button.

The script for the background button Show Matches should look like this:

```
on mouseUp
  set the cursor to 4
  lock screen
  global counter,hitList,totalMatches, viewCard
  put word counter of hitList into viewCard
  go to card id viewCard
  put "Viewing match" && counter && "of" &&
  totalMatches into background field "matches
  monitor"
  show background button "next match"
  show background button "previous match"
  show background button "browse"
  hide background button "next"
  hide background button "previous"
  hide background button "show matches"
  unlock screen with visual effect
  iris open to gray
end mouseUp
```

This button reverses the action of the Browse button and returns the user to the current card in the hitList. The only time that this button will be visible is after the user has activated the Browse button.

There should be several handlers in the stack script in addition to the handlers already discussed. The first should be the following:

```
on closeCard
  put empty into bkgnd field "matches monitor"
end closeCard
```

This handler ensures that the background field "matches monitor" does not have any text in it after a card closes, regardless of whether the user is looking through a list of matches, exiting the stack, or browsing.

The stack script should also include the following handler, which is executed whenever a user leaves the stack:

```
on closeStack
```

```
hide background button "next"
hide background button "previous"
hide background button "next match"
hide background button "previous match"
hide background button "browse"
hide background button "show matches"
end closeStack
```

This handler hides all but two of the background buttons, preparing the database stack to be opened by one of the two buttons in the search stack that will reveal the appropriate set of navigational buttons.

The database stack, like the search card, should also have a handler to prevent unwary users (especially DOS users who like to use arrow keys instead of electronic mice!) from using the arrow keys and ignoring your carefully crafted scripts:

```
on arrowKey
  beep
end arrowKey
```

You may not want to add this handler until the data for your database have been entered into the stack. Instead of telling the computer to beep when the arrow buttons are pressed, you could write a script that clicks on the Next or Previous button, depending on which arrow was pressed; but be careful because these two buttons are sometimes replaced by the Next Match and Previous Match buttons.

You can also add a short handler to aid in the input of data into this database:

```
on newCard
  send tabKey
end newCard
```

This will place the cursor in the first field of a new card so that the person inputting data can start typing immediately. If you trap the return key and give it the new card command, the typist will not have to remove his or her hands from the keyboard at all:

```
on returnKey
  doMenu "new card"
end returnKey
```

Finally, after all of the data have been entered, you should add a handler to the stack script that makes sure that the menu bar is hidden and puts users in HyperCard's read-only mode (called "browse"), so that they do not make changes to the database.

```
on openStack
  set the userLevel to 1
  hide menubar
end openStack
```

ENHANCEMENTS

The steps I have described in this article should enable anyone with a basic working knowledge of HyperCard to build a HyperCard stack that allows end users to conduct Boolean AND or OR searches in an in-house HyperCard database. The scripts as I have described them operate in what I hope is a smooth and seamless fashion, but there are undoubtedly improvements that could be made. One of the beauties of HyperCard programming is that anyone can build on what others have done by writing scripts that run more efficiently and creating buttons to provide more features.

For example, the variations on the Find command I described in the first section of this article could be incorporated into radio buttons to give users the option of searching for full words or phrases. A print button could be added to the database stack so that users can print out the results of their searches. You could also write an openCard script to display the card number on each card in the database stack, so that the user gets a better idea of the size of the database. A "time out" feature that resets the computer to the search card or to an introductory card could also be added. These are but a few of the alterations that could be made to these stacks to make them more useful. So go ahead, create my stacks, make some changes, and see what HyperCard can do for your library.

REFERENCES AND NOTES

1. A good introduction to HyperCard is Monica Ertel and Jane Oros, "A Tour of the Stacks: HyperCard for Libraries," *Online* 13:45–53 (Jan. 1989). The authors include an extensive bibliography of HyperCard books, magazines, learning tools, and sources. Another bibliography and examples of HyperCard in academe can be found in the HyperCard anniversary issue of *Wheels for the Mind* 4, no.3 (1988).
2. Two articles I used extensively are Chris Pelkie, "Hypersearch Toolkit," *Nibble Mac* 3, no.4:38–46 (1988), and Chris Pelkie, "Text

Scanner," *Nibble Mac* 3, no.6:64–71 (1988).
3. The script I have used here is by Warren Michelsen and is included in Dan Shafer, "More on Buttons; HyperCard v. 1.2," *The APDAlog* 3:24–27 (July 1988). ■■

Conversion of a Non-MARC Database to MARC

Frank Winter

In November 1984 the University of Saskatchewan Library signed an agreement with Geac Computers Ltd. to implement the Geac Library Information System. The library, in common with most other libraries in this situation, was immediately faced with a major retrospective conversion project but of an uncommon type. Almost 100 percent of the library's 586,000 titles were in machine-readable form, with only a few nonroman alphabet shelflist cards remaining to be transliterated and converted. The majority of the library's serial and monograph item level holdings were also in machine-readable form and linked to barcodes. The conversion problem was that the bibliographic database was not in MARC format. This article describes how the conversion was accomplished.

THE TESA DATABASE

The TESA (an acronym for Technical Services Automation) system was locally developed and implemented in February 1971. TESA ran on an IBM mainframe and was almost entirely a batch system. The system supported monograph acquisitions, monograph and serials cataloging, and card production.

Although the library extracted MARC records from its local source record database for use in TESA, the record format used in the master bibliographic database, the Library Holdings File or LHF, was considerably simpler than the MARC format. There were several reasons why a simplified LHF format was adopted. First, storage space was at a premium, and the overhead of the MARC record (the directory and fields that were not used by the library) was significant. Second, TESA programs were written in COBOL. For reasons particular to COBOL, there would have been needless programming complexity if different record formats were used for acquisitions and cataloging files. Third, the primary use of the database record was to produce catalog cards. Only enough information to format these cards accurately was necessary. Finally, the policy of minimal tagging also had the benefits of less staff training, a relatively low error rate, and freeing catalogers to concentrate on content rather than content descriptors. However, the format was designed from the beginning to be convertible back to the more complex MARC format.[1] The record format was modified over time to be consistent with developments in MARC and AACR and to meet the changing needs of the library.

The LHF format consisted of three sections—a fixed length directory, fixed length control fields, and variable length data fields. Simple indicators and delimiters were also used to further indicate the nature of the contents of a variable field and to mark sections within a variable field or to mark the end of the field. Figure 1 illustrates the complete LHF record format. Figure 2 shows a record in LHF format.

In 1976 the library received approval to begin in-house development of an online circulation control system to replace its existing punch card-based system. Although retrospective conversion had been discussed since 1971, this decision gave new urgency to the need to convert the library's 220,000 pre-1971 accessions. The choices facing the library were the usual ones: either to create a skeletal circulation record with holdings or to convert the entire shelflist record to the LHF format. It was decided to do a full conversion. Here again the usual decisions of whether to contract with an outside vendor, to match records against an outside database, or to keyboard in-house were investigated. It was decided to use library staff to rekey the entire shelflist card. The other critical choice was that

Frank Winter is Associate Librarian, University of Saskatchewan Libraries, Saskatoon, Canada.

although it was decided that there would be no attempt to upgrade the shelflist records to AACR levels, the retrospective conversion team would upgrade the punctuation and spacing of certain fields to conform with the ISBD standard. This decision was consistent with the alterations made to coding when the library adopted the revised chapter 6 of AACR in 1974.

CONVERSION TO MARC

During the negotiations with Geac it was realized that the database format conversion would be a critical element in a successful implementation. The MARC record format was central to every procedure where bibliographic records were used in the integrated system. The Geac system could handle "nonstandard" MARC tags and delimiters, should the library so choose, but in that case the full power of the software would not be utilized. Online and batch-edit checking of a cataloger's work and duplicate checking of incoming MARC source records depended on system-supplied tables of authorized tags, indicators, and subfields. The content and therefore the look and the effectiveness of indexes would be affected. For instance, if the statement of responsibility in the title field (Tag 245 $c) was not properly delimited, searches on the title and title keyword indexes would be affected. The display of bibliographic records in the online catalog and in other products such as COM fiche or printed reports and the usefulness of the ability to generate bibliographic reports from the CIRC file would be hampered. In the acquisitions module the usefulness of the interface that loaded records received from approval plan vendors and then transferred them into the cataloging edit file depended on a standard MARC format which would be consistent throughout the local system.

