Volume 12,
Number 4

ISSN 0730-9295

Information Technology and Libraries

December 1993

CONTENTS

ITLBDC 12(4) 377–464 (1993); Inf Tech Libr

ISSN 0730-9295

Information Technology and Libraries

Volume 12, Number 4: December 1993

CONTENTS

ITLBDC 12(4) 377–464 (1993); Inf Tech Libr

Information Technology and Libraries

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

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 88 of the March 1992 issue. Manuscripts of articles, communications, and news items should be addressed to the editor: Thomas W. Leonhardt, *Information Technology and Libraries,* Bizzell Library, 401 W. Brooks, University of Oklahoma, Norman, OK 73019-0528; (405) 325-2611; Bitnet QC6305@UOKMVSA. Copies of books for review should be addressed to: Susan B. Harrison, *ITAL Book Reviews,* The New York Public Library, 455 Fifth Ave., New York, NY 10016. Copies of software for review should be addressed to: George S. Machovec, *ITAL Software Reviews,* Colorado Association of Research Libraries, 3801 E. Florida Ave., Ste. 370, Denver, CO 80210-2542. Advertising arrangements should be made with Stuart Foster, Advertising Sales Manager, and Dolores LaPointe, Advertising Coordinator, c/o *Choice,* 100 Riverview Center, Middletown, CT 06457; (203) 347-6933.

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

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology Association. Subscription price, $22.50, is included in membership dues. Nonmembers may subscribe for $45 per year in the U.S.; $50 in Canada, Mexico, Spain, and other PUAS countries; $55 in other foreign countries. Single copies, $15. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. *Postmaster*: Send address changes to Information Technology and Libraries, 50 E. Huron St., Chicago, IL 60611.

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

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

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

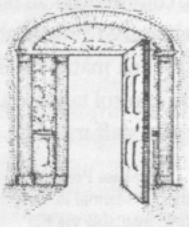
The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI 239.48-1984. ∞

A Better Way To Search Databases

We started in 1985, database searchers committed to better search software design. We became the premier vendor of Medline, then expanded our catalog to other databases. Last year we won Information World Review's PRODUCT OF THE YEAR for faster, easier search software. But a better way means meeting the evolving needs – individual and campus wide – of today's library users.

Announcing OVID: a database interface so flexible it molds itself to your search environment.

With OVID you're free to move from one operating system to another without retraining. OVID's Common User Interface assures identical functionality in DOS, Windows and UNIX.



A haven for beginners, OVID's Easy Mode has on screen prompts. The more experienced can pull-down menus showing an array of search options. Experts will feel at home using online syntax.

Search with natural language if you like. OVID mapping cuts through the mystery of controlled vocabularies, homing in on precisely matching subject headings.

There's an unprecedented array of search tools – indexes, thesauri, limits and fields – many never before available in an interface. They're all standard OVID features.

HELP for every search function is context-sensitive and on screen, never more than a keystroke or mouse click away.

OVID. A better way to search ERIC, Current Contents®, PsycINFO®, Medline, Readers' Guide Abstracts, EMBASE and more.

CD PLUS Technologies

New York 800-950-2035/212-563-3006 London 44-(0)81-748-3777 Amsterdam 20-672-0242

Authority Control

Consistent and current authority headings are the key to the effectiveness of any library catalog. The *WLN MARC Record Service (MARS)™* compares your headings to LC Name and Subject, LC Children's, NLM medical subject and/or WLN authority files; verifies or upgrades all headings to current, authorized forms; and supplies a complete and accurate set of matching USMARC authority records including notes and cross references. At WLN, we offer personalized service throughout your project, strict adherence to project schedules, comprehensive bibliographic record and authority heading upgrading programs, customized programming based on your specifications, auxiliary authority files for improved match rates, and three levels of manual review. Most of all, we offer authority control tailored to your needs at a price you can afford. Let the *MARS* team be your extra staff member, invisible yet indispensable.

WLN MARC Record Service: Authority Control . Database Preparation . Customized Programming . Smart Barcode Labels . Bibliographic Record Upgrading . Manual Review . CD-ROM Public Access Catalogs . Ongoing Authority Updating/Database Updating, available via FTP.

Yes, I want to find out more about authority control and how the *WLN MARC Record Service (MARS)* can help my library.

Name_____ Title_____

Institution_____

Address_____

_____

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

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

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

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

Projected start date_____ Completion date_____

Comments_____

Signature_____

Please mail to: WLN, P.O. Box 3888, Lacey, WA 98503-0888. Or FAX to (206) 923-4009.
Just call us at 1-800-DIALWLN for more information about authority control.

WLN

TURNING INFORMATION INTO KNOWLEDGE

Cataloging Standards and Machine Translation: A Study of Reformatted ISBD Records in an Online Catalog

Gregory J. Wool, Bart Austhof, Anita Breckbill, and B. Larry Mozer

Labeled bibliographic display screens in online catalogs can repackage records created for card catalogs in ways that restructure the record, redefine data categories and contexts, and add or omit selected categories of data. The impact of automated display on catalog records in a medium-sized research library was studied by comparing the card and online versions of 1,005 records created according to the International Standard Bibliographic Description (ISBD) conventions. Thirty-eight types of changes ascribable to the catalog software were identified, and the extent of occurrence in the sample was tabulated for each. Changes found involve rearrangement, inaccurate labeling, repetition, addition, and omission of data elements, along with the elimination of the traditional distinction between descriptive data and access points. The findings suggest that current descriptive cataloging standards may be ill-suited to the creation of records for online display.

During the past few years, online catalogs have become ubiquitous in large and medium-sized libraries. The capabilities that these catalogs have and that card catalogs lack have numerous revolutionary implications for the future of bibliographic control and access. One of these capabilities is the automatic reformatting of bibliographic records into a labeled, tabular display format, based on the mapping of labels to MARC fields, with the labels, the mapping, the choice of data elements displayed, and the order in which they are displayed determined locally. This feature has proven extremely popular with librarians, to the point that few if any systems now on the market do not offer it. It is probably safe to say that the resulting line-by-line field displays beside brief, locally standardized verbal tags have replaced the familiar paragraph-style ISBD arrangement in most libraries by now. In a very short time, and with very little fanfare, the face of the catalog record has been transformed.

Although the benefits of customizing a library's bibliographic records in this way are self-evident to many, the use of machine reformatting for display raises a number of issues. These include the optimal length of record displays, the relative appropriateness of the various data elements for display, and the

Gregory Wool is Assistant Professor and Monographs Cataloger, Iowa State University Library; Bart Austhof is Assistant Professor and Head, Documents and Maps Section, Donald O. Rod Library, University of Northern Iowa; Anita Breckbill is Assistant Professor and Music and Architecture Cataloger, University Libraries, University of Nebraska-Lincoln; B. Larry Mozer is a Library Technician II in the Serials Cataloging Section, University Libraries, University of Nebraska-Lincoln. A condensed version of this paper was presented on June 28, 1993, at a Research Forum sponsored by the Library Research Round Table at the American Library Association Annual Conference in New Orleans.

way bibliographic items are to be identified and cited, all of which are issues specifically related to the selection and arrangement of bibliographic information. Issues related to labeling include the clarity to the user and accuracy of the labels chosen (or, put another way, the adequacy of the labels for conveying the bibliographic concepts implicit in *AACR2* and the MARC formats) and what to do about any inaccuracies, unacceptable ambiguities, clutter, or other noise that may result from an overly literal machine translation from ISBD and *AACR2* to a library's chosen version of "patronese."[1]

The above reference to machine translation is neither idle nor hyperbolic. In describing the nature of bibliographic records, Hagler writes:

> Bibliographic data are communicated by means of a language. Since they are expressed in the words of a language such as English or French, it may seem that no other language is involved, but bibliographic language is also present with its own syntax. The "sentences" of bibliographese have a rigid structure and their expression is concise and often elliptical. As with any other language, the beginner can quickly comprehend much, but instruction and experience are needed before all the details and subtleties are either noticed or understood. Like any other language, the language of bibliography is nothing but a set of commonly accepted conventions.[2]

Hagler also calls the format of a record "the grammatical structure, or syntax, of the bibliographic language."[3] Neville says much the same thing and describes cataloging as itself a translation process, one that "cannot be mechanized, since it contains intuitive and pragmatic steps for which watertight algorithms could not be written."[4]

Because of the complexity and subtlety involved in any translation process, machine translation remains an activity fraught with problems. Its goal, after all, is to reproduce meaning from one system in another. In the editor's introduction to a volume of essays on the topic, Nirenburg writes:

> The task of MT [machine translation] can be defined very simply: the computer must be able to obtain as input a text in one language (SL, for source language) and produce as output a text in another language (TL, for target language), so that the meaning of the TL text is the same as that of the SL text. It is clear that finding a way of maintaining invariance of meaning is the crucial problem in MT research.[5]

Of course, in a technical sense, machine reformatting of bibliographic records has been commonplace since the 1970s. Libraries using bibliographic utilities such as OCLC produced their card catalog records by inputting data in one format—the MARC format (as defined by their utility)—and having the utility translate the data into a standard *AACR* arrangement for output on cards. When a library goes online with labeled displays, those same MARC records (now in the automation vendor's version of MARC) are simply translated again into a different format. What, then, is the problem? It is this: all those records were created following *AACR* or *AACR2*, both of which specify a paragraph-style, unlabeled display format with the tracings displayed separately, if at all; in other words, the language of card displays. The *AACR* data were then translated by inputting into the MARC format and were translated back, by a card printing program, into the prescribed *AACR* display. The process of breaking up the card record, rearranging it, and adding labels based on a predetermined mapping of labels to MARC fields is really a translation, not so much from MARC, as from *AACR* or *AACR2* to public display.

What this means is that for much of the North American library community, *AACR2* and the International Standard Bibliographic Descriptions on which it is based no longer govern cataloging output (i.e., bibliographic display). However, they continue to govern cataloging input (i.e., the content of records). Catalogers are still creating records designed for unlabeled, proselike, description-here-and-tracings-over-there display—in other words, for catalog cards—and their work is transformed by computer software into something quite different, even though the data remain the same. Is *AACR2* really so flexible, though, that records created in the language of card-style display can be reformatted any number of ways without encountering the problems that attend machine translation?

The broader question is, Just what is the relationship between labeled record displays in OPACs and the current descriptive cataloging rules and standards? The study reported below represents one approach to this ques-

tion. Because the revised *AACR2* mandates the ISBD display format and prescribes a certain order for the display of bibliographic data elements, many libraries are abandoning these elements of *AACR2* in their use of OPAC labeled displays. But data recorded in a format that uses position, prescribed punctuation, spacing, and context to give them meaning have been transferred intact to a format where meaning is supplied by one- or two-word labels and the data themselves are rearranged.

Are the standardized labels an adequate substitute for the grammar of the ISBD display? Does the clustering of similar data items beside a single label help or hinder their interpretation? Can the labels themselves change the meaning of the data? This study was undertaken to shed light on these and other related questions. Its specific objectives, pursued within the confines of a single catalog, were to (1) identify the types of difference that exist between ISBD card catalog records and their labeled, online counterparts and (2) document the extent of these changes in the catalog (i.e., the percentage of records affected).

LITERATURE REVIEW

The authors were able to find only one previous study of the impact of machine reformatting on bibliographic records. This study, as reported by Crawford, Stovel, and Bales for the Research Libraries Group (RLG), considers display issues only, ignoring effects on the content and the possibility of translation problems.[6] In fact, although much has been written about online catalogs, the impact of automation on cataloging, and the potential impact of automation on cataloging rules and standards, very few references to labeled bibliographic displays appear in the literature. Accounting for this, at least in part, is the newness of the phenomenon. However, it may also reflect a widespread assumption that any change represented by labeled displays is no more than cosmetic.

Even the literature on OPAC design issues contains but a few passing references to record reformatting, concentrating instead on help screens, indexing, and the display of lists of records. The reformatting of catalog records is considered almost exclusively in terms of its potential for improving the catalog's visual appeal and making existing catalog data easier to understand. Matthews postulates the qualities of user-friendly screen displays, including the consistency of display, the use of labels for all information variables, and the suppression of unnecessary data.[7] (However, he also calls for standardization of labeling terms.)[8] Allen focuses on the need for a variety of displays to better serve users with differing perceptual abilities and information needs.[9]

The above-mentioned RLG study found that most of the users who were questioned preferred labeled to card-style displays, but also that labels, wide margins, and blank space between lines—while each dramatically improving readability and visual appeal—forced records that would otherwise fit on one screen onto multiple screens, making information beyond the first few lines much less likely to be consulted.[10] Stephens expresses concern about excessive use of screen space, suggesting: "As display standards develop for OPACs, librarians should also consider utilizing display position as an alternative to text labels for certain kinds of data."[11] He also stresses the importance of blending "information . . . from a variety of sources . . . into a single coherent display."[12]

More recently, however, a very different problem with labeling was noted by Crawford (one of the authors of the RLG study) in an address at the 1991 ALA Annual Conference:

> And, of course, if you use labels, you need to determine what these labels are. Not all that easy. I have yet to see a name for the 265 field or the 300 field that causes me to jump up and say, "That's it!" And I'm not sure there is a single best name for 6xx fields or 7xx fields or some others. Arcane fields, particularly coded fields, are very difficult to label—but they may be inappropriate for display in any case.[13]

Other authors consider issues pertinent to labeled record displays, though usually without reference to labeled displays themselves. Carpenter argues that bibliographic data are most useful when arranged in customary citation order.[14] Shoham and Lazinger describe advantages of clustering all authority controlled headings, including the uniform title and the main entry, together at the end of the record.[15] The function and future of the main entry heading in an automated catalog is, of course, extensively discussed by Gorman, Patrick Wilson, Carpenter, Boll,

Shoham and Lazinger, Bell-Gam, Madison, and others.[16]

Carpenter and Gorman present separate proposals for removing authority controlled headings from bib record displays but keeping them accessible through modular structures and linking capabilities.[17] Tillett also advocates linking bibliographic and authority records, but she suggests that displaying a controlled heading immediately following the title-page name it refers to will make the linkage clear to the user.[18] Duke, Boll, and Tillett note the role of keyword searching in breaking down *AACR2*'s distinction between descriptive data and authority controlled "access points."[19] Madison asserts that an increased emphasis on analytic access to materials "poses questions regarding the relationships between access and description," citing the recent debate over the retention of the 440 field (one of the few that serves both functions) in the MARC formats.[20]

The role of ISBD—with its prescribed punctuation, spacing, and display order—as a "bibliographic lingua franca" is articulated by Gorman, Schmierer, Hagler, Bryant, Süle, and Tillett.[21] Roberts and Bourne, on the other hand, note the decline of ISBD's usefulness in an era of simplification, both of the cataloging process and of its product (i.e., labeled screen displays).[22] Berg and Hasund Langballe describe efforts in the Scandinavian countries to reform ISBD and the MARC formats by introducing the principle that no piece of information (e.g., an author's name) should be recorded more than once in a record, even in different forms.[23] Rosenberg and Borgman demonstrate what might well be the future beyond ISBD, in a prototype that presents bibliographic information in a screen display formatted to look like a title page.[24]

For the foreseeable future, Rowley and Boll each consider the need for a new cataloging code that takes advances in OPAC technology into account; labeled screen displays appear to play little role in their deliberations.[25] Boll, however, recognizes the need for a code that not only takes into account the burgeoning capabilities of automated catalogs, but exercises a measure of control over their use in catalog design. Coral describes the inadequacy of index displays for music materials as stemming from insufficient control by catalogers over online displays of their work.[26] Also, a bit of serendipity: in an article showing that the main entry concept is firmly embedded in *AACR2*, Baughman and Svenonius demonstrate the dangers of "term mapping"—the automatic substitution of one term for another with a somewhat different meaning.[27] The implications for labeling of bibliographic data, though unstated (and probably unintended), are nonetheless palpable.

There is, in addition, a small but growing body of evidence that reformatting issues related to bibliographic description are beginning to be addressed. Hagler, in recent works, notes the separation of output from input standards, brought on by online display technology, and the need for their reintegration.[28] The 1992 NOTIS Online Users' Group Meeting included a session entitled "Issues Related to Serials Cataloging on NOTIS," which featured presentations by Wilson on the process and impact of machine reformatting at Vanderbilt University, and by Gago on the relationship of locally reformatted records to national standards.[29] During the second half of 1992, occasional comments from Randall, Clarke, McMillen, Ercelawn, and others on the electronic bulletin board AUTOCAT expressed concerns about, or reported instances of, "changing how we record cataloging data to suit the current version of our online system."[30] Finally, a task force of the International Federation of Library Associations' Standing Committee on Cataloguing has launched a comprehensive study of the nature and function of bibliographic records in a context of multiple formats and applications.[31]

One last item of note is a 1990 paper by Wool that outlines several issues facing anyone making bibliographic display decisions for an OPAC.[32] The present study was undertaken as a follow-up to that paper in an attempt to gather data bearing on the validity of the concerns expressed there. By comparing a representative sample of ISBD card-catalog records with their reformatted, labeled versions in the online catalog of a medium-sized research library, the authors sought to determine the nature and extent of changes to those bibliographic records and to assess the relationship of such changes with existing cataloging standards.

CONTEXT OF STUDY

IRIS (Innovative Research Information System) is the online library catalog at the University of Nebraska-Lincoln (UNL). It operates on INNOPAC software from Innovative Interfaces, Inc. At the time that the university's records were printed out for study (late December 1991 and early January 1992) the system was running on Release 6. IRIS serves a campus library system (the University Libraries) that consists of a main library and nine branches with holdings of nearly 1 million titles, or just over 2 million volumes, plus the Law Library with holdings of about 37,000 titles, or 281,000 volumes. (Records created after March 1990 for Law Library materials do not appear in the University Libraries' card catalog shelf list, and thus are not included in this study.)

The INNOPAC system offers two labeled display formats: one for the public catalog, the other available to staff with a password. The public display format was chosen for this study. The screen displays were designed, within INNOPAC parameters, by a task force composed of volunteers from the library staff. No professional catalogers were on this task force, but the task force actively sought input from all units in the library system, and several concerns brought up by the serials catalogers were addressed in the final product.

The task force was limited to twenty-five field labels at that time. Twenty-three were used in the staff display and eighteen in the public display. (Appendixes A and B contain lists of the field labels, their descriptions, and the MARC tags to which they refer.) The task force decided to display all searchable fields in the public display. (See appendix C for information on fields indexed in IRIS.) This decision ensured that the patron can verify why the bibliographic record was retrieved, but a duplication of information in the record can result. (For instance, the 490 field—a nontraced series statement—is keyword searchable and thus displayed, but it often duplicates information in the 830 series tracing, which is title indexed and thus also displayed.) Since the UNL catalog came up on IRIS, it has become possible to base the labels on MARC field indicators, and the limitation to twenty-five field labels has been lifted. Accordingly, several labeling changes have been considered for IRIS, and a few of them have been implemented.

Other display decisions for IRIS included:

1. TITLE is the first label displayed; AUTHOR is second.
2. Uniform titles (MARC fields 130 or 240) display along with title statements (field 245) beside the label TITLE. Since INNOPAC displays fields mapped to a single label in field number order, in records where a uniform title is present, it is the first field displayed.
3. The entire 245 field (including the subtitle and statement of responsibility) displays beside the label TITLE.
4. The author main entry (if any) and all author tracings display together beside the label AUTHOR.
5. Series statements (440 or 490) and series tracings (440 or 8xx fields) display together beside the SERIES label but after the note fields (if any).
6. ISBNs and ISSNs are suppressed from the public display.
7. Title tracings display at the end of the record beside the label OTHER TI; they are immediately preceded by the subject headings (SUBJECT).

Figures 1–3 show a record from the sample in OCLC/MARC format, in ISBD/*AACR2* card display format, and in the IRIS public display format.

METHODOLOGY

On December 12, 1990, the authors identified a sample of 1,017 records from the shelflist. This represented the number of drawers in the LC classification and the sound-recordings file. The first card showing ISBD punctuation was pulled from each drawer. Where a drawer appeared to have no ISBD records, the last ISBD card in a nearby drawer was pulled instead. This method, admittedly, was not without problems, since it sought a sample representative of the entire shelf list rather than of the subset of ISBD records, but it took much less time to pull and photocopy than a more precise sample would have.

Originally, the authors intended to study the online records at the terminal, but that proved impractical, so around the end of 1991 (December 27, 1991, to January 7, 1992), the

```
OCLC:  7276463          Rec stat:  c
Entered:  19810305      Replaced:  19900210      Used:  19920508
Type: a          Bib lvl: m      Source:          Lang:  eng
Repr:            Enc lvl:        Conf pub: 0      Ctry:  nyu
Indx: 1          Mod rec:        Govt pub:        Cont:  b
Desc: a          Int lvl:        Festschr: 0      Illus: c
                 F/B:     0a     Dat tp:   s      Dates: 1981,
  1  010      81-2087//r90
  2  040      DLC ¦c DLC ¦d LDL
  3  020      0374185042 : ¦c $30.00
  4  050 00   PS3503.Y45 ¦b Z48 1981
  5  082 00   811/.52 ¦a B ¦2 19
  6  049      LDLL
  7  100 1    Bynner, Witter, ¦d 1881-1968.
  8  240 10   Correspondence. ¦k Selections
  9  245 10   Selected letters / ¦c edited, and with an
introduction, by James Kraft.
 10  260      New York : ¦b Farrar, Straus, Giroux, ¦c c1981.
 11  300      xxiv, 275 p. : ¦b port. ; ¦c 22 cm.
 12  490 1    The Works of Witter Bynner
 13  504      Includes bibliographical references and index.
 14  600 10   Bynner, Witter, ¦d 1881-1968 ¦x Correspondence.
 15  650  0   Poets, American ¦y 20th century ¦x Correspondence.
 16  700 10   Kraft, James.
 17  800 1    Bynner, Witter, ¦d 1881-1968. ¦t Works. ¦f 1978.
```

Figure 1. Record for Witter Bynner's Selected Letters, *OCLC/MARC Format (Width Reduced by Twelve Characters for WordPerfect).*

online versions of the sampled ISBD records were printed out and photocopied. Roughly 10 percent of the sample records were analyzed, and a list of ways in which the card and online versions differ from each other was compiled. A code was assigned to each category of difference. Following this, the sample was divided into thirds, and three of the authors each took a portion, coding each of its records for the kinds of difference found. The fourth author separately analyzed and coded the entire sample. Finally, the authors met to examine differences in coding (many of which resulted from problems in defining certain categories of difference) and to reach agreement on data definitions.

A small number of records were found to have been recataloged by the time the online records were printed out, and on closer examination a handful were found not to be ISBD records. These records were deleted from the study, leaving a sample of 1,005 records. Of these, sixty-three are for serials (including a CD-ROM), twelve are for scores or sheet music, ten are for sound recordings, and one is for a set of slides. The rest are for monographs (broadly defined, including sets and loose-leaf services).

FINDINGS

As mentioned above, a list of differences attributable to the software was developed from examining a subset of the sample. This included nineteen types of difference found in all record formats, ten types of difference found (or expected to be found) only in records for serials, and three types found in records for music scores or sound recordings. During the coding process, nine additional types of difference were found, affecting a small number of records. These categories of difference are listed in tables 1 through 3, along with the number and percentage of records affected by each. A more detailed summary follows here.

Two titles (codes 1a and 1b). Nearly 7 percent of the online records studied show this anomaly, which is exacerbated by the precedence given the uniform title. On nineteen of the sixty-nine records affected, the uniform title has been suppressed from card display through the use of a field indicator (a one-digit code conveying display instructions or other information for reformatting) in the MARC format. The INNOPAC software, however, ignores all field indicators. Chiefly affected

```
LOVE      Bynner, Witter, 1881-1968.
PS          [Correspondence. Selections]
3503        Selected letters / edited, and with
Y45Z48    an introduction, by James Kraft. -- New
1981      York : Farrar, Straus, Giroux, c1981.
            xxiv, 275 p. : port. ; 22 cm. -- (The
          Works of Witter Bynner)
            Includes bibliographical references
          and index.
            ISBN 0-374-18504-2 : $30.00

            1. Bynner, Witter, 1881-1968--
          Correspondence.  2. Poets, American--
          20th century--Correspondence.
          I. Kraft, James.  II. Title
          III. Series: Bynner, Witter, 1881-1968.
          Works. 1978.
```

Figure 2. Record for Witter Bynner's Selected Letters, *ISBD/Card-Catalog Format.*

```
TITLE          Correspondence. Selections.
               Selected letters / edited, and with an introduction,
                 by James Kraft.
AUTHOR         Bynner, Witter, 1881-1968.
               Kraft, James.
PUBLISHER      New York : Farrar, Straus, Giroux, c1981.
DESCRIPT.      xxiv, 275 p. : port. ; 22 cm.
NOTE           Includes bibliographical references and index.
SERIES         The Works of Witter Bynner.
               Bynner, Witter, 1881-1968. Works. 1978.
SUBJECT        Bynner, Witter, 1881-1968 --Correspondence.
               Poets, American --20th century --Correspondence.
    LOCATION          CALL #                                STATUS
1 > LOVE              PS3503 Y45Z48 1981                    AVAILABLE
```

Figure 3. Record for Witter Bynner's Selected Letters, *OPAC Display Format (IRIS, University of Nebraska-Lincoln, Jan. 1992) (Width Reduced by Seven Characters for WordPerfect; Second TITLE Field Affected).*

are records for music materials, religious scriptures, literary classics, and serials.

Statement of responsibility appears beside TITLE label (code 2). The decision to display the statement of responsibility (MARC field 245, subfield c) can be justified in that it preserves the integrity of the ISBD title statement and maintains access to information about the nature of responsibility for a work. (Ninety percent of the records in the study were affected.) However, this design also guarantees a degree of redundancy when tracings are brought into the description. At present, INNOPAC lacks the capability to split off subfields for separate labeling.

Author headings following title statement (codes 3–8). These changes involve bringing authority controlled author headings into the body of the description. Ninety-seven percent of the sampled records contain author headings and are thus affected by the basic "order-of-display" decision (code 3). Seventy-two percent have an "author main entry" that was effectively demoted (code 4). Just over 50 percent have name-added entries that also display online between the title and publisher statements (code 6).

