journal of
library automation

september, 1981

journal of library automation

CONTENTS

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

Journal of Library Automation 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. The *Journal* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 143 of this volume.

Manuscripts of articles, communications, and news items should be addressed to: Brian Aveney, Editor, *Journal of Library Automation*, Blackwell North America, 10300 S.W. Allen Blvd., Beaverton, OR 97005. Copies of books submitted for review should be addressed to: David L. Weisbrod, *JOLA Book Reviews*, Systems Office, Yale University Library, Box 1603A, Yale Station, New Haven, CT 06520. Advertising arrangements should be made with Judith G. Schmidt, 1408 D St., SE, Washington, DC 20003.

Journal of Library Automation is a perquisite of membership in the Library and Information Technology Association. Subscription price, $7.50, is included in membership dues. Nonmembers may subscribe for $15 per year. Single copies, $4.

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

Publication of material in the *Journal of Library Automation* 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 *Current Contents, Current Index to Journals in Education, Education, Library Literature,* and *Quarterly Bibliography of Computers and Data Processing.* Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

Second-class postage paid at Chicago, Illinois, and at additional mailing offices. *Postmaster: Send address changes to* Journal of Library Automation, *50 E. Huron St., Chicago, IL 60611.*

Who Rules the Rules?

"Why can't the English teach their children how to speak?" wondered Henry Higgins, implying that a lack of widely and consistently followed rules of usage created linguistic backwardness and anarchy. Higgins' question might be rephrased today as: "When will the code teach its founders how to catalog?"

The Library of Congress has historically fitted catalog codes to its own practices rather than following them slavishly. The best example is the lamentable policy of superimposition: continued use of preestablished forms of names that are not in compliance with the *Paris Principles* or *AACR1*. This was a cause of widespread confusion and complaint and the practice was eventually discontinued . . . well, sort of discontinued. The various interpretations of *AACR1*, the inclusion of new rules, and pressure for further modifications eventually led to the drafting of *AACR2*, a code that was supposed to end variance and controversial practices.

One might assume that including LC as a principal author of the new text and an LC official as one of the editors might result in a code that it could actually follow. Judging by the spate of exceptions and interpretations made so far (more than 300), this has not been the case. In the place of superimposition, we have new impositions known as "compatible headings." They may not be readily ascertained according to the rules, but have been granted a sort of bibliographic squatter's rights.

Although it would be simpler for catalogers to follow the rules consistently, they must instead check several *Cataloging Service Bulletins* and *Name Authorities* to see whether LC has determined that a given personal, corporate, or serial name is already "compatible" with *AACR2*. This can result in cataloging delays, higher processing costs, and inconsistent entries. *AACR2* and uncertainties regarding its application by LC have been widely credited with lower cataloging productivity.

This is not to imply that LC is behaving in a strictly arbitrary or capricious manner vis-à-vis the code. They can be seen as caught on the horns of a trilemma, with vast internal needs and increasing external demands competing for a shrinking budget. President Reagan may have whispered sweet nothings during National Library Week, but during budget hearings it became clear that libraries are not as "truly needy" as impoverished generals and interior decorators.

Decisions to depart from *AACR2* have been based primarily on cost factors. The decision by the RTSD Catalog Code Revision Committee and the Joint Steering Committee not to consider cost and implementation factors has led both to widespread opposition to the code resulting in a one-year delay in implementation, and to the modifications that LC has made and is making. Some variations such as using "Dept." for "Depart-

ment" and "House" for "House of Representatives" make fiscal and common sense. Many other LC changes are simply bibliographic nit-picking, minor irritants to catalogers who must flip back and forth between the text of *AACR2* and half a dozen *Bulletins* to settle a minor point of description. Why didn't LC representatives attempt to say, "Wait a minute—we just can't do that now," while the code was being considered rather than after it was published? Anyway, considering that LC was starting up a whole new catalog and closing the old one, one wonders why rules not to be applied retrospectively had to be tinkered with to such an extent.

Major questions still to be resolved include not only the compatible-name quandary, but the treatment of serials, microform reproductions, establishment of corporate names and determination of when works "emanate from" corporate bodies, and the romanization of Slavic names.

The decision to use title entry for serials and monographic series even in the case of generic titles has been controversial. There are, of course, exceptions to the rules, and there will be differences in how uncertain catalogers construct complex entries with parenthetical modifiers. Unfortunately, rules establishing entries for serials have sometimes been muddied rather than clarified in the *Bulletin*. Consider the example in the Winter 1981 issue wherein the bulletin of the Engineering Station of West Virginia University is entered under "Bulletin," while the same publication for the entire university is entered under "West Virginia University Bulletin." Also, consider the complex cross-reference structure required to direct users between the two files, both of which may well be split again, historically, between author/title and title main entry. This is a special problem in the case of large monographic series generated by corporate bodies.

The LC position on microform reproductions of previously published works is clearer, but is still a point of controversy. They have decided to provide the imprint and collation (er, make that "Publication, distribution, etc., area" and "Physical description area") of the original work, with a description of the microform in a note. In other words, they're sticking to *AACR1*. The RTSD CCS Committee on Cataloging: Description and Access is currently trying to resolve this conflict, one in which many research libraries have sided with LC. This body is also trying to unravel the mystique of "corporate emanation" introduced in *AACR2*.

Another sore point has been the LC decision to follow an alternative rule, which prefers commonly known forms of romanized names over those established via systematic romanization. That LC is correctly following the spirit of the general principle for personal names is little comfort to research libraries with large Slavic collections.

How are other libraries responding to the murky form of *AACR2*? Some are closing old card catalogs and continuing them with COM or temporary card supplements. Some of these are establishing cross-reference links between variant forms of names between catalogs, while others are not.

Some are keeping their catalogs open and shifting files, while others are splitting files. Some are shifting some files and splitting others. *AACR2* was intended to provide headings that could be easily ascertained by the user. Ironically, the temporary result is scrambled catalogs: access systems involving multiple lookups and built-in confusion. Until most bibliographic records are in machine-readable form under reliable authority control this will continue to be the case. Authority control, it would seem, has long been an idea whose time has come but whose application is yet to be realized.

The cooperative efforts of the Library of Congress and the major bibliographic utilities to establish reliable automated authority control will do much to ameliorate the problems presented by *AACR2*. It would also be helpful if LC, perhaps with the financial assistance of other libraries, networks, and foundations, would publish what might be called *AACR2½*—not a new edition of the code but one accurately reflecting actual LC practice. Finally, future code makers would be wise to consider cost and other implementation factors in their deliberations. Professor Higgins, ever the optimist, would rather sing "Wouldn't it be loverly" than hear another verse of "I did it my way."

JAMES R. DWYER

EDITOR'S NOTES

Title Change

It often seems that the only things that change their names as often as library publications are standards organizations. Not to be left out, *JOLA* will be called *Information Technology and Libraries* beginning with Volume 1, Number 1, the March 1982 issue. This name was approved by the LITA Board in San Francisco this June as more accurately reflecting the true scope of the journal.

New Section

With this issue, we are initiating a new section: "Reports and Working Papers." This is intended to help disseminate documents of particular interest to the *JOLA* readership. We solicit suggestions of documents, often developed as working papers for a specific purpose or group but of interest and value to our readership. In general, documents in this section are neither refereed nor edited.

Mitch

I take great personal pleasure in publishing Mike Malinconico's speech upon presenting the 1981 LITA Award to Mitch Freedman.

Readers' Comments

We do continue to solicit suggestions about the journal but receive few. Is anybody reading it? If you have any thoughts about what we should or shouldn't do, we would welcome your sharing them.

The British Library's Approach to AACR2*

Lynne BRINDLEY: British Library, Bibliographic Services Division, London, England.

The formal commitment of the British Library to AACR2 *and* Dewey 19 *entailed substantial changes to the U.K. MARC format, the* BLAISE *Filing Rules, and a variety of products produced for the British Library itself and for other libraries, including the* British National Bibliography. *The British Library file conversion involved not only headings but also algorithmic conversion of the descriptive cataloguing.*

Along with the U.S. Library of Congress and the national libraries of Australia and Canada, the British Library was formally committed to the adoption of the *Anglo-American Cataloguing Rules, Second Edition (AACR2)* and *Decimal Classification, 19th Edition (DC19)* in 1981. This entailed fairly substantial changes to the MARC format as published in the U.K. *MARC Manual, 2nd Edition* as well as the implementation of the new and more sophisticated *BLAISE* (British Library Automated Information Service) *Filing Rules.*[1]

There is, of course, never an ideal time for making major changes—politically, economically, or technically; and the Bibliographic Services Division (BSD) found itself having a large number of preexisting separate systems, particularly for our batch processing work, which had grown up over a long period of time and had in most cases been tailor-made to the individual products. Whilst relatively small, BSD is nonetheless responsible for a multiplicity of products and services, almost all of which were to be affected to some extent by the change to *AACR2/DC19*. Briefly, then, a comment on the different services and the degree to which they were affected, thus setting the scene for our decisions on machine conversion.

---

*Based on a talk given at the Library Association seminar "Library Automation and AACR2," held in London on January 28, 1981. The views expressed in this paper do not necessarily represent those of the British Library or the Bibliographic Services Division.

SERVICES AND IMPACTS
Printed Publications

The major printed publication of the division is the *British National Bibliography*. It is arguable that for the printed publications (especially the weeklies) there would have been little justification for retrospective conversion. The files could have been cut off at the end of 1980 and started afresh for 1981—it might, however, have precluded, or certainly have made more messy, the possibility of any multiannual cumulations across this period.

Microform Products

These are mostly individual COM catalogues, both within the BL, especially the Reference Division, and externally, provided through LOCAS (BSD's Local Catalogue Service) to some sixty libraries in the U.K. In many ways those libraries that plunged into automation early, building up files of records derived from central U.K. and LC MARC, were likely to be worst affected. Individual machine-readable files had grown very large and exploited not only relatively *current* cataloguing data, but also full *retrospective* U.K. holdings back to 1950. Also we foresaw *no lessening of use* by libraries taking our catalogue service of the U.K. retrospective 1950–80 file after *AACR2* implementation. Therefore the grounds for attempting automatic retrospective conversion of records were indisputable.

Tape Services

U.K. exchange tapes, either as a weekly service or through the Selective Record Service, are supplied to nearly one hundred organisations. The same arguments that there will be continuing selection from the retrospective files apply—therefore, for compatibility and ease of use we needed to consider conversion. The weekly exchange tape service makes a clean *AACR1/AACR2* break, but obviously libraries have back files of *AACR1* records. Mindful of our responsibility to other organisations and agencies utilising our records, we decided to make our own converted tapes of LC and U.K. MARC records available to tape-service customers to aid their own conversions.

Online Services

Regarding the BLAISE Online Information Retrieval System for U.K. and LC MARC, our concern was to ensure continued easy searching and printing across the total span of files. Without automatic conversion it would have been difficult, if not impossible, to ensure consistency in search elements and index entries (e.g.: In U.K. MARC, series fields 400, 410, and 411 no longer exist, so without conversion a searcher would have to remember specific search qualifiers for pre-1981 records, and different ones thereafter). Without conversion the searcher would need a lot more knowl-

edge of MARC and the history of cataloguing practices to formulate effective strategies.

Outside Users of MARC

Last and very much *not* least was a consideration of what we could do to help the now large community of U.K. MARC users in coping with the changeover. This is now a very large and diverse group relying on BSD for the provision of bibliographic records for whatever purpose. Our own conversion enabled us to provide a multiplicity of aids to libraries. Of particular note are (1) U.K. and LC exchange tapes of converted records, and (2) machine-readable and microfiche versions of our own Name Conversion File, which is being used as the basis for the new Name Authority Fiche.

So, in the context of the variety of our services the case for conversion was strong.

RETROSPECTIVE CONVERSION

The extent of the retrospective conversion exercise is discussed below. In conjunction with this work we were faced with the necessity of rationalising our COM and print product software (Library Software Package), both to enable it to drive each of the previously separate print applications and to ensure that it had sufficiently sophisticated output facilities to cope with the complexity of *AACR2*/U.K. *MARC 2* records, with their increase in numbers of subfields, their repeatability, all or some, and varying sequences, to produce the specified layout and punctuation across our services.

Extent of Conversion

We are now in a position to discuss the retrospective conversion exercise. Having decided in principle to become involved with conversion, the extent of our involvement had to be established. British libraries have never had the tradition of building and utilising name authority files, and certainly the concepts fit more easily in the North American primarily online system context rather than in the predominantly batch cataloguing systems established in the U.K. The BL therefore found itself without a machine-readable authority file and began to create one from scratch to enable the important heading changes required by *AACR2* to be handled automatically.

Again because of the overriding importance of COM catalogues in the U.K., considerable attention was paid not only to automatic heading changes but also to automatic MARC coding and text conversions bringing the descriptive cataloguing elements also into line with *AACR2*/U.K. *MARC 2*, so that catalogue records could be consistent on output whether derived from the conversion or newly created.

The third consideration for conversion was our Library of Congress file

(*Books All 1968–*), used in the U.K. as part of our cataloguing services and as a file in the BLAISE online system. We had always performed certain conversions on LC records to bring them more into line structurally with the U.K. MARC format. However, U.K. libraries using these records for cataloguing purposes still had to undertake substantial editing. It was therefore decided to use the opportunity to enhance this conversion and bring LC records into line with U.K. *MARC 2* to make them of maximum use to British librarians.

To summarise, then, the retrospective conversion comprised three main parts:

1. That part which utilised information stored in the *Name Conversion File*, which records the *AACR2* and *AACR1* forms of names. This enabled the automatic conversion of major, commonly occurring personal and corporate headings.

2. *Automatic MARC coding and text conversions*—this consisted of specifications at MARC tag and subfield level of algorithms for automatic MARC coding and some bulk text conversions. It resulted in records being converted to a pseudo–*AACR2*/U.K. *MARC 2 format*, so that all output specifications, whether by profile or by online inversion, had only to cater for the new format.

These two parts of the conversion are inexorably linked, both conceptually and in programming terms, with frequent references to alternative courses of action dependent on whether a match has been found on NCF. The details of conversion are in "Specification for Retrospective Conversion of the UK MARC Files 1950–1980,"[2] prepared in the Computer Services Department.

3. The third facet of conversion was to our *Library of Congress files* (*Books All 1968–*), to bring records in line with U.K. *MARC 2* as far as possible. Only conversions of tags, indicators, subfield marks, punctuation, and order of data elements have been included; no attempt has been made to bring textual data into conformity with BSD practice. The converted records are therefore in *AACR2* form to the extent that LC applies *AACR2* to a particular record.

The next section highlights major points of each part of the conversion, commenting particularly on aspects of programming and testing.

Name Conversion

The Name Conversion File was built up by BSD's Descriptive Cataloguing Section over nine months of 1980 and comprises authenticated *AACR2* headings with the *AACR1* form where different. It will form the basis of an authority file of headings and references for future BSD cataloguing and will be the first publicly available U.K. authority file. The file was maintained using existing LOCAS facilities. Pseudo-MARC records were created recording the *AACR1* and *AACR2* forms of headings in the format shown in example 1.

FIELD

001	(control number)
049	(source code)
110.1	\$a Great Britain \$c Accidents Investigation Branch (Name Heading in AACRII Form)
710.1	\$a Great Britain \$c Department of Trade \$c Accidents Investigation Branch (Name Heading in AACRI Form)
910.1	\$a Great Britain \$c Department of Trade \$c Accidents Investigation Branch \$x See \$a Great Britain \$x Accidents Investigation Branch (Reference for AACRII Name Heading)

Name Conversion File Record

Example 1

The file being used for conversion comprised some 12,000 records, of which 4,000 had *AACR2* heading changes. The remaining records were authenticated by BSD as correct *AACR2* headings without alteration. Of the changed headings most were prolific personal and corporate (particularly U.K. government) headings.

The first stage of the conversion process for U.K. MARC records (1950–80) involved all records being processed against the Name Conversion File to replace *AACR1* with *AACR2* headings and associated references.

In programming terms, the name conversion was relatively easy— relatively, that is, in the context of bibliographic programming. The matching program used was not particularly sophisticated. It took each NCF record, identified the 7xx (*AACR1*) field, created a key of fifty characters stripping out all blanks, embedded punctuation and diacriticals, and then tried to match the key against each 1xx heading in whatever file was being converted. If there was a match on the key, then the program proceeded to match character by character through the data looking for an *exact* match. If this was not found, then the NCF record was not processed. Example 2 shows this procedure more clearly.

Of course, this file has not converted *all AACR1* headings, but it has ensured that the majority of headings *likely to recur* (i.e., of any significance in catalogue collocation of headings) have been automatically changed.