The library's Geac implementation team had to wrestle with decisions on whether to use the Geac system's capabilities for display constants, format codes, and stopword lists to solve some of the conversion difficulties or whether to try to make the database record as complete a MARC record as possible. The library was also aware that it must look beyond the immediate conversion issues to position itself for migration to some undetermined future automated system.

The decisions finally made were that:

- the library would convert its LHF database in such a manner as to conform as closely as possible to the standards of the Canadian MARC communications formats for monographs and serials;
- any local data not clearly equivalent to the meaning and content of standard MARC tags, indicators, or subfields would be kept distinct by the use of local tags, local subfields, and local indicators;
- no data in the LHF would be discarded; and
- certain leader values would be stored in a local tag 901 in order that they would be available for display and editing online or on printouts. (A secondary reason for this decision was that library staff had heard stories of other libraries' difficulties in obtaining complete records in MARC format when moving from one vendor's system to another vendor's.)

BIBLIOGRAPHIC CONVERSION

During pre-contract negotiations with Geac that company had agreed to do the conversion of the bibliographic database to the MARC format. However, the library's investigation of the work necessary to map the LHF format to MARC revealed that the entire process would be very detailed and very complex. There was some feeling that Geac had underestimated the complexity of the task, but more fundamentally, it came to be accepted that the conversion effort must be the library's responsibility, not Geac's, if there was to be as detailed and complete a conversion to MARC format as possible.

Investigation of the issues began in April 1984. Information about TESA and the LFH was very scattered. Staff assigned to the conversion used whatever partial documentation was available as well as procedural change memoranda and individual memory. For instance, the initial source to develop a table of all the General Material Descriptors that had been used by the library was the list of terms contained in AACR2. This was supplemented by local documentation from the Cataloging Policy Manual and finally added to by consulting

1. Control Fields

FIELD NUMBER	FIELD ID	SIZE	POSITION	TYPE	DESCRIPTION
1		8	1-8	X	Library holdings number
2	AC	12	9-20	X	LC card number/Canadiana Serial number
3	AF	1	21	X	MARC indicator
4	AJ	1	22	X	type of acquisitions
5	AK	1	23	X	form of publication
6	AN	1	24	X	Kardex indicator
7	AP	4	25-28	P	net price
8	AQ	2	29-30	X	country
9	AR	3	31-33	X	language
10	AB	10	34-43	X	ISBN/ISSN
11	AH	1	44	X	source library
12	AM	1	45	X	treatment indicator
13	AN	2	46-47	A	cataloguer's initials
14	AG	1	48	X	serials indicator
15	AT	2	49-50	I	date of publication
16		4	51-54	P	date last modified
17		1	55	B	update to Circulation system indicators
18	AR	1	56	X	AACR2 indicator
		6	57-62		not used
19		2	63-64	I	length of variable portion
20		54	65-118	I(27)	offset from position 121 for each of the 27 variable fields
21		up to 9878	119-9996	X	variable fields

2. Variable Fields

FIELD NUMBER	FIRST INDICATOR	SECOND INDICATOR	NAME
01	P,C,M,U		Main Entry (author)
02	A,N		Title
03			Imprint
04	T,D,U		Series Statement
05			Edition Statement
06			Purchase Order Comment/Serial receipt info
07			Document Number
08			Cheque Number
09			Location/Copy Info (serials only)
10			Departmental Interest Profile
11			Holdings Field
13			Call Number
14			Collation
15	U,R,T		Supplied Titles
16			Bibliographic Notes
17		Sequence Number	Subject Added Entries
18	P,C,M,U,T	Sequence Number	Other Added Entries
19			Series Added Entries
20			Location Note (Replaced)
21			Supplements
22			Shelf List Note (Replaced)
25			Continuation of ISBN
26			Cataloguing Instructions
27			Links for the Search Code File

Figure 1. TESA Libhold Format.

3. Indicator Codes

FIELD NUMBER	FIRST INDICATOR	DESCRIPTION
01	P	Personal name
01	C	Corporate name
01	M	Conference or Meeting name
01	U	Uniform Title Heading
02	A	Title (Added entry)
02	N	Title (Not traced)
04	T	Series name or title (traced)
04	D	Series (traced differently)
04	U	Series (untraced)
15	U	Uniform title
15	R	Romanized title
15	T	Translated title
18	P	Personal name
18	C	Corporate name
18	M	Conference or meeting name
18	U	Uniform title heading
18	T	Title traced differently

Figure 1. Cont.

```
        Library Holdings File Number:  0808403C        LAST MODIFIED 08/04/81
        "A" Card Fields                                (A2 - NOT PRESENT)
AB (SBN) ƀ AC (LC CARD NO.) ƀ AF (ON MARC TAPE) N
AG (SERIALS IND.) ƀ AH (ORIGINAL SOURCE LIBRARY) D AJ (TYPE OF ACQUISITION) ƀ
AK (FORM OF MATERIAL) B AM (ORIGINAL HANDLING CODE) ƀ AN (KARDEX IND.) ƀ
AP (NET PRICE) $.00 AQ (COUNTRY CODE) CN AR (LANGUAGE CODE) ENG
AT (PUB. DATE) 1932
             CATALOGUER'S INITIALS MC
               CARD ID              DIAGNOSTIC
                   90 345678901234567890123456789012345678901234567890123456789012345678
FIELD
                      I
MAIN ENTRY         01  RJenness, Diamond, #1886-$
                      I
TITLE              02 AThe Indians of Canada/#by Diamond Jenness ...$

IMPRINT            03 Ottawa : F. A. Acland, King's Printer, 1932.$
                      I
SERIES NOTE        04 TCanada. National Museum.#Bulletin ; 65$

HOLDINGS FIELD     11 Ød CØ:40690##  c.2:294486$

CALL NUMBER        13 E78 .C2J5$

COLLATION          14 x, 446 p. : ill. ; 26cm.$
                      I
SUBJECT ADDED
ENTRIES            17 1Indians of North America - Canada.$
```

Figure 2. LHF Record.

the cataloger who had had responsibility for handling most of the relevant material over the preceding decade. In May 1984 the initial specifications for monographs and serials were drafted. The general patterns and techniques that would have to be used in the conversion program emerged early on in this investigation, but these specifications were continually refined as testing revealed deficiencies.

There were several levels of complexity in the mapping. The most straightforward mapping was from a single LHF field to a single MARC tag. Slightly more complex was the case of a single LHF field containing data that could be mapped to one of several MARC tags. Sometimes these MARC tags were the same tags such as the subject headings field, field 17, which could contain multiple subject headings. Sometimes a single LHF field could be mapped to several different but essentially similar MARC tags. The most complex situation occurred where a single LHF field contained data that had to be mapped into several quite different MARC tags.

Author main entry was an example of the most straightforward conversion. Field 01 Indicator P was mapped to Tag 100 $a. Field 01 Indicator M was mapped to Tag 111 $a.

An example of a single LHF field containing multiple but essentially similar MARC tags was field AC. This field might contain a Library of Congress Card Number, a Canadian number, or a UK MARC number. The data in this field were analyzed by the length and position of the numbers and correctly assigned to either tag 010, 015, or 016.

Where MARC tags, subfields, or indicators had quite specific meanings that could not be correctly assigned by the conversion program, local values were substituted. For instance, it was not possible to distinguish the subject headings in field 17 with sufficient precision to assign tags 600, 610, 611, 630, or 650 or their associated indicators or subfields. A local tag 699 $a was assigned to all subject headings, with local first and second indicators of b̸b̸. Subsequent subfields in field 17 were identified by the presence of " b̸ - b̸." This character string was eliminated, and a local $i subfield code substituted. The conversion program was also set to look for the presence of a single delimiter "#." If one was found, a $t subfield was assigned.

