ARTICLES
Marion Shemberg 179 From Citation to Piece in Hand: The Search for Efficiency in Accessing Monographic Series

David Miller 190 Identical in Appearance but Not in Actuality: Headings Shared by a Subject-Access and a Form/Genre Access Authority List

Deborah A. Mohr and Anita Schuneman 205 Changing Roles: Original Cataloging by Paraprofessionals in ARL Libraries

Kathlin L. Ray and Mary S. Lang 219 Analyzing Search Styles of Patrons and Staff: A Replicative Study of Two University Libraries

NOTES ON OPERATIONS
Lynne M. Hayman 236 Database Design for Preservation Project Management: The California Newspaper Project

Linda Groom 254 Converting Wade-Giles Cataloging to Pinyin: The Development and Implementation of a Conversion Program for the Australian National CJK Service

Zahiruddun Khurshid 264 Cooperative Cataloging: Prospects and Problems for Libraries in Saudi Arabia

FEATURES
273 Letters

Robert M. Hiatt 274 Celebrating C. Sumner Spalding

275 Index to Advertisers
Human Resources Strategies in Information Services Management
by Meg Paul

This hands-on guide to human resource management looks at the most effective strategies for motivating employees and maintaining a successful information center in the age of the virtual library. How will the information center change? How will database searching, electronic ordering, and document delivery impact the industry? How can information services professionals like you use your expertise to ensure your library's success?

1998 ■ 1-85739-118-7 ■ c. 200 pp. ■ $50

The Corporate Memory: Information Management in the Electronic Age
by Ken Megill

Modern technology has created a new challenge in the area of document retention: How do you decide which records stay and which ones go? Managing corporate memory — that is, all the active and historical information in an organization that is worth preserving — is key to effective operations. This comprehensive sourcebook provides a coherent plan for integrating and managing records of all types — print, microform, and magnetic media — in the electronic age.

1997 ■ 1-85739-158-6 ■ 131 pp. ■ $50

Order any or all of these indispensable volumes today
by calling 1-888-BOWKER2 (1-888-269-5372).
Or fax your order to 1-908-508-7696.
DDC®21
Keeping Pace with Knowledge

The Books

- Major revisions in 350-354 Public administration, 370 Education, and 560-590 Life sciences
- Revised table for countries of the former Soviet Union
- Updated terminology and new topics: rap music, virtual reality, in-line skating, and snowboarding
- Over 150 pages added to the Relative Index
- Manual expanded by over 100 pages and indexed in the Relative Index


The Compact Disc

- DDC 21 in a Microsoft Windows environment
- LAN compatible to support multiple users
- More ways to search and display Dewey
- Multiple records from anywhere in DDC displayed on a single screen
- Drag and drop information between windows
- Annotation feature to reflect local classification decisions
- Over 4,000 new entries in the electronic index

Dewey for Windows. Compact Disc with documentation. 1996. Single workstation $400.00; annual replacement disc $150.00 per year. Local area network or multi-copy site license $500.00 per year (includes annual replacement discs); each additional compact disc $20.00 per year.

The Guide

Written by classification experts, the Practical Guide explains how to apply DDC schedules, and assign and build class numbers. Completely revised for DDC 21. Includes exercises and answers designed to reinforce, through practice, the examples and explanations given in the text.


To order, write OCLC Forest Press, 6565 Frantz Road, Dublin, OH 43017-3395
or fax your order toll-free to: 1-888-DEWEY21
We accept MasterCard and VISA
Windows is a registered trademark of Microsoft Corporation in the U.S. and a trademark in other countries.
From Citation to Piece in Hand: The Search for Efficiency in Accessing Monographic Series

Marian Shemberg

The author of this article looks at access problems that occur when the citations from indexes for articles in series differ from the way the items are cataloged. To illustrate the difficulties encountered, a search was conducted in the online catalogs of various libraries and a comparison made between the bibliographic entries found to citations both in electronic indexes and in authors’ references. The results indicate that the ramifications of local cataloging decisions affect all areas of librarianship. In this electronic age, the need is great to integrate the various sources of information needed for a patron to go from citation to piece in hand.

Beginning with Poole’s original *Index to Periodicals* in 1848, libraries, of sheer necessity, have outsourced indexing of articles in journals to producers of databases. Tyckoson (1989, 10) states that “the failure of the catalog to incorporate the growing number of journals and other special formats has resulted in an entire industry of commercial services designed to supplement the catalog in this area.” Yet until recently, the catalog was not designed to assist the patron in the transition from an indexing citation to the bibliographic record in order to locate and retrieve the desired item. There has been limited standardization or coordination between the abstracting and library worlds.

The use of the catalog as a finding aid was addressed by Tate (1963), who compared main entries in the Library of Congress (LC) card catalog with citations in references. Relevant to the present study, Tate (§190) also compared the citations “with the corresponding catalog entries for data regarding useful secondary entries.” Though only peripheral to her study, she (p.191) did point out that one type of “secondary entry that may have some utility as a finding device is the series entry.” Tate also warned that secondary entries vary over time and among libraries, with the needs of the individual library taking precedence over standardization. Even at that time, she mentioned the problems of growth of the catalog and the expense to libraries in adding secondary entries. It appears that more than thirty years later, libraries are still grappling with these issues.

Only one other article (Tagliacozzo, Rosenberg, and Kochen 1970) was found in which the researchers examined the information that patrons brought to the catalog in searching for known items. They did not focus on series but rather directed their attention toward the search...
process and patrons' ability to find the item sought. They state (p. 231): "the catalogue has to contain at least one written record representing the item . . . that the user is trying to find; the user has to find one such record; and he has to recognize it as corresponding to his mental representation of the object of his search." It is this last aspect that is examined in the present research.

The difficulties that patrons face when they go from known-item series entries to the catalog is an area in which little has been written. No further research studies were found that related bibliographic citations of series to finding these items in the catalog. In this article, I review access problems that face the catalog user as a result of differences in the way an item in a series is represented in various forms in indexes and in the catalog.

It is acknowledged that cataloging monographic series is complex. Authors have raised this topic in the cataloging literature as an ongoing issue. Steinhagen (1989) enumerates some of the complicated issues with which catalogers have struggled cataloging serials and series. McBride (1985) gives a history of the attempts to deal with monographic series.

However, catalogers are not the only ones struggling with this issue. Representation of an item in the catalog, especially a serial or series, not only affects the patron; it affects all areas of librarianship. In this article, I review the issue from the perspective of the reference librarian and try to identify some options to meet the different needs of publishers, indexers, and librarians in light of the state of technology today.

The Issues

To illustrate the problems of access that a user might encounter in moving from a citation to a piece in hand, we will look at a specific series that has kept the same title, the same ISSN, and the same editors since 1973, yet is still elusive to identify. The title of this series is New Directions for Community Colleges (NDCC), a series published by Jossey-Bass that is important in the field of education and is used heavily at Ohio State University (OSU). With continuity of title, ISSN, series editors, and place of publication, one would think it would be easy to identify individual pieces. However, reference librarians at the Education/Psychology Library at OSU have been aware of problems in accessing this series for years. It is especially troublesome because of its double numbering system, and because at OSU it had been cataloged as an unanalyzed serial until 1989. At that time, the cataloging treatment was changed to individual monographic records, to provide better access.

It seems reasonable to posit that access is facilitated if an item is cataloged in the same way in which it is indexed in databases and used in references or bibliographies. This hypothesis is based on the ideas of standardization and consistency. Crawford (1987 and 1992) emphasizes that a catalog should be clear and consistent. He states (1987, 68) that "usage should be consistent from a patron's view, not from a designer's view." Consistency is also beneficial when moving from an electronic index to the online catalog. Snavely and Clark (1996, 50) point out that the "process, from having the citation in hand to having the article in hand, has been largely overlooked in the literature." Awareness of problems in retrieving items from citations is a first step to improving success in retrieval.

Problems in identifying the example series center around the following factors:

1. Inconsistency in cataloging treatments due to human error
2. Dual numbering systems on the source
3. Local decisions over time and in multiple branch libraries to catalog item as a serial, a monograph in a series, or both
4. Whether or not the field containing the series entry is indexed
5. Citation differences in electronic indexes
6. Ways in which articles are cited in references
7. Screen display

Although these factors refer to one title in a specific library, the same factors
can be found in part in any larger library, especially those with large serial holdings.

**Methodology**

Each issue of *NDCC*, beginning with number 1 (1973), was examined for citation information, numbering systems, and changes in description on the title page and its verso.

The online catalog at OSU (OSCAR) was checked to see how each issue was handled over time. OSU belongs to the OhioLINK consortium, which is a statewide consortium linking academic, community college, and research libraries in the state of Ohio. OhioLINK's computerized union catalog is accessible through the Internet. The title was also searched in the central OhioLINK catalog and individually in the catalogs of member libraries. When there were gaps in issue holdings at member institutions, each catalog was checked by individual issue title to determine whether the item was owned but either not identified or not indexed by series title. Of the 30 libraries in the system at the time of this writing, 17 libraries had at least 10 issues of *NDCC*.

When it was discovered that cataloging practices of member libraries varied greatly, catalog records for this title were also examined at large research libraries outside the consortium via a Z39.50 interface to see whether large research institutions cataloged these items at a more consistent level.

The libraries that were searched included LC and four members of the Association of Research Libraries (ARL) outside the state of Ohio. When the Z39.50 search interface being used to access other catalogs proved too restrictive at LC and one ARL library, those catalogs were searched directly via the Internet using the remote catalog's native software. These methods were used to identify consistency and differences in cataloging practices of the sample of large research libraries. Then comparisons were made across all catalogs on the treatment of this series.

Many patrons get citations to articles in the *NDCC* by way of electronic indexes. So the form of citation in several electronic databases that index this series were searched by series title to determine the form of the citations found there. The indexes used were ERIC, Education Index, ArticleFirst, and CARL UnCover. Particular attention was paid to how the citations meshed with bibliographic records in the various online catalogs.

Another patron source for *NDCC* citations is bibliographic references in the education literature. This was examined through a Dialog search of Social Sciences Citation Index (SSCI). The first search of SSCI was for the series title using “New Directions Community” as the search terms; 120 references to *NDCC* were found for which copies of the actual articles could be obtained. Because the series title was used to retrieve the references, it seems logical that all 120 matches contained the series title in the citation. This was true, with some variations that could also affect retrieval in a local catalog: 109 had the correct title; 10 had *New Directions in Community Colleges* instead of “for,” and one left off the final *s* from “Colleges.”

Finally, multiple access points were compared among the databases—the citations in references and the various catalogs. These included series title, issue title, sequential number, volume and issue numbers, ISSN, and dates. The time period of the study was September 1995 through June 1996.

**Findings**

**How the Publisher Identifies These Items**

*NDCC* has been published quarterly since 1973. Each issue is identified both by a sequential number and by volume and issue numbers. Each issue also has a unique title in addition to the series title, as well as issue editors and series editors. From its inception in 1973 until mid-1981, each of these issues had two separate title pages. The first is the series title page, which includes the editors of the series, the sequential number, and the date. The second title page gives only the unique issue title, the editors of that
theme issue, and the publisher. Additionally, the verso of the second title page lists a volume number in roman numerals and an issue number, different from the sequential numbering of the series. The sequential number is not repeated on the verso from 1973 to 1982. In 1983, the sequential number was added to the verso of the title page. Also, ongoing from issue number 5 (1974) through number 68 (1989), a supplementary note on the verso of the title page states, "New Directions is numbered sequentially. . . . The volume and issue numbers above are included for the convenience of libraries" (NDCC 1974, title page verso).

The date designations on the issues of NDCC also vary. Issues number 1, spring 1973, through number 22, summer 1978, have dates of spring, summer, autumn, or winter, plus the year. Then only the year is indicated on the title page until issue number 35, where the date changes to the month and year, e.g., September 1981. Each issue then gives a month and year through issue number 54, June 1986. From issue number 55 to the present, the dates are fall, winter, spring, or summer and the year. Again, although these variations refer specifically to NDCC, any public or technical service librarian (and some sophisticated patrons) know this problem is by no means unique.

**BIBLIOGRAPHIC RECORDS IN ONLINE CATALOGS**

Various libraries have cataloged this title as an unanalyzed serial, as an analyzed serial, or as a monograph in a series, both with series entries indexed and not indexed. LC's series authority record for the series indicates that it is considered a monographic series (MARC field 646) for issues 1–45 and a serial for number 46 on. For dates of publication and volume designations, the record shows number 1 (spring 1973). Therefore, from 1973 to 1984, issues number 1 through number 45, LC separately cataloged this title as a monograph with a series entry. A MARC 515 note (numbering peculiarities) states that issues from number 46 are also called volume 12, number 2, and so on. All of the issues have volume and issue numbers in roman numerals on the verso of the title page, yet the authority record indicates only that a dual numbering system is used from number 46, which is when LC started cataloging NDCC as a serial. Therefore, patrons checking LOCIS, the online catalog at LC, will find separate records for each of the first 45 issues in the monographs catalog, but will find one serial record in the serial catalog by series title only. As Hagler and Simmons (1982, 105) state, "It should be no surprise that the policies of the Library of Congress, like those of any other library, change from time to time." Most information in a bibliographic record is taken from the item itself; in the case of this 515 note, it appears to indicate local practice treatment (monograph series versus serial), not the history of the item.

In monographic records of this title, librarians have either used the sequential number, the volume and issue numbers, or a combination of the two. One way librarians identify the dual numbering system is to trace by one number but to indicate in a series field or a note field an alternate numbering system. For example, number 37 is identified as "number 37=v10, number 1 (Mar. 1982)." This redundancy is a great service to patrons and public service librarians alike, who might bring only partial information to the catalog. The note information might not be an access point in all catalogs; however, it may ease identification of a particular issue once the bibliographic record is retrieved.

Another way that librarians identify the dual numbering systems is with notes indicating that the series is also numbered by a certain volume and issue on the verso of the title page. Occasionally, due to errors, the identifying issue numbers are omitted. This makes access and identification extremely difficult for the patron.

The addition of the Blackwell North America Table of Contents service at OhioLINK has improved identification and retrieval of NDCC articles. Many monographic records in this series include notes or contents that identify chapter titles and chapter authors. These addi-
TABLE 1

<table>
<thead>
<tr>
<th>Number of Libraries</th>
<th>Cataloged as Serial</th>
<th>Cataloged as Mono in Series</th>
<th>Cataloged Both as Mono and Serial</th>
<th>Volume/Issue Number Traced</th>
<th>Sequential Number Traced</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>X (analyzed)</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Citations to the records are retrievable by keyword search and by title or author if “Table of Contents” is a feature of the search interface. Such additions help enormously in creating an efficient and effective system for accessing this type of publication.

BIBLIOGRAPHIC RECORDS IN A UNION CATALOG

All member libraries of OhioLINK use the Innovative Interfaces system (INNOPAC). Patrons can access the central union catalog or any of the individual member catalogs through the Internet.

The central union catalog has 97 titles cataloged under NDCC. All are monographs in a series except for one serial record. The screen displays make identification of the serial record difficult. Cataloging an item as part of a serial makes “one record that will stand for all the individual issues of a serial title, with little ongoing maintenance” (Inner 1990, 4). Cataloging series as an unanalyzed serial is convenient for processing purposes, but users might need greater expertise to recognize a serial record and interpret issue holdings attached to the record. In OhioLINK, full bibliographic records retrieved for the monographic entries have a unique issue title and issue editor. Each title is identified as part of the series. In the central union catalog, the series field indicates the sequential number. Most records do not identify the volume and issue number. Some records, as in the LC records referred to above, show both the sequential number and the volume and issue numbers, but the series is indexed by the sequential number. If the local system sorts numbered issues separately from volumes, an added complication for patrons is introduced. Table 1 provides a summary of citations to NDCC retrieved from OhioLINK libraries.

SEARCHING OTHER ONLINE CATALOGS

LC and four ARL libraries were also searched. However, access to catalog information via a Z39.50 interface can vary. Holdings, in most cases, were not indicated, which prevents identification of items in a serial record.

The first ARL library was searched by series title. The first record retrieved was the serial authority record. In a 515 field is the note “Volume numbering ceased with the December 1982 issue.” This seems in conflict with the 515 note from LC, which indicates that, from number 46, items are also called volume 12, number 2, and so on. This note reflects local information, not what indexes may cite as being on the actual item. Up through number 40, some records are identified by the sequential number, while others are identified by volume and issue numbers. Starting with number 41 (1983), the rest of the monograph records use the sequential number. The second and third libraries treat this series both as a serial and as a monograph with a series added entry. This dual treatment occurs because one branch or departmental library may catalog the series as monographic, while another catalogs it as a serial. One might be analyzed, while another might not.

Searching the last library via Z39.50
TABLE 2
HOW LC AND ARL LIBRARIES CATALOG THIS SERIES

<table>
<thead>
<tr>
<th>Library</th>
<th>Cataloged both as Mono and Serial</th>
<th>Cataloged as Mono in Series</th>
<th>Traced by Sequential Number</th>
<th>Traced by Volume/Issue Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>vl1n1</td>
</tr>
<tr>
<td>Library A</td>
<td>X (analyzed)</td>
<td>X</td>
<td>X</td>
<td>vl1n1</td>
</tr>
<tr>
<td>Library B</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Library C</td>
<td>X (analyzed)</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Library D</td>
<td>X (analyzed)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

was not as successful. Three serial records were located, but no monographic records were retrieved when searching by series title. A search by each issue title was then performed. This search yielded all issues of the series. It was found that the remote search interface provided for a search by series, separate from a title search. This search produced the monographic records. In this library's catalog, the series cannot be searched as a title, and the Z39.50 interface does not have a series search option.

At LC, the Z39.50 search interface allowed for searches only of the book catalog. The first 45 issues were retrieved. To find the serial record at LC, LOCIS was accessed through the Internet. A summary of the category treatment by LC and the ARL libraries is shown in table 2.

RESULTS

What becomes clear after searching all of these catalogs is that standards vary or are unclear, especially to the remote user. Local practice determines how this series will be cataloged. One branch of a library might catalog it as a serial and another as a monograph; one might index with the volume and issue numbers, another might index by the sequential number. Two or three records for different copies of the same issue might be present in the catalog. Few libraries used one bibliographic record and attached all holdings to that one record. This proved to be a major impediment to access and retrieval, an impediment that has significance for reference, interlibrary loan, acquisitions, cataloging, and binding. Consistency and compromise are needed in cataloging centers at each university or research center.

It would be simpler if there were a national policy on series identification and citation that was followed by publishers, abstractors, and libraries. LC has provided much guidance in the past. Yet Hagler and Simmons (1982, 105) point out that, "the policies of the Library of Congress concerning such matters as series cataloguing, the provision of analytics and contents notes, etc., are those best suited to its own particular needs as a large research library." However, most libraries use shared catalog records often derived from LC copy.

MONOGRAPHIC TREATMENT

A number of observations can be made with regard to monographic treatment of this series.

- OhioLINK member libraries, LC, and one ARL library that catalog this series as a monograph (or the monograph records of the analyzed serial) display the records in sequential number order.

- The other ARL libraries that catalog this series as a monograph (or as an analyzed serial) display the monographic records in alphabetical order by unique issue title or main entry. Some records appear in the multiple title list by series title, some by issue title, and one appears by the editor of the issue, depending on the main entry. This arrangement has two draw-
TABLE 3
HOW NDCC IS CITED IN INDEXES

<table>
<thead>
<tr>
<th>Index</th>
<th>Series Title Shown</th>
<th>Issue Title Shown</th>
<th>Sequential Nos. Shown</th>
<th>Volume/Issue Nos. Shown</th>
<th>Issue Editor(s) Shown</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC journals</td>
<td>X</td>
<td></td>
<td>81, 82, 84, 85</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ERIC Fiche Issues</td>
<td>X</td>
<td>X</td>
<td>With exceptions</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Education Index articles</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Education Index issues</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ArticleFirst</td>
<td>X</td>
<td>Some issues</td>
<td>X</td>
<td>Volume only</td>
<td></td>
</tr>
<tr>
<td>CARL UnCover</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

backs. One is that the issue titles are typically not present in records of the electronic indexing services. The second is that the sequential number and volume number may not be displayed in the local catalog's multiple title screen.

- In addition to inconsistencies among catalogs, some show internal variations. Monographs are sometimes indexed with the sequential number and other times with the volume and issue numbers in the same library. At times, these “missing” items are owned but are cataloged by volume and issue numbers, so are located further down in the multiple display screen. Other times they are not owned. It is difficult for the patron to determine which is the case.

- Records of some issues are not indexed by the series, though they may have a series statement. They are not retrieved by series title.

- Errors occur. For example, issue number 2 at one library has two separate item records attached to it. One is the correct issue number 2 (1973) Meeting the Financial Crisis. The other is number 34 (volume 9, no. 2), Women in Community Colleges. There is no separate title entry for issue number 34 for that library.

- The number of the issue is not indicated. By searching for these monographs by known issue title, the sequential number was ascertained.

SERIAL TREATMENT

A number of observations can be made with regard to serial treatment of this series.

- Some holdings in the serial records are identified by volumes while others are shown as numbers. Often the sequential number is incorrectly identified by the volume label. At times, no identifying number is included.

- A serial record may exist, but only one issue is listed as a holding. The rest of the issues are cataloged as monographs. Or isolated issues are cataloged as monographs, while the rest of the collection is cataloged as a serial.

- Libraries catalog these items both as a serial and as monographs in a series. Sometimes the same issues are cataloged in two separate ways by different catalog centers at the same institution. Dual cataloging may be helpful to patrons or may be confusing.

Consistency is a feature that patrons assume about an online catalog, and when that consistency fails to materialize, patrons are poorly served. However, citations may also add to patron confusion.

CITATIONS FROM INDEXES

Various indexes treat this title differently. Each treatment is described below; an outline is provided in table 3.

ERIC. ERIC has the longest history of citing NDCC, having cited it continuously.
TABLE 4

<table>
<thead>
<tr>
<th>SSCI Series Title Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=120</td>
</tr>
<tr>
<td>Series title cited</td>
</tr>
<tr>
<td>Issue title cited</td>
</tr>
<tr>
<td>Sequential number cited</td>
</tr>
<tr>
<td>Volume number cited</td>
</tr>
<tr>
<td>Issue number with volume</td>
</tr>
</tbody>
</table>

since 1973. Until 1983, it was the only service indexing this series. ERIC citations to NDCC are consistent over time—they include the chapter author and chapter title, series title, volume and issue numbers, and date. When the ERIC microfiche format of whole issues of this series is cited, the sequential number is usually added to the title entry.

Through 1993, ERIC citations continued to be by author and title of chapter, with the journal entry giving the volume and issue numbers and the date. However, in 1993, a note was added, giving the sequential numbers. Issues number 84 and number 85 also include the issue title in the Note field. However, this ceased after issue number 85, winter 1994. The issue number is no longer furnished. Given that most libraries in this sample cataloged the series by sequential number, this omission would make it very difficult for patrons to make the leap from the ERIC index to items owned by libraries.

**Education Index.** Education Index has cited NDCC from 1983 (number 42) to the present. All of these issues are cited by sequential number only. In addition to indexing the individual articles from each issue, Education Index also cites the entire issue, giving the issue titles and the issue editors. No volume and issue numbering is provided. The format of this index coincides with most of the cataloging practices at the libraries studied.

**Carl UnCover.** Carl UnCover indexes NDCC back to number 65, spring 1989, by sequential number only. Each full record includes author and chapter title, series title, season, year, and sequential number. No issue titles or issue editors are given.

**ArticleFirst.** ArticleFirst started indexing NDCC in the summer of 1992. It indexes by chapter author and title, series title, volume number, issue number, and date. The volume number listed varies between arabic numbers and roman numerals. The issue number is actually the sequential number. Thus, ArticleFirst combines the two numbering systems. Sometimes an issue description is included, the description being the unique issue title. However, that field is not searchable by title or keyword. In the fall of 1995, issue number 91 is identified by sequential number only. The volume and issue numbers are absent. With issue number 92, the unique issue title is omitted.

**Citations from Social Sciences Citation Index**

The Social Sciences Citation Index (SSCI) was examined for treatment of this title. The treatments are described below; an outline is provided in table 4.

An online search through Dialog of SSCI yielded the following results. Seventy-four (68%) of the citations listed the issue title, often including the editor of the issue. Ninety-one (76%) gave the sequential number, while 24 (20%) listed the volume number. Of those citing the volume number, 9 also included the issue number. Five gave no identifying number, only the year. One author gave the sequential number but called it volume 77 (1992). Most used only the numbers and year, not indicating whether it is the volume or the sequential number.

A second online search of SSCI through Dialog was done by author and year for the 5 most cited authors. This
search produced 17 citations; all were cited by series title and sequential number.

**Implications for Libraries**

How can finding known items in catalogs that were cited in indexes be made more efficient? What are our options? Though solutions are difficult to implement and may be prohibitively costly, it is productive to discuss possible resolutions. Because the catalog has the greatest impact on access for patrons, it would be productive if the user were kept in mind when records were created and decisions were made about numbered monographic series.

The first implication deals with the design of catalogs and decisions made about treatment of series. It should be an explicit goal of catalogs to provide an easy transition to the catalog from bibliographic indexes in which library titles are indexed.

Reference librarians can help with their knowledge of patron behavior. Because each series title is unique, this knowledge usually comes after the fact when referring to a specific title. When reference librarians see the types of difficulties patrons are experiencing, they should have a way of communicating effectively with catalogers to resolve the difficulties. At this stage, it means enhancing or changing a catalog record. Or it might mean a policy change.

Would access be eased if NDCC were cataloged as a serial or as a monographic series? On the basis of this study, when a series is well known by its series title and when it has a dual numbering system with the sequential number prominent, it should be accessible by series title with the identifying sequential number. If possible, both numbers should appear in the record.

The best choices for cataloging a series are as an analyzed serial or as a monograph in a series. There are advantages and disadvantages to each for the patron. Cataloging NDCC as an analyzed serial would give it a shared call number, and all issues at a library would be shelved together. In addition to the serial record, an individual bibliographic record would exist for each item. However, because some libraries restrict usage of their serials, and many libraries do not allow serials to be checked out, access would be curtailed. Patrons at OhioLINK libraries could not initiate borrowing of NDCC issues from a terminal using patron-initiated circulation modules. They would have to get a copy of an article through interlibrary loan, which entails time delays.

In some libraries, serial records are in a separate catalog. When libraries catalog a series both as a serial and as a monograph, how is one to know which catalog to search for a particular issue? It would be helpful to have cross-linking between these types of catalogs. The patron has already spent time searching indexes and now may need to decide which title to use, which number to use, and which catalog to access.

When libraries catalog these issues as an unanalyzed serial, searching by unique title is fruitless. However, if a series is cataloged as an analyzed serial, a user can retrieve entries either by unique issue title or by series title if these are indexed fields. If it is cataloged as a monograph in a series and it is indexed in the title field by series, it is again retrievable by either title. When it is not indexed by series, searching by series title is futile. To be helpful to patrons who get citations from indexes or references, numbered monographic series should be indexed by series. However, indexing series and analyzing serials may involve greater initial expense for libraries.

The second implication for cataloging concerns the numbering system. The primary source of citation information, the title page, clearly states the issue number and the date. Catalogers are aware that “some data elements have gradually drifted to locations other than the title page, particularly to the verso of the title leaf where it is now common to find date and edition information” (Hagler and Simmons 1982, 85), but they add that the title page is still the primary source of this information. Looking beyond the title page for further numbering is time-consuming. Libraries are looking to simplify cataloging, not to create complexity. Titles in numbered monographic series with dual numbering systems where the index-
ing service uses only the volume and issue numbers create complexity. Access by the series title, with the date and both numbering systems noted in the series field, might facilitate access. This is especially true when the dominant number is the sequential number and most authors are used to citing volume and issue numbers.

Repercussions for interlibrary loan (ILL) loom large. Seaman (1992) explains that ILL requests for books can be generated as a result not only of user error but by catalog failure for items actually owned by the requesting library. One category of catalog failure is that no catalog entry exists for series titles cataloged as serials. One serial entry for numerous issues of this series is common practice at many of the libraries studied. Seaman dealt with requests for ILL in error that were found by ILL staff and were identified as being owned by the borrowing library.

Two other situations make ILL more elusive. Through OhioLINK all registered users can borrow monographs by requesting them without ILL or staff intervention. This feature is called patron-initiated circulation. If an item is located in the central catalog, any member library user can initiate a request that will then be accepted by the system as valid when the same item is not owned or is unavailable at the user's home library. Where the cataloging treatment is the same, the system can easily make this determination. However, when one library has cataloged a title as a serial and another has it as a monographic series, the system cannot determine the sameness. If a monograph in a series is requested despite the fact that the user's host library owns the serial, the system will allow the request.

The second situation assumes that the staff of ILL can make all the leaps from volume number to sequential number and know that the item is actually owned but is cataloged differently. Can the ILL office interpret all of this information for monographic series with dual numbering schemes, given the variety of cataloging apparent at various libraries?

Acquisitions is also affected. When an item is not cataloged in the way it is being requested, libraries might purchase additional copies, thinking they do not own a particular issue. Some libraries in the study have extra copies of various issues of this title. These extra copies might be due to its popularity, or they might be due to mistakenly purchasing a copy when the catalog record for an issue could not be found in preorder searching.

In the public service area, public service librarians must be knowledgeable about these issues to teach patrons to be flexible, to question whether the citation in hand is complete, correct, or needs modification to find whether the item is owned at a particular library. We need to instill in our patrons the willingness to ask questions, to appear foolish in order to save hours of searching futilely. We need to learn about the problems facing serial catalogers and to inform them how patrons search for items.