Several anomalies result from this treatment of author headings. The authors of this article or study determined that in 55 percent of the records, the 245 statement of responsibility is effectively repeated in its entirety by the AUTHOR label and one or more of the headings beside it (code 5). (Differences in the form of the name were counted as immaterial, except in the handful of instances in which the connection between a title page name and a heading was unrecognizable;

Table 1. Categories of Difference Defined before Coding of Records (All Types of Materials)

Code	Description	No. of Records	% of Total (N=1,005)
1a	Two titles in TITLE (one is uniform title appearing on card)	49	4.88
1b	Two titles in TITLE (one is nonprinting uniform title)	20	1.99
2	TITLE includes statement of responsibility	909	90.45
3	Author headings follow title statement	978	97.31
4	Author main entry "demoted"	724	72.04
5	AUTHOR area repeats statement of responsibility in TITLE	545	54.23
6	AUTHOR includes name-added entries	508	50.55
7	AUTHOR includes nonauthors	389	38.71
8	AUTHOR includes author-title tracings	25	2.49
9	Contents note precedes other notes	40	3.98
10	Order of NOTE notes different than on the card record	23	2.29
11	A note repeats or clarifies AUTHOR information	98	9.75
12	Series information follows notes	323	32.14
13	Nontraced series statement and series tracing appear together beside SERIES label	39	3.88
14	Series tracing repeats series statement in SERIES (identical wording)	11	1.09
15	A note repeats or clarifies OTHER TI information	54	5.37
16	ISBN/ISSN on card but not online	549	54.63
17a	More cards in record than screens	17	1.69
17b	More screens in record than cards	77	7.66

however, any phraseology indicating something other than "just plain authorship" was counted as not repeated by the AUTHOR label.) In 38 percent of the records, one or more headings displayed beside the AUTHOR label name a person or entity falling outside the study's working definition of *author*—essentially, a person or entity named in the card record as the actual creator of a literary, musical, or artistic work (code 7). Twenty-five records display work citations (author-title headings) beside the AUTHOR label (code 8).

Order of notes (codes 9 and 10). The only consistent change in note order involves the contents note (field 505), which is given its own label and placed ahead of notes bearing the NOTE label (code 9—4 percent of records affected). Local notes (field 590) often display "out of order" online, but they are not included in this study. All other notes display in the same order online as on cards, except in records for sound recordings, and—somewhat mysteriously—a handful of monograph and serial records. Note fields mapped to the label NOTE display in the order in which they appear on the MARC record, which in these few cases does not correspond to how they display on cards (code 10).

Notes justifying tracings (codes 11 and 15). In a small but not insignificant percentage of the records, a note justifies an author or title tracing (or in a few cases, the main entry). In IRIS, such notes have the effect of restating (often in more detail) information provided by the AUTHOR or OTHER TI label.

Series information (codes 12–14). The decision to display series information following the notes affects nearly a third of the records in the sample (code 12). In about 4 percent of the records, a nontraced from-the-piece series statement (field 490) and the corresponding series heading (field 830 or 810) display together, creating an apparent redundancy (code 13). In about a fourth of these records (1.09 percent of sample) the

Table 2. Categories of Difference Defined before Coding of Records (Serials and Music Materials)

Code	Description	No. of Records	% of Total (*N*=1,005)
s1	BEGAN WITH includes last-issue as well as first-issue information	17	1.69
s2	BEGAN WITH includes last-issue information only	0	0.00
s3	CONTINUES label substitutes for "Absorbed:," "Formed by the merger of:," or "Separated from:" at beginning of note	4	0.40
s4	CONT. BY label substitues for "Absorbed by:," "Merged with:," or ". . . to form:" in note	3	0.30
s5	CONTINUES or CONT. BY precedes other notes	25	2.49
s6	Note in NOTE area repeats CONTINUES/CONT. BY information	0	0.00
s7	Note in NOTE area clarifies CONTINUES/CONT. BY information	0	0.00
s8	OTHER TI includes "key title"	27	2.69
s9	Key title in OTHER TI identical to title in TITLE	19	1.89
s10	Frequency data omitted online	18	1.79
m1	Nonprinting relator codes for names appear online	7	0.70
m2	DESCRIPT. includes duration code	6	0.60
m3	MUSIC NO label for publisher's number; precedes PUBLISHER	19	1.89

Table 3. Categories of Difference Defined during Coding of Records

Code	Description	No. of Records	% of Total (*N*=1,005)
n1	Only a date appears beside PUBLISHER label	16	1.59
n2	OTHER TI includes title of a related document	7	0.70
n3	A note repeats or clarifies the uniform title	21	2.09
n4	Variant title notes for serials omitted online	11	1.09
n5	Summary notes online do not begin with "Summary:"	18	1.79
n6	Serials indexing information omitted online	4	0.40
n7	DESCRIPT. includes score note	2	0.20
n8	Nonprinting place-name heading appears online	1	0.10
n9	SuDoc number appears twice online	2	0.20

wording of the nontraced statement and the heading are identical (code 14).

Omission of standard numbers (code 16). As noted above, IRIS suppresses ISBNs and ISSNs from public display. Fifty-five percent of records in this sample are affected.

Number of display surfaces (codes 17a and 17b). Because of concerns expressed in the literature about the impact of multiple screen displays on record use, records displaying on more or fewer screens than cards were counted. Records with more screens than cards (code 17b) outnumber those with more cards than screens (code 17a) by seventy-seven to seventeen. However, these together account for less than 10 percent of the sample.

Serials (codes s1-s10, n4, n6). In the preparation of the initial list of categories, ten types of changes were identified that could be expected to appear in records for serials. Two involved the mapping of field 362 (numeric/chronological designation) to the label BEGAN WITH. In about a fourth of the serial records, the last-issue as well as the first-issue

designation appears beside BEGAN WITH (code s1). However, no records in the sample have last-issue information only beside this label, as might be expected in records lacking first-issue information (code s2).

Of the five expected change types involving fields 780 (earlier title) and 785 (later title), only one—they display ahead of notes labeled NOTE (code s5)—is found in a large number of records (twenty-five of sixty-three). The labels these fields are mapped to are, respectively, CONTINUES and CONT. BY; however, relationships other than straight title changes are also expressed in these fields (such as mergers, splits, and absorptions), and such relationships would thus be incorrectly labeled. In the sample, this occurs in only four records for CONTINUES (code s3) and in three records for CONT. BY (code s4); that could, however, represent a significant fraction of the serials records. Categories s6 and s7 were set up to count records in which both a 580 note field and a 780 or 785 containing essentially the same information are present. (In such records, the 780 or 785 would display online but not on the cards.) No records with this type of anomaly are found in the sample.

Nearly half the serials records (twenty-seven of sixty-three) contain a key title assigned by the International Serials Data System. On cards, it appears alongside the ISSN, and online it displays beside the label OTHER TI (code s8). On most of these records, this assigned title is identical to the "real" title (code s9). On eighteen records, data on the frequency of publication that appear on the cards are omitted online (code s10). However, this does not occur consistently within IRIS; other records retain their frequency information in the new format.

Two other types of changes were found during coding. Notes generated by field 246 (variant title) on cards do not appear online; this field also generates title tracings that display beside the OTHER TI label (code n4). This is one situation in which a decision was made to avoid partial repetition of information; eleven records were affected. Only four records contain indexing information (field 510) that is suppressed online (code n6). However, the field in question has not been used in serials cataloging at UNL for some years now.

Scores and sound recordings (codes m1-m3, n7). On significant numbers of records for these materials (admittedly from a very small sample), "relator codes" in name headings, indicating type of responsibility, are suppressed from card display but appear online (code m1); duration codes (field 306) display online but not on cards (code m2); and the publisher's number, which displays on cards in a note, appears online from field 028, while the note is suppressed (code m3). On two records, the note "Full score" appears online beside the DESCRIPT. label (code n7).

Miscellaneous changes (codes n1-n3, n5, n8, n9). A variety of other changes were found in the sample that had not been anticipated. On a number of records for theses, no publisher or place of publication had been recorded, so that online only a date displays beside the label PUBLISHER (code n1). In seven records, a title of a related work displays beside the label OTHER TI (code n2). About 2 percent of the sample contains a note justifying a uniform title (code n3). Eighteen records, mostly for children's and young adults' literature, contain a summary note (field 520) that, on the cards, begins with the OCLC print constant "Summary:"; this print constant is omitted in the online records, and no special label has been supplied to take its place (code n5). The other two changes, described in table 3, are barely perceptible in the sample.

The definition of some of these categories (especially 5 and 7) proved very tricky, and the initial coding produced a number of disagreements that had to be negotiated away, resulting in more closely defined categories. Larger issues raised in this process include:

1. What constitutes authorship? This had to be resolved before determining how many records include name headings that are inappropriately labeled AUTHOR (code 7). Conceivably, one could count editors, illustrators of children's books, performers in sound recordings, and authors of introductions as authors. After much debate, we chose to include none of these and in fact defined *nonauthor* rather broadly. (Likewise, prominently named corporate bodies—as defined by *AACR2*—are considered nonauthors unless named in the statement of responsibility.)[33] Undoubtedly, a narrower definition would produce a much lower count in this category.

2. Is the "item" or the "work" being cataloged? This age-old cataloging issue especially impinges on category 7. Again we

leaned, in our consensus decisions, toward cataloging the work, meaning that translators, annotators, musical performers, and so on, are counted as nonauthors.

3. What constitutes unnecessary repetition? This question speaks to the issue of clutter in online records and is relevant to categories 5, 11, 13, and 15 in particular. For example, deciding whether to code "5" on a record meant answering the question, are "AUTHOR Smith, John D." and "reported by John D. Smith" equivalent statements? For this question, as it happens, we decided the answer is no (although "by John David Smith" or "written by John Smith" are considered equivalent to "AUTHOR Smith, John D."), so we coded 5 only when the statement of responsibility is a simple statement of authorship, not adding information found nowhere else on the record. For other categories, however, we counted all instances of repetition and partial repetition without judging whether they are in fact unnecessary.

ANALYSIS

To get a clearer picture of what machine reformatting has produced, the various categories can be grouped under the following broader headings:

- Change in order of presentation
  (codes 3, 4, 6, 9, 10, 12, 13, s5, m3)
  986 records (98.11 percent) affected
- Questionable accuracy of labeling
  (codes 1a, 1b, 2, 7, 8, s1, s2, s3, s4, s8, s9, m2, m3, n1, n2)
  985 records (98.01 percent) affected
- Repetition or partial repetition of data
  (codes 2, 5, 11, 13, 14, 15, s6, s7, s9, n3, n9)
  959 records (95.42 percent) affected
- Addition or omission of data
  (codes 1b, 16, s10, m1, m2, n4, n5, n6, n7, n8, n9)
  591 records (58.81 percent) affected
- More or fewer cards than screens
  (codes 17a and 17b)
  94 records (9.35 percent) affected

Tables 4 through 8 present a more detailed breakdown. It should be noted that a few categories were placed into two of these broader groups.

In addition, it should be obvious that the kinds of difference found are not all of equal significance. Changes likely to produce misinformation, confusion for the user, or apparent clutter are of much more concern than those that involve rearrangement of data items with no further effect, simplistic but not necessarily misleading labeling (e.g., the title statement as TITLE), widely separated redundancy of information, and planned addition or deletion of data. By this standard, the categories of difference can be classified as "major" and "minor" as follows:

- Major: codes 1a, 1b, 5, 7, 8, 14, s1, s2, s3, s4, s6, s7, s9, s10, m2, n1, n2, n4, n9
- Minor: codes 2, 3, 4, 6, 9, 10, 11, 12, 13, 15, 16, 17a, 17b, s5, s8, m1, m3, n3, n5, n6, n7, n8

Tables 9 and 10 depict these groups more fully. On this basis, analysis of the data shows that 920 records (91.54 percent of the sample) contain at least one major anomaly produced by machine translation.

DISCUSSION

In an assessment of the significance of the findings, it is well to keep two things in mind: (1) the study is of a single library catalog as it reflects a particular system's software and the local display decisions in place at the beginning of 1992; (2) while the sample of records studied is very large and substantially representative of the variety of ISBD records created at UNL from around 1980 through the fall of 1990 (within eight months of the catalog's going online), the sampling technique used may nonetheless preclude authoritative generalization. Even so, these findings raise a number of issues relevant to all online catalogs with labeled bibliographic displays.

The authors identified thirty-eight types of changes in the sampled records after reformatting for online display. Most of these changes are purely the intended results of display decisions, including nearly all changes in order of data display, as well as most additions and omissions of data. The principal side effect here is the commingling of authority-controlled tracings with elements of the description, which will be discussed further below. The number of screens versus the number of cards is an issue of some interest, due to indications that many catalog users seldom look beyond the first screen of an online record; but more than 90 percent of the sampled records had the same number of

Table 4. Categories of Difference Involving Changes in the Sequence of Data Presentation

Code	Description	No. of Records	% of Total (*N*=1,005)
3	Author headings follow title statement	978	97.31
4	Author main entry "demoted"	724	72.04
6	AUTHOR includes name added entries	508	50.55
12	Series information follows notes	323	32.14
9	Contents note precedes other notes	40	3.98
13	Nontraced series statement and series tracing display together beside SERIES label	39	3.88
s5	CONTINUES or CONT. BY precedes other notes	25	2.49
10	Order of NOTE notes different than on the card record	23	2.29
m3	MUSIC NO label for publisher's number; precedes PUBLISHER	19	1.89
Any of above	Change in order of presentation	986	98.11

screens as cards, indicating that the display format, at least in IRIS, has little if any significant impact on display length.

Changes involving less-than-accurate labeling, however, appear to constitute a form of unintended noise resulting from machine translation. The most extensive of these anomalies is the display of the statement of responsibility along with the title beside the TITLE label. To what extent this violates truth in labeling can, of course, be debated. If the label is seen as a mere pointer indicating the general nature of the information, concerns about its descriptive adequacy can seem overly fastidious. To those familiar with MARC and *AACR2*, what displays beside this label is a close-knit title statement, and all of it belongs. But most catalog users lack this perspective. To them, the label TITLE says "this is the title," and the inclusion of authorship information probably seems strange at best.

Much more troublesome, though, is the use in this catalog of the label AUTHOR. The mapping of added-entry fields 700, 710, and 711 to this label along with the main-entry fields 100, 110, and 111 produces some bizarre (but predictable) side effects, especially the designation of Festschrift honorees as "authors" and the presence of author-title citations. This is the result of taking a category of bibliographic data meant to provide access to a broad variety of names connected with a work and redefining it as a very specific type of descriptive information—but without changing what kind of information goes into it. The problems with treating named meetings as authors (a hallmark of *AACR2*) also come into sharp focus here. Whether editors, performers on sound recordings, illustrators, and so on, are legitimately seen as authors is probably a matter of personal opinion, but if they are denoted as such, they should be recorded as such intentionally (i.e., not simply as a desired access point) and consistently.

The findings also indicate that some form of data repetition is to be found in almost all records in the OPAC studied. Some readers may argue here that this constitutes no change, that the same repetitiveness exists in the card records as well and has simply been rearranged. They have a point; in traditional, card-based cataloging practice, name and title headings are usually justified in the body of the description, and since the advent of the unit record, these tracings have been printed on all card entries following the description.[34]

In the IRIS public display, however, the tracings are, for all practical purposes, part of the description. They are no longer separated and isolated in a kind of appendix, where they can be safely ignored by most catalog users, but fully integrated with the notes, the collation, and the title-page transcription in what might be called an enriched record. Even in such a record, though, the juxtaposition of a

Table 5. Categories of Difference Involving Questionable Accuracy of Labeling

Code	Description	No. of Records	% of Total (*N*=1,005)
2	TITLE includes statement of responsibility	909	90.45
7	AUTHOR includes nonauthors	389	38.71
1a	Two titles in TITLE (one is uniform title appearing on card)	49	4.88
s8	OTHER TI includes "key title"	27	2.69
8	AUTHOR includes author-title tracings	25	2.49
1b	Two titles in TITLE (one is nonprinting uniform title)	20	1.99
s9	Key title in OTHER TI identical to title in TITLE	19	1.89
m3	MUSIC NO label for publisher's number; precedes PUBLISHER	19	1.89
s1	BEGAN WITH includes last-issue as well as first-issue information	17	1.69
n1	Only a date appears beside PUBLISHER label	16	1.59
n2	OTHER TI includes title of a related document	7	0.70
m2	DESCRIPT. includes duration code	6	0.60
s3	CONTINUES label substitutes for "Absorbed:," "Formed by the merger of:," or "Separated from:" at beginning of note	4	0.40
s4	CONT. BY label substitues for "Absorbed by:," "Merged with:," or ". . . to form:" in note	3	0.30
s2	BEGAN WITH includes last-issue information only	0	0.00
Any of above	Questionable accuracy of labeling	985	98.01

Table 6. Categories of Difference Involving Repetition or Partial Repetition of Data

Code	Description	No. of Records	% of Total (*N*=1,005)
2	TITLE includes statement of responsibility	909	90.45
5	AUTHOR area repeats statement of responsibility in TITLE	545	54.23
11	A note repeats or clarifies AUTHOR information	98	9.75
15	A note repeats or clarifies OTHER TI information	54	5.37
13	Nontraced series statement and series tracing display together beside SERIES label	39	3.88
n3	A note repeats or clarifies the uniform title	21	2.09
s9	Key title in OTHER TI identical to title in TITLE	19	1.89
14	Series tracing repeats series statement in SERIES (identical wording)	11	1.09
n9	SuDoc number appears twice in online record	2	0.20
s6	Note in NOTE area repeats CONTINUES/CONT. BY information	0	0.00
s7	Note in NOTE area clarifies CONTINUES/CONT. BY information	0	0.00
Any of above	Repetition or partial repetition of data	959	95.42

Table 7. Categories of Difference Involving Addition or Omission of Data

Code	Description	No. of Records	% of Total (*N*=1,005)
16	ISBN/ISSN on card but not online	549	54.63
1b	Two titles in TITLE (one is nonprinting uniform title)	20	1.99
s10	Frequency data omitted online	18	1.79
n5	Summary notes online do not begin with "Summary:"	18	1.79
n4	Variant title notes for serials omitted online	11	1.09
m1	Nonprinting relator codes for names appear online	7	0.70
m2	DESCRIPT. includes duration code	6	0.60
n6	Serials indexing information omitted online	4	0.40
n7	DESCRIPT. includes score note	2	0.20
n9	SuDoc number appears twice online	2	0.20
n8	Nonprinting place-name heading appears online	1	0.10
Any of above	Addition or omission of data	591	58.81

Table 8. Categories of Difference Involving More or Fewer Cards Than Screens

Code	Description	No. of Records	% of Total (*N*=1,005)
17b	More screens in record than cards	77	7.77
17a	More cards in record than screens	17	1.69
Any of above	More or fewer cards than screens	94	9.35

from-the-title-page statement of authorship with the label AUTHOR followed by the person's name heading can seem unnecessarily redundant. The Scandinavian ISBD-reform efforts mentioned earlier attest to this, as does the suppression of the 245 "statement of responsibility" in the default version of NOTIS 5.0's public display.

Under *AACR2* the descriptive data and the authority-controlled headings serve different functions (namely, description and access) and are created under separate sets of rules. As the headings take on the descriptive function as well, a competition ensues in which the headings, being better suited to indexing and searching online, have the advantage. One might well foresee that under growing pressures for a less-cluttered record and for simplification of cataloging, much from-the-piece descriptive data will first be suppressed from public display and not too much later cease to be recorded altogether. If this happens on a local-option basis, however, the international exchangeability of records made possible by ISBD will be progressively eroded.

In recent years, the recording and the display of bibliographic information have become separated by technology. Data recorded under the internationally accepted standards for one format are, in more and more libraries, being displayed in a different, locally determined format with additional information (labels) that reinterpret the data and can change their meaning. As authority-controlled name and series headings become descriptive as well as collocative in function, the usefulness of free-text descriptive data comes under question. It is hard to believe that the recording of bibliographic data (i.e., descriptive cataloging) will not change soon to accommodate the new display capabilities. It is in the interest of shared cataloging that this change be managed through a new code that will bring recording and display standards together again.

Table. 9 Categories of Difference Involving Major Anomalies

Code	Description	No. of Records	% of Total (*N*=1,005)
5	AUTHOR area repeats statement of responsibility in TITLE	545	54.23
7	AUTHOR includes nonauthors	389	38.71
1a	Two titles in TITLE (one is uniform title appearing on card)	49	4.88
8	AUTHOR includes author-title tracings	25	2.49
1b	Two titles in TITLE (one is nonprinting uniform title)	20	1.99
s9	Key title in OTHER TI identical to title in TITLE	19	1.89
s10	Frequency data omitted online	18	1.79
s1	BEGAN WITH includes last-issue as well as first-issue information	17	1.69
n1	Only a date appears beside PUBLISHER label	16	1.59
14	Series tracing repeats series statement word for word in SERIES	11	1.09
n4	Variant title notes for serials omitted online	11	1.09
n2	OTHER TI includes title of a related document	7	0.70
m2	DESCRIPT. includes duration code	6	0.60
s3	CONTINUES label misleading	4	0.40
s4	CONT. BY label misleading	3	0.30
n9	SuDoc number appears twice online	2	0.20
s2	BEGAN WITH includes last-issue information only	0	0.00
s6	Note in NOTE area repeats CONTINUES/CONT. BY information	0	0.00
s7	Note in NOTE area clarifies CONTINUES/CONT. BY information	0	0.00
Any of above	Major anomalies	920	91.54

RECOMMENDATION

As more and more libraries rearrange and label their bibliographic records for online display, they are likely to face problems similar to those found in this study. Such problems can be ameliorated through labeling that is more careful and precise, avoiding the clustering of several fields under a single label, and keeping descriptive and collocative data apart. System features that could help include a generous allowance of defined labels, the capability to label subfields separately, and a display header showing not only the search term entered, but the data item within the record matching that search term.

Unfortunately, such measures, which would serve to mimic existing *AACR2* data structures, can easily result in longer, more complicated, and jargon-riddled displays—something no one wants in an increasingly user-oriented service environment. At the same time, suppressing fields identified as "redundant" can backfire, in that those fields often contain, as well, information not found elsewhere. But as records continue to be created according to card-display principles, the number of records that display awkwardly or confusingly online can only increase.

It is time, therefore, to begin working toward a new cataloging code. Whatever form it takes, whatever old or new practices it sanctions, this code should differ from *AACR2* in at least one important respect: it should result in records flexible enough to be added to, subtracted from, and rearranged without loss or garbling of meaning. What is needed is a modular record structure, in which every

Table 10. Categories of Difference Involving Minor Anomalies.

Code	Description	No. of Records	% of Total (*N*=1,005)
3	AUTHOR after TITLE	978	97.31
2	TITLE includes statement of responsibility	909	90.45
4	Author main entry "demoted"	724	72.04
16	ISBN/ISSN on card but not online	549	54.63
6	AUTHOR includes name-added entries	508	50.55
12	Series informatin follows notes	323	32.14
11	A note repeats or clarifies AUTHOR information	98	9.75
17b	More screens in record than cards	77	7.66
15	A note repeats or clarifies OTHER TI information	54	5.37
9	Contents note precedes other notes	40	3.98
13	SERIES includes both statement and tracing	39	3.88
s8	OTHER TI includes "key title"	27	2.69
s5	CONTINUES or CONT. BY precedes other notes	25	2.49
10	NOTE notes—different order	23	2.29
n3	A note repeats or clarifies the uniform title	21	2.09
m3	MUSIC NO label precedes PUBLISHER	19	1.89
n5	"Summary:" missing online	18	1.79
17a	More cards in record than screens	17	1.69
m1	Nonprinting relator codes for names appear online	7	0.70
n6	Serials indexing information omitted online	4	0.40
n7	DESCRIPT. includes score note	2	0.20
n8	Place-name heading added online	1	0.10
Any of above	Minor anomalies	999	99.40

segment of data can stand on its own with appropriate labeling and which can support all possible display lengths and combinations of data elements.

A possible foundation for such a structure can already be found in the MARC formats. The new code, then, would have its basis in USMARC (or a successor communications format). It would specify not what data should appear in a particular sequence on three-by-five cards, but what data should go into each MARC field. While bibliographic displays would, of necessity, be deregulated, useful guidelines for labeling and content could have a place in the code. Whoever develops this code would also have to consider online indexing issues and determine to what extent descriptive data should be controlled or standardized. Faithful title page transcriptions could well become a thing of the past—or else raw material for invisible cross-references.

The most efficient way to get this project off the ground might be for the Joint Steering Committee for the Revision of *AACR* to create a task force that would study all relevant issues (including expected future capabilities of information technology) and ultimately produce a draft code for consideration. Such a task force would need to work closely with MARBI (the ALA committee that considers revisions of the USMARC formats) and the committees in IFLA (International Federation of Library Associations) responsible for ISBD. Certainly, significant changes in

USMARC and ISBD would be a necessary part of the process. But a record structure that can take advantage of recent and future advances in information processing would be well worth the effort required.

CONCLUSIONS

The comparison of a sample of ISBD records with their reformatted, labeled versions in an online catalog provides evidence that the machine reformatting of bibliographic records can easily change the content of these records. The nature of such change can include subtly altering the meaning of data through term mapping and labeling; creating clutter by juxtaposing similar, but functionally separate, data items and dissolving the functional distinction; and adding or omitting data. While many anomalies can be regarded as minor, major changes, which compromise either the functionality of the record or the integrity of the data, affect more than 90 percent of the sample.

As a form of machine translation, reformatting constitutes a new element in the cataloging process, a final cut (to draw a moviemaking analogy) that is likely to be beyond the control either of catalogers or of cataloging standards. The importance of the end product, however, practically ensures that sooner or later, considerations of "how this will look online" will influence how catalogers do their work. This argues for a substantial rethinking of the role and nature of cataloging standards, as well as a reexamination of the existing rules. Further studies of the impact of machine reformatting on the cataloging product (and process) can be invaluable in this effort.