Conversion of series entries in field 04 illustrates some other techniques. It was critical in this case that the conversion program performed the examination and tag and subfield substitution in the correct order. If there was a LHF indicator of D or U present, the MARC tag 490 could be assigned, with a first indicator value of 1 (if D) or Ø (if U). Then the program would search for the presence of "%#" and assign $x. If "b̸;b̸" was found, then this string was replaced with "b̸;$v." By searching for personal pronoun character strings: "%Itsb̸," "%Its#," "%Hisb̸," "%His#," "%Hersb̸," or "%Hers#," tag 400 could be correctly assigned, as could tag 440 where there was a TESA indicator T and no title delimiter (#). Any other series entries in field 04 would have to be either tag 410 or 411. These tags could not be accurately assigned, and a local tag 434 was substituted.

The other major technique was format recognition using ISBD punctuation, text strings, and language tables. This technique was used frequently to assign subfield codes and to assign the proper values for nonfiling indicators in tags 130, 240, 241, 242, 245, 434, 730, and 740. For instance, if the language code in field AR was FRE, the program would consult a table of French language definite and indefinite articles to assign the proper nonfiling indicator.

Searching for ISBD punctuation worked very well in assigning correct subfield codes to tags 245, 260, and 300 for most monograph records. It did not work for monographs catalogued between 1971 and 1974 unless they had been subsequently revised. These items had been added to the collection after TESA was implemented but before the library adopted the ISBD coding standard.

Searching for text strings was also used very successfully when searching field 16 (bibliographic notes) to assign the proper indicators for tags 777, 780, and 785 of the serial records and various 5xx tags. For instance, if the phrase "Issued with" was found, tag 777 could be assigned. Words and phrases such as "Continues," "Supersedes in part," and so on would enable tag 780 to be assigned with the correct second indicator.

Figure 3 illustrates the MARC conversion format of the record in figure 2.

One element in the LHF record not converted was the item level holdings data in field 11. This field was a complexly coded, highly compressed holdings format. It was realized that the CIRC file, which contained holdings in an "exploded" format and which in some cases was more completely coded than the LHF record, would be a better source for item level information. It was decided to merge the item level information from the CIRC database with the bibliographic information in the LHF database before creating a MARC format record. This aspect of the database conversion is not discussed here.

TESTING

The development of file conversion specifications lasted from April 1984 to January 1985. In May 1985 twenty LHF records were converted into MARC format to check the program. The challenge facing library staff at this point was to devise testing methods that would reveal deficiencies and that could be used to report back to the programmers. This process was hampered upon occasion by system operators who were learning the new system and who inadvertently corrupted files. Eight staff worked in four teams for two weeks to check every character in each corresponding LHF and MARC record to ensure that the conversion worked as specified and also to ensure that no data were lost in conversion. The LHF-to-MARC specifications were revised as necessary, and a second test run of 243 records was produced. Further modifications to the conversion program were necessary, and new specifications for some tags were developed. Fifty dummy LHF records were added to the test file in order to create conversion conditions that had not been present in the 243 record file. Four more iterations were necessary before the program was approved in October 1985.

MARC Tag	Seq No.	Create Info	1st Ind	2nd Ind	Data Len.	Field Data
001	001		ƀ	ƀ	0008	808403cƀ
008	001		ƀ	ƀ	0041	000000u19320000cnƀuƀƀƀƀƀƀƀƀƀƀ0333uuengur0
099	001		ƀ	ƀ	0011	$aE78 .C2J5
100	001		ƀ	ƀ	0025	$aJenness, Diamond$d1886-
245	001		1	4		$aThe Indians of Canada /$cby Diamond Jenness ...
250	001		ƀ	ƀ		$aOttawa :$bF. A. Acland, King's Printer,$c1932.
300	001		ƀ	ƀ		$ax, 446 p. :$bill. ;$c26 cm.
434	001		9	0		$aCanada. National Museum.$tBulletin ;$v65
699	001		ƀ	ƀ		$aIndians of North America$iCanada.
966	001		ƀ	ƀ		$1dƀ$m$b9010000192936
966	002		ƀ	ƀ		$1ƀƀ$c2$m
901	001		tam5u			
902	001		mcn			

Figure 3. LHF Records Converted to MARC.

The next tests were of the Geac program which loaded the conversion records into the new system. This lasted from October to December 1985. In December 1985 work began on testing the transfer of bibliographic and holdings data to the Geac circulation module. This was completed in January 1986.

SEQUENCE OF CONVERSION

The division of work finally agreed upon was that:

- the library would create a Library-To-Geac File (LTG) from its LHF and CIRC files. The LTG file would consist of one record for each tag in MARC format;
- Geac would assemble the LTG records into a complete Geac MARC record containing bibliographic and holdings information. Geac would also normalize the call numbers and complete the information in the summary holdings tags (tag 930) from information contained in the individual holdings tags (tag 966);
- the library would load the completed records into its new local master bibliographic database (MBIB); and
- the new circulation database (CIRC) would be created using the standard MBIB to CIRC transfer program.

Figure 4 illustrates the sequence of conversion. Figure 5 illustrates the format of one LTG record.

MAINTAINING TWO SYSTEMS

The initial snapshot of the TESA database took place on October 7, 1985. It was known that several supplemental loads would be necessary because of the time necessary to create, load, and index the various databases. Six subsequent loads were performed between October 7, 1985, and April 22, 1986.

During this period, the library had to determine how the two systems were to be maintained. In TESA, the library wanted to continue to produce catalog cards and to transfer new records and holdings to the CIRC module that had to function until the Geac CIRC module went live. Database maintenance staff were given detailed instructions on the types of work which would be allowed. New accessions and holdings updates continued on TESA and CIRC. Record and holding deletions for records added then deleted between one load and the next were permitted, but deletions of records already in the LTG file had to be held back until the live file could be accessed on Geac. Skeletal record creation on CIRC and the assignment of replace-

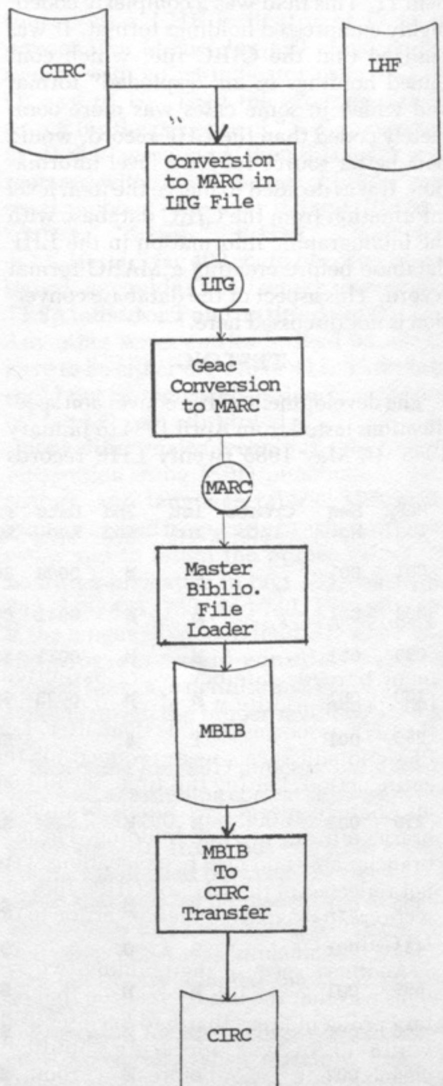


Figure 4. Sequence of Programming and Database Conversion.