Automatic MARC Coding and Text Conversions

This is commonly known as the format conversion program and forms the bulk of the "Specification for Retrospective Conversion." The original specification was extremely complex, particularly bearing in mind the tight time scales that we were working to. The major difficulty throughout all parts of this facet of conversion was having to specify procedures to accommodate the variety of usage of MARC across thirty years, including previously automatically converted 1950–68 U.K. MARC records; it has

NCF RECORD

710 (AACRI) $a Great Britain $c Civil Service Department
               $c Central Computer Agency#

110 (AACRII) $a Central Computer Agency#

910 (AACRII) $a Great Britain $c Civil Service Department
               $c Central Computer Agency $x See $a Central
               Computer Agency#

KEY: 10$AGREATBRITAIN$CCIVILSERVICEDEPARTMENT$CCENTRALC

Matching on data—would match Central Computer Agency
               would not match Central Cataloguing Agency

N.B. KEY EQUALS 50 CHARACTERS (Upper Case)

NCF RECORD

700 (AACRI) $a Walker $h David Esdaile#

100 (AACRII) $a Walker $h David E. $q David Esdaile $r
               1907 -#

900 (AACRII) $a Walker $h David $c 1907 - $x See $a
               Walker, David E.#

KEY: 10$AWALKER$HDAVIDESDAILE

BOOK RECORD

Before:

100   Walker $h David Esdaile#

900 -

After:

100            $a Walker $h David E. $q David Esdaile
               $r 1907 -#

900            $a Walker $h David $c 1907 - $x See $a
               Walker, David E. $z 100#

N.B.   Addition of new reference

Name Conversion Matching
Example 2

been almost impossible to verify absolutely that any of the automatic changes would cover all cases.

Not surprisingly, this was an extremely complex program. It had to allow for manipulating in fairly precise ways nonstandard and variable data, and had to be designed to cope with occurrences in many different combinations. The programmer had to code for these combinations, some of which may possibly never have been used. It is probably the case that certain combinations do not exist, but this could not be guaranteed over such a large number of records until the total file had been converted.

A good example of the complex logic of this kind of processing is found in the 245 field, where seven complex conditions were allowed for:

FIELD 245

(1)   If $e _____ then _____ else _____

(2)   If $f _____ then _____ else _____

(3)   If $d or $e _____ or _____ or _____
       or _____ or _____ or _____ or _____

then _____
else if $d or $e _____ or _____ or _____
or _____ or _____ or _____ or _____
or _____
then _____
(4) If tags _____ then _____
(5) If 008 _____ and _____ or _____ or _____
then _____
(6) If $h _____ then _____ and _____
(7) If $e _____ then _____
else if first $e _____ then _____
else _____
else _____

Repeat for all levels of 245.

Another variation on this theme is that the specification catered for what it expected to find. Again, because of the volume and span of data the expected was not always found. For example, a lot of processing of references is dependent on the presence of a $x. What do you do when you find a record accidentally without one?

A third problem was that of *interdependency of fields* and subsequent actions. A good example of this is found in 110s and related 910s. If a 110 is changed, you may have to *create* a 910, replace a 910 with another one, or reorganise existing subfields. Then you may have to reorder the field and also flag the action to come back to later in the program. Hence you are switching back and forth across fields throughout the program. You cannot simply start at field one, process sequentially, and then stop. Clearly this makes program testing that much more complicated.

However, those were the problems—really a very small percentage of the whole. From all that has been seen of the converted files so far it has been a highly successful exercise. All of the major MARC changes and many of less significance have been converted automatically by this program—Treaties, Laws, Statutes, Series, Conferences, Multipart works—the resulting records being consistent in MARC tagging structure and in significant headings and areas of text.

Library of Congress File Conversion

It has already been stressed that the automatic MARC coding and text conversions for U.K. MARC were very complex programs. Perhaps even more complicated was the conversion program written to transform LC into U.K. MARC format. The main reason for this is that the U.K. and NCF conversions are one-off programs and a great number of the manipulations could be hard-coded. However, it is intended that the LC conversion program will be used on an ongoing basis against each weekly LC tape. Thus each conversion has been treated as a separate parameter to the

program so that it is general purpose and easily alterable in the light of changes of practice by LC. To give you some idea of the complexity, there are well over 600 separate parameters to the program. I say separate, but in fact they are interrelated parameters, so that if a minor change is made to one it can potentially affect many others.

Many of the problems relating to this program could again only be really apparent in volume testing, not in writing. Each parameter written and tested in isolation was satisfactory, but when they began to be put together in modular form, then the problem of unusual combinations began to show.

Although the conversion parameters for LC records are extensive, they cannot touch the cataloguing data, certainly not nearly as much as in the U.K. MARC conversion. There are added problems in the fact that the records coming to us from LC do not show the clean *AACR1/AACR2* break that BSD is adopting. We are having to allow for mixed records from LC at least in the foreseeable future. Details of the LC–to–U.K. MARC conversion are published in a detailed specification.[3]

COMMON ISSUES IN CONVERSION
Testing

It is possible to draw out common problems applicable across all the conversion work, particularly in testing. They are as follows:

1. Variability of records;
2. Complexity of records;
3. Volume of data;
4. Nonstandard data;
5. Repercussions throughout system.

Variability

This is an obvious problem in the handling of MARC records, but particularly pertinent when trying to do such complex manipulations. The record format itself is of course variable—there are very few essential fields or data elements; most need not be present at all; if they are present, they can be there once or ten times. Standards of cataloguing, and therefore MARC coding, have changed considerably over the period in question, adding to the variability. In some exceptional cases BSD practices are different from those prescribed in the MARC manual, e.g., nonstandard use of title references. All of this results in additional difficulties from specification, through programming and testing. On average we found that one conversion process took two to three times the amount of coding required for more normal computer processing.

Complexity

This is linked with variability and was manifest particularly in the fact that it was extremely difficult to ensure that the programs catered for all

conditions. We found that testing threw up oddities not allowed for in the original specification. In an ideal situation with no time constraints a totally tailored and comprehensive test file should have been drawn up for each facet of conversion. This exercise alone would have taken a good year and would still not have catered for the unexpected data problems. In practice, whilst BSD's Descriptive Cataloguing staff were able to provide several hundred records that tested the majority and most important of the conversions, we always faced the possibility of coming across exceptions. This soon became apparent when volume testing commenced and each new file threw up another combination and a different program route not previously tested.

Volume

The third major factor adding to the complexity of the whole operation was the sheer volume of data to be processed. Approximate figures are as follows:

U.K. MARC	0.7 million records
LC MARC	1.4 million records
LOCAS	2.5 million records

The combination of these three factors—variability, complexity, and volume of data—made testing extremely difficult and expensive in machine terms, in that large test batches of material had to be processed.

Nonstandard Data

Like any large file, U.K. MARC has its share of incorrect data, most of it of no particular significance. However, some problems arose in conversion testing resulting occasionally in corrupted records. One example that springs to mind was the incorrect spelling of months in Treaties, giving problems in the 110 $b conversion to 240.

Repercussions throughout System

A cautionary note, really: we made a decision that postconversion records should not be put back and overwrite existing master files until they had been through validation programs (i.e., those used for validating new input for BNB and LOCAS); it was felt that this was a necessary safeguard against reintroducing any structurally incorrect records postconversion. It was here again that testing threw up timely reminders of just how much the validation programs had been upgraded and changed since many of the original records had been input through the system.

Scheduling

The scheduling of such a large, complex exercise was extremely difficult, with interdependency of processing related to the success or otherwise of overnight runs. A lot of time was spent before the conversion period in

discussion with our computer bureau to ensure maximum cooperation throughout the difficult time. They were extremely helpful in ensuring operator coverage throughout weekends and priority for our work. One of the problems we encountered was having to forecast the approximate number of machine hours that would be required throughout January 1981 when the bulk of conversion work was carried out. At the time the figures were needed we were still in early stages of programming so no volume tests could be run. Equally, although we were experienced in large-volume processing it was difficult to draw any direct comparisons with production work. Additionally, we had to allow for a heavier than normal production work load towards the end of the year, which always sees annual volumes, cumulations, online file reorganisation, and so on. Scheduling therefore was a fine art to ensure correct priorities for production, the bureau's own work, and conversion, and to minimise contentions for files and peripherals.

Staffing

Of interest is a picture of the human resources involved in this project. What is striking is the magnitude of the task achieved by very few people. The overall management of the project was taken on by existing line management within BSD's Computer Services Department. Two project leaders were appointed, one a librarian and one a systems analyst.

The librarian had a team of four temporarily seconded staff who were totally responsible for all output profile specifications (printed products and COM), testing, and implementation. They also did a considerable amount of checking of test file conversion runs.

The systems analyst was a project leader for three analyst-programmers and one JCL writer. Between them they were responsible for LC and U.K. conversion programming and the new filing rules.

Existing operations staff and others as appropriate within the division were called upon for other tasks.

Disruption to Services

Whilst disruption to our normal production services was kept to an absolute minimum, it was decided that it would be necessary to temporarily suspend certain services through the month of January 1981 while the bulk of the file conversion took place. Throughout the period, the BLAISE online information retrieval system continued to be operational: associated online facilities that would normally allow the despatch of MARC records to catalogue files were suspended to avoid any non-*AACR2* or nonconverted records inadvertently updating converted LOCAS files. The production of COM catalogues through LOCAS was suspended for a single month, and the first issue of *BNB* for 1981 was not scheduled until early in February. The schedule for the conversion exercise was adhered to with no major slippage except in the case of our LC file conversion; this exercise

stretched on into the spring for a variety of technical reasons largely concerned with the characteristics of the LC data.

CONCLUSIONS

Having been so closely involved in this project it is difficult to draw out general conclusions as yet. However, there are some already obvious benefits both for BSD and the wider library community: the rationalisation of our software for COM/printed products will lead to easier maintenance and future upgrading; the introduction of the *BLAISE Filing Rules* across all our products is an improvement; the new LC conversion will make our LC files much more easily usable by the British library community; we have the basis of a U.K. Name Authority File for the first time.

This was a vast and sophisticated conversion exercise and will result in U.K. MARC files probably more uniform in structure than they have ever been. It forms an excellent basis for the continuation of BSD services, especially those based on utilising records across the whole time span, e.g., BLAISE information retrieval, Selective Record and cataloguing services. Equally, because our conversion has been so extensive we have been able to share it: the specification, the Name Conversion File, and the converted U.K. and LC files were all available at minimal cost to libraries in the U.K.

Of course, it is not the 100 percent solution—it was never intended to be—so of course if you look hard enough you will find inconsistencies. However, it has proved that very extensive automatic conversion is possible even with today's state of the art of computing and that BSD had led the way, indeed eased the path of transition to *AACR2* for British libraries.

REFERENCES

1. British Library, Filing Rules Committee, *BLAISE Filing Rules* (BL, 1980).
2. British Library, Bibliographic Services Division, Computer Services Department, "Specification for Retrospective Conversion of the UK MARC Files 1950–1980" (unpublished with limited distribution).
3. British Library, Bibliographic Services Division, "Specification for Conversion of LC MARC Records to UK MARC" (unpublished with limited distribution).

Lynne Brindley is head of customer services for the British Library Automated Information Service (BLAISE).

AACR2: OCLC's Implementation and Database Conversion

Georgia L. BROWN: OCLC, Inc., Dublin, Ohio.

OCLC's Online Union Catalog (OLUC) contains bibliographic records created under various cataloging guidelines. Until December 1980, no system-wide attempt had been made to resolve record conflicts caused by use of the different guidelines. The introduction of the new guidelines, the Anglo-American Cataloguing Rules, Second Edition (AACR2), *exacerbated these record conflicts. To reduce library costs, which might increase dramatically as users attempted to resolve those conflicts, OCLC converted name headings and uniform titles in its database to AACR2 form. The purpose of the conversion was to resolve record conflicts that resulted from rule changes and to conform to LC preferred forms of heading if possible.*

BACKGROUND

In May 1978, upon receiving an advance copy of the *Anglo-American Cataloguing Rules Second Edition (AACR2)*, OCLC formed an internal task force of librarians who were professional catalogers to study the new rules. The AACR2 Task Force was charged with identifying differences between *AACR2* and *AACR1* as applied by the Library of Congress. The task force compared the two sets of rules on a rule-by-rule basis to determine: (1) effects of rule changes on the MARC record formats, (2) who benefited from the changes, and (3) relative costs of the changes on both a one-time and a continuing basis. Each change was assigned a number from 0 to 5 to represent the cost to libraries (0 being no cost and 5 being maximum cost).

The task force identified a total of 454 significant rule changes or new rules. The task force categorized each rule's effect, and in its judgment, 56 percent of the changes would benefit neither the librarian nor the patron, 23 percent would benefit librarians, and 21 percent would benefit patrons. The estimates of the percentage of changes along the cost spectrum are illustrated in table 1.

Table 1. *Estimates of AACR2 Changes in Terms of Costs*

Cost Range	Percentage of Changes— One-Time	Percentage of Changes— Continuing
0	18	20
1	54	56
2	13	20
3	9	0
4	4	2
5	2	2

Identification of Conversion Requirements

Originally, the findings of the task force were to be used to adjust the OCLC online system and card production programs to accommodate *AACR2* changes. However, in light of estimated costs to individual libraries to convert existing headings and uniform titles to *AACR2* form, the task force studied the requirements for an OCLC machine conversion. The machine conversion required that information within the record be consistently identifiable.

The task force used work sheets to record and keep track of its findings. The first column of each row on the work sheet represented one rule. The row was completed with the rule number, the *AACR2* form with tagging, the pre-*AACR2* form with tagging, instructions, and comments. Figure 1 illustrates a work sheet.

An analysis of the work sheets indicated that one method to convert to *AACR2* form was to develop an OCLC authority control system based on

AACR2 Rule	AACR2 Form with Tagging	Pre-AACR2 Form with Tagging	Instructions	Comments
22.5D1	100 10 Zerotina, Karel z	100 10 z Zerotina Karel	Within ‡a z could be searched, deleted, and added at end of field	For Czech and Slovak names only
25.4B	1xx ‡a . . . 240 ‡a Theaetetus	1xx ‡a . . . 240 ‡a Theaitētos OR 240 ‡a Theaetetus	Set up table of uniform titles where Greek forms change to Latin forms. Change 240 ‡a Greek form to Latin form	This would require reading of ‡a . . . checking against table
21.26 and 22.14	100 10 Parker, Theodore ‡c (Spirit) 700 10 Ramsdell, Sarah A.	100 10 Ramsdell, Sarah A. 700 10 Parker, Theodore	No way to automatically recognize those records requiring change
25.9	240 ‡a Selections	240 ‡a Selected works	If text of 240 ‡a is "Selected works" change to "Selections"	This will require reading text of ‡a

Fig. 1. *Task Force Conversion Worksheet*

the LC name authority file. Due to time constraints and the complexity of developing such a system, however, OCLC decided on a second method: to convert Online Union Catalog headings and uniform titles using the LC name authority file and some additional data manipulation techniques that would detect changes not done by the authority processing.

Preconversion Testing

Using the work sheets, the task force assigned the rule changes to pattern sets. Pattern sets were defined as combinations of character strings, punctuation, subfield coding, and other characteristics that indicate that the heading could be algorithmically changed to conform to the new rules. These changes were further divided into those that could be converted by machine and those that could not be converted by machine. Approximately 100 pattern sets were initially identified.

Before making a commitment to convert all 100 of these pattern sets, tests were run to determine the approximate number of bibliographic records that would be changed. A test file obtained by selecting records at random from the Online Union Catalog as of September 2, 1978, already existed at OCLC. The test file represented a 1 percent sample of the database on that date, or 41,212 records. Programs run on the test file identified the patterns within the bibliographic records and counted the number of times each pattern occurred in the test file. Table 2 illustrates selected results of pattern set sampling. Patterns not found in the test file were later eliminated from those to be applied against the entire Online Union Catalog. "U.S." was found in qualifying fields 754 times, and "Covenant" was found only once. "University of" was found 486 times on the test sample; however, it could be incorrectly converted frequently enough to eliminate it from the list of pattern matching to be done. Tests also indicated that some changes that appeared straightforward, when applied, introduced further errors that would have to be resolved after the conversion.

Of the 41,212 records, 100 records were manually checked for system changes that would need to be made for the existing bibliographic records

Table 2. Selected Results of Pattern Sampling

Rule Number	Number Matched	Comments
21.39A	32	‡a . . . ‡k Liturgy and ritual
21.39C	7	‡a Jews ‡k Liturgy and ritual
24.1B	71	State University
21.33	28	Constitution
	3	Charter
	1	Covenant
21.35	27	Treaties
25.15	206	Laws, etc.
25.6B1	0	Books, Parts, Numbers
25.9	19	Selected works
24.27B2	0	Pope

to comform to *AACR2*. General findings included:

Change	Number of Records
None	33
More than one	21
Minor personal name change	19
Personal name modification	13
Single change other than personal name	14

Specific changes that would be needed are shown in table 3. As noted in the table, personal name changes account for more than two-thirds of all required conversion changes.

As a final note, name headings to be converted by authority processing could not be estimated by sampling, since the LC name authority file was not available online when the tests were run.