REFERENCES AND NOTES

1. Gregory J. Wool, "Not Just a Pretty Face: Redesigning the Bibliographic Record for an Online Catalog," in *Academic Libraries: Reaching Up and Stretching Out: Proceedings from the 1990 Spring Meeting of the Nebraska Library Association, College and University Section*, ed. Joan Giesecke (ERIC ED 332 703), 79–89.
2. Ronald Hagler, *The Bibliographic Record and Information Technology*, 2d ed. (Chicago: American Library Assn., 1991), 27.
3. Ibid., 221.
4. H. H. Neville, "Computers and the Language of Bibliographic Descriptions," *Information Processing & Management* 17 (1981): 137–38.
5. Sergei Nirenburg, "Knowledge and Choices in Machine Translation," in *Machine Translation: Theoretical and Methodological Issues*, ed. Sergei Nirenburg (Cambridge, England: Cambridge Univ. Pr., 1987), 2.
6. Walt Crawford, Lennie Stovel, and Kathleen Bales, *Bibliographic Displays in the Online Catalog* (White Plains, N.Y.: Knowledge Industry, 1986).
7. Joseph R. Matthews, "Suggested Guidelines for Screen Layouts and Design of Online Catalogs," *Library Trends* 35 (Spring 1987): 564.
8. Ibid., 565.
9. Bryce Allen, "Individual Differences, Values and Catalogs," *Technicalities* 11 (July 1991): 8–9.
10. Crawford, Stovel, and Bales, *Bibliographic Displays*, 3–4, 19, 75.
11. Irving E. Stephens, "Getting More Out of Call Numbers: Displaying Holdings, Locations and Circulation Status," *Cataloging & Classification Quarterly* 13 (1991): 101.
12. Ibid., 98.
13. Walt Crawford, "Starting Over: Current Issues in Online Catalog User Interface Design," *Information Technology and Libraries* 11 (Mar. 1992): 71.
14. Michael Carpenter, "Does Cataloging Theory Rest on a Mistake?" in *Origins, Content, and Future of AACR2 Revised*, ed. Richard P. Smiraglia (Chicago: American Library Assn., 1992), 96.
15. Snunith Shoham and Susan S. Lazinger, "The No-Main-Entry Principle and the Automated Catalog," *Cataloging and Classification Quarterly* 12 (1991): 64.
16. Michael Gorman, "*AACR2*: Main Themes," in *The Making of a Code: The Issues Underlying AACR2*, ed. Doris Hargett Clack (Chicago: American Library Assn., 1980), 45–46; Patrick Wilson, "The Catalog as Access Mechanism: Background and Concepts," in *Foundations of Cataloging: A Sourcebook*, ed. Michael Carpenter and Elaine Svenonius (Littleton, Colo.: Libraries Unlimited, 1985), 261–63; Michael Carpenter, "Main Entry," in *The Conceptual Foundations of Descriptive Cataloging*, ed. Elaine Svenonius (San Diego, Calif.: Academic, 1989), 73–95; John J. Boll, "The Future of *AACR2*," *Cataloging and Classification Quarterly* 12 (1990): 14–16; Shoham and Lazinger, "The No-Main-Entry Principle," 51–67; Ruby A. Bell-Gam, "Problems with the Basis for the Title Main Entry of Motion Pictures According to the Anglo-American

Cataloging Rules," *Cataloging and Classification Quarterly* 13 (1991): 90–91; Olivia M. A. Madison, "The Role of the Name Main-Entry Heading in the Online Environment," *The Serials Librarian* 22 (1992): 371–91.

17. Carpenter, "Does Cataloging Theory Rest on a Mistake?" 98–100; Michael Gorman, "After *AACR2*R: The Future of the Anglo-American Cataloging Rules," in *Origins*, 91–93.
18. Barbara B. Tillett, "Bibliographic Structures: The Evolution of Catalog Entries, References, and Tracings," in *Conceptual Foundations*, 161–62.
19. John K. Duke, "Access and Automation: The Catalog Record in the Age of Automation" in *Conceptual Foundations*, 119–21; Boll, "The Future of *AACR2*," 11–12; Tillett, "Bibliographic Structures," 160.
20. Olivia M. A. Madison, "Issues in Bibliographical Access and Control in the Online Information Environment," *Journal of Library Administration* 15 (1991): 102.
21. Michael Gorman, "Bibliographic Description: Past, Present and Future," *International Cataloguing* 16 (Oct./Dec. 1987): 43; Helen F. Schmierer, "The Impact of Technology on Cataloging Rules," in *Conceptual Foundations*, 109, 115; Hagler, *Bibliographic Record*, 228–37; P. Bryant, "What Is That Hyphen Doing, Anyway?" *International Cataloguing and Bibliographic Control* 18 (1989): 27–29; Gisela Süle, "Bibliographic Standards," in *Bibliographic Access in Europe: First International Conference*, ed. Lorcan Dempsey (Brookfield, Vt.: Gower, 1990), 252; Barbara B. Tillett, "Future Cataloging Rules and Catalog Records," in *Origins*, 112. The quoted phrase is from Gorman.
22. Winston Roberts and Ross Bourne, "Challenges to Bibliographic Standards: A View From IFLA," in *Bibliographic Access*, 278.
23. Øivind Berg, "Current Problems with the MARC/ISBD Formats in Relation to Online Public Access of Bibliographic Information," *International Cataloguing and Bibliographic Control* 20 (Jan./Mar. 1991): 12–14; Anne M. Hasund Langballe, "Brief Cataloguing in Card and Online Catalogues," *International Cataloguing and Bibliographic Control* 20 (Apr./June 1991): 31.
24. Jason B. Rosenberg and Christine L. Borgman, "A Report on the Loading of MARC Format Bibliographic Records into HyperCard," *Information Technology and Libraries* 10 (Dec. 1991): 292–97.
25. J. E. Rowley, "Towards *AACR3*: A Review of the Implications of OPACs for Cataloguing Codes and Practices," *Library Review* 38 (1989): 7–18; Boll, "The Future of *AACR2*," 3–34.
26. Lenore Coral, "Indexing and Retrieving Special Materials in Online Catalogues," *International Cataloguing and Bibliographic Control* 21 (Apr./June 1992): 29–31.
27. Betty Baughman and Elaine F. Svenonius, "*AACR2*: Main Entry Free?" *Cataloging and Classification Quarterly* 5 (Fall 1984): 1–15.
28. Ronald Hagler, "The Consequences of Integration," in *Conceptual Foundations*, 212; Hagler, *Bibliographic Record*, 235–36.
29. *1992 NOTIS Users' Group Meeting* (Evanston, Ill.: NOTIS Systems, Inc., 1992), 486–98, contains outlines and illustrations only for these presentations. None of the authors of this study attended the meeting.
30. AUTOCAT-L postings from Kevin Randall, Aug. 19, 1992; Sherman Clarke, Sept. 9, 1992; Sophia McMillen, Sept. 11, 1992; Ann Ercelawn, Nov. 6, 1992. The quoted phrase is from Ercelawn.
31. "Terms of Reference for a Study of the Functional Requirements for Bibliographic Records." 1992-09-04. International Federation of Library Associations, Standing Committee on Cataloguing, 1992. Photocopy.
32. Wool, "Not Just a Pretty Face," 75–100.
33. Michael Gorman and Paul W. Winkler, eds., *Anglo-American Cataloguing Rules*, 2d ed., 1988 rev. (Chicago: American Library Assn., 1988), 2.
34. Richard Joseph Hyman, *Information Access: Capabilities and Limitations of Printed and Computerized Sources* (Chicago: American Library Assn., 1989), 51. ■■

APPENDIX A.
IRIS INFORMATION SHEET, 1990
(PUBLIC DISPLAY)

Order of Tags in the Public Display

Label	Description	MARC Tags
TITLE	Title-page title, uniform title	130, 240, 245
AUTHOR	Personal author, corp. author, joint authors	100, 110, 111, 700, 710, 711
EDITION	Edition, revised ed., etc.	250
MUSIC NO	Music publishing no.	028
SUDOC NO	Government documents class no.	086
PUBLISHER	Imprint statement, publisher, place, date	260
DESCRIPT.	Physical descriptions, page count	254, 255, 256, 300, 302, 306
CONTENTS	Specifies all or part of the contents of a work	505
FREQUENCY	How often a serial is published	310, 315
BEGAN WITH	Date and vol. number for a serial, usually indicates the first issue of the title	362
CONTINUES	Preceding title for a serial	780
CONT. BY	Succeeding title for a serial	785
NOTE	Various notes to help describe the work	500–504, 507–508, 511, 513, 515–516, 518, 520–523, 525, 530, 533–534, 536–538, 545–547, 550, 556, 561, 565, 570, 580, 590
SERIES	All forms of series titles	400, 410, 411, 440, 490, 800, 810, 811, 830
SUBJECT	All subject headings	600–651, 690
PRE-1801	Place of printing for rare books	752
FORMER TITLE	Previous title for a serial	247
OTHER TITLE	Varying titles for an item; includes uniform title used as an added entry and other titles from the piece, added-entry title.	222, 246, 730, 740

APPENDIX B.
IRIS INFORMATION SHEET, 1990
(STAFF DISPLAY)

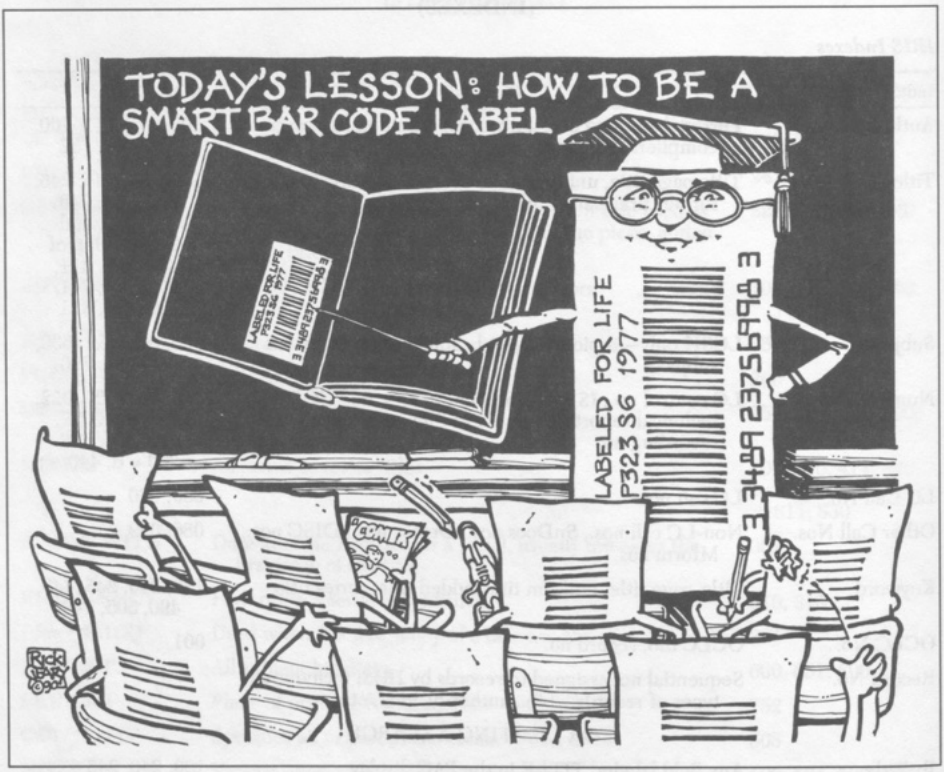
Order of Tags in the Staff Display

Label	Description	MARC Tags
TITLE	Title page title, uniform title	130, 240, 245
FORMER TITLE	Previous title for a serial	247
OTHER TITLE	Varying titles for an item; includes uniform title used as an added entry and other titles from the piece, added-entry title	222, 246, 730, 740
AUTHOR	Personal author, corp. author, joint authors	100, 110, 111, 700, 710, 711
EDITION	Edition, revised ed., etc.	250
PUBLISHER	Imprint statement, publisher, place, date	260
DESCRIPT.	Physical description, page count	254, 255, 256, 300, 302, 306
SERIES	All forms of series titles	400, 410, 411, 440, 490, 800, 810, 811, 830
BEGAN WITH	Date and vol. number for a serial, usually indicates the first issue of the title	362
FREQUENCY	How often a serial is published	310, 315
PREV FREQ	Used when the frequency of a title has changed	321
SUBJECT	All subject headings	600–651, 690
PRE-1801	Place of printing for rare books	752
CONTENTS	Specifies all or part of the contents of a work	505
NOTE	Various notes to help describe the work	500–504, 507–508, 511, 513, 515–516, 518, 520–523, 525, 530, 533–534, 536–538, 545–547, 550, 556, 561, 565, 570, 580, 590
CONTINUES	Preceding title for a serial	780
CONT. BY	Succeeding title for a serial	785
OCLC NO	OCLC number	001
ISBN/ISSN	ISBN and ISSN numbers	020, 022
ALT NO	Various report numbers, film no., recording no.	023, 024, 027
MUSIC NO	Music publishing no.	028
SUDOC NO	Government documents class no.	086
MARC	CODEN no. from *Chem. Abst.*	030
	Vendor stock no.	037
	Language of the work	041
	Library of Congress card no.	010, 011
	OCLC cross ref. no.	019

APPENDIX C.
IRIS INFORMATION SHEET, 1990
(INDEXES)

IRIS Indexes

Index	Description	MARC Tags
Author	Personal author, corp. author, conf. author, joint author, compilers, editors, illustrators, performers	100, 110, 111, 700. 710, 711, 752
Title	Title page title, uniform title, added title, variant title, former title, series title	130, 240, 245, 246, 247, 440, 740, 773, 830, and title portion of 400, 410, 411, 700, 710, 711, 800, 810, 811
Subject	LCSH only—topical subject headings, names used as subjects	600, 610, 611, 630, 650, 651, 690
Numeric	LC control no., ISBN, ISSN, music pub. no., CODEN technical report no., standard film no., standard recording no.	010, 011, 020, 022, 023, 024, 027, 028, 030, 088, 260 = d, 440 = x
LC Call No.	LC call nos.	050, 090
Other Call Nos.	Non-LC call nos., SuDocs nos., Dewey nos., DISC nos., Mform nos.	086, 099
Keyword	Title-page title, uniform title, added title, series title	130, 240, 245, 440, 490, 505, 740
OCLC No.	OCLC bib. record no.	001
Record No.	Sequential no. assigned to records by IRIS; includes all types of records—bib., authority, item, etc.	
	FOR LIMITING A SEARCH	
By Title	Any field labeled TITLE in the PAC display	130, 240, 245
By Author	Any field labeled AUTHOR in the PAC display	100, 110, 111, 700, 710, 711
By Subject	Any field labeled SUBJECT in the PAC display	600, 610, 611, 630, 650, 651, 690
By Publisher	Publisher in the imprint field	260 = b
By Date	Date in the imprint field	260 = c



Why enter bar code numbers into your circulation system one at a time, when Blackwell can automatically assign bar code numbers to your entire MARC database and provide your library with smart bar code labels in shelflist order. Call or write today for more information on Blackwell's smart bar code label service.

Technical Services Division
BLACKWELL
NORTH AMERICA, INC.

Technical Services Division, Blackwell North America, Inc., 6024 S.W. Jean Road, Building G, Lake Oswego, Oregon 97035
Phone Toll Free: 1-800-547-6426 or 503-684-1140 (in Oregon) FAX: 1-503-639-2481 Canadian Toll Free: 1-800-626-1807

An Overview of Applications of Automation to Special Collections: Maps and Archives

Bobs M. Tusa

Automation has traditionally been viewed as an inappropriate means of control of the unique items in special collections libraries because of automation's fundamental requirement of standardization. It has, however, come to be viewed as an excellent means of control of and access to these items by curators who view the special collections library as a system, precisely because of standardization. This study looks at the issue of standardization in the application of computerized automation to some materials found in special collections libraries—specifically, map collections and archives—and at some recent examples of those applications in both North America and Western Europe.

This is the second part of a two-part article presenting an overview of recent applications of automation in special collections libraries. The first part, published in the previous issue of this journal, dealt with rare books and art collections. The second part considers maps and archives. Central to the applications selected is the systems approach brought to bear by international curators on all aspects of the management of their collections.

MAP COLLECTIONS

"If automation has come slowly to libraries, it has come at a snail's pace to the administration of map collections. Libraries, and their directors, continue to be very book-oriented with many special collections, not just maps, suffering because of it."[1] The preceding statement was made in 1988 by David A. Cobb of the Map and Geography Library at the University of Illinois at Urbana, the editor of ALA's *Guide to U.S. Map Resources* (1986). According to Cobb, a 1986 study conducted by ALA's Map and Geography Round Table showed that "of 919 libraries, only 55 of them catalogued their maps. . . . Few libraries—only 26 percent—took advantage of the two large bibliographic utilities in the United States for improved map access and control. Perhaps not surprising then is the fact that only 5 percent of these map collections had an online catalog" (p. 50).

The approach to automating the University of Illinois' map collection is indicative of similar approaches. The university's library system was already committed to OCLC, which Cobb also preferred because it contained more map records than RLIN: 186,165 map titles as of June 1, 1987, of which 104,561 had been contributed by the Library of Congress Geography and Map Division (p. 51).

Cobb received a U.S. Office of Education Title II-C grant to catalog some 525,000 maps and aerial photographs, taking advantage of OCLC's holdings statements permitting listings of individual holdings for small sets. However, the problem of providing automated access to individual sheet maps within large series—the majority of maps in map collections—remained unsolved. Cobb was aware of the automated graphic index to maps in series developed by the University of

Bobs M. Tusa is Head of Special Collections, Manuscripts and Archives, Prescott Memorial Library, Louisiana Tech University. Her e-mail addresses are tusa@latech.bitnet and tusa@vm.cc.latech.edu.

Edinburgh (see below) and made the decision to limit Illinois' holdings statements on OCLC to map series of less than 100 sheets while awaiting further technological developments.

Cobb felt that the benefits of OCLC cataloging of maps were substantial, for example, in the area of access: "Because the . . . system is available statewide, it means a user can query the local system for a particular map, choose the item needed and check it out while never leaving Chicago 135 miles away!" (p. 53). However, he sounded the familiar warning about the need for standardization, specifically in the area of map coordinates: "There continues a strong voice within the cartographic community that coordinate indexing and searching [are] the answer to accurate location of geographic information. . . . The 'guessing' of coordinates within the OCLC library system continues to frustrate users and catalogers alike. . . . Such standardization requires a great deal of cooperation and a central clearinghouse such as the national library" (p. 49).

A similar approach was taken by the French Bibliothèque Nationale's Maps and Plans Department,[2] which took advantage of the library's GEAC software to catalog some million maps, globes, and other items in UNIMARC format and according to ISBD(CM) [Cartographic Materials] standards. Because the department is the national depository for the registration of copyrights for atlases, maps, and plans, it receives some 10,000 items annually, to which access had been provided in the past by means of manual index cards. Before deciding on the required data elements and on adjustments to fields in a computer system originally designed for monographs, the department met with representatives of France's cartographic community, such as the National Geographic Institute and the Mediterranean Geosciences Institute, in order to achieve national cooperation in the creation of standards.

The map records in the resulting online database are accessible directly, via a menu, and by searching with Boolean operators. The problem of maps in series has been handled by a three-level approach: the first level containing data common to all the sheets in the series, the second level referring to subseries data, and the third level containing sheet-specific data, with links between the levels. Although the Bibliothèque Nationale has no plans to digitize items from its map collection, it has conducted tests in the areas of videodisc production and graphic indexing.

A non-MARC, non-AACR2, in-house system for indexing maps on a microcomputer has been developed for the American Geographical Society Collection at the University of Wisconsin, Milwaukee Library. The software, GEODEX: GEOgraphic InDEX System for Map Series, is available for sale, for operation on IBM PCs. Data for individual sheet maps are entered via a template so that coordinates for an entire series need be entered only once. An index is created by responding to a series of prompts. Searching is accomplished by entering the region, the series file name, coordinates, sheet names, and/or record numbers, but not, however, via place names, which is the form in which most geographical reference questions are framed. Although a review of the software was favorable, the reviewer pointed out the lack of standardization: "For these reasons of not meeting international standards, GEODEX probably would serve well as a medium of communication but not as the method of building a national union list of holdings or publisher output."[3]

One map system that appears to solve the above problems of maps in series, graphic indexing, and international standards is CARTO-NET, developed in 1981 by the Map Library of the University of Edinburgh's Department of Geography and funded since 1985 by the British Library. CARTO-NET's software is ORACLE, a fourth generation relational database management system. Cataloging follows *AACR2* standards in the UKMARC format. Data about maps in series are handled in a fashion similar to that of the Bibliotheque Nationale's three levels, with validation and verification tables to ensure consistency of input. Data are entered via downloads from MARC tapes, interactively, and by digitizing paper records.

CARTO-NET can be linked to gazetteer and air photograph information and to satellite imagery on other ORACLE tables; it can produce catalog cards in UKMARC format, reports, and indexes on tape, disk, or microfiche; it is the basis of a loans system; and it can be interfaced via GEOLINK to GIMMS software so that diagrams may be used to index the database. In fact, "experience gained so far with users shows that graphic

searching is the preferred route into the database."[4] By entering the coordinates of a mapped region, a diagram appears on the screen; the researcher then zooms in to query the area more closely. "The maps which lie within the chosen area, and fulfill the other search conditions such as scale or publication date, are displayed on a selective graphic index, either on a VDU-screen or on a graphic plotter" (p. 100). An option for accessing air photograph information permits calling up maps or photographs of an area measuring a specified distance from a line, useful for road planning.

Map databanks available on CARTO-NET include the Great Britain Ordnance Survey and the CIA's World Databank II. The menu search options are by U.K. national grid, latitude/longitude, country-scale-publication date, urban area-scale-date, subject, shelfmark, interest, topographic feature, town name, and series. The British Library is currently studying the establishment of a national cooperative network for map cataloging and making plans for extending the network abroad, an idea made possible by having originally developed the program according to international standards.

Additionally, a microcomputer version of CARTO-NET is now available for IBM PCs that support ORACLE. According to Barbara Morris of Geographical Information Systems at the University of Edinburgh, one of the major differences between the mainframe and the microcomputer software is that PC-CARTO-NET was designed to take advantage of windows technology. The other difference is the lower price of the microcomputer version, with growth available via networking. PC-CARTO-NET records may be uploaded into the mainframe CARTO-NET and vice-versa. Again, "the standardisation of the main System means that there is a real possibility of national and international exchange of records of individual map sheet holdings being achieved in the UK, Europe and further afield."[5] Currently the CARTO-NET team is discussing with the Research Libraries Group (RLG) the creation of a cartographic network in the United States.

The familiar warning about the need for cooperatively established standards comes from Mary Lynette Larsgaard, author of the standard textbook of map librarianship: "Any library considering computerization should begin by looking at the national standards: any U.S. library not using AACR2 and MARC had better have a very good reason for cutting off its users from what has become a national network."[6]

ARCHIVES

Historically, the reaction of archivists to the key concept of standardization was that it was inappropriate to the archival environment. Fredric M. Miller, author of the Society of American Archivists' most recent manual on arranging and describing archives and manuscripts, summarizes the archivists' traditional attitude: "For many years, archivists thought that archival description could not be standardized and information could not be shared across repository lines because each repository's holdings are unique. This attitude changed under the influence of both library techniques and automation."[7] The irony is that it is the computer, with its demand for standardization, that has enabled archivists to achieve better access to their unique items than ever before, without giving up their essential tools, like finding aids, and without distorting descriptions of their holdings to fit librarians' bibliographic descriptions.

According to Michael Cook, archivist at the University of Liverpool and frequent writer on the subject of automation of archives and records, "the computer has provided an instrument which for the first time offers the possibility of improving radically both the rate of output of finding aids and the depth of access they give to the raw materials."[8]

It has been a long struggle. The first MARC format for manuscripts, "developed by the MARC office at LC without the input of LC's own Manuscripts Division, was deemed unsuitable by most archivists. Even LC's own Manuscripts Division refused to use it,"[9] according to Janet Gertz and Leon J. Stout, respectively special collections cataloger and university archivist at Pennsylvania State University. However, archivists were experimenting with their own automated mainframe systems, such as SPINDEX at the National Archives. Additionally, after the debacle of the first MARC format, the Society of American Archivists (SAA) assumed a leadership role and created in 1977 the National Information Systems Task Force, which directed its attention not to specific

hardware or software but to the fundamental problem of standards. In 1982, the task force brought out its data element dictionary.

Their efforts bore fruit in the new MARC AMC (Archives and Manuscript Control) format released in 1983. "This was the only format developed outside the library community and the only one now jointly maintained by MARBI and a nonlibrary professional organization, the Society of American Archivists . . ." (p. 9).

MARC AMC has been generally accepted by the archivistic community because it enables unique groups of documents that lack author, title, and publisher to be identified according to archivistic standards. The main entry has become the creator of the records, and a fuller, non-AACR2 form of name may be used. The title is that supplied by the archivist and includes the characteristic span of dates. There is neither edition nor publication statement nor call number fields, and physical description is in terms of linear or cubic feet. Numerous notes fields allowing for free-text description permit the archivist's "scope and content" note to be incorporated into the record. Linking fields permit the hierarchical relationship of archival materials to be reflected. Fields have been added for genre, occupations, and functions of corporate bodies.

There are still problems with implementing the MARC AMC format, notably in the area of standardization: of names, subject headings, and other fields to which a thesaurus is applicable. One major advantage is that "in addition to providing a vehicle for bibliographic description of archival and manuscript materials, the AMC format, as its name implies, offers archivists a control mechanism with which to record and monitor accession, processing, records scheduling, and use of materials" (p. 18). More importantly, according to Richard J. Cox of the School of Library and Information Science at the University of Pittsburgh, "the use of . . . the AMC format has . . . revolutionized the archivist's attitude toward standards; archivists are now purposefully determined to work within the more established and well-supported library standard world."[10]

As many archives and libraries with manuscript holdings converted to MARC AMC, the costs of retrospective conversion became an issue. As Patricia D. Cloud, assistant university archivist in the Northwestern University archives, points out:

> Retrospective conversion of archival and manuscripts material is similar to retrospective conversion of published material in that for both the basic idea is to create machine-readable records to replace preexisting manual records; however, the procedures followed and, not coincidentally, the costs incurred are very different. . . . Only a small percentage of any library collection is "unique," or so rare as to not be cataloged in MARC. In comparison, virtually all archives and manuscripts collections are, by definition, unique.[11]

Cloud reports on the results of the retrospective conversion of some 21,000 archival and manuscript records carried out from 1984 to 1986 at twelve major RLG research libraries. The project, funded by the NEH and the Pew Memorial Trust, concentrated on materials available for scholarly research and significant in the field of the humanities, avoiding, for example, administrative archives. The conversion met the standards described in *RLG Standards for Use of the AMC Format* with collection-level descriptions according to *AACR2* and *Archives, Personal Papers, and Manuscripts* (*AAPM*). Names were established according to *AACR2* and the Library of Congress Name Authority File. At least one subject heading was required to conform to either the Library of Congress or the National Library of Medicine Subject Headings.

The time-consuming problems most frequently encountered during the RLG conversion project were the inadequacy of existing finding aids, which were the source of the descriptive information about the collections, and the familiar problem of standardization, specifically the authority work involved in establishing names and subject headings. "The relative rigor that standardization imposes will have a significant impact on record creation time" (p. 581).