Type Code

N = Numeric character

A = Alphabetic character

X = Alpha-numeric character

FIELD #	NAME	SIZE	POSITION	TYPE
1	Control Information	16	1-16	X
2	Libhold Number	8	9-24	X
3	MARC Tag	3	25-27	N
4	Sequence Number (control)	3	28-30	N
5	Run Date Created (control)	6	31-36	N
6	Run Hour (control)	2	37-38	N
7	MARC First indicator	1	39	X
8	MARC Second indicator	1	40	X
9	Field Data Length	4	41-44	N
10	Field Data	Var	45-5000	X

Control Information fields may contain data useful for the conversion process, but which itself is not to be included in the MBIB database.

Note that the 4 character record length preceding this record is not shown in the record.

Figure 5. LTG Record Layout.

ment barcode numbers were permitted, but special priority was given to updating the LHF as soon as possible thereafter. Log files and edit sheets were retained for the entire period. During this time, the library also had to merge two small branch libraries into one in a new building. The new branch name and new book locations were not reflected in the LHF or CIRC files but were created as part of the conversion to the LTG file.

Another aspect of maintaining two systems was that the library realized that certain problem areas that had been set aside while TESA was operating should be cleaned up in TESA before conversion, either because the solution was easier or because clean-up on the Geac system would have to wait until maintenance staff were thoroughly familiar with a new system. In some cases, this clean-up meant that certain items had to be recataloged.

AFTERMATH

In general, the conversion to MARC format was very successful. All of the objectives set for the conversion were met. However, problems have arisen in certain specific areas.

Country Codes

The country codes in the LHF were two letter abbreviations, as they are in MARC for most nations. However, Appendix B of the Canadian MARC communications format for monographs specifies that three letters be used for the United States, Canada, the United Kingdom, and the U.S.S.R. The country codes are one of the data elements which the library has chosen to carry from the MBIB record to the circulation files in order to write reports. The programming to create the data element in CIRC is site specific. Either the library must alter its programming at some cost and then rebuild the relevant circulation file or each incoming MARC record's country code must be edited to conform with the existing database. The library has chosen to do the latter and hopes to find an opportunity during the next major system upgrade to alter the data. (A similar problem was created during the call number normalization, where it proved impossible to eliminate the second period that was used to bound .I. and .O. in cutter numbers appearing on cards. The call number index is sensitive to the presence of periods because of the mixture of call number schemes which the library uses. Staff must continue to put in the second period to keep the index consistent and to permit accurate retrieval.)

Database Maintenance

The library's catalogers adhere to the standards of the Canadian MARC formats for current coding. However, for many years to come most of the bibliographic records will be in the conversion format. While doing database maintenance there is often an opportunity to upgrade the coding of tags that are used as access points. Despite exhortations to the contrary, cataloging staff sometimes unncessarily alter the coding on converted records that are being updated. This creates needless work and unnecessary system load in an area where there already is a throughput bottleneck. This type of cosmetic change will probably increase as a turnover and time distance the cataloging staff from the LHF format and familiarity with the conversion decisions.

Conflicts between Local and Source Record Tags

Much of the local TESA data which could not be assigned to a standard MARC tag was put into local tags 901, 902, 906, 910, and 911. Tags 910 and 911 conflict with Canadian MARC tags 910 and 911, which carry French or English language equivalencies, cross-references, and history notes. Consequently, these tags must be stripped out of incoming source records. As the library begins to use more source records from other services, such as UTLAS or OCLC, it is possible that there will be more conflicts between other local and source record tags, coding and/or content that will require manual inspection to detect and correct.

Authorities

The inability at the time of conversion to correctly assign standard tags or subfield codes to certain author and series tags and all subject headings will increase the amount of manual intervention necessary when the library is able to integrate an authority control system with the online catalog.

CONCLUSIONS

The successful conversion to the Canadian MARC format was primarily made possible by three conditions at the library that assisted the programmers. First, catalogers and shelflist conversion staff were expected to adhere to conventions of coding (that is, ISBD) and wording from the Anglo-American Cataloging Rules. Database quality in this area and also in the areas of spelling and content was ensured by rigorous checking. Secondly, there was a great deal of effort devoted to keeping the bibliographic and circulation databases clean and compatible one with the other. Finally, the library "knew" its databases, either because of adherence to standards, quality control and such documentation as existed, or through the use of computer diagnostics which enabled staff to identify records that fulfilled certain criteria needed for testing. In the excitement and confusion of implementing a new system, when policies and procedures are being made on the fly, it is necessary to remind staff constantly to document decisions in some central source, such as a Cataloging Policy Manual. Similarly, in times of budget cutbacks it is necessary to resist the temptation to cut back on database quality control. The lessons the University of Saskatchewan Library learned from the conversion illustrate how critical the consequences of neglect could have been.

Bibliographic databases in non-MARC formats abound. The LHF-to-MARC conversion experience is still relevant as the library decides whether to incorporate other databases into its master bibliographic database. For instance, one of the colleges federated with the university would like to merge its library records with the master bibliographic database on Geac. Although most of this college library's catalog is not in machine-readable form, for the last two years it has been using the Librarian's Helper software package to produce cards. This product has a record format remarkably similar to the LHF. Thus, the issues surrounding the question of format conversion were easily resolved. In another instance, the library's Government Publications Department has begun the creation of a machine-readable database of records using the BUCAT software package. Source records in the CODOC format for retrospective conversion purposes were obtained from the CODOC Consortium but had to be converted to a Canadian MARC format to allow for the possibility of eventually merging these records into the master bibliographic database.[2] Cataloging staff were able to use their experience with the LHF conversion to advise on the CODOC conversion.

It is now not unusual to find library automation systems that can handle multiple bibliographic files in such a manner that users can conduct a search across these databases using a single argument. When Boolean and other relational operators can be used in combination with authority control, some of the concerns about quality control and adherence to standards may diminish. However, in an integrated library system such as Geac it is very important to keep the format as standard as possible in order to maximize the power of the software.

ACKNOWLEDGEMENT

The work of Peter Burslem, Frank Coxford, and Rose-Anne McCrory at the University of Saskatchewan Library and of Terry Cohen and Paul Nixon at Geac in achieving a successful conversion is gratefully acknowledged. Rose-Anne McCrory and Frank Coxford also reviewed earlier drafts of this paper and made many helpful suggestions.

REFERENCES AND NOTES

1. Edwin J. Buckinski and William L. Newman, "The Cataloguing System at the University of Saskatchewan, Saskatoon," in *Automation in Libraries: Papers presented at the CACUL Workshop in Library Automation* (Canadian Association of College and University Libraries, 1971) p.3.3–3.4.
2. For a description of another CODOC to MARC conversion, see Peter Hajnal, Valentina de Bruin, and Dale Biteen, "MARC and Codoc: A Case Study in Dual Format Use in a University Library," *Journal of Library Automation* 10:358–73 (Dec. 1977). ■■

INDEX TO ADVERTISERS

Recent Publications

Book Reviews

As the book review editor's six-year term comes to a close, she would like to thank Eileen Fenton for her invaluable editorial assistance.—Karin Trainer, Editor.

Informetrics 87/88: Select Proceedings of the First International Conference on Bibliometrics and and Theoretical Aspects of Information Retrieval, Diepenbeek, Belgium, 25–28 August 1987. Ed. by Leo Egghe and Ronald Rousseau. Amsterdam and New York: Elsevier, 1988. 329 p. $92 (ISBN 0-444-70425-6).

This book has great value, but its value has been very well hidden from the nonexpert reader. In form it is simply an assemblage of twenty-six papers, about half of the papers presented at a conference that the editors (who were also the organizers) describe as the first international conference on bibliometrics. For no apparent reason the papers are arranged in alphabetical order by author, without an introduction or an adequate preface to give the reader any sort of overview of the topics that are discussed and how they interrelate. Even a canny reader who makes an initial reconnaissance of a difficult book by skimming its index will be frustrated here; the index is merely an unhelpful alphabetical listing of key terms selected rather casually and inconsistently from the papers, and lacking any grouping or linkage of terms or any hierarchical framework that would suggest an underlying scope and structure for the subject matter in the book.