**Conclusions**

Regardless of how many titles in series exist in academic library collections, it would seem that indexing the series entries and adding series entries would entail more work if more series were analyzed and dual numbering schemes were always recorded. However, from the perspective of efficient use of the catalog, this effort would reduce the work of users, many of whom are library staff in acquisitions, ILL, reference, and cataloging services. Reference librarians believe the time and effort would be well worth the value, not only in more accurate and efficient access but also in reduced frustration and greater user satisfaction with the catalog. The sharing of resources is highlighting historic inconsistencies and the lack of national standards. Retrospective conversion has been focused on journals with title changes. Now that many libraries have been able to update these records, it might be time to ease access for series.

A national policy on series cataloging appears to be evolving. In the *Cataloging Service Bulletin* (Library of Congress, 1996, 11), it is stated that: the Library of Congress Rule Interpretations (LCRIs) provide for a common practice, sometimes called 'national' practice...
Moreover, it is incumbent upon the Library of Congress to make decisions to provide common practice in order to share bibliographic and authority records for the common good, which may at times be at odds with local needs. In the case of series, the practice applies to LC and to those NACO libraries that contribute series authority records.

CSB includes LC practice guidelines for tracing analyzed series. LC is taking a leadership role in advancing the practice that series should be traced by series statement. Complying with the common practice is expected to result in more series entries and serial numbering in the catalog, but it will provide more efficient access. The trend to limit the number of access points to speed up the cataloging process may have had the unintended consequence of making it more difficult for patrons to access items from citations taken from electronic indexes. The vast resources of electronic indexes need to be supported by strengthening citation forms in the catalog.

Can librarians influence publishers? Publishers, first and foremost, want to remain in business and remain profitable. Librarians can inform them of difficulties their publishing practices are creating. Perhaps this awareness will open dialogue for greater understanding, even if, for example, it does not change a dual numbering system for a currently published journal. In addition, some journals include recommended forms for citation with the articles in each issue. Making this practice a model and integrating it with cataloging practices would bring clarity where there was confusion.

To standardize a catalog to agree with citation indexes would be impossible given the lack of standardization of these services. However, as more library catalogs link citation indexes to their bibliographic records, it would appear that a solution must be found for this linkage to succeed. Effort may need to be placed on system solutions to integrate index citations with the library’s catalog. As more companies seek solutions to these issues, it seems feasible that an algorithm will be developed to link the index citations to the correct entry in the library catalog. Having consistent and enhanced access points for series entries would make linking catalogs with similar access points in indexes useful and productive. Having standards for series entries that are followed by publishers and libraries that incorporate additional access points that are possible with electronic catalogs could make the link from citation to piece in hand a quick and efficient process for patrons and librarians alike.

**Works Cited**


New directions for community colleges. 1974. no. 5.


Identical in Appearance but Not in Actuality: Headings Shared by a Subject-Access and a Form/Genre Access Authority List

David Miller

Authority records were compared for established headings that are identical in Library of Congress Subject Headings (LCSH, 18th edition) and Moving-Image Materials: Genre Terms (MIM). First, the Use For, Broader Term, Narrower Term, and Related Term references in the LCSH authority file were compared with their counterparts in MIM, to determine the proportion of duplication existing between them. Fewer than 10% of these reference headings are identical. Second, a qualitative comparison was made of the "semantic spaces" inhabited by identical headings in different contexts: a general-purpose subject access list as compared with a medium-specific form/genre access list. It was found that, in many cases, headings that are identical as character strings have markedly different meanings in different contexts. The conclusion offered is that, both quantitatively and qualitatively, pairs of identical headings differ sufficiently from each other that the creation of authority records for each usage represents no duplication in any meaningful sense. The striking divergence, in many instances, between semantic spaces provides food for thought on the differences between the naming of subjects and of forms/genres.

At the American Library Association (ALA) Annual Conference in June 1995, the Machine-Readable Bibliographic Information Committee (MARBI) considered Proposal 95-11, "Definition of X55 Fields for Genre/Form Terms in the USMARC Authority Format" (Stockton-San Joaquin County Public Library 1995). MARBI, an interdivisional joint committee of ALA, is concerned with the creation, review, and development of standards for the presentation of bibliographic information in machine-readable form. Proposal 95-11 establishes authority heading fields (MARC fields 155, 455, 555, 755) corresponding to field 655 (Index Term—Genre/Form) of the USMARC Bibliographic Format (Library of Congress, Network Development and MARC Standards Office 1991, 126). The history of this issue, first raised in 1993 and pursued through two discussion papers and two
proposals, is covered in the text of Proposal 95-11. The driving force was the desire on the part of some librarians to create authority records for form/genre headings, as subject authority records are now. This would be the case even in instances where headings of each type, drawn from different thesauri, appeared to be identical, in order to allow for cross-references and notes that apply only to one thesaurus.

The new fields proposed were accepted as part of the USMARC Authority Format, but concern was expressed, in the proposal itself and in committee discussion, about the apparent duplication mandated by creating separate authority records for identical headings. In the proposal, the question was raised whether it would be preferable to create two authority records for the same term—one a topical heading (tagged as 150), the other a genre/form term (tagged as 155). It was also asked whether the duplication of identical terms would prove confusing to the user (Stockton-San Joaquin County Public Library 1995). The potential adverse consequences of duplication in indexing were mentioned in committee discussion, but these were not regarded as sufficient to prevent acceptance of the proposal (ALCTS/LITA/RASD 1995).

These speculations about potential indexing and display problems raise the question of whether separate authority records for the same heading, used for subject and form/genre access respectively are in fact duplicates. Besides the established headings, which are necessarily identical as character strings, what sort of parallelisms between such pairs of authority records would actually exist? This question, which can be addressed both quantitatively and qualitatively, was investigated by comparing identical headings drawn from the Library of Congress Subject Headings (LCSH) and Moving-Image Materials: Genre Terms (MIM).

LCSH is a well-known list of subject headings applied by catalogers in the United States and other countries, developed and managed by the Library of Congress (LC). MIM was developed by the National Moving Image Database Standards Committee of the National Center for Film and Video Preservation at the American Film Institute. The stated purpose of the list was “to standardize terms used to designate genres and forms of moving image materials” (Yee 1988, 11). It is a general list, best suited to collections that contain many different types of moving-image materials, and has been designed to apply across media. Headings from MIM are entered in USMARC bibliographic field 655.

**Background of the Study**

The present study is the third in a series of writings originating from the close comparison of MIM and LCSH. Miller (1992) examined the levels of compatibility between the lists, using the LC Subject Authority File to represent LCSH. The study was conducted in order to determine the potential for conflict or ambiguity if the two lists were used and indexed together in a general academic library catalog. It was determined that, while most of the approximately 190 MIM headings presented no conflicts with LCSH, either conceptually or as character strings, there were five categories of headings that presented potential problems of overlap and synonymy. These were (1) headings that might be synonymous with those in LCSH, (2) headings that are similar in wording to related headings in LCSH, (3) headings that are identical in form between the two lists, (4) headings that serve as “see” references in LCSH, and (5) homographs: headings identical in form but referring to different areas of knowledge.

Miller (1995) began with a set of LC subject headings drawn from the first three categories described above, that is, headings that might be considered synonymous, similar, or identical with those in MIM. The OCLC Online Union Catalog was searched for records that contained those subject headings in an attempt to determine whether they were used for subject or form/genre access and in what proportions. The major finding of that study was that this set of “subject” headings was used for form/genre access between 6% and 99% of the time, with a mean of just over 50%. In only a few instances
did authority record information mandate form/genre use. Nevertheless, it was clear that form/genre access to the cataloged materials was regarded as essential and was provided by catalogers. Subject headings from LCSH were therefore being pressed into service as a kind of moving-image form/genre thesaurus—however partially and inconsistently.

It is the third category of headings—the 27 headings that are identical in form between the two lists—that we consider in the present study. The entire population of these headings, which are applied to works used for the study of moving-image materials, therefore serves as a basis for the comparisons that follow (the list of 27 headings is included as appendix A). These headings, which are identical in form and apparent duplicates, were compared to determine the extent to which they are synonymous in meaning.

Together, these three studies can be regarded as bases for considering several questions associated with using multiple thesauri for both subject and form/genre access. What kinds of compatibility exist among thesauri, and how might the incompatibilities be managed? Can a mainstream list, intended primarily for subject access, be adapted for effective form/genre use in lieu of specialized thesauri? What does it mean to say that headings are identical when drawn from lists developed for different purposes? The usefulness of the present study, then, lies both in what it might say about LCSH or MIM per se and in its potential for replication with other subject heading populations and lists.

ASSUMPTIONS AND LITERATURE REVIEW

Three assumptions underlie these studies. The first is that an item's subject is conceptually distinct from the form or genre it exemplifies. That is, an item's "aboutness" is separate from its nature as a physical, intellectual, or aesthetic entity. While some genres, such as Western novels, presuppose subject matter in a general way, this should not lead us to assume that subject and form/genre are similar concepts. The second assumption is that explicit, separate provision of access to subjects and forms/genres is desirable as a service to the end user. The third assumption is that end users intuitively understand the difference between "is" and "is about," even if they do not articulate that understanding in searching behavior or reference interviews. These assumptions are subject to debate, but are here taken as givens.

Miller (1995) provides a brief literature review of writings concerned with subject and form/genre access issues. For recent developments in this area, a good source can be found in Taylor (1997). This site includes documents related to the work of the ALA/ALCTS/CCS Subject Analysis Committee Subcommittees on Form, from 1991 to the present. At LC, a Form/Genre Working Group, headed by the assistant chief of the Cataloging Policy and Support Office, has been working since 1995 to develop this access throughout LC's collections (Yee 1997). This complex, long-term effort has great potential to bring multiple benefits to the broader library community in terms of rationalized vocabularies, systems development, and redeveloped cataloging practices. Another source of current developments is the electronic discussion list GSAFD (Library Subject Access to Fiction). Discussions in this forum frequently include debates about literary genre headings, their scope and appropriateness for different collections.

OBJECTIVES OF THE STUDY

In the study described here, two types of comparison were made:

1. Quantitative comparison. Authority headings were compared for those headings whose authorized forms are identical in LCSH and MIM to determine the nature and percentages of references shared by each pair of records. It was assumed that the only elements of subject authority records that could be compared quantitatively (besides the established headings), were the reference headings, e.g. Use For (UF). Other elements, such as note fields or classification
number fields, are rarely if ever identical in authority records based on different lists, and no attempt was made to score similarity of these elements.

2. Qualitative comparison. Authority headings were compared to determine the extent to which identical headings in the two lists designate identical or analogous concepts. This aspect of the study clearly involved subjective evaluation, which, although not susceptible to counting and scoring, nevertheless can provide a basis for discussion about the use of established headings as means of naming particular concepts in specific contexts.

**METHOD**

For the quantitative comparison, authority records for identical LCSH and MIM headings were compared for the presence of identical references. Two categories were derived: identical references in identical positions and identical references in different positions.

“Identical references” are reference headings that have identical character strings as the established headings. “Identical positions” refers to the functions that the references serve in authority records: Use For, Narrower Term (NT), Broader Term (BT), Related Term (RT), or See Also From (XX). (See Also From is in MIM only and is rarely used.) As an example of the first category, both lists include Beauty pageants as a UF reference for the established heading Beauty contests. As an example of the second category, LCSH includes Erotic films as an NT for Erotica; MIM, by contrast, lists Erotic films as a UF reference for Erotica.

Two further categories of references were also derived: similar references in identical positions and similar references in different positions.

Similar references are reference terms that are similar, though not identical, character strings. Synonymy, or conceptual overlap considered generally, is not indicated. As an example of the third category, the established heading Talk shows is given two RTs in LCSH: Interviewing in radio and Interviewing in television. These are similar to Interviews, an RT given for this heading in MIM. The fourth category can be demonstrated with the established heading Concerts, a NT used in LCSH, which is similar to Concert films, a UF included in MIM.

It should be clear that the four types of matched references represent decreasingly rigorous categories of duplication. In fact, the first category is the only one that can be said to represent genuine duplication of elements in authority records for identical headings. The other three categories were derived, however, in order to determine the proportion of overlap between pairs of authority records, even under less precise matching conditions. This can be stated as a question: As the criteria for duplication are loosened, how is overlap affected?

For each type of reference (UF, NT, RT, BT, XX) examined, the total number of references of that type was counted for each list. For example, there were a total of 20 BTs in the MIM records examined, as compared with 51 in the related LCSH records (see table 1). The number of matches obtained in each of the above categories was divided by each total to determine the percentage of overlap existing in each case.

For the qualitative comparison, conceptual overlap between pairs of identical terms was evaluated by comparing the semantic content of pairs of authority records. This comparison included reference terms as well as scope notes, classification numbers, and examples. Taken together, these elements provided the context for analyzing the meaning of the established heading.

**RESULTS AND DISCUSSION**

The results of the quantitative comparison of reference headings are presented in tables 1 through 7 (all percentages are rounded to two decimal places). In table 1, we find that 3.78% of the MIM reference headings fall into Category 1 with respect to their counterparts in LCSH. The complement for LCSH-to-MIM comparison is 1.34%. In tables 3, 4, and 6, the equivalent percentages for Categories 2-4 are presented.

In tables 2 and 5 we find information...
TABLE 1
CATEGORY 1: IDENTICAL REFERENCES IN IDENTICAL POSITIONS

<table>
<thead>
<tr>
<th>MIM</th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overlap/Total</strong></td>
<td><strong>%</strong></td>
</tr>
<tr>
<td>UF</td>
<td>5/67</td>
</tr>
<tr>
<td>NT</td>
<td>0/46</td>
</tr>
<tr>
<td>RT</td>
<td>2/50</td>
</tr>
<tr>
<td>BT</td>
<td>0/20</td>
</tr>
<tr>
<td>[XX]</td>
<td>0/2</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>7/185</td>
</tr>
</tbody>
</table>

Note: XX (see also from) reference terms are listed separately, rather than combined with Related Terms, in order to respect the authority record structure used in MIM.

about the nature and number of equivalences noted, when either identical or similar references were found in different positions. For example, table 2 shows that, in the case of identical references, 6 UF references in MIM corresponded to NT references in LCSH. One additional MIM UF reference corresponded to a LCSH BT reference.

In table 7 the percentages of overlap are totaled for the first and second categories, the third and fourth categories, and all categories considered together. When identical references alone are considered, whether in identical or different positions, less than 10% of the MIM reference headings are found in the corresponding LCSH authority records. The total for LCSH-to-MIM comparison is much smaller, at 3.26%. When similar references are considered, the respective figures are 13.52% and 4.79%. The totals for all categories of overlap stand at 22.71% of MIM references identical or similar to LCSH references, but only 8.05% of LCSH references corresponding to those in MIM.

The meanings of these figures are not self-evident and will depend on the uses to which they are put. It would seem that, from a database management perspective,

TABLE 2
CATEGORY 2: IDENTICAL REFERENCES IN DIFFERENT POSITIONS (EQUIVALENCES)

<table>
<thead>
<tr>
<th>MIM</th>
<th>=</th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF</td>
<td>=</td>
<td>NT</td>
</tr>
<tr>
<td>RT</td>
<td>=</td>
<td>UF</td>
</tr>
<tr>
<td>NT</td>
<td>=</td>
<td>RT</td>
</tr>
<tr>
<td>BT</td>
<td>=</td>
<td>RT</td>
</tr>
<tr>
<td>UF</td>
<td>=</td>
<td>BT</td>
</tr>
</tbody>
</table>

TABLE 3
CATEGORY 2: IDENTICAL REFERENCES IN DIFFERENT POSITIONS (PERCENTAGES)

<table>
<thead>
<tr>
<th>MIM</th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overlap/Total</strong></td>
<td><strong>%</strong></td>
</tr>
<tr>
<td>UF</td>
<td>7/67</td>
</tr>
<tr>
<td>NT</td>
<td>1/46</td>
</tr>
<tr>
<td>RT</td>
<td>1/50</td>
</tr>
<tr>
<td>BT</td>
<td>1/20</td>
</tr>
<tr>
<td>[XX]</td>
<td>0/2</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>10/185</td>
</tr>
</tbody>
</table>
only the first two categories have any value, at least for their impact on clerical work. Let us assume a situation in which catalogers create authority records for MIM headings, for a database that already contains records for the identical LCSH headings. The act of keying a reference heading that is also identical might be regarded as an unjustifiable waste of resources. In this case, the database manager will want to remember that this situation obtains for less than 4% of all LCSH reference headings. This is not only a trivial percentage of duplication, but is easily addressed with basic word-processing functions (e.g., copy and paste).

By contrast, the totals of all categories of overlap might be of interest from a broader perspective. If one regards similarity of headings as character strings as indicating similarity in meaning—a plausible but not inevitable assumption—then the less precise matches of categories 3 and 4 might serve as evidence that authority records for identical headings provide redundant access to the same concept. If this is true, then maintaining separate authority records that only differ in function (i.e., subject vs. form/genre) might be regarded as a theoretical nicety at best. Worse, it would give rise to the concern voiced in the June 1995 MARB discussion, where patrons might be required to perform double lookups for

**TABLE 4**

**Category 3: Similar References in Identical Positions**

<table>
<thead>
<tr>
<th></th>
<th>MIM</th>
<th></th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overlap/Total</td>
<td>%</td>
<td>Overlap/Total</td>
</tr>
<tr>
<td>UF</td>
<td>1/67</td>
<td>1.49</td>
<td>1/60</td>
</tr>
<tr>
<td>NT</td>
<td>8/46</td>
<td>17.39</td>
<td>8/385</td>
</tr>
<tr>
<td>RT</td>
<td>2/50</td>
<td>4.00</td>
<td>2/26</td>
</tr>
<tr>
<td>BT</td>
<td>2/20</td>
<td>10.00</td>
<td>2/51</td>
</tr>
<tr>
<td>XX</td>
<td>0/2</td>
<td>0.00</td>
<td>0/0</td>
</tr>
<tr>
<td>Totals</td>
<td>13/185</td>
<td>7.03</td>
<td>13/522</td>
</tr>
</tbody>
</table>

**TABLE 5**

**Category 4: Similar References in Different Positions (Equivalences)**

<table>
<thead>
<tr>
<th></th>
<th>MIM</th>
<th>=</th>
<th>LCSH</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF</td>
<td>=</td>
<td>NT</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>UF</td>
<td>=</td>
<td>RT</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>UF</td>
<td>=</td>
<td>BT</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>=</td>
<td>NT</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>BT</td>
<td>=</td>
<td>UF</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 6**

**Category 4: Similar References in Different Positions (Percentages)**

<table>
<thead>
<tr>
<th></th>
<th>MIM</th>
<th></th>
<th>LCSH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overlap/Total</td>
<td>%</td>
<td>Overlap/Total</td>
</tr>
<tr>
<td>UF</td>
<td>8/67</td>
<td>11.94</td>
<td>1/60</td>
</tr>
<tr>
<td>NT</td>
<td>0/46</td>
<td>0.00</td>
<td>5/385</td>
</tr>
<tr>
<td>RT</td>
<td>3/50</td>
<td>6.00</td>
<td>3/26</td>
</tr>
<tr>
<td>BT</td>
<td>1/20</td>
<td>5.00</td>
<td>3/51</td>
</tr>
<tr>
<td>[XX]</td>
<td>0/2</td>
<td>0.00</td>
<td>0/0</td>
</tr>
<tr>
<td>Totals</td>
<td>12/185</td>
<td>6.49</td>
<td>12/522</td>
</tr>
</tbody>
</table>
TABLE 7
SUMMARY TOTALS

<table>
<thead>
<tr>
<th>Category</th>
<th>MIM: Overlap with LCSH</th>
<th>LCSH: Overlap with MIM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overlap/Total</td>
<td>%</td>
</tr>
<tr>
<td>Category 1</td>
<td>7/185</td>
<td>3.78</td>
</tr>
<tr>
<td>Category 2</td>
<td>10/185</td>
<td>5.41</td>
</tr>
<tr>
<td>Subtotal</td>
<td>17/185</td>
<td>9.19</td>
</tr>
<tr>
<td>Category 3</td>
<td>13/185</td>
<td>7.03</td>
</tr>
<tr>
<td>Category 4</td>
<td>12/185</td>
<td>6.49</td>
</tr>
<tr>
<td>Subtotal</td>
<td>25/185</td>
<td>13.52</td>
</tr>
<tr>
<td>Total</td>
<td>42/185</td>
<td>22.71</td>
</tr>
</tbody>
</table>

the same concept. This conclusion, though, is countered by the fact that, in this case at least, a maximum of just over 8% of LCSH reference terms are even similar to their MIM counterparts. That is, most of the terms present in LCSH authority records have either no MIM equivalents or are, at best, more or less synonymous with them. Regardless, discussions of synonymy take us out of the realm of quantitative comparison and into the second part of this study.

In any event, a string-matching argument for similarity of authority records cannot be supported in this instance. This is even more true when record elements not compared here (such as notes) are taken into consideration. On the quantitative plane, authority records for identical MIM and LCSH headings must be considered to have no significant similarities.

QUALITATIVE COMPARISON

The task of determining the possibility of overlap in meaning between identical headings found in different lists is not one amenable to counting or scoring. What is involved is a consideration of the microcosmic worlds of subject authority records as semantic spaces. In such a semantic space, each element (including fixed-field elements in machine-readable records) can potentially contribute to understanding the meaning of the established heading, within the context of a particular list. We are not concerned, for the moment, with whether or not such a space expresses the meaning intended by the list's creators. Rather, we focus on analyzing the meaning connoted by authority record elements, as actually present in any given record. The subjectivity involved in this task seems analogous with that involved in literary criticism. While different critics might arrive at markedly different interpretations of a text, the text in its concrete form serves as an objective point of reference for those different readings (let us set aside, for the time being, the mutability of electronic texts). This preserves the second part of this study from what might otherwise be considered mere impressionism.

The examination of authority records as semantic spaces has received very little attention in the literature. This is probably because most of the literature on authorities has focused on macrocosmic issues such as syntactic structure and the control of synonyms, issues that are critical to the construction and application of any thesaurus or term list. In addition, most authors writing about authorities have considered only a single list, in which case the question of comparative meanings does not arise.

In his study of the webs of meaning created by See Also–linked headings, Sinkankas (1972, 8–9) writes:

"Words are ambiguous, meaning is slippery, connections are based upon a system of facets of meaning that are felt but not stated. LC does not help in this manner [matter?], having failed to explain a term or its subject coverage about ninety per-
cent of the time. Scope notes and examples are given only when the situation cries [sic] out for them, and many times not even then. It may be possible to get around this problem by going to the LC classification tables, which are many times more comprehensive in their delineation of subject coverage. Often the classification notation (alphanumeric) is included after the subject heading in the LC list, which will speed matters. Where they are not, educated guessing will have to suffice.

Schmierer (1980) describes the "authority data for particular different access points" as typically including the established heading, variant forms, information both about the established heading considered and related established headings, and information sources consulted. She writes (p. 601) that the data "are maintained together so that when an established authorized form is used as an access point, its use may be understood in a context." A working group of the International Federation of Library Associations and Institutions (IFLA) developed the Guidelines for Subject Authority and Reference Entries, which provides definitions of the terms "cataloguer's note," "information note," and "scope note," as well as examples of application (International Federation of Library Associations and Institutions 1993).

Bearman and Szary (1987) discuss authorities as reference files in a multidisciplinary setting. They contend (p. 73) that "as soon as we accept the concept of authority files as information sources in themselves, apart from the 'bibliographic' or other files whose fields they 'authorize,' it becomes necessary to reexamine the types of information they contain, how they are structured, and the integrity of the data they hold." Indeed, Bearman and Szary (p. 75) propose a situation more complex than is considered here, with "multiple, independent authorities at the field level [where] conflicting authorities may control disputed values of the same field within a record!" This transformation of the authority record clearly complicates the idea of semantic space; the context in which an established heading is embedded would be broadened to reveal its connotations across disciplines. Olszewski (1994) more modestly advocates the use of the OCLC Authority File as a valuable information source in its current state. While he focuses on biographical and historical information contained in name authority records, his discussion emphasizes the value of record elements beyond the established heading.

It should be noted that there are also a number of authors who have examined the macrocosmic semantic spaces generated by the relationships between established headings in a list. Sinkankas (1972) performed an extended investigation of see-also relationships. Richmond (1973) discovered an extensive library science thesaurus contained within the Engineers Joint Council's Thesaurus of Engineering Terms. Johnson and Cochrane (1995) described a technology for creating a spatial array of terms from the INSPEC Thesaurus. Here, the semantic space was made visual, as a hierarchical array of narrower, broader, and "top" terms contrasts with a cloudlike display of related terms, described as a "space of loose associations" (abstract). This spatial display is a feature of the search tool itself, in contrast to Sinkankas' large graphical representation and Richmond's conceptual extrapolation of their respectively studied macrocosms.

Two sets of semantic spaces, as demonstrated in authority records for the headings Gossip and Drama, will be analyzed here. As used in LCSH, Gossip is situated in the realm of ethical problems (see figure 1). It is related to Slander, has the narrower term Talebearing, and is assigned the Library of Congress classification number BJ1535, "Special vices." Gossip is unambiguously regarded in this context as a vice, a near neighbor to, if not identical with, outright lying. In MIM, by striking contrast, Gossip is regarded as information about a person, particularly a celebrity, which is essentially true (see figure 2). Its broader term is Documentaries and factual films and video, and it is assigned the related terms
Biographies, News, and Magazines. The scope note (Yee 1988, 53) reads “Use for nonfiction films and programs which relay rumors, anecdotes, likes and dislikes, personal history, etc. of celebrities.” The inclusion of “rumors” modifies the truth value of Gossip to an extent, but not enough to neutralize its documentary, nonfictional, biographical nature. We are also reminded that the heading is used for instances of broadcast gossip, rather than studies of communication ethics, in its relation to Magazines (e.g., 60 Minutes).

The heading Drama presents a subtler distinction in usage. The use of Drama in LCSH encapsulates the amb-

Figure 1. Gossip: LCSH
Context: Ethics

Figure 2. Gossip: MIM
Context: Infotainment
“Nonfiction films and programs”
guity that pervades the study of dramatic literature, a study poised between literary analysis and stage practice (see figure 3). The heading is grounded in Literature, the encompassing BT, reinforced by three of its four UF terms, Plays, Drama—Philosophy, and Drama, Modern. At the same time, it is related to the headings Acting and Dialogue, the former not fundamentally a literary matter, and its fourth UF term is Stage, upon which Acting happens. Its fifty-three NTs fall primarily into three categories: dramatic genres such as Horror plays, subjects such as Religion in drama, and elements of playwriting such as Prologues and epilogues. None of the NTs relate directly to the practice of theater produc-
tion, with the possible exception of Promptbooks. The heading Drama, then, is intended to designate a literary form, closely connected with stage production.

In MIM, Drama designates a genre of televised or filmed fiction, without direct connection to the stage (see figure 4). Its BT is Fiction, which has to do with "imaginary characters and events" rather than literature (Yee 1988). Its scope note (44), "use for fictional works of serious intent, emphasizing conventions of character and narrative development through conflict and cathartic resolution," marks Drama as a type of fiction distinct from Comedies, one of its RTs. The other RT, Plays, is reserved for documentation of live play performance. Drama in MIM thus already pertains to the work as produced for a viewing audience, rather than a text that serves as an object for critical study. Additionally, while a viewing audience is implied at a secondary level in LCSH, the medium and nature of production, live theater, is different.

In each of the 27 cases studied, comparative examination of the semantic spaces occupied by the "identical" headings in each list reveals more or less marked differences of meaning. Not every comparison shows as dramatic a distinction as that between the two meanings of Gossip or as subtle as that of Drama. While limitations of space prevent a discussion of every pair of headings, two additional cases can be mentioned. The authority records for the heading Talk shows point to very similar meanings; that is, the idea of what a "talk show" is seems to be similar whether considered as subject or as genre. By contrast, the records for Music create quite distinct spaces. In LCSH, Music connotes Western classical music primarily, but the entire sphere of music and its impact on human culture by extension. In MIM, the heading is used exclusively for recordings of live performances. But even where comparative meanings are superficially similar, the distinction between subject and form/genre access underlying the two lists means that the entities indexed by "identical" terms will, in most cases, differ significantly from each other. If an established heading is given meaning by the elements of its authority record, that meaning is also grounded in a list's intended application.

**Summary**

A quantitative comparison was made to determine the percentage of reference headings that were shared between authority records for 27 established headings identical in LCSH and MIM. When the comparison was limited to reference headings identical in both form and function, it was found that fewer than 4% of MIM references were shared with LCSH. Conversely, fewer than 2% of LCSH references were shared with MIM. The criteria for comparison were incrementally loosened, finally to include reference headings only similar as character strings, and which may serve different functions in their authority records. Under these liberal conditions, approaching synonymy rather than identity, it was found that approximately 23% of MIM headings were shared with LCSH. However, the converse LCSH-to-MIM figure was just over 8%.

A qualitative study was also made of the semantic spaces created by authority records for identical headings. It was found that identical headings often have different meanings in the context of their respective thesauri. In short, the creation and maintenance of separate authority records for "identical" headings cannot automatically be assumed to be redundant effort. Additionally, the sets of resources correctly indexed by headings identical in form are likely to differ.

**Management Considerations**

Library managers responsible for authority control will naturally be concerned about how best to minimize conflicts between headings from multiple thesauri. The specificity of discipline-based vocabularies, while of potential benefit to users, carries with it the need to clarify the scope of similar or identical headings, among other challenges. Authors of the literature on the subject are primarily concerned with management of multiple-
subject thesauri, a different situation from that giving rise to this study. Approaches to the multiple-subject thesaurus problem can, nevertheless, provide insights on management of form/genre lists in combination with LCSH or other subject lists.