For repositories of archives and manuscripts lacking access to a mainframe computer and/or too small to justify mainframe automation, the answer has been creating databases using the MARC AMC format on microcomputers.

One of the early experiments used dBase III software. Ronald J. Zboray, microfilm editor of the Emma Goldman Papers at the In-

stitute for the Study of Social Change at the University of California at Berkeley, describes four ways of using dBASE III Plus, beginning with the warning that "the programming required to use dBASE for AMC most efficiently may well be used as an argument against using dBASE at all for this purpose."[12] The various programming efforts are aimed at overcoming dBASE disadvantages: the requirement that a single database file be maintained on one disk, the wasted storage space of fixed-length fields, and the inability to search variable-length fields.

Zboray's first experiment with dBASE III Plus involved minor adaptations to the software: using forty-eight instead of the total seventy-seven MARC AMC fields, adding the dBASE III Memo Searcher utility program to enable the free-text searching of memo fields, and storing indexes on a separate disk to save storage space. His second experiment, in simple dependency, split the single database into two dependent databases, linked through a common field, which was indexed in one of the databases. This doubled the storage to 10,000 records. The third approach, multiple dependency, featured short codes in the main database record that were translated in dependent databases, which then functioned as authority files. This approach was feasible in a homogeneous archives described with repetitive contents of several fields. The fourth approach used a menu with yes/no flags rather than short codes in each field, translated by the dependent databases stored on other disks. This approach achieved the most efficient utilization of the software although sharing disks over a network slowed the performance of the program.

The most outstanding recent contribution to the application of MARC AMC on microcomputers is MicroMARC:amc, an IBM PC based software designed by Frederick L. Honhart, director of the University Archives and Historical Collections at Michigan State University (MSU) and for which he received the Society of American Archivists' C.F.W. Coker Award. When Honhart set out in 1983 to develop a microcomputer-based software using the recently released MARC AMC format, it was not then clear that this format would gain the general acceptance that it has. In fact, one of the reasons for the acceptance of MARC AMC among archivists, especially those with small to medium-sized archives, is their acceptance of Honhart's MicroMARC:amc software.

Honhart's intention, "to develop a system that would be useful profession-wide, not just for MSU,"[13] was pursued systematically by analysis of the functions of an archives repository. "Functional goals identified in the [functional requirements] documents included automating all manual processes in an integrated fashion on a stand-alone microcomputer, better reference access to the records for both staff and researchers, and the promotion of standardization, looking to the future for data exchange between repositories and within the individual user's community" (p. 82). The software was developed in a public manner, with frequent SAA updates, and tested at MSU and additional test sites before its release in 1986. Since that time, a second version has been released, with improvements such as search capacity in any part of the record, conversion to ASCII format, the MARC I/O module for inputting and outputting MARC AMC records, and shorter index production time.

Honhart feels that MicroMARC:amc combines the advantages of a local microcomputer database—low cost and local control—with the advantages of national standards and the capacity of input/output to the national bibliographic utilities. However, he warns that local control means that the concern for standards must begin at home: "The lack of system controls over the entry of illegal information, however, does pose potential problems. If the user contemplates transferring AMC records to another AMC system, then care must be taken to ensure that the records created locally meet the same standards as the system in which they will be entered."[14]

An early favorable review of MicroMARC:amc was given by Jon K. Reynolds of Georgetown University, who criticized only the inability to reformat text on the screen although the report module corrects problems of spacing and word-wrapping. He summarizes: "The MARC format does not lend itself well to microcomputer file structures, yet microcomputers are likely to be the only automation that the majority of archives will ever see. Thus the availability of a workable MARC format program is a significant contribution to archives management."[15]

While such efforts at designing automated systems such as MicroMARC:amc and apply-

ing published software such as dBASE to specific archival collections continue, recent developments toward creating national and possibly international standards have taken place. Two of the most important recent projects are being conducted by the Society of American Archivists and the [British] Society of Archivists.

The SAA's Working Group on Standards for Archival Description first met in December 1988 to discuss "whether the archival profession should seek to establish a standard for archival information management systems that would identify those areas in which all of the elements in an archival repository's information system operate."[16] The group's findings were recently reported by H. Thomas Hickerson, assistant director of the Olin Library for Rare Books, Manuscripts, and Archives at Cornell University.

The group first considered then rejected the idea of a technical standard primarily because it was not cost effective. They then considered the desirability of "a set of guidelines that would list the specific components to be included in an archival information management system and define the interrelationship between the components" (p. 26). They studied the functional requirements of documents produced by five very different archivistic environments: the National Archives and Records Administration, a federal agency with multiple divisions; the New York State Archives, a centralized state agency with a strong records management component; Cornell University's Department of Archives and Manuscripts, a department of a major research library; the Research Libraries Group, a national bibliographic network limited by the design of its database; and MicroMARC:amc, the USMARC AMC compatible system designed for use on microcomputers.

The SAA group concluded that, based on these documents, there is currently a remarkable level of uniformity of practice in American archives; however, they felt that a written standard would be beneficial in providing a set of base guidelines for vendors and others interested in developing an automated system. Such a system would comprise six interrelated functional components: appraisal, processing, space management, preservation management, disposal management, and reference. They also felt that a set of national guidelines would serve as "a foundation for a system of description standards" (p. 28).

Such a national system of description standards is concurrently under development in Great Britain. Since 1984, the Archival Description Project of the Society of Archivists has been underway at the University of Liverpool under the direction of Michael Cook, university archivist and author of publications on the application of automation to archives.

While the efforts of North American archivists, under the influence of libraries' automated systems, have centered on developing standards for bibliographic descriptions of archival materials, the British have concentrated instead on standardizing the finding aid—that essential, and frequently idiosyncratic, tool that represents and describes a collection of manuscripts or archival material. While the Americans were producing the USMARC AMC format and the APPM, the British were producing the *Manual of Archival Description* (*MAD*) and now *MAD2*.[17]

Michael Cook brings up to date the British Archival Description Project and introduces the second edition of *MAD*. According to Cook, the *MAD2* standards are based on the principle that "archival finding aids (sets of descriptions at any level) are essentially structured databases: that is, they are not simply sequences of free text without the repeating patterns that reveal the existence of an underlying structure which is common to all similar archival descriptions."[18] Common to all finding aids are the arrangement of archives and any pertinent classification scheme; the level of arrangement (collection, series or class, item, piece); the depth of description; and the provision of adding new material to an existing collection.

MAD2 provides guidelines for identifying the various components and establishing their interrelationships. For example, a standard level number is allocated to each level of arrangement and description. Patterns and models are provided along with examples of practice. Cook points out that "nothing in the *MAD* standard is innovation. The project team's work has been simply to codify and restate what has become general practice" (p. 137).

Additionally, the British team circulated a discussion draft of the UKMARC AMC for-

mat. Interestingly, in contrast to the U.S. experience, Cook says that "the outlook for MARC-based bibliographic databases in Britain and Europe appears doubtful" (p. 137). In Great Britain, according to Cook, neither major academic libraries nor the British Library Bibliographic Services are interested in promoting their development, and the main national database, BLAISE, contains no archival data. What is happening instead is that local databases using nonstandard software and hardware, such as the University of Southampton's Wellington Papers Database, are proliferating. Some of these databases are accessible via Bitnet, but the user must master each database's set of commands.

However, as both Cook and Hickerson point out, the U.K. approach to standards for archival description and the U.S. approach to standards for archival information management systems are complementary and could form the basis for international data exchange.

REFERENCES AND NOTES

1. David A. Cobb, "Online Bibliographic and Circulation Systems: The Illinois Example," *INSPEL* 22 (1988): 49–50.
2. Pierre-Yves Duchemin, "Automation at the Maps and Plans Department of the Bibliothèque Nationale," *INSPEL* 22 (1988): 246–51.
3. Jenny Marie Johnson, "GEODEX: A Review," *Western Association Map Libraries Information Bulletin* 22 (Nov. 1990): 57.
4. Barbara Morris, "CARTO-NET: Graphic Retrieval and Management in an Automated Map Library," *INSPEL* 22 (1988): 97.
5. Barbara Morris, "PC-CARTO-NET: A Microcomputer Version of the CARTO-NET Map Library Automation System," *INSPEL* 24 (1990): 73.
6. Mary Lynette Larsgaard, *Map Librarianship: An Introduction*, 2d ed. (Littleton, Colo.: Libraries Unlimited, 1987), 150.
7. Fredric M. Miller, *Arranging and Describing Archives and Manuscripts* (Chicago: Society of American Archivists, 1990), 81.
8. Michael Cook, *Archives and the Computer*, 2d ed. (London: Butterworth, 1986), 2.
9. Janet Gertz and Leon J. Stout, "The MARC Archival and Manuscripts Control (AMC) Format: A New Direction in Cataloging," *Cataloging & Classification Quarterly* 9, no. 4 (1989): 8.
10. Richard J. Cox, "The American Archival Profession and Information Technology Standards," *Journal of the American Society for Information Science* 43, no. 8 (1992): 571.
11. Patricia D. Cloud, "The Cost of Converting to MARC AMC: Some Early Observations," *Library Trends* 36 (Winter 1988): 576.
12. Ronald J. Zboray, "dBase III Plus and the MARC AMC Format: Problems and Possibilities," *American Archivist* 50 (Spring 1987): 212.
13. Frederick L. Honhart, "MicroMARC:amc: A Case Study in the Development of an Automated System," *American Archivist* 52 (Winter 1989): 81.
14. Frederick L. Honhart, "The Application of Microcomputer-Based Local Systems with the MARC AMC Format," *Library Trends* 36 (Winter 1988): 590.
15. Jon K. Reynolds, "MicroMARC:amc," *American Archivist* 50 (Spring 1987): 267.
16. H. Thomas Hickerson, "Standards for Archival Information Management Systems," *American Archivist* 53 (Winter 1990): 25.
17. Michael Cook and K. Grant, *Manual of Archival Description* (London: Society of Archivists, 1986); Michael Cook and Margaret Procter, *A Manual of Archival Description*, 2d ed. (Aldershot, England: Gower, 1990).
18. Michael Cook, "The British Move Toward Standards of Archival Description: The MAD Standard," *American Archivist* 53 (Winter 1990): 133. ■■

Communications

Evaluating Commercial Text Search-and-Retrieval Packages

Richard A. Glassco

This article presents a guide for evaluating commercial text search-and-retrieval programs. An overview of text searching is followed by a discussion of operating environments, features commonly found in commercial text-search packages, and features that distinguish one text-search package from another, such as platforms, size of database, query-formulation capabilities, hypertext capabilities, real-time data entry and classification, and browsing aids.

Over fifty vendors sell commercial off-the-shelf (COTS) text search-and-retrieval products. COTS text search packages with a wide variety of features and user interfaces are available for most computer hardware and software platforms. The information presented here is a guide for evaluating the variety of COTS programs. The concepts presented apply to individuals, universities, corporations, and government agencies.

OVERVIEW OF TEXT RETRIEVAL

The process of searching for text strings in a large collection of documents can be divided into four stages. The first three stages take place while the user is performing the text search. First, the user formulates a query, specifying in some way the material for which the text database is to be searched. Second, the system interprets the user's query, performs the search, and returns to the user a list of documents meeting the search criteria. Text systems usually perform the search by comparing search terms with an index file containing a sorted list of words found in the document database. The list of documents returned to the user is generally called a hit list. Third, the user selects documents from the hit list and browses them, reading and perhaps printing selected portions of the retrieved documents.

The fourth stage consists of entering documents into the system and creating indexes and pointers to facilitate subsequent searches. This process often takes place in batch mode during off-hours so that system performance is not degraded during working hours. Some systems, such those used by news and intelligence-collection agencies, add items to the database in real time with a live data feed. The process of loading documents into the system and updating indexes is normally not a concern of the average text-search system user.

Clearly these four stages are highly interdependent. The user interface should provide a consistent method for entering search queries and for browsing matched documents. The database should be structured to expedite the type of searching permitted by the queries, and the data-entry procedures must work within the structure of the database and the search indexes.

PLATFORMS AND SYSTEM ARCHITECTURE

The first decisions likely to be made when choosing a COTS text-search package concern the computer platform and operating system to be used and the system architecture. Text-search packages are sold for nearly every computer platform and operating system, from Macintosh and IBM-compatible personal computers (PCs) to mainframes and massively parallel supercomputers. Answers to the following questions will help determine the platform and architecture:

Richard A. Glassco is a member of the technical staff at the MITRE Corporation in McLean, Virginia. This article originated as a study for his client, the U.S. Patent and Trademark Office (PTO), to help the PTO's system architect evaluate vendor proposals for an automated system to search the text of U.S. and foreign patents, patent applications, and related documents. His other areas of interest are system engineering, simulation and modeling, and requirements analysis.

- Must the text-search system run on an existing platform and operating system?
- How much text data is to be searched?
- Will the system be used by a single user, or will many users share the same data?
- What are the cost constraints?
- What are the response-time constraints?
- Text-search packages for personal computers are generally the most versatile, with the widest range of features. The price is usually much less than for mainframe versions, but numerous copies may need to be purchased. The drawbacks are that the search process runs at a slower speed than on a mainframe and that a much smaller amount of data can be stored on a PC. At an average size of four thousand characters per page plus a typical overhead of 30 percent for index files, ten thousand pages of text would fill up more than fifty megabytes of a PC disk drive. If multiple users need access to the same document database, the PCs must be networked to a central server, or multiple copies of the database must be maintained.

A large number of packages have recently become available for UNIX-based workstations. They take advantage of workstation speed, standard windowing interfaces, and UNIX's multitasking capability. These packages often feature networking and distributed database capability.

The client-server approach is based on the distributed approach in a networked environment. The document database and the text search-and-retrieval software reside on a central server, while sophisticated data presentation and user-interface software reside on the user's workstation. The power of the server is used for the data-intensive job of comparing search terms with text files or indexes, while the workstations are best suited for graphical interfaces, local searches, and possible data decompression.

Another use of the client-server approach is to break the index file into pieces corresponding to work groups and to maintain a server for each group. This approach provides fast response time for documents "owned" locally. Searches of portions of the index stored on other servers can be performed in the background while the user is retrieving and studying locally owned documents. This system architecture is illustrated in figure 1. A disadvantage of this approach is that each subindex has to be updated individually each time the master file is updated. A partitioning of the index file to correspond to different user groups could also be done logically on a mainframe, without any difference in the physical architecture of the system.

A mainframe-based system is generally more expensive and less flexible than the previous architectures, but provides for large amounts of storage, fast response time, and standard data management and configuration control. The mainframe may also handle query and display formatting, enabling searches to be conducted from nonintelligent character-based terminals.

In the parallel-processing approach, many processing units conduct searches simultaneously. Typically, the file to be searched is broken up into many pieces, and each processor searches its segment of the index file. The processors may or may not share memory and storage. If the processor and file segments are balanced correctly, each processor can operate independently of the others, and all processors will complete processing at approximately the same time. The resulting multiple list of hits is merged before being presented to the user for retrieval. This process is portrayed in figure 2.

The number of parallel processors can range from two to several thousand. A typical architecture with multiple processors features a complex of ninety search engines (using Motorola 68020 microprocessors) controlled by two search-engine complex managers connected to each other and the mainframe over a local area network (LAN). When a very large number of processors work in parallel, the term *massively parallel* is used.

The greatest problem in a multiprocessor machine is that of coordination among the processors. If the pieces of the problem are interdependent, significant effort is required to make sure that one processor does not get ahead of the point where it needs data from another processor. The application of searching index files or text files, however, is ideal for multiprocessing, because the files can be segmented so that there is no interdependency among them. Each segment of the index file can be searched independently.

Some commercial systems use parallel processors to search original text files in their entirety. They are optimized for speed in character comparison and data flow through the processors. Index files are not required,

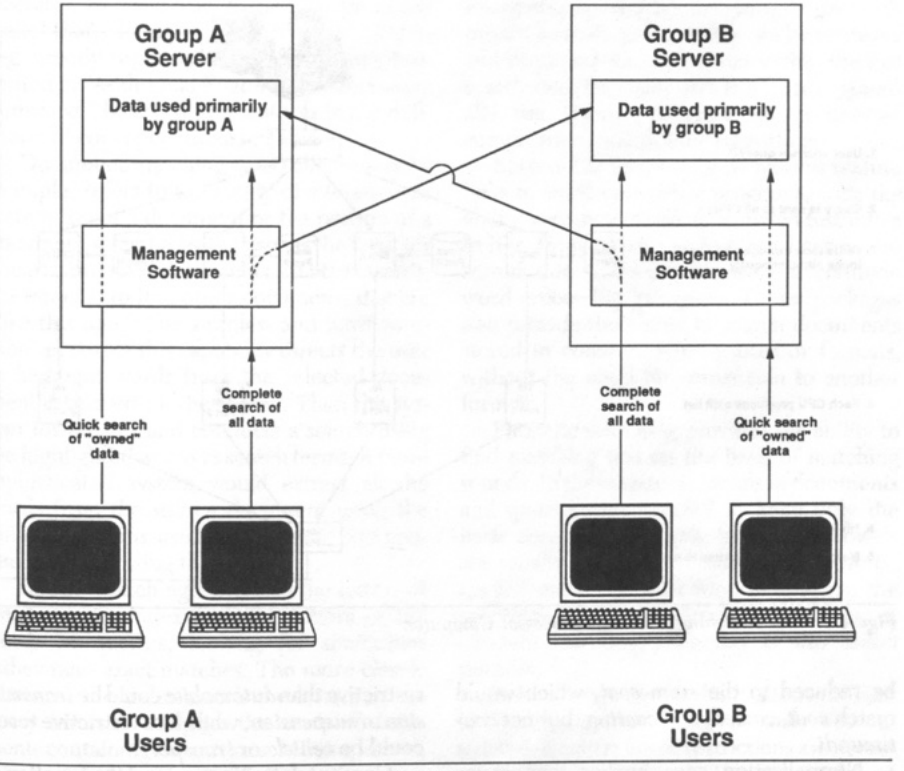


Figure 1. Partitioning a Text Database Across Group Servers.

and the search time grows linearly with the size of the database.

FEATURES FOR QUERY FORMULATION

The most basic function of a text-search system is single-word searching to find documents that contain a specified word or phrase. Most text-search packages, however, offer additional capabilities to combine or extend single-word searches to help the user find relevant documents and exclude irrelevant ones. This section presents text-search features in order ranging from most common and least powerful to least common and most powerful. Evaluators of COTS packages should consider each of these features to determine the value to the particular application.

The Boolean operators—AND, OR, and NOT—connect search terms to limit or expand queries, such as *term1 OR term2*. More complex logical constructions can be built using parentheses to group terms, such as *(term1 OR term2) OR (term3 AND NOT term4)*. While it is possible to target a search accurately using a Boolean expression, the complexity of logical operators can be daunting to untrained users.

Proximity operators specify that search terms occur within a specified number of words, sentences, or paragraphs of one another. Often the user can select whether the terms may occur in any order or must occur in a specified order.

Wildcard characters are used when the exact form of a search term is not known. Usually there is one wildcard character that matches any single character (e.g., "dis?" matches *disk* and *disc*) and another wildcard character that matches any number of characters, including zero (e.g., *colo*r* matches *color*, *colour*, and *colocator*).

Stemming takes the root stem of each query term and searches for all words beginning with that root and having normal English suffixes. For example, the word *coated* would

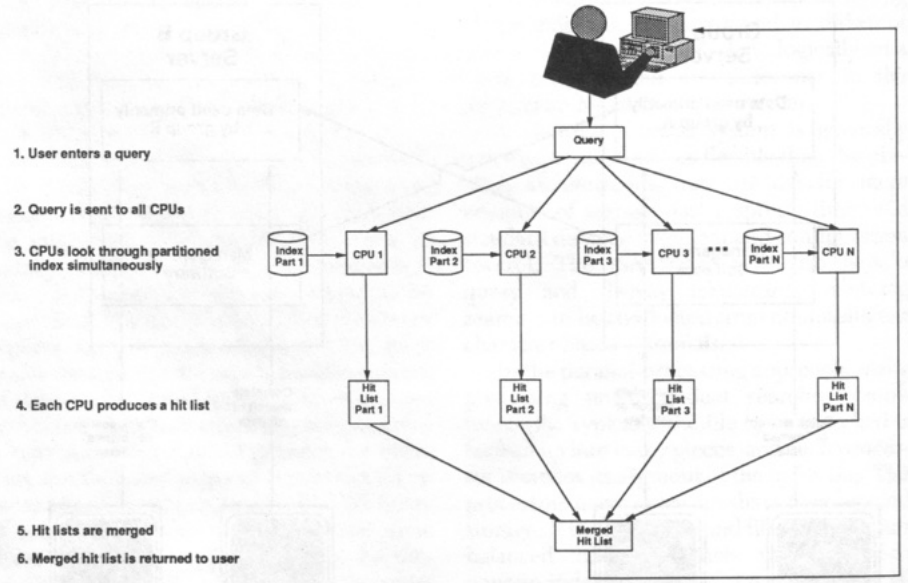


Figure 2. Text Searching on a Multiprocessor Computer.

be reduced to the stem *coat*, which would match *coat*, *coated*, and *coating*, but not *coatimundi*.

Normalization goes further than stemming, automatically supplying irregular forms of words, such as plurals and past tenses, and searching for those terms as well. For example, the search term *frozen* would also initiate a search for *freeze*, and the search term *mouse* would also initiate a search for *mice*. These irregular forms must be looked up in a stored dictionary.

Synonym files go a step further than normalization, supplying synonyms of search terms, retrieved from a stored dictionary. For example, the search term *car* would initiate searches for *automobile* and *motor vehicle* as well as for *car*. The files could also supply spelling variants such as *flavor* and *flavour*. A third possibility is the inclusion of full words represented by acronyms and abbreviations.

Thesaurus capability goes beyond synonym files to contain a hierarchy of terms either more or less specific than the given word. The user can be prompted interactively to specify whether to narrow the search by choosing a more restrictive term or to broaden the search by supplying a less restrictive term. For example, suggested terms more restrictive than *automobile* could be *transmission* or *suspension*, while less restrictive terms could be *vehicle* or *transport*.

Concept definitions extend the function of a thesaurus. Concepts (also called topics) are rule-based objects that represent areas of research. Each concept contains a hierarchy of related terms, weighted by importance to the concept. The collection of concepts constitutes a knowledge base. The knowledge base must be constructed in advance by an automated tool, with input from experts. The more specialized the information in a database is, the more straightforward is the process of building and using the knowledge base. A few commercial text-search packages offer concept-based searches, and the number is growing. They come with prepackaged knowledge bases that are, of necessity, very general.

Term weighting introduces a gradation in matching responses, recognizing that some search terms are more central, or relevant, to the desired concept than others. The score of a matched document is an arithmetic function of the weights, determined by the pattern of Boolean connectors between terms. Weights can similarly be attached to the distance between pairs of terms related by a proximity

operator to yield the degree of proximity rather than a binary one or zero. Term weighting directly supports relevance ranking, presented in "Additional Features for Document Browsing," below. It also features in the definition of concepts, discussed above.

Document matching, also called query by example, refers to the ability of a text-search system to use a document or the portion of a document selected by the user as the basis for a search on other documents. In other words, the intent is to find other documents that are "like this one." The simplest and most common version of this capability directs the user to highlight words from the selected document displayed on the screen. Then the system formulates and conducts a search using the highlighted words as search terms. A more sophisticated system would extract all the words from the source document using the same algorithms used for indexing and conduct a search using them all.

Pattern matching compares the letters of words in the index with the letters of the words in queries, looking for similarities rather than exact matches. The more closely letters in the index match the letters in the query terms, the higher the score of documents containing those terms. Missing letters, incorrect letters, or extra letters will reduce the score, but will not eliminate the document as a hit until the score falls below a threshold score. As a result, the user does not have to guess ahead of time where to put wildcard symbols; the system will select anything that comes close. This feature is very helpful for searching text data that entered the system via optical character recognition (OCR), where occasional misspellings caused by misinterpretation of characters may not be caught and corrected. A sample pattern match is illustrated in table 1.

SPECIAL FEATURES

Once the platform, system architecture, and basic features for the text-search package have been determined, the selection should proceed on the basis of special features, that is, capabilities that the desired system must have, and that are found in some, but not all, COTS packages. If the desired system must have one or more of the capabilities described in this section, the search can be narrowed to a smaller number of COTS packages.

Some COTS packages feature links to image-display capabilities. Images generally cannot be searched, but they can be retrieved and displayed when a pointer within the text is activated. Packages with this feature generally use industry-standard image storage, compression, and display algorithms.

Several COTS text-search systems feature links to word-processing programs, with the ability to copy sections of retrieved text into a buffer, from which the text can later be inserted into a document using a common word-processing program. These packages also provide the ability to search documents stored in common word-processor formats, without the need for conversion to another format.

Phonetic searching provides the ability to find matching text on the basis of matching sounds. In these systems, words in documents and queries are encoded to keep only the basic consonants; double letters and vowels are usually dropped. Because sounds can be spelled many different ways in English, the precision of this type of search is often low. A phonetic-searching capability is also called soundex.

Natural language capability allows the user to specify queries using normal, but restricted, English. Some restrictions are necessary to avoid ambiguities. The queries do not require training in a formal query syntax, so the system can be used more easily by the general public. On the other hand, natural language queries do not permit great complexity. A system capable of interpreting natural language queries correctly must have a broad knowledge of syntax and synonyms. The system must be able to recognize the subject, verb, object, and any qualifiers, and to supply synonyms if the words used do not correspond to search terms.

Some COTS products advertise the ability to search in foreign languages. There are two categories of problems associated with foreign languages. The first is the storage, entry, and display of accented and non-Roman characters, unless the text is transliterated. The second is that advanced search methods that depend on grammar, such as automatic plurals and synonyms, must be rewritten using the grammar of each foreign language.

Standard Generalized Markup Language (SGML) is a markup language defined by the International Organization for Standardization (ISO) standard 8879. SGML contains

Table 1. Example of Pattern Matching

Search Term: Salvage Threshold: .9

Index Term	Evaluation	Score	Pass
Salvage	Exact match	1.00	Yes
Selvage	A valid word, not a likely typo	0.80	No
Saivage	Not a valid word, "i" resembles "l," likely scanner error	0.90	Yes
Savage	A valid word, a possible typo	0.82	No
Salbage	Not a valid word, "b" near "v" on keyboard, likely typo	0.90	Yes
Salvag	Not a valid word, no other word begins that way, probable typo	0.92	Yes
Salvager	Contains desired term, regular form, probable hit	0.96	Yes

rules for specifying a treelike structure of within-text tags to define sections of documents such as headers, footnotes, quotes, abstracts, tables, lists, and pointers to graphics files. It usually specifies only the structural elements of the document; subsequent programs applied later determine how those elements are displayed or printed. A text-search program using SGML tags can search for words within specific element types such as figure titles, tables, and headings. SGML fully supports hypertext applications that handle documents containing both text and images. SGML may be required in some situations for compatibility with standards or other organizational users.