Early estimates, based on the tests and anticipated name authority matches, called for conversion of 8 percent of the Online Union Catalog, or 560,000 records, to *AACR2*. However, samplings done by the Library of Congress indicated that 17 percent of all MARC records contained one or more headings that needed to be converted. OCLC assumed that this statistic would also apply to its database. The task force's study, in general, showed that OCLC could convert by machine a large portion of its bibliographic records to conform to *AACR2*.

DESIGN METHODOLOGY

OCLC formally initiated the *AACR2* project to: (1) accommodate the use of *AACR2* format in the online system, and (2) convert existing bibliographic records to *AACR2*. Accommodating *AACR2* formats required validating additional content designators, modifying card printing to allow for the new content designators, and training users. Also, the seven bibliographic format documents (*Books*, *Serials*, *Audiovisual Media*, *Scores*, *Sound Recordings*, *Maps*, and *Manuscripts*) were rewritten to include the new content designators and *AACR2* input conventions and

Table 3. *Modifications Needed for AACR2 Conversion (Based on a Sample of 100 Records)*

Modification	Occurences per 100 Records	Percent of Modification
Personal name	57	69
Parenthesize geographic location	8	10
U.S.—United States, Gt. Brit.—Great Britain	3	4
Uniform title modification	3	4
Drop geographic location from corporate name	5	6
‡k dropped	2	2
University heading	2	2
Conference date and place inverted	2	2
U.S. Congress	1	1
Total	83	100

examples. The remainder of this paper will deal with the conversion of existing bibliographic records in the Online Union Catalog, OCLC's bibliographic database. The purpose of the conversion was to resolve record conflicts that resulted from rule changes affecting name headings and uniform titles.

Functional Specifications

Two sets of functional specifications were written based on the preproject studies by the AACR2 Task Force. Set 1 functional specifications addressed the conversion of bibliographic records to *AACR2* by matching the records in the LC name authority file and then incorporating data into the bibliographic records. Set 2 functional specifications addressed the machine manipulation of character strings that formed a given pattern.

Set 1 Functional Specifications

Three constraints were placed on the conversion described in set 1 functional specifications. First, the pre-*AACR2* form of a converted field must be retained. Second, the bibliographic record must be retrievable by both pre-*AACR2* and *AACR2* forms. Third, the field that was changed must be identified to users, and the record must indicate that it had been modified by machine conversion.

Set 1 functional specifications listed the fields in the bibliographic and authority records that should be considered in the conversion, grouping bibliographic fields that should be matched with given authority fields. For each field, characters were eliminated that might inadvertently cause a no-match result. Subfield codes and delimiters, multiple blanks, and diacritics were eliminated from the character string used for matching. All alphabetic characters were converted to uppercase letters and certain subfields were eliminated from the matching strings. This process was applied to both bibliographic and authority records. The resultant matching strings, for a bibliographic and an authority field, were compared on a character-by-character basis. If any character was different, there was no match.

Matches were treated differently depending on the contents of the name authority field. Four cases for matching were defined:

Case 1. Bibliographic field matches AACR2 authority field. In case 1, the only change needed was to indicate in subfield *w* of the bibliographic field that it conformed with *AACR2*.

Case 2. Bibliographic field matches non-*AACR2* authority field; *AACR2* form present in authority record. Case 2 called for the following changes: (1) replacing the bibliographic field with the *AACR2* form from the authority record; (2) moving the replaced bibliographic data to another field (an 87x field); and (3) indicating in the converted bibliographic field that conversion had been done.

Case 3. Bibliographic field matches non-*AACR2* authority field; *AACR2* form not present in authority record. In case 3, the authority record contained the form preferred by LC, but not the *AACR2* form. If the bibliographic field matched a "see from" reference (4xx authority field), case 3 called for the following changes: (1) replacing the bibliographic field with the authoritative field (1xx authority field); and (2) moving the replaced bibliographic data to another field (an 87x field). No indication was added that the field was machine-converted, since the form supplied was not *AACR2*.

Case 4. Bibliographic field tagged as personal name matches authority field tagged as corporate name. In case 4, the bibliographic tag was corrected to a corporate-name tag. Case 4 was used to clean up the database and to allow more fields to be converted.

Set 2 Functional Specifications

For set 2 functional specifications, the pre-*AACR2* form of the entry also must be retained and the record retrievable by both pre-*AACR2* and *AACR2* forms. These functional specifications called for conversion of six pattern sets. Each pattern set might apply to multiple fields and, within the fields, to multiple character strings.

Some of the pattern sets were further subdivided into various conditions. For example, pattern set 2 specified the conversion of form subheadings. This pattern set looked only at one field, the 110 field, but held two conditions. In the first condition, any one of ten character strings might be matched. In the second condition, either of two character strings qualified for matching. Pattern set 2 was actually one of the easier sets to work with since it involved minimum data manipulation and testing.

The most complicated pattern set concerned music uniform titles where only two fields were involved but six possible conditions had to be considered. One of these conditions required conversion of forty-two character strings, provided other information was present.

Development Plan

After reviewing the two sets of functional specifications, a development plan was established. This plan outlined the steps involved in software development for the project, named an individual responsible for each step, estimated the duration of each step, identified the objectives of software development, and identified potential time conflicts for staff and machine resources. The time estimates were constantly monitored and revised during the project cycle to ensure that the work would be completed on time.

Development Method

Based on a thorough analysis of the functional specifications, the following basic design was chosen:

1. Read a bibliographic record.
2. Identify a field in the bibliographic record for potential conversion.
3. Derive a key from that field. The key derivation used would be the same as that used for the online system, except that it would be extended to include fields not normally indexed but that needed to be converted to *AACR2*.

   Derived search keys are formulated by extracting a certain number of characters from the words in a name. For personal names, a 4,3,1 key is used; i.e., the first four characters from the surname, the first three characters from the forename, and the middle initial.
4. Perform a keyed search of the LC name authority index files.
5. For each hit on an index record, read the corresponding name authority record and check for a match of the authority and bibliographic fields. When a match is found, merge the bibliographic and authority data.
6. Repeat steps 2 through 5 for every field in the bibliographic record that qualifies for conversion.
7. Scan the bibliographic record for fields that might be converted using the machine-manipulation pattern matching and compare these fields with the various patterns. Should a match occur, manipulate the string accordingly.
8. If a record has been converted, add the 040 field if it is not already present in the record; or, edit the 040 field to include a subfield *d* indicating that OCLC has modified the record.
9. Repeat the entire process for every record in the Online Union Catalog.

Design Method for Conversion

The method presented a complex design. Because it required indexing fields not normally indexed by the OCLC system, the search keys would have to be specified. Also, the 130, 430, 530 uniform and variant title fields in the name authority file would have to be indexed and the keys defined. This could be done by adding the search keys to the existing name index file, which contains indexes to the LC name authority file, or by creating a separate file. Adding to the existing name index file would result in inconsistent data within the file, mixed names and titles, and, more important, interference with the online system. Using a separate file would mean more maintenance, necessitate slightly more machine space, and require two searches to cover all search possibilities for derived name authority search keys. (It should be noted that currently online system users cannot search the name authority file using a derived title search key.)

SOFTWARE DEVELOPMENT

Project software design defined activity along the two lines of conversion, corresponding to the functional specifications: conversion of name

headings by matching bibliographic headings with headings in the LC name authority file, and conversion of name and uniform title headings through machine manipulation of existing bibliographic data. Conversion by matching name authority headings was broken down into subactivities as specified by cases 1 through 4 in the functional specifications. Conversion by machine manipulation was subdivided into:

1. Conversion of conference name headings.
2. Conversion of uniform titles—music.
3. Conversion of uniform titles—general.
4. Conversion of form subheadings.
5. Conversion of general material designators.
6. Conversion of "United States" and "Great Britain" abbreviations.

The entire conversion was designed to be directed by a series of run-time parameters that specified which subactivities were to be performed, whether the conversion was to be run concurrently with the online system, the names of files to be used (including audit and checkpoint files), and the range of OCLC numbers to be processed. The run-time parameters allowed multiple processes (programs) to be run simultaneously, with each process running against a different part of the Online Union Catalog.

The design also included use of an audit trail, where a record is written to a file every time a change is made to a bibliographic field. The trail consisted of the OCLC number and the type of subactivity applied to the field.

Conversion restarts were specified to be automatically controlled through a checkpoint file. Checkpoint records in this file contained the latest OCLC number processed, total number of records processed, total number of records, and time stamps to calculate elapsed time. To effect a restart, the conversion was simply rerun and the checkpoint file handled file positioning to ensure against duplicate reprocessing of records.

An overriding development priority was to design the software to be flexible enough to handle both the conversion of the Online Union Catalog and the conversion of incoming MARC tapes. In this way, pre-*AACR2* headings would be converted (if they met the specifications) before being loaded into the database.

Growth Requirements

At the same time that the coding began, the project staff studied the design to determine its effects on the existing system. Additional disk space was projected based on the estimate of bibliographic records to be converted. Based on past research of field lengths, project staff estimated that 66.42 bytes (characters) would be added to each converted record. Based on earlier samplings by the Library of Congress, it was assumed that 17 percent of the database would be converted (a figure that turned out low). Therefore, 79.04 additional megabytes would be used. Because an additional 13 percent of this would be needed for file management, the total

requirement for the expansion of the bibliographic file was projected as 89.3 megabytes.

The bibliographic index files would also expand with the conversion. Not only would the old index keys be retained but new keys would be added. It was estimated that 4 percent of the bibliographic records would generate new keys (duplicate keys are not added to the files), for an additional requirement of ten megabytes. It was also calculated that six megabytes would be required for the new name authority index file. The total additional space required for the expansion of the bibliographic file, the expansion of the bibliographic index file, and the addition of the name authority index file was thus 105.3 megabytes. This space would have to be available before the conversion could be run.

Testing

As coding progressed into the testing phase, it became obvious to the project staff that existing testing methods were not well suited to testing the conversion software. Therefore, a utility program was developed to enter bibliographic records in a test file using techniques similar to those used by the online system. These test bibliographic records included both good and bad data, and records that should and should not be converted. An attempt was made to cover as many situations as practicable. For example, a given record might have multiple fields that would convert and, within a given field, multiple conversions might apply. Images of the converted test records were manually compared with the original entry. At the same time, the accuracy of the audit trail was verified. The conversion process was tested using a utility debugger to simulate error conditions that did not occur as a result of other tests. Changes to the online system code were tested using a simulator. All testing and development work was done on a development machine.

Calibration tests were made on the Data Base Processor (DBP), the database management portion of the online system. The calibrations were taken in a stand-alone environment to calculate the length of time needed to run the conversion and to test the conversion software on a larger database than the test files. At the time of the calibration tests, the LC name authority file held about 250,000 records; it was not distributed across the various disk packs but rather restricted to a fairly small number of packs. Between the time of the calibration and the conversion run, the LC name authority file grew to 450,000 records and was distributed evenly across the disk packs on the DBP. According to the calibration tests, the conversion to *AACR2* was expected to take ninety-two hours, with sixteen copies of the software processing different ranges of the bibliographic file. The calibration tests also estimated that 28 percent of the bibliographic records would be converted, much higher than originally estimated.

After the calibration tests, the software underwent quality assurance tests. The conversion software was run against test files on the DBP to

verify the conversion process and to provide the data for the next phase of quality assurance, the regression test. During regression testing, each subsystem in the online system, with new software changes included, was tested by OCLC staff. Additional tests were made of normal work flows to ensure that all functions that previously worked still functioned correctly and all functions that should not work still did not work. No problems were uncovered during these tests and no software changes were made.

CONVERSION OF THE OCLC ONLINE UNION CATALOG

The conversion was designed to run either in a stand-alone mode or concurrently with the online system. The major drawback to running in a stand-alone mode was that the online system would be unavailable to users for some period of time. However, this was not deemed as great a problem as running the conversion while the online system was operational. With the online system operational, the conversion would have to "lock" the bibliographic record as it is processed, thus potentially affecting system performance. For example, if a user wanted to retrieve a record that was locked, he or she would have to wait until the record was unlocked. Since the *AACR2* conversion process locks the bibliographic record when it reads it and keeps it locked until the conversion for that record is complete, the record could stay locked for several seconds.

Before the conversion could be run, various files had to be created on the DBP. The bibliographic file on the DBP is partitioned across twenty-nine disk packs, each pack holding 250,000 records within a range of OCLC control numbers. The start-up commands and parameters were put into one file for execution. One audit file was created for each process to be run. The conversion began running with sixteen processes. Ten of the processes were run against two disk packs, with four processes running against a single disk pack. At the time of conversion, the DBP contained fourteen CPUs; twelve of the processes ran alone in a CPU, and two processes ran in each of two CPUs.

As soon as the conversion began, on December 13, 1981, at 4:00 a.m., another calibration test was done to estimate completion time. The results showed that the file redistribution that was expected to lower the time estimates significantly had not produced the expected result. Attempts were made to explain the discrepancies, but it was concluded that the processes simply were slow. The I/O rate and CPU utilization rate were high. Based on these calibration test findings, it was decided to start up additional processes so that one process would be run on a single disk pack, with two processes per CPU. The original sixteen processes had to be stopped, the range of OCLC numbers processed redistributed, and additional audit files created. Twenty-eight processes were then started up. All records in the twenty-ninth disk pack, records with control numbers greater than seven million, were to be handled by the twenty-eighth process.

The conversion ran smoothly until some of the processes encountered a problem they could not handle. The conversion was then stopped. Because the problem was not immediately obvious, the records being processed at the time of the error were skipped and the conversion restarted using the checkpoint file. The problem was later identified—if the converted field held more than 255 characters, the length of the field was incorrectly calculated—and software was corrected. The audit files were saved up to the point of the correction to identify the problem records. Using these audit files to find records that had been converted, a preconversion copy of the bibliographic file was scanned for records that would need correction. Fifty-six records were identified and sent to the Bibliographic Maintenance Section, User Services Division, of OCLC for manual correction.

From this point on, the conversion ran smoothly but slowly, processing an average of 28,500 records per hour. The checkpoint files were read every two hours to monitor the speed of the conversion. Because this monitoring in itself proved to be quite cumbersome, a program was written to format the checkpoint data for easier readability. The resultant reports showed a breakdown by process of how much of the conversion had been done, the rate at which it had been done, and how much remained. By using these reports, as a process would finish, another slower process could be divided and started up to balance the load and finish faster. Periodically, converted records were written on hard-copy printers for OCLC staff to use to check the accuracy of the conversion.

The checkpoint reports showed that 39 percent of the records in the Online Union Catalog were being converted to *AACR2*. This percentage was much higher than anticipated by the calibration tests, and consequently the disk space needed for expansion was more than anticipated. Files not used by the conversion were deleted and index files were moved to other disk packs to allow the bibliographic files to expand.

The last record was converted and all processes stopped by 10:45 a.m. on December 21, after 246 hours of work. The bibliographic file and its indexes were reorganized, slack space squeezed out, and all files that had been deleted were put back. The online system was made available to users at 7:00 a.m., December 23, 1980. A total of 3,704,440 changes had been made on more than 2,767,000 records. Table 4 lists the number of fields converted for each activity.

SUMMARY

Some records could not be converted because: (1) the data within the field were incorrect or inadequate, or (2) the record would have exceeded field number and record length limits.

OCLC has made a continuing effort since the conversion to correct problems. The most difficult and numerous problems involved the LC name authority file. In some cases the data within the authority records are incorrect, while in other instances multiple authority records exist. The

Table 4. Fields Converted for Each Activity

Activity	Number of Fields Converted
Mistagged corporate name fields	1,268
Direct *AACR2* match	2,685,211
Match where *AACR2* form is elsewhere in the authority record	614,333
Match on LC preferred form	23,611
Conversion of conference name headings	96,382
Conversion of uniform titles—music	68,905
Conversion of uniform titles—general	2,263
Conversion of form headings	31,278
Conversion of general material designators	49,978
Conversion of "United States" and "Great Britain" abbreviations	131,211

conversion used the first matching authority record it encountered. The most desirable record, as it turned out, was sometimes not the first encountered.

A series of eight fixes was programatically applied to the OLUC to correct problems, using either the audit file or database scans to select the record to be corrected. Fixes 1 and 2 were similar in that each was the result of a bad authority record and the original form was restored. Headings converted to "Voice of America (Radio program)" were changed back to "United States. Dept. of State" by fix 1. "United States Bureau of the Census. Census of construction industries (1972)" was changed back to "United States. Bureau of the Census" by fix 2.