Nevertheless, no librarian, educator, researcher, or practitioner in the information field should summarily dismiss the book. The papers in it range widely and deeply over an increasingly important field; they are of good currency and technical quality and are written with commendable clarity for the most part. In aggregate, this material could, with the addition of some organized commentary, serve quite well as a textbook on bibliometrics. (The editors would like to see the term "informetrics" accepted for this field in place of "bibliometrics" and "scientometrics.") In their brief preface, the editors state their conviction that the field should be defined broadly, to cover all statistical and mathematical aspects of library, documentation, and information problems, and explicitly include theoretical aspects of information retrieval. The papers selected for publication here reflect a deliberate effort to present a very wide variety of bibliographic methodologies and concerns. For example:

Brookes (London) reviews the development of the bibliometrics concept and some of its component statistical ideas (e.g., sampling, frequency, ranking, goodness of fit), and points out that some of the apparently competing bibliometric distribution models were devised for application to quite dissimilar problems and cannot be regarded as equivalent. Van Poucke (Limburg/Diepenbeek) presents a clear statement of the need for, and the promise of, bibliometric research to cope with the growing complexity of decision making for information providers and information seekers. Tague (Western Ontario) illustrates some practical considerations that are important in order to make bibliometric studies useful to practitioners and not simply interesting to theoreticians. Bookstein (Chicago) shows that bibliometrics should be taken seriously, by examining the long-term consequences for a library of adopting book acquisition policies based on different models of the relative values of different sections of the collection; a policy incorporating the notion of a bibliometric distribution of such values yields distinctive consequences that appear to be more realistic than the consequences of other types of models. Egghe (Limburg/Diepenbeek) reports a mathematical analysis of a simplified technique devised by Fussler to obtain a sampling of books in a shelved collection, to determine how much variation from true random sampling might result from adopting that technique (very little). Ravichandra Rao (DRTC, Bangalore) discusses bibliometric analysis of library circulation data. Radecki (Nebraska) and Salton (Cornell), in separate papers, discuss the merits of theoretical models of document retrieval processes. McGrath (Buffalo) describes the use of cluster analysis for comparing library collections. Various papers discuss the properties or comparative performance of specific types and variants of bibliometric models.

Quite a few of the papers focus on methods for descriptive analysis of large information source populations, as opposed to focusing on the attainment of information service objectives. Moed (Leiden) discusses ways in which online databases may and may not be used for bibliometric analysis. Others discuss actual bibliometric studies involving, for example, frequencies of word occurrence or co-occurrence, or frequencies of citation or cocitation, that have yielded better understanding of authorship patterns, literature development, terminology development, the structures of scientific disciplines as revealed in publication patterns, and so on. Appropriately, there is even a bibliometric analysis of the literature of bibliometrics, by Peritz (Hebrew University).

Each paper carries an abstract, which is helpful. The volume includes lists of the conference participants and the sponsoring organizations. The quality of editing and production is less than one would expect in so expensive a publication. There are numerous misspellings and instances of exotic word usage; they hurt, but do not destroy, the understandability of the papers. More serious are some failures to edit for completeness and consistency. For example, at the start of a paper by Griffith (Drexel) on goodness of fit of bibliometric models, there is no text that relates to his first two references; whether the problem was with the manuscript or with loss of text during composition, it should have been caught and corrected.—*Ben Ami Lipetz, School of Information Science and Policy, State University of New York at Albany.* ■■

Kibirige, Harry M. *Local Area Networks in Information Management.* New Directions in Information Management, no.18. New York: Greenwood, 1989. 177p. $39.95 (ISBN 0-313-26191-1).

At a glance, this is the kind of book that you want on your shelves. It is part of a series that includes several well-respected authors. The chapters have numerous references, and there is a bibliography, a glossary, an index, and a number of figures and tables. The book is physically well built, and the paper complies with the NISO permanent paper standard.

It still looks good when you start reading. The author states that his perspective is that of an information scientist but that the book should be useful for most people who want to "get a good grasp of the LAN technology and its applications." The text is based on three years of research that included a one-year research grant from the City University of New York (CUNY).

But then small details begin to suggest that all is not well. Microcomputer "memories have advanced to . . . 512K"; what happened to 640K? Megabyte memories are reserved for "super micros" (32-bit machines by the author's definition). Throughout the book are examples of incomplete or out-of-context statements. In a confusing section on coaxial cable, the author states that "baseband coaxial cable may sell from 50 cents to $3 per foot, whereas regular broadband coaxial cable ranges from 35 cents to $1 per foot." This misleads the reader because it omits the fact that the broadband environment typically uses miles of cable whereas the baseband environment typically uses hundreds of feet. In other words, the total broadband cable cost is usually higher. The author talks about X.25 in a LAN context but makes no mention of TCP/IP (protocols for packaging and moving data). In a discussion about CSMA (network traffic protocol), the author states, "for real-time applications like item check-out systems, its unpredictable nature is unsatisfactory." I will hate to have to tell this to our circulation staff, since the circulation terminals are connected to the computer over Ethernet LAN using CSMA in one of its incarnations!

This leads me to my biggest concern with this book. The author confuses the communication aspect of a LAN with applications that run on distributed systems. He repeatedly laments that there are no library systems that run on a LAN; I think he is trying to say that there are no distributed systems. The point is that a LAN creates a communications and network environment in which the applications may or may not be distributed.

The report of the CUNY study omits many important details but supplies page after page of irrelevant statistical analysis. Based on the incomplete explanation regarding the samples, I have to conclude that there is a significant potential for bias, especially in the area of vendors of library products. There were seventeen statistical factors reported for each user category regarding a count of microcomputers; means are reported to four decimal points. This may look impressive, but it conveys very little information. Do you understand what a kurtosis of 6.44194 looks like? Lacking the questionnaire, the reader has to guess the meanings of the various categories, for example, "shrdev," "preslan," "oth_obj." The tables have extraneous unlabeled numbers whose meaning is impossible to determine. It is meaningless to report cumulative frequencies and percentages for yes/no responses.

Because of this book's potential to mislead a naive reader and its lack of content for the knowledgeable reader, I cannot recommend it. However, in reward for reading this far, I highly recommend that you do buy *Local Area Networks: The Second Generation* by Thomas W. Madron (Wiley, 1988). At 256 pages for $24.95 (paperbound but of very good quality), there is much more content for your money. The information is very accurate. There is enough technical detail to make this useful as a reference work, but the presentation is lucid and easy to understand. Although published a year earlier, this book has references that are more current than those in Kibirige's book and that also provide some good examples of the diverse, specialized LAN publications that are available. The glossary has useful definitions, there is an index, and the illustrations provide graphic examples to support the technical detail. *Local Area Networks: The Second Generation* could be used as a state-of-the-art review or as a textbook.—*R Bruce Miller, University of California, San Diego, La Jolla.* ■■

Saffady, William. *Optical Disks vs. Micrographics as Document Storage and Retrieval Technologies*. Westport, Conn., and London: Meckler, 1988. 106p. paper, $27.50 (ISBN 0-88736-345-8).

When forklifts become part of document handling systems, record managers look for alternate means of archiving or storing materials. The march of technology has complicated choices in this area, and the assurance of further rapid change will not quiet the sleep of document management planners. William Saffady's book, *Optical Disks vs. Micrographics as Document Storage and Retrieval Technologies*, is intended as a guide to the choice between the trusted, stable, film-based technology and the immature but promising optical filing technology. The intended audience of this book includes planners, designers, or implementers of document or records management systems; planners will benefit most.

The text is broken into two sections. The first is a brief survey of the literature addressing the competitive relationships between micrographics and optical filing systems. It provides a functional comparison and contrast of the two technologies as well. The second part, fully 85 percent of the text, is a detailed comparison of computer-assisted microfilm retrieval systems and their optical filing counterparts (including examples of hybrid systems).