Machine classification is the ability of a text-analysis system to classify documents based on their word content. Given a sample set of documents belonging to a classification, the computer can determine statistically which words are common in documents in that class and are not common in documents of other classes. These words provide the best discriminators to use when a new document is to be classified. Thus the computer can be an aid to classifying documents into existing classifications or to the definition of classifications.

The practice of routing incoming documents to interested parties based on the word content of the documents is well established in news- and intelligence-gathering communities. The process is the reverse of the text-search process; instead of comparing a query with multiple documents, it involves comparing an incoming document with multiple standing queries, called profiles. There is a profile for each potential user of incoming information, containing terms or topics in which the user is likely to be interested.

ADDITIONAL FEATURES FOR DOCUMENT BROWSING

The features described in this section are very useful in helping the document-browsing process. Evaluators of COTS packages should consider each of these features to determine the value to the particular application.

Relevance ranking (also known as relevance ordering) is the ability to sort the hit list into descending order of probable relevance before presenting the list to the user. Many COTS packages now claim the ability to do relevance ranking. Of course the machine can only guess which documents the user will find most relevant, but there are several schemes by which the machine can assign relevance scores to documents. These schemes include using term weights, giving greater weight to more recent documents, giving greater weight to documents containing the highest number of hits, giving greater weight to documents containing the most different search terms, giving greater weight to documents with the highest search-term density (number of search-term hits divided by the total number of words in the document), giving greater weight to documents with search terms occurring most closely together, and giving greater weight to those search terms that occur least frequently in the inverted index (these terms are assumed to have the best ability to discriminate between relevant and irrelevant documents).

Once a relevance score has been computed for each matching document, the sys-

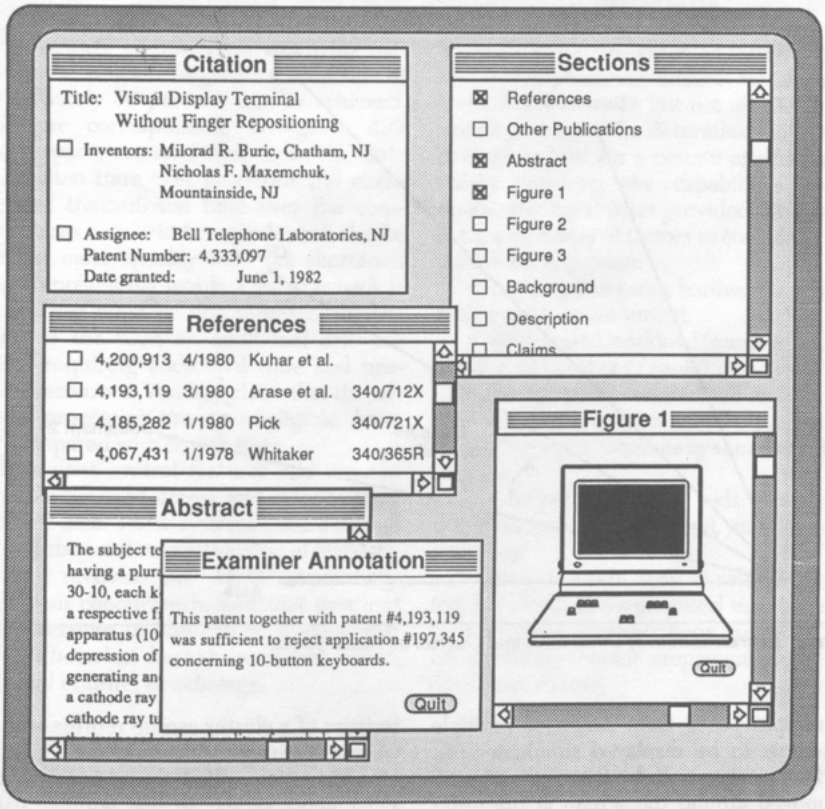


Figure 3. Sample Screen for Hypertext Capability.

tem sorts the document titles before displaying them to the user. The user may specify a relevance threshold, above which all documents are retrieved and displayed and below which documents are not retrieved unless requested. The user then starts at the top of the list and keeps examining documents until he or she decides there is no more relevance to the search at hand.

Hypertext systems link documents in the database by pointers to other documents with similar characteristics. The links can be based on words, phrases, or concepts. The links can facilitate jumping from one document to another within a complex web of interrelationships. The key to a hypertext system is establishing the links (often called hyperlinks), either manually or automatically with expert assistance. The end of a link may be a specific place within a document, an annotation attached to the document, a span of text within the document, or the document itself. Links can be made between a node in one document and another node in the same or a different document. Hypertext is often associated with the capability to retrieve and display nontext information such as images or recorded sounds attached to documents. Full hypertext capability is more common in PC programs than mainframe-based programs. Figure 3 shows a possible screen for a hypertext-searching system for U.S. patents.

Many popular programs and operating environments use a graphical user interface (GUI) that features multiple windows and a pointing device (usually a mouse). Among other capabilities, windows can enable text and images to be displayed on the same screen

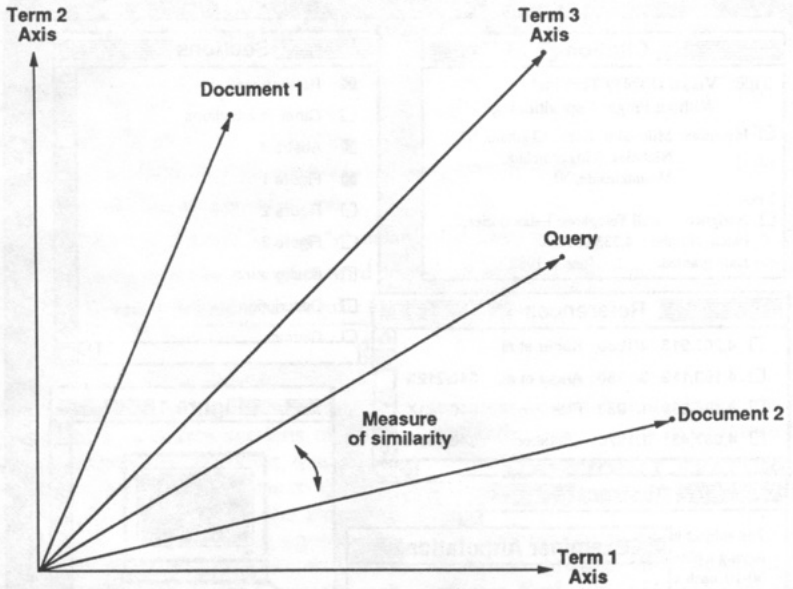


Figure 4. Representation of Documents and Queries in Vector Space.

or multiple pages from single or multiple documents to be displayed simultaneously. Windows can generally be enlarged, reduced, and moved around the screen at the user's convenience.

FEATURES RELATING TO DOCUMENT STORAGE

The features described in this section do not visibly affect functionality. They are of concern to system designers, however, because they affect storage. Evaluators of COTS packages should consider each of these features to determine the value to the particular application.

The vector-space model of storage is one way of quantifying how similar documents are to each other, to a group of documents, or to a query. If there are T possible indexing terms, then each document can be represented as a T-dimensional vector in T-space, with the nth coordinate being the number of times index term n occurs in the document. The similarity between two documents can be defined as the cosine or dot product of the angle between the two vectors. A group of similar documents is called a cluster, and the characteristics of a cluster can be represented by a weighted average of the individual document characteristics, called the cluster centroid. Documents similar to one another will be stored in the same virtual or physical area of database storage. Queries can be represented in the same way as documents. At query time, the characteristics of a query are compared with the characteristics of cluster centroids until the closest match is found. Then all the document members of the cluster can be considered as matches to the search. The search process is spared the trouble of searching the entire index file for individual occurrences of specific search terms. A representation of documents and queries in vector space is portrayed in figure 4.

The disadvantage of the vector-space approach is the large amount of computation required to construct the vectors and to do mathematical operations on them, and the storage required to store them. With the decreasing cost of central processor power (especially with massively parallel architecture) and storage, this disadvantage is not as formidable as it once was.

Text data compression can produce signif-

icant savings in storage space. By using fixed-length or variable-length coding schemes to take advantage of frequently occurring patterns in the English language, a reduction of approximately 50 percent can be achieved. There are corresponding savings in disk storage space, retrieval time from the disk, transmission time over processor I/O channels, and transmission time over the communications network. A text index can also be searched more quickly with the shortened strings representing words. One drawback is that the text must be decompressed before display to the user, an additional software process requiring additional time and processing resources. Vendors claim that the additional processing time is negligible, however, compared with search time.

Document control features add security from unauthorized access and misplaced or destroyed data. These systems ensure that all users see the most current version of text data, that any revisions are made sequentially rather than concurrently, and that past and current versions of files can be re-created if data have been lost due to hazards such as fire, accidental erasure, or sabotage.

SUMMARY

This article has presented an introduction to text-search platforms, architectures, and features found in some but not all COTS text-search packages. The determination of which package is best for a certain application requires weighing the capabilities needed against the capabilities provided. The following is a summary of factors to consider during the selection process:

- Existing computer hardware and operating-system environment
- Anticipated workload (number of documents and number of users)
- Distributed or centralized architecture
- Mission requirements, such as integration with images, searches by sound or foreign language
- Query-formulation aids, such as thesaurus, concept definition, and document matching
- Browsing aids, such as relevance ranking, hypertext, and a graphical user interface
- Storage considerations, such as data compression, vector representation, and document control

BIBLIOGRAPHY

Croft, W. B. "An Overview of Information Retrieval Techniques." Paper presented at the Fourteenth Annual International Association for Computing Machinery (ACM)/Special Interest Group in Information Retrieval (SIGIR) Conference on Research and Development in Information Retrieval, Chicago, October 1991.

Delphi Consulting Group. *Text Retrieval and Imaging: An In-Depth Exploration of the Technologies and Alternatives*. 6th ed. Boston: Delphi Consulting Group, 1991.

Saffady, W. *Text Storage and Retrieval Systems: A Technology Survey and Product Directory*. Westport, Conn.: Meckler, 1989.

Salton, G. *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Reading, Mass: Addison-Wesley, 1989. ■■

Special Section: A Decade of East Asian Scripts on RLIN

Introduction

Joan M. Aliprand

September 1983 was a significant month for the East Asian libraries of North America. With the introduction of RLIN CJK, they were now able, for the first time, to create machine-readable records containing vernacular Chinese, Japanese, and Korean (CJK). But cataloging is not done just for its own sake. RLIN CJK meant that bibliographic descriptions in the authentic writing system were now available online, not only for reuse by other catalogers, but to meet the needs of scholars, bibliographers, and reference librarians.

The following papers link the past and the future of East Asian cataloging. Hideo Kaneko recounts past efforts to standardize and share East Asian cataloging. This vision of cooperation was not fully achieved until East Asian cataloging was automated with RLIN CJK.

Ai-Hwa Wu writes about retrospective conversion, the ongoing effort by East Asian libraries to have all their cataloging information in machine-readable form. This paper brings together the past and the future: the past, when the cataloging was done on cards, and the future, when all the records will be online.

RLIN CJK also demonstrated that a single unified repertoire of ideographs could be used to encode and retrieve text in the Chinese, Japanese, and Korean languages. The concept of a "unified Han" ideographic repertoire is embodied in the international character set published earlier this year and will be the cornerstone for future developments in the automation of East Asian scripts. ■■

RLIN CJK and the East Asian Library Community

Hideo Kaneko

RLIN CJK became operational in September 1983. At the end of March 1993, there were well over a million records with vernacular Chinese, Japanese, and/or Korean (CJK) in the RLIN database. Most of these records were entered online by thirty-seven participating institutions. The remainder, 142,391 records as of March 1993, were batch loaded. Most of these came from OCLC. The six most productive participants are the Library of Congress (209,511 records), Yale University (98,488 records), Columbia University (93,728 records), the University of Michigan (72,933 records), Princeton University (64,494 records), and the University of Toronto (63,304 records).[1]

What has this meant to the East Asian library community? Obviously, the advent of RLIN CJK meant that for the first time the community was able to bring under control all Chinese, Japanese, and Korean works within a nationwide system of automated processing, thereby making them accessible to the scholarly community.

Ever-advancing electronic technology and communications networks have revitalized cooperative and coordinated collection development and shared access. Bibliographic control, the development of a comprehensive bibliographic database, is a prerequisite to any cost-effective sharing of resources, that is, to collection development and access.

Hideo Kaneko is Curator, East Asian Collection, Yale University Library.

RLIN CJK has also produced additional dividends for the East Asian library community; it has brought the technical processing and bibliographic practices of the East Asian libraries into the mainstream of American librarianship. CJK was developed as an integral part of the main bibliographic record file of a national bibliographic utility. The design ensures that CJK records are accessible in the same manner and same instance as Roman-alphabet records. Furthermore, since machine manipulation of bibliographic records requires greater standardization and precision in record creation, RLIN CJK forced the East Asian library community to conform, literally overnight in some cases, to the bibliographic practice of mainstream American librarianship. Cataloging rules addressing the needs of East Asian libraries had been developed over the years, but they had not been hitherto closely adhered to. The effect of RLIN CJK on East Asian collections can be seen if we consider previous efforts to coordinate East Asian cataloging.

COORDINATION OF EAST ASIAN CATALOGING

The first cooperative cataloging project of the East Asian library community was the Oriental Card Reproduction Project. The Library of Congress (LC) reproduced (without editing) and sold by subscription catalog cards for Chinese, Japanese, and Korean works cataloged by American libraries. During the life of the project (1949–58), cards for more than 88,000 works were reproduced and distributed: 54,278 cards for works in Chinese, 32,532 for works in Japanese, and 1,985 for Korean titles.[2] Some of these cards, produced in an abundant exercise of freedom of expression, are still in the catalog trays of many libraries.

In 1953, the Oriental Processing Committee (OPC) was established by a directive from the Librarian of Congress to study problems and make recommendations on all aspects of the cataloging of Oriental materials. From 1953 to 1957, the OPC held almost one hundred meetings.

In 1954, the Cataloging and Classification Division of the American Library Association appointed the special Committee on Cataloging Oriental Materials. For the next four years, a heavy and continuous correspondence took place between this committee and OPC, resulting in additions and changes to the two basic codes of 1949: *ALA Cataloging Rules for Authors and Title Entries* and the *Rules of Descriptive Cataloging of the Library of Congress*.[3]

These modifications and additions for the cataloging of Chinese, Japanese, and Korean materials were subsequently incorporated into the *Anglo-American Cataloging Rules* (*AACR*) of 1967 (with the exception of the sections on romanization, word division, etc.). Romanization rules for Chinese, Japanese, and Korean have since been issued separately by the Library of Congress, most recently in the *ALA/LC Romanization Tables*.[4]

To their credit, the ALA special committee and LC OPC made two crucial basic decisions early on:

1. There should be no separate cataloging code for East Asian materials, but the existing ALA and LC rules should be modified and expanded in such a way as to make the cataloging of these materials feasible.
2. The cataloging practice should be such that the produced cards would be capable of being interfiled with cards for works in Western languages.

The decisions were followed in all subsequent works of the OPC and have also been the underlying principles for cataloging in most of the major East Asian collections of the United States. At the 1957 spring meeting of the Association for Asian Studies, it was announced that "the newly developed cataloging rules had been adopted by all major American library collections of East Asian publications."[5]

For the first time, there was a standard format for the cataloging of East Asian–language materials. The rules opened the way for the development of national cooperation in the cataloging of these materials. This had not been possible previously because of the lack of standards for romanization practice, for the citation of personal and corporate authors in a uniform manner, and for the listing and order of arrangement of the elements of bibliographic description, among other things.

SOURCES OF EAST ASIAN CATALOGING COPY

In 1957 the Far Eastern Language Section was established in the Descriptive Cataloging Division of the Library of Congress. The original intention was that the section not only

would catalog LC holdings but also would edit catalog cards prepared by other institutions for publication or adoption for LC's own use. It soon became apparent that the complexities of cataloging materials in East Asian languages, the printing of their vernacular scripts, and the great variance in compliance with cataloging rules found in the submitted cards made "cooperative cataloging" a slow, expensive, and impractical undertaking.

Instead, the Library of Congress began the large-scale printing and distribution of LC cards for Chinese, Japanese, and Korean monographs and serials in 1958. The availability of Japanese cards was greatly facilitated in 1968 when LC established an overseas office in Tokyo and the National Program for Acquisitions and Cataloging (NPAC) was extended to Japanese materials. Through a shared cataloging arrangement with the National Diet Library, LC was quickly able to provide cataloging copy for current Japanese works to American libraries. No similar arrangement was ever made for Chinese and Korean materials. NPAC for Japanese books ended in December 1985.

When Japan MARC data are available for loading into the RLIN database, the cycle will be repeated, with cataloging for Japanese materials in machine-readable form. Increasingly, retrospective records are also available in Japan MARC. It is hoped that sources for Chinese and Korean materials will also be acquired.

Because of the complexity of cataloging CJK materials, the catalog cards produced locally by libraries other than LC for East Asian–language works were not incorporated into the printed *National Union Catalog* if LC did not hold the title. In response to the East Asian library community's strong request for better and faster bibliographic control of Chinese-language materials, LC published the *Chinese Cooperative Catalog* (*CCC*) from 1975 to 1982.[6] The *CCC* listed LC preliminary cards and printed cards along with unedited cards from selected libraries; there were no added entries, cross-references, or additional holding-location identification.

Book catalogs of many of the major East Asian collections were published in the 1960s and 1970s by G. K. Hall, providing improved access to the holdings of institutions such as the University of California at Berkeley, the Hoover Institution at Stanford University, the University of Chicago, the University of Michigan, the New York Public Library, and, most recently, the Harvard-Yenching Library.

SUBSEQUENT EVENTS

With one or two exceptions, the cataloging of most major East Asian collections had varying degrees of compliance with cataloging rules. In most cases, CJK cards were not interfiled with Western-language cards in the main library's catalog. Such variances and peculiarities created no major problems, so they were tolerated.

East Asian cataloging activities (as well as acquisitions work and public services) often took place in a pocket, away from the main library. In many cases, CJK catalogers reported to a supervisor who specialized in Chinese, Japanese, or Korean; because of this division based on linguistic competence, cataloging practice sometimes differed within an East Asian library.

The East Asian library community continued to deal with the increasingly complex problems relating to the cataloging of East Asian library materials. In fall 1980, shortly after *AACR2* had become the focus of attention among the nation's catalogers, the Subcommittee on Technical Processing of the Committee on East Asian Libraries, Association for Asian Studies, began compiling a workbook that would illustrate the new rules with East Asian examples for East Asian catalogers.[7] The workbook was needed for the following reasons:

1. *AACR2* does not provide sufficient guidelines and examples for treating East Asian materials.

2. The ISBDs on which Part 1 ("Description") of *AACR2* is based do not have provisions and examples for the unique situations in the bibliographic description of materials in East Asian languages.

3. National, regional, and local *AACR2* training institutes had not touched upon the special problems East Asian catalogers face.

AACR 2 Workbook for East Asian Publications was published in March 1983 and was probably helpful to catalogers embarking on automated CJK cataloging.[8]

EFFECT OF RLIN CJK

After RLIN CJK was introduced in September 1983, many CJK catalogers faced a double trauma of having to learn machine-based

cataloging and USMARC tagging, and, for the first time, having to follow all the existing cataloging rules closely. On the whole, CJK catalogers made an excellent transition to the automated environment, with most of them now reporting to a single CJK-cataloging supervisor instead of to individual language specialists. The introduction of automation and the necessary training that went with it demonstrate that some cataloging issues apply to all languages. Today, some CJK cataloging units are even more productive than their counterparts in Roman-alphabet languages, despite the fact that CJK cataloging requires reading and manipulating CJK scripts in addition to performing as much Roman-alphabet work as in Western-language cataloging.

RLIN CJK bibliographic records, in romanized form in most cases, are now integrated into local databases for all to see. As a consequence, there has been better communication among catalogers throughout the country, and an enhanced sense of camaraderie has emerged within individual libraries. Happily, this pleasant breeze is also reaching up to clear the air between East Asian collections and their respective main libraries, creating better communication and mutual understanding. Whenever local CJK enhancements arrive and East Asian scripts start appearing in the local online database, the East Asian collection of that institution sails even deeper into the waters of the mainstream of American librarianship.

REFERENCES

1. Research Libraries Group, "Cumulative RLIN CJK Statistics as of 06/30/93," distributed by RLG.
2. Edwin J. Beal, Jr. "The Committee on East Asian Libraries: A Brief History," *Association for Asian Studies Committee on East Asian Libraries Newsletter* 41 (Sept. 1973): 43.
3. Beal, "Committee on East Asian Libraries," 44–45.
4. Randall K. Barry, comp. and ed., *ALA/LC Romanization Tables* (Washington, D.C.: Cataloging Distribution Service, Library of Congress, 1991).
5. Beal, "Committee on East Asian Libraries," 45.
6. Gloria H. Hsia, "From CCC to NUC: Changing with Time," *Association for Asian Studies Committee on East Asian Libraries Bulletin* 73 (Feb. 1984): 35–36.
7. Association for Asian Studies, Committee on East Asian Libraries, Subcommittee on Technical Processing, *AACR2 Workbook for East Asian Publications*, ed. Thomas H. Lee and Beatrice Chang Ohta (Madison, Wis.: University of Wisconsin Memorial Library, 1983).
8. *AACR2 Workbook for East Asian Publications*, preface. ■■

With Characters: Retrospective Conversion of East Asian Cataloging Records

Ai-Hwa Wu

Just as the automated cataloging of East Asian collections in North America lagged behind the automation of general libraries, so retrospective conversion (recon) of cataloging records in East Asian languages has lagged years behind retrospective conversion for material in other languages. While Schottlaender noted that "the American library community has been actively engaged in converting its manual cataloging records to machine readable form since the mid-1970s," it was not until the late 1980s that recon at East Asian libraries began to occur.[1] In her survey conducted in early 1990, Tsiang reported that in 1989–90 only nine libraries and one consortium had either completed or started recon projects of their East Asian cataloging records.[2] Even though the number of East Asian libraries that have begun to convert their cataloging records had increased to twenty-four by April 1993,[3] that number constitutes only 40 percent of the total number of libraries reported as having an East Asian collection in 1992.[4] The retrospective conversion of East Asian collections is still a relatively new application of library automation.

Recon of East Asian library cataloging records completes the move into the mainstream of all library services and endeavors having to do with materials in East Asian languages. The genesis of the integration of East Asian cataloging into the mainstream of automated library processing began in 1983 with the introduction of RLIN CJK. Furthermore, as more and more East Asian cataloging records are converted into machine-readable form, the wealth of earlier East Asian scholarship is increasingly brought under better bibliographic control and made accessible for research both locally and globally.

Very little has been written on the retrospective conversion of East Asian materials. I consulted Hseuh's study of the professional literature for 1980–90,[5] and conducted an article search through library journals for the period since 1990. To my surprise and dismay, I discovered only two articles that deal specifically with recon of East Asian cataloging records.[6,7] A few reports have been posted and shared on listservs (such as the Committee on East Asian Libraries' EASTLIB and the Research Libraries Group's East Asian Member Network [EAMEMNET]), but the majority of the efforts and valuable experiences of the twenty-four East Asian libraries mentioned above has simply not been made widely known through publication. It is my hope that in this report I can share some of my thoughts and, more important, the experience and findings that my colleagues have shared with me.

Ai-Hwa Wu is Chinese Studies Librarian, Arizona State University, Tempe.

OPTIONS FOR RECON

To facilitate a more meaningful discussion on recon at the Research Libraries Group's East Asian Studies Roundtable in March 1993, a survey was conducted and two written reports were submitted for posting on EAMEMNET.[8–10] A total of twenty-two responses from twenty libraries were received. Eleven libraries reported a recon project. Four of these projects were done by UTLAS and Retro Link Associates. The remaining seven were done in-house.

Reporting libraries gave a number of reasons for being able to undertake a recon project:

1. The decreasing number of new materials to be cataloged due to a reduction in the acquisitions budget
2. Participation in a librarywide recon project
3. The need to get fuller bibliographic records into a local online catalog to facilitate an integrated local system
4. The performance of recon on an as-encountered basis (e.g., when adding volumes or noting dead numbers) or when staff was available

The nine libraries that had not undertaken a recon project stated that a lack of funding was the main impediment.

Tsiang's report on East Asian recon included libraries that do not use RLIN CJK. Six libraries did recon under contract with a vendor, and four had in-house projects. The vendors selected were OCLC and Asian Shared Information & Access (ASIA).

In my second survey of the recon activities

of East Asian libraries (on EASTLIB in April 1993), fourteen libraries responded. Six of these libraries had neither responded to the RLG survey in January 1993 nor been included in Tsiang's report. Four of these six libraries had some recon activity, and one was in the process of submitting a grant application for a contract-out recon project.

Although Schottlaender identified three recon options, there are essentially only two options for East Asian libraries: in-house recon or use of a contractor. Among the in-house projects, the primary methods adopted by East Asian libraries are direct online conversion or a combination of off-line batch recon with some direct online conversion.

Direct online conversion means searching for matching records in either the RLIN or the OCLC databases, requesting a copy of any source record found (e.g., "derive" in RLIN), and creating new records for titles not found. An excellent example of this method (using RLIN) is reported in detail by Seely.[11]

A combination of offline batch recon and direct online conversion involves initial use of a batch recon method to machine search the titles against the bibliographic database. The subsequent step is for staff to create and input new records online for titles not found in the first step. For example, with RLG's Batch Recon (BRCON) software, searches can be keyed on a PC. The searches are forwarded to RLG (either by uploading of a file or on diskette), where they are run against the RLIN database. An exact match causes a record for the library doing the recon project to be added to the RLIN database; inexact matches and failures to retrieve a result are reported back to the library. Cataloging records for titles with no cataloging copy are transcribed directly into the RLIN database by library staff. Reports by Cheng and Wu posted on RLG's Conferral Bulletin Board provide detailed descriptions of how this works.[12]

For libraries that contracted out the conversion of their East Asian cataloging records, two different forms of machine-readable records have been produced: romanized only, and both romanized and vernacular (CJK). The former will need upgrading with East Asian characters by direct online input. The latter are completed records, since their essential bibliographic data are also provided in the East Asian scripts, and they are accessible and discernable that way.