Mandel (1987, v) describes "four basic approaches to providing access to databases indexed by different vocabularies": creating segregated files, mixing vocabularies, integrating vocabularies, and providing enhanced front-end navigation. These comments will focus on the first two approaches.

The major disadvantage of the segregated-file approach, where subject access is concerned, is that users are required to perform more than one search for the same or similar concepts. However, since we are here concerned with distinguishing applications of identical headings rather than integrating them, the segregated file approach may be the simplest. Provision of a form/genre index, with headings drawn from USMARC 655 fields, could provide clearly distinguishable access to items of particular types, rather than works about those types. Of course, separate indexing of 655 fields still requires that decisions be made about form/genre headings in existing records coded in 650 fields. Can these be recoded, either automatically or manually? What are appropriate candidates for recoding? Managers must also consider what to do about headings, such as Science fiction in LCSH, which have been used both for subject and form/genre access. Automatic rather than manual recoding of these headings might not be feasible, depending on the size of the catalog and nature of the collection.

At Curry College, the Levin Library's system was profiled in 1995 with a separate form/genre index. Headings have been added to the index in three ways: deliberate provision of form/genre headings (primarily through video cataloging using MIM), recoding of 650 to 655 fields on an ad hoc basis (primarily via current cataloging), and acceptance of 655 fields on copy cataloging (primarily for works of fiction). We have plans to retrospectively convert and provide genre headings for particular literary genres, such as detective stories, for which classes are taught regularly at the college. As of May 1997, a section of the form/genre index contains these headings:

- Medical films and video (Nonfiction)
- Medical novels
- Mock heroic literature
- Musical fiction
- Nature films and video (Nonfiction)
- Nonsense verses
- Pastoral fiction
- Picaresque literature
- Police films and programs
- Political fiction
- Political films and programs
- Prison films and programs
- Psychological fiction
- Questions and answers

In the second approach, terms from all vocabularies are indexed in a single list and retrieved through a single search. Mandel notes (v), "the two major problems caused by this approach are: (1) obvious vocabulary clashes (e.g., the same term is postable in one vocabulary and non-postable in another), and (2) degradation of access to specialized collections." This is no less true where subject and form/genre vocabularies are mixed than it is with multiple-subject vocabularies. Two primary techniques for minimizing the difficulties with this approach are qualification of terms and editing of headings. Automated systems that allow qualification of terms, based on coded information in authority records, make it possible to clarify the scope of identical or similar headings used in the same index. For example, the unqualified heading Soap operas might be juxtaposed in a subject index with Soap operas (Video/Film). Qualifiers might also be manually added to headings from alternative lists, though this is more labor-intensive and might have to be undone after a system migration or upgrade.

It is also possible to create a policy mandating that specific headings from alternative lists not be used. This also implies an ongoing investment in policy maintenance and review and could result in a loss of headings that, again, would be useful with an improved system. Despite
the drawbacks of the mixed-vocabulary approach, for many libraries this will be the most feasible in the short term.

The third approach, integrated vocabularies, involves different means of relating vocabularies through mapping and the creation of syndetic structures, such as micro- and macrothesauri. This is an area in which a great deal of research has been done (see Chaplan 1995 and Olson and Strawn 1997 for two recent reports). The creation of maps, which would allow machine-assisted switching between related headings, implies a more sophisticated relationship between thesauri than either segregated files or mixed vocabularies. Though the initial editorial effort can be significant, and ongoing review is required, the results can be shared among libraries and systems and need not be redeveloped by every agency. Where subject and form/genre vocabularies are mapped, however, it is essential that systems be able to distinguish among the proper objects of redirected searches. As an example, a subject search for studies of Melodrama (LCSH) should not be directed to examples of Melodrama (MIM) on film or video.

**AREAS FOR FURTHER RESEARCH**

The rational, deliberate provision of access to genres and forms raises a number of questions requiring further investigation. The series of comparisons that have been made between LCSH and MIM should be replicated with other vocabularies. Do other form/genre thesauri overlap with LCSH to a greater or lesser extent than does MIM? How true is this, not only for established headings, but for the elements in their authority records? How do these degrees of overlap affect the choice to establish separate indexes, provide automatic qualification of headings within a single index, or otherwise make explicit the meanings of similar headings? The same questions apply to overlaps among form/genre thesauri themselves, not just in comparison with LCSH or other subject-access lists. There are a number of opportunities for vocabulary-mapping studies in these questions.

Research into the capabilities of automated systems—and communication with systems vendors regarding support for multiple thesauri—will be a continuing need for the foreseeable future. As of spring 1997, we are far from having universal support for either X55 MARC Authority fields or qualification of terms based on authority record coding. Enhanced front-end navigation to help the user determine whether subject or form/genre searching is desired is an area for investigation. In addition, algorithms will need to be developed so that systems can apply the mapped terms to appropriate bibliographic items, if mapping between subject and form/genre vocabularies is developed.

Researchers might also want to consider how library patrons will be instructed in the use of form/genre terms. While library instruction is generally not considered a technical services function, the introduction of form/genre indexes or inclusion of new vocabularies might be resisted due to the challenge of educating users. It was asserted above that library users probably have an intuitive understanding of the difference between "is" and "is about." Nevertheless, libraries have not provided explicit access to forms and genres in a consistent fashion (except to works of Western classical music) through LCSH. Introducing such explicit access will require additional instruction in many situations, and technical services librarians might be able to share insights with their colleagues in public services.

**CONCLUSION**

The 27 pairs of headings studied, while identical in form between LCSH and MIM, diverge in application. That is, on the evidence of the authority records that support them, they cannot be regarded as duplicative. The quantitative comparison, of potential interest to the database manager responsible for managing an authority file, revealed that fewer than 10% of reference terms are shared between pairs of records. The qualitative comparison indicated the likely reason for this lack of duplication: the headings have different
meanings, sometimes to a startling degree, and are intended to apply to different bibliographic objects.

The lack of actual identity expressed by erstwhile-identical headings should remind us that headings are names, as Julia Pettee (1954, 18) pointed out. My first name, and that of one of Michelangelo's statues, are identical only as character strings. The entities designated are very different. It is impossible that a sane person would regard them as the same thing, and yet the names are identical. The fact that different concepts may share the same name presents significant management challenges when multiple thesauri are used in one catalog. Nevertheless, if we consider a heading to be not only a character string, but also what that string designates, we might come to the conclusion that Gossip and Gossip, Drama and Game shows have little in common. This is all the more true when one of the two designates what an item is, and the other, what it is about.

**Works Cited**


(gopher://marvel.loc.gov/00/.listarch/usmarc/95-11.doc).

APPENDIX A.
IDENTICAL HEADINGS IN LIBRARY OF CONGRESS SUBJECT HEADINGS AND MOVING-IMAGE MATERIALS: GENRE TERMS

Advertising
Anthologies
Beauty contests
Computer animation
Drama
Editorials
Erotica
Fiction
Film noir
Game shows
Gossip
Interviews
Melodrama
Music
Music videos
Musicals
Newsreels
Parodies
Press conferences
Propaganda
Quiz shows
Science fiction
Soap operas
Talent shows
Talk shows
Television
Vaudeville
Changing Roles: Original Cataloging by Paraprofessionals in ARL Libraries

Deborah A. Mohr and Anita Schuneman

The role of the paraprofessional cataloger in academic libraries is rapidly changing. The authors in this study investigated the nature of paraprofessionals’ work in original cataloging activities at ARL institutions and compare their findings with those of an earlier survey. Original cataloging was defined to encompass a variety of activities including description, the creation of name and uniform title headings, subject analysis, and classification. Findings reveal that 77.1% of the cataloging department heads at the responding ARL institutions report paraprofessional involvement in one or more of these activities, with original description the most common and subject analysis the least common. Among the reasons commonly cited for such involvement were paraprofessional career development and cost savings. The respondents also noted advantages and disadvantages of paraprofessional participation in original cataloging as well as reactions they had observed among both professional and paraprofessional catalogers to this participation.

The role of the paraprofessional cataloger in academic libraries is changing at an ever-increasing rate. Both experience and a review of the literature demonstrate that paraprofessionals are taking on tasks previously performed only by professional librarians. While at one time there may have been general agreement that librarians performed the intellectual work of cataloging and paraprofessionals the routine tasks, it is no longer clear that this is where the distinction between “professional” and “paraprofessional” work lies. Although philosophical issues rarely arise on the job, catalogers and cataloging managers deal with the practical issues on a daily basis, and our interest arose from our awareness of increasing paraprofessional involvement in original cataloging. In this study, we examine the daily operations to determine what original cataloging activities are being performed by paraprofessionals.

Research Questions

We hypothesized that paraprofessionals participated widely throughout the
Association of Research Libraries (ARL) libraries in all areas of original cataloging, with the greatest level of involvement in original description and in literature classification and the smallest in subject analysis and in classification of nonfiction works. It was further hypothesized that comparison with past studies would reveal a continuation of the trend of increasing paraprofessional involvement in original cataloging over time.

Third, it was hypothesized that ARL cataloging department heads would most frequently cite cost savings as a reason for paraprofessional involvement in original cataloging, given the current difficult economic situation of many libraries. Finally, we projected that revision of original cataloging work would more often continue beyond the initial training period for paraprofessional catalogers than for professionals.

**Literature Review**

Authors of previous surveys have documented the involvement of paraprofessionals in original cataloging. Further, they have shown that paraprofessional participation has been increasing steadily at least since 1977, six years after OCLC's debut, when Braden, Hall, and Britton (1980) surveyed OCLC member libraries about their cataloging practices. They defined "original cataloging" as "cataloging without copy" and found that respondents in 16.6% of the larger libraries stated that support staff handled original cataloging either alongside professional catalogers or exclusively.

A few years later, in 1983–84, Eskoz (1990) surveyed catalog departments in 160 academic libraries chosen for "regional balance and a varied range of campus sizes and types" (p. 380). She subsequently conducted follow-up interviews with catalog department heads in 40 of the libraries in 1986 and 1987. She found that both professionals and support staff did original descriptive cataloging in 30% of the libraries in 1983–84, which increased slightly to 35% of the libraries by 1986–87. However, support staff were assigned this original cataloging in only two of the libraries in the surveys, while no professionals were assigned this responsibility. Eskoz found similar rates of paraprofessional participation in assigning call numbers (32.5% in 1983–84, 35% in 1986–87) and assigning subject headings (27.5% in 1983–84, 35% in 1986–87).

By 1990, Oberg et al. (1992) found that paraprofessionals regularly did original descriptive cataloging in 51% of the ARL institutions that responded to their survey. However, the percentages of ARL libraries in which paraprofessionals assigned subject headings and call numbers remained about the same (36% for each task), a figure similar to that found in Eskoz's second study. In Conturbia's 1991 study (1992) of cataloging of foreign-language materials, although more specific in scope, the author found similar levels of paraprofessional involvement in these tasks.

Despite differences in the populations surveyed, definitions, and methodology, these authors pointed to a general trend in cataloging that was occurring throughout the library world during the same period: the increasing delegation to paraprofessionals of tasks formerly considered professional. This trend reflected emerging thought about the proper role of the professional librarian in technical services—and in the profession in general—as managers, leaders, and innovators, less involved than previously in day-to-day operations (Veaneer 1982; Bishoff 1987; Younger 1991; Rider 1996).

There were other factors that had an effect on the responsibilities of cataloging librarians and staff respectively. During this period, automation made many library tasks more routine, and in cataloging departments, the increased availability of copy for copy cataloging pushed broad implementation of copy cataloging. At the same time, copy catalogers gained experience and skills over the years (Benaud 1992), which could be used for more complex work. In addition, economic pressures pushed for the containment of costs through the delegation of tasks to the lowest levels of staff possible while maintaining an acceptable degree of quality and without harming morale (Williams 1991).

With these developments came con-
cern for paraprofessionals in their changing roles as well. Oberg (1994, 1995; Oberg et al. 1992) led the call for equitable treatment of paraprofessionals in libraries and indicated the need to imbue paraprofessionals with the profession's traditional values as support staff play a greater role in libraries. Rider (1996) approved of the new roles for copy catalogers as a means for them to use their experience and skills appropriately, while Dyckman (1992) emphasized attention to recruitment and retention of paraprofessionals in their developing role.

Nettlefold's sociological view (1989, 523) of these trends posited that the rise of paraprofessionalism in libraries resulted from the maturing of the library profession:

Aspiring professions . . . have often resisted paraprofessionalism initially, afraid of a claim on their "exclusive skill jurisdiction." . . . As the aspirants have redefined and upgraded their core of knowledge and skills, confidence in their professionalism has grown. They have therefore begun to act as established professions and now foster paraprofessionalism. . . . Librarians at first opposed paraprofessionalism, but as their professionalism has grown, opposition has turned to support.

Representing another sociological perspective, Harris (1992, 9–10) argued that: some of the female areas of specialization, especially cataloging, are undergoing a process of 'deprofessionalization' or 'deskilling'. . . [which] can occur when a field loses control over its knowledge base. In the case of cataloguers, this loss of control has come about largely because of the widespread use of cataloguing networks . . . [which] not only shifts most of the in-house cataloguing work in libraries to non-professional staff but . . . also alters the working patterns of the remaining professional cataloguers, [who can no longer specialize in particular subject areas.]

**METHODOLOGY**

We developed an original survey instrument, which was mailed in April 1995 to the heads of the main cataloging departments of the 119 ARL institutions. ARL libraries served as the target population because their specialized collections meant higher rates of original cataloging than for many other academic libraries and for ease of comparison with some past studies. After two follow-up mailings, 83 respondents had returned their questionnaires for a total response rate of 69.7%.

We developed our own survey instrument because none of the existing questionnaires of which we were aware dealt with all of the issues we wished to explore. As would be expected, the questionnaire from the earlier study of Braden, Hall, and Britton included questions about now-outdated cataloging policies. The survey by Eskoz, which was closest in scope to ours, did not contain the same level of detail with regard to the reasons for paraprofessional involvement in original cataloging or the reactions of professionals. Oberg’s 1990 survey focused on all library paraprofessionals, while Conturbia limited her survey to foreign language cataloging.

Methodologies differed as well among these surveys as to the libraries surveyed, survey respondents, and what followup measures were employed. Two of the 4 studies were specifically targeted at ARL libraries, although all surveyed academic libraries. All but Oberg’s broader study were surveyed cataloging department heads rather than cataloging practitioners. One researcher, Eskoz, used telephone interviews as a means of follow-up.

**DEFINITIONS**

For the purposes of the study, “original cataloging” encompassed original bibliographic description, establishment of name and uniform title headings, subject analysis, and both nonfiction and literature classification, when any of these exceeded verification of data in existing records. This definition included tasks sometimes performed in complex copy cataloging, such as assigning subject headings to a bibliographic record that lacked them. “Revision” meant a review of the cataloging decisions made in a record, as opposed to general proofreading.
“Professional” catalogers were defined as those who held positions that required a master’s degree in library or information science. “Paraprofessional” catalogers were defined as those who held positions that did not require an M.L.S., although many paraprofessional library employees do have the degree. The definitions of all these terms were included in the appropriate survey questions so that respondents were aware of them.

The term *paraprofessional* was used despite its drawbacks. Benaud (1992, 84) notes that *paraprofessional*, while one of the most prevalent terms, along with *support staff*, is used to describe library workers without an M.L.S., “no longer reflects the sophisticated skills needed on the job.” Oberg (1995), while acknowledging the drawbacks, finds the use of the term *paraprofessional* highlights comparisons with similar phenomena in fields such as law and medicine.

Definitions used in previous studies differ from the definitions used here, although not so significantly as to prevent later comparison. For example, Oberg (1992) defined “paraprofessional” to include office support and other staff activities, which were omitted from the definition here because only paraprofessional staff in cataloging departments were considered. However, the definitions of *original cataloging* were very similar—either “cataloging without copy” or a listing of tasks—and again support subsequent comparison among survey findings.

**Survey Findings**

Eighty-three librarians responded. The findings are reported question by question (13 questions in all). The questions covered the following topics: department size, original cataloging activities performed by paraprofessionals, what revision of original cataloging was done, reasons for paraprofessional involvement in original cataloging, reactions observed among cataloging employees to this involvement, and advantages and disadvantages of the paraprofessional involvement from the perspective of the cataloging department heads surveyed. The survey findings are reported generally in the order of the questions on the survey, with the exception of department size which is reported after the question of what original cataloging activities are performed by paraprofessionals. Immediately following the findings on each question, we have included a brief discussion of these findings in the context of other studies.

**Distribution of Original Cataloging Activities**

Overall, 64 libraries (77.1% of those responding) had paraprofessional employees who performed some original cataloging tasks, while 19 (22.9%) did not. Of the 64 responding libraries where paraprofessionals were involved in original cataloging, 63 (98%) also indicated the tasks that paraprofessional catalogers performed. The distribution of original cataloging tasks is shown in figure 1.

**Original Description**

Original description was the original cataloging activity most commonly performed by paraprofessionals in the responding libraries. Paraprofessional catalogers did original descriptive cataloging in 62 of the 63 libraries in which cataloging department heads responded to this question. While this result agreed with the hypothesis, the extent of the participation was astonishing; paraprofessionals do original description in nearly 100% of the libraries in which they participate in original cataloging activities.

The result is perhaps less surprising given the paraprofessionals’ copy cataloging experience. Also, a wealth of documentation—including the Anglo-American Cataloguing Rules second edition; Library of Congress Rule Interpretations; guidelines; and manuals—provides a solid base for training in original descriptive cataloging.

The level of paraprofessional involvement in original descriptive work has increased notably since the survey of ARL libraries conducted in 1990. That survey found paraprofessionals were regularly assigned original descriptive cataloging in
51% of the responding libraries (Oberg et al. 1992), versus 74.7% of all respondents to the present survey. In the present study, “regular” participation was not specified, but even with this difference in definitions, there appears to be widespread acceptance of paraprofessional participation in original descriptive cataloging in ARL libraries.

**Classification**

Paraprofessionals did original literature classification in 53 (84.1%) and original nonfiction classification in 42 (66.7%) of the 63 libraries in which paraprofessionals were involved in original cataloging. The first finding reflects long-standing practice, as described by Foster (1987, 96) ten years ago: “In at least a few libraries non-professionals are being assigned certain original cataloging operations . . . Non-professional cataloging of fiction books has proven especially successful.” This success rests in part on the relatively similar classification schemes and common use of standard tables for classifying literature.

Although authors of previous studies have not differentiated between literature and nonfiction classification, no study shows as high a percentage of paraprofessional involvement in classification in general as the present survey. The rates of participation in the studies by Eskoz, Oberg et al., and Conturbia ranged from 32.5% to 37.5% of responding libraries. Even the lower rate of involvement for original nonfiction classification in the present survey, 49.4% of all respondents, represents a 13.4% increase since the 1990 study by Oberg et al., although again definitions of participation differ somewhat.

The trend might be attributed to increasing acceptance of paraprofessionals’ contributions and abilities in this area as well as to time and financial pressures. In addition, Rider (1996, 29) predicts that “subject headings and classification numbers will continue to be assigned by greater numbers of staff with appropriate training because the information can often be adapted quickly from other online sources.”

**Subject Analysis**

Consistent with the hypothesis, 41 (65.1%) of the 63 libraries in which para-
professionals are involved in original cataloging. This is, however, close to the number of libraries (42) in which paraprofessionals do original nonfiction classification. This finding may reflect the fact that subject analysis and nonfiction classification are the two cataloging tasks most often cited as professional activities that require an M.L.S. (Benaud 1992). From our own experience, these areas are also less codified and documented than the others and require the most creative application of the sketchier rules that do exist.

Comparisons with findings from past studies show a substantial increase in the rate of paraprofessional involvement in original subject analysis, a result similar to that found in the area of original classification. In the current study, the rate was 49.4% of all respondents; in earlier studies, it ranged from 36% of responding ARL libraries in Oberg et al. (1992) to 30% for foreign-language materials in Conturbia (1992). The trend might be due to increasing acceptance of paraprofessionals’ contributions and abilities in this area, economic and time pressures, and better documentation and rationalization of subject analysis practices over the past few years.

Establishment of Name and Uniform Title Headings

Cataloging department heads indicated that paraprofessional catalogers establish some name and uniform title headings in 48 (76.2%) of the 63 libraries in which paraprofessionals do some original cataloging, which is 57.8% of all responding libraries. Authors of past studies did not distinguish between this activity and original description in general, so there are no grounds for comparison. As shown in figure 1, this task is the third most commonly performed by paraprofessionals in the responding libraries, after original description and literature classification and before original nonfiction classification and subject analysis.

Heading establishment and authority record encoding are well documented, which may aid in training along with paraprofessionals’ exposure to headings and authority work in copy cataloging. In our experience, a copy cataloger was the biggest contributor of name and series headings to the NACO (Name Authority Cooperative) project in the library in which we formerly worked, showing the potential importance of paraprofessional contributions in this area.

Department Size and Composition

We sought to identify what constitutes a typical configuration for an ARL cataloging department in which paraprofessionals do some original cataloging. Questionnaire respondents gave the numbers of paraprofessional and professional positions in their cataloging departments in full-time equivalents (FTEs). Total department sizes ranged from 4 to 124 FTEs in the libraries in which paraprofessionals performed some original cataloging, with an average of 24.9 FTEs per department. In libraries where paraprofessionals did no original cataloging, department sizes ranged from 7 to 657 FTEs, with an average of 53.9 FTEs.

The wide range in size of the second group is due to the Library of Congress (LC), which is an exceptional library because of its FTE count of 657, which raises the average size by 33.2 FTE. For that reason, median department size might represent a more meaningful basis for comparison between these two groups of libraries. For libraries in which paraprofessionals participated in original cataloging activities, median department size was 23 FTEs, versus 19 FTEs for the libraries in which paraprofessionals did no original cataloging. Paraprofessional involvement in original cataloging thus appears slightly more likely with greater department size (with the exception of LC), perhaps because of the larger pools of skills upon which to draw as well as the potentially greater quantities of materials that require original cataloging.

Cataloging departments for both groups of libraries generally had more paraprofessional than professional positions. Among libraries in which parapro-
fessionals performed some original cataloging, the overall ratio of paraprofessional cataloging positions to professional cataloging positions in FTEs was 1.79 to 1. For the other group, the average ratio was 1.46 paraprofessionals to 1, although this reversal disappears if LC is eliminated, which changed the overall ratio to 1.4 paraprofessional FTEs to 1 professional FTE. Departments in which paraprofessionals participated in original cataloging had, on average, about 1.25 times as many paraprofessional FTEs for each professional FTE as the departments in which paraprofessionals did no original cataloging.

This difference in configuration makes sense. With a higher ratio of professional cataloging positions to paraprofessional positions, the need for paraprofessional involvement in original cataloging is minimal, with the larger number of professionals as well as the large amount of cataloging copy available to all but the most specialized libraries. In larger departments with a lower ratio of professional to paraprofessional positions, more pressure occurs for paraprofessionals to perform original cataloging. Paraprofessional participation in original cataloging can both assist in getting the original cataloging done when there are fewer professional catalogers, who are also apt to be required to spend more time on activities other than title-by-title cataloging, and keep the paraprofessionals supplied with work. Later questions in the survey identify other reasons for paraprofessional participation in original cataloging activities.

**Comparison of Revision Rates**

Revision of original cataloging work is one method of quality control. Respondents from ARL libraries in which paraprofessionals did some original cataloging indicated whether or not paraprofessionals as a group and professionals as a group were generally revised in their original cataloging activities after training. If the answer was yes, respondents indicated whether professional catalogers, paraprofessional catalogers, or both did the revising. "Revision" was defined as a review of the cataloging decisions made in a bibliographic record, rather than mere proofreading.

Responses were nearly evenly split between departments in which paraprofessionals continued to undergo revision of their original cataloging even after full training: in 33 (55%) of the responses revision continued after training, and in 27 (45%) revision ceased with training. For professional catalogers, however, continuing revision of original cataloging occurred much less often, taking place in only 17 (27.9%) of the responding libraries, as opposed to 44 (72.1%) of libraries in which professional catalogers worked with no revision beyond the training period. Although small in scale, authors of a survey done in 1988–89 of eleven ARL libraries learned that "support staff are more likely than professionals to be evaluated on the basis of statistics kept on material processed or work completed or on the basis of 'qualitative' checks of their output" (Estabrook, Mason, and Suellflow 1993, 239).

In most cases, if continuing revision of the work of either paraprofessionals or professionals was done, professional catalogers performed it (paraprofessionals: 27, or 81.8% of the respondents to this question; professionals: 14, or 87.5%). Nowhere did only paraprofessionals revise original cataloging, although both paraprofessional and professional staff did so for paraprofessionals in 6 (18.1%) of the responding libraries and for professionals in 2 (12.5%). This finding lends some support to Benaud's observation (1992, 86) that "experienced catalogers, professional or not, act as 'resource' persons and train new catalogers. As a result, it is often the case that paraprofessional catalogers train [professional catalogers] and revise [their] work." The findings of this study lend only partial support in that it does not appear to be often that paraprofessionals revise the work of professionals (2, or 12.5% of the libraries). Our finding also indicates that while acceptance and even advocacy of paraprofessional involvement in original cataloging might be widespread in ARL libraries, there still exists one line of demarcation.
between “paraprofessional” and “professional” roles in many cases.

**Reasons Underlying Paraprofessional Involvement**

Most questionnaire respondents indicated one or more reasons why paraprofessional employees did or did not do original cataloging in their libraries. These were not mutually exclusive. Almost all (62, or 96.9%) of the respondents from libraries in which paraprofessionals performed original cataloging tasks answered this multiple-choice question. The results are presented in figure 2. The respondents most frequently chose “career development for the paraprofessional employees” as a reason (39, or 62.9% of those who answered the question), which does not support the hypothesis that cost savings would be the most commonly cited reason.

It is heartening that so many of the cataloging department heads who responded were concerned with the growth and development of their paraprofessional staff. This concern echoes discussion in the literature (Oberg 1994, 1995; Rider 1996). As Rider (1996, 29) notes, paraprofessionals “should have more opportunities for advancement and individual recognition, and continue to play more visible roles within the library profession.” Indeed, this finding suggests that paraprofessional involvement in original cataloging benefits both the libraries and the paraprofessional catalogers.

Concerns for cost and getting the original cataloging done followed closely as cited reasons for involving paraprofessionals in original cataloging. “Cost-saving measure” or “original cataloging volume too great for professional employees alone” were cited by 38 (61.3%) of the respondents. And nearly half the respondents (27, or 43.7%) selected both cost savings and volume as reasons. In some cases, the cost savings may have arisen from lower salaries for paraprofessionals, raising an equity issue, but Williams (1991, 33–34) has also noted that “the contribution of librarians to technical services operations is relatively expensive. While their salaries may not be high, the amount of time that librarians are asked to spend on [activities outside their departments] means that much of their time is spent outside the units in which they work.” In many academic libraries, para-
professionals may spend much more of their time actually cataloging than do professional catalogers. This has been our own experience as catalogers.

Only 12 (19.4%) cited “the division between professional and paraprofessional cataloging is artificial” as a reason, and none as the sole reason. For the most part, cataloging department heads focused on the more practical issues of career development of paraprofessional staff members and the exigencies of budgets and workloads.

**Other Reasons**

Half the respondents from the 64 libraries in which paraprofessionals did some original cataloging specified additional reasons. 16 (25.8%) mentioned the paraprofessionals’ experience and general qualifications, their subject, language, and format expertise, and the appropriateness of delegating the less complex, more routine aspects of original cataloging to paraprofessionals. Other reasons, cited less frequently, included institutional policies that professional librarians should not catalog (6.3% of those who gave other reasons), the difficulty of recruiting a sufficient quantity or quality of professional catalogers (6.3%), the demand for professionals in other areas of the libraries (3.1%), and the historical lack of division in some libraries between professional and paraprofessional cataloging tasks (3.1%). These other reasons also focused largely on the practical concern of getting the work done.

Among respondents from libraries in which paraprofessionals did no original cataloging, 17 (89.5%) gave one or more reasons why not. Again, multiple reasons were mentioned. These respondents most often selected “original cataloging should be a professional activity” (9, or 52.9% of respondents to this question) as a reason. While the choice of this reason did not indicate whether the opinion was the respondents’ own or an institutional policy, in effect it served as the latter. The question of what constitutes professional work in cataloging was as much a practical issue as it was a philosophical one for this group of respondents.

“Civil service rules, union contract, or other such restrictions” ran a close second, selected by 8 (47.1%). Only 3 (17.6%) indicated “available paraprofessional employees [are] inexperienced or otherwise unsuited,” and none gave this as the sole reason. No one chose “[a] lack of professional employees to train and revise the paraprofessionals” as a reason why the paraprofessionals did no original cataloging.

Three (17.6%) of the respondents to this question specified other reasons why paraprofessionals did not participate in original cataloging in their libraries. These other reasons included having sufficient professional catalogers to do all the original cataloging or, conversely, enough copy cataloging to keep the paraprofessional catalogers occupied; the value of professional activities in enriching original cataloging abilities; and “tradition—difficult to break resistance from both professionals and paraprofessionals.” As more libraries have involved paraprofessionals in more aspects of original cataloging, however, this tradition has become less entrenched and will continue to do so. Resistance from paraprofessionals who would prefer not to do original cataloging without an increase in pay has merit, and this inequity should be addressed by the profession.