Fixes 3 through 7 were needed to correct programming problems, omissions in the functional specifications, and changes in LC procedures. Subfields x, y, and z were deleted from 600 fields by the conversion. Fix 3 moved the subfields back into the 600 fields. Fix 4 reordered the e and q subfields in personal name headings that had been moved into the field in the wrong order by the conversion. The conversion had supplied a subfield g between the word "Manuscript" and the following text in 110 fields. Fix 5 changed subfield coding g to n when LC began using the n. Fixes 6 and 7 restored some fields to the original form, which had been unintentionally converted. Fix 6 corrected form subheadings, and fix 7 corrected music uniform titles. "Constitutional" had been treated as "Constitution," i.e., it was deleted from the field. Some terms within music uniform titles were to have been pluralized. However, the conversion did not differentiate between terms needing pluralization and those that were already plural. "Masses" ended up as "Masseses." Fix 7 corrected this problem.

Forty-six headings, including Chopin, Shakespeare, and Beethoven, were identified as unconverted headings, resulting from the multiple authority record problem. Fix 8 adjusted the name authority file so the desired record would be the first encountered, scanned the OLUC to select records containing the forty-six headings, and ran those selected records through the conversion process. Approximately 80,000 records were converted by fix 8.

Other problems were expected to filter in, although the stream has slowed to a trickle. These problems continue to be dealt with by OCLC librarians. On the whole, problems were expected, planned for, and handled in a timely fashion. OCLC originally envisioned the conversion of its large database to encompass 8 percent of the total records online; 39 percent of the records were converted, and they were available to OCLC users before the January 1, 1981, deadline set by the library community.

Georgia L. Brown is manager, Cataloging Section, in the Development Division of OCLC.

OCLC's Database Conversion: A User's Perspective

Arnold WAJENBERG and Michael GORMAN: University of Illinois Library, Urbana-Champaign

This article describes the experience of a large academic library with headings in the OCLC database that have been converted to AACR2 form. It also considers the use of LC authority records in the database. Specific problems are discussed, including some resulting from LC practices. Nevertheless, the presence of the authority records, and especially the conversion of about 40 percent of the headings in the bibliographic file, has been of great benefit to the library, significantly speeding up the cataloging operation. An appendix contains guidelines for the cataloging staff of the University of Illinois, Urbana-Champaign in the interpretation and use of LC authority records and converted headings.

The library of the University of Illinois, Urbana-Champaign, is the largest library of a publicly supported academic institution, and the fifth largest library of any kind, in the United States. In the last year for which figures are available (1979–80), the library added more than 180,000 volumes representing more than 80,000 titles. The library is currently cataloging more than 8,000 titles a month; more than 80 percent of the records for these titles are derived from the OCLC database (Library of Congress and OCLC member copy).

Because our cataloging is of such volume and because we are actively engaged in the development of an online catalog, we decided to use the second edition of the *Anglo-American Cataloguing Rules* (*AACR2*) earlier than the "official" starting date of January 1981. We began to use *AACR2* for all our cataloging in November 1979. This early use of *AACR2* has led to two consequences. First, we now have OCLC archival tapes representing about 150,000 titles cataloged according to *AACR2*. This represents a valuable and continuously growing bibliographic resource that can be used without modification in our future online catalog. Second, we have a considerable and unique collective experience in the practical application of *AACR2*. The minor problems of working with *AACR2* in an *AACR1*

plus superimposition environment (until January 1981) were more than compensated for by these two positive results.

OCLC CONVERSION

With our practical background in the use of *AACR2* and our continuing need for a high volume of cataloging, we were, naturally, keenly interested in the (to our mind) progressive decision of OCLC to use machine matching techniques to convert the form of name and title headings in its database—the Online Union Catalog (OLUC)—to conform to *AACR2*. We recognized the limitations of the project, essentially those defined by the capabilities of the computer for matching character by character, but felt that this was a major venture that would, when completed, produce major benefits.

What follows is an assessment and analysis of the results of the project in the light of the experience of a library that is dedicated to achieving high-volume, quality cataloging. We deal with the LC authority file as well as the OCLC headings because the LC file was the basis of the project and because, from the practical point of view, the two files are complementary aspects of the same service.

The greatest value of the conversion, and its greatest claim to uniqueness, lies in the sheer size of the project in terms of headings checked and changed. Our catalogers, and others who work with current materials, estimate that more than 40 percent of the name and title fields we use in our current cataloging have a w subfield indicating that the name or title has been changed to its *AACR2* form. Since OCLC estimates that 39 percent of the name and title fields were affected by the conversion, it would appear that the headings that were changed are the headings that we are more likely to use. In other words, the project has brought us more than a 39 percent benefit. We are also greatly encouraged to find that the number of headings coded *dn* (meaning *AACR2* "compatible," or, more bluntly, LC's modifications of the provisions of *AACR2*) is a very tiny minority of all converted headings. This means that when, in the future, this policy of "compatibility" is lessened or dropped, there will be relatively few changes to be made.

LC AUTHORITY RECORDS

We also benefit from the presence of LC authority records in the OCLC database when we establish headings that are new to our catalogs. There is one problem with the use of these records, which was revealed by a sample of new University of Illinois authority records (see table 1). This sample of 368 new University of Illinois records reveals that LC authority records are available relatively rarely for new headings. This is not surprising as these new headings are established most often as part of the process of original cataloging, which, almost without exception, occurs in our library only when OCLC copy is not available. It seems to us to be unfortunate that

Table 1. Recently Established Headings

	No Authority Record	Record Coded c*	Record Coded d*	Record Coded n*
Given name headings	13	5	0	1
Single surname headings	212	26	2	2
(Number of this category with initialisms expanded in parentheses)	(132)	(7)	(1)	(2)
Compound surname headings	29	12	0	1
(Number of this category with initialisms expanded in parentheses)	(2)	(0)	(0)	(0)
Single surnames plus uniform titles	3	0	0	0
General corporate headings	34	12	0	0
General headings with subdivisions	7	2	0	0
Government headings	4	2	0	1
Total	302	59	2	5

*Key:

c—in subfield *w*, indicates an AACR2 form, as established by Library of Congress.

d—in subfield *w*, indicates an AACR2 "compatible" form, as established by Library of Congress.

n—in subfield *w*, indicates that the input operator could not determine which set of rules governed the form of the heading.

member libraries cannot contribute their authority records to the OCLC database. Our experience suggests that the online authority file would grow very rapidly if that were the case. To put it another way, the OCLC conversion provides an enormous and valuable resource of *AACR2* headings. It did not, and could not, provide new authority information. OCLC will be complementing its valuable work in upgrading the retrospective file when it devises and implements a scheme for making available authority records for new headings derived from a wide range of sources. Since so many headings were converted to *AACR2*, it may seem churlish and ungrateful to complain that more was not done. The following descriptions are not intended to form part of an attack on OCLC's project or to minimize its achievement.

FORM SUBDIVISIONS

The project failed to delete form subdivisions (such as "Liturgy and ritual" and "Laws, statutes, etc.") from added entry headings and subjects. The program correctly deleted them from main entry headings, but the inconsistencies resulting from their retention elsewhere makes the job of ensuring consistency in a large copy cataloging operation that much harder.

This inconsistency in treatment is illustrated by examples 1 and 2. Example 1 originally was entered under

110 10 Illinois. k Laws, statutes, etc.

The program correctly changed the main entry heading to
110 10 Illinois

and added a subfield *w*, coded *mn* (the *m* indicates a conversion by machine to the *AACR2* form; the *n* means "not applicable," and indicates that there is no title element in the heading).

Example 2 has as main entry

110 20 Illinois Community College Board

but has as added entry

710 10 Illinois. k Laws, statutes, etc. t Illinois
public community college act

Under *AACR2*, the subfield *k*, "Laws, statutes, etc.," should not be present in the heading. Unfortunately, the program looked only at 110 fields, not at 710 fields, and so the heading was not corrected in the conversion. It must therefore be edited manually by every library that uses the record.

PROGRAM PROBLEMS

Our direct use of the online authority file is somewhat hampered by the programming oversight that makes it impossible to search uniform titles. Of course, uniform titles that are accompanied by a 100 field (notably in music) can be retrieved by an author search, but those without 100 fields (anonymous classics, sacred scriptures, etc.) are virtually inaccessible.

There were a handful of specific instances in which the specifications were inadequate or the programs seem to have malfunctioned. These resulted in some oddities such as the conversion of the subject "Jesus Christ" to "Sermon on the Mount" and the (*surely* not politically motivated) switch from "U.S. Department of State" to "Voice of America." OCLC has been scrupulous in identifying and publicizing these errors. They are few in number and, though conspicuous, have rarely caused us many problems.

As can be seen, the problems caused by what we see as failures on OCLC's part are few and affect few cataloging circumstances. The remaining problems either result from the decisions and actions of the Library of Congress and, hence, are wholly or mostly out of OCLC's control, or are of such a nature that they cannot be solved by computer matching techniques without extensive editorial intervention. Whether such human intervention is possible and, if possible, cost-beneficial is not for us to say, though it must be recognized that to transform the OLUC to pure *AACR2* conformity would be a herculean task. That task would undoubtedly involve many of the hundreds of thousands of records that are seldom or never used.

```
Screen 1 of 2
UIU - FOR OTHER HOLDINGS, ENTER dh DEPRESS DISPLAY RECD SEND
 OCLC: 7044295      Rec stat: c Entrd: 810102        Used: 810120
Type: a Bib lvl: m Govt pub: s Lang:  eng Source: d Illus:
 Repr:    Enc lvl: I Conf pub: 0 Ctry:  ilu Dat tp: s M/F/B: 10
 Indx: 0 Mod rec:   Festschr: 0 Cont:
 Desc: i Int lvl:   Dates: 1980,
  1 010
  2 040      SPI c SPI
  3 092      346.045 b I29L, 1980
  4 092      b
  5 049      UIUU
  6 110 10   Illinois. w mn
  7 245 10   Illinois laws relating to planning and development / c prepared
by Division of Government and Community Services, Office of Housing and
Community Development.
  8 260 0    [Springfield] : b State of Illinois, Dept. of Commerce and
Community Affairs, 1980.
  9 265      Illinois Dept. of Commerce & Community Affairs, 222 So. College
St., Springfield, IL 62706.
 10 300      89 p. : c 28 cm.
 11 490 1    Reference series - Department of Commerce and Community Affairs :
1980
```

```
Screen 2 of 2
 12 500      "An update and revision of the publication, 'Reference series no.
3, Laws relating to planning and development,' June 1976."
 13 650  0   Regional planning x Law and legislation z Illinois.
 14 650  0   City planning and redevelopment law z Illinois.
 15 650  0   Zoning law z Illinois.
 16 710 10   Illinois. b Dept. of Commerce and Community Affairs.
 17 710 10   Illinois.  Office of Housing and Community Development.
 18 810 1    Illinois. b Dept. of Commerce and Community Affairs. t Reference
series - Department of Commerce and Community Affairs : v 1980.
 19 871 19   j 110/1 a Illinois. k Laws, statutes, etc.
```

Example 1

```
Screen 1 of 2
UIU - FOR OTHER HOLDINGS, ENTER dh DEPRESS DISPLAY RECD SEND
 OCLC: 6606264      Rec stat: n Entrd: 800811        Used: 810325
Type: a Bib lvl: m Govt pub: s Lang:  eng Source: d Illus:
 Repr:    Enc lvl: I Conf pub: 0 Ctry:  ilu Dat tp: s M/F/B: 10
 Indx: 0 Mod rec:   Festschr: 0 Cont:
 Desc: i Int lvl:   Dates: 1979,
  1 010
  2 040      SPI c SPI d m.c.
  3 043      n-us-il
  4 092      344.074 b I29Le
  5 092      b
  6 049      UIUU
  7 110 20   Illinois Community College Board w cn
  8 245 10   Legislation enacted into law from the 1978 and 1979 sessions of
the Illinois General Assembly that affect the Illinois public community college
system.
  9 260 0    Springfield : b Illinois Community College Board, c 1979.
 10 263      Illinois Community College Board, 3085 Stevenson Dr., Springfield,
Ill. 62703
 11 300      44, [21] p. : c 28 cm.
 12 500      Cover title.
```

```
Screen 2 of 2
 13 500      Appendix includes supplements to the Illinois public community
college act.
 14 650  0   Community colleges x Law and legislation z Illinois.
 15 710 10   Illinois. k Laws, statutes, etc. t Illinois public community
college act.
```

Example 2

SERIALS

The most troublesome example of the kind of problem that cannot be resolved by machine matching is that of serials. The OCLC conversion project was, quite properly, not concerned with choice of entry (*AACR2*, chapter 21). This seems a simple and clearly defined decision. When we come to consider serials, this clear distinction between choice and form of entry becomes blurred. The major change brought about by *AACR2* (rule 21.1B2) is that many serials previously entered under the heading for a corporate body are to be entered under their titles. In fact, the great majority of serials will now be entered under title. The upshot of this is that the citation (or form of heading) for a serial changes from, for example,

National Society for Medical Research. Bulletin

to

Bulletin / National Society for Medical Research

The restriction of the OCLC project to forms of heading means that most serials in OLUC will be found under headings the form of which may be correct but are inappropriate for citations. This problem, which, of course, cannot be resolved by computer matching, has led to difficulties for us in copy cataloging, because a degree of expertise is needed to apply *AACR2* rule 21.1B2 and to distinguish between the majority of serials where the 110 field should be changed to a 710 and the small minority where the 110 field should remain as it is. Since most serials are to be entered under their titles, it occurs to us to suggest that the OCLC conversion project could have changed all 110 fields in records identified as relating to serials to 710. By that method, the majority of serials would be correctly entered and the potential for mistaken citations greatly reduced.

MULTIPLE PERSONAL NAMES

Persons who write under more than one name (real names, pseudonyms, etc.) and who are not primarily identified by one of those names (*AACR2* 22.2C3) pose a special problem. Under the provisions of *AACR2*, such persons are to be represented in the catalog (and the database) under two or more names. Despite the fact that "Creasey, John" and "Marric, J. J." and "Ashe, Gordon" are all names used by the same man, they will appear as separate headings from now on. Under *AACR1* plus superimposition one of those names ("Creasey, John") was used as the heading for all works. Within the confines of the OCLC project, there was no method available to distribute the various records under the various headings. It occurs to us that some method based on matching the name found in the 245 $c subfield with the 100 field might, at least, have resulted in the project recognizing probable cases calling for multiple headings. For example:

100 a Hibbert, Eleanor
245 a Bride of Satan / $c Jean Plaidy

could alert the system to a case for change. We recognize that this would call for more sophisticated computer matching techniques and that it would call for editorial intervention. A good example of the problem this has caused for us is the case of the Danish author Karen Blixen. She wrote under that name and under the pseudonyms Isak Dinesen and Pierre Andrézel. Records in the database that were added before 1981 will use "Blixen, Karen, 1885–1962" as the heading for all her works including those published under pseudonyms. Since the Blixen heading is a perfectly acceptable *AACR2* form, the conversion program codes it as an *AACR2* heading, which it is for the Blixen books but is not for those published under other names.

The authority record (example 3) includes a note identifying both pseudonyms as valid *AACR2* headings, but, of course, the programs as written cannot interpret such a note and match them with appropriate records.

CORPORATE NAME CHANGES

Corporate bodies present a similar problem when one is dealing with those that have changed their name. Until 1967, the Library of Congress used the latest name of such bodies with see references from the earlier names. Both editions of *AACR* require that works issued under the earlier names be entered under those names and works issued under the latest name be entered under that name, the various names being connected by *see also* references. However, records in the OLUC for earlier works cataloged before 1967 will show those works entered under a later name.

```
FOR BIB RECORD ENTER bib DISPLAY RECD SEND
Rec stat: n Entrd: 801121 Used: 801121
Type: z Bib lvl: x Govt Agn:  Lang:  Source:
Site: 004 InLC: a Enc lvl: n Head ref: a Head: cc
Head status: a Name: a Mod rec:  Auth status: a

 1 010     n  79007719
 2 100 10  Blixen, Karen, d ,885-1962. w n001790215aacann----nnnn
 3 400 10  Andrbezel, Pierre w n002790215aanann-----nnnd
 4 400 10  Dinesen, Isak w n003790215aanann----nnnd
 5 667     The following pseudonyms are valid AACR 2 headings: a Andrbezel,
Pierre, 1885-1962; a Dinesen, Isak, 1885-1962 w n004790215aanann----nnnn


NO HOLDINGS IN UIU - FOR HOLDINGS ENTER dh DEPRESS  DISPLAY RECD SEND
OCLC: 1443879     Rec stat: c Entrd: 750711     Used: 810725
Type: a Bib lvl: m Govt Pub: _ Lang: eng Source: u Illus:
Repr:    Enc lvl: I Conf pub: _ Ctry:  ___ Dat tp: _ M/F/B: ___
Indx: _ Mod rec: _ Festschr: _ Cont:
Desc:    Int lvl:  Dates: 1963,____
 1 010     63-11618
 2 040     c ORL d OCL d m.c.
 3 050 0   PZ3.B62026 b Eh
 4 092     Fic
 5 092     b
 6 049     UIUU
 7 100 10  Blixen, Karen, d 1885-1962 w cn
 8 245 1   Ehrengard c [by] Isak Dinesen [pseud]
 9 260 0   New York, b Random House, c [1963]
10 300     111 p. c 22 cm.
```

Example 3

Because those later names are valid *AACR2* headings in terms of their form, they are coded *cn* (i.e., *AACR2* validated) by the program, even though they may not be the right headings for the records to which they are attached. A good example of this problem is that of the "Lutheran Church-Missouri Synod." An earlier name of this religious body is "Deutsche Evangelisch-Lutherische Synode von Missouri, Ohio, und Andern Staaten." Unfortunately, the authority record (example 4) does not even show that the earlier name is valid according to *AACR2*. The conversion program, on encountering the earlier name used as a heading, would change it to the later name and code that form as being the *AACR2* heading.