The University of Hawaii, the University of Toronto, and the University of Michigan are examples of institutions with East Asian libraries that converted their East Asian cataloging records in romanized form only. The romanized records of the University of Toronto were added to the RLIN database soon after the introduction of RLIN CJK. Many of them now have vernacular script access, either because Toronto has added CJK data or because another library's CJK record has clustered with the romanized record. In the preceding three cases, the records of the East Asian libraries were converted by vendors under a retrospective conversion project for the university library as a whole. Hawaii used Retro Link Associates as its vendor; Michigan and Toronto used UTLAS.

Some institutions have contracted with vendors to do complete record conversion with both romanized and East Asian vernacular data. Examples are Cornell University (in 1993 with Retro Link Associates) and the University of California at Los Angeles (since 1989, with OCLC and ASIA).

PROBLEMS OF EAST ASIAN RECON

Many problems related to recon for East Asian cataloging records are similar to those for non-East Asian records, as in the following examples:

- Cataloging rules and local policy and practice change over the years. This requires converters to have an adequate historical knowledge of these rules and practices and how they have changed, and an awareness of the current version of AACR2 and LC Rule Interpretations.
- Name, uniform title, and subject heading change.
- The local systems have idiosyncrasies.

These problems have been well addressed by Hart and in several articles in *Retrospective Conversion: History, Approaches, Considerations*, a special issue of *Cataloging & Classification Quarterly*.[13,14]

Although there is not a great deal of difference from the current cataloging of East Asian materials, East Asian recon presents unique challenges:

- East Asian recon calls for language expertise, a good command of the romanization scheme for the language(s), and a suitable cataloging background.

- There have been some variations in romanization practices over the years.
- Hyphenated and unhyphenated names (in romanization) present a challenge for the searcher.
- Works published in one East Asian country are often republished in another.
- Records produced prior to 1958 (when a national standard for East Asian cataloging was introduced) may present special problems.
- ISBD punctuation must be supplied.
- Handwritten data on cards may be difficult to read.
- Some notes are particular to the non-Roman environment.
- A limited number of vendors can provide East Asian recon.

For a full-level, standard East Asian bibliographic record, data must be in both romanized and East Asian vernacular forms in at least the 245 and the 260 fields (and the 250 and 4xx fields where appropriate). This demands not only expertise in the relevant language(s) but also a good command of the romanization scheme for the language. A suitable cataloging background is also essential.

The romanization for some East Asian characters has changed over the years. Also, some characters have more than one way of interpretation (reading) depending on how and when they are used. For example, the Chinese character for "happiness" was romanized originally as *lo* but as *le* in recent years; the same character is romanized as *le* when it means "happiness" but as *yüeh* when it means "music." In addition, as Morimoto and Seely duly point out, Japanese romanization and word division practices have also changed since 1983.[15,16] For example, the original romanization meaning "Japan" was *Nippon*, but it is *Nihon* now. The need to search under alternative romanization and readings to ensure a better hit rate cannot be emphasized enough.

Hyphenated and unhyphenated names present yet another challenge for a successful match. Chinese names, for instance, call for a good knowledge of the romanization rules and the ability to recognize when the data element in question is a name or just a descriptive word.

Works may be republished in different countries. For example, when a Chinese work is republished in Japan, the same title may sometimes be given an added Japanese reading. Chinese and Japanese are both written in ideographs, but the ideographs may be romanized according to the rules for Chinese or those for Japanese. Depending on the judgment of the cataloger, a Japanese reprint of a Chinese original may be romanized according to either the rules for Chinese or those for Japanese. To retrieve cataloging copy, the searcher must try both possibilities.

Familiarity with the characteristics of East Asian cataloging done prior to 1958 is required, because these records were produced when a common cataloging standard for East Asian materials did not exist. They often contain a transcription of bibliographic data from anywhere in the work without the use of brackets to indicate their origin.

Understanding of the pre-ISBD rules and earlier practices for East Asian bibliographic description is important, since the cataloging data on an East Asian card may be in a long running paragraph without conventional punctuation to separate the descriptive areas. Brackets are sometimes used to show that particular data were taken from someplace other than the title page or title page substitute.

A large number of East Asian catalog cards have handwritten data in the body of the entry, series statement, and notes. Because of the different styles of handwriting used, they are hard to read, and the book itself must often be checked. This slows down the retrospective conversion process and adds to its cost.

It is important that the specifications for a recon project include additional instructions to deal with certain notes found on East Asian catalog cards, for example, "Title romanized," "Series romanized," and also series tracings.[17]

Scarcity of vendors able to do East Asian recon has made the market less competitive. This may be one of the factors contributing to higher pricing for East Asian recon than for non-East Asian recon. The unit cost for East Asian recon is almost three times higher than for a non-East Asian recon record.[18]

Only four vendors in North America are currently in the business of full East Asian recon: Asian Shared Information & Access, in Alhambra, California; Fastcat-Asia of the Pacific Automation Systems Services, in Beverly Hills, California; OCLC, in Dublin, Ohio; and Retro Link Associates, in Provo, Utah.

SOLUTIONS TO THE PROBLEMS OF EAST ASIAN RECON

As advocated by many authors of recon articles, a well-thought-out plan and an acceptable recon standard, together with support and commitment from the institution, will ensure a smoother project and a greater rate of success. This is also true for an East Asian recon project.

The problem of overly high standards for recon was addressed by the former RLG East Asian Cataloging Subcommittee. The subcommittee prepared a set of guidelines and in 1991 submitted them for approval to RLG's East Asian Program Committee (RLG's program committee structure has since been replaced by focused task forces of limited duration). These guidelines identified the basic elements to be included for three different levels of fullness for an East Asian recon record in RLIN. They can be easily adopted for use in an East Asian recon project either in-house or under contract by any vendor.

Local circumstances will determine whether recon is done in-house or under contract. A library that decides to embark on East Asian recon needs to have a set of written specifications precise and illustrative enough to "provide a consistency and accuracy in converting the records."[19] These specifications need to address the level of record quality that is acceptable to the library (with measures for quality) and how much authority control is to be exercised on the access points.

As the classic advice goes, reading the literature and surveying other libraries that have experience in East Asian recon will undoubtedly provide good insight into how to avoid or handle particular problems, as well as valuable information on the performance of a potential vendor. For a new vendor who has not been in the market long enough to establish a track record, a tryout conversion of fifty to one hundred catalog cards (done by the vendor according to the library's written specifications) will markedly help determine the ability of the vendor to do the job. The sample must be carefully chosen to include different East Asian languages and to reflect different levels of bibliographic complexity and difficulty, e.g., uniform titles, corporate versus personal name entries, social science versus literary works.

Tsiang and Hart have both stressed the need to know the characteristics and capabilities of one's local online system.[20,21] By knowing how one's local online system works and what it can provide, one can avoid unnecessary frustration or duplication of effort. For example, if the local online catalog will be able to provide authority control using LC authority tapes in the next year or two, it may be worthwhile to limit authority checking in the recon project to cases such as an entry name that has a different form elsewhere on the card or a name that can be read in more than one way.

For in-house recon, adequate training of recon staff with immediate proper follow-up pays off. Depending on the experience of the staff, a concise briefing on historical practices and old cataloging rules may be sufficient. A basic rule of thumb is to convert data unless there is an obvious conflict in the information that appears on the card.

During times of budget constraints and money shortage, a joint effort in grant applications and shared distribution of the recon load (by subject or by imprint date) could be advantageous. In 1991, a joint project involving ten RLG libraries was funded.[22] Five University of California campuses are currently planning to apply for a California State Library grant to do East Asian recon.[23]

CONCLUSION

Despite all the problems that might have inhibited more East Asian libraries from engaging in recon of their cataloging records, the success and struggle of those who have converted their records definitely serve as encouragement and a remedy to those who have not. Most of the libraries surveyed did indicate that they were planning to convert as soon as funding became available.

As more and more East Asian cataloging records are converted into machine-readable form, the cost of East Asian recon will inevitably decline due to the increasing number of records available for matching in the source databases. This will, in turn, encourage and enable more East Asian libraries, both in North America and overseas, to convert their cataloging records. With these earlier records converted and online, East Asian libraries will further join the ranks of other libraries in making bibliographic information widely available to the scholarly community by electronic means.

REFERENCES AND NOTES

1. Brian Schottlaender, "Introduction," *Cataloging & Classification Quarterly* 14, nos. 3–4 (1992): 1.
2. Amy Ching-fen Tsiang, "Retrospective Conversion of East Asian Materials," *Committee on East Asian Libraries Bulletin* 91 (Oct. 1990): 10–15.
3. Tsiang "Retrospective Conversion," augmented by two surveys conducted by the author electronically via RLG's EAMEMNET in February 1993 and the Committee on East Asian Libraries' EASTLIB List in April 1993.
4. "Current Status of East Asian Collections in American Libraries, 1991/1992," *Committee on East Asian Libraries Bulletin* 98 (Feb. 1993): 35–36.
5. Daphne C. Hseuh, "Recon Road Maps: Retrospective Conversion Literature, 1980–1990," *Cataloging & Classification Quarterly* 14, nos. 3–4 (1992): 5–22.
6. Tsiang "Retrospective Conversion."
7. Ai-Hwa Wu, "Experiences and Problems in Retrospective Conversion of East Asian Language Materials," *Committee on East Asian Libraries Bulletin* 98 (Feb. 1993): 3–7.
8. "Survey on CJK Recon Projects/Plans," RLG's Conferral Bulletin Board, Notice 462 (Jan. 1993).
9. Doris Seely, "Minnesota's East Asian Recon Project," RLG's Conferral Bulletin Board, Notice 476 (Feb. 1993).
10. Jane Cheng and Holly Fisher, "Recon Project at the Nelson-Atkins Museum of Art," RLG's Conferral Bulletin Board, Notice 483 (Mar. 1993).
11. Seely, "Minnesota's East Asian Recon Project."
12. Notices 483 and 328, respectively.
13. Amy Hart, "Operation Cleanup: The Problem Resolution Phase of a Retrospective Conversion Project," *Library Resources & Technical Services* 32, no. 4 (Oct. 1988): 378–86.
14. *Cataloging & Classification Quarterly* 14, nos. 3–4 (1992).
15. Response from Hideyuki Morimoto, University of Iowa, to the author's April 1993 survey via EASTLIB.
16. Seely, "Minnesota's East Asian Recon Project."
17. Response from Kengo Yasumatsu, University of Hawaii, to "Survey on CJK Recon Projects/Plans."
18. Wu, "Experiences and Problems in Retrospective Conversion," 4.
19. Ibid., 6.
20. Tsiang, "Retrospective Conversion," 14–15.
21. Hart, "Operation Cleanup," 379–80.
22. "The Research Libraries Group Receives $190,000 NEH Grant for East Asian Retrospective Cataloging Project," *Technical Services Quarterly* 9, no. 1 (1991): 83–84.
23. Response from Phyllis Wang, University of California at Davis, to the author's April 1993 survey via EASTLIB. ■■

Tutorial

The Net Result: Enthusiasm for Exploring the Internet

Constance L. Foster, Cynthia Etkin, Elaine E. Moore, Sandra L. Staebell, and Peggy Wright

What began as a vision and committee charge from the dean of libraries at Western Kentucky University (WKU) ended as a successful faculty retreat focusing on the Internet. For the third year all University Libraries faculty from public and technical services, special collections, and the Kentucky Museum set aside one day before the beginning of the fall semester to retreat from their normal library responsibilities and environment to explore, in-depth, a single topic, and in 1993 the retreat focused on the Internet.

TEACHING THE INTERNET TO LIBRARIANS

The role of academic librarians as information professionals in using Internet resources and identifying key pathways for users is still in its infant stages. Despite warnings that one cannot teach the Internet and that people must learn it by doing, the retreat planning committee looked for ways to present complex ideas to library faculty who had varying degrees of computer expertise and diverse attitudes about computers in the workplace. The ultimate goal was to generate enthusiasm for using the Internet. Although previous retreats had featured "outside" speakers, the committee members felt reasonably secure in featuring themselves as leaders this year. Besides, it was a great opportunity to hone or develop Internet skills. The committee was composed of the supervisors of reference services, government services, and serials, as well as the research instruction coordinator and museum registrar/collections curator. The approach that the committee took to introduce the Internet to librarians might also be used with support staff and patrons.

A pre-workshop survey of those who planned to attend the retreat indicated that most participants had experience with online catalogs and OCLC or other online databases. Five of twenty-two respondents indicated that they were very familiar with computers, while seventeen felt somewhat familiar with computers.

One challenge in preparing for the Internet workshop was that the campus had yet to become connected directly to the Internet, connection being slated for the fall of 1993. Accounts could be obtained through the Kentucky Educational Computing Network (KECNET), allowing access to the University of Kentucky's gateway to the Internet. Consequently, each of the committee planners had to be assigned an account and a password for practicing and for developing workable exercises. In anticipation of the direct connection to Internet, the retreat committee plunged headlong into hours of meetings, individual practice sessions, and more meetings to refine the agenda; create exercises, bibliographies, and glossaries; and rehearse the presentation. The total time frame for this entire venture was four weeks.

PRELIMINARY TRAINING PLAN

The research instruction coordinator encouraged the committee to develop a logo and theme that could be used for general faculty sessions. The idea evolved for an Internet backboard and a basketball net into which went global balls with key Internet acronyms: FTP, WWW, TCP/IP, WAIS, Archie, VERONICA, and Gopher (see figure 1). As

Constance L. Foster is Serials Supervisor, Cynthia Etkin is Government Services Supervisor, Elaine E. Moore is Reference Services Supervisor, Sandra L. Staebell is Kentucky Museum Registrar and Collections Curator, and Peggy Wright is Research Instruction Coordinator at the Helm-Cravens Library, Western Kentucky University, Bowling Green. For additional information contact Cynthia Etkin (etkin@wkyuvm.bitnet), Constance Foster (rfosterc@wkyuvm.bitnet), or Elaine E. Moore (moore@ wkyuvm.bitnet).

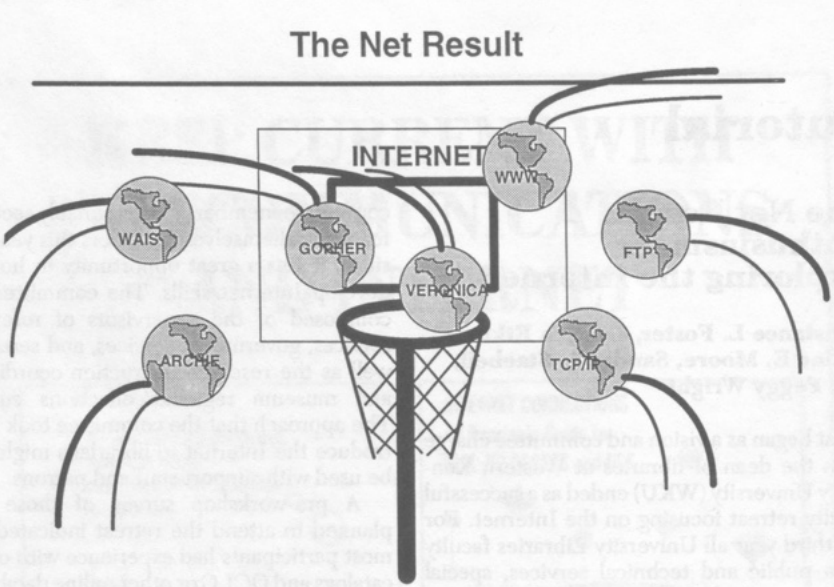


Figure 1. Workshop Logo.

ideas evolved for presentations, so did an abundance of analogies to explain and define the intricacies of the Internet.

THE RETREAT

The retreat opened with comments from the dean of libraries, followed by an update on campus Internet progress from the associate vice-president for finance and administration. The associate vice-president indicated that a direct connection operating through the University of Kentucky in the fall semester would mean Internet access for all faculty. This update was promising news for everyone who recognized the essential role that the Internet has in today's world and today's libraries.

Classroom Presentations: Analogy and Image

The morning session began with an introduction by the committee chair, the reference services supervisor, who piqued everyone's interest with a display of kudzu, a furry gopher, and Archie comic books in front of the lectern. She outlined the agenda, identified handouts in the folders with special reference to the evaluation form that would help determine future workshop topics and strategies, and clearly stated what each person should be able to accomplish by the end of the day: (1) log onto the Internet; (2) use basic Internet navigational tools to find various pieces of information; (3) search other library catalogs; (4) search for book reviews; (5) search for manuscript collections; and (6) explore information sources from a variety of educational, scientific, and government databases. Participants' folders contained an agenda, a summary of the pre-workshop survey on computer expertise, handouts to accompany various presentations, a bibliography of Internet books, a special list of items on reserve in WKU's main library, a glossary, and a keyboard template for the terminals used in the afternoon hands-on session at one of the campus computing labs.

The next presenter, the government services supervisor, explained the what, how, where, and why of the Internet: what it is and does, how it evolved, where it is going, and why librarians need it. She then cited a few statistics on its phenomenal growth (like kudzu) and shared some historical background. Articles in the workshop folder supplemented this overview. A transparency of an eighteenth-century Kentucky road map

compared with one published in 1993 illustrated the complexity of travel and the need for using current guides to plan for driving trips, a perfect analogy for navigating the Internet.

An explanation of the Internet address system (the Domain Name System) followed. The serials supervisor noted in similar fashion the pitfalls a traveler would face if a local travel agent designed a route by using a 1970 road map for a trip and prepared the total itinerary in one bulky document instead of a flip-chart format featuring small segments that eventually carry the traveler to the destination of choice. She also highlighted the importance of how information travels across the Internet, the basic pieces of an Internet address, and how a user can recognize from what domain the information comes.

Throughout the morning session the speakers relied on many forms of audiovisuals: a chalkboard, transparencies, slides, and an actual log-on demo using a laptop computer and a projection display panel. Equally important was the intent to keep all discussions and explanations as free of acronyms as possible. When an acronym was used, the speaker was required to explain it. Keeping everything simple and presenting concepts slowly and carefully while holding everyone's attention were fundamental goals of the retreat committee.

The committee had decided early in its planning not to introduce downloading files to local workstations or to dwell on WAIS or the World-Wide Web. Those topics could wait for other workshops. Also, since electronic mail and discussion groups had been discussed the previous year, the focus this year was strictly on accessing information on the Internet. To succeed with this objective, the museum registrar, a bona fide Internet novice, assumed the most difficult and longest part of the presentation. Her ability to understand concepts and then explain them to the audience was a key motivational element.

Her presentation unfolded the mysteries of remote log-in, file transfer protocol, Gopher, VERONICA, and Archie. The comparison of basic Internet navigational tools to those required by a handy person who uses different tools to install carpet, fix plumbing, or wire a house helped participants understand what happens when a user initiates a search. A by-product of this segment was proof that lack of Internet experience was not a barrier to exploring the information universe.

A break after the last formal presentation allowed the workshop leaders to set up the equipment for subsequent online demonstrations through KECNET. Speakers gave demonstrations that used the navigational tools of the Internet and illustrated possible uses of the Internet in the daily work environment. Book reviews, library catalogs, manuscript collections, and other sources would become integral parts of exercises for the afternoon hands-on session in the computing lab. A slide summary, narrated by the research instruction coordinator, served as a wrap-up of what had been covered in the past three hours.

Hands-On Session

The afternoon hands-on session took place in a computing lab in order to give participants a chance to explore the Internet firsthand. In retrospect, with the knowledge that seventeen participants and five committee members had to share seven computers, the committee should have split the afternoon into two practice sessions. Due to the shortage of workstations, participants had to wait for a chance to work through exercises. Some returned to practice further when computers became available, and everyone remained congenial and patient. The committee members' reward came in watching people practice, succeed, come back for more, invent their own searches, and succeed again.

Feedback

The evaluations turned in after the retreat reflected seventeen responses (100 percent). The very positive assessments of the workshop actually overwhelmed the committee. Thirteen rated the workshop as excellent overall; four as good. Twelve people agreed that additional workshops were necessary. Comments ranged from suggestions about narrowing the focus of the exercises to more library catalogs to a definite plea for continuous updating of a "vigorously growing system."

The committee chair summarized these evaluation results and sent them to all library faculty along with a list of the retreat planners who will serve as liaisons for their various constituencies. In this way all librarians have a contact and consultant when the Internet

becomes part of their daily routines. Already some librarians are spending time on personal computers, signing on via KECNET, and traveling down the Internet highway.

CONCLUSION

The committee identified four key factors as contributing to the positive response to this workshop: (1) maintaining a high level of enthusiasm during all stages of the retreat; (2) using an abundance of analogies that allowed for a ready grasp of concepts; (3) keeping acronyms at a minimum and striving for an uncomplicated path of instruction; and (4) reinforcing the examples and comments of the other presenters to give a thread of unity to the agenda. These factors brought about a real change in attitude toward computer use and hardware barriers. Before the workshop began, one participant commented, "This better be good." No doubt the committee members felt that they had achieved the net result when this same person later complimented the committee for its program and expressed interest in the potential of the Internet.

Whether a person chooses to travel, surf, navigate, or explore cyberspace, the basic tools of Internet access can be used with a minimum of frustration. Last year's theme of the American Library Association challenged librarians to empower people by providing access to information. A librarian's ease and expertise with Internet tools and enthusiasm for Internet resources are essential parts of that empowerment. Analogy, image, creativity, and colleagues' support can create a climate conducive to successful Internet training. ■■

News and Announcements

Notes of RLIN Record Task Force Discussions Meeting at ALA, June 27, 1993—Teleconference Number 3, July 7, 1993

Participating were Ross Atkinson, Cornell; Paul Conway, Yale; Christina Perkins Meyer, Minnesota; Anne Myers, Boston University Law; Betsey Patterson, Emory; Diane Perushek, University of Tennessee; Julia Van Haaften, NYPL; Robin Wendler, Harvard; Robert Wolven, Columbia; Karen Smith-Yoshimura, RLG. Myers and Van Haaften were absent from the ALA meeting.

At ALA we reviewed, as time permitted, the issues and questions raised by members' e-mail submissions on the bibliographic requirements for records loaded into RLIN from a functional perspective and barriers that need to be overcome, realizing that some outstanding questions may need wider consultation. During the July 7, 1993, teleconference, we reviewed the notes of the ALA meeting and discussed issues we had not covered.

Provisional or Precataloging Records

Knowing what item has been ordered or acquired but not yet cataloged is important for collection development. With the emphasis on speed in technical processing services, it is also valuable to have a source record at the point of ordering. However, such records need to be clearly marked as "on order" or "not processed," so that they are not used as the basis for ILL requests.

Most of the discussion focused on the costs and feasibility of loading precataloging records from local systems. Besides the bibliographic information, would information such as date ordered, date received, and possibly vendor (if not confidential) be output? How do local systems export archival processing information? Can local systems easily differentiate between cataloging and on-order or received items that are not cataloged? Would precataloging records still be useful without distinguishing between "on order" and "received"? Do all local systems have the capability of outputting MARC (or MARC-like) records for such items, or would the overhead be too costly? How would this information be represented in RLIN records? (Would we need another RTYP?) Would RLG be able to handle the increased volume if both precataloging and cataloging records were sent?

A quick poll around the table indicated that the combined cataloging and acquisitions records that may be output from a site may range from about 150 titles to 60,000 titles (including maintenance and recon) a week. Ultimately, loading precataloging records into RLIN depends on (1) a critical mass of contributors of such records who agree that sharing such information is valuable, (2) the ability of local systems to send such records in a standardized format, (3) the ability of RLG to identify these records as precataloging records and set CC and RTYP values accordingly and to retain order/receipt information, preferably without requiring the creation of an ACQ segment. Because of the perceived constraints for all libraries to contribute on-order records, it was deemed unlikely that RLIN could ever serve as a centralized ordering agency.

Outstanding Issue. Feasibility and identification of precataloging records for loading into RLIN, and whether a new definition of "RTYP a" records or even a new RTYP value would be needed.

Preservation

The task-force members reviewed what would be required to make RLIN serve as "an international bibliographic database of record for preserved materials, regardless of the format of the materials." We agreed that it was important to include the (1) full bibliographic information for preserved materials (one medium to another), including 007s and 533 as appropriate; (2) location of printing negatives (or access instructions for image files)—Paul Conway reported that a small, informal survey showed that 50 percent of printing masters currently cannot be found; (3) action taken

(e.g., reformatting, digitization, deacidification) noted in the 583 field of the record for the original. There was less agreement about requiring that the nature of the reformatting project be specified (e.g., the funding agency for the project), since it is not clear whether all funding is meant to be public and whether it is necessary to record the funding support on the national record rather than the local record. The intention to reformat (queue) onto film, into digital image, or onto paper in the 583 field was considered highly desirable.

The use of RLIN to support cooperative preservation efforts is especially important to the RLG PRESERV service. The task force thought it unlikely that RLG could require the inclusion of the above information for RLIN libraries not participating in RLG's PRESERV service.

Recommendation. Refer to RLG's PRESERV service participants whether the inclusion of information about reformatting, conservation, and intentions in records participants load into RLIN should become part of the PRESERV service policy.

Outstanding Issue. Resolution on handling multiple versions of a single item (probably linked to the handling of USMARC holdings, see below).

ILL and End-User Access

The value of recording restrictions on materials for ILL was acknowledged, but we agreed that such information may be more appropriate as part of an "ILL profile" and should be referred to SHARES participants to consider as part of the planned ILL workstation development.

Generally, the use of RLIN records for ILL and resource sharing demands considerably more detailed information than other functions, such as cataloging and collection development. More and more frequently, end users approach their online environment as the computer equivalent of browsing the stacks, or they look for items using unconventional approaches. For end-users, being able to search records with a wide variety of access points becomes essential.

Minimal-level cataloging records can restrict searching to such a degree that ILL staff and end-users may view such records as problematic, not helpful.

Standards. With the tension between the need to create cataloging records to full standard and pressures to catalog as many titles as fast as possible, it is unlikely we could ever require only standard records be loaded into RLIN—almost any source record is better than none. It may be more important to mandate that the encoding level accurately reflect the level of cataloging, which also indicates whether cataloging was done with the item in hand. We agreed that RLG should continue to accept cataloging records regardless of the standard used, but the encoding level should be accurately represented.