**Reactions to Paraprofessional Involvement**

The 64 cataloging department heads who responded to the survey that paraprofessionals in their libraries participated in original cataloging were asked in an open-ended question to describe briefly any reactions they had observed among either paraprofessional or professional catalogers to this development. 39 (73.6% of the respondents to this question) reported they observed positive reactions among paraprofessionals, while 7 (13.2%) observed negative reactions. In reporting reactions among the professionals, 30 (58.8%) pointed to positive reactions, and 14 (27.4%) found negative reactions. 7 respondents reported mixed reactions.
from both paraprofessional (13.2%) and professional (13.7%) catalogers.

Reactions of Paraprofessionals

On the positive side, paraprofessionals were observed to "enjoy" original cataloging and to find it "more difficult, challenging, and job rewarding." The mixed reactions reported usually included both themes of resentment and enjoyment, which is understandable in situations where paraprofessionals received additional training and responsibility without attendant increases in compensation and status.

According to the department heads, a recurring negative reaction among paraprofessionals was "resentment of doing professional work for lower pay" than the professionals. Many paraprofessionals were long-term employees who had absorbed other formerly professional duties over the years. These sources of resentment have been discussed by Oberg (1995, 1994). Williams (1991, 39) states "support staff members who learn new skills will have to be paid more, and, if these employees are to be given managerial slots, mechanisms must be developed to allow them to sit with librarians in policy making sessions when it is appropriate and to pay them at levels commensurate with their responsibilities."

Reactions among Professionals

Many department heads observed positive reactions among professional catalogers. Professionals in some libraries welcomed help with the routine aspects of original cataloging and in others appreciated the paraprofessionals' expertise in particular languages, formats, or subject areas. This range of reactions is to be expected as the profession's focus continues to shift from day-to-day operations to responsibility for management, leadership, and innovation.

Other respondents to this question stated that the professional catalogers in libraries in which paraprofessional catalogers participated in original cataloging were sometimes "upset at seeing original cataloging 'demeaned' as inappropriate for professional attention." They complained that "administrators don't understand or value the work of catalogers." Paiste and Mullins (1990) offer some suggestions for job enrichment for professional catalogers that might indirectly reduce some of these tensions. Their lengthy list is varied but with a common theme: the activities do not focus on cataloging individual works. Their suggested activities call for more involvement in the overall work of the cataloging department in areas such as planning, goal setting, establishing standards, hiring and training new library assistants and catalogers, supervising department work units, consulting with those outside the department, instructing library employees and users in effective use of the catalog, working on assignments outside of cataloging, managing cataloging projects, and developing specialized knowledge in handling particular subjects or types of materials.

Advantages and Disadvantages of Paraprofessional Involvement

For the 64 libraries in which paraprofessionals participated in original cataloging activities, 62 (96.9%) listed the advantages they saw in involving paraprofessionals in this way in response to an open-ended question. Many of the advantages mentioned appeared to be elaborations of the reasons for involving paraprofessionals in original cataloging in the first place.

Advantages

Respondents emphasized the advantages both to paraprofessionals and to the cataloging departments and libraries. The main advantage to professional catalogers was that involving paraprofessionals in original cataloging freed the professionals' time for more complex cataloging, for management and problem-solving, for professional activities, and for duties in other areas of the library, such as bibliographic instruction. A few respondents also mentioned better morale and job satisfaction among professional catalogers—perhaps because they felt better able to
fulfill truly professional roles—as well as among paraprofessional catalogers.

For the most part, however, respondents focused on the advantages to the paraprofessional catalogers of participating in original cataloging activities. According to respondents, paraprofessionals gained job satisfaction and better morale from applying their experience and their language, format, and subject skills to what the paraprofessionals perceived as the challenging, varied, and interesting tasks of original cataloging. The training and experience in original cataloging enhanced their understanding of cataloging issues and even their copy cataloging work. Some earned tangible rewards of promotions or pay increases, while others experienced intangible gratification such as the “reward of contributing to [a] national database so that other catalogers can use their work.”

Increased job satisfaction led, in turn, according to the respondents, to a better retention rate for paraprofessional catalogers. One pragmatic department head also commented that “technological changes are eliminating the more routine technical services activities. Unless paraprofessionals are performing more advanced activities, their jobs will be eliminated.” As Rider (1996, 29) has noted, “the growing trend to accept a higher percentage of cataloging copy with minimal editing” will provide opportunities for paraprofessionals to perform more complex duties.

ADVANTAGES FOR THE LIBRARY

Respondents also identified many benefits of having paraprofessionals contribute original cataloging. Chief among these were greater productivity and the reduction of backlogs, in part because paraprofessionals generally had more time for cataloging than did professionals. In addition, Benemann (1993, 19) has asserted that “the more complex you make a cataloger’s task, the more productive he or she becomes.” In several cases the paraprofessionals’ original cataloging work filled gaps left by dwindling personnel budgets or by the difficulty of recruiting qualified professional catalogers. Materials were made available more quickly, and access was provided to some materials that would otherwise not have received full cataloging treatment, thereby benefiting library users.

Paraprofessional involvement in original cataloging improved efficiency, which, as one respondent noted, “requires that work be done at the lowest level that is capable of doing it well.” Paraprofessional participation also increased the cost-effectiveness of original cataloging and “allow[ed] the librari[es] to concentrate resources on other things.”

Other advantages to the cataloging departments were the sense of teamwork and partnership that resulted, the increased contributions of paraprofessionals to departmental decision making, and professional development for both paraprofessionals (as catalogers) and professionals (as supervisors). One respondent found paraprofessional participation in original cataloging mirrored the situation in other library departments, where paraprofessionals handled most of the daily operational tasks.

Even in libraries where paraprofessionals did no original cataloging, 17 (89.5%) of those respondents answered the question, although for their libraries the advantages are hypothetical. 14 (82.3%) saw the same potential advantages as reported by cataloging department heads from libraries where paraprofessionals did participate in original cataloging. Three respondents (17.6%) either saw no advantages or found the question irrelevant in their situations.

DISADVANTAGES

The cataloging department heads from 38 (59.4%) of the libraries in which paraprofessionals did some original cataloging described the disadvantages of paraprofessional involvement in original cataloging. They most commonly mentioned the time and personnel resources needed to train and supervise the paraprofessionals and to provide ongoing quality control of their original cataloging work. As one respondent pointed out, “Total person-hours can
be more per title (although since some of the hours are from lower-paid paras, the cost is less)." The time needed for training was a particular drawback if the library's paraprofessional cataloging staff underwent high turnover. Training and supervision required time from the professional catalogers, who did not always like or feel prepared for a management role.

The respondents' other major concern was to avoid exploiting the paraprofessionals who did original cataloging, reflecting the reactions some had observed among the paraprofessional catalogers themselves. Many department heads mentioned the issue of equity for paraprofessional and professional catalogers in terms of both job descriptions and levels of compensation. Some commented that this had not been an issue in their own libraries, but "in some situations some paraprofessionals might [feel] that their positions should be [up]graded and their salaries increased" when they began assuming original cataloging tasks. Professional catalogers, according to the respondents, also reacted negatively at times to the "role blurring" that occurred and feared the downgrading of their own role in cataloging.

Respondents to this question mentioned several other disadvantages or concerns. The educational backgrounds of the paraprofessional catalogers varied from that of professionals, particularly in that some paraprofessionals did not have the foreign-language or subject knowledge needed for original cataloging. This opinion was not unanimous, however; one respondent articulated that the best use of the talents of each employee was critical and that many of the paraprofessionals in that library have advanced degrees, language expertise, and specialized subject knowledge. Some department heads also saw the paraprofessionals' lack of the theoretical background gained from formal study of cataloging principles (usually gained in an M.L.S. program) as well as their lack of involvement in professional activities and lack of awareness of cataloging issues and trends as another factor in the amount of training required. Several respondents reported the difficulty that paraprofessionals faced in switching back and forth between original cataloging and copy cataloging policies and procedures: "in copy cataloging they check [so] much of the bib record that they defeat the purpose of copy cataloging." Another drawback was the inflexibility of civil service rules or union contracts as to the duties paraprofessionals could perform; while department heads in some libraries might have liked to offer paraprofessionals more opportunities to do original cataloging—presumably with appropriate compensation—this was not always possible.

Of the cataloging department heads from the 19 libraries in which paraprofessionals did not perform any original cataloging activities, 14 (73.7%) described the potential disadvantages of having paraprofessionals do original cataloging. This group raised many of the same points as the other group of respondents, but some additional concerns reflected the reasons this group had indicated for not involving paraprofessionals in original cataloging work. Some stated that original cataloging should be a professional task, that having paraprofessionals participate would not be cost-effective, or that "cataloging productivity [might] suffer." One respondent noted that the original catalogers in his or her department could not agree among themselves about which duties the paraprofessionals should perform. In a related concern, other department heads were afraid of possible negative effects on the professionals' morale from the blurring of the line between professional and paraprofessional tasks and the potential downgrading of the profession.

**Conclusion**

The results of this survey highlight continuing changes in assignment of original cataloging responsibilities. Paraprofessionals are assigned some original cataloging tasks in the majority of all ARL libraries receiving surveys—64 of 119 libraries. When the 64 libraries are counted among responding libraries, the percentage becomes much higher—77.1% of the respondents to the survey. Original subject analysis and nonfiction classification are
sometimes done by paraprofessionals in nearly half of all ARL libraries, a clear sign of the trend given that these two activities have traditionally been assigned exclusively to professionals. Original cataloging responsibilities alone are clearly no longer the distinguishing characteristic between professional and paraprofessional cataloging work.

The respondents to this survey expressed many valid concerns with increased paraprofessional involvement in original cataloging. These include paraprofessionals’ lack of the theoretical background that is traditionally obtained by librarians in a graduate degree program, possible lesser commitment to the profession, extensive training time required, and also the issue of equity for the paraprofessionals who do advanced work.

The changes identified by this survey will have far-reaching effects on cataloging departments, but this trend has implications for the profession as well because cataloging and technical services departments have often led the way in employing paraprofessionals in work previously considered professional. The trend will no doubt continue as economic pressures require that libraries become ever more cost-effective, but the desire to do more with less must be weighed against the need for appropriate compensation and recognition for paraprofessionals as they learn new skills and perform new tasks.

The challenge presented to librarians now is to define clearly these new roles for both professionals and paraprofessionals, continuing the groundwork laid by Veanner (1982), Bishoff (1987), Younger (1981), and Rider (1996). A survey among both professionals and paraprofessionals of current perceptions of and satisfaction with their roles might prove enlightening. Comparison with other professions that employ substantial numbers of paraprofessionals could offer some guidance in creating realistic definitions.

More research is also needed to examine the extent to which the profession is maturing (Nettlefold 1989) or its most professional functions are being deskilled (Harris 1992). These contrasting sociological perspectives on our work offer respectively positive and negative views of the increasing paraprofessional involvement in progressively more skilled cataloging activities. Is one or the other more correct? At the same time, Veanner, Bishoff, Younger, and Rider continue to see important roles for professionals in bibliographic access and technical services.

Despite the concerns raised by respondents to this survey, paraprofessionals continue to be more involved in all facets of original cataloging over time. Many heads of cataloging in ARL libraries have clearly discovered that, with training and supervision, paraprofessional catalogers can do a good job on the time-consuming work of original cataloging.

**Works Cited**


Harris, Roma. 1992. Information technology and the de-skilling of librarians; or, the erosion of a woman’s profession. *Computers in libraries* 12, no. 1: 8–16.


Analyzing Search Styles of Patrons and Staff: A Replicative Study of Two University Libraries

Kathlin L. Ray and Mary S. Lang

Librarians at the University of the Pacific (UOP) designed a replication of an earlier transaction log study. We hypothesized that library staff would use a feature that allows the searcher to limit a search by location or material type more often than patrons. We also hypothesized that staff and reference librarians would have a higher success rate than public users. Our third hypothesis was that UOP patrons would perform keyword searches more often than library staff. Studies were conducted in 1995 and 1996 to test these hypotheses. In our two-year endeavor to provide comparative data on the search styles of patrons and staff, we discovered that replicating a study is not nearly as straightforward as we had initially thought. We also found it surprisingly difficult to compare year to year data at the same institution. This was primarily due to a continually changing technological environment.

In a recent article on patron and staff searching styles at Adelphi University Library, Ballard (1994) urged librarians using Innovative Interfaces Inc. Innopac software to collect and analyze system data in order to compare results. Accepting the challenge, librarians at the University of the Pacific designed a replication of Ballard's transaction log study. The decision to replicate this research was a conscious attempt to produce data that would allow us to compare results of two similar institutions. UOP has 4,100 students and Adelphi has 5,600. Both universities have a main library and a science branch. The library collections are roughly similar in size (UOP, 410,000 volumes, Adelphi, 303,000), and both have fewer than 2,700 periodical titles. Staffing is considerably different, however; UOP employs 10.5 librarians, 20 support staff, and 115 student assistants while Adelphi employs 23 librarians, 35 support staff, and 11 student assistants (data from American Library Directory, 49th ed. 1996–97).

While we were interested in comparing UOP and Adelphi patron search patterns, we were particularly eager to compare staff search styles. Until Ballard's work, transaction log studies of library staff searches were virtually nonexistent (Peters 1993). Concurring with Ballard, we hypothesized that staff would use a feature that allows the searcher to limit a search by location or material type more often than patrons. We also hypothesized

KATHLIN L. RAY (kray@uop.edu) is Head, Access Services, University of the Pacific Libraries, Stockton, CA. MARY S. LANG (mlang@uop.edu) is Catalog Librarian, University of the Pacific Libraries, Stockton, CA. Manuscript received MMMM DD, 1997; accepted for publication MMMM DD, 1997.
that staff and reference librarians would have a higher success rate—i.e. fewer zero retrievals—than public users. Our third hypothesis, contrary to Ballard, was that UOP patrons would perform keyword searches more often than library staff.

We assumed that search patterns of UOP patrons and staff would be similar to the search patterns of Adelphi patrons and staff. Ballard (1994) found that patrons and library staff had very different searching strategies, with staff going beyond the standard author, title, subject, and keyword searches eight times more often than patrons. Ballard found that patrons at Adelphi searched most often by subject (35.5%) with title searches a close second (31%), author third (21%), and keyword a distant fourth (8%). Library staff searched most often by title (46%), then by other options such as international standard number or barcode (32%), third by author or subject (16.6%), and fourth by keyword (5%). Staff used the limit feature more often (3%) than the public (2%). In our study, surprisingly, we found that UOP patrons searched the system quite differently than users at Adelphi, preferring keyword searches to subject, title, or author searches. UOP staff, like staff at Adelphi, searched most often by title and least often by subject (public services staff) or keyword (technical services staff).

Search Strategies

Confounding initial assumptions, two decades of analysis have shown that users often prefer to search by subject rather than by author or title (Ballard 1994; Hunter 1991; Larson 1991b; Simpson 1989; Tillotson 1995; Tolle 1983; Zink 1991). However, users also find subject searches to be the most frustrating, primarily due to their ignorance of Library of Congress Subject Headings (LCSH). As Markey (1984) and many others (e.g., Hunter 1991; Wallace 1993; Zink 1991) have pointed out, users do not understand the concept of a controlled vocabulary, the composition of LCSH, or the relationships among the headings. Subject searches are also notorious for retrieving either too many records or none at all (Hunter 1991; Larson 1991a; Peters 1989). And while “zero hit” searches do not necessarily denote an unsuccessful search, a high number of zero retrievals prompts concern about user failure.

Although researchers of many TLA analyses have concluded that users prefer subject searching, others indicate a user preference for keyword. Indeed, there is evidence that the frequency of keyword searching outstrips other types of searching and is more effective at retrieving results than are subject searches. Tillotson (1995) found that searching by keyword retrieved all relevant materials about 50% of the time while subject searches often retrieved nothing. She determined that the large record sets retrieved by keyword searches did not interfere with the quick discovery of useful citations. Larson (1991b) observed a consistent decline in the use of the subject index and a corresponding rise in the use of title keyword in his study of the University of California's MELVYL catalog. Martin, Wyman, and Madhok (1983) found that users of the SULIRS system at Syracuse University preferred to search by keyword (36%)
more often than by subject (26%).

Furthermore, many researchers have recommended the use of keyword searching as a way to circumvent the frustrations of searches based on LCSH terms. Cherry (1992) suggested that catalogs be programmed to automatically run a keyword search on any subject search that retrieves zero hits. Zink (1991) advocated a “truth-in-labeling” approach to inform patrons of the difference between a keyword and an exact subject search. He recommended that the subject search screen include a statement instructing searchers to use LCSH to construct queries.

Despite the popularity of subject and keyword searches, however, other researchers show title searching to be the preferred search strategy (Peters 1989; Cherry 1992; Connaway, Budd, and Kochtanek 1995). Cherry (1992) found that users performed title searches 52% of the time (16% of these were keyword title searches). Cherry found that keyword searching accounted for 16% and subject searching 17% of the total searches conducted. In another recent study, Connaway, Budd, and Kochtanek (1995) found that title searching exceeded LCSH subject and subject browse searches combined—which accounted for nearly one-third of all searches.

Transaction log research has played a major and vital role in increasing librarians’ knowledge of patron search patterns and preferences. Unfortunately, data gathered from individual studies are not usually generalizable to a larger population. This is due to a variety of factors, including: the diversity of systems and search options, type of data collected, the fields searched, and time periods studied. For example, simply attempting to compare keyword search results is fraught with difficulties. A keyword search in some systems will look in subject, title and content notes fields; in other systems, keyword may invoke title fields only. One system may have separate keyword indexes for title, author and subject and another a browse mode that permits a quasi-keyword search. We hoped to overcome these obstacles by replicating a transaction log study conducted at an institution of similar size, using the same automated system, and following the same methodology.

**Methodology**

When we initiated the study in February 1995, public access to Innopac was available through twelve networked stations in the university’s main and branch libraries. Five dial-up ports were available for remote users. Library staff accessed Innopac through 12 terminals and numerous PC workstations in offices. Staff occasionally searched at public terminals but agreed to restrict their work to non-public terminals during the three months of the study. Utilizing Innopac’s ability to collect terminal-specific transaction data, we followed Ballard’s example and created terminal groups according to the categories of users we wanted to study. We expanded Ballard’s categories to include a separate category for reference librarians, hypothesizing that the search strategies of knowledgeable and experienced librarians would be very different than those of public users and even clerical staff. We created four terminal groups: (1) terminals used by public users, (2) terminals used by technical services staff, (3) terminals used by public services staff, and (4) terminals used by reference librarians. Innopac’s data management system compiled detailed transaction data according to the terminal groups we designated.

Transaction logs for each terminal group were printed out and the file cleared every Wednesday. The first three weeks of data were used as a pretest in order to detect errors and fine tune the terminal groupings. By the last week in February, we were ready to begin the study. Every Wednesday at 7:00 a.m. for 14 weeks (from the fourth week of February through the third week of May) we printed out search data for each terminal group. The data were then entered into a spreadsheet and organized by month and type of search (see table 1). During the
### TABLE 1

**1995**

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Limit</th>
<th>No Hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>March</td>
<td>29,841</td>
<td>670</td>
<td>9,356</td>
<td>10,657</td>
<td>155</td>
<td>4,212</td>
<td>1,476</td>
<td>12,658</td>
<td>838</td>
</tr>
<tr>
<td>April</td>
<td>27,962</td>
<td>672</td>
<td>10,120</td>
<td>8,940</td>
<td>110</td>
<td>4,477</td>
<td>1,009</td>
<td>12,815</td>
<td>721</td>
</tr>
<tr>
<td>May</td>
<td>26,238</td>
<td>560</td>
<td>8,094</td>
<td>8,917</td>
<td>116</td>
<td>3,894</td>
<td>1,043</td>
<td>11,479</td>
<td>905</td>
</tr>
<tr>
<td>Average</td>
<td>28,013.67</td>
<td>634.00</td>
<td>9,190.00</td>
<td>9,504.67</td>
<td>127.00</td>
<td>4,194.33</td>
<td>1,176.00</td>
<td>12,317.33</td>
<td>821.33</td>
</tr>
<tr>
<td>Total</td>
<td>84,041</td>
<td>1,902</td>
<td>27,570</td>
<td>28,514</td>
<td>381</td>
<td>12,583</td>
<td>3,528</td>
<td>36,952</td>
<td>2,464</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>2.26</td>
<td>32.81</td>
<td>33.93</td>
<td>0.45</td>
<td>14.97</td>
<td>4.20</td>
<td>43.97</td>
<td>2.93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patrons</th>
<th>Total</th>
<th>Limit</th>
<th>No Hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>24,339</td>
<td>524</td>
<td>7,835</td>
<td>9,803</td>
<td>139</td>
<td>3,714</td>
<td>941</td>
<td>9,450</td>
<td>431</td>
</tr>
<tr>
<td>April</td>
<td>21,038</td>
<td>396</td>
<td>7,084</td>
<td>8,283</td>
<td>103</td>
<td>3,217</td>
<td>732</td>
<td>8,372</td>
<td>434</td>
</tr>
<tr>
<td>May</td>
<td>18,611</td>
<td>318</td>
<td>5,686</td>
<td>8,147</td>
<td>108</td>
<td>3,204</td>
<td>701</td>
<td>6,141</td>
<td>418</td>
</tr>
<tr>
<td>Average</td>
<td>21,329.33</td>
<td>412.67</td>
<td>6,868.33</td>
<td>8,744.33</td>
<td>116.67</td>
<td>3,378.33</td>
<td>791.33</td>
<td>7,987.67</td>
<td>427.67</td>
</tr>
<tr>
<td>Total</td>
<td>63,988</td>
<td>1,238</td>
<td>20,605</td>
<td>26,233</td>
<td>350</td>
<td>10,135</td>
<td>2,374</td>
<td>23,963</td>
<td>1,283</td>
</tr>
<tr>
<td>%</td>
<td>100.00</td>
<td>1.93</td>
<td>32.20</td>
<td>41.00</td>
<td>0.55</td>
<td>15.84</td>
<td>3.71</td>
<td>37.45</td>
<td>2.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Staff</th>
<th>Total</th>
<th>Limit</th>
<th>No Hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>5,502</td>
<td>146</td>
<td>1,521</td>
<td>854</td>
<td>16</td>
<td>498</td>
<td>535</td>
<td>3,208</td>
<td>407</td>
</tr>
<tr>
<td>April</td>
<td>6,940</td>
<td>276</td>
<td>2,682</td>
<td>660</td>
<td>7</td>
<td>1,264</td>
<td>277</td>
<td>4,452</td>
<td>287</td>
</tr>
<tr>
<td>May</td>
<td>7,630</td>
<td>242</td>
<td>2,409</td>
<td>770</td>
<td>8</td>
<td>690</td>
<td>342</td>
<td>5,341</td>
<td>487</td>
</tr>
<tr>
<td>Average</td>
<td>6,690.67</td>
<td>221.33</td>
<td>2,204.00</td>
<td>761.33</td>
<td>10.33</td>
<td>817.33</td>
<td>384.67</td>
<td>4,333.67</td>
<td>393.67</td>
</tr>
<tr>
<td>Total</td>
<td>20,072</td>
<td>664</td>
<td>6,612</td>
<td>2,284</td>
<td>31</td>
<td>2,452</td>
<td>1,154</td>
<td>13,001</td>
<td>1,181</td>
</tr>
<tr>
<td>%</td>
<td>100.00</td>
<td>3.31</td>
<td>32.94</td>
<td>11.38</td>
<td>0.15</td>
<td>12.22</td>
<td>5.75</td>
<td>64.77</td>
<td>5.88</td>
</tr>
<tr>
<td>PS Staff</td>
<td>Total</td>
<td>Limit</td>
<td>No hit</td>
<td>Keyword</td>
<td>Kword &quot;W&quot;</td>
<td>Author</td>
<td>Subject</td>
<td>Title</td>
<td>Other</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>March</td>
<td>4,751</td>
<td>146</td>
<td>1,307</td>
<td>818</td>
<td>16</td>
<td>472</td>
<td>202</td>
<td>2,902</td>
<td>357</td>
</tr>
<tr>
<td>April</td>
<td>6,144</td>
<td>268</td>
<td>2,405</td>
<td>614</td>
<td>5</td>
<td>1,218</td>
<td>146</td>
<td>3,921</td>
<td>245</td>
</tr>
<tr>
<td>May</td>
<td>5,520</td>
<td>242</td>
<td>1,594</td>
<td>730</td>
<td>8</td>
<td>576</td>
<td>242</td>
<td>3,542</td>
<td>430</td>
</tr>
<tr>
<td>Average</td>
<td>5,471.67</td>
<td>218.67</td>
<td>1,768.67</td>
<td>720.67</td>
<td>9.67</td>
<td>755.33</td>
<td>196.67</td>
<td>3,455.00</td>
<td>344.00</td>
</tr>
<tr>
<td>Total</td>
<td>16,415</td>
<td>656</td>
<td>5,306</td>
<td>2,162</td>
<td>29</td>
<td>2,266</td>
<td>590</td>
<td>10,365</td>
<td>1,032</td>
</tr>
<tr>
<td>%</td>
<td>100.00</td>
<td>4.00</td>
<td>32.32</td>
<td>13.17</td>
<td>0.18</td>
<td>13.80</td>
<td>3.59</td>
<td>63.14</td>
<td>6.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TS Staff</th>
<th>Total</th>
<th>Limit</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>751</td>
<td>0</td>
<td>214</td>
<td>36</td>
<td>0</td>
<td>26</td>
<td>333</td>
<td>306</td>
<td>50</td>
</tr>
<tr>
<td>April</td>
<td>790</td>
<td>8</td>
<td>277</td>
<td>45</td>
<td>2</td>
<td>46</td>
<td>131</td>
<td>526</td>
<td>42</td>
</tr>
<tr>
<td>May</td>
<td>2,107</td>
<td>0</td>
<td>814</td>
<td>40</td>
<td>0</td>
<td>114</td>
<td>100</td>
<td>1,796</td>
<td>57</td>
</tr>
<tr>
<td>Average</td>
<td>1,216.00</td>
<td>2.67</td>
<td>435.00</td>
<td>40.33</td>
<td>0.67</td>
<td>62.00</td>
<td>188.00</td>
<td>876.00</td>
<td>49.67</td>
</tr>
<tr>
<td>Total</td>
<td>3,648</td>
<td>8</td>
<td>1,305</td>
<td>121</td>
<td>2</td>
<td>186</td>
<td>564</td>
<td>2,628</td>
<td>149</td>
</tr>
<tr>
<td>%</td>
<td>100.00</td>
<td>0.22</td>
<td>35.77</td>
<td>3.32</td>
<td>0.05</td>
<td>5.10</td>
<td>15.46</td>
<td>72.04</td>
<td>4.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Total</th>
<th>Limit</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>2,203</td>
<td>22</td>
<td>587</td>
<td>404</td>
<td>12</td>
<td>144</td>
<td>129</td>
<td>1,490</td>
<td>36</td>
</tr>
<tr>
<td>April</td>
<td>2,677</td>
<td>17</td>
<td>896</td>
<td>411</td>
<td>3</td>
<td>261</td>
<td>93</td>
<td>1,895</td>
<td>17</td>
</tr>
<tr>
<td>May</td>
<td>2,722</td>
<td>29</td>
<td>755</td>
<td>322</td>
<td>5</td>
<td>167</td>
<td>192</td>
<td>1,999</td>
<td>42</td>
</tr>
<tr>
<td>Average</td>
<td>2,534.00</td>
<td>22.67</td>
<td>746.00</td>
<td>379.00</td>
<td>6.67</td>
<td>190.67</td>
<td>138.00</td>
<td>1,794.67</td>
<td>31.67</td>
</tr>
<tr>
<td>Total</td>
<td>7,602</td>
<td>68</td>
<td>2,238</td>
<td>1,137</td>
<td>20</td>
<td>572</td>
<td>414</td>
<td>5,384</td>
<td>95</td>
</tr>
<tr>
<td>%</td>
<td>100.00</td>
<td>0.89</td>
<td>29.44</td>
<td>14.96</td>
<td>0.26</td>
<td>7.52</td>
<td>5.45</td>
<td>70.82</td>
<td>1.25</td>
</tr>
</tbody>
</table>
study period there were a total of 84,041 staff and patron searches.

1995 Pilot Study Results

The results were puzzling. The first anomaly was a substantially higher percentage of title searches by UOP library staff than at Adelphi (see table 2). Upon investigation, we found that the high number of title searches might have been caused by the sizable one-time supplemental funds for collection development received by the library in early 1995. Many staff members, in public services as well as technical services, were involved in checking the catalog for holdings, searching duplicate titles, etc. We also discov-
Welcome to University of the Pacific

You may search for library materials by any of the following:

A> AUTHOR
T> TITLE
K> KEYWORDS in the title, subject and contents
L> Library of Congress Subject Headings
C> CALL NO
I> INTERNATIONAL standard number
H> Library HELP and Information
R> RESERVE Lists
D> Connect to DATABASES, libraries and the Internet
Q> QUIT
Choose one (A,T,K,L,C,I,H,R,D,Q)

Older books (prior to 1982) may not be here; check the card catalog.
Please select the Library Help and Information screens <H> to learn more about PacifiCat. For more assistance, ask at the Reference Desk.

Figure 1. UOP 1995 Introductory Search Screen

It was observed that several reference librarians had been heavily involved in the collection development project and had routinely searched Innopac for titles while at the reference desk and when working in their offices. Since terminals in their offices and at the reference desk had been included in the reference terminal group, the data we collected did not accurately reflect the typical search strategies used by reference librarians.