Another example of the problem is:

> Chamber of Commerce of the United States of America.   International Department

This is identified as the *AACR2* form (example 5) but, in fact, the department has changed its name to "International Division."

LC PRACTICE

Another problem we have encountered is that of the literal-mindedness of the computer programs in matching like with like. This problem is compounded by inconsistencies in cataloging practice resulting from variations in LC cataloging practice. An example of this problem is that of the Nigerian author Chinua Achebe. The heading "Achebe, Chinua" is marked as being *AACR2* despite the fact that the authority record shows

```
Screen 1 of 2
FOR BIB RECORD ENTER bib DISPLAY RECD SEND
Rec stat: n Entrd: 801228 Used: 801228
Type: z Bib lvl: x Govt Agn:    Lang:        Source:
Site: 011 InLC: a Enc lvl: n Head ref: a Head: c
Head status: a Name: a Mod rec:    Auth status: a

1  010      n  80057065
2  110 20  Lutheran Church--Missouri Synod. w n001801105aacann----nnnn
3  410 20  Evangelical Lutheran Synod of Missouri, Ohio, and Other States w
n002801105aanann----nnna
4  410 20  Missouri Synod w n003801105aanann----nnna
5  410 20  German Missouri Synod w n004801105aanann----nnna
6  410 20  Deutsche Evangelisch-Lutherische Synode von Missouri, Ohio und
Andern Staaten w n005801105aanann----nnna
7  410 20  Evangelisch-Lutherische Synode von Missouri, Ohio und Andern
Staaten w n006801105aanann----nnna
8  410 20  German Evangelical Lutheran Synod of Missouri, Ohio, and Other
States w n007801105aanann----nnna
9  510 20  Lutheran Church in the U.S. w n008801105aanann----nnna

Screen 2 of 2
10  510 20  Evangelical Lutheran synodical conference of North America w
n009801105aanann----nnna
11  667      AACR 1 form: a Lutheran Church--Missouri Synod w
n010801105aanann----nnnn
12  667      The following subdivision has not been used as a heading: a
Lutheran Church--Missouri Synod. Fiscal Office w n011801105aanann----nnnn
```

Example 4

```
Screen 1 of 2
FOR BIB RECORD ENTER bib DISPLAY RECD SEND
Rec stat: c Entrd: 801207 Used: 810531
Type: z Bib lvl: x Govt Agn:   Lang:   Source:
Site: 007 InLC: a Enc lvl: n Head ref: a Head: c
Head status: a Name: a Mod rec:   Auth status: a


1 010      n  80102217
2 110 20  Chamber of Commerce of the United States of America. b
International Dept. w n001801021aacann----nnnn
3 410 20  Chamber of Commerce of the United States of America. b Foreign
Commerce-Foreign Policy Dept. w n002801021aanann----nnna
4 410 20  Chamber of Commerce of the United States of America. b
International Division w n003801021aanann----nnna
5 510 20  Chamber of Commerce of the United States of America. b Foreign
Commerce Dept. w n004801021aanann----nnna
6 510 20  Chamber of Commerce of the United States of America. b
International Relations Dept. w n005801021aanann----nnna

Screen 2 of 2
7 667     The following heading for the earlier name is a valid AACR 2
heading: a Chamber of Commerce of the United States of America. Foreign
Commerce-Foreign Policy Dept. w n006801021aanann----nnnn
8 670     An introduction to doing import ... 1947- w n007801021aanann----
nnnn
```

Example 5

that he was born in 1930. LC's announced policy is to give dates "whenever the information is readily available," but only for headings established after December 1980. This restriction creates inconsistencies in LC practice that are hard to predict. The result is that we often establish a personal heading with a date, only to discover that LC is not using it. The definition of "readily available" is clearly elastic and does not provide clear guidance to other libraries. It is irritating and occasionally burdensome but does not create a quantitatively serious problem.

One unfortunate result of LC's machine conversion of its authority file to *AACR2* forms has been to make notes on the authority records harder to understand. This is because only headings and references were changed; notes were not affected. This means that the wording of the notes may refer to a state of affairs that has altered as a result of the *AACR2* conversion.

Example 6 is the authority record for the *AACR2* form of heading for the University of Illinois prior to the change of name in 1966. The history note (field 667) incorrectly says that the heading for works published before 1966 is "Illinois. University" (the pre-*AACR2* form). Since the *AACR2* form as established by LC looks very much like the new name, "University of Illinois at Urbana-Champaign," the authority card is very difficult to understand. Nothing short of revising the note, and/or the use by LC of a less confusing qualifier than "(Urbana-Champaign campus)," will make the authority record intelligible.

An example of how LC practice has affected the OCLC program adversely is in the area of the so-called compatible headings. These are instances of when LC has chosen to depart from the provisions of *AACR2* for one reason or another. Leaving aside the utility and morality of such a policy, it presents a considerable problem to those of us who use OCLC

records. The example that follows is of the worst of these "compatible" practices. LC has decided to ignore the common form of name for persons who are not "famous or published under an American imprint."[1] Thus, the writer P. C. Boeren would be recorded as "Boeren, P. C. (Petrus Cornelis), 1909– " under the provisions of *AACR2*, but, because Boeren is neither famous nor American, the "compatible" heading will be "Boeren, Petrus Cornelis, 1909– ." This heading is not acceptable in an *AACR2* catalog.

```
   Screen 1 of 4
FOR BIB RECORD ENTER bib DISPLAY RECD SEND
Rec stat: c Entrd: 801122 Used: 810718
Type: z Bib lvl: x Govt Agn:  Lang:      Source:
Site: 038 InLC: a Enc lvl: n Head ref: a Head: c
Head status: a Name: a Mod rec:  Auth status: a


  1 010       n  79049104
  2 110 20  University of Illinois (Urbana-Champaign campus) w
n029801115aacann----nnnn
  3 410 20  Illinois Industrial University w n002790829aanann----nnna
  4 410 20  University of Illinois w n027800317aanann----nnna
  5 410 10  Illinois. b University. w n030801115aaaana----nnna
  6 410 10  Urbana (Ill.). b University of Illinois (Urbana--Champaign
campus) w n031810616aanann----nnna
  7 410 10  Illinois. b University (Urbana--Champaign campus) w
n038810616aanann----nnna
  8 510 20  University of Illinois at Chicago Circle w n007790829aanaen----
nnna
  9 510 20  University of Illinois at the Medical Center w n008790829aanaen--
--nnna


  Screen 2 of 4
 10 510 20  University of Illinois (System) w n009790829aanaen----nnna
 11 510 20  University of Illinois at Urbana-Champaign w n032810616aanaen----
nnna
 12 510 20  University of Illinois at Congress Circle, Chicago w
n033810616aanaen----nnna

Screen 3 of 4
 13 665      The Illinois Industrial University was chartered in 1867. In 1885
the name was changed to University of Illinois and in 1966 to University of
Illinois at Urbana-Champaign. a Works by this body published before the change
of name in 1966 are found under a University of Illinois (Urbana-Champaign
campus) a Works published after that change of name are found under a
University of Illinois at Urbana-Champaign. a The Chicago Undergraduate
Division of the University of Illinois was established in 1946. In 1962 the
name was changed to University of Illinois at Congress Circle, Chicago, and in
1964 to University of Illinois at Chicago Circle. a Works by this body
published before the change of name in 1964 are found under a Universty of
Illinois at Congress Circle, Chicago. a Works published after that change of
name are found under a University of Illinois at Chicago Circle. a In 1966
the University of Illinois at Urbana-Champaign, the University of Illinois at

  Screen 4 of 4
  13 (cont)
Chicago Circle, and the University of Illinois at the Medical Center, were
reorganized into equal administrative campuses within a university system with
a central administrative staff in Urbana. a Works published by these bodies
after the reorganization in 1966 are found under a University of Illinois at
Urbana-Champaign. a University of Illinois at Chicago Circle. a University of
Illinois at the Medical Center. a University of Illinois (System) a SUBJECT
ENTRY: Works about these bodies are entered under the name or names in
existence during the latest period for which subject coverage is given. In the
case where the required name is represented in this catalog only under a later
form of the name, entry is made under the later form. w n010810616aanann----
nnnn
 14 667      The following earlier heading is also a valid AACR 2 heading: a
Illinois Industrial University. w n004790829aanann----nnnn
```

Example 6

More, it is quite possible that if Boeren's works are published in America or if LC suddenly decides that Boeren is "famous," the heading will be changed. This is an infrequently encountered problem for us but one where LC's peculiar policies have created problems that have nothing to do with OCLC or *AACR2*.

CONCLUSION

The problems that we have cited above are real but not numerically significant (except in the case of serials and multiple personal names— neither of which are under OCLC's control). They are far outweighed by the tremendous value of the more than 40 percent of OCLC headings that have been converted to their *AACR2* form.

The OCLC conversion has made it possible for us to do *AACR2* cataloging more quickly than in the period November 1979–December 1980. We have issued guidelines to our professional, paraprofessional, and clerical cataloging staff who deal with all the headings we encounter in using OCLC (see appendix).

Problems such as those we have described are dealt with in our guidelines, and in practical terms now in day-to-day work. They may take some extra time, but overall our cataloging operation has been greatly speeded by OCLC's conversion.

REFERENCE

1. *Cataloging Service Bulletin*, no.6:6(Fall 1979)

APPENDIX

University of Illinois Library at Urbana Champaign
Copy Cataloguing Guidelines

Authority Records

LC authority records, now available on OCLC, can be very helpful in determining the correct AACR 2 form of headings, and should be cited on authority cards we prepare, when we use them in establishing headings.

The tag numbers used on authority records sometimes have different meanings from the numbers used on bibliographic records. The meanings are:

    1xx Heading
    4xx See reference (i.e. from the form in this field to the form in the 1xx field)
    5xx See also reference (i.e. from the form in this field to the form in the 1xx field)
    6xx Notes (e.g. the authority used by the LC cataloguer)

Each field concludes with a w subfield, consisting of 24 characters indicating in coded form various types of information about the heading. The 13th character, the 3rd past the six-character date, consists of one of five letters indicating the rules governing the form of heading in that field. The codes are:

    c AACR 2
    d Compatible with AACR 2
    b AACR, 1967 ed.
    a Earlier rules (e.g., ALA rules of 1949, etc.)
    n Not applicable or not applied

Here is an example of an LC authority record, omitting the fixed field and some of the references:

    010   n 790558820
    110 20 State University of New York at Buffalo. w n008801115aacann----nnnn
    410 10 Buffalo. b University w n002791105aaaann----nnna
    410 10 New York (State). b State University, Buffalo. w n009801115aaaann----nnna
    667   The following heading for an earlier name is a valid AACR 2 heading:
           University of Buffalo. w n007791105aanann----nnnn

When OCLC carried out its AACR 2 conversion project, the data about the rules encoded in subfield w was added to headings in bibliographic records, if those headings were altered by the conversion. For bibliographic records in OCLC, subfield w contains 2 characters, each of which must be one of the following:

    c (for AACR 2 heading)
    d (for ACCR 2 compatible heading)
    m (for machine converted heading)
    n (not applicable or not applied)

The first character applies to the name portion of the heading; the second, to the title portion. Obviously, in many cases there is no title portion, in which case the second character will be n. The code m (machine converted heading) is used when a heading is altered directly by program, rather than being extracted from an authority record. An example would be the elimination of subfield k Laws, statutes, etc.

1. USE OF SUBFIELD W IN CATALOGUING

Since OCLC does not want member libraries to apply the letter codes in subfield w for their original input, the presence of a c or d in subfield w should always indicate an LC decision identifying an AACR 2 or AACR 2 compatible heading.

Supply subfield w for all cataloguing to be added to OCLC's data base. The codes to be used are given in ILLINET's Information Bulletin #92, from which this table is copied:

    1 AACR 2 form found in on-line LC Name-Authority File
    2 AACR 2 compatible form in on-line LC Name-Authority File
    3 AACR 2 form supplied by inputting institution with copy in hand and piece *not* in hand
    4 AACR 2 form supplied by inputting institution with piece in hand
    5 Author or title portion of heading not converted to AACR 2 form.

This subfield (#w) is always the last subfield in the field. It must contain a two character code. The first character applies to the name portion of the heading; the second character applies to the title portion of the heading. If the heading is a name heading and does not include a title portion, use "n" as the second part of the code. If the heading is a uniform title heading, use "n" as the first part of the code. Examples:

    700 10 Day Lewis, C. #q (Cecil), #d 1904-1972 #w In
    600 10 Schmidt, H. R. #q (Heinrich Rudolf) #w 4n
    130 00 Bible. #p N.T. #s Authorized. #f 1974. #w n4

Accept headings coded c in subfield w as correct AACR 2 headings, unless the heading is for an author entered under surname who writes in a non-Roman alphabet language. For such

authors, use the form given only if it is a standard romanization of the name in the original alphabet. If a form other than the standard romanization is used, substitute the standard romanization, and trace an x ref. from the form coded c.

2. LC AUTHOR HEADINGS WITHOUT DATES

LC recently announced that it will not add dates to a heading already established without dates, unless the dates are needed to resolve a conflict. When there is no conflict, the dates will be recorded in the authority record in a 6xx field, but will not be added to the heading. Dates will be routinely added to newly established headings at the time the headings are established, if the information is readily available. LC codes such headings c, not d, because AACR 2 does not require that a date be added to the heading, except to resolve a conflict. If such an LC authority record is available when a heading is being established, use the LC form, without adding dates to the heading, unless dates are needed to resolve a conflict in the new catalogue. Record the dates on an authority card. If LC authority is not available when a heading is being established, use dates in the heading if the information is readily available. If, later, LC authority is found that omits date from the heading, do NOT change the heading as already established for the UIUC new catalogue. Since records in OCLC may contain headings without dates for persons we have established with dates, some conflicts will be generated. These should be resolved by catalogue maintenance staff, who will add dates in pencil to headings on new cards that lack dates, but are otherwise identical with headings in the new catalogue. Such conflicts in the machine record will be cleaned up gradually, after FBR is up.