Recommendation. RLG should mandate accurate EL values. The current practice of loading records with errors and flagging them, rather than not loading them, should be continued.

Holdings and Other Local Data

Holdings information, especially for serials, is important for both ILL and preservation (including 533 $m). RLG's and RLIN libraries local systems' implementation of USMARC holdings needs to be in sync. (Refer to the results of the survey *USMARC Holdings in RLIN Records*, posted on RLIN-L on June 14, 1993.) We discussed the possible problems of eliciting and supporting contributions when so much copy and volume maintenance is done locally, and we considered whether detailed holdings information is really necessary in an international database.

Recommendations. Summary holdings information for serials is sufficient for RLIN records (work needs to be done to see if summaries can be derived from elements of local systems' detailed holdings information).

RLG should not delete the last record for an item. (We need to consider how to show that the item represented by the record is no longer held and remains as a "resource" record only.)

Since detailed information about items may always be available only in a local system, the task force also recommended that RLG consider as a possible future enhancement providing Z39.50 client functionality that would allow users to go from a record in the RLIN database to the holding institution's local OPAC (e.g., a referred search via ISBN, ISSN, local system record number, etc.).

Barriers to Be Overcome

Of primary concern are how to define policies that can be supported in an automated way and how to ensure as much as possible that RLG can support the full range and volume of data expected. Any additional overhead or costs introduced to contribute records would likely hinder, if not preclude, record contributions. We also need to document the benefits of contributing records and address constraints that may be imposed by third-party vendors.

Hit rates may still be a consideration for libraries choosing data resources other than RLIN for source copy. Recent studies seem to indicate that RLIN hit rates are no longer a major consideration, but there may still be a negative perception dating from when there was a holdup in loading NOTIS records.

RLG staff are particularly interested in hit-rate studies, especially those that identify specific areas in which RLIN hit rates need to be improved, so that we can be more proactive in strengthening those areas. If anyone has such information, please send it to Karen Smith-Yoshimura at bl.kss@rlg.stanford.edu. One constraint that was mentioned about using RLIN was that not all VIM records that have already been sent to RLG for loading have been loaded. Brigham Young's VIM records have recently been loaded; Karen Smith-Yoshimura agreed to pursue the loading of any others. Chris Meyer and Anne Myers suggested that identifying AACR2 records on the PRI and MUL displays could help technical processing staff more quickly identify the best source record in a cluster; Karen Smith-Yoshimura agreed to submit a change request for this.

The task force will be discussing possible pricing structures for FY95 and what criteria should be used for basing possible rebates or credits for record contributions submitted by tape or FTP. RLG currently loads records in the ongoing dataloads list the weekend after they are received; RLG needs to publicize its commitment to loading these ongoing dataloads in a timely fashion and the conditions that affect the turnaround time for loading initial files.

Recommendation. Include turnaround times for ongoing record loads as part of the RLG Service Level Statement and in the monthly reports include the success on meeting the stated levels, as with all other service levels.

Actions to be Discussed at the Next Teleconference

Action 1. Karen Smith-Yoshimura will distribute information on current and planned RLIN record-loading capabilities.

Action 2. Task force members are to draft two surveys: one for all RLIN institutions, to elicit their expectations of the use of the RLIN database within their local environments (a conceptual view of the RLIN database by its users), and one for local systems experts, in a subset of RLIN institutions, who can advise on their local systems' capabilities, so that the task force can more accurately predict the feasibility of any recommendations to expedite record contributions.—*Notes by Karen Smith-Yoshimura, June 30, 1993, revised, July 14, 1993.*

SOFTWARE AG and Professional Sales Solutions Sign Marketing/ Distribution Agreement

SOFTWARE AG has entered a partnership with Professional Sales Solutions (PSS), which will gain exclusive third-party distribution rights to resell SOFTWARE AG's Tapestry product set. PSS will also resell and market SOFTWARE AG's CON-NECT, REVIEW, NATURAL DOCUMENT MANAGEMENT, ADABAS TEXT RETRIEVAL SYSTEM, SUPERNATURAL, and COMPLETE.

PSS, under this agreement, will have sole responsibility for the development and support of the Tapestry Library System. Tapestry is a collection of advanced tools that assist librarians in activities such as archiving, online cataloging, circulation, searching, and inventory.

House Committee Approves Networking Applications Bill

On June 30, the House Science, Space and Technology Committee approved Representative Rick Boucher's (D-Va.) networking applications bill. It was ordered reported (H.Rept. 103-173) and placed on the Union

Calendar on July 13. Now called the National Information Infrastructure Act, H.R. 1757 would define the National Research and Education Network as a program; establish network connections, training, and research programs for schools and libraries; and authorize networking applications in the areas of education, health care, libraries, and government information. On June 30, the ALA Council passed a resolution supporting this bill.

LC Announces LC MARVEL

The Library of Congress Machine-Assisted Realization of the Virtual Electronic Library (LC MARVEL) is now available over the Internet. LC MARVEL uses Gopher software and is therefore most easily accessed from another Gopher server or by using a PC-based Gopher client.

Point to: marvel.loc.gov PORT 70 140.147.2.15

Although direct telnet access is available, only ten simultaneous external connections will be supported initially. We recommend accessing LC MARVEL through other Gopher servers or by using a PC-based Gopher client, because no usage restrictions are imposed. For direct connection, Telnet to *marvel.loc.gov* and log on as *marvel*.

The goal of LC MARVEL is to serve as the Campus-Wide Information System for the Library of Congress staff and additionally to offer service to the U.S. Congress and constituents throughout the world. Although LC MARVEL is still experimental, it is now being made publicly available over the Internet. Please address all comments and reports of any technical problems experienced when using the system to lcmarvel@seq1.loc.gov.

Examples of Information Available from LC Marvel

Anyone interested in government information will find a wide variety of material available from LC MARVEL. A quick search turned up the following:

- Congressional phone and fax numbers
- Congressional committee assignments
- Presidential documents
- White House press releases
- Supreme Court decisions
- Text of Clinton administration's FY93 budget proposal (by section)
- FCC documents
- Listing of federal bulletin-board systems
- Articles from the *Library of Congress Information Bulletin Catalog of Federal Domestic Assistance*, searchable by keyword
- *Federal Register* (excerpts only; full text available only by subscription)

Since LC MARVEL is a Gopher system, articles can be mailed directly to your e-mail address.

ALAWON Back Issues Renamed; Selected ALA Council Resolutions Now Available Online

The back issues of *ALAWON* available from the listserv have been renamed. The file name for all issues is now ALAWON; the file type is still the volume and issue number. For example, *ALAWON*, volume 2, number 1, is now called ALAWON 201. ALA Council documents sponsored by the Committee on Legislation are also available from the listserv. The following text is from the file RES-93 INDEX:

> In 1993, the ALA Council passed 13 resolutions sponsored by the Committee on Legislation related to federal legislation. The full text of the resolutions can be retrieved by sending a message to listserv@uicvm (Bitnet) or listserv@uicvm.uic.edu (Internet) with the command *get filename filetype* in the body of the message. For example, to retrieve the Resolution on GPO appropriations passed during the 1993 Midwinter Conference, send the command *get res-m93 gpo-app*. Each resolution has its Council Document number at the bottom of the text. Council Document no. 27.2 was not a resolution and is not in this list. Questions concerning the resolutions can be sent to the ALA Washington Office at alawash@alawash.org.

To retrieve the full list of documents available from the ALA Washington Office, send the command *get ala-wo filelist* to listserv@uicvm (Bitnet) or listserv@uicvm.uic.edu (Internet).

ALAWON is available free of charge and is available only in electronic form. To subscribe, send the message "subscribe ala-wo [your name]" to listserv@uicvm (Bitnet) or listserv@uicvm.uic.edu (Internet). Back issues of *ALAWON* are available from the list server. To find out what's available, send the

ALA Council Document on the Listserv

File Name	File Type	Full Title
RES-M93	GPO-APP	Resolution on GPO appropriations (1992–93 Council Document #27.1)
RES-M93	GOV-INFO	Resolution on the right to federal government information (1992–93 Council Document #27.3)
RES-M93	INT-LIB	Resolution on the possible closing of the Natural Resources Library of the Department of the Interior (1992–93 Council Document #27.4)
RES-A93	GPOACCES	Resolution on Public Law 103-40 Government Printing Office Electronic Information Access Enhancement Act of 1993 (1992–93 Council Document #27.5)
RES-A93	PTOCDROM	Resolution urging the Patent and Trademark Office to provide government information on CD-ROMs to depository libraries (1992–93 Council Document #27.6)
RES-A93	UNESCO	Resolution urging UNESCO to reconsider recent changes in their depository program (1992–93 Council Document #27.7)
RES-A93	CLASINFO	Resolution regarding the classification of U.S. Government information (1992–93 Council Document #27.8)
RES-A93	DODBENEF	Resolution on portability of benefits for non-appropriated fund employees who converted to the civil service system (1992–93 Council Document #27.9)
RES-A93	NHPRC	Resolution in support of the National Historical Publications and Records Commission (1992–93 Council Document #27.10)
RES-A93	LC-APP	Resolution on the restoration of funding for the Library of Congress (1992–93 Council Document #27.11)
RES-A93	GPO-APP	Resolution on FY94 Government Printing Office appropriations (1992–93 Council Document #27.12)
RES-A93	NETAPPS	Resolution on networking applications legislation (1992–93 Council Document #27.13)
RES-A93	OMB-A130	Resolution on the revision of OMB Circular A-130 (1992–93 Council Document #27.14)

message *send ala-wo filelist* to the listserv. The ALA-WO file list contains the list of files with the exact file name and file type. To get a particular file, issue the command *send filename filetype* to the listserv.

RLG to Add British Library's New Table-of-Contents Database to CitaDel

The Research Libraries Group (RLG) will mount the British Library's new table-of-contents database on CitaDel, RLG's citation and document delivery service. Called Inside Information, the new database will be launched by the British Library Document Supply Centre (BLDSC) in the United Kingdom during July and will be available on CitaDel as of September 1. Inside Information provides author, title, and journal citation information for articles appearing in 10,000 of the most requested titles in BLDSC's collection of 50,000 of the world's principal journals and magazines.

As of September 1, Inside Information will contain approximately 900,000 citations from October 1992 onward, with BLDSC estimating that more than 1 million citations will be added to the file every year. Daily updating is planned.

Full text of every article cited will be available by document delivery through CitaDel. Delivery methods include airmail, fax, and RLG's Ariel (document transmission software for the Internet).

Inside Information will be available free to institutions purchasing annual subscriptions to RLIN (the Research Libraries Information

Network). For institutions not purchasing RLIN subscriptions, Inside Information will be available for a yearly fee as a stand-alone CitaDel file. (RLG member institutions automatically receive a 5 percent discount on all subscription fees.)

Institutions purchasing annual subscriptions to RLIN or Inside Information will receive a free copy of the Ariel software.

For more information, please contact the RLIN Information Center at 1-800-537-7546; e-mail bl.ric@rlg.bitnet or bl.ric@rlg.stanford.edu (Internet); fax (415) 964-0943.

NEH Awards Funds to Continue RLG Chinese Rare-Book Project

The National Endowment for the Humanities has awarded the Research Libraries Group $404,536 to support the fourth phase of its highly successful cooperative project to establish an international union catalog of Chinese rare books on RLIN. This brings to $683,000 the total contributions NEH has made over the past five years to this RLG project.

The funds will enable six North American libraries that participated in phase three of this project to continue cataloging their rare Chinese materials and will allow four North American libraries that were unable to participate in phase three to contribute cataloging records of their rare Chinese holdings. (A list of these libraries appears below.)

The grant will also support the continued participation of Peking University and the Chinese Academy of Sciences, both located in Beijing, and the Liaoning Provincial Library in Shenyang. Other China-based libraries may contribute records of their rare materials to the project if funds permit.

Staff assigned to the project by the participating libraries will attend training sessions before phase four commences in the fall of 1993. All catalog records generated by the thirteen libraries will be submitted to the project's central editorial office, located at Princeton University, where they will be coordinated, emended, and entered into RLIN according to guidelines and standards established in the project's earlier phases.

A significant by-product will be the publication of the project's cataloging guidelines, a document that standardizes the descriptive cataloging of Chinese rare books and will be of particular interest to libraries following North American library practices.

North American institutions participating in phase four are Columbia University, the New York Public Library, Princeton University, the University of British Columbia, the University of California at Berkeley, the University of Chicago, all of which also participated in phase three; and the Freer Gallery of Art/Arthur M. Sackler Gallery, the University of Minnesota, the University of Pennsylvania, and the University of Toronto, who are new participants as of phase four. All are members of the Research Libraries Group.

For further information, please contact Karen Smith-Yoshimura, RLG director of library support and services, at (415) 691-2270, E-mail bl.kss@rlg.bitnet or bl.kss@rlg.stanford.edu.

WLN to Provide Full Service Internet Connection for Seattle Public Library

The Seattle Public Library and the Western Libraries Network (WLN) have signed an agreement under which WLN will provide Internet services for the library. WLN will supply a TCP/IP Internet connection beginning June 1993. WLN will also provide other services, including electronic-mail service for the library staff on the WLN Internet server.

Humanities Scholars Urged to Take Action Influencing New Information Technologies

The creation of a ten million-volume digital library is among the actions advocated by leaders of the scholarly community in the *Summary of Proceedings* of *Technology, Scholarship, and the Humanities: The Implications of Electronic Information*, recently released by the Getty Art History Information Program (AHIP) and mailed to 10,000 decision-makers whose policies affect the humanities. The report highlights the results of a groundbreaking national conference organized by AHIP and the American Council of Learned Societies and cosponsored by leading national organizations: the Coalition for Networked Information, the Council on Library Resources, and the Research Libraries Group.

The conference, held in Irvine, California, in fall 1992, convened an unusually varied

group of distinguished scholars, academic administrators, technologists, librarians, and leaders of national institutions and learned societies from around the nation to discuss how current trends in information technology will affect the humanities. They concluded with plans of action and calls for humanities scholars to participate more aggressively in information technology developments that bear upon education and research. In order to involve greater numbers of humanities scholars, the publication of the *Summary of Proceedings* was supported by a grant from the Andrew W. Mellon Foundation to the American Council of Learned Societies.

The *Summary of Proceedings* presents far-reaching recommendations for national action, such as calls to do the following:

1. Initiate a national collaborative effort to pursue an active advocacy role for the humanities in today's rapidly expanding electronic environment and, working with existing advocacy organizations, enter the current dialogue, both inside and outside the academy, on the development and direction of new information technologies to serve the humanities.

2. Promote, as a national priority, the creation of a ten million-volume digital library, broadly conceived to encompass the full spectrum of humanities research collections.

3. Identify and develop exemplary collaborative programs and projects that demonstrate the effective creation, sharing, and distribution of electronic information among institutions, organizations, and individuals in the humanities.

4. On the individual, disciplinary, and institutional levels, collaborate within and outside the humanities in the development of standards for exchange of, access to, and description and preservation of electronic research.

5. Investigate how the humanities can use information technology to increase, reallocate, examine, and generate resources in new ways.

6. Adjust the current definition of scholarly research and instruction to reward innovative uses of electronic information and media.

7. Enlist humanities scholars to interpret the impact of information technology on society and promote critical understanding of the role that information technology can play in both research and teaching.

8. Sponsor initiatives—workshops, fellowships, and summer institutes—that provide opportunities for training and that enrich the mixture of information technology and the humanities.

The *Summary of Proceedings* also includes excerpts from the keynote address delivered by Vartan Gregorian, president of Brown University and former president of the New York Public Library; summary reports from conference working groups; and resumes of position papers commissioned from five distinguished experts, which served as a basis for working group discussion.

To receive additional information about the conference, to order the *Summary of Proceedings*, or to receive a brochure outlining the recommendations of conference participants, please contact Susan Siegfried, the Getty Art History Information Program, 401 Wilshire Blvd., Suite 1100, Santa Monica, CA 90401-1455; (310) 451-6366; fax (310) 451-5570; e-mail enq91ss@uclamvs (Bitnet) or enq9Iss@mvs.oac.ucla.edu (Internet).

Full texts of the keynote address, papers, and *Summary of Proceedings* are available through the Internet at ftp.cni.org; select the directory \CNI\documents\tech.schol.human\papers [for the papers] and \CNI\documents\tech.schol.human [for the summary]. ■■

Recent Publications

Book Reviews

The FISCAL Directory of Fee-Based Research and Document Supply Services. 4th ed. Chicago: American Library Assn., 1993. 445p. paper, $65 (ISBN 0-8389-2161-3).

The *FISCAL Directory*, produced by the Fee-Based Information Service Centers in Academic Libraries Discussion Group of ACRL, "provides detailed information on over 445 fee-based information services available from major research universities; public, special, and corporate libraries; and selected commercial firms from around the world" (p. xi).

The "Service Profiles" are arranged alphabetically by institution, one page for each profile. Information includes the service name, address, phone, fax, e-mail, Internet, and Telex numbers and addresses; time zone in which the service is located; hours to Greenwich Mean Time (GMT); year founded; name of a contact person; and a one-paragraph narrative description of the service. The profile pages are well designed and easy to read. Seven indexes facilitate use.

The introduction and user guide is informative, but the *FISCAL Directory* is so well designed that most users won't need it. All of the pages except the soft covers are perforated and three-hole punched so that users can arrange—or rearrange—the *Directory* in their own binders, which nearly disqualifies it as a tool for public reference or circulating collections.

The data were gathered from ten thousand surveys conducted between December 1991 and May 1992. The services included met two criteria: they provide fee-based research or document delivery and they provide services to the public at large. *FISCAL* was composed using dBase IV and DB Publisher.

The *FISCAL Directory* appears to be directed to businesses and researchers who do not have access to the full range of information services offered by most academic, many public, and some special libraries. It also serves corporate librarians with access to information or databases outside their corporate interests. For those who operate fee-based services, it is a guide to their colleagues—and competitors.

Sixty-one of the profiles describe commercial document delivery services. Many of the small but growing number of new, fee-based searching and research services developed in U.S. academic and large public libraries are included. Almost none of the many long-established archival or genealogical research services are entered, however, and the publication is biased toward "modern," technology-based services.

In some areas, the information in the directory is scanty or potentially misleading. For example, the "Research Areas Index" contains only thirty U.S. entries for "Patent Research," and only one U.S. Patent Depository Library is listed in the "Areas of Special Expertise" index. Most states have at least one full-service patent depository library, and it is difficult to consider why one should consult this directory for assistance in researching U.S. patents. The "Services Offered Index" lists seventy-four entries for "Research Services, Faculty or Expert Consultation," missing thousands of small consulting firms and faculty who perform these services.

At least some of the information is available in other sources libraries may have or use. In fact, the committee identified potential survey candidates by consulting other directories, some of which are named in the introduction. Myriad existing directories, guides, compendia, etc., provide comprehensive coverage of the universe of information resources without reference to fees.

In the fast-paced world of fee-based services, fees change. The newest data in this work are now one and one-half years old, and a small, nonrandom sampling uncovered many fee changes. This volume covers a topic for which a directory is needed, but *FISCAL* is a partial list, not a directory. Library acquisitions budgets are reeling, and ALA com-

mittees that collect and disseminate information ought to find more creative—and library-friendly—methods to perform these tasks that they have chosen. *FISCAL* should be on a Gopher, not on a reference shelf. This fourth edition is recommended only for libraries that require a single list of fee-based services and recognize that it is far from comprehensive.—*John Webb, Washington State University.* ■■

Franklin, Janice R. *Database Ownership and Copyright Issues among Automated Library Networks: An Analysis and Case Study.* Norwood, N.J.: Ablex, 1993. 181p. $45 (ISBN 0-89391-752-4); paper, $24.50 (ISBN 1-56750-016-1).

This is required reading for anyone interested in database ownership and copyright of large, automated library networks. Unlike most writings on library networks, which have taken a developmental or technical stance, this analysis and case study provides an objective view of library networks from a sociological perspective. Using network newsletters, reports, and network conferences, Franklin has studied the relationship between two major network players—OCLC and the Southeastern Library Network (SOLINET). Both entities are shown to have a vested interest in ownership of bibliographic records and copyright of a compilation of records. The study of relationships between players is used to expose the strengths and weaknesses of the networks themselves.

The second chapter provides an overview of library cooperation in the context of external factors such as economics, the growth of centralized services, the development of regional networks, and the establishment of the MARC standard for exchanging machine-readable records. As a precursor to discussing the networks in the context of social network theory, chapter 3 is devoted to an overview of various models of network architecture.

The most extensive section of the book is chapter 4—"Database Ownership and Copyright Issues among automated Library Networks." In this section Franklin looks as the changing roles of OCLC and regional networks and the member libraries. She traces the history of negotiations between OCLC and the regional network SOLINET over the issue of copyright of the national database. The granting of a copyright in March 1984 to OCLC is seen as a barrier to access and cooperative information exchange by most of the members of the library networking community. While cooperation has continued between OCLC, libraries, and the regional networks since March 1984, the future of regional networks is seen as volatile—even risky.

Balancing the extensive discussion of OCLC's effort to copyright the national database is chapter 5, a case study of SOLINET that details its rise as a regional network and defines its goals as a nonprofit entity established to benefit its member libraries. Following the descriptive section is an analysis of SOLINET according to social network principles. Central to this analysis is a description of the social environment in which SOLINET evolved and the relationships between players. As in the OCLC study, these factors clearly show why SOLINET took the position it did with regard to the copyrighting of bibliographic records in the union catalog.

Franklin concludes that the conflict between OCLC and SOLINET over the issue of database ownership and copyright can be viewed in a broader context as a positive development since it served to redefine relationships and improve communication and has resulted in a "balance of power at the national level." In an epilogue, Franklin suggests a number of additional library network issues that would benefit from social network analysis. It is hoped that the author will turn her expert analysis to these topics so that we may better understand library networking and, as a consequence, plan better for the future.—*Dawn Talbot, University of California, San Diego.* ■■

Gill, Suzanne L. *File Management and Information Retrieval Systems: A Manual for Managers and Technicians.* 3d ed. Englewood, Colo.: Libraries Unlimited, 1993. 267p. $29.50 (ISBN 1-56308-050-8).

Despite the growing use of computer equipment by businesses, paper remains the dominant medium for recording and storing business information, according to Gill. An established records and information management program can ensure that paper is man-

aged efficiently for successful retrieval and provide for maintenance and disposal of information according to corporate needs, legal guidelines, and security requirements.

This text provides a practical introduction to records management. The emphasis is on developing "the skills necessary to create a procedures manual while learning file organization theory" (p. xv). An introductory section covers the background of records and information management and the history of business records. The major section of the book covers the steps involved in preparing a procedures manual. This section provides general information and covers classification methods (including extensive treatment of different filing methods and how to handle problems in filing), processing materials, retention of records, circulation procedures, equipment and supplies, and the choices involved in deciding between centralization and decentralization.

Paperless files are discussed in chapters on micrographics, computers, and optical disks. A section on case histories discusses five representative companies, including the types of records found and procedures developed for each. Chapters conclude with exercises for the student, with answers provided at the end of the book.

The book has been updated from the second edition with some additional text in each chapter and a section added on disaster planning. The section on paperless files contains new exercises and a new case history. Some illustrations have been replaced by photographs of newer equipment, and a number of additional illustrations have been included. Gill states that the edition is intended for those requiring an efficient file-management system in companies employing fewer than two hundred people.

In the preface, Gill points out the similarities between libraries and records centers, noting that skills taught in library science programs are transferable to record- and information-management programs. This book would be useful in either setting.—*Nancy B. Olson, Mankato State University.* ■■

Harris, Michael A., and Stan A. Hannah. *Into the Future: The Foundations of Library and Information Services in the Post-Industrial Era.* Norwood, N.J.: Ablex, 1993. 182p. paper, $19.95 (ISBN 1-56750-01503).

The purpose of this book is to survey and assess critically the large base of interdisciplinary literature on the "information society" and to rally librarians and information scientists to participate in and expand on the prior arguments regarding the mission of the profession in the postindustrial era. The authors review "the historical, economic, political, and cultural forces that are influencing the restructuring of library and information services in the post-industrial era, and . . . the complex political economy, or landscape, of the information age."

There are six chapters and an extensive bibliography as well as author and subject indexes. In chapters 1 and 2 the reader is brought up to speed on the historical arguments and debates that have occurred on the topic since the 1960s. A major area of focus in these chapters and throughout the book is Daniel Bell's work, although the works of numerous other authors are highlighted and compared and contrasted to Bell's writings.

Chapter 3 addresses the economic and political forces of the information age and the complex issue of a national information policy in which there is an inevitable tension between "information" as a commodity and free access or the democratization of access. This discussion is brought closer to home in chapter 4, where a warning is sounded for all would-be information professionals to confront the " 'schizoid dilemma' between the notion of 'open communication' and 'knowledge as property.' " Chapter 4 goes on to summarize the changes that would be necessary to form a new professional identity and to articulate the issue that most of these changes contradict long-standing practices of traditional librarianship.

The implications of the postindustrial era on the workplace and the nature of work are examined in chapter 5. In the final chapter the authors share their own rethinking of the impact on librarians of the postindustrial landscape and the changing mission of the library.

Into the Future: The Foundations of Library and Information Services in the Post-Industrial Era is far from light reading and requires determination on the part of the reader. As a critical analysis of the literature, the text is full of references, quotes, and opinions from numerous writers from over a

thirty-year span that are compared and contrasted. At times *Into the Future* is difficult to follow, and through most of the text one is never quite sure where the authors stand on issues, since the emphasis is on providing the reader with background from other writers. However, this is not to say it is not worth the effort to forge through the book. *Into the Future* raises important philosophic and economic issues of the postindustrial era that represent the changing nature of the "information marketplace" in the twenty-first century. We cannot afford to ignore these and the impact on the profession.—*Bernice Ginder, Rutgers, The State University of New Jersey.* ■■

Higginbotham, Barbra Buckner, and Sally Bowdoin. *Access Versus Assets: A Comprehensive Guide to Resource Sharing for Academic Librarians.* Frontiers of Access Series, no. 1. Chicago: American Library Assn., 1993. 399p. $60; $54 ALA members (ISBN 0-8389-0607-9).

This 399-page monograph is an impressive compendium of quotations: 584 of them in nine chapters and an appendix. While there is not a great deal of original thinking presented, the book does describe and explore a broad range of approaches academic librarians can use to provide their readers with access to information that exists outside the walls of the local library.

The focus is on "access," which the authors define simply as "access to that which one does not own." There is little reference to "assets," but it is clear that the term is used to mean locally held resources.