The second and most dramatic anomaly was the unbelievably low number of subject searches (3.7%) at public terminals and the unusually high number of keyword searches (41%) in comparison to any study we had seen, especially in comparison to the one at Adelphi (see table 2). It was difficult to believe that our students really did search by keyword eleven times more often than by subject. We also looked closely at patron title searches. While the number of title searches performed by our users was not substantially higher than Adelphi's, we had observed a staff member using a public terminal during the course of the study. We were concerned that staff had skewed the patron data by using public terminals. While subsequent conversations calmed many of our misgivings, a couple of staff admitted they had forgotten and had used public terminals during the three months of the study. Searching infrequently and almost exclusively by title, staff impact on total patron transactions was presumably minimal but remained a concern.

Screen Menus

From observation and user feedback, we suspected Innopac's search screens might have contributed to the atypical patron search patterns. On the main menu the keyword search option appeared just below author and title options (see figure 1). The subject search option, below keyword, was not labeled "Subject" but "Library of Congress Subject Headings" to emphasize the need to use a specific vocabulary. The naming and order of the search options had not been accidental. Following Zink's (1991) suggestion, the menu had been designed to encourage keyword searching and thereby reduce student frustration with LCSH searches. Reference librarians had strongly empha-
Welcome to University of the Pacific \hspace{20pt} PacifiCat

You may search for library materials by any of the following:

A>AUTHOR
T>TITLE
K>KEYWORDS in the title, subject and contents
S>SUBJECT (Library of Congress Subject Headings)
C>CALL NO
I>INTERNATIONAL standard number
H>Library HELP and Information
R>RESERVE Lists
V>VIEW your circulation record
P>Repeat PREVIOUS Search

Q RETURN TO GATEWAY MENU

Older books (prior to 1982) may not be listed; check the card catalog.
Please select the Library Help and Information screens <H> to learn more about PacifiCat. For more assistance, ask at the Reference Desk, 2433.

Figure 2. UOP 1996 Introductory Search Screen.

sized keyword searching in instruction sessions and one-on-one at the reference desk. Perhaps the effort had succeeded all too well; subject searching at the library had all but disappeared.

\textbf{1996 STUDY}

Due to the obstacles encountered in the 1995 study, we decided to repeat the study one year later in February 1996. In addition to correcting any errors in the 1995 staff and patron search data, we wanted to look at the influence of the search menu and search option labels on patron search strategies. Given our findings in the 1995 study, we developed a hypothesis that the labeling of search options would have an effect on the way in which patrons searched the catalog. If the naming of an option influenced use patterns, then a change in search labels should cause a shift in the way the public used the system. The subject search option on the main menu was simply relabeled "Subject" followed by "Library of Congress Subject Headings" in parentheses (see figure 2). We retained the order of the options—keyword preceding subject—on the main search screen. Another factor, which had nothing to do with the design of our study but influenced the results nonetheless, was a change from an initial online catalog screen to a gateway screen in November 1995. This gateway menu highlighted periodical indexes and other libraries' catalogs and was the initial screen on all of our public stations (see figure 3).

In addition to making a change in the main search screen, we created a separate transaction group for terminals located at the reference desk only. Thus, we separated staff into three groups: reference staff on the one hand, and public and technical staff on the other. By excluding terminals in librarian offices we hoped the data would more clearly identify the search patterns of librarians at the reference desk, particularly their interactions with patrons. Terminals at the reference desk were only operational during the hours the desk was open. We added another transaction group to capture search patterns of remote users and began tracking searches that yielded record sets of 500 or more. Finally, we reminded staff
often not to use public terminals during the study period.

In 1996, we again monitored Innopac's transaction logs for 14 weeks from mid-February to mid-May (see table 3). We assumed that the 1996 data would confirm our hypothesis with regards to menu search labels as well as permit us to take a more accurate look at staff, on-site patron and remote patron search patterns.

1996 Results

Patron Searches

The results of the 1996 study were dramatically different than those of the pilot study (see table 4). Subject searches at public terminals increased 600% (from 3.7% to 23.7%) while keyword searches dropped almost 12% and title searches dropped 10%. Despite the decline, however, patrons still searched most often by keyword with title searches a close second. In both 1995 and 1996, UOP students used keyword more often than any other type of search. However, the dramatic rise in subject searching in 1996 appeared to confirm our hypothesis that the labeling of options on the search menu had an influence on the way users searched the catalog. The order in which the search options appeared on the screen (keyword before subject) was probably a contributing factor as well. Ballard (1994) found that keyword searching increased when it was listed as the first option on the search menu. If the subject option had come before the keyword option on the menu, perhaps keyword usage would have dropped even further.

Use of the online catalog’s advanced features was quite low. Consistent with Ballard’s findings, patrons were more likely than staff to rerun a failed title or subject search in keyword by pressing the “w” key but they still used this feature less than 1% of the time (see table 4). More patrons at UOP utilized the limit feature (which narrows a search by material type, location, etc.) than did patrons at Adelphi, and remote users used it more often than onsite users.

Ballard (1994, 303) found that the biggest difference between staff and patron searches was that patrons were ten times more likely to get “unmanageably large” (over 500) search results. Patrons in our study retrieved record sets with 500 or more items 3.7% of the time, remote users retrieved large sets 4.7% of the time, and staff retrieved large sets 2.8% of the time (see table 5). However, while our analysis confirmed that public users are much more likely than staff to retrieve large sets, the actual numbers are small and did not seem to indicate a major problem. Public and staff alike retrieved zero hits about one third of the time (34% and 35% respectively) with remote users faring slightly worse than onsite searchers (see table 6). UOP users were moderately more successful than Adelphi users, who
### TABLE 3

**1996**

<table>
<thead>
<tr>
<th>All</th>
<th>Total</th>
<th>Limit</th>
<th>Over 500</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword “W”</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>16,642</td>
<td>3,571</td>
<td>521</td>
<td>5,586</td>
<td>3,578</td>
<td>96</td>
<td>3,648</td>
<td>2,488</td>
<td>6,459</td>
<td>469</td>
</tr>
<tr>
<td>April</td>
<td>15,202</td>
<td>696</td>
<td>404</td>
<td>5,287</td>
<td>3,219</td>
<td>96</td>
<td>2,485</td>
<td>2,720</td>
<td>6,218</td>
<td>560</td>
</tr>
<tr>
<td>May</td>
<td>10,824</td>
<td>591</td>
<td>329</td>
<td>3,509</td>
<td>2,358</td>
<td>78</td>
<td>1,855</td>
<td>2,115</td>
<td>4,168</td>
<td>328</td>
</tr>
<tr>
<td>Average</td>
<td>14,222.67</td>
<td>1,619.33</td>
<td>418</td>
<td>4,794</td>
<td>3,051.67</td>
<td>90</td>
<td>2,662.67</td>
<td>2,441</td>
<td>5,615</td>
<td>52.33</td>
</tr>
<tr>
<td>Total</td>
<td>42,668</td>
<td>4,858</td>
<td>1,254</td>
<td>14,382</td>
<td>9,155</td>
<td>270</td>
<td>7,988</td>
<td>7,323</td>
<td>16,845</td>
<td>1,357</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>11.39</td>
<td>2.94</td>
<td>33.71</td>
<td>21.46</td>
<td>0.63</td>
<td>18.72</td>
<td>17.16</td>
<td>39.48</td>
<td>3.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patrons</th>
<th>Total</th>
<th>Limit</th>
<th>Over 500</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword “W”</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>9,760</td>
<td>346</td>
<td>370</td>
<td>3,214</td>
<td>2,945</td>
<td>85</td>
<td>1,796</td>
<td>2,126</td>
<td>2,714</td>
<td>179</td>
</tr>
<tr>
<td>April</td>
<td>9,207</td>
<td>259</td>
<td>326</td>
<td>3,301</td>
<td>2,740</td>
<td>90</td>
<td>1,577</td>
<td>2,277</td>
<td>2,465</td>
<td>148</td>
</tr>
<tr>
<td>May</td>
<td>7,505</td>
<td>245</td>
<td>276</td>
<td>2,611</td>
<td>2,039</td>
<td>72</td>
<td>1,437</td>
<td>1,871</td>
<td>2,046</td>
<td>112</td>
</tr>
<tr>
<td>Average</td>
<td>8,824</td>
<td>283.33</td>
<td>324</td>
<td>3,042</td>
<td>2,574.67</td>
<td>82.33</td>
<td>1,603.33</td>
<td>2,091.33</td>
<td>2,408.33</td>
<td>146.33</td>
</tr>
<tr>
<td>Total</td>
<td>26,472</td>
<td>850</td>
<td>972</td>
<td>9,126</td>
<td>7,724</td>
<td>247</td>
<td>4,810</td>
<td>6,274</td>
<td>7,225</td>
<td>439</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>3.21</td>
<td>3.67</td>
<td>34.47</td>
<td>29.18</td>
<td>0.93</td>
<td>18.17</td>
<td>23.7</td>
<td>27.29</td>
<td>1.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Remote</th>
<th>Total</th>
<th>Limit</th>
<th>Over 500</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword “W”</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>1,944</td>
<td>71</td>
<td>82</td>
<td>730</td>
<td>527</td>
<td>29</td>
<td>405</td>
<td>511</td>
<td>451</td>
<td>50</td>
</tr>
<tr>
<td>April</td>
<td>1,722</td>
<td>78</td>
<td>96</td>
<td>629</td>
<td>476</td>
<td>27</td>
<td>355</td>
<td>486</td>
<td>361</td>
<td>44</td>
</tr>
<tr>
<td>May</td>
<td>769</td>
<td>41</td>
<td>29</td>
<td>276</td>
<td>200</td>
<td>6</td>
<td>188</td>
<td>177</td>
<td>185</td>
<td>19</td>
</tr>
<tr>
<td>Average</td>
<td>1,478.33</td>
<td>63.33</td>
<td>69</td>
<td>545</td>
<td>401</td>
<td>20.67</td>
<td>316</td>
<td>391.33</td>
<td>332.33</td>
<td>37.67</td>
</tr>
<tr>
<td>Total</td>
<td>4,435</td>
<td>190</td>
<td>207</td>
<td>1,635</td>
<td>1,203</td>
<td>62</td>
<td>948</td>
<td>1,174</td>
<td>997</td>
<td>113</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>4.28</td>
<td>4.67</td>
<td>36.87</td>
<td>27.13</td>
<td>1.4</td>
<td>21.38</td>
<td>26.47</td>
<td>22.48</td>
<td>2.55</td>
</tr>
<tr>
<td>PS Staff</td>
<td>Total</td>
<td>Limit</td>
<td>Over 500</td>
<td>No hit</td>
<td>Keyword</td>
<td>Kword &quot;W&quot;</td>
<td>Author</td>
<td>Subject</td>
<td>Title</td>
<td>Other</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>March</td>
<td>5,155</td>
<td>3,216</td>
<td>135</td>
<td>1,860</td>
<td>491</td>
<td>7</td>
<td>1,739</td>
<td>157</td>
<td>2,528</td>
<td>240</td>
</tr>
<tr>
<td>April</td>
<td>3,876</td>
<td>429</td>
<td>61</td>
<td>1,427</td>
<td>309</td>
<td>4</td>
<td>714</td>
<td>106</td>
<td>2,401</td>
<td>346</td>
</tr>
<tr>
<td>May</td>
<td>2,128</td>
<td>339</td>
<td>43</td>
<td>613</td>
<td>211</td>
<td>3</td>
<td>331</td>
<td>90</td>
<td>1,313</td>
<td>183</td>
</tr>
<tr>
<td>Average</td>
<td>3,719.67</td>
<td>1,328</td>
<td>79.67</td>
<td>1,300</td>
<td>.337</td>
<td>4.67</td>
<td>928</td>
<td>117.67</td>
<td>2,080.67</td>
<td>256.33</td>
</tr>
<tr>
<td>Total</td>
<td>11,159</td>
<td>3,984</td>
<td>239</td>
<td>3,900</td>
<td>1,011</td>
<td>14</td>
<td>2,784</td>
<td>353</td>
<td>6,242</td>
<td>769</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>35.7</td>
<td>2.14</td>
<td>34.95</td>
<td>9.06</td>
<td>0.13</td>
<td>24.95</td>
<td>3.16</td>
<td>55.94</td>
<td>6.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TS Staff</th>
<th>Total</th>
<th>Limit</th>
<th>Over 500</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>1,178</td>
<td>9</td>
<td>10</td>
<td>363</td>
<td>20</td>
<td>1</td>
<td>72</td>
<td>138</td>
<td>914</td>
<td>34</td>
</tr>
<tr>
<td>April</td>
<td>1,621</td>
<td>0</td>
<td>11</td>
<td>408</td>
<td>29</td>
<td>0</td>
<td>152</td>
<td>251</td>
<td>1,134</td>
<td>55</td>
</tr>
<tr>
<td>May</td>
<td>801</td>
<td>3</td>
<td>5</td>
<td>195</td>
<td>26</td>
<td>0</td>
<td>64</td>
<td>95</td>
<td>588</td>
<td>28</td>
</tr>
<tr>
<td>Average</td>
<td>1,200</td>
<td>4</td>
<td>8.67</td>
<td>322</td>
<td>25</td>
<td>0.33</td>
<td>96</td>
<td>161.33</td>
<td>878.67</td>
<td>39</td>
</tr>
<tr>
<td>Total</td>
<td>3,600</td>
<td>12</td>
<td>26</td>
<td>966</td>
<td>75</td>
<td>1</td>
<td>288</td>
<td>484</td>
<td>2,636</td>
<td>117</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>0.33</td>
<td>0.72</td>
<td>26.83</td>
<td>2.08</td>
<td>0.03</td>
<td>8</td>
<td>13.44</td>
<td>73.22</td>
<td>3.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ref Desk</th>
<th>Total</th>
<th>Limit</th>
<th>Over 500</th>
<th>No hit</th>
<th>Keyword</th>
<th>Kword &quot;W&quot;</th>
<th>Author</th>
<th>Subject</th>
<th>Title</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>492</td>
<td>0</td>
<td>7</td>
<td>141</td>
<td>122</td>
<td>3</td>
<td>38</td>
<td>67</td>
<td>250</td>
<td>15</td>
</tr>
<tr>
<td>April</td>
<td>500</td>
<td>9</td>
<td>6</td>
<td>130</td>
<td>141</td>
<td>2</td>
<td>39</td>
<td>86</td>
<td>223</td>
<td>11</td>
</tr>
<tr>
<td>May</td>
<td>395</td>
<td>4</td>
<td>5</td>
<td>92</td>
<td>84</td>
<td>3</td>
<td>23</td>
<td>61</td>
<td>221</td>
<td>6</td>
</tr>
<tr>
<td>Average</td>
<td>462.33</td>
<td>4.33</td>
<td>6</td>
<td>121</td>
<td>115.67</td>
<td>2.67</td>
<td>33.33</td>
<td>71.33</td>
<td>231.33</td>
<td>10.67</td>
</tr>
<tr>
<td>Total</td>
<td>1,387</td>
<td>13</td>
<td>18</td>
<td>363</td>
<td>347</td>
<td>8</td>
<td>100</td>
<td>214</td>
<td>694</td>
<td>32</td>
</tr>
<tr>
<td>%</td>
<td>100</td>
<td>0.94</td>
<td>1.3</td>
<td>26.17</td>
<td>25.02</td>
<td>0.58</td>
<td>7.21</td>
<td>15.43</td>
<td>50.04</td>
<td>2.31</td>
</tr>
</tbody>
</table>
conducted zero hit searches 39% of the time. Librarians at the reference desk were the most successful, with a zero hit retrieval rate of 26%.

Remote users performed 10% of the public searches in 1996. Because we had not gathered data separately for dial-up users in 1995, we could not analyze the impact of menu changes on this group's search strategies. Like onsite library users, remote users searched most often by keyword followed closely by subject (see table 4). They used Innopac's special features more often than on-site users, especially the limit feature. Not surprisingly, remote users relied on system help screens more often than other users (one-third of all help sessions were called up by remote users). Due to the fact that all remote users enter the system through one port, we were not able to determine whether the majority of our remote users were UOP students or unaffiliated users. We plan to look more closely at our remote users in a future transaction log study.

Reference Desk

Data gathered from the reference desk terminals showed that 50.4% of the catalog searches conducted by reference librarians were title searches (see table 7). The preponderance of title searches is understandable given the number of inquiries received about library holdings. What was less clear was why librarians (15%) searched by subject less often than patrons (24%). When asked, several librarians said they preferred teaching students keyword searching because the Boolean logic capability was more important than the conceptually inclusive overview gained by browsing subject headings. They also felt that many students are confused by the list of subheadings retrieved with a subject search and found the keyword results more straightforward. Reference librarians were also less likely than patrons to use the limit and rerun-as-keyword features.

**TABLE 4**

<table>
<thead>
<tr>
<th></th>
<th>1995 %</th>
<th>1996 %</th>
<th>Remote 1</th>
<th>Adelphi %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author</strong></td>
<td>15.8</td>
<td>18.2</td>
<td>21.4</td>
<td>21.0</td>
</tr>
<tr>
<td><strong>Title</strong></td>
<td>37.5</td>
<td>27.3</td>
<td>22.5</td>
<td>31.1</td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td>41.0</td>
<td>29.2</td>
<td>27.1</td>
<td>8.3</td>
</tr>
<tr>
<td><strong>Subject</strong></td>
<td>3.7</td>
<td>23.7</td>
<td>26.5</td>
<td>35.5</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>2.0</td>
<td>1.7</td>
<td>2.6</td>
<td>4.1</td>
</tr>
<tr>
<td><strong>Limit</strong></td>
<td>1.9</td>
<td>3.2</td>
<td>4.3</td>
<td>1.8</td>
</tr>
<tr>
<td><strong>Kword &quot;w&quot;</strong></td>
<td>.6</td>
<td>.9</td>
<td>1.4</td>
<td>1.8</td>
</tr>
</tbody>
</table>

1 Subset of 1996 data

**TABLE 6**

<table>
<thead>
<tr>
<th></th>
<th>1995 (%)</th>
<th>1996 (%)</th>
<th>Adelphi (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patrons</td>
<td>3.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remote</td>
<td>4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech. staff</td>
<td>.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public staff</td>
<td>2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference Desk</td>
<td>1.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 5**

<table>
<thead>
<tr>
<th>Over 500 Retrieved - 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patrons</td>
</tr>
<tr>
<td>Remote</td>
</tr>
<tr>
<td>Tech. staff</td>
</tr>
<tr>
<td>Public staff</td>
</tr>
<tr>
<td>Reference Desk</td>
</tr>
</tbody>
</table>

**TABLE 7**

<table>
<thead>
<tr>
<th></th>
<th>1995 (%)</th>
<th>1996 (%)</th>
<th>Adelphi (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patrons</td>
<td>32</td>
<td>34</td>
<td>39</td>
</tr>
<tr>
<td>Remote</td>
<td>n/a</td>
<td>37</td>
<td>n/a</td>
</tr>
<tr>
<td>Tech. staff</td>
<td>36</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td>Public staff</td>
<td>32</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>Reference Desk</td>
<td>n/a</td>
<td>26</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Staff Searching

While investigating the low use of other indexes by our technical services staff, we discovered that Innopac's transaction logs collected only catalog search data, not database search data. This was the biggest surprise (and disappointment) of our effort to replicate Ballard's study. Innopac's transaction logs do not capture staff searching as we normally think of it. Innopac transaction data, although extensive and detailed, are limited to catalog searches and do not include searches of the database, which typically constitute the bulk of staff searching activity at UOP.

The difference between the Innopac catalog and the Innopac database was not a distinction we were expecting to confront in our analysis. Innopac's bibliographic catalog, in addition to the public mode, can be searched in a staff mode that allows viewing of all fields in the Machine-Readable Cataloging (MARC) format, as well as attached order records, item records, check-in records and so on, much of which understandably is hidden from public display. While searching the catalog is a quick and convenient way for library staff to retrieve certain records or information, it does not allow them to update any information or fields in retrieved records, and thus is seldom the preferred mode of searching by staff other than those doing reference or collection development work. Most staff searching, particularly in technical services, takes place within database update functions—in acquisitions, circulation, serials, and cataloging modules. These searches—to receive an ordered book, to check-in a periodical issue, to enter a new barcode and item status for a book returned from the bindery, to look for an existing record before creating an on-the-fly record within circulation—are ignored by Innopac's transactions logs. Yet these are the searches that would reveal the most about staff searching styles.

Ballard is vague on this distinction. He does not assert that his analysis included these types of staff searching, but neither does he explain how it is possible to analyze staff searching styles meaningfully with logs that, in our view, exclude the bulk of searches that staff might perform on Innopac. Ballard includes a graph that illustrates technical services staff search percentages, and we can only surmise what would lead technical services or circulation staff to do so much title searching in the catalog and not in a database update mode. At UOP our technical services processes are much more tightly integrated than they were a few years ago. For instance, pre-order searching generally takes place seamlessly as part of the order entry transaction. Perhaps at Adelphi, pre-order searches are conducted in a separate process using the staff mode of the catalog. This type of difference in pre-order processes would account for the substantial differences in technical services staff search patterns at the two libraries. It appears necessary to understand the organizational and procedural arrangements at libraries being compared before any meaningful conclusions about their staff searching patterns can be drawn.

Lacking database search data, an analysis of staff catalog searches is likely to be of limited value. However, a few observations are in order, if only to compare with Adelphi's findings. Our initial assumption that staff in public service areas (excluding the reference desk) searched primarily for specific items when searching the catalog proved correct (see table 8). Title and author searches accounted for 80.8% of their searches. Technical services staff also searched most often by title and author but performed four times as many subject

<table>
<thead>
<tr>
<th>Reference Desk* - 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author</td>
</tr>
<tr>
<td>Title</td>
</tr>
<tr>
<td>Keyword</td>
</tr>
<tr>
<td>Subject</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Limit</td>
</tr>
<tr>
<td>Kword Avg</td>
</tr>
</tbody>
</table>

* 1995 not included due to skewed data
TABLE 8
PUBLIC SERVICES STAFF CATALOG SEARCHES

<table>
<thead>
<tr>
<th></th>
<th>1995</th>
<th></th>
<th>1996</th>
<th></th>
<th>1996</th>
<th>Adelphi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>Total</td>
<td>%</td>
<td>Total</td>
<td>%</td>
<td>Total</td>
</tr>
<tr>
<td>Author</td>
<td>13.8</td>
<td>2,266</td>
<td>24.9</td>
<td>2,784</td>
<td>9.3</td>
<td>1,695</td>
</tr>
<tr>
<td>Title</td>
<td>63.1</td>
<td>10,365</td>
<td>55.9</td>
<td>6,242</td>
<td>48.8</td>
<td>8,859</td>
</tr>
<tr>
<td>Keyword</td>
<td>13.2</td>
<td>2,162</td>
<td>9.1</td>
<td>1,011</td>
<td>7.9</td>
<td>1,432</td>
</tr>
<tr>
<td>Subject</td>
<td>3.6</td>
<td>590</td>
<td>3.2</td>
<td>353</td>
<td>10.6</td>
<td>1,918</td>
</tr>
<tr>
<td>Other</td>
<td>6.3</td>
<td>1,032</td>
<td>6.9</td>
<td>769</td>
<td>23.5</td>
<td>4,261</td>
</tr>
<tr>
<td>Limit</td>
<td>4.0</td>
<td>656</td>
<td>35.7</td>
<td>3,984</td>
<td>4.8</td>
<td>876</td>
</tr>
</tbody>
</table>

searches as public staff (see table 9). Because librarians in technical services often use the catalog for collection development purposes, a higher incidence of subject searching was expected. Public services staff searched by keyword far less often than patrons. They also used the limit feature far more often than technical services staff. We speculate that this is attributable primarily to the staff in the Music/AV area who use the limit option to identify scores, videos, compact discs and other audio-visual materials.

A curious difference between staff searches at Adelphi and UOP was in the "Other" category. UOP public services staff seldom searched by call number, international standard number, bibliographic record number, or any other Inopac field. At Adelphi, these searches represented a sizable portion of public and technical services staff searches (approximately 31.9%); at UOP these were a miniscule part (less than 5%) of all staff search patterns. It is difficult for us to draw any conclusions about the discrepancy without more knowledge of the staffing arrangements and assignments at Adelphi. In the case of UOP, the low use of other indexes by public staff did not surprise us because the nature of their work—primarily circulation and reference activities—would not generally require them to employ these indexes.

IMPACT OF ONLINE INDEXES
A closer look at the 1996 data revealed an unexpected result. In the process of comparing searches in the two years, we found that total searches had dropped nearly 50%: from 84,000 in 1995 to 42,700 in 1996 (see tables 1 and 3). Searches at public terminals had declined by almost 60% while staff searches fell 20%. The number of searches by Technical Services staff remained nearly the same in 1996 but public services searching declined. However, because public service staff had been heavily involved in the 1995 collection development project, the subsequent decline in catalog searching made sense.

But it was unclear what could have accounted for the drastic decline in the number of patron searches. If the 1995 search total had been inflated by staff using public terminals for the collection development project, the 1996 total might have been expected to be down by a couple of thousand searches—but not 40,000. We rechecked our methodology, examined the transaction log setups, checked and double-checked the numbers, but found no detectable errors. There was no decrease in library users. In fact, the gate count and the reference question tallies indicated an increase in foot traffic during this time. Public terminals had been in constant use and we often observed students waiting to use a terminal.

The only change between the periods studied in 1995 and 1996 that could have had an impact on patron search preferences was the addition of the Information Access Company (IAC) full text Expanded Academic Index ASAP and Busi-
ness Index ASAP to our library's catalog gateway in December 1995. The IAC indexes were first made available through the online catalog in December 1994, but student use was minimal. Student awareness and enthusiasm increased considerably when the full text version with the article print station became available in late 1995. We knew from observation and public feedback that the online indexes had become tremendously popular with our students. We examined the IAC connect logs and the Innopac gateway logs to see whether they confirmed our suspicion that the indexes were the cause of such a drastic decline in library catalog searching. Unfortunately, the number of connections made to outside databases had not been tracked until 1996, and even this data was inconclusive. Both IAC and Innopac log the number of connections made to IAC databases from Innopac. However, the number of connections does not accurately reflect the number of students using the database during one session. Students often walk up to a terminal already connected to IAC. The number of searches performed on the databases are not tracked, only the initial connection and the number of minutes the connection is maintained. There was no objective way to determine whether students had performed 40,000 searches in the IAC indexes rather than in the library's catalog during the three month study. However, we had observed students who formerly had to be led to periodical indexes now beginning their research with the IAC indexes. We could only deduce that students were searching the online databases at least as often as they searched Innopac.

Growing student use of the Expanded Academic Index and Business Index might have also contributed to the subsequent increase in subject searching on Innopac. Because the default search for both indexes is a subject search, students were becoming familiar with LCSH, which is the thesaurus of both indexes. Students in library instruction classes were increasingly taught how to understand and use LCSH headings and subdivisions in order to use the IAC databases. There are no hard data to support the idea that our students' increasing familiarity with the subject searching of online indexes increased their subject searching of the catalog but there is anecdotal evidence.

### Replication Difficulties

In our two-year endeavor to provide comparative data on the search styles of patrons and staff, we discovered that replicating a study is not nearly as straightforward as we had initially thought. Despite the similarities in size of the institution, library holdings, systems, and research methodologies, it was unexpectedly difficult to compare data from the two libraries with a high degree of confidence. This was due to a variety of
factors. Differences in menu design and labels, for example, appear to have an impact on patron search patterns. We were not aware of nor did we think to consider that the differences in the search screens of the two institutions might affect the results so dramatically. Another factor was bibliographic instruction. At UOP, library instruction appeared to have some influence on student search strategies, yet we knew nothing of Adelphi's instructional programs. Most of our students participated in hands-on instruction sessions where librarians taught both subject and keyword search strategies. Even the way we staffed the reference desk (one librarian per shift; no staff or para-professionals) may have had an impact on the search strategies of our users.

Comparing data from studies of two libraries is understandably problematic. But we also found it surprisingly difficult to compare year to year data at the same institution. This was primarily due to a continually changing technological environment. At UOP, the number of terminals and PCs available for student and staff searching have changed constantly. For example, in 1996 we added four networked PCs for public access to Innopac (three after the period of the study) and removed a dumb catalog terminal. In the coming year, three networked workstations will feature Innopac's new graphical interface. This will make the task of tracking and interpreting patron search patterns even more challenging.

We also found that it was not easy to isolate the terminal groups or gather valid data or ensure staff compliance with our instructions. And while it might appear straightforward and gratifyingly scientific to analyze reams of computer-generated data, there are multiple variables that have an influence on the numbers collected by the system. These variables must be controlled to ensure accurate results. If librarians were scientists working in laboratories, we could closely control or isolate the variables and regulate our environments. But we work in living laboratories, vibrant and dynamic and elastic. As such, replications of transaction log studies will never be as scientifically rigorous as we might wish.

Despite the problems, a transaction log analysis remains one of the best and most accurate ways to examine basic user search patterns. While the number of variables that might affect results cannot all be controlled, the results are still valuable and illuminating. As we discovered in 1995, a pilot study will help expose the most serious problems. The replication of previous TLA studies is important, not only to build up a body of reliable research but also to test the validity of the original research. Through our study, new information about the Innopac's inability to collect full staff searching statistics came to light, which is important for future studies.

By replicating Ballard's study of Adelphi users we uncovered a wealth of information about our own library users, how they search our system, and how their strategies are affected by a variety of elements. We learned how their search patterns compared with those of users at a similar institution and sought explanations for similarities as well as differences. By comparing our data with that of another library, we were able to question our results more rigorously and use our knowledge more effectively.