3. ACCEPTABLE dn FORMS

Headings coded d in authority records (dn in bibliographic records) are the AACR 2 "compatible" forms. In many cases, the difference from AACR 2 is trivial, and the form can therefore be used. In such cases, if LC authority is available, use the form as established by LC, and record the information on an authority card. If LC authority is not available when a heading is being established, follow AACR 2. If, later, LC authority is found that establishes a "compatible" form, do NOT change the form in the UIUC new catalogue to the LC "compatible" form. It will sometimes happen that "compatible" forms will be found on records in OCLC (coded dn, usually). Such headings may be used *only* if they fall into one of the categories listed below. This will sometimes result in "compatible" forms and true AACR 2 forms both being used in the new catalogue. In some cases, the two forms can be interfiled; in other cases, Catalogue Maintenance staff will need to correct "compatible" headings in pencil. Acceptable dn forms are:

a. LC will omit hyphens between forenames if the heading has been established without hyphens, even though rule 22.1D2 would require hyphens. Use the LC form, if found. Catalogue Maintenance will interfile headings identical except for the presence or absence of hyphens.

b. LC will continue to place the abbreviation *ca.* after a date in the heading for a person, if the heading has already been established in that form, even though rule 22.18 specifies that the abbreviation should precede the date. Use the LC form, if found. Catalogue Maintenance will interfile headings identical except for the placement of the abbreviation *ca.*

c. LC will not correct the language of an addition to a personal name heading; i.e. will not change to the language used in the person's works. (E.g., a heading already established as *Louis Antoine, Father* will not be changed to *Louis Antoine, père*, even though the latter is the author's usage.) Use the LC form, if found. Catalogue Maintenance will correct conflicts in pencil, to the LC form.

d. LC will not change a personal name heading to a fuller form of the name, even if the shorter form is not predominant. Use the LC form, if found. Catalogue Maintenance will correct conflicts created by personal name headings that vary in fullness to the form to which a "see" reference has been made. If there is no "see" reference, Catalogue

Maintenance will refer the conflict to the appropriate cataloguing service.

e. LC will continue to use additions to surname headings supplied by cataloguers, for headings already established with such additions. Use the LC form, if available. Catalogue Maintenance will resolve conflicts by adding qualifiers in pencil to headings that are otherwise identical with the forms with qualifiers.

f. LC will continue to use titles of honor, address, or nobility with headings that have already been established with such titles, even though the authors do not use such titles. Use the LC form, if found. Catalogue Maintenance will resolve conflicts by adding qualifiers in pencil to headings that are otherwise identical to the forms with the qualifiers.

g. LC will not use initial articles in uniform title and corporate headings, even when they are required by AACR 2. We will follow LC practice in this, and use the LC form when found. Catalogue Maintenance will interfile uniform title and corporate headings that are identical except for the presence or absence of initial articles.

h. LC will continue to use the abbreviations Bp. and Abp. for personal name headings that have already been established with those abbreviations used as qualifiers, instead of spelling out the qualifiers in full. Use the LC form, if found. Otherwise, follow AACR 2 and spell out "Bishop" and "Archbishop". Catalogue Maintenance will resolve conflicts by correcting in pencil to the form spelled out in full.

i. LC will not add terms of incorporation to corporate headings already established without them, nor delete them from corporate headings already established with them, even though LC interpretation of AACR 2 would require such adjustment. Use the LC form, if available. Otherwise, retain terms of incorporation in corporate name headings only if the term is an integral part of the name, or if, without the term, it would not be apparent that the heading is the name of a corporate body. Catalogue Maintenance will resolve conflicts by adding, in pencil, terms of incorporation to headings identical to established forms except for the absence of such terms.

j. LC will not add geographic qualifiers to corporate headings established previously without such qualifiers, even though they have chosen to apply the option in rule 24.4 that allows qualifiers to be added when there is no conflict. Use the LC form, if available. Catalogue Maintenance will resolve conflicts by adding qualifiers in pencil to headings identical to established headings except for the absence of such qualifiers.

k. LC will not reduce the hierarchy of Far Eastern corporate headings, established before 1981, even though AACR 2 rules would require that intervening superior bodies would be omitted from the heading. Use the LC form, if available. Catalogue Maintenance will refer conflicts to the appropriate cataloguing agency for resolution. The Asian Library Cataloguer is the final authority for such headings.

l. LC will not change the capitalization of acronyms and initialisms to conform to the usage of the corporate body, if the acronym has already been established with a different capitalization. Use the LC form, if available. Catalogue Maintenance will resolve conflicts by interfiling acronyms and initialisms that are identical except for variations in capitalization.

m. LC will not supply quotation marks around elements in a corporate heading that has already been established without quotation marks, even though this varies from the usage of the body. Use the LC form, if available. Catalogue Maintenance will resolve conflicts by interfiling headings identical except for the presence or absence of quotation marks.

n. If LC is attempting to resolve a conflict (i.e. two different people with identical author statements), and neither dates nor expanded initials are available to resolve the conflict, LC will add an unused name in parentheses to the heading if the information is available. E.g.:

established heading: Smith, Elizabeth
new author: Elizabeth Smith

(New author's full name, Ann Elizabeth Smith, is available)
LC heading: for new author: Smith, Elizabeth (Ann Elizabeth)

Use LC forms if found in name authority file. Catalogue Maintenance will refer problems to the appropriate cataloguing agency.

4. UNACCEPTABLE dn FORMS

In a few cases, the AACR 2 "compatible" forms, coded d in authority records and dn in bibliographic records, are unacceptable in the UIUC Library. Instead, we will follow AACR 2 in constructing these headings, and record the LC form on authority cards when they are found. We will also make references from the LC forms, if they would file differently from the forms we use. For many of these, Catalogue Maintenance will have to refer conflicts to the appropriate cataloguing agency. In a few cases, Catalogue Maintenance can make the corrections on the cards. The unacceptable dn forms are:

a. LC will sometimes, but not always, continue to use headings established prior to 1981 with names spelled out in full, when the authors represent some of those names with initials. Follow AACR 2 in constructing headings for these names. Use initials in conformity with the authors' usage, and add the corresponding full names in parentheses, in subfield q, when the information is available. Whenever an element in a compound surname or a first forename is represented by an initial, make a reference from the fuller form. Usually, a reference will not be needed if a forename other than the first is represented by an initial.
b. LC will continue to add "pseud." to personal name headings already established with that qualifier. Do not use the qualifier "pseud." when establishing personal name headings, and delete the term from OCLC records that use it, including records added by LC. Catalogue Maintenance will resolve conflicts by lining out the qualifier "pseud." in headings.
c. LC will continue to add 20th century fl. dates to personal name headings already established with such dates. Do not use 20th century fl. dates when establishing personal name headings, and delete such dates from OCLC records that use it, including recorded added by LC. Catalogue Maintenance will resolve conflicts by lining out 20th century fl. dates in headings.

5. 87x FIELDS

One part of the AACR 2 conversion project by OCLC was the addition of fields tagged 870, 871, 872, or 873. These fields contain the pre-AACR 2 forms of headings that were changed by the conversion. OCLC participants can add 87x fields to records they enter into the data base. However, we will not supply these fields in our cataloguing.

6. AUTHORITY CARDS

Prepare authority cards whenever references are needed, and whenever an LC authority record for the heading is found, even if we do not use the LC form. Citation of the authority record takes the form: "LC Auth. Rec." followed by the record number and the indication, in parentheses, of the code for rules given in subfield w. Example:

Akademie der Wissenschaften und der Literatur (Mainz, Germany)

LC Auth. Rec. 80076417 (cn)

If the LC form differs from the form used as the heading in UIUC, give the LC form in parentheses, following the subfield w code. Example:

Abrahamson, Max W. (Max William)

LC Auth. Rec. 78064817 (dn) (Abrahamson, Max William)

It will sometimes happen, when establishing the heading for a corporate body, that an LC

authority record for a subdivision of the body you are establishing will give you the AACR 2 form of the body you are setting up. Precede the citation to the authority record with the word "From". Example:

United States. Environmental Protection Agency. Region V.

From LC Auth. Rec. 80159375 (cn)
(The LC authority record is for the Water Division of Region V)

7. REFERENCES

The basic rule for making references is given in AACR 2, rule 26.1: "Whenever the name of a person or corporate body or the title of a work is, or may reasonably be, known under a form that is not the one used as a name heading or uniform title, refer from that form to the one that has been used. Do not make a reference, however, if the reference is so similar to the name heading of uniform title or to another reference as to be unnecessary." Ultimately, this decision depends on the cataloguer's judgement. Usually, make a reference only if it would file differently from the established heading and from all other references. Refer from variant forms found in works catalogued for this library, and in standard reference sources. LC authority records will often suggest useful references. However, we may need references not traced by LC, and we may not need all of the references LC traces. Notice especially that LC authority records will often give a reference from the pre-AACR 2 form, even when it would file with the AACR 2 form. For example, the authority record for Akademie der Wissenschaften und der Literatur (Mainz, Germany) traces a reference from Adakemie der Wissenschaften und der Literatur, Mainz—the pre-AACR 2 form. These two forms would file together, so we do not need the reference.

We will trace "see also" references from forms that can legitimately be used as headings, whether or not they have been used yet in the UIUC library. We will no longer observe the former restriction, which allowed "see also" references to be made only if both headings had been used.

For further information on authority records and references, see the cataloguing manual, section A79.

AW:lgo

Arnold Wajenberg is principal cataloger and **Michael Gorman** is director, technical services, at the University of Illinois Library.

Communications

Automation and the Service Attitudes of ARL Circulation Managers

James R. MARTIN: University of Rochester Library, Rochester, New York.

The circulation function in our large academic libraries has undergone two important transformations since the turn of the century. The first of these is departmentalization; the second, automation. The departmentalization of the circulation function has tended to separate the circulation department from the library's educational and information functions, the more "professional" aspects of librarianship. Laurence Miller makes this point in his dissertation, "Changing Patterns of Circulation Services in University Libraries," which focuses on the rise of circulation departmentalization.[1] Miller surveyed large academic libraries to determine if certain services—reference, interlibrary loan, orientation, catalog assistance—were being withdrawn from the circulation function. After verifying a withdrawal of these services and identifying them as the "professional" ones, Miller drew the conclusion that circulation is therefore suspect as a professional activity.[2] His are generally held conclusions as Robert Oram suggests:

> Until recently, librarians have been reluctant to deal with circulation problems on an organized basis. The belief that circulation was, in part at least, custodial and clerical rather than managerial and professional underlies much of the reluctance to solve mutual circulation problems through a professional group.[3]

Paralleling this change in the circulation function's organizational setting, the mechanization of the circulation process has continued to move from the laborious and slow use of manual procedures and book cards toward the immediate updating and record keeping of the online system. Circulation automation has passed from the early days of simply mechanizing files (represented by the batch system) to the present, where libraries have the potential capacity to perform the complete circulation control process with real-time systems.[4] Sophisticated online systems have begun to truly control the complete circulation function.

The metamorphosis of circulation automation—from simple mechanization to full computerization—has had a tremendous impact on the technical side, the processes, of the circulation department. Likewise it may well have had impact on the service attitudes, priorities, and leadership of the department. The level of automation may relate to the circulation manager's attitudes and priorities, and in the words of an American Library Association committee, "the impact of automation might change the image of the circulation librarian."[5] As it automates, gaining control over its own processes, the circulation department and its manager may actually become more responsive to its users—more service oriented, more "professional."

In February 1980, a questionnaire was sent to circulation managers of all the ninety-eight academic libraries that hold membership in the Association of Research Libraries.[6] It sought to (1) identify the degree and state of automation of the circulation function, classified by the three system categories of manual, batch, and online systems, and (2) to capture opinions on the circulation manager's view of his management role and his attitudes on service issues and user demands. These attitudes were related to the three types of systems. Seventy-six questionnaires were returned, for a 78 percent response rate.

CIRCULATION DEPARTMENT CHARACTERISTICS

Circulation departments ranged in size from 4 to 78 FTE employees. The average department size was 18, the median 14.25. The number of students employed ranged from 0 to 175. Twenty-nine percent of managers said staffing was not adequate and 45 percent said they had to depend too heavily upon students. Fifty-seven percent of managers of manual systems responded that they had to depend too heavily upon students, versus 27 percent of batch and 50 percent of online managers. (Because of variations in what is counted, transaction volume figures are not particularly informative.)

CIRCULATION SYSTEM CHARACTERISTICS

The seventy-six responding libraries reported approximately thirty-two different system configurations. Thirty-nine percent of these systems were manual, 34 percent were batch, and 26 percent were online. Nineteen percent of the total were manual McBee systems and 15 percent were LIBS100 online systems. Manual systems had been in use an average of twenty-six years, batch systems an average of eight years (range: ten months to eighteen years), and online systems an average of three years (range: three months to eight years).

CIRCULATION MANAGER CHARACTERISTICS

Typically, the circulation manager in an ARL library is the head of a department. ARL circulation managers had held their positions from six months to twenty years. Five years was the average, but 68 percent listed five years or less. Gender was evenly distributed: thirty-eight males and thirty-eight females. The managers of manual systems were 43 percent male/57 percent female, those of batch systems were 54 percent male/46 percent female, and of online systems 55 percent male/45 percent female. Seventy percent of all managers had an MLS, and 30 percent did not; 40 percent of managers of online systems did not have an MLS. A majority of circulation managers (57 percent) reported spending over 25 percent of their time on matters outside of strictly circulation concerns. In fact a substantial minority, 23 percent of all managers, spent over 50 percent on extracirculation matters.

SATISFACTION WITH CIRCULATION SYSTEM

As a group, ARL circulation managers are not satisfied with their systems, as table 1 shows. Online-system managers consistently rate their systems most highly.

Asked if their systems were "close to ideal," only 17 percent of all respondents were affirmative. Only 3 percent of manual-system managers agreed that their system was "close to ideal"—as compared to 12 percent of batch managers and 45 percent of online managers. Hidden in these averages is the fact that three managers gave their systems perfect scores on all four questions and those systems were all online: GEAC, LIBS100, and an IBM-based online system. (Table 2 summarizes responses on the four system-performance statements.)

HARDWARE, SOFTWARE, AND DOWNTIME

Circulation managers with automated systems also reported on their experience with equipment, software, and downtime. Batch-system managers were more satisfied with hardware and software (74 percent for both) than were online managers (60 percent satisfied with hardware and 65 percent with software). However, open-ended questions revealed that dissatisfaction with online-system hardware and software centered around limitations of the LIBS100 system (used by 55 percent of online-system managers). The LIBS100 system was panned for "inflexible software," "poor fines system," and "lack of reserve book features." (These are all long-recognized limitations that were partially addressed in the relatively recent Release 24.) The downtime situation was more satisfactory, however, for online managers than batch managers. Seventy-five percent reported downtime was not a problem as against more than 63 percent of batch-system managers.

Table 1. Responses by Type of System (N = 30 Manual, 26 Batch, 20 Online)

	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
"Our Circulation System Is Completely Adequate"					
Manual	1(3%)	4(13%)	1(3%)	12(40%)	14(40%)
Batch	1(4%)	5(19%)	1(4%)	13(50%)	6(23%)
Online	3(15%)	7(35%)	1(5%)	6(30%)	3(15%)
"Our Circulation System Is Reliable"					
Manual	1(3%)	15(50%)	1(3%)	10(33%)	3(10%)
Batch	3(12%)	9(35%)	-	11(42%)	3(12%)
Online	5(25%)	11(55%)	-	3(15%)	1(5%)
"Our Circulation System's Records Are Very Accurate"					
Manual	2(7%)	7(23%)	2(7%)	16(53%)	3(10%)
Batch	3(8%)	12(46%)	-	9(35%)	3(12%)
Online	4(20%)	10(50%)	-	6(30%)	-
"Our Circulation System Is Close to Ideal"					
Manual	-	1(3%)	-	7(23%)	22(73%)
Batch	-	3(12%)	-	8(31%)	13(50%)
Online	3(15%)	6(30%)	3(15%)	4(20%)	4(20%)

Table 2. Summary of Responses on Four System Questions (Detail Given in Table 1)

	Mean	Median	Standard Deviation	Minimum Value	Maximum Value	Variance
Manual	9*	9	3.27	4	16	11
Batch	10.08*	8.5	3.81	4	18	15
Online	13.45*	14	4.57	5	20	21

*20 = strongly agree, 16 = agree, 12 = no opinion, 8 = disagree, 4 = strongly disagree.

SERVICE ATTITUDES

Respondents were asked to mark attitude statements on a five-point scale: "strongly agree," "agree," "no opinion," "disagree," and "strongly disagree." Attitude statements fell into four categories: (1) specific service concerns, (2) the importance of the managerial role, (3) user problems, contacts and complaints, and (4) user demands and expectations. The averages of the last three groups were used to explore the question of association between level of automation and manager service attitudes (see table 3).

SPECIFIC SERVICE CONCERNS

Ninety percent of circulation managers agreed that "speed of service is very important to users," and no online-system manager disagreed. Forty-three percent of manual-system managers agreed that "control of circulating books tends to be inade-

quate." This compares to 16 percent of batch managers and 15 percent of online-system managers. Asked whether "users tend to expect more service than the department can give," 56 percent of manual managers agreed, as did 46 percent of batch managers and 40 percent of online-system managers.

ATTITUDES TOWARD MANAGEMENT ROLE

The study found that circulation managers are uniformly strong in their affirmation of the importance of their role, with a slight tendency for online managers to be more affirmative. In fact, 100 percent of respondents agreed with the statement that the "management of the circulation function is important." Ninety-three percent agreed that "circulation management should rank high among the library's priorities." Ninety-five percent disagreed with the negative statement that "circulation

Table 3. *Attitude Responses, Averages*

	Management Role (9 Questions)	*Demands and Expectations* (6 Questions)	*Contacts and Complaints* (6 Questions)	*Totals*
Manual	4.38	3.48	3.88	3.913
Batch	4.34	3.52	3.9	3.92
Online	4.48	3.46	4.03	3.99

5 = most positive response.
1 = least positive response.

management offers little opportunity for the exercise of initiative." Ninety-four percent of all managers disagreed that "circulation management lacks complexity."

ATTITUDES TOWARD USER PROBLEMS, CONTACTS, AND COMPLAINTS

The study found that circulation managers are uniformly strong in their desire to respond to user complaints and problems, but with a slight tendency for online managers to be more favorable to the user. One hundred percent of online managers regarded user contacts as pleasant, as did 93 percent of manual and batch managers. Ninety-five percent of online managers, 92 percent of batch managers, and 87 percent of manual managers affirm that patron contact provides the challenge in circulation work. Eighty percent of online managers and 73 percent of manual and batch managers rejected the statement that "complaints tend to be unfounded." Sixty-five percent of the respondents of online systems were more likely to favor the user by thinking "complaints are most often substantive," as compared to 50 percent of manual managers and 48 percent of batch managers. Ninety percent of online managers disagreed that users "complain far too much," compared with 84 percent of batch managers and 79 percent of manual managers.