Access Versus Assets identifies the cost centers associated with each approach (copyright and other fees, staff time, equipment, communication charges), as well as the advantages and disadvantages of each method. The authors suggest how librarians can evaluate and assign costs to individual approaches to access such as interlending and commercial document supply. No actual costs are quoted. The authors do not directly concern themselves with cost-reduction techniques.

Topics that relate more directly to ownership than to access (coordinated collection development, storage centers, CD-ROM, or other electronic full-text data available in the local library) are treated only briefly, if at all.

Among the strongest chapters in the book is chapter 2, "A Point of Embarkation." The authors have woven together sixty quotations to describe the demise of library self-sufficiency, the impact of technology, the hidden costs of ownership, and the viability of resource sharing. They stress that even the largest research institutions can purchase only 5 to 10 percent of the world's publishing output. The authors argue that the two situations that have most strongly influenced the shift of emphasis from ownership to access are the spiraling costs of journal subscriptions and the decline of the American dollar in the foreign markets where many scholarly periodicals are published. They argue further that paper-based periodical indexes never created the demand for journal literature that CD-ROMs have generated and that demand for journal literature is rising rapidly with the introduction of CD-ROM products.

Also very useful, but only temporarily, is an appendix of case studies in resource sharing. Examples of thirty shared catalogs, electronic document-transmission projects, reciprocal borrowing agreements, and document-delivery programs are examined, including such recent ones as the OhioLINK effort to link a large number of systems and create an online union catalog.

Although much of the book's material is available elsewhere, the volume pulls together in one place information from scores of books, journals, and unpublished consulting reports. The book is recommended for the professional staffs of college and university libraries and for faculty and students in library schools who want the convenience of having a great deal of information assembled in an attractively designed and printed monograph.—*Richard W. Boss, Information Systems Consultants Inc.* ■■

Interlibrary Loan Of Alternative Format Materials: A Balanced Sourcebook. Ed. Bruce E. Massis. New York: Haworth, 1992. 196p. $21.95 (ISBN 1-56024-394-5). (Also published as *Journal of Interlibrary Loan and Information Supply*, vol. 3, nos. 1 and 2, 1992).

This work combines the practical utility of an international directory of libraries lending alternate-format library materials with a collection of essays detailing the practice and

philosophy of international interlibrary lending. This directory will be essential for interlibrary loan departments serving blind or visually handicapped patrons, particularly if they are non-native speakers of English.

Recognizing the dual trends of growing demand for alternate-format materials (recordings, braille, and large print) by patrons with print-related disabilities and increasing multilingual populations, the IFLA (International Federation of Library Agencies) section of Libraries for the Blind initiated an international survey in 1990 of 260 libraries known to provide services to blind and visually impaired patrons. Respondents from libraries in twenty-nine countries answered questions about tools such as national union catalogs of alternate-format materials, borrowing trends both within their country and internationally, and finally, whether there exists one library designated or willing to serve as a clearinghouse for international requests. The Section of Libraries for the Blind also coordinated its data-gathering efforts with simultaneous work being done by the Conference of Directors of National Libraries (CDNL). Data from the CDNL's 1991 survey are reported and add to the coverage of the scope of international lending.

The responses provided an enlightening survey of the state of the art in international lending of nonprint library materials. While some problems are unique to a country's political or economic position, other problems are familiar to the interlending arena: the lending-borrowing imbalance, copyright issues, media format variations, delivery system speed and costs, and the issue of the value of media. Coeditors Massis and Vitzanksky "expect this publication to act as a catalyst" for the creation of international standards and initiatives to overcome the problems highlighted.

The essays include a practical discussion of problems of lending and borrowing audiovisual materials; a discussion of the sale, loan, and gift and exchange programs at the National Library Service for the Blind and Physically Handicapped in the United States; and an excellent essay by Graham P. Cornish on "The Philosophy Behind International Interlending and Its Implications for the Visually Handicapped." The volume concludes with sample interlibrary loan forms from various countries and the IFLA form and instructions for sending and receiving international interlending requests. The editors are to be commended for this balanced sourcebook of practical data and philosophical foundation.—*Kathleen M. O'Connor, Gonzaga University.* ■■

McNulty, Tom, and Dawn M. Suvino. *Access to Information: Materials, Technologies, and Services for Print-Impaired Readers.* Chicago: American Library Assn., 1993. 162p. paper, $28; $25.50 LITA members. (ISBN 0-8389-7641-7).

The authors define print-impaired persons as individuals who are literate and are blind or visually impaired or have a learning disability. Those with physical or mobility disabilities that inhibit reading are excluded from coverage. The authors' stated intent is "to increase the reader's awareness of the variety of systems available and the advantages and disadvantages of each" as well as "to help unravel the paradox of how best to integrate the print-impaired individual into an environment where integration depends so profoundly on access to print documents."

The book's scope places limits on its utility. Its definition of print-impairment describes only a portion of the disabling conditions that inhibit access to printed information as defined in the Americans with Disabilities Act or as found in any given population. Most libraries are responsible for providing access to populations with a much broader range of disabling conditions.

The descriptions of current technologies and devices designed to assist the visually impaired are clear, and the assessments thoughtful and useful. Within its stated scope, the book offers a comprehensive snapshot in time. The final chapter describes an existing accessible online catalog and two proposed model programs. These are good examples of applications of the technologies and devices designed to make print materials accessible and to integrate the print-impaired. These portions of the book successfully fulfill the author's stated intent.

Other material presented is less supportive of the central focus. A chapter is devoted to a crash course in computer literacy, and considerable attention is given to the history and development of the various nonprint formats and technologies. These are presented

first by type of format (tactile or audio), then by one of the three categories of print-impairment as defined by the authors (individuals with low vision). Much of this information is not essential to understanding the formats or technologies described and does not contribute substantially to the reader's awareness of what is available or to a library's ability to integrate the print-impaired. Much is also available from a variety of other resources.—*Donna Dziedzic, New York Public Library.* ■■

Saffady, William. *Electronic Document Imaging Systems: Design, Evaluation, and Implementation.* Westport, Conn.: Meckler, 1993. 183p. paper, $40 (ISBN 0-89736-840-9).

With the current emphasis on adding full-text databases to the well-dressed library menu, a good book on document imaging seems especially pertinent for librarians and other information professionals. This reviewer found Saffady's coverage of document imaging systems to be concise and readable, providing sufficient detail to enable the reader to tackle a complex document imaging project. Saffady defines the scope of this book as covering systems that record and store document images in electronic formats on optical or magnetic media for reference purposes. "Image" is the key word here, as it is important to remember that imaging systems described in the book store documents as images, not character-code text with which we are most familiar in library applications. Preserving the original appearance of documents is the foremost concern of true imaging systems.

Primary chapter topics are input, storage, and output. An introductory chapter includes an overview of the development of imaging systems, and a concluding chapter enumerates some of the management and operations issues inherent with imaging systems.

A lengthy discussion on indexing issues is included in the chapter on input. Saffady points out that while much literature is available on indexing issues related to journal articles and publications, very little exists on indexing needs for business documents. What follows is a detailed description of indexing parameters, values, subject term selection, vocabulary control, and depth and consistency. Saffady concludes the indexing section with references to studies on the performance differences of various levels of indexing personnel. The author cites studies that compare the results of indexing prepared by subject specialists and professional indexers to that prepared by less experienced and trained staff. Don't worry if your indexing data entry is performed by clerical staff; studies show that the results should be satisfactory!

Input devices are covered in detail, with scanning speed, resolution, device design, and digitization modes among the topics addressed. Indexing gets more coverage in this chapter, with descriptions of key-entry and bar-code entry of indexing values. Autoindexing, as bar-code entry is known, is described in detail. Optical and magnetic storage devices are also described and compared in some detail. The formulas provided for determining space requirements should be helpful.

The "Issues" chapter covers media stability, legal issues concerning digitized documents, system integration concerns, and the all-important cost determinations for both startup and ongoing support.

Generous bibliographies are provided, and a cost calculation worksheet should help in the planning and evaluation process. Sixteen and one-half pages of hardware, software, and turnkey system vendors, service bureaus, and systems integrators should be enough to get anyone started on a document imaging project. This title will be well worth the purchase price if implementing a document imaging system appears on your to-do list.—*Janet Woody, Virginia Commonwealth University.* ■■

Woodsworth, Anne, and James F. Williams. *Managing the Economics of Owning, Leasing, and Contracting Out Information Services.* Brookfield, Vt.: Ashgate, 1993. 204p. $49.95 (ISBN 1-85742-018-7).

This book is designed to help decision makers and managers think about the best ways to use human, technological, and informational resources, taking into account the new paradigms for traditional library and information services that already exist. This is, I believe, one of the first books to explore the subject of how a library administrator can

evaluate the alternatives of leasing or contracting for information services as ways of delivering some or all library services. Chapter titles include "In-house Information Services," "Information Workers," "Resources, Services and Delivery," "Information Technologies," "Shared or Cooperative Routes," "Costs and Charging Strategies," and "Shaping Core Services in the Future." The book includes a well-selected bibliography and index.

The subject of outsourcing the provision of information services in an organization is new to many library administrators and managers, but it is not new in business and government. This well-developed trend, together with the technology-driven changes that are sweeping our profession, is changing the basis of many of the decisions we make about funding of and budgeting for information services. Woodsworth and Williams' book is a well-organized and thoroughly researched discussion of the economics of the provision of these services. The authors cover all aspects of the subject, basing their analysis on current information and delivering their thoughts in clear, succinct language.

I particularly like the way the authors have used their impressive group of contributors. At appropriate points in each chapter the authors have inserted lengthy quotations from one or another of the contributors. The quotations are always relevant to the subject being discussed at that point in the chapter and in every case they enhanced my understanding of the material. I found this use of contributors to be very effective. The information provided by the contributors is valuable, and this method of presentation enhances its value and improves the readability of the book.

My only criticism of this book is that its dry-sounding title belies the excitement within. Many of the issues discussed in this book are controversial, e.g., fee-based services in libraries. The excitement for me is in the confident and straightforward approach these authors take with these issues and their ability to communicate that confidence so effectively. My conclusion after reading the book is that Woodsworth and Williams are advocates of clear thinking, and their book helps make clear thinking possible about a wide range of difficult issues.

With this book Woodsworth and Williams have made a significant contribution to the management literature in our profession. Not only do I recommend the purchase of this book, I recommend that it be read by all librarians, especially those who have or aspire to have management responsibilities.—*George Rickerson, University of Missouri System.* ■■

Other Recent Receipts

Anderson, Carol L., and Robert Hauptman. *Technology and Information Services: Challenges for the 1990s*. Norwood, N.J.: Ablex, 1993. 234p. paper, $24.50 (ISBN 1-56750-021-8).

Browne, Mairead. *Organizational Decision Making and Information*. New York: Neal-Schuman, 1993. 272p. paper, $26.95 (ISBN 1-56750-017-X).

Decision Support Systems in Academic Libraries. By Roy Adams and others. New York: Bowker, 1993. 141p. $45 (ISBN 1-85739-047-4).

The Electronic Journal: The Future of Serials-Based Information. Ed. Brian Cook. New York: Haworth, 1993. 106p. $15.95 (ISBN 1-56024-452-6).

European Directory of Software for Libraries and Information Centres. Comp. Joanna Wood. Brookfield, Vt.: Ashgate, 1993. 251p. paper, $69.95 (ISBN 1-85742-092-6).

Format Integration and Its Effect on Cataloging, Training, and Systems. ALCTS Papers on Library Technical Services and Collections, no. 4. Ed. Karen Coyle. Chicago: American Library Assn., 1993. 100p. paper, $20; $18 ALA members (ISBN 0-8389-3432-3).

If We Build It: Scholarly Communications and Networking Technologies: Proceedings of the North American Serials Interest Group, Inc. Ed. Suzanne McMahon, Miriam Palm and Pam Dunn. New York; Haworth, 1993. 326p. $24.95 (ISBN 1-56024-450-X).

Information Policy: A Framework for Evaluation and Policy Research. Norwood, N.J.: Ablex, 1993. 193p. $45 (ISBN 0-89391-890-3); paper, $23.50 (ISBN 1-56750-018-8).

Intner, Sheila S. *Interfaces: Relationships between Library Technical and Public Services*. Englewood, Colo.: Libraries Unlimited, 1993. 231p. $32 (ISBN 1-56308-059-1).

Journal of Information Networking. Published three times per year. London: Taylor Graham. Vol. 1, 1993– . Subscription rate: U.S. $107; $51.50 to Institute of Information Scientists members (ISSN 0966-9248).

Journal of Document and Text Management. Published three times per year. London: Taylor Graham. Vol. 1, 1993– . Subscription rate: U.S. $103/year; $51.50 to Institute of Information Scientists members (ISSN 0969-9325).

Letters

To the Editor:

I'd like to clarify two minor points in Fae Hamilton's excellent review of *The Online Catalog Book* (*ITAL*, June 1993):

1. The "inexplicable" variation between three-to-a-page screen captures and two-to-a-page displays was purely economic. The book could not exceed 560 pages without raising the price substantially; I used three-to-a-page displays in some cases in order to tell the whole story.

2. Plans for annual updating and continued publication of *The Catalog Collection* were always posited on achieving satisfactory sales levels and were always somewhat tentative.

In fact, sales of *The Catalog Collection* have not been sufficient to warrant updating or ongoing publication; the title will probably be out of print by the time this letter appears.—*Walt Crawford, RLG.* ■■

To the Editor:

In a recent paper, Bryce Allen (*ITAL*, June 1993, p. 203–8) reported on a study to "determine how changing the format of browse displays might affect information retrieval performance for end users." In the study, many searchers were given the same search problem and conducted their searches on an experimental subject headings list that some searchers could browse as a single-alphabet list of all terms regardless of level and that other searchers could browse by selecting the desired hierarchic level (i.e., by expanding from a higher-level term to see the terms at the next-lower level, or by contracting from a lower-level term to see terms at the next-higher level in the subject list). The study obviously entailed a great deal of organization and effort to conduct, and its topic is important for rational improvement of library systems. Regrettably, because of inconsistencies in the research design and ambiguities in the use of terminology, it is not possible to accept that the conclusions stated by the author have been established by the data he reported.

The summary conclusion given in the paper is that the study showed that searching with the hierarchical type of display improved efficiency but did not affect effectiveness. Nowhere in the paper are the terms *efficiency* and *effectiveness* defined; however, one can infer intended meanings from the context.

With respect to the assertion that efficiency improved, one can see from the data that the only clear difference between the searches with the two different displays was that, on average, only about half as many headings were scanned (i.e., displayed) in the hierarchical searches as in the single-alphabet searches. One infers that this is what was meant by improved efficiency. But there is ambiguity here, because the rationale for experimenting with hierarchical display that was introduced early in the paper was not to reduce the number of *headings* scanned, but rather to reduce the *time* required to find appropriate headings. The average search times reported in Allen's study did not decrease with hierarchical search and may even have increased somewhat. With such findings, did the study really demonstrate an increase in efficiency? Apparently not, even though an increase in efficiency would surely have been expected.

The improved efficiency to be gained from hierarchical searching was recognized and exploited back in the dark ages of manual file systems (card catalogs with divider cards; correspondence files with folder tabs) and had been explored theoretically as well (e.g., B. A. Lipetz and C. T. Song, *Journal of the American Society for Information Science* 21, no. 2, [Mar. 1970]: 140–1). Barring the occurrence of certain types of collateral differences, a hierarchical file should be faster for a person to search for a desired term than a single-alphabet file. The failure of Allen's study to show this is almost certainly attributable to weakness in research methodology. In addition, weakness in methodology made it virtually a foregone conclusion that the study would find no difference in "effectiveness" (which, although not defined, appears to

mean the average number of headings, or the average number of highly relevant headings, retrieved in the search).

The difficulties with methodology are found mainly in the study's ambiguity in defining for the searchers what is meant by a completed search. The searchers were first told to try to find at least two subject headings that were related to the topic they had just read something about. They were then "also instructed that once they had found the two subject headings required, they could continue to search the interface for additional headings if they wanted. They were told that they should govern the amount of this additional searching activity by their perceptions of how much they normally would search an index of this sort when preparing a term paper." In other words, although the study was supposed to detect differences in searching behavior with the experimental displays, the searchers were pretty clearly instructed, or at least encouraged, *not* to deviate from their usual behavior. The searchers were given no hint that speed in completing the search task was desirable; rather, they could easily infer that they were expected to spend about as much time as usual on the search effort and to retrieve about as many headings as usual from the search—that is, to maintain their usual effectiveness. There is support in the literature for the readiness of users to maintain customary practices when encountering new systems; it was observed, for example, that the average amount of time spent in catalog search sessions did not change when the catalog system was changed from manual to computer (B. A. Lipetz and P. J. Paulson, *Library Trends* 35, no. 4 [Spring 1987]: 597–617). In Allen's study, the average time spent on a search (about 6 to 6.5 minutes) was more than enough to allow most searchers to select the apparently "usual" number of terms (about 9 or 10) from the 747-heading subject file regardless of the type of browsing interface provided; single-alphabet searchers had the time to scan the entire list, and, from the data Allen reported on number of lines scanned, it appears that many of them did just that.

Despite the methodological mismatch between research objective and instructions given to searchers, it might well have been possible with Allen's methodology to detect the expected reduction in search time with hierarchical searching. Instead of measuring only the total time expended in a search, one might have measured the time that a searcher took to select only the first (or the first two or three) headings (or highly relevant headings). With the hierarchical display, searchers would presumably make their initial selections more quickly than single-alphabet searchers and then do more dawdling later in the search session. An alternative way to detect faster searching, involving modification of methodology, would be to simply inform searchers that speed is desirable and omit any instruction or encouragement to maintain normal searching behavior.

Finally, Allen states in his conclusions that larger and more realistic files will be necessary in order to determine whether the hierarchical display promotes more effective searches. Larger files can indeed be more realistic and therefore desirable. But they may not be really necessary. For example, using just the smallish file from Allen's study, the searchers could be asked to find, quickly if possible, as many headings as they considered relevant. But their searches would be cut off after only two or three minutes; this would assure that a searcher could not do an exhaustive brute-force scan of the experimental file, which could be analogous to the situation in "unrestricted" searching of much larger files. The headings actually retrieved during the allowed search time could be studied and compared for quantity and quality ("effectiveness"), with the expectation that the hierarchical display would produce superior results.—*Ben-Ami Lipetz, School of Information Science and Policy, State University of New York at Albany.* ■■

Professor Allen responds:

Professor Ben-Ami Lipetz raises some important points. His suggestion that the time taken to complete various components of a search be examined in detail has a great deal of merit, and I have the data with which to undertake that analysis. If, as he suggests, overall search time is a good measure of search efficiency, then the data I presented in my research note may be taken to suggest that a hierarchical display of subject headings can reduce one kind of cognitive load (the effort required to scan headings), but may replace that cognitive load with another: perhaps the

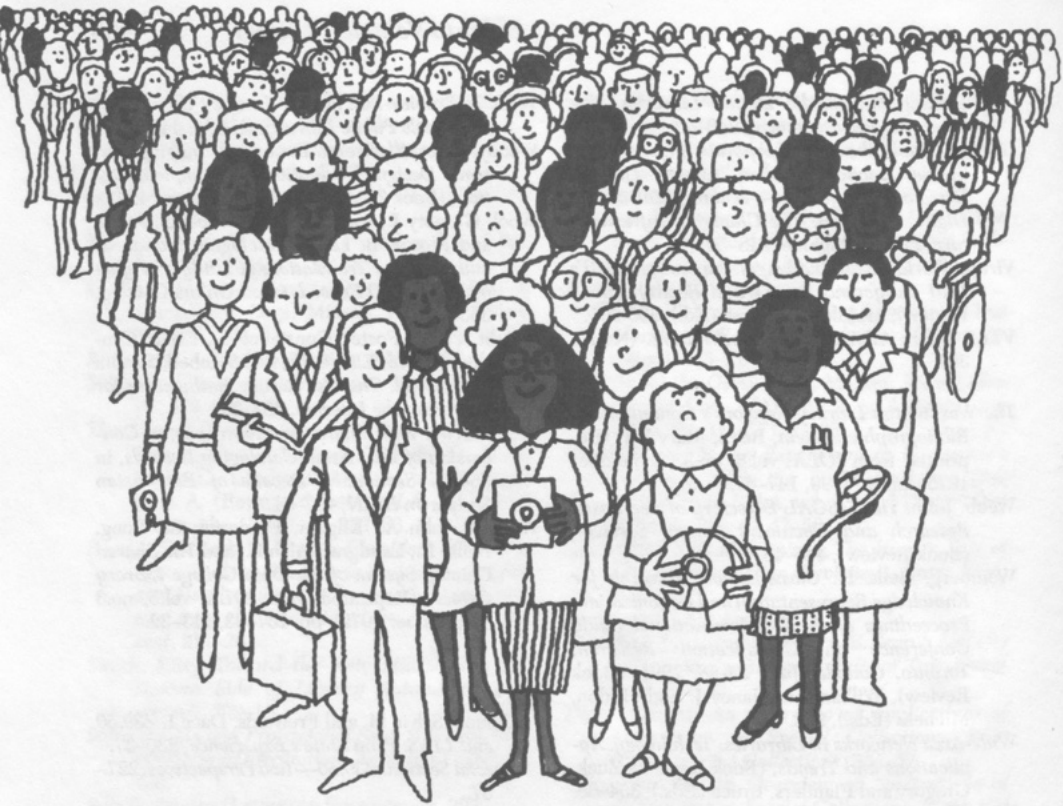
effort needed to navigate up and down the hierarchical list. This is an area that requires additional research, and Lipetz's suggestions are very helpful.

I am less certain about his idea that speed of searching should be emphasized in instructions to searchers. The persistence of "normal behavior" he notes is so prevalent that we might regard the task of the system designer as creating systems that will enhance search performance when searchers behave normally. If a system enhancement produces better search performance only when searchers are instructed to behave in an unusual manner, we may question the contribution of that enhancement in an operational, "normal-user" context. The alternative approach Lipetz suggests of calculating the time users took to find the first two headings seems to be more ecologically valid, and well worth the effort to reanalyze the transaction logs.—*Bryce Allen, Associate Professor, Graduate School of Library and Information Science, University of Illinois, Champaign.* ■■

Information Technology and Libraries Index to Volume 12 (1993)

Compiled by Kieth C. Wright

INDEX TO ADVERTISERS



SETTING STANDARDS IS EVERYONE'S BUSINESS.

Establishing effective technical standards. It makes good business sense, no matter what business you're in. Because clear, cohesive practices and procedures help things run smoothly.

Which is why we need your help.

Who are we? We're NISO. The National Information Standards Organization. We develop technical standards for the entire range of information services and products: online information services, librarianship, bookselling, indexing, and library equipment and supplies. In fact, our standards are an integral part of many work environments.

NISO has developed over fifty information standards. Here are just a few of the standards that benefit you.

- Standard identifiers for books and serials
- Standard format and vocabulary for information sharing
- Standard computer-to-computer protocol linking local, regional, national, and international bibliographic systems
- Standards for permanent and durable book papers and bindings

How can you help us? By helping NISO set the standards in your profession. Because only with the input of successful organizations like yours can we develop proper standards. We also need your help with funding. You see, we depend on the support of our members. Members from all types of organizations including libraries, publishers, large and small companies, and government agencies. There are a range of membership options available; one is right for your organization.

What's in it for you? A chance to influence your profession's direction and growth by setting its standards. A chance to work closely with your peers. A chance to plan for the new standards that will directly affect you.

What should you do now? Call, write, or fax us. We'll tell you more about ourselves, how we can work for you, and how you can help us.

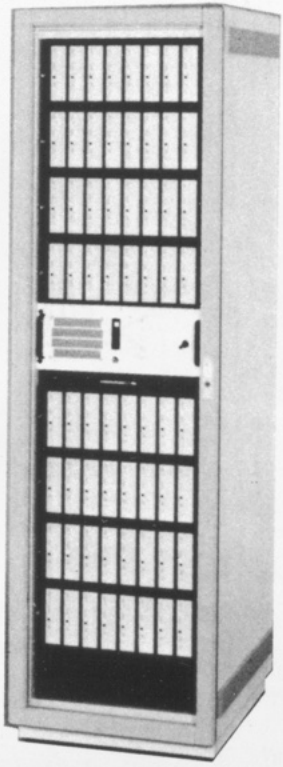
NISO

P.O. Box 1056 • Bethesda, MD 20827
Telephone: 301-975-2814 • Fax: 301-869-8071 • Bitnet: NISO@NBSENH

TODD ENTERPRISES INC.
NETWORK SOLUTIONS

from 4 to 56+
CD-ROM DRIVES in ONE CABINET

Todd massive data retrieval Total Access Systems feature dockable drives and security keylocks. Todd uses Hitachi drive mechanisms which have an access time of under 300 msec and are known for reliability. They have a two-year warranty. The security lock feature of Todd Total Access Systems is useful in public areas. Todd services include hardware design and development, customized systems, and network installation and management. Todd Total Access units will operate within your workstation environment with all administrating software. Call **800 445-TODD** *or (718) 343-1040 for more information.*



Todd's Total Access systems interface with your workstation environment.

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

established
1 9 7 1
Todd Enterprises
over two decades of excellence

TODD ENTERPRISES INC.
224-49 67th Avenue · Bayside, NY 11364
718 343-1040 · 800 445-TODD
FAX 718 343-9180 · BBS 516 829-0212

available on GSA contract

There's More To Know.

Today's patrons expect easy access to information sources via our growing electronic networks. And they want full text and image capabilities. Advances in connectivity, the Z39.50 protocol, and the Internet have created a global network from which to draw information.

There's More In Store.

Truly open systems, which provide access to this global network, are becoming a reality. Open systems utilizing client/server technology offer greater flexibility through the integration of components from multiple vendors. And easy-to-use graphical interfaces (GUIs) are increasing productivity for both patrons and staff. Open systems are reshaping the way we work.

There's More To NOTIS.

NOTIS Systems has been a leader in library automation for over 10 years, and today supports over 200 libraries. We're developing open systems that will help you provide expanded patron access and give you greater system flexibility and cost-effective solutions. NOTIS was the first to offer Z39.50-compliant products like *InfoShare*,™ a database access server for UNIX platforms; *ProPAC*,™ a windows-based client for PCs and workstations; and *PACLink/PACLoan*,™ a collection sharing system. Keep an eye on the future. Keep an eye on NOTIS.

NOTIS SYSTEMS, INC.
AN AMERITECH COMPANY

708.866.0150