Saffady's text is not shy of detail. He writes about his subject with the detail with which Melville wrote of whales; one can be forgiven a certain sympathy with Ahab. Nonetheless, the myriad details about equipment, standards, capacities, work flow, retrieval probabilities, media costs, indexing, legality, and storage longevity will reward the careful reader who is planning a substantial investment in a document storage or archival system. There is even mention of paperclip removal time—no trivial concern when documents come in pallet loads. The waters here are treacherous, and Saffady's long experience in the micrographics realm can be ignored only at one's peril.

The major users of this book will be those choosing (or planning for a choice) among competing systems in this evolving arena. Unfortunately, it is not well organized to fill this reference role. One would like an extensive table of contents and, considering the many tables, a list of table captions as well; neither is present. The reader is left largely to the search-by-scan method of fact retrieval.

The index is somewhat uneven as well. One might expect a book comparing high-density storage media to have an entry for recording density. There is no listing under *density*, *capacity*, *recording density*, or *recording capacity*. How about *media characteristics*? This term points to an undifferentiated thirty-six-page range. If you look under *areal density*, you start to get somewhere, but this seems an unlikely first choice of access points. The entry for *aperture cards* leads the reader to two textual occurrences (neither of which alludes specifically to aperture cards) but not to a sequence of tables (2–18 through 2–26) that summarize optical storage media capacity and various microform equivalences, including aperture cards.

Assessing a moving target such as optical storage systems is necessarily fraught with danger. The book contains examples of out-of-date information due to marketplace changes, but this problem is hard to avoid in the analysis of automation technology. Thus, for example, costs for mastering CD-ROMs are out of date, although in general the decision principles are valid.

The bibliography provides a useful entry to the relevant literature. The references include 145 citations from 1979 to 1987; 80 percent date from 1984 to 1987 inclusive. Saffady's analysis of the literature provides an overall flavor of the consensus of scholars and pundits. His summaries of the literature reflect a field that is characterized by change and uncertainty, offering no clear path for the document manager. This is realistic given the immaturity of optical media in particular. There is very little in the way of critical analysis or commentary; the reader has the job of evaluating the import of published opinion. Neither is there any attempt to suggest the future trajectory of this field. Saffady's book is not the place to look for trends; he is firmly grounded in the here-and-now.

The typeface and layout is typewriter quality; it is large and easy enough to read, despite uneven printing quality, but seems somewhat primitive in this age of desktop publishing. This volume is obviously not intended for the ages, but our expectations about presentation quality are increasing, even for transient information such as this. This is a small, but increasingly important issue to those who spend much of their professional lives with the written word. Typeface cues such as differentiated section and subsection headings are especially important in a volume that lacks suitable access tools to navigate the details and tables that comprise the meat of this book.—*Stuart Weibel, OCLC Office of Research, Dublin, Ohio.* ■■

Samuels, Alan R. *Essential Guide to the Library IBM PC: Shareware for Library Applications*. Essential Guide to the Library IBM PC, v.10. Westport, Conn.: Meckler, 1988. 234p. spiralbound, $24.95 (ISBN 0-88736-184-6).

Mr. Samuels divides his treatment of this topic into three sections, "Introduction to Software Alternatives," "Applications," and "The Shareware." In the final section, he has done us a service by compiling a descriptive bibliography of shareware products that he has found useful in automating a variety of library functions. The fifty-four products included are reviewed fairly, with both positive and negative comments about cost, availability, features, stability, support, etc. While this chapter constitutes slightly less than half the book, it is in my opinion the most valuable portion of the book.

The preface and first chapter establish the purpose of the book and its intended audience. The author discusses the lack of a comprehensive "fair shake" treatment of shareware as a viable alternative to commercial software and delineates his methodology for evaluating not only individual products, but also appropriate applications. In the remaining chapters, he further relates the often undocumented process of obtaining useful shareware and highlights sources of additional information. By offering guidelines for what to consider and what to avoid, Mr. Samuels provides valuable insight. Unfortunately, there are several editorial oversights in the text and an unacceptable amount of repetition and fragmentation, indicating a need for additional editing before press. I found this disturbing.

Overall Mr. Samuels addresses some important aspects of small-scale library automation and finding alternative solutions given individual library constraints. While implementing a shareware alternative may be less expensive than commercial software when comparing only the licensing fee, the author suggests a major commitment to staying on top of rather elusive software sources, a process that can easily consume a major portion of one's time, not to mention money for long-distance phone calls to remote bulletin boards. Not even the largest of libraries is likely to favor devoting FTEs to this activity, nor would it be necessarily less expensive than obtaining commercial software. This point brings to light the question of whether or not a latest-version fixation is necessary or worthwhile; the author assumes that every application requires the latest available version, an opinion that confuses technological and functional obsolescence.

In dealing with the issue of cost, the author chooses to emphasize the "reasonable" cost of acquisition (usually $1 to $10 per disk) over the cost of registration. If these programs are truly shareware, then the user is entitled to use them only to the extent that permits reasonable product evaluation. Beyond that the user is expected to register if he or she wishes to continue using a product. While the author is quick to state that he has registered his copies, more space is devoted to condemning programmers who protect their livelihood (through limited demo models, time-released use, undocumented features, etc.) than is given to recommending compliance. There is a difference between seeking a less-expensive satisfactory alternative and infringing on copyright law. This distinction should be made much clearer, particularly in a text suggested as a guide to professionals whose obligation it is to provide access to materials while upholding the protection of creative invention.

Evaluating a need and finding the appropriate tool to meet that need is an important process. If a more cost-effective means can be found, then of course it should be investigated, as Mr. Samuels suggests. There are, however, several other issues that receive scant attention in his book. For example, how much training is needed with each particular solution and what costs are associated with each? What options are available for a site license? How many types of programs does one really need? Is it more important to pay less or to do less? Is it necessary to contend with existing standardization? Will all suppliers accept purchase orders? Or should one's staff be expected to fund software acquisition? And what is the risk of viral infection and how can it be minimized? The answers to these types of questions and their relative importance are likely to vary from library to library.

While I find the author's approach a bit narrow at times, I do think that there are useful discussions in this book, although some of them are given more thorough treatment elsewhere. The value of Samuels' work is his description of shareware acquisition, a frequently circuitous process, and his outline of shareware potential in a library setting. With regard to one of his initial objectives he also succeeds—the legitimizing of applying systems analysis and design theory to alternative software solutions. It is unfortunate that the unedited, informal style of the text hides much of this accomplishment.—*Thomas C. Wilson, University of Houston Libraries, Texas.* ■■

Townley, Charles T. *Human Relations in Library Network Development.* Hamden, Conn.: Shoe String, 1988. 161p. $25 (ISBN 0-208-02086-1).

Townley states his purpose in undertaking this 1979 dissertation research as being "to present information on how organization theory and development techniques can be used to improve the effectiveness of library networks." He does succeed in fulfilling this purpose. The theories he presents are usually clearly defined, but undefined jargon has a tendency to pop up in the text, especially in discussions of organization theory. Seven chapters give an overview of library networks while intermingling the application of concepts of organizational behavior with discussions of the organizational intervention development process and organizational intervention techniques. The remaining chapters cover theory and related case studies, with a final chapter summarizing the research and findings and recommending further study.

A number of specific subjects and examples are contained in this work, including: why libraries establish networks; an analysis of human behavior as it applies to networks; organizational development as a process for changing perceptions in network relations; and organizational development as it applies to library networks. Although Townley reiterates the tried-and-true reasons for library network creation, he sheds new light on the subject by describing these organizations and their goals as "providing synergetic opportunities for libraries, library users, and related organizations." Cost-effectiveness, resource sharing and risk sharing as applied to library networks are also mentioned, as are perceptual variables regarding interorganizational relationships. These variables, such as goal compatibility and importance, agreement, influence, and performance, must be addressed if the network is to survive and flourish. The summary chapter touches on the high points of the preceding chapters and is followed by a thirteen-page bibliography and an index.