ATTITUDES TOWARD USER DEMANDS AND EXPECTATIONS

Circulation managers are generally favorable in their attitudes toward user demands and expectations. Several statements in this area, however, ran contrary to the tendency of online managers to agree slightly more with attitudes favorable to the user than managers of batch and manual

systems. For example, while 93 percent of manual-system managers and 85 percent of batch managers agreed that "the circulation department should be oriented towards users' expectations," only 70 percent of online managers did. On the statement, "Users should be more tolerant of limitations in circulation services," manual managers disagreed by 34 percent, batch managers by 40 percent, and online managers by 20 percent.

These responses against the trend of the online manager as more user oriented may be due to the fact that the study was not completely successful in differentiating between responses based on general attitudes and those based directly on the specific system in use. In other words, the relative quality of each circulation system or even the "bugs" peculiar to a specific system may affect one's attitude toward the user's need to tolerate the limitations of that system. Manual-system managers know the limitations on their service are keyed to inefficient systems, whereas online-system managers know their systems and services are already at a high level. This knowledge of the system in use colors service attitudes.

CONCLUSION

The study found a depressed state of circulation-system development and support in ARL libraries. Seventy-four percent of circulation managers, on average, rated their systems negatively on basic system integrity, as shown in table 2.

The thirty manual-system managers gave their systems an average score of 9, to the effect that their systems were ideal, adequate, reliable, and accurate. The twenty-six batch managers gave their systems an average score of 10.08, the twenty online managers an average of 13.45. Recognizing the considerable constraints under which

today's large academic libraries struggle, there is, nonetheless, room for criticism of library priorities.

This study must be viewed as only a first step (largely tentative and exploratory) in relating automation with service attitudes. It suggests that online systems may be associated with managers more positive in their view of the management role and more positive in their attitudes toward users than batch- and manual-system managers. Further research would be useful at this point to compare levels of automation (manual, batch, and online) with circulation-staff service attitudes or those of patrons using the systems.

REFERENCES

1. Laurence Miller, "Changing Patterns of Circulation Services in University Libraries" (Ph.D. dissertation, Florida State University, 1971), p.iii.
2. Ibid., p.149.
3. Robert Oram, "Circulation," in Allen Kent and Harold Lancour, eds., *Encyclopedia of Library and Information Science*, V.5 (New York: Marcel Dekker, 1971), p.1.
4. William H. Scholz, "Computer-Based Circulation Systems—A Current Review and Evaluation," *Library Technology Reports* 13:237 (May 1977).
5. Robert Oram, "Circulation," p.2.
6. James Robert Martin, "Automation and the Service Environment of the Circulation Manager" (Ph.D. dissertation, Florida State University, 1980), p.22.

Statistics on Headings in the MARC File

Sally H. McCALLUM and James L. GODWIN: Network Development Office, Library of Congress, Washington, D.C.

In designing an automated system, it is important to understand the characteristics of the data that will reside in the system. Work is under way in the Network Development Office of the Library of Congress (LC) that focuses on the design requirements of a nationwide authority file. In support of this work, statistics relating to headings that appear on the bibliographic records in the LC MARC II files were gathered. These statistics provide information on characteristics of headings and on the expected sizes and growth rates of various subsets of authority files. This information will assist in making decisions concerning the contents of authority files for different types of headings and the frequency of update required for the various file subsets. The National Commission on Libraries and Information Science supported this work.

Use of these statistics to assist in system design is largely system-dependent; however, some general implications are given in the last section of this paper. In general, counts were made of the number of bibliographic records, headings that appear in those records, and distinct headings that appear on the records. The statistics were broken down by year, by type of heading, and by file.

In this paper, distinct headings are those left in a file after removal of duplicates. Distinctness will not be used to imply that a heading appears only once in a source bibliographic file, although distinct headings may in fact have only a single occurrence. Thus, a file of records containing the distinct headings from a set of bibliographic records is equivalent in size to a MARC authority file of the headings in those bibliographic records.

METHODOLOGY

These statistics were derived from four MARC II bibliographic record files maintained internally at LC: books, serials, maps, and films. The files contain updated versions of all MARC records that have been distributed by LC on the books, serials, maps, and films tapes from 1969 through October 1979, and a few records that were then in the process of distribution. The files do not contain CIP records. A total of 1,336,182 bibliographic records were processed, including 1,134,069 from the books file, 90,174 from the serials file, 60,758 from the maps file, and 51,176 from the films file.

A file of special records, called access point (AP) records, was created that contains one record for the contents of each occurrence of the following fields in the bibliographic records:

Type of Heading	*Fields*
personal name	100, 700, 400, 800, 600
corporate name	110, 710, 410, 810, 610
conference name	111, 711, 411, 811, 611
topical subject	650
geographic subject	651
uniform title heading	130, 730, 830, 630

Only the 6XX subject fields that contained LC subject headings (i.e., second indicator = 0) were selected as AP records. The main entry data string was substituted for the pronoun in the series (4XX) fields that contained pronouns. The AP records also contained information from the bibliographic records that assisted in making the counts, such as the date of entry of the record on the file, the identity of the type of bibliographic file, and the language of the bibliographic record.

A third file was derived from the AP file that contained a normalized character string for each AP record heading. These normalized AP records were used to produce the counts of distinct headings by clustering like data strings. Normalization included conversion of all characters to uppercase, and masking of diacritics, marks of punctuation, and other characters that do not determine the distinctness of a heading, but would interfere with machine determination of uniqueness. The subfields included in the normalized string, hence used for all heading comparisons, are given below. Only use-dependent subfields, such as the relator subfield, and those that belonged to title clusters in author/title headings were excluded. Examples of the AP file field contents and the normalized forms are:

AP field contents:

Chuang-tzu
Chuang-Tzŭ

[Blaeu, Joan] 1596–1673
Blaeu, Joan. 1596–1673
Blaeu, Joan, 1596–1673

Byron, George Gordon Noël Byron, Baron, 1788–1824
Byron, George Gordon Nöel Byron, baron, 1788–1824
Byron, George Gordon Noel Byron, baron, 1788–1824
Byron, George Gordon Noël Byron, Baron, 1788.1824

normalized forms:

CHUANG TZU
BLAEU JOAN 1596 1673
BYRON GEORGE GORDON NOEL BYRON BARON 1788 1824

Distinct headings for this study were determined by comparing on the following subfields:

Type of Heading	*Subfields*
personal name	a, b, c, d
corporate name	a, b, k, f, p, s, g
conference name	a, q, e
topical subject	a, b, x, y, z
geographic subject	a, b, x, y, z

All occurrences of repeating subfields were included. The relator data of subfields were dropped from personal and corporate name headings as were the title subfields in author/title headings. A separate study will examine the occurrence of author/title headings. Approximately 8 percent of the name headings in the files carry title subfields: 6 percent are series and 2 percent are author/title subjects or added entries.

Two types of distinct heading counts were generated for topical and geographic subject headings. One takes account only of main terms, the a and b subfields, excluding all subject subdivisions. The other compared the complete heading strings, including subject subdivisions.

CHARACTERISTICS OF THE FILES

The four bibliographic files from which the statistics were derived were begun in different years and are of unequal size. Table 1 presents the number of bibliographic records added to each of the MARC files by the year that the record was first entered into the file. The records added in the first months of 1979 have been eliminated from tables 1–3, thus the total number of records under consideration is 1,210,809. In the combined file, the records for books dominate the contributions from other forms of materials, representing 85 percent of the combined file records. After the addition of the films and serials records in 1972 and 1973, the total number of records added each year leveled off to around 115,000 but jumped to an average of slightly more than 150,000 records per year following the ad-

Table 1. Number of Records Added to Each File by Year

Year Entered	Book	Serial	Map	Film	Total
1968	11,812	0	0	0	11,812
1969	43,874	0	1,104	0	44,978
1970	86,004	0	3,467	0	89,978
1971	105,390	0	8,857	6,280	114,247
1972	73,437	0	4,665	6,280	84,382
1973	92,512	3,720	5,566	8,929	110,727
1974	99,004	10,682	6,246	8,457	124,389
1975	86,527	15,866	6,721	8,604	117,718
1976	120,106	19,098	6,876	5,432	151,512
1977	140,011	17,999	7,011	4,797	169,818
1978	169,044	12,643	5,584	4,464	191,735
Total	1,027,721	80,008	56,117	46,963	1,210,809

Table 2. Numbers of Headings and Distinct Name Headings Added to All Files by Year

	Number of Headings			Number of Distinct Headings		
Year Entered	Personal Names	Corporate Names	Conference Names	Personal Names	Corporate Names	Conference Names
1968	14,526	3,138	155	12,620	2,139	143
1969	53,134	21,206	1,027	39,184	9,364	909
1970	104,365	42,798	2,175	63,037	14,286	1,769
1971	129,617	57,496	2,742	64,029	15,216	2,158
1972	91,040	45,768	1,942	41,246	9,891	1,402
1973	118,188	57,847	2,625	48,703	12,653	1,862
1974	127,588	73,303	2,972	51,623	17,129	1,983
1975	113,622	76,417	2,519	50,291	18,135	1,742
1976	154,718	88,207	3,454	73,182	23,120	2,306
1977	182,860	87,985	3,487	89,353	23,906	2,333
1978	218,535	97,042	4,192	99,780	24,280	2,831
Total	1,308,193	651,207	27,290	633,048	170,119	19,438

Table 3. Numbers of Subject Headings and Distinct Subject Headings Added to All Files by Year

			Number of Distinct Headings			
	Number of Headings		First Terms Only		Full Headings	
Year Entered	Topical Subjects	Geographic Subjects	Topical Subjects	Geographic Subjects	Topical Subjects	Geographic Subjects
1968	10,615	1,857	4,390	489	7,775	1,512
1969	45,161	9,047	8,104	1,980	23,617	5,426
1970	89,304	21,054	8,170	4,263	34,526	10,179
1971	115,220	31,278	6,853	5,417	36,689	12,862
1972	92,247	20,760	4,236	2,597	26,201	7,074
1973	121,161	27,890	4,460	3,105	33,061	9,819
1974	137,843	31,814	4,524	3,553	39,262	11,413
1975	130,980	30,650	4,203	3,417	40,129	11,818
1976	168,840	39,886	5,125	4,142	55,468	15,472
1977	185,331	44,973	5,718	4,194	59,529	16,676
1978	222,565	49,923	7,151	4,034	69,856	17,855
Total	1,319,267	309,132	62,934	37,191	426,113	120,106

dition of major non-English roman alphabet language records in 1976. The increase is noticeable primarily in the books and serials files since the maps file had been adding those languages since 1969 and only a limited number of non-English-language audiovisual materials are cataloged. The unusually large number of records added to the books file in 1971 resulted from a special project to add retrospective titles to the file. The large increase in books records in 1978 was due to the COMARC project in which retrospective LC records that had been converted to machine-readable form by other libraries were contributed to the LC MARC file. Approximately 12,000 COMARC records were added in 1977 and 28,000 in 1978. The fall in numbers of film records produced in 1976–1978 reflects a general fall in production of instructional films in the United States.

Counts of items cataloged that are compiled by LC processing services from catalogers' statistics sheets show that LC cataloged approximately 225,000 titles in 1978; thus, approximately 73 percent of LC cataloging is currently going into machine-readable form. The principal exclusions are records for most nonroman material (only nonroman records for maps have been transliterated and added since 1969) and a few records for music, sound recordings, incunabula, and microforms. The portion being put into machine-readable form should rise significantly as the romanized records for items in several nonroman alphabets are added in the next year.

NAME HEADINGS

Table 2 presents the number of occurrences of name headings in the MARC bibliographic files and the number of distinct name headings, both by type of heading and by year. The number of distinct headings that were new to the file in a year was determined by comparing the headings added in a given year against those added in all previous years. It is not surprising to find that 66 percent of name-heading occurrences are personal names, 33 percent are corporate, and only 1.4 percent are conference. The figures shift when considering the distinct names, where 77 percent are personal and only 21 percent are corporate.

Looking at the total figures in table 2, while 1,308,193 of the headings that appeared on the records were personal names, only 633,048 or 48 percent of these were distinct. Of the rest, 52 percent were duplicates of the distinct headings. Similarly, 26 percent of corporate names were distinct, with 74 percent being duplicates; and 71 percent of conference names were distinct, with only 29 percent being duplicates.

In 1968, 87 percent and 68 percent of personal and corporate names, respectively, were distinct, i.e., 13 percent and 32 percent "had been used previously" when they appeared on a bibliographic record during the year. As the base file of names grows, the percentage of names appearing on new records but which "had been used previously" rises, to 60 percent and 77 percent in 1974. While the figures reported in table 2 indicate that the percentage of headings used that were repeats fell slightly again in 1977 (51 percent and 73 percent), this is probably due to the influx of new names with the addition of new languages in 1976–77. Additional statistics gathered on English-language items show the percentage of repeating headings becoming steady after 1974.

SUBJECT HEADINGS

Statistics concerning distinct topical and geographical subject headings were collected for main terms, excluding subdivisions, and for full subject heading strings. Table 3 gives the numbers of headings and the numbers of distinct headings of each type found in the MARC file. Looking at the total figures, only 4.8 percent of topical first terms are distinct, the rest are duplicates. This indicates an average occurrence of 20.8 times for each first term. Slightly more, 12 percent, of the geographic first terms are distinct.

When the full headings with topical, period, form, and geographic subdivisions are considered, the percentage of headings that are distinct rises to 32.3 percent for topical subjects and 38.8 percent for geographic subjects. Thus, 67.3 percent of topical and 61.2 percent of geographic are duplicates of existing headings. In the yearly figures, sub-

ject headings show the same tendency as name headings in that the percentages of headings that appear on new records but which "had been previously used" rises as the stock of headings increases and then levels off. Subjects were also affected by the addition of other roman alphabet languages in 1976–77 but not to a very large degree.

For all access points, name headings and full string subject headings, name headings account for 55 percent of the headings that occur in the bibliographic records, with only 45 percent attributable to topical and geographical headings. It should be noted that 12 percent of the name headings that appear on the bibliographic records are names used as subjects.

FREQUENCIES OF OCCURRENCE

Counts were also made of the frequency with which name headings occurred in the bibliographic files. Table 4 summarizes the frequency data: 66 percent of distinct personal names, 62 percent of distinct corporate names, and 84 percent of distinct conference names occur only once in the files. The percent of corporate names with single occurrences is surprisingly close to that for

personal; however, the percent of names having multiple occurrences falls more slowly for corporate than for personal names. While 5.47 percent of corporate names occur ten or more times, only 1.92 percent of personal names occur ten or more times. The figures for personal names roughly correspond to those obtained by William Potter from a sample taken from the main catalog at the University of Illinois at Urbana-Champaign. That study showed 63.5 percent of personal names occurred only once.[1]

The number of occurrences of different types of headings are compared in figure 1. The bars show the numbers of personal, corporate, conference, topical, and geographic headings that appear in the bibliographic files. The shaded areas represent the number of headings that are distinct, thus the upper part of each bar represents additional occurrences of the headings from the shaded area. For personal, corporate, and conference headings a further distinction is made between distinct headings that occur only once, the crosshatched area, and those that have multiple occurrences. Thus the multiple occurrences of corporate names may be seen to come from a small

Table 4. Frequency of Occurrence of Name Headings in All Files

Number of Occurrences	Distinct Personal Names		Distinct Corporate Names		Distinct Conference Names	
	Number	Percent	Number	Percent	Number	Percent
1	456,328	65.65	116,250	62.02	18,021	83.90
2	119,681	17.22	30,185	16.10	2,049	9.54
3	46,247	6.65	11,563	6.17	587	2.73
4	23,951	3.45	6,814	3.64	289	1.35
5	13,820	1.99	4,109	2.19	163	.76
6	8,790	1.26	2,958	1.58	98	.46
7	5,827	.84	2,175	1.16	56	.26
8	4,056	.58	1,673	.89	48	.22
9	2,998	.43	1,395	.74	36	.17
10	2,153	.31	10,037	.55	18	.08
11–13	4,116	.59	2,180	1.16	44	.20
14–20	3,748	.54	2,632	1.40	41	.19
21–50	2,678	.39	2,901	1.55	23	.11
51–100	448	.06	936	.50	4	.02
101–200	149	.02	374	.20	2	.01
201–300	47	.01	109	.06	1	.00
301–400	19	.00	46	.02	0	.00
401–500	11	.00	21	.01	0	.00
501–1000	5	.00	53	.03	0	.00
1001 +	2	.00	18	.01	0	.00
Total	695,074	99.99	187,429	99.98	21,480	100.00

number of distinct corporate headings, as was indicated by the slow decrease of the multiple-heading occurrence rate (i.e., a small group of corporate names have a very large number of occurrences).