**Conclusions**

The most unexpected and intriguing aspect of our study was the discovery of the rise in student use of online periodical databases and the subsequent impact on catalog searches. We were surprised by the dramatic increase and had we had any hint of the apparent impact of the online indexes on student search preferences, we would have closely tracked the number and length of connections made to the indexes from the beginning of the study. Admittedly, the number of connections made to the index database is not comparable to the number of Innopac searches. However, it would have provided a baseline from which to measure subsequent use. A substantial increase in the number of connections made in 1996 compared to the number made in 1995 might have supported our supposition that the online
indexes were the primary cause of the decrease in total searches in our catalog.

Our data also suggest that the shift from a traditional online catalog menu to a gateway menu might have been a contributing factor in the increased use of online indexes. On the previous menu, the option to connect to online databases was buried deep in a long list of search options and did not specify periodical indexes (see figure 1). On the new gateway menu, the IAC indexes were given high visibility and a more precise label (see figure 3). As stated previously, Ballard noted an increase in query searching after it was placed first on the list of search options. In addition, we found that the labeling of individual search options had an impact on search styles. Subject searching rose dramatically when the search option was changed from its more precise and descriptive label, “Library of Congress Subject Heading,” to the less precise but more comprehensible “Subject” label. While the former label was more accurate, it obscured the meaning for the typical user.

As with most research projects, our study prompted more questions than answers. We discovered how little we knew about our remote users, how impossible it was to study staff search patterns with incomplete transactions logs, how technology had unintended impacts on systems and the users of those systems. Despite the obstacles and setbacks, our attempt to replicate a transaction log study was of great benefit to our library, and we encourage other libraries to improve on our effort.

**Works Cited**


In this article, database design for preservation project management is addressed. The system described manages a serials preservation project in which issues from multiple repositories are gathered and collated for preservation microfilming. The system accommodates the ongoing standardization of bibliographic data in the MARC format to facilitate the comparison of holdings among collections. It imports holdings records in the OCLC format and provides for the addition of condition reports and information about existing microfilm service copies, filmers, and the locations of master negative microfilm. Boolean search strategies, employing data from both bibliographic and holdings records, facilitate the identification of newsprint available for microfilming. Management information, added to bibliographic and holdings records, supports the production of reports tracking preservation activity.

The computer is a relative newcomer to preservation management, although the pace of implementation is accelerating. Applications were adopted first for binding preparation activities either as stand-alone proprietary software or as components of integrated library systems in the 1980s. Since then, an increasing number of functions have been automated, including library and archival condition surveys—specifically CALIPR (Ogden and Jones 1991) and PRENAPP (RLG Needs Assessment Task Force 1993); data logging and analysis to measure and report on environmental storage conditions; selection and implementation of physical conservation treatments; and the tracking of reformatting queues for microfilming.

Some early applications were developed as stand-alone modules employing general-purpose database management software, such as dBase, FoxBase, and Revelation. Recognition of the advantages of linking preservation information to bibliographic records and associating preservation with other library functions came quickly and is increasingly in evidence (Bruer 1995). One means of making preservation information available within as well as beyond the local institution is the creation of fields in Machine-Readable Cataloging (MARC) bibliographic and item records. Many libraries have imple-
mented posting mechanisms to record preservation decisions in nationally accessible bibliographic records, storing these records in the Research Libraries Information Network (RLIN) or the OCLC Online Computer Library Center, Inc. databases.

The system application described in this article was designed to manage the California Newspaper Project (CNP), a component of the U.S. Newspaper Program (USNP) funded by the National Endowment for the Humanities (NEH). The database design was intended to facilitate the collection and incorporation of preservation information during the inventory and cataloging phase of the project, which was the first phase. Preservation data are entered into cataloging and holdings records in the system and become the content of system-produced management reports that direct, track, and inform subsequent preservation phases.

THE U.S. NEWSPAPER PROGRAM

The USNP represents an ongoing collaboration of the NEH and the Library of Congress (LC) to identify and preserve newspapers of the United States and its territories. Within the scholarly community, the American Council of Learned Societies reported that access to newspapers, an important yet impermanent resource, was one of its highest research needs (Eugene C. Barker Texas History Center 1990). Implementation of the USNP began in 1982 with grants to six repositories with holdings encompassing most of the fifty states. By 1996, all states were represented, and 29 projects representing U.S. states and territories had been completed. "Completed and current projects will produce bibliographic records for 245,000 newspaper titles and microfilm for 55 million pages of newsprint. To date, the NEH has spent $36 million in support of USNP projects, with non-federal contributions totaling $14 million" (National Endowment for the Humanities 1996). Cataloging is entered into the CONSER (Cooperative Online Serials) database in OCLC, and records are distributed on tape by LC. Holdings entered into OCLC are profiled in the USNP Union List.

USNP projects typically consist of three distinct phases: planning, inventory and cataloging, and preservation. In practice, the inventory and cataloging and preservation phases of a project can overlap. In a small number of projects, microfilming precedes cataloging. Logic would dictate that full inventory and bibliographic control of these items would be achieved prior to filming in order to identify the most complete set of holdings; however, the large number of newspapers involved, combined with the extremely vulnerable and fragile condition of many, means that many newspapers would be lost to accelerating deterioration if they were not filmed quickly. Consequently, a number of USNP projects have adopted a modified regional approach, in which details of the state's major collections are noted and then filming on a county-by-county basis proceeds as fieldwork is conducted. This was the approach envisioned for the CNP.

CNP GOALS

The bibliographic phase of the CNP began at the Center for Bibliographical Studies and Research, University of California (UC) Riverside. Although the initial planning was undertaken by the California State Library (CSL) almost a decade earlier (in 1983), the project began in earnest in 1992. The project's goals were determined at this time by the project management team and were modeled on those of the other state newspaper projects. Although the exact statements were modified during the life of the project, the goals as expressed below reflect what the project intended to accomplish. The project set out to:

- Survey and inventory U.S. newspaper collections in the state
- Construct cataloging records in the CONSER database in OCLC in accordance with national standards and guidelines
- Construct holdings records in the OCLC Union List Subsystem, USNP Union List
• Provide information about California newspapers to scholars, historians, genealogists, and the public by distributing access to the project's database and producing finding aids
• Produce microfilm of endangered newspapers, meeting national preservation standards for filming and storage
• Place service copies of microfilm in libraries for public use
• Ensure continuing access to California's newspapers into the future

CNP PLANNING AND CHALLENGES

Despite the existence of impressive files of bibliographic information at CSL and other research libraries, as well as the existence of numerous published scholarly bibliographies and histories, no single source was found to serve as a comprehensive bibliography of California newspapers. A similar situation held for the state's serial holdings in that no single database could serve as the predominant union list from which a major set of newspaper holdings could be extracted. The three largest union lists are CALLS, the California Academic Libraries List of Serials, which is a University of California MELVYL system database that lists holdings for the University of California and several other large academic library collections; CULP, the California Union List of Periodicals, that lists holdings for the State Library and public, special, and academic library periodical collections of all types; and the OCLC Union List Subsystem, representing academic and public libraries, including the California State University. (As of this writing, the California State Serials Union List Project is building a merged union list in the MELVYL system.) Newspaper holdings are underrepresented in each of these databases.

In addition to the absence of a single comprehensive file of bibliographic and holdings information, California's newspapers are scattered throughout the state in the collections of libraries, historical societies, newspaper publishers, commercial microfilmmers, and private collectors. Although estimates of the total number of titles are reputedly as high as 15,000, no more than 6,000 titles are represented in the two largest repositories—CSL and UC Berkeley. Also, these collections often house limited runs of titles, while more complete runs are held in local or county library, historical society, and publishers' collections.

Further complicating the task of collecting and organizing information was the fact that, even where cataloged, very few of these newspapers had been subjected to adequate bibliographic control or, in many cases, even inventoried or listed. In the past, national cataloging standards were not applied to resources of local or ephemeral interest (as newspapers were too often categorized), and even the largest university collections did not furnish full bibliographic records, track title changes, or record the newspapers in the library's catalog. Frequent title changes, mergers, and varying frequency of editions, if listed, were not presented in a uniform manner. Issues listed under a single title by one repository might be listed under five titles in another, due to title changes (see figure 1). The same newspaper might be represented by different forms of its title in different union lists, due to the generic nature of newspaper titles and their appearance on mastheads where the placement and size of words can make it difficult to define the title.

The challenge facing the CNP was to bring together information about California's newspapers to enable strategic planning and continual updating as the fieldwork was conducted. The body of information that was organized was complex in both its bibliographic and physical manifestations. On the project, information about newspapers from a large number of sources was collected, organized and standardized so that bibliographic and holdings identification among repositories could be compared. This was an ambitious cataloging project, requiring years for completion and a flexible database management system to support the overlay of full MARC bibliographic records on existing records as cataloging proceeded. Notes and holdings had to be moved from one record to another on an
**Figure 1. Title Changes for the San Bernardino Sun**

ongoing basis as title changes were identified. And to support the preservation microfilming, information on the condition, location, producers, quality, and technical specifications of existing microfilms as well as the condition of physical newsprint on site had to be collected and associated with the bibliographic and holdings data.

Several categories of information sources were identified. The previously mentioned serial lists, together with other regional union lists of newspapers and local library systems, such as the catalog at the University of California, Los Angeles (UCLA), furnished a significant portion of the initial data for the project. The initial survey of approximately 700 collections done by CSL in 1983 and the CNP survey of 2,700 California repositories and
publishers undertaken in 1993 also provided a substantial amount of data. The CNP survey included a questionnaire that solicited information about newspapers collected, microfilmed, and indexed, as well as policies for cataloging, interlibrary loan, storage, and collection development. Additionally, a work form for reporting individual titles and issues held, their format, and condition was attached to the survey.

Other research files came from individuals. Physicist Richard L. Lingenfelter supplied the project with his extensive record of California newspaper publishing. This list (compiled for Lingenfelter’s personal research) includes publishing details for over 7,000 California newspapers. Additional inventories of newspapers on microfilm were supplied by Bay (Library) Microfilms and Custom Microfilming Corporation. Last, secondary historical sources and bibliographies were used to identify additional newspapers.

A DATABASE MANAGEMENT SYSTEM

The choice of a system best suited to management and manipulation of bibliographic and preservation data was considered essential to the project's success, and the desired functionality was comprehensive. Throughout the project, the system supported the ongoing work of field teams based at three locations: UC Riverside, UC Berkeley, and CSL. It was necessary to generate lists of repositories and collections sorted by zip code, city, county, or region to assist the field teams in planning for canvasses. It was necessary to produce cataloging work forms selected by library or region, which carried all information previously gathered by the project staff, so that as additional issues of a title were encountered in the field, cataloging could be upgraded to incorporate new information. When the preservation phase of the project got underway, it was necessary to produce management reports that could identify candidate titles for preservation microfilming and track work throughout the preservation process.

It was important that the project adopt a system that could be used immediately in support of data entry and cataloging but which could be further developed to add functionality. This was needed because catalog records were being created at the same time that survey responses were being keyed and other data were being received from other sources. The system had to facilitate the import of MARC and non-MARC records and allow record overlay in either direction. Records were often supplemented with information keyed from secondary sources, such as extensive notes on editors and publishers, unverified information about publication patterns, other frequency editions, and related titles. Such information would later be confirmed by catalogers in the field and added in defined MARC fields to cataloging records. Holdings records were supplemented with information on format and condition that would later assist in the selection of candidate titles for preservation microfilming.

Once the fieldwork was completed, future needs required that the database be able to produce reference lists of titles and holdings selected by library, city, county, region, subject, language, or ethnic audience. Finally, the system had to serve as a production platform both for a comprehensive bibliography of California newspapers and a research database covering California's newspapers at the conclusion of the project.

SELECTION CRITERIA

Although a formal request for proposal was not required, criteria for the system selection were established by project management for use in evaluating software packages and reflected the functional needs that project management foresaw. The system needed to

- import data in both MARC and ASCII text formats,
- export USMARC bibliographic records
and ASCII text files,

• provide relational capabilities for flexible linkage of records in different databases,

• accommodate large text fields,

• permit ongoing changes to record structures and indexes when new needs were defined,

• enable project staff to define search keys and design specialized reports to serve the needs of field canvassers, librarians, and researchers,

• be accessible from the same PC-compatible computers used by project staff to access the OCLC system and run Windows applications, and

• be accessible by project staff located at remote sites as well as by catalogers in the field.

Following a comparative evaluation of software, the Cuadra STAR system was selected. STAR is an information management and retrieval system used by many librarians, museum curators, and information professionals, and database producers and publishers (Cibbarelli 1992). While not strictly a relational database system, STAR has strong relational functionality. Unlike many relational databases, it supports variable field lengths and repeatable fields. It supports user design of record structures and local definition of indexes, reports, and global operations for importing and overlaying records and changing record content. In addition to the basic package, the CNP also purchased MARCPLUS, a MARC database record template, with facilities for import and export of USMARC records. Initially, the system ran on a 486sx PC under the SCO UNIX operating system. Later, when the application had grown, both in size and value, it was converted to a more powerful Sun SPARC platform running under the Solaris UNIX operating system.

CNP DATABASES

The CNP system application was defined to consist of four primary databases: MARCNEWS, NEWS, HOLD, and CONTACT.

MARCNEWS was the MARC database composed of authoritative serial CONSER records imported in batch mode records from the OCLC-held CONSER database (see figure 2). Originally, CONSER records are created and entered by CONSER-participating library catalogers, which includes the CNP project catalogers.

The second database, NEWS, contained a record for each California newspaper title identified by the project, as well as out-of-state titles cataloged at libraries by the field teams. Titles in this file were keyed or imported from a number of sources, including the project's survey of libraries and publishers, secondary historical sources, union lists, and local library systems (see figure 3). When these newspapers were cataloged in OCLC, the newly created MARC records were imported into STAR and overlaid on existing records in the NEWS file. In this process, certain defined non-MARC note fields were retained (see figure 4). The NEWS file was larger than MARCNEWS, as it included titles that had not been fully cataloged by the project and, in some cases, those for which issues had not been identified.

The third database, HOLD, was composed of holdings records, supplemented with information about format and condition. Holdings data were keyed into STAR from survey responses, union lists, and inventories (see figure 5). These records were later replaced by OCLC local data records that were prepared by project catalogers after the issues were inventoried and cataloged on-site (see figure 6). While local data records in OCLC represent library collections only, records in the HOLD database included the collections of microfilmmers, publishers, and private collectors as well. This made it possible to compose a more complete list of issues filmed, as well as those available for filming. Holdings records were linked to bibliographic records by NEWS record numbers and, once titles were cataloged, by OCLC record numbers. The system's linked design (and a STAR system function named Supersearch) meant that holdings could be reformatted and displayed with associated bibliographic records retrieved in a search (see figure 7).
It also meant that NEWS record numbers could be edited in the holdings records to link holdings to a different bibliographic record when a title change was identified. A defined global operation then matched the edited NEWS record number in HOLD with the related bibliographic record and replaced the OCLC control number in HOLD with a revised OCLC control number (see figure 6).

Last, the CONTACT database included a record for each library collection, museum, publisher, microfilmer, and individual contacted in the survey or identified for contact in the course of fieldwork. The database supported the organization of fieldwork, tracked progress, and recorded action items. Records carry names, addresses, and phone numbers, responses to the CNP survey questionnaire, and the dates and results of phone contacts and fieldwork on-site (see figure 8).

This database was employed to produce mailing labels, which could be selected by county, type of institution, or other search criteria. By means of a designed report, labels could be sorted by zip code, institution, or another defined field. A code assigned by the project to each CONTACT record links it to holdings records. A second linking field, the OCLC four-character collection code, was added to the CONTACT record once the CNP profiled a library for participation in the U.S. Newspaper Program Union List in OCLC. By employing this linking field, a global cross-load operation...
Record: 2515

**TITLE**
The Californian

**PUBLISHER**
Colton & Semple

**PLACE**
San Francisco, San Francisco County, California

**PLACE**
Monterey, Monterey County, California

**FREQUENCY**
Weekly

**DATES PUBLISHED**
Vol. 1, no. 1 (Aug. 15, 1846)—v. 3, no. 15 (Nov. 11, 1848).

**GENERAL NOTE**
First newspaper in California. Moved to San Francisco in May 1847; sold by Semple, and became a competitor of the California Star. Suspended from May 26, 1848 to late June 1848 during the Gold Rush; resumed publication in Aug. 1848.

**PUBLISHER NOTE**
Rev. Walter Colton, editor and Dr. Robert B. Semple, publisher, 8/15/1846-4/7/1848; H.I. Sheldon, 8/1848-Extras 8/3/1847, 6/2/1848; Flysheet 5/29/1848.


**Figure 3. NEWS Record (Data Keyed from Secondary Sources)**

Supplied the collection name to OCLC-style local data records imported into the HOLD database once OCLC cataloging and union listing took place (see figure 6). Imported local data records were defined by the project as verified holdings, in contrast to keyed holdings, which were viewed as preliminary information subject to verification. Keyed holdings were dropped from the HOLD file at the conclusion of work with a collection by means of a global delete operation. This was done after condition statements, filmers, storage locations, and notes were cross-loaded to imported local data records.

**SYSTEM SUPPORT OF FIELDWORK**
Prior to conducting fieldwork, which includes inventory and cataloging operations, the CONTACT database was searched to generate a list of libraries, publishers, and individuals identified in the place to be visited. For work in a large city, such as Los Angeles or San Francisco, a search might be defined for a specific zip code or codes, while for a sparsely populated county, records for the entire county might be retrieved subsorted by city. Appointments were made to visit each library, museum, publisher, or individual holding newspapers, and records in the CONTACT file were updated to reflect current information, such as updated addresses, number of titles held, hours of operation, or interlibrary loan policy, acquired during the phone conversation. A note was added to record the date of the phone contact and any subsequent field visit. In the course of phone interviews and field visits, information about additional collections and contacts in the county were often uncovered, and this information was added to the CONTACT database for later pursuit. Once the collections to be canvassed were identified,
Record: 2515

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCLC NUMBER</td>
<td>8260154</td>
</tr>
<tr>
<td>SN NUMBER</td>
<td>84-25963</td>
</tr>
<tr>
<td>UNIFORM TITLE</td>
<td>Californian (Monterey, Calif. : 1846)</td>
</tr>
<tr>
<td>TITLE</td>
<td>The Californian</td>
</tr>
<tr>
<td>PUBLISHER</td>
<td>Colton &amp; Semple</td>
</tr>
<tr>
<td>PLACE</td>
<td>San Francisco, San Francisco County, California</td>
</tr>
<tr>
<td>PLACE</td>
<td>Monterey, Monterey County, California</td>
</tr>
<tr>
<td>FREQUENCY</td>
<td>Weekly</td>
</tr>
<tr>
<td>LANGUAGE</td>
<td>Spanish</td>
</tr>
<tr>
<td>BEGINNING YEAR</td>
<td>1846</td>
</tr>
<tr>
<td>ENDING YEAR</td>
<td>1848</td>
</tr>
<tr>
<td>GENERAL NOTE</td>
<td>First newspaper in California. Moved to San Francisco in May 1847;</td>
</tr>
<tr>
<td></td>
<td>sold by Semple, and became a competitor of the California Star.</td>
</tr>
<tr>
<td></td>
<td>Suspended from May 26, 1848 to late June 1848 during the Gold Rush;</td>
</tr>
<tr>
<td></td>
<td>resumed publication in Aug. 1848.</td>
</tr>
<tr>
<td>PUBLISHER NOTE</td>
<td>Rev. Walter Colton, editor and Dr. Robert B. Semple, publisher,</td>
</tr>
<tr>
<td></td>
<td>8/15/1846-4/17/1848; H.I. Sheldon, 8/1848- Extras 8/3/1847, 6/2/1848;</td>
</tr>
<tr>
<td></td>
<td>Flysheet 5/29/1848. Moved from Monterey. United with California Star</td>
</tr>
<tr>
<td></td>
<td>to form California Star and Californian. Ed: Robert B. Semple,</td>
</tr>
<tr>
<td></td>
<td>5/22/1847-7/12/1847; Benjamin R. Buckelew, 7/17/1847-9/1/1847;</td>
</tr>
<tr>
<td></td>
<td>Robert Gordon, 9/8/1847-12/10/1847; Benjamin R. Buckelew, 1/1848-</td>
</tr>
<tr>
<td></td>
<td>3/1848; J.D. Hoppe, 3/1848-5/10/1848; Henry L. Sheldon, 5/17/1848-</td>
</tr>
<tr>
<td></td>
<td>6/2/1848, 7/15/1848, 8/14/1848-11/4/1848; Edward C. Kemble, 11/11/1</td>
</tr>
<tr>
<td></td>
<td>1848. Pub: Robert Semple; 5/22/1847-7/12/1847; Robert Gordon, 9/8/1</td>
</tr>
<tr>
<td></td>
<td>1847-12/1847 (B.R.B. and J.D.H. Mar-May?) Benjamin R. Buckelew, 7/1</td>
</tr>
<tr>
<td></td>
<td>17/1847-9/1/1847; B.R.B. 17/1848-5/10/1848; J.D. Hoppe and Co. (Jacob</td>
</tr>
<tr>
<td></td>
<td>D. Hoppe, Henry L. Sheldon and Joseph Dock (I) 5/17-29/1848, 7/14/18</td>
</tr>
<tr>
<td></td>
<td>48-10/17/1848; Henry L. Sheldon, Benjamin F. Foster and W.E. Weaver,</td>
</tr>
<tr>
<td>OCLC NOTE</td>
<td>Published in San Francisco, Calif., May 22, 1847-Nov. 11, 1848.</td>
</tr>
<tr>
<td>OCLC NOTE</td>
<td>Suspended with May 29, 1848 notice; resumed with July 15, 1848 issue.</td>
</tr>
<tr>
<td>GEOGRAPHIC</td>
<td>Monterey (Calif.) – Newspapers.</td>
</tr>
<tr>
<td>GEOGRAPHIC</td>
<td>San Francisco (Calif.) – Newspapers.</td>
</tr>
</tbody>
</table>

Figure 4. NEWS Database (Overlayed with Cataloging Data)

and, in some cases, after an inventory had been conducted by the project, work forms were produced by searching on collection codes carried in holdings records in the NEWS database (see figure 9). When a title had previously been cataloged and a full MARC record was available, a MARCNEWS record was also printed. This work was conducted by student assistants at the CNP headquarters office. When titles were thought to have changed or be related, but the evidence was insufficient to establish linking entries in the cataloging record, that information was put in note fields of the NEWS records to inform field teams of research to be conducted in the field. Although the project's original plan was to have field teams access the system online via dial access or telnet to the system server in Riverside, the uniform difficulty of securing phone lines or network connections at the site of newspaper collections where the work took place (often in storage or remote areas of a library or publisher’s office) made this means of access impractical. Instead, work forms were printed and carried into the field. If a reasonably accurate or comprehensive inventory of a collection had previously been done by
the library or the CNP field team, the forms were sorted in shelllist order using that inventory. Otherwise, forms were sorted by city of publication, then title. A separate printout of the county’s newspapers was produced, so that the team could update the list with any new information that is gathered in the course of fieldwork (see figure 10).

At the conclusion of work at a library (after records have been loaded into the STAR system), the project delivered a printed reference list of newspaper titles and holdings (see figure 11).

**REFERENCE USES OF THE NEWS DATABASE**

As the project has progressed, and the NEWS database has become the most comprehensive source of information about California newspapers, an increasing number of reference queries have been received from librarians, researchers, and genealogists from across the country. If the CNP is unable to identify newspaper issues sought by a researcher, the query is posted to a record in the NEWS database so that a response can later be provided when additional issues are located in the field. In responding to reference queries, the project has fostered contacts with persons knowledgeable about California newspapers and their locations; this information can be used to identify additional issues for microfilming by the project. The NEWS and HOLD databases have proven so valuable that the CNP plans to make them more widely available on CD-ROM or via the Internet.

**PRESERVATION MICROFILMING**

One of the most important functions of the system in the support of the preservation process has yet to be realized. The selection process, as well as the management, organization, and tracking of microfilming, will be greatly facilitated by the selection of relevant bibliographic records, associated holdings, condition reports, and management information in the form of specialized reports. The CNP databases do serve in planning for the project’s preservation phase. Throughout
Figure 7. NEWS Record Display with Associated Holdings from the HOLD Database

the bibliographic phase, holdings records representing microfilm were supplemented on an ongoing basis with information about the condition of film, filming agencies, and the location of master negative microfilm. Searches against the database were used to generate lists of each city's microfilmed newspapers (see figure 12), beginning with the 100 largest California cities and county seats. Such reports have assisted the project to determine the geographic distribution of microfilming to date.

Other more comprehensive reports will support the project's selection of titles for microfilming. Holdings records for newspaper carry information about the condition of the original, specific issues held, whether the issues are bound, and their availability for filming. As selection for preservation microfilming is undertaken, titles that are candidates for filming will have a field marked in the bibliographic record. Searching on this field will permit project staff to select a list of newspapers for filming from each county, to evaluate attached newsprint holdings, and to make the best selection for microfilming (see figure 13).

Another field added to the bibliographic record has been defined for the notation of issues wanted for microfilming (see figure 14). A report has been defined as a finders' list, selected by searching "not null" values in this field (see figure 15). The selection of titles can be limited to a specific city or county for distribution to the local media of a specific community or communities. Other fields in holdings records have been defined to carry information about the specific issues selected for filming, the date on which issues are retrieved from a collection, whether the owner needs newsprint returned after filming, and the date on which issues are returned (see figure 16). This information will assist project staff in tracking the gathering, collating, and delivering of issues.
Record: 3233

STAR CODE: paspl
INSTITUTION: Henry Hoover Public Lib.
CONTACT: John Smith, Library Director
CONTACT: Betty Jones, Serials Asst.
ADDRESS: 30 Sunny Way
CITY: Hooverville
COUNTY: Riverside
ZIPCODE: 90001
PHONE: (619) 232-8294
NEWSPAPERS?: YES
SIZE: 00026
MAIL?: YES
CATALOGED?: NO
OCLC?: YES
UNION LIST: Inland Library System Union List of Serials
ILL?: YES
FILMING?: YES
FILMING CO.: Data Microfilming Corp.
MASTER NEG: Data Microfilming Corp.
INDEXES: The Desert Sun
INDEXES: Riverside Press-Enterprise (Desert Section only)
CLIPS: The Desert Sun
CLIPS: Riverside Press-Enterprise
NOTE: Newspapers are indexed both manually and on computer.
NOTE: The Desert Sun is a Gannett paper and may be indexed online.

Figure 8. Record from the CONTACT Database

Record: 24

OCLC NUMBER
SN NUMBER 24
UNIFORM TITLE
TITLE: American and Chinese Commercial Newspaper
VARYING TITLES
AUTHOR
PUBLISHER: Kee
PLACE: San Francisco, San Francisco County, California
FREQUENCY: Weekly
PAST FREQUENCY
LANGUAGE: Chinese
DATES PUBLISHED 2/15/1883-1901
RELATED TITLES: As Weekly News Ap 94-Ap 95
OTHER NOTES
SUBJECTS
GEOGRAPHIC
PREVIOUS TITLE
LATER TITLE
OTHER EDITION

Figure 9. Cataloging Data Sheet
<table>
<thead>
<tr>
<th><strong>OCLC NUMBER</strong></th>
<th>30337734</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SN NUMBER</strong></td>
<td>94-51085</td>
</tr>
<tr>
<td><strong>TITLE</strong></td>
<td>Colton daily courier</td>
</tr>
<tr>
<td><strong>VARYING TITLE</strong></td>
<td>Colton courier</td>
</tr>
<tr>
<td><strong>PUBLISHER</strong></td>
<td>F.S. Hosfelt</td>
</tr>
<tr>
<td><strong>PLACE</strong></td>
<td>Colton, San Bernardino County, California</td>
</tr>
<tr>
<td><strong>FREQUENCY</strong></td>
<td>Daily (except Sat. and Sun.)</td>
</tr>
<tr>
<td><strong>PAST FREQ.</strong></td>
<td>Daily (except Sun.)</td>
</tr>
<tr>
<td><strong>BEGINNING YEAR</strong></td>
<td>1uuu</td>
</tr>
<tr>
<td><strong>ENDING YEAR</strong></td>
<td>19uuu</td>
</tr>
<tr>
<td><strong>DATES PUBLISHED</strong></td>
<td>Published 3/5/1912-1942</td>
</tr>
<tr>
<td><strong>OCLC NOTE</strong></td>
<td>&quot;Daily&quot; appears within masthead ornament.</td>
</tr>
<tr>
<td><strong>RELATED TITLES</strong></td>
<td>Continued by Colton courier Continues Colton chronicle</td>
</tr>
<tr>
<td><strong>GEOGRAPHIC</strong></td>
<td>Colton (Calif.) - Newspapers.</td>
</tr>
<tr>
<td><strong>HOLDING</strong></td>
<td>Colton Area Museum (J. Hofer reports holdings here)</td>
</tr>
<tr>
<td><strong>HOLDING</strong></td>
<td>Colton Public Library Dates: 1912-1942.</td>
</tr>
</tbody>
</table>

**Record: 73**

<table>
<thead>
<tr>
<th><strong>SN NUMBER</strong></th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td>Colton page</td>
</tr>
<tr>
<td><strong>PLACE</strong></td>
<td>Colton, San Bernardino County, California</td>
</tr>
</tbody>
</table>

**Record: 83**

<table>
<thead>
<tr>
<th><strong>SN NUMBER</strong></th>
<th>83</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td>Crestline courier-news</td>
</tr>
<tr>
<td><strong>VARYING TITLE</strong></td>
<td>Crestline courier news</td>
</tr>
<tr>
<td><strong>PUBLISHER</strong></td>
<td>Desert Community Newspapers</td>
</tr>
<tr>
<td><strong>PLACE</strong></td>
<td>Crestline, San Bernardino County, California</td>
</tr>
<tr>
<td><strong>FREQUENCY</strong></td>
<td>Weekly</td>
</tr>
<tr>
<td><strong>BEGINNING YEAR</strong></td>
<td>1991</td>
</tr>
<tr>
<td><strong>ENDING YEAR</strong></td>
<td>9999</td>
</tr>
<tr>
<td><strong>DATES PUBLISHED</strong></td>
<td>Began in 1991.</td>
</tr>
<tr>
<td><strong>GEOGRAPHIC</strong></td>
<td>Crestline Courier-News Dates: 1974-1986. (Info provided by publisher)</td>
</tr>
<tr>
<td><strong>HOLDING</strong></td>
<td>San Bernardino County Lib., Lake Arrowhead Dates: 6 mos. +. Condition: good.</td>
</tr>
</tbody>
</table>

*Figure 10. First Page of Printout of San Bernardino County Newspapers*
Natural History Museum of Los Angeles Newspapers

El Aguacero. Los Angeles, California. Weekly. Began in 1878.—
Cf. Gregory, W. Amer. newspapers. OCLC #10213884
Amigo del pueblo (Los Angeles, Calif.) Los Angeles, California. Weekly. Began in 1861.—Cf. Gregory, W. Amer. newspapers. OCLC #10213494
American sentinel (Oakland, Calif.: 1886) Oakland, California. Weekly. Began in Jan. 1886; ceased in 1890.—Cf. Gregory, W. Amer. newspapers. OCLC #13881968
Alhambra (Alhambra, Calif.) Alhambra, California. Weekly. OCLC #30210793
Acton rooster. Acton, California. Monthly. Published 1891-1915 OCLC #30306105
American and Chinese commercial news = Hua Mei hsin pao. San Francisco, California. Weekly. OCLC #34582275

Figure 11. First Page of Sample Reference List
Record: 201

**TITLE** Carta editoriale.
**PLACE** Riverside, Riverside County, California

Record: 198

**UNIFORM TITLE** Citrograph (Redlands, Calif.)
**TITLE** The citrograph.
**PUBLISHER** Redlands, San Bernardino County, California.
**FREQUENCY** Weekly.
**BEGINNING YEAR** 1887
**ENDING YEAR** 1908
**DATES PUBLISHED** Vol. 1, no. 1 (July 16, 1887) - Ceased with Nov. 14, 1908 issue. - Cf. Gregory, W. Amer. newspapers.