The case studies are related to theoretical models of specific situations. Stress management in library network management is one such situation, with advice given that moderate stress is good, because it serves to motivate innovation and learning. Other

case studies include a variety of intervention situations. In particular, the author addresses technostructural interventions, where the structure of the work is changed to address the needs of the organization as well as those of the people. The two major facets of technostructural intervention as explained by Townley are work design, where tasks are designed for high personal fulfillment and high organizational fulfillment, and job enlargement, where all tasks are divided into natural units that can be identified and measured. The network case studies are interesting to read and help to communicate the aspects of organization theory, and so "by improving our ability to deal with the perceptual component of network relationships, we can develop library networks that are more rewarding organizations in which to participate, and at the same time achieve a higher level of effectiveness in their operations."

The author strongly stresses the need for networks to seek the aid of outside consultants, with these consultants providing leadership. This position implies that library network participants are unwilling and unable to identify problems and attempt change without outside help. Because of this orientation, the intended audience is assumed to be either novices to librarianship or individuals engaged in research. Librarians with network experience may gain insight into interpersonal and behavior techniques and their application, but network librarians looking for a definitive guide to the human element as it reacts with the network environment had best look elsewhere.—*Denise A. Garofalo, Mid-Hudson Library System, Poughkeepsie, New York.* ■■

Non-Book Reviews

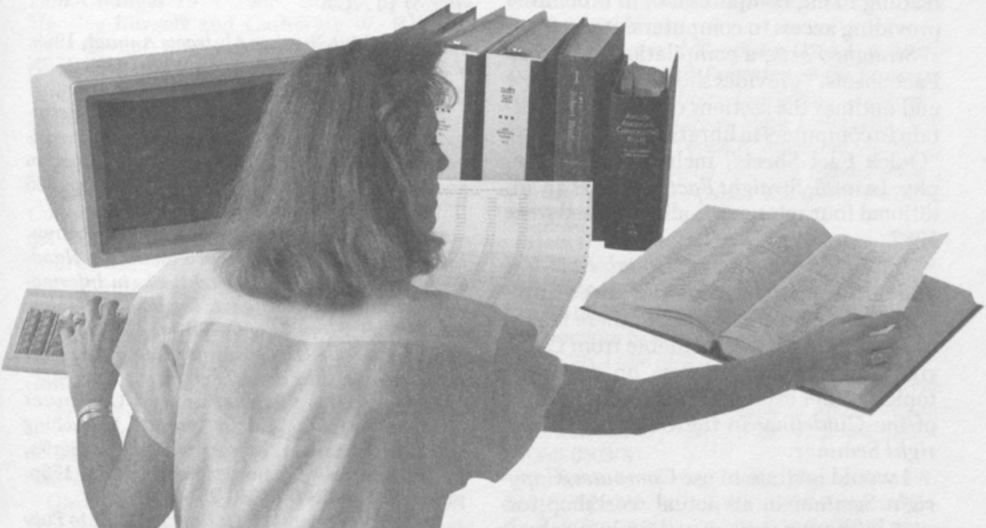
Miller, Jerome K. *Computer/Copyright Seminar*. 4th ed. Friday Harbor, Wash.: Copyright Information Services, 1989. Text and 60-minute audiotape, $24.89.

Copyright law as it pertains to computer use in libraries is a topic of current interest to any library now in the business of providing online services to patrons or staff. The growth of computer information access by both librarians and patrons in all types of libraries has increased interest in copyright issues as applied to many different types of computer applications, such as downloading information, duplicating software, protecting library-generated databases, and identifying misuse in the library.

Computer/Copyright Seminar is intended for librarians and administrators providing computer access in either public or staff areas. The fourth edition provides the latest information on copyright and may be purchased on an annual standing order if a library chooses to have this type of information on hand each year. The 1989 edition was taped on February 17, 1989. The tape is twenty-six minutes long and is accompanied by text. The publisher, Copyright Information Services, estimates the time to complete the seminar to be 60 minutes. The publication is housed in a binder that can be shelved, with the tape stored on one side of the binder and the text filed in a pocket.

This publication is a combined format that explains the current copyright law for computer software. The combined format is successfully used to give the user both textual and audio explanations of the current law. Millers covers fourteen topics of interest to libraries, ranging from simple operations such as legally producing backup copies to the more complicated implications of software distribution through local area networks. The publication consists of one audiocassette of a seminar workshop; *Straight Facts*, a collection of papers distributed at the workshop; and *Official Fair-Use Guidelines: Complete Texts of Four Official Documents Arranged for Use by Educators*, third edition.

The text portions of the *Computer/Copyright Seminar* provide an explanation of the various topics discussed on the tape. Miller tells the listener when to stop the tape and read textual materials in order to better understand the explanation on the tape. The topics are presented in clear, precise language providing important information to the user in layman's terms. Miller stresses that the text portions are not suitable to be used alone and that they should be used with the audiotape. The *Guidelines*, however, can be purchased sepa-



rately from Copyright Information Services.

Miller provides examples of library use and misuse of computer programs and databases. The federal enforcement of misuse is also described with library examples. The problem of applying the law to library settings is covered. Dividing the presentation into types of libraries is unnecessary; the information is valuable to any library, reading room, computer lab, or other area providing access to computer data.

Straight Facts, a compilation of "Quick Fact Sheets," provides short excerpts of law and outlines the sections of a law that pertain to computers in libraries. Five of the six "Quick Fact Sheets" include a bibliography. In total, *Straight Facts* provides an additional four references, all published since 1987.

The eighteen-page *Official Fair-Use Guidelines* booklet provides three additional references. In addition, there is a list of other publications available from Copyright Information Services on copyright topics. Miller explains the use and meaning of the *Guidelines* in the *Computer/Copyright Seminar*.

I would hesitate to use *Computer/Copyright Seminar* in an actual workshop format without someone to conduct who is comfortable discussing law. If a workshop format is not used, the materials are easily read alone at the reader's own pace. This publication could be routed to all who are interested and then used to plan local policy. This is not to imply that effective policy be written using the *Computer/Copyright Seminar* alone. Policy writing is discussed in only one section. Writing a policy involves research on policy writing plus literature on copyright applications. In addition, it will be necessary to research the rights of the institution and its current use of computers and software to write a complete policy statement.

I recommend this publication for any library now involved or planning to be involved in computer software/database applications. I also recommend it as an informative, current publication on computer copyright law for any collection that is purchasing materials on the subject. It serves as a tool on copyright law as well as an example of successful combined format publishing.—*Carol M. Kelley, the University of Texas at El Paso.* ■■

Other Recent Receipts

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

Artificial Intelligence Abstracts Annual, 1988. V.5. New York: Bowker, 1988, 422p. $395 (ISBN 0-8352-2639-5, ISSN 0000-1244).

Beaumont, Jane, and Joseph P. Cox. *Retrospective Conversion: A Practical Guide for Libraries*. Computers in Libraries, no.7. London and Westport, Conn.: Meckler, 1989. 198p. $35 (ISBN 0-88736-352-0).

Becker, M., R. Haberfellner, and G. Liebetrau. *Electronic Data Processing in Practice: A Handbook for Users. Ellis Horwood Books in Information Technology. Chichester, England: Ellis Horwood, 1989. 615p. $49.95 (ISBN 0-470-21402-3).* [Translation of *EDV-Wissen for Anwender—Ein Handbuch für die Praxis*, 7th ed.]

Bibliographic Instruction and Computer Database Searching. Ed. by Teresa B. Mensching and Keith J. Stanger. Library Orientation Series, no.17. Ann Arbor, Mich.: Pierian, 1988. 173p. paper, $30 (ISBN 0-87650-251-6).

Books and Periodicals Online: A Guide to Publication Contents of Business and Legal Databases, Jan. 1989, v.2, no.1. Ed. by Nuchine Nobari. New York: Books and Periodicals Online, 1989. 434p. paper, $135 (ISSN 0951-838X).