**OCLC NUMBER** 24451032

Record: 8969

**UNIFORM TITLE** Enterprise (Riverside, Calif. : 1977)
**TITLE** The enterprise
**PUBLISHER** Press-Enterprise Co.
**PLACE** Riverside, Riverside County, California
**FREQUENCY** Daily
**BEGINNING YEAR** 1977
**ENDING YEAR** 1981
**OCLC NUMBER** 27823090

---

**Figure 12.** First Page Printout of Riverside County’s Microfilmed Newspapers

Search: COUNTY=RIVERSIDE AND FORMAT=PHOTOGRAPHIC

---

Search: COUNTY=RIVERSIDE AND FORMAT=PHOTOGRAPHIC
<table>
<thead>
<tr>
<th>Record: 7732</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
</tr>
<tr>
<td><strong>VARYING TITLE</strong></td>
</tr>
<tr>
<td><strong>VARYING TITLE</strong></td>
</tr>
<tr>
<td><strong>PUBLISHER</strong></td>
</tr>
<tr>
<td><strong>PLACE</strong></td>
</tr>
<tr>
<td><strong>FREQUENCY</strong></td>
</tr>
<tr>
<td><strong>BEGINNING YEAR</strong></td>
</tr>
<tr>
<td><strong>ENDING YEAR</strong></td>
</tr>
<tr>
<td><strong>DATES PUBLISHED</strong></td>
</tr>
<tr>
<td><strong>OCLC NUMBER</strong></td>
</tr>
<tr>
<td><strong>CANDIDATE</strong></td>
</tr>
<tr>
<td><strong>ISSUES NEEDED</strong></td>
</tr>
<tr>
<td><strong>HOLDING</strong></td>
</tr>
</tbody>
</table>

**Figure 13.** Title Marked as Candidate for Preservation Microfilming in the NEWS Database

<table>
<thead>
<tr>
<th>Record: 8784</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
</tr>
<tr>
<td><strong>PUBLISHER</strong></td>
</tr>
<tr>
<td><strong>PLACE</strong></td>
</tr>
<tr>
<td><strong>FREQUENCY</strong></td>
</tr>
<tr>
<td><strong>BEGINNING YEAR</strong></td>
</tr>
<tr>
<td><strong>ENDING YEAR</strong></td>
</tr>
<tr>
<td><strong>DATES PUBLISHED</strong></td>
</tr>
<tr>
<td><strong>OCLC NUMBER</strong></td>
</tr>
<tr>
<td><strong>CANDIDATE</strong></td>
</tr>
<tr>
<td><strong>ISSUES NEEDED</strong></td>
</tr>
</tbody>
</table>

**Figure 14.** Issues Needed for Filming Noted in NEWS Record
San Bernardino County
Newspapers Issues Needed for Microfilming


The Colton chronicle. Semiweekly. Colton, California. Published from 1888 to 1911. ISSUES NEEDED: 1888; 1898-1904; other scattered issues.

Figure 15. Finders List Produced from the NEWS Database
Search: COUNTY=SAN BERNARDINO AND NEED NOT=*

for filming, seeing that they are filmed in a timely fashion, and disposing of newspaper once filming has taken place. Examples of potential reports are:

- Titles delivered for filming before a certain date, but not yet filmed (DELIVERED<8/96 AND FILMED="")
- A listing of issues retrieved for filming from a specific library (LIBCODE=XXX AND RETRIEVED NOT="")
- Filming completed after a specified date (FILMED>8/15/96)

The last report will serve as a notification and queue for cataloging preservation masters subsequent to filming. Other reports, such as one comparing turnaround time among microfilmmers, can be defined as needed. New fields, indexes, and reports will be required, which highlights the advantage that new data structures, indexes, and reports can be modified or created in the CNP system.

CONCLUSION

The CNP system has the capability of managing and associating bibliographic, holdings, and preservation data created in a preservation microfilming project. The flexible system design, the integration of MARC and non-MARC data, and a high level of user control over indexing decisions and report definition model a new system design for managing a broader spectrum of library functions. The system

<table>
<thead>
<tr>
<th>TITLE</th>
<th>California chronicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUBLISHER</td>
<td>Frank Soule &amp; Co.</td>
</tr>
<tr>
<td>PLACE</td>
<td>San Francisco, San Francisco County, California</td>
</tr>
<tr>
<td>FREQUENCY</td>
<td>Semimonthly</td>
</tr>
<tr>
<td>BEGINNING YEAR</td>
<td>1853</td>
</tr>
<tr>
<td>ENDING YEAR</td>
<td>1858</td>
</tr>
</tbody>
</table>

Figure 16. Issues to be Filmed Displayed in Association with the NEWS Database
is an indispensable tool for directing current project day-to-day operations, and, in the future, the resulting databases will be the most comprehensive resource for information about California's newspapers.

WORKS CITED


Converting Wade-Giles Cataloging to Pinyin: The Development and Implementation of a Conversion Program for the Australian National CJK Service

Linda Groom

In 1996, the National Library of Australia processed over 500,000 Chinese, Japanese, and Korean (CJK) USMARC records through a program that identified data in the Wade-Giles romanization scheme and converted it to the more user-friendly Pinyin romanization scheme. The conversion program was developed by the National Library in response to the needs of member libraries of the Australian National CJK Service. In this article, the author describes the development of the Pinyin conversion program, the functions of the program itself, and how it has been used in Australia. She also outlines the Pinyin word division standards adopted in Australia and briefly discusses the feasibility of an alternative methodology for the automatic creation of Pinyin data.

Australian libraries with collections of Chinese material have been faced, as have many libraries in other countries, with the problem of two competing romanization schemes for Chinese: Wade-Giles and Pinyin (Lu 1995; Tao and Cole 1990; Young 1992). In essence, the problem for Australian libraries has been that, while users of Australian Chinese collections generally have a strong preference for Pinyin romanization, most of the Chinese data in Australian library catalogs are in Wade-Giles. For some library users, especially users in public libraries and undergraduates learning Chinese, Wade-Giles romanization is almost incomprehensible. The cost of manually converting library catalogs from Wade-Giles to Pinyin has, however, been prohibitive.

In June 1996, the Australian National CJK Service was established. The CJK Service is an online service maintained by the National Library of Australia. It provides bibliographic data for copy cataloging and location information for CJK materials held in Australian libraries. Around 90% of the bibliographic records

LINDA GROOM (lgroom@nla.gov.au) is Digital Services Manager, National Library of Australia, Canberra, and was formerly the National CJK Project Manager. Manuscript received January 30, 1997; accepted for publication March 13, 1997.
contain the original ideographic characters, as well as romanization.

During the three-year project to establish the CJK Service, the National Library took the opportunity to develop a conversion program that automatically converts Wade-Giles bibliographic records to Pinyin. This program has been used to process over 500,000 bibliographic records, including all the CJK records available through the Library of Congress Cataloging Distribution Service, and selected records from the OCLC Online Computer Library Center, Inc., and Research Libraries Information Network (RLIN) databases. The conversion program allowed the creation of parallel Pinyin and Wade-Giles databases, each containing CJK records and ideographic characters, but using different Chinese romanizations.

**The Decision to Create a National Pinyin Database for Australia**

Through the late 1980s, Australian libraries debated with some vigor the merits of the two romanization schemes, notably at the annual user meetings of the Australian Bibliographic Network (ABN). In response to this, the National Library of Australia commissioned a study by MacDougall (1991). After considering MacDougall’s study and comments from Australian libraries, the ABN Standards Committee felt there was no option but to allow both romanization schemes on the ABN database.

In 1995, as the implementation date of the National CJK Service database drew nearer, the debate was rekindled. The forums for the debate were the meetings of the CJK Technical Committee, the committee charged with developing cataloging policy for the National CJK Service. Chaired by Chooi Hon Ho of Monash University, the committee comprised representatives from each of the initial eight member libraries of the CJK Service.

At the Technical Committee’s meeting in Canberra on February 14 and 15, 1995, much time was spent on the question of which scheme should be adopted for the romanization of Chinese in original cataloging on the National CJK Service database. As might be expected, the committee members representing the libraries with the largest and oldest Chinese collections and with large catalogs in Wade-Giles, such as the National Library itself and the Australian National University, argued in favor of adopting Wade-Giles as the National CJK Service standard. Committee members representing libraries with more recently developed collections argued in favor of Pinyin. The issue was not resolved at the meeting.

The member libraries were, however, unanimous on one point: that Australian libraries should do everything possible to move to a single standard for the romanization of Chinese. There was a strong view that the dual romanization policy that had been used for ABN cataloging was not working. The inefficiencies of having two standards, and the failed searches that resulted when these two standards occurred in either a national shared database or a library’s own catalog, were clear and urgent problems.

It was with the knowledge that a solution was needed that I attended the First International Seminar on Chinese Document Database, in Guangzhou, February 27, 1995, to March 4, 1995. There I had the good fortune to meet Karl Lo, International Relations and Pacific Studies Librarian at the University of California, San Diego, and author of a program that converted Wade-Giles words to Pinyin (Lo and Miller 1991). The program was designed to work with input strings that were guaranteed to be in Wade-Giles, converting these strings letter by letter. Machine-Readable Cataloging (MARC) records for Chinese material in Western library catalogs contain a great deal more than Wade-Giles—for instance, alphabetic subfield codes and notes and subject headings in English. I had unpleasant visions of what such a program might do to phrases such as “edited by” or “title page.”

I posed this problem to Lo during the conference. Was it feasible for a program to distinguish Wade-Giles words from all other words in a MARC record and to convert those words, and only those
words, to Pinyin? In Lo's view, it was entirely feasible. He responded to every concern raised with an indication of what could be accomplished in machine conversion by this program, and in the end I was convinced. He has since told me that if he had seen some of our actual data, he would have been far less optimistic. Still, I came away from the conference inspired. I am tempted to wonder how many other projects have had origins that include an element of frank and friendly international misunderstanding!

In March, a small team at the National Library began work on creating a program that would read USMARC records, identify any Wade-Giles words in them, convert the Wade-Giles to Pinyin, and reassemble a valid USMARC record. The team consisted of two National Library staff: Ching-Ping (CP) Tang and Murthy Manchella, with Karl Lo as consultant. We experimented with letter-by-letter conversion and word-by-word conversion and chose the latter. By May, we had a working program and were able to produce sample data acceptable to the CJK Technical Committee. Two months later, the CJK Steering Committee approved the establishment of a Pinyin database that would be a mirror image of the Wade-Giles database.

**DESCRIPTION OF THE WADE-GILES TO PINYIN CONVERSION PROGRAM**

The program is written in C. It accepts files of USMARC records, in batches of up to 20,000 records. It first checks that the MARC record contains the 001 (record number) field, since the program needs this field to keep track of each record. It then reads the language code in the 008 field, if present, and divides the file into Chinese, Japanese, Korean, and other. This division by language is necessary because later steps in the program use different logic for records whose main language is not Chinese.

Readers of this article may wonder why we bothered at all with MARC records for Japanese, Korean, and other material. We needed to put all records through the conversion to catch, for example, Japanese translations of Chinese works, where the author name and uniform title were in Wade-Giles. This requirement added considerably to the complexity of the program.

**MARC TAG TABLE**

After dividing the file by language, the program processes each MARC record in turn. For each record, it consults a table that lists all the USMARC fields and subfields that, in the team's view, are likely ever to contain a Wade-Giles word. Some fields, such as 008 and 040, which contain codes or symbols, should never contain any Wade-Giles words and so can be safely excluded from the conversion process. The table of MARC fields and subfields grew with each version of the program, as Wade-Giles data turned up in unexpected places. For example, field 020 subfield a (ISBN), which one might expect to be totally numeric, sometimes contains a Chinese publisher name. It might have been simpler in the end to list only those fields in which Wade-Giles would not occur. It would have been a much shorter list.

**MAIN PROCESSES**

Each word in each subfield is then put through a sequence of tests, which aim to determine whether the subfield contains Wade-Giles or not. If the subfield contains Wade-Giles and only Wade-Giles, it is converted. If it contains no Wade-Giles, it is left untouched. If the program cannot decide whether the subfield contains Wade-Giles, the program converts it and flags the preconversion image and the postconversion image for human review.

To make the sequence of tests, the program refers to the “word table” that lists all valid Wade-Giles words. The following attributes are specified for each Wade-Giles word:

- Its Pinyin equivalent.
- Whether its Pinyin equivalent is the same. For example, the Wade-Giles word “ang” converts to Pinyin “ang”; these words are called “same” words.
- Whether there is a Pinyin word that contains the same letters, but has a different meaning. For example, “chao” is valid in both Wade-Giles and
Pinyin, but Wade-Giles "chao" converts to Pinyin "zhao." These words are referred to as "common" words.

• Whether the Wade-Giles word also happens to be valid in Hepburn (Japanese romanization) or McCune-Reischauer (Korean romanization).

The tests use a process of elimination. The broad structure of the tests is listed below.

1. **Personal name processing.** If a personal name field contains a non-Wade-Giles word, e.g., "Jimmy Chang" where "Jimmy" is non-Wade-Giles, and "Chiang Kai-shek" where "shek" is non-Wade-Giles, the program leaves the entire name unconverted.

2. **Pure Wade-Giles processing.** If a subfield contains only pure Wade-Giles words—i.e., words that are unique to Wade-Giles—it is converted. The program also tests for subfields that contain other combinations involving pure Wade-Giles words. For example, subfields that contain both a pure Wade-Giles word and a pure Pinyin word are flagged for human review. Subfields that contain a pure Wade-Giles word and an English word (the program categorizes English words as "nontable" words) are also flagged for human review, unless they are in fields such as subject (6xx) fields. Subject fields often contain phrases like "T’ang dynasty," where the presence of the English word does not cast doubt on the fact that "T’ang" is Wade-Giles. In these cases, the pure Wade-Giles word is converted, and the phrase becomes "Tang dynasty."

3. **Pure Pinyin processing.** We found it necessary to put in a test aimed at weeding out Pinyin data, because databases that are supposedly Wade-Giles contain a surprising amount of Pinyin. If the subfield consists entirely of pure Pinyin words, then it is not converted. The program also tests for subfields that contain other combinations involving pure Pinyin words, for example, a pure Pinyin word and some "common" words or a pure Pinyin word and an English word. Because of the frequency of English words in statements of responsibility—e.g., "translated by Wen Zhong"—special processing occurs for statements of responsibility.

4. **Processing of non-Chinese words.** If the subfield consists entirely of words that are not listed in the word table (i.e., none of the words are Wade-Giles), then it is not converted. The program also tests for subfields that contain other types of words, such as nontable words plus "same" or "common" words. Fields such as general notes (500) or chronological subfields in subject headings (6xx subfield y), where English words are often mixed with Wade-Giles words, are flagged for human review. Human review seems the only reliable way to ensure that notes data, such as "Running title Shang-hai kuan pao," are correctly converted to "Running title Shanghai guan bao."

5. **"Common" words processing.** If the subfield contains any words that occur in both Wade-Giles and Pinyin but with different meanings, and if the subfield is in a field such as a title or note field, it is flagged for human review. Other subfields containing "common" words are converted.

6. **"Same" words processing.** If the subfield contains any words that occur in both Wade-Giles and Pinyin and have the same meaning, it is converted. Since these words are the same in Wade-Giles and Pinyin, conversion does not appear to take place, except where two Wade-Giles words are hyphenated. Thus "Yang-wen" converts to "Yangwen."

7. Remaining subfields are not converted.

**PROCESSING OF THE WADE-GILES WORDS “TI” AND “LO”**

Most Wade-Giles words have a one-to-one conversion to Pinyin. The Wade-Giles words "ti" and "lo," however, each have more than one Pinyin equivalent. The word "ti" can convert to either "di" or "de," and "lo" can convert to either "le" or "luo." To make the correct conversion, the program searches the linked 880 field to
find the Chinese character that is linked to the “ti” or “lo.” From the code for that character, the program can correctly convert the word. If the record does not contain any Chinese characters, the program checks the field and subfield in which the word occurs. A “ti” occurring in a personal name (100 $a, 600 $a, etc.), for instance, can be converted to “di” with a high degree of reliability. In some subfields, however, the correct conversion cannot be predicted; these are flagged for human review.

**Human Review of Doubtful Cases**

Each batch of records produces some cases that are flagged for human review. These files can be transferred to the reviewer’s computer using file transfer protocol (FTP) and reviewed using a commercial editing package such as Microsoft Write or reviewed on the machine on which the conversion program resides, using GNU Emacs, a UNIX-based editor.

The human reviewers are presented with the field before conversion, the field after conversion, and a brief message saying why the field has been flagged for review. The reviewer does not see the entire bibliographic record. This produced some disadvantages, in that human reviewers had to open another window to the CJK Service database to check a romanized field against the Chinese characters carried in the linked 880 field. A design that involved processing entire MARC records, however, would have produced slow response times both in calling up files and in other processing.

Reviewers choose to keep either the unconverted or the converted version. They can also edit the converted version of the field if they wish.

**Final Processing**

The reviewed fields are then merged back with the other fields in the postconversion bibliographic record. The program checks to see whether there is a title added entry (740) that, after conversion, is now the same as the main title (245). If it is exactly the same, the 740 field is deleted. The program finally recalculates the MARC directory and reassembles a valid US-MARC record.

**Word Division Standards**

In preparing for the development of the Pinyin database, the CJK Technical Committee faced a situation in which there was no internationally recognized standard for Pinyin word division. In 1995, the committee decided to set the word division standards for use by catalogers using the National CJK Service (1996, resolution 5) stating:

That where a cataloger inputs Pinyin data into the National CJK System, each Chinese character should be input as one Pinyin syllable, except for proper and geographic names, where the syllables should be joined. Thus the Wade-Giles Mao Tse-tung is converted as Mao Zedong.

Considerable discussion preceded this decision. Pinyin word division standards have been a controversial topic for many years in Australia. The Technical Committee considered a range of issues, including the effects of single-syllable Pinyin on keyword searching, the extent to which proximity searching could help, the need for an authoritative dictionary if the joined-syllable approach was adopted, and the emerging practice of the National Library of China, in which recent MARC record samples contained single-syllable Pinyin. Another aspect that influenced the Technical Committee was the Committee’s enthusiasm for “automatic romanization,” a prospective enhancement to the Innopac software under which catalogers would enter Chinese characters only, and Innopac would automatically create the parallel romanized fields. Automatic romanization is easier to implement if each character is romanized as a single syllable. Perhaps the crucial factors were that, although the Committee recognized that there were good arguments on both sides, they felt that a decision was better than further years of debate, and that for practical reasons a standard was needed that was simple and easy for catalogers to apply. The CJK Technical Committee also decided (1996, resolution 12) that the
Record source: ABN

Before conversion:
245 $a Ch’un ts’un

After conversion:
245 $a Ch’un ts’un

**Figure 1.** Unexpected Diacritics.

Policy for authorities on the Pinyin database should be as follows:

That . . . in the Pinyin database:
- where the preferred form in LCNA is a pure Wade-Giles form, the preferred form in the Pinyin database should be that Wade-Giles form converted into Pinyin.
- where the preferred form in LCNA is an “established” i.e. not a pure Wade-Giles form, for instance Chang Kai-shek and Confucius, the preferred form on the Pinyin database should be that “established” form.

The word division standard has since been proposed by the Library of Congress for adoption by American libraries (Melzer 1996).

**Conversion Errors**

The National CJK Service Pinyin database contains some incorrect data. Although we have now corrected the causes of many of these errors in the program, we have not been able to correct all of the records on the database that contain errors. Some examples of error types are listed below:

1. **Unexpected diacritics.** The diacritics in “ch’un” and “ts’un” in figure 1 had been incorrectly keyed as alifs; they should have been keyed as ayns. The presence of the alifs meant that the conversion program did not recognize “ch-n” and “ts-n” as Wade-Giles words and left them unconverted. (Note that the conversion program does predict the common cataloger shortcut of typing apostrophes instead of ayns and makes correct conversions. We had not predicted the possibility of a cataloger typing an alif instead of an ayn.)

2. **Unexpected tag and subfield combinations.** The “chih shih” in figure 2 was not converted because we had not listed 710 $d in the table of data elements to be processed. It has now been listed in the table.

3. **Unexpected subfield coding.** The $h subfield in figure 3 was incorrectly left in Wade-Giles. The $h subfield should have been coded as a $b; $h was not listed in the MARC tag table. We have now added 245 $h to the MARC tag table, to catch cases such as this in future.

4. **Records with Chinese titles but with language code “jpn.”** Although the uniform title field (130) in figure 4 was converted correctly, the main

Record source: LC (90216819)

Before conversion:
710 $a Tseng, Tso-chou, $d chih shih 1844

After conversion:
710 $a Zeng, Zuozhou, $d chih shih 1844

**Figure 2.** Unexpected Tag and Subfield Combinations.
title (245) was not converted. The program does not process 245 fields when the language code is "jpn." If the record had been correctly coded as “chi” instead of “jpn,” the conversion would have occurred without error. The program could be amended to process the 245 fields even when the language code is “jpn,” but this would cause a large number of fields to be flagged unnecessarily for human review. Incorrect language coding is one of the most common sources of conversion errors.

5. Japanese records with parallel titles in Chinese. In figure 5, the language coding was correct. However, for records with language code “jpn,” we had not listed the 245 $b subfield for processing. If 245 $b was listed, many subtitles in Japanese that contained words also valid in Wade-Giles would be needlessly flagged for human review. However, the result of not listing 245 $b for jpn records is that Chinese parallel titles in Japanese records are skipped by the conversion. This was a pragmatic decision to save reviewing time. The program could be easily amended to check 245 $b’s in Japanese records.

6. Diacritics omitted. “Tien-chin” in figure 6 was incorrectly converted to “Dianjin,” and “yu” was incorrectly converted to “you,” because the cataloger omitted diacritics.

7. Some notes become incongruous. Some Wade-Giles records contain a note such as “Added title also in Pinyin...” Once the record has been converted to Pinyin, such notes are incongruous. Although it would be possible to write software to detect and delete many of these notes, we decided it was not worth the effort. It is tempting to dismiss many of the above problems as resulting from cataloger error in the original data and to say that the conversion program is not at fault. Catalogers are only human, and errors are part of life, however. Our policy has been to code in expectation of cataloger errors where possible. In creating the Pinyin database, the reviewers...
also picked up a number of cataloger errors and corrected them as part of the reviewing process.

**MANAGEMENT PROBLEMS**

A large problem in management of a conversion using this program is the provision of staff to review the data flagged as doubtful conversions. At present, the Wade-Giles to Pinyin program flags around 12% of all fields for review. For the conversion of the National CJK Service database of over 500,000 bibliographic records, a total of 920 hours of human reviewing, not including supervisory time, over a period of two months was required. It would be easy to change the program's logic so that fewer data were flagged for human review, but this would produce more conversion errors.

Another major management problem for the National Library of Australia has been the workload and complexities of maintaining parallel Wade-Giles and Pinyin databases. With the help of the software supplier Innovative Interfaces, we have solved a number of the problems of cross-updating the two databases, but some problems remain.

**HOW THE CONVERSION PROGRAM HAS BEEN USED BY THE NATIONAL CJK SERVICE**

Libraries that wish to convert to Pinyin can use the Pinyin database as a migration tool. For the many libraries that added holdings to the ABN for their CJK material, the migration path is relatively smooth. Their holdings have been copied to the Pinyin database, allowing those libraries easily to extract a Pinyin version of MARC records for all the CJK material they own. These MARC records can then be loaded to the library's local system, where they replace Wade-Giles records. A number of libraries are planning to use the Pinyin database in this way. Libraries whose holdings are not on the CJK Service database will need to add their holdings to it retrospectively in order to be able to extract the Pinyin records. Alternatively, they could perform a local conversion to Pinyin using the Wade-Giles to Pinyin conversion program directly on their local data. The
National Library will make the conversion program available to any CJK Service member libraries that wish to use it.

The CJK Technical and Steering Committees have agreed that the existence of the Wade-Giles database will be reviewed in mid-1998. At that point, it will be decided whether the workload of supporting two databases is worthwhile. The ultimate aim is to meet the member libraries' requirement for the adoption of a single Chinese romanization scheme in Australia; given the trend among the member libraries, it seems almost certain that the single romanization scheme will be Pinyin. A likely option thus seems to be to cease support for the Wade-Giles database; the timing of that would, however, depend on the extent to which member libraries are able to migrate their local systems to Pinyin.

OTHER METHODOLOGIES

Some readers of this article may wish to question the wisdom of putting such programming effort into romanization. Romanization is, after all, an extremely poor substitute for the original characters. For many Australian libraries with Chinese collections, however, it will be some years until they acquire local systems that can support Chinese characters. These libraries need a source of MARC data that will allow them to present their cataloging in the romanization scheme most familiar to their users. Even in universities where the library may have installed a library system that supports CJK characters, library users often connect to the library catalog from their departments and offices, and few of these users yet have the software on their desktop PCs to display CJK characters. My personal view is that the debate over whether to use Wade-Giles or Pinyin has consumed librarians' energies too long; by moving to a single standard, librarians can focus on the more productive task of presenting cataloging data in the original script.

A second argument against expending effort on converting one romanization scheme to another is that it would be better to store the data in its original character form only and to build software that will create a view "on the fly" in whatever romanization scheme the user prefers. This approach is very attractive. It puts the emphasis where it should be—on the original characters. It covers all user preferences by offering choices. It would allow any new romanization schemes to be supported by simply adding another table to the software that creates the "views." It would allow catalogers to create records containing only the original characters, without the double work of having to create romanization. It would also allow wider use of the MARC records created in countries such as China and Korea; such MARC records normally do not contain any romanization.

On the other hand, there are a number of issues that need to be faced before libraries can create records containing only the original characters and rely on software to meet all romanization requirements. Some of these issues are listed below.

- All the catalog data for non-Roman material would need to contain the original characters; retrospective conversions to add original characters will need to be done by some libraries.
- Many users will still wish to use romanization in search terms, in cases where they find it quicker to type a romanized word than to create an ideographic character. In some cases, where a Chinese character with the same appearance has more than one equivalent code, searches on romanized terms produce more reliable results. To meet this need for romanized searching, library systems that stored only the original characters would need to develop the ability to search a romanized search term against indexes that contained the original characters.
- Although some library systems have developed software that will convert Wade-Giles search terms to Pinyin (such as OCLC) or assist catalogers by creating romanization from ideographic characters (such as Innopac, Dynix), there are no library systems, to my knowledge, that yet offer or plan to offer the full views capability.
- Decisions would need to be made on what to do with the romanized data
that presently exist, along with original character data, in MARC records; should they be discarded?

- Cataloging issues—such as when is a view of a heading a view, and when is it an unpreferred form?—would need to be addressed.

- The USMARC format, which enshrines romanization in the main fields (1xx, 245, etc.) might need to be changed, with consequent changes to library systems.