Changes to the Anglo-American Cataloguing Rules, Second Edition, *As Published in the 1988 Revision*. Comp. by Edward Swanson. Lake Crystal, Minn.: Soldier Creek, 1989. 164p. paper, $20 (ISBN 0-936996-35-8).

Computer-Based Group Communication: The AMIGO Activity Model. Ed. by Uta Pankoke-Babatz. Ellis Horwood Books in Information Technology. Chichester, England: Ellis Horwood, 1989. 320p. $74.95 (ISBN 0-470-21461-9).

Defining and Applying Effective Teaching Strategies for Library Instruction: Papers Presented at the Fourteenth Library Instruction Conference Held at Ohio State University, 7–8 May 1987. Ed. by Mary Beth Bunge and Teresa B. Mensching. Library Orientation Series, no.18. Ann Arbor, Mich.: Pierian, 1989. 133p. paper, $30 (ISBN 0-87650-252-4).

Gateway Software and Natural Language Interfaces: Options for Online Searching. Ed. by James A. Benson and Bella Hass Weinberg. Library Hi Tech Special Studies, no.2. Ann Arbor, Mich.: Pierian, 1988. 204p. paper, $45 (ISBN 0-87654-213-3).

Intellectual Property Issues in the Library Network Context: Proceedings of the Library of Congress Network Advisory Committee Meeting March 23–25, 1988. Network Planning Paper, no. 17. Washington, D.C.: Library of Congress, Network Development and MARC Standards Office, 1989. 98p. paper, $7.50 (ISBN 0-8444-0636-8, ISSN 0160-9742).

Library of Congress Subject Headings: Significant Changes 1974–1988. Comp. by Teresa Hensley Burgett and Catherine W. Roberts. Lake Crystal, Minn.: Soldier Creek, 1988. 74p. paper, $15 (ISBN 0-936996-33-1).

Management Issues in the Networking Environment. Ed. by Edward R. Johnson. New York and London: Haworth, 1988. 141p. $24.95 (ISBN 0-86656-692-9).

Miller, Rosalind E., and Jane C. Terwillegar. *Commonsense Cataloging: A Cataloger's Manual*. 4th ed. New York: Wilson, 1989. 179p. $38 (ISBN 0-8242-0776-9).

Mumford, Enid, and W. Bruce McDonald. *XSEL's Progress: The Continuing Journey of an Expert System*. John Wiley Information Systems Series. Chichester, England: Wiley, 1989. 241p. (ISBN 0-471-92322-2).

Nauratil, Marcia J. *The Alienated Librarian*. New Directions in Information Management, no.20. New York: Greenwood, 1989. 129p. $35 (ISBN 0-313-25996-8).

Olson, Kent C., and Robert C. Berring. *Practical Approaches to Legal Research*. New York and London: Haworth, 1988. 150p. $29.95 (ISBN 0-86656-253-3).

O'Neill, Margaret, and Anne Morris. *The Role of Information Scientists in the Development of Expert Systems: Interim Report*. British Library Research Paper Series, no.57. London: British Library, 1989. 97p. paper (ISBN 0-7123-3190-5).

The Online Catalogue: Developments and Directions. Ed. by Charles R. Hildreth. London: Library Association, 1989. 232p. (ISBN 0-85365-708-4).

Penn, Ira A., and others. *Records Management Handbook*. Aldershot, England, and Brookfield, Vt.: Gower, 1989. 249p. $58.95 (ISBN 0-566-05666-6).

Proceedings of the Conference on Application of Scanning Methodologies in Libraries; November 17–18, 1988, National Agricultural Library, Beltsville, Maryland. Ed. by Donald L. Blamberg, Carol L. Dowling, and Claudia V. Weston. Beltsville, Md.: National Agricultural Library, 1989. 140p. paper.

Ravden, Susannah J., and Graham I. Johnson. *Evaluating Usability of Human-Computer Interfaces: A Practical Method*. Ellis Horwood Books in Information Technology. Chichester, England: Ellis Horwood, 1989. 126p. $37.95 (ISBN 0-470-21496-1).

Researching and Teaching Diverse Library User Groups: Papers Presented at the Sixteenth National LOEX Library Instruction Conference Held at Bowling Green State University, 5–6 May 1988. Ed. by Teresa B. Mensching. Library Orientation Series, no.19. Ann Arbor, Mich.: Pierian, 1989. 165p. paper, $30 (ISBN 0-87650-258-3).

Ross, Catherine Sheldrick, and Patricia Dewdney. *Communicating Professionally: A How-To-Do-It Manual for Library Applications*. How-To-Do-It Manuals for Libraries, no.3. New York and London: Neal Schuman, 1989. 293p. paper, $35 (ISBN 1-55570-031-4).

Saffady, William. *Introduction to Automation for Librarians*. 2d ed. Chicago and London: American Library Assn., 1989. 363p. $40 (ISBN 0-8389-0503-X).

Software for Optical Storage. Ed. by Brian A. Berg and Judith Paris Roth. Supplements to *Optical Information Systems*, v.5. Westport, Conn., and London: Meckler, 1989. 230p. $47.50 (ISBN 0-88736-379-2).

Teaching the Online Catalog User: Papers and Work Session Notes Presented at the Second Biennial LOEX . . . Workshop . . ., 9–10 May 1985. Ed. by Carolyn A. Kirkendall. Library Orientation Series, no.16. Ann Arbor, Mich.: Pierian, 1988. 256p. paper, $30 (ISBN 0-87650-250-8).

Telecommunications Abstracts Annual, 1988. V.4. New York: Bowker, 1988. 449p. $395 (ISBN 0-8352-2647-6, ISSN 0000-1252).

Who's Who in Technology. 6th ed. Ed. by Amy L. Unterburger. Detroit, Mich.: Gale, 1989. 2V. 2742p. $380/set (ISBN 0-8103-4950-7, ISSN 0877-5901). ■■

Correction

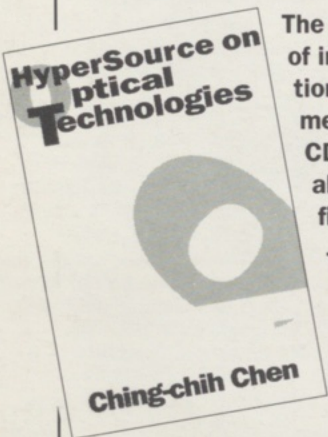
Authorship of the software review of BibBase/ACQ, which appeared in the June 1989 issue of *Information Technology and Libraries,* was incorrectly attributed to David Buxton. The correct author is Dave Ritchie, Head of Cataloging and Ordering at the State University of New York College at Cortland, New York. The editor regrets the error.

Information Technology and Libraries Index to Volume 8 (1989)

Compiled by Kieth C. Wright

HyperSource on Optical Technologies

by Ching-chih Chen



The most comprehensive and up-to-date source of information on the development and applications of optical technology and interactive media, including interactive videodisc, CD-ROM, CD-V, CD-IV, CD-I, DVI, WORM, erasable disc, high-density optical card, optical film, digital paper, and hypertext/hypermedia.

This indispensable tool makes available to librarians and information specialists the resources utilized by Dr. Chen in her considerable research, development and application work related to these technologies.

HyperSource on Optical Technologies, **a comprehensive print-based publication containing over 3,000 citations not readily found in library literature, will form a base which will link to other potential hypertext/hypermedia publications on the subject. This will include a hypermedia product, "HyperIntro to CD-ROM Technology & Products," on floppy disks for the Macintosh, expected to be released concurrently with this book.**

$29.50 pbk. Approx. 275p. ISBN: 0-8389-7359-0 Oct 1989

Library and Information Technology Association

a division of the American Library Association
50 East Huron Street • Chicago, IL 60611
Voice 312.944.6780 • Fax 312.440.9374

Tollfree 1.800.545.2433
In Illinois 1.800.545.2444