**CONCLUSION**

The conversion of the 500,000-record National CJK Service database has demonstrated that it is possible to convert bibliographic data from Wade-Giles to Pinyin with a high degree of automation of the conversion process. It has also demonstrated that the libraries of one Western nation, at least, can reach the agreement on Pinyin word division standards that is necessary for a large conversion project.

Some problems remain—including the conversion of authority data and the conversion of general catalogs where Wade-Giles data are scattered among many Western languages. Although it would be foolish to underestimate these remaining problems, Australian libraries now have a clear path toward their vision of a single Chinese romanization scheme, and toward its consequent benefits of more reliable searching for users, and greater opportunities for exchange of Chinese cataloging data.

**WORKS CITED**


Cooperative Cataloging: Prospects and Problems for Libraries in Saudi Arabia

Zahiruddin Khurshid

Despite witnessing the enormous success of cooperative cataloging programs in Europe and North America, the libraries in Saudi Arabia and other Arabian Gulf states have made no concerted efforts in initiating similar programs. The author emphasizes the need for cooperation in cataloging—especially of local materials. Due to unavailability of cataloging copy from the Library of Congress and other sources for a high percentage of imprints from Arab countries, and in the absence of bibliographic cooperation, libraries are doing redundant cataloging and thus wasting both time and money. The author reviews the current state of bibliographic cooperation, discusses prospects for further cooperation, and highlights difficulties in cooperation. The author concludes with recommendations that King Fahd National Library develop a national cooperative cataloging program with specific goals, set guidelines, and prepare rules and procedures for participating libraries to contribute in various cooperative cataloging programs in the Kingdom of Saudi Arabia.

Cooperative cataloging is "the original cataloging of bibliographic items through the joint action of a group of independent libraries which make the bibliographic records accessible to group members and sometimes to non-participating libraries as well. Sometimes called shared cataloging because the cataloging responsibility and cataloging product are shared" (ALA 1983, 59). The driving force for cooperative cataloging in North America has been to reduce cataloging costs. Librarians were as much concerned about the cataloging cost in 1885 when it was just $0.3575 (Tillett 1993) per item as they are now when it is $48 (Allan 1990). The cost factor becomes even more serious when it is expended over and over again in libraries for the same title. For example, if one hundred libraries acquire the same title and catalog it at an average cost of $48, the libraries together would spend $4,800 to catalog a single title. The other reasons that warranted bibliographic cooperation were the growing backlog of uncataloged materials and the lack of catalogers with language expertise or adequate training.

Impressed by the success of various

Zahiruddin Khurshid (khurshid@dpc.kfupm.edu.sa) is Senior Manager, Cataloging Operations Division, and Acting Library Systems Analyst, King Fahd University of Petroleum and Minerals Library, Dhahran, Saudi Arabia. This is a revised version of a paper presented at the Third SLA/AGC Conference on Strengthening Resource Sharing in Libraries and Information Centers in the Arabian Gulf region, Al-Ain, United Arab Emirates, March 15–17, 1995. Manuscript received March 17, 1997; accepted for publication May 2, 1997.
cooperative cataloging programs in North America, and especially the success of computerized networks, libraries in other parts of the world have built upon, and perhaps even improved, the North American model (Holley 1993). The bibliographic networks in Europe, Australia, and New Zealand are renowned for their shared cataloging activities. But the situation in Asia in general and the Arabian Gulf region in particular is not very encouraging. The prospects and problems of cooperative cataloging in Saudi Arabia discussed in this article can be generalized to the entire region.

BIBLIOGRAPHIC COOPERATION IN SAUDI ARABIA

Saudi Arabia has achieved phenomenal development in all socioeconomic sectors during the last twenty-five years. The government has spent huge sums of money to develop human resources by providing free education to its citizens from primary level to graduate level and beyond. The education program is well supported by physical facilities—including libraries. All seven university libraries have developed information resources in support of their academic programs. Both government and private organizations have established over 40 special libraries to fulfill information needs of the community in business, commerce, and industry. According to the author of one recent survey (Abbas 1994), there are 69 public libraries throughout the kingdom, of which 59 are controlled by the General Directorate of Public Libraries under the Ministry of Education. The rest are controlled by other ministries.

Several academic and special libraries are currently undertaking automation projects, and at least one academic library at King Fahd University of Petroleum and Minerals (KFUPM) is fully automated. However, if we review the numerous articles written on the state of libraries and librarianship in Saudi Arabia, we find that most authors have talked about the lack of cooperation among libraries (Khurshid 1979; Line 1983; Ashoor 1989). Each library has generally followed its own course in achieving the goal of collection building, automation, library services, etc. Not only have the individually projected goals not been achieved, but duplication of effort has wasted a great deal of time and money. While there is a comprehensive need for cooperation in each area of library activity, I will attempt to address a more pressing need for bibliographic cooperation in Saudi Arabia and how it can be achieved.

The collections of libraries in Saudi Arabia in general are comprised mostly of Arabic materials. Cataloging copy for about 75% of Arabic materials is not available from any sources, local or external. The following statistics show the representation of various language materials in CDMARC Bibliographic, the Library of Congress database of approximately five million Machine-Readable Cataloging (MARC) records on CD-ROM. Arabic records form less than 1% of the entire database:

<table>
<thead>
<tr>
<th>Language</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>34,750</td>
</tr>
<tr>
<td>Italian</td>
<td>126,227</td>
</tr>
<tr>
<td>Spanish</td>
<td>288,388</td>
</tr>
<tr>
<td>French</td>
<td>348,630</td>
</tr>
<tr>
<td>English</td>
<td>3,205,798</td>
</tr>
</tbody>
</table>

The average number of Arabic records added to the CDMARC Bibliographic database every year (based on a ten-year period, 1986–95) is 1,107, which represents approximately 12% of the Arabic titles produced in the entire Arab world (United Nations 1995; International Marketing Data and Statistics 1996). Aside from representing a small percentage of Arabic records, CDMARC Bibliographic (or for that matter any other Western database) has two additional problems. First, the Arabic records are presented in romanized (transliterated in English) form, which is not acceptable to Arabian local users and therefore generally not used by Arab libraries. Second, Arab libraries use their own name and subject headings, which means cataloging copy from external sources would require a great deal of changes in headings.

It is clear from the above facts that the cataloging copy of Arabic materials available from external sources is not only limited but also is not fully compatible with
local cataloging practices. That is why the percentage of original cataloging of Arabic materials in Saudi libraries is around 90%. In the absence of any shared cataloging program, each library is doing its own cataloging of many titles that have been cataloged by other libraries.

**CURRENT STATE OF COOPERATION**

There has always been some awareness among Saudi libraries about the need for shared cataloging, but it did not bring about any tangible results. The primary reason is the absence of a forum or an agency that could initiate, organize, and coordinate the activities. However, some activities have been initiated by a few individual libraries as well as one group of libraries, which can be considered as forerunners of bibliographic cooperation in Saudi Arabia. Of them the three major activities are discussed below.

1. **Cataloging-in-Publication.** Several publishers are currently providing Cataloging-in-Publication (CIP) data with their publications. Notable among them are the university presses of King Abdulaziz University, King Saud University, Imam Mohammad Ibn Saud University, and Umm Al-Qura University. The CIP data included in university publications are generally prepared by the library staff of the respective universities. Although the number of publications with the CIP data produced by these presses is quite small, it still helps to reduce the expense of redundant cataloging.

   The most significant development that has taken place in the area of CIP in recent years is the assignment of a new role to King Fahd National Library (KFNL) in Riyadh as the central agency in preparing CIP for books published in Saudi Arabia. The national library was created in 1990 through a royal decree. The library is governed by a Trustee Council consisting of a chairman and five members. While the members are appointed by the King, the council appoints the director, who is responsible for administering the library in the light of the goals set by the council. Before KFNL assumed its recent role, only a small group of publishers—including the university presses—were preparing CIP data. The KFNL has hired some Arabic and non-Arabic catalogers to cope with the increasing amount of CIP cataloging. (Non-Arabic catalogers are generally hired from India and Pakistan and can read Arabic but may have difficulty in understanding and speaking Arabic. They are normally responsible for cataloging non-Arabic materials because of their English-language skills.) If the publishers throughout the kingdom cooperate by sending the title page and preliminaries of books nearing publication to the National Library so that preliminary cataloging data can be prepared, it can bring about the same results as have been produced by the CIP program of the Library of Congress.

   According to a survey, the CIP program saved the U.S. libraries an estimated $1.15 billion in 1990 by providing authoritative and timely bibliographic access to materials that libraries otherwise would have had to catalog themselves (Thomas and Younger 1993). No such survey has been done in Saudi Arabia to estimate the national saving from the CIP program of the KFNL. However, the program is growing quickly, and with the increasing participation of all major publishers of Saudi Arabia in the program, it is expected that approximately 80% of the titles produced in the kingdom (between 350 to 400 titles per year) will include CIP data, thus saving the Saudi libraries approximately $1 million a year (the cost of 400 common titles cataloged in fifty libraries at an average cost of $50 per title).

   Another important benefit of the program will be the possible improvement in the quality of cataloging, which has been poor due to
nonadherence to cataloging standards by libraries and the lack of qualified and trained catalogers.

2. Remote access to online catalogs. In the early 1980s, a few libraries started automation projects, and by the end of the decade sixteen Saudi libraries had been engaged to a varied extent in automation enterprises (Al Dosary and Ekrish 1991). DOBIS/LIBIS and MINISIS are the most widely used software packages. In fact, there are five users of DOBIS/LIBIS (excluding Umm Al-Qura University [UQU], which leased DOBIS/LIBIS in 1986 but has not yet installed it) and an equal number of users of MINISIS. Some libraries are using in-house-developed software.

Of the libraries engaged in automation, 75% have online catalogs in operation, and all of them have created databases ranging in size from a few thousand records to more than 300,000. Although there has been no exchange of records between automated libraries, the possibilities of exchange do exist—at least among the libraries that have the same system. Some years back a couple of DOBIS/LIBIS user libraries approached the KFUPM Library with a request to allow them to import KFUPM records into their local databases. The KFUPM Library could not accept this request for the reason that about 60% of its records were downloaded from the LC MARC tapes, and copyright regulations do not allow libraries to redistribute LC MARC records. However, the library was and is still willing to exchange records created locally. Last year, as a result of several years of discussions between three DOBIS/LIBIS user libraries (King Abdulaziz, King Fahd, and King Saud Universities), an arrangement was made to provide remote access to the online catalog of each library from the other two. As of 1996, the KFUPM Library online catalog was accessible from King Abdulaziz and King Saud University Libraries. However, the KFUPM Library has access to the King Saud University Library catalog only. Access to the King Abdulaziz University Library catalog will, it is hoped, be available soon.

The database of these three large libraries together can provide access to more than half a million records, which include duplicate and triplicate records for approximately 30% of the titles (identified in accessions lists of the three libraries) that are available in all the three libraries. If the libraries were to engage in shared cataloging via remote access to their respective catalogs, there could be a significant drop in the original cataloging needed by each library.

3. Access to KACST bibliographic databases through Gulfnet. Gulfnet is a computer network established by King Abdulaziz City for Science and Technology (KACST) in Riyadh, Saudi Arabia. It was started in May 1985 to facilitate exchange of data, information, and programs for use by scientists, researchers, librarians, and information professionals in Saudi Arabia and the Gulf states. At present 11 universities and research centers in Saudi Arabia, Kuwait, and Bahrain are connected to Gulfnet through public leased telephone lines. There is also a dial-up communication network, KACSTNET, which connects 60 research institutions and libraries in the kingdom with the central computer at KACST.

Among the services offered by the network is the searching of various databases, including a union list of serials and the KACST library catalog. The union list of serials database was developed by KACST and contains 20,000 titles held by academic and research libraries in Saudi Arabia. It has become a primary source of locating materials for interlibrary loan. Another online union list of Arabic periodicals held by university libraries of Saudi Arabia is under preparation. The KACST library
catalog (which can be searched through Gulfnet and KACSTNET) provides information about the KACST collection, which includes over 11,000 books, 350,000 technical reports on microfiche, and over 60,000 source documents relating to the national databases produced by KACST (Al-Tasan 1992).

The use of Gulfnet to share cataloging data has been very limited. Only the KACST catalog is available for searching by member libraries. Access to catalogs of other libraries is not yet available, and there is not much hope that Gulfnet services will be expanded further once full Internet connectivity becomes available in the kingdom.

**PROSPECTS FOR FURTHER COOPERATION**

The current state of cooperative cataloging reviewed in the preceding section cannot be considered satisfactory. The programs have been undertaken in isolation and are very limited in scope. In developing these programs, no efforts seem to have been made to learn from the experiences of cooperative cataloging programs developed by Western libraries. Some of the renowned programs might produce similar results if adapted to local requirements. In the following section, we will discuss the prospects for developing cooperative cataloging programs in the Kingdom of Saudi Arabia.

**UNION CATALOGS**

As mentioned earlier, DOBIS/LIBIS and MINISIS are the two major systems currently in use in Saudi Arabia. Of the two systems, DOBIS/LIBIS has excellent capabilities for networking, which have not been fully exploited. The DOBIS/LIBIS network structure allows participating libraries to share a union catalog in which a record will be created once by the library that first acquired an item. If copies of the same item are acquired by other libraries, they can either add their location information to the existing bibliographic record, or they can create a separate file of local copies.

Of the five DOBIS/LIBIS user libraries, the KFUPM Library is in the best position to act as the central node in the network. It has highly qualified and motivated staff who can prepare quality records and also review and upgrade the quality of records from participating libraries according to international standards. The "correct cataloging" subfunction of DOBIS/LIBIS can allow authorized catalogers to review and correct catalog entries of other catalogers who have lower authorization levels (Khurshid 1994).

Another union catalog can be developed by MINISIS user libraries with the central node at KFNL. This library is best suited to be the central node because most of the libraries that use MINISIS are located in Riyadh, which would make communication easier between member libraries and the central node. This does not mean that communication with the central node cannot be made from remote locations. However, that would require high-speed communication lines as well as some additional hardware.

If the new versions of DOBIS/LIBIS and MINISIS support the Z39.50 protocol, then it would be possible to link the two union catalogs and expand the scope of the database for cooperative cataloging. While the successor product of DOBIS/LIBIS is expected to have Z39.50 support, we do not know whether future versions of MINISIS will provide this support.

**NATIONAL BIBLIOGRAPHIC AGENCY**

Each country in the Gulf region is expected to create a National Bibliographic Agency (NBA) as an integral part of the national library or as an independent agency with the responsibility for preparing authoritative and comprehensive bibliographic records for all new publications issued in the country. They are further expected to create those records in accordance with accepted bibliographic standards (IFLA 1979). King Fahd National Library is already carrying out some of the tasks of a national
bibliographic agency, such as preparing a national bibliography and CIP data. However, in order for libraries to take advantage of these services, some improvements need to be made. These include:

1. The National Library should distribute its cataloging records in machine-readable form so that libraries can download them into their local systems and save time and money.
2. The National Library should implement a comprehensive legal deposit system so that the library gets each and every title published or issued (in all languages and formats) in the country. This will also increase the availability of cataloging copy of local materials.
3. The CIP program should be expanded to include publications from all major publishers—including those from university presses. This would especially benefit small libraries, which cannot make use of machine-readable records due to lack of computer infrastructure.

ACCESS TO ONLINE CATALOGS THROUGH THE INTERNET

Saudi Arabia is expected to join the Internet soon with a central node at KACST in Riyadh. Among the numerous information resources available on the Internet are hundreds of online catalogs of academic and special libraries that offer opportunities for any library catalog to be treated as a bibliographic utility. Following the example of Western libraries, the libraries in Saudi Arabia could offer their catalogs on the Internet both for use by local libraries for cataloging and for locating materials for interlibrary loan. Those libraries that cannot make their catalogs available on the Internet could at least make them accessible through dedicated or dial-up lines.

DIFFICULTIES IN COOPERATION

There are some difficulties that must be overcome in order for cooperation to work successfully.

1. Leadership. To start a nationwide cooperative cataloging program, a strong central agency is needed that can provide leadership to the libraries participating in the program. Being a national agency, KFNIL is in a better position to provide this leadership. In addition to acting as a depository library of all Saudi publications, it has a mandate to compile a national bibliography, prepare CIP data, assign ISBN and ISSN numbers, etc. To carry out these and other activities, the library must muster the support of all libraries that can make effective contributions to the program. The Library of Congress provides an excellent model of leadership for any national agency that aims to engage in a cooperative cataloging program.

2. Standardization. The cataloging records produced by the participating libraries in a cooperative program must be consistent, which requires application of standards. While all libraries claim to follow standards in preparing records, the real situation is quite different. Authors of a KFUPM study in 1987 revealed serious inconsistencies in the cataloging practices of university libraries of Saudi Arabia (Khurshid 1988). The libraries did not fully adhere to cataloging standards. Librarians at several libraries have modified, revised, or interpreted cataloging rules to satisfy local needs. There is no coordination among them to prepare national-level standards.

Whatever standards have been prepared are mostly individual efforts and are out-of-date. For example, the Arabic subject headings list prepared by Swaydan (1985) and adopted by all university libraries in Saudi Arabia as a standard list has not been revised since its first publication. As a result, many libraries have taken liberty in adding subject headings for new disciplines on their own—without consulting or notifying other libraries.

Another problem in the area of standardization is the absence of a
national MARC format for the exchange of bibliographic information. At present libraries are using a variety of formats, including USMARC, extended USMARC, and locally developed MARC. In the case of non-Arabic data, DOBIS/LIBIS and MINISIS user libraries are using interface programs to import bibliographic records from Western sources. These programs might facilitate exchange of records between libraries having the same system. However, the exchange of records between libraries with different systems has been a real problem. How this problem can hamper the progress of a cooperative program can be seen in the case of the KACST union list of serials project. When the project was initiated in 1985, the KFUPM Library and some other automated libraries provided their serials records in machine-readable form, but they could not be loaded into the KACST system because of incompatibility with the KACST record format. As a result, KACST had to rekey these records into their system manually. Updating this list is also difficult for the same reason.

The lack of a national MARC format that handles Arabic script is yet another part of the problem. Many libraries have either modified the original software, for example DOBIS/LIBIS, or have developed in-house programs to create Arabic records. These programs are not compatible with the Arabic script program developed by the Research Libraries Information Network (RLIN) and therefore cannot accept records from the RLIN database or from the Library of Congress, which also uses the RLIN program to catalog Arabic materials.

The development of the system at RLIN to record titles in both the romanized and Arabic script (which was underwritten by the Kuwait Foundation for the Advancement of Sciences) did not gain acceptance as a standard system by Arab libraries for two reasons. First, it treats Arabic information and its romanization as paired fields, and second, it places the romanized data before the original script representation in an online record. In the USMARC equivalent of the same record, the romanized data are given in a regular field and their alternate graphic representation in an 880 field, with a sub-field 6 in each field to create a link between them (Aliprand 1992).

Arab librarians do not wish to see Arabic information in script form subordinated to romanization in an online record; rather they would prefer to see the complete record in their own script. They feel that instead of asking Western agencies to prepare a MARC format to accommodate Arabic script for them, they should come forward and prepare their own format. In this regard librarians at the KFUPM Library prepared a study in 1994 to review the current state of Arabization by individual libraries and institutions at local, regional, and international levels and to seek answers to several questions that will help to provide the framework for developing an ARABMARC (Al-Muhtaseb, Ashnoor, and Khurshid 1994). Further progress in this area depends on the willingness of a national or regional organization to coordinate or provide funding for the project.

The use of a varied array of character sets in commercial applications, such as ASMO 449 in DOBIS/LIBIS and MINISIS, and an RLIN extended Arabic character set in VTLS will continue to be a major problem in the exchange of Arabic data unless software vendors decide to adopt Unicode as a common standard. Unicode might not be a perfect standard for processing Arabic characters, but it can be improved through participation of individuals as well as libraries from Arab countries. Unfortunately no one from Saudi Arabia or the Gulf
region is known to have been participating in the development of Unicode. Unicode may also be a better solution for Arab libraries that are currently addressing the issue of the Arabic character set to be implemented in ARABMARC.

3. Telecommunications and computer infrastructure. Cooperative cataloging requires a certain level of telecommunications and computer infrastructure (Holley 1993). This level has not been fully achieved in Saudi Arabia. The telecommunication system in particular is not well developed. As a result, libraries are not taking full advantage of Internet resources. The Internet is being used in cooperative cataloging projects in the United States both to communicate in a timely fashion between librarians and to access catalog records in other online catalogs (McCombs 1994). Without the Internet or other telecommunications networks, libraries in Saudi Arabia cannot achieve much success in bibliographic cooperation.

The present communication link among libraries in Saudi Arabia is very limited and is through dedicated or dial-up lines, which are slow and not very reliable. This also hampers any computerized cooperative cataloging program in the country, which requires quick access to information and the exchange of large amounts of bibliographic data between libraries. Except for most university and special libraries, which have mini- or mainframe systems, other libraries do not have the level of computer infrastructure that is essential for participating in computerized cooperative cataloging programs. This level needs to be improved before any national-level cooperative program is implemented.

Having reviewed the prospects for and difficulties in bibliographic cooperation, it is appropriate to highlight the major steps that need to be taken to make the cooperation work successfully. The national library should implement the legal deposit comprehensively, expand the CIP program, develop and adopt cataloging standards including the ARABMARC format, and prepare and distribute records in machine-readable form. Telecommunications and the computer infrastructure should be improved to allow libraries to provide remote access to their online catalogs through Gulfnet and dial-up lines and eventually through the Internet once full connectivity becomes available in the kingdom.

CONCLUSION

The vision of cooperative cataloging in Saudi Arabia should parallel that established in other countries: "to put into place the necessary support to catalog, once and only once, every item owned or made accessible by libraries and to share that information with all others who need it" (Thomas and Younger 1993, 257). Librarians in Saudi Arabia and elsewhere in the Gulf region have some understanding of the advantages of cooperative cataloging programs. But they do not have a clear idea about how these programs can be developed and implemented and in what way they can participate in these programs. It would therefore be appropriate for a national agency—preferably King Fahd National Library together with some notable academic and special libraries in the kingdom—to initiate a program similar to the renowned Program for Cooperative Cataloging (PCC) of the United States, with a clear mission and goals. The Saudi Arabian PCC equivalent could set guidelines and prepare rules and procedures for participating libraries to contribute through various programs, including the creation of bibliographic databases, union lists of serials, Arabic name authorities databases, etc. The bibliographic component of the program could also define a core-level bibliographic record containing a uniform level of detail, as is done by the PCC. The scope of the program could be widened to develop a set of Arabic cataloging standards for preparing uniform records (including the ARAB-
MARC format), provide training, and seek funds to support cooperative library projects in Saudi Arabia and in the entire Gulf region.

**Works Cited**


To the Editor:


I hope LRTS will accept a non-scholarly letter from a non-scholar. The following caught my eye.

"While not diminishing the importance of shelf browsing, catalog departments must weigh the cost of shelving and reviewing classification carefully against its perceived benefits."

This seems to be an example of Dewey's Law (Harry Dewey): "What saves the librarian's time frequently costs the reader his."

As a librarian who labored before the widespread use of OPACs, my information may be dated, but it is my understanding that most books in public and academic libraries are selected by browsing.

If this is true, one of the most important tasks librarians perform is to place a book where a browser might find it.

To my mind it's a user's use of a book that justifies the cost of classification.

A study needs to be done to determine if a book's circulation is affected one way or another by local shelisting.

Herb White wrote, as I recall, if his library had two copies of the same book, he would use different classification numbers to "tap" different browsers.

If cutting catalog costs is paramount, perhaps we should close the stacks and shelve books by size and save space as well.—Marvin H. Scilken, M.L.S., editor, The Unabashed Librarian

To the Editor:

I would first like to thank Crystal Graham for her correction of my error regarding serials records. I am guilty indeed. My statement regarding added entries for earlier and later titles for serials was justified, incorrectly, as follows. AACR2 rule 21.30G for added entries for related works directs the cataloger to make added entries for works closely related to the work being cataloged. Rule 21.30G refers to rule 21.28 (Related works) for guidance. Because rule 21.28 includes "continuations and sequels," I assumed that serials would also require added entries for title changes.

Second, please note that an important reference is missing in the published article:


This reference should follow the reference to Svenonius. Thanks are due to John M. Cys, Moffett Library, Midwestern State University, for pointing out this omission.—Allyson Carlyle, Graduate School of Library and Information Science, University of Washington
Celebrating C. Sumner Spalding

Robert M. Hiatt

More than 150 admirers gathered March 15 at St. Barnabas Episcopal Church, Temple Hills, Maryland, to celebrate the life of Charles Sumner Spalding, who died March 8 in Baltimore from complications of pneumonia. This remarkable man was an internationally renowned authority on cataloging and an accomplished organist, composer, and band and choir director. Most of the music played at the celebration was Spalding’s own. He was also a good friend whose sense of humor, devotion to others, and ability to laugh at himself were hallmarks.

Born in Somerville, Massachusetts, in 1912, Spalding received bachelor’s and master’s degrees in music from Harvard University in 1933 and 1934 and a master’s degree from the School of Library Service, Columbia University, in 1940 before joining the Library of Congress as a music cataloger later that year. With the outbreak of World War II, Spalding took a leave of absence from the Library and joined the U.S. Army, where he served as bandmaster of the 362d Army Band. Although he admitted that he had no formal training to serve as a conductor, his hard work, leadership, and concern for his troops resulted in lifelong friendships that were renewed at periodic reunions of his band members.

After the war, Spalding returned to the Library of Congress, where he served as assistant chief and chief of the Catalog Maintenance Division, chief of the Serial Record Division, chief of the Descriptive Cataloging Division, and assistant director (cataloging) (equivalent to the current position of director for cataloging), the position that he held when he retired from the Library in 1975. Upon his retirement Spalding received the Library’s highest award, the Distinguished Service Award.

During his professional life, Spalding represented the Library of Congress at many national and international conferences and served on many boards and committees. The 1961 International Conference on Cataloguing Principles in Paris, at which he represented the Library, adopted an internationally accepted set of principles of choice and form of entry. These principles led to the first edition of the Anglo-American Cataloguing Rules, for which Spalding served as general editor and for which he received the Margaret Mann Citation, the highest award in cataloging and classification awarded by the American Library Association. He also represented the Library at the 1969 International Meeting of Cataloguing Experts in Copenhagen, at which the machinery was set in motion for a series of International Standard Bibliographic Descriptions. At the national level, Spalding served as the Library’s representative to the United States Board on Geographic Names (BGN), an interagency organization responsible for providing uniformity in geographic nomenclature and orthography throughout the federal government. At the time of his retirement he was chairman of BGN. Former director for acquisitions Glen A. Zimmerman was heard saying that Spalding was “an icon in the library world, particularly in the area of cataloging. Sumner

Robert M. Hiatt is Senior Cataloging Policy Specialist, Cataloging Policy and Support Office, Library of Congress.
played a major role in the Cataloging in Publication Program right from the beginning. What struck me was his willingness to try something entirely different, apart from the cataloging status quo, apart from a very well-established reputation in the field.”

During his retirement Spalding continued to pursue all his interests, including continuing analysis of professional cataloging matters, full participation in the musical life of several churches, and musical composition. At 81 he was granted a patent for a specialized electronic device to enhance the sound of an organ.

C. Sumner Spalding will long be remembered not only as a giant in the cataloging and classification community but also as a true Renaissance man. His two sons, Charles Spalding, Jr., and Guy W. Spalding, make their homes in the Washington area. Donations to celebrate the life of Charles Sumner Spalding can be given to St. Agnes Hospital, Baltimore, Maryland.
Mark It On Your Calendar Today!

PUBLIC LIBRARY ASSOCIATION’S
SEVENTH NATIONAL CONFERENCE
MARCH 10-14, 1998 / KANSAS CITY, MISSOURI

PUBLIC
LIBRARIES

Vital, Valuable, Virtual

MORE THAN 100 PROGRAMS / 200 PLUS EXHIBITORS
PRECONFERENCES / AUTHOR LUNCHEONS / AND MORE!

PUBLIC LIBRARIANS: THE PUBLIC LIBRARY ASSOCIATION IS FOR YOU!
PLA members receive substantial price breaks and advance registration
and housing information. Join today!
For information call 1-800-545-2433, ext. 5PLA

PLA is a division of the American Library Association
A New Service on the Information Superhighway

Authority Express

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

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

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

“Authority Control for the 21st Century”

LIBRARY TECHNOLOGIES, INC.
2300 Computer Avenue, Suite D-19 Willow Grove, PA 19090
(215) 830-9320 Fax: (215) 830-9422
(800) 795-9504 email: LTl@LibraryTech.Com
SPEED YOUR CATALOGING WITH TWO DYNAMIC PRODUCTS FROM THE LIBRARY OF CONGRESS

Classification Plus & Cataloger's Desktop

- Increase accuracy. Pinpoint the right schedule or rule.
- Speed cataloging. Find references quickly.
- Classify and describe your collections more easily.
- Keep up-to-date with quarterly cumulations.

FTP A DEMO FROM OUR WEB PAGE!
http://www.loc.gov/cds/cclroms1.html#demo

Call (800) 255-3886 in the U.S.
or (202) 707-6100

Library of Congress
Cataloging Distribution Service
101 Independence Avenue, S.E.
Washington, DC 20541-4912

Products may be copied onto a LAN hard drive for multiple users.

Classification Plus
Get the latest LC Classification schedules plus the complete Library of Congress Subject Headings.

Single User
$470
Each Additional User: $16

Cataloger's Desktop
Access the most popular and authoritative Library of Congress cataloging publications.

Single User
$870
Each Additional User: $18

Discount Combination Price:

Single User
$1,150
Each Additional User: $32

Ask about prices outside North America