FILE GROWTH

As a bibliographic file grows and the stock of names and subjects that are contained in the associated authority file increases, the number of new-to-the-file headings that are required for the new bibliographic records would be expected to fall. Figure 2 illustrates that tendency and shows that there is a leveling off of the number of new-to-the-file headings per new bibliographic record after the bibliographic file reaches a certain size. For example, after approximately 700,000 bibliographic records are in the file, for every additional 100 bibliographic records approximately 298 name and subject headings
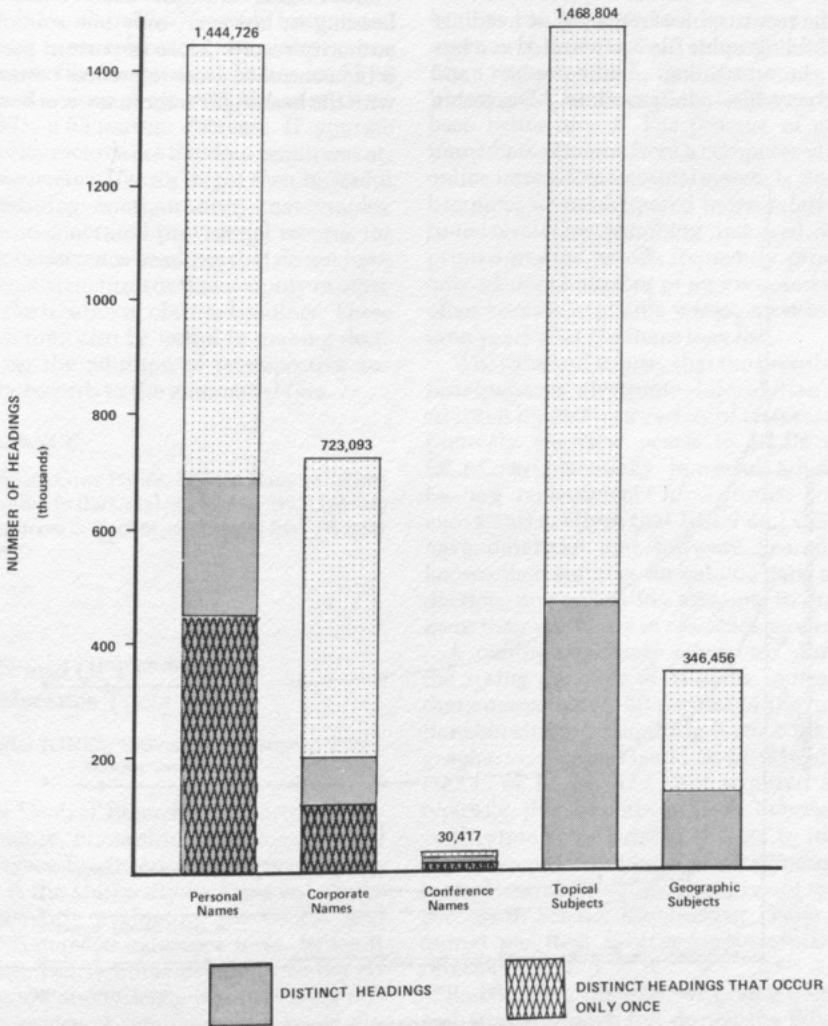
Fig. 1. Number of Headings by Type.

will be assigned, and, of these, approximately 53 will be new personal names, 14 new corporate names, 2 new conference names, 35 new topical subjects, and 10 new geographic subjects; the remaining 184 headings used will already be established in the authority file. Thus after a certain bibliographic file size is reached, the growth of the authority file is approximately a linear function of the growth of the bibliographic file.

IMPLICATIONS

The reoccurrence frequency of headings in a bibliographic file is often cited as a factor in designing bibliographic and authority-file configurations. Discussion

centers on the necessity of carrying authority records for headings that occur only once in a bibliographic file. With reference to the name-heading data in table 4 and figure 1, carrying authority records only for headings that occur more than once could potentially reduce the size of the authority file from that indicated by the whole shaded area (including shaded and crosshatched) to the plain shaded area, i.e., from 903,983 records to 310,123, a 66 percent decrease.

Controlling multiple occurrences of a heading is, however, only one role of the authority record. More important perhaps is the control of cross-references connected with the heading. Preliminary work with a
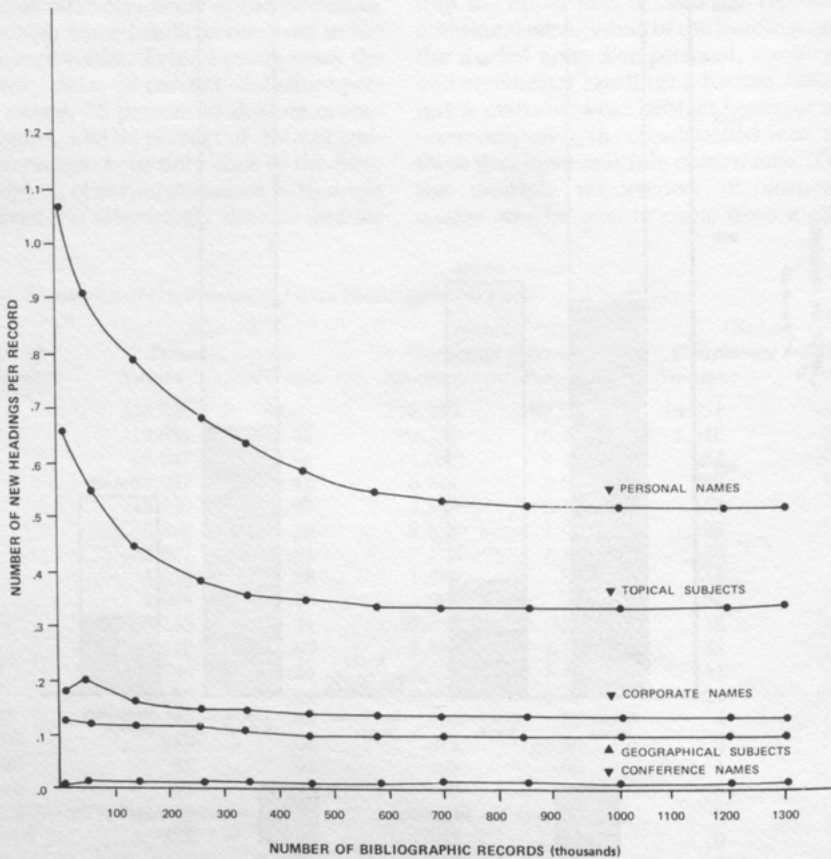

Fig. 2. Number of New Headings per Record for All Files.

random sample of personal names in the LC file indicates that less than 17 percent of personal names require cross-references. Thus the personal name headings that occur only once but would require authority records because of cross-references could be less than 17 percent. The frequency data combined with reference structure data could have a significant impact on design.

Out of a total of 695,074 personal names in the authority files associated with the MARC bibliographic files examined here, 456, 328, or 66 percent, occur only once. Of these, fewer than 77,575 would be expected to have cross-references, thus the name-authority file for personal names could be reduced in size from 695,074 records to 316,321, a 55 percent decrease. If separate authority records are a system requirement, the occurrence figures might then be useful for defining configurations that employ machine-generated provisional records for single-occurrence headings that do not have reference structures or that simplify in other ways the treatment of these headings. These figures may also be useful in making decisions on the addition of retrospective authority records to the automated files.

REFERENCE

1. William Gray Potter, "When Names Collide: Conflict in the Catalog and AACR2," *Library Resources & Technical Services* 24:7 (Winter 1980).

RLIN and OCLC as Reference Tools

Douglas JONES: University of Arizona, Tucson.

The Central Reference Department (social science, humanities, and fine arts) and the Science-Engineering Reference Department at the University of Arizona Library are currently evaluating the OCLC and RLIN systems as reference tools, to see if their use can significantly improve the effectiveness and efficiency of providing reference service. A significant number of the questions received by our librarians, and presumably by librarians elsewhere, in-

volve incomplete or inaccurately cited references to monographs, conference proceedings, government documents, technical reports, and monographic serials. If by using a bibliographic utility a librarian can identify or verify an item not found in printed sources, then effectiveness has been improved. Once a complete and accurate description of the item is found, it is a relatively simple task to determine whether or not the library has the item, and if not, to request it through interlibrary loan.

Additionally, if the efficiency of the librarian can be improved by reducing the amount of time required to verify or identify a requested item, then the patron, the library, and, in our case, the taxpayer, have been better served. The promise of near-immediate response from a computer via an online interactive terminal system is clearly beguiling when compared to the relatively time-consuming searching required with printed sources, which frequently provide only a limited number of access points and often become available weeks, months, or even years after the items they list.

We realize, of course, that the promise of instantaneous electronic information retrieval is limited by a variety of factors, and presently we view access to RLIN and OCLC as potentially powerful adjuncts to—not replacements for—printed reference sources. Given that RLIN and OCLC have databases and software geared to known-item searches for catalog card production, our evaluation attempts to document their usefulness in reference service.

A preliminary study conducted during the spring semester of 1980–81 indicated that approximately 50 percent of the questionable citations requiring further bibliographic verification could be identified on OCLC or RLIN. The time required was typically five minutes or less. Successful verification using printed indexes to identify the same items ranged from 20 percent in the Central Reference Department to 50 percent in Science-Engineering. Time required per item averaged approximately fifteen minutes.

Based on our findings, we plan a revised and more thorough test during the fall semester of 1981–82, which will include an assessment of the enhancements to the

RLIN system scheduled to be operational this summer. The proposed test will involve eight members of the reference staff—four from each department—who will be trained to search on OCLC and RLIN. Those selected will include both librarians and library assistants who regularly provide reference assistance. The results obtained from such a representative group will better enable us to assess the impact on the whole reference staff should we later decide to fully implement the service. They will be the only ones involved in sampling questions and conducting comparative searches.

The test will have two components, the first of which will be a twenty-week period to collect at least 400 sample questions. During their regularly scheduled reference hours, the eight specially trained librarians will collect samples of reference requests for materials that, based on the information initially given by the patron, cannot be identified in the card catalog. After checking the catalog, the librarian will then complete the top portion of a two-page self-carbon form with all of the information that is known about the requested item. Then, at regular intervals during the semester, the pages of each form will be separated and distributed to other members of the test staff for batch-mode searching. The manual OCLC and RLIN searching for each query will be done by different staff members to eliminate crossover effects. Each request will be searched on both OCLC and RLIN with the following information being recorded:

1. Date of the material requested (if known).
2. Type of material (e.g., conference proceeding).
3. Amount of time required to do the search.
4. Success or failure of the search.

This information will then be cumulated in a statistical table, and the results of each search will be keypunched for computerized analysis using the BMDP (BioMedical Computer Programs) statistical package to determine whether or not effectiveness and efficiency have been improved significantly.

In addition, on twenty-four randomly selected days during the semester the trained searchers will count the total number of questions received by them on that day that would have been appropriate to search on RLIN or OCLC. By using these data it will be possible to extrapolate the potential usefulness of the systems for the entire semester.

The second component of the test will be a two-week real-life test during which all questions requiring further verification would be searched immediately on RLIN, OCLC, and in the appropriate printed sources to compare time required, success rate, and type of material requested. This sort of test would permit the searcher to continue to negotiate with the patron as the search progressed, which is the usual situation. Also, this would provide the only opportunity to have the patron judge the value of subject searches done on RLIN.

If funding is received, preliminary results should be available in early 1982. Anyone conducting similar or otherwise relevant studies is asked to contact the author.

Replicating the Washington Library Network Computer System Software

Thomas P. BROWN: Manager of Computer Services, and Raymond DeBUSE: Manager of Development and Library Services, Washington Library Network, Olympia.

The Washington Library Network (WLN) Computer System supports shared cataloging and catalog maintenance, retrospective conversion, reference, COM catalog production, acquisitions, and accounting functions for libraries operating within a network. The system offers both full MARC and brief catalog records as well as linked authority control for all traced headings. It contains more than 250,000 lines of PL/1 and IBM BAL code in more than 1,100 program modules and runs on IBM or IBM-compatible hardware with IBM operating systems (MVS,OS/VS1). All database management functions are provided by ADABAS, a product of Software A.G. of North America. The online system runs un-

der CICS/VS 1.5. A set of assembler codes called the TP Monitor Interface defines a standard service interface between the applications programs and the TP monitor. This allows easy upgrade to different TP monitors and convenient points for collecting performance statistics.

The majority of the Bibliographic Subsystem updating is done in batch mode to conserve online resources. A new version of the system with interactive updating is currently being planned, for use in special environments.

The applications software was designed and implemented with a number of important conventions:
1. Top-down design.
2. Standard use of IBM environments.
3. Structured coding techniques.
4. Interfaces to a database management system (ADABAS) and teleprocessing monitor (currently CICS).
5. Standard naming and formatting.
6. Use of a standard set of data structures and assembler subroutines to manipulate data.
7. Identification of maintenance changes in source programs.

In addition, programming for the online functions meets other conditions:
1. Load modules less than 20K bytes.
2. No PL/1 run-time subroutines.
3. Reentrant coding.
4. Standard services for the TP Monitor Interface.
5. Applications are kept as terminal independent as possible, with the TP Monitor Interface performing input and output translations.

REPLICATION

A system with these characteristics, even though large, can easily be transported to a different site. While WLN was not designed with multiple replications in mind, a policy decision made by the network a few years ago made replication an attractive possibility. Recognizing that it had a capability that would be highly competitive with other online shared bibliographic services, WLN expanded its service area beyond the state of Washington. It set limits to its expansion, however, having determined that it would remain a small, responsive organi-

zation providing what it hoped would be superior service to its participating libraries. Having set such limits, however, created two impediments to its achieving superior service: WLN would have a smaller base of libraries from which its participants could obtain the benefits of shared cataloging, and there would not be the fiscal resources necessary to support a large continuing development effort. Both would penalize libraries for joining WLN, the first with a lower hit rate against the database and the second with fewer added capabilities.

Replication provided a possible answer to both of these problems, as well as a potential source of income. In its software license agreements, WLN asks the licensee to agree to bibliographic data sharing. All cataloging done by a licensee or its participants would thus be available for loading on WLN's own database; likewise, all WLN participant cataloging will be made available to the licensee. WLN, at least, would accept catalog records only from libraries that follow its bibliographic standards; that is, the standards of the Library of Congress. Currently this sharing is accomplished via magnetic tape, but in the future, online record interchange may be possible, given WLN's current work in this area.

WLN also explicitly asks in its software license agreements that the replicating institution carry out an organized program of development to meet the latter's particular needs. Such development is monitored by WLN in order that redundant work is not undertaken and to ensure that the various efforts relate coherently. There is a built-in constraint upon major modification and redesign: WLN is packaging enhancements and changes into periodic releases of the source code and requiring that the replicants install these releases within twelve months of the date issued.

Because of the interest in shared development and because WLN itself is not in a position to provide first-level program maintenance, the system is distributed in source-code form. The initial installation, however, is of load modules (programs in a form efficiently read and executed by machine), allowing immediate testing of the system's capabilities in its new environ-

ment. WLN is currently negotiating a contract with a new firm, Biblio-Techniques, that will offer a more nearly turnkey version of WLN, packaged with ADABAS and Software A.G.'s TP monitor, COMPLETE, and, if necessary, with the required hardware.

NATIONAL LIBRARY OF AUSTRALIA

The first replication of the system was made at the National Library of Australia (NLA) in Cranberra in early 1979. NLA had its own IBM 370/148 and an established data processing staff. ADABAS had been installed prior to the arrival of WLN's installation consultant. Minor changes are necessary in CICS to support dedicated WLN terminals, and these were quickly made and the system was up within days. Further work allowed searching on the system from the library's 3270 terminals. After a couple of weeks of shakedown, a WLN staff member spent about two weeks training NLA staff in the use of the system. It has been in full production for in-house production cataloging for more than a year now, and this spring is being extended to other libraries around the country on a pilot basis, testing the concept of the newly defined Australian Bibliographic Network (ABN). NLA has replaced the 370/148 with a larger machine.

UNIVERSITY OF ILLINOIS

The second installation occurred earlier this year at the University of Illinois, where the system was obtained to carry out a pilot project in which the Urbana campus and the River Bend Library System will use it as an online public-access catalog in conjunction with the LCS circulation system. Again, load modules were installed and the system was up within a few days, running on the University's administrative computer at the Chicago Circle campus. Illinois staff have had some difficulties in recompiling all of the source code, but these problems are being worked out. WLN will warrant that the source code supplied corresponds to the load modules it installs. The system as presently distributed by WLN can in no way be considered turnkey. Local computer operations or JCL require-

ments as well as differing levels of staff expertise can create problems. Furthermore, WLN handles source management through WYLBUR, a text-editing system, and this is not included with the WLN software. The module descriptions, programming language, mode, link-edit information, etc., maintained through this facility are supplied, however. Either a test or, if contracted for, a full database is also supplied. WLN has had some difficulties in creating a valid test database for Illinois, but has now defined procedures to better control the process.

WLN has distributed its second release to Australia as a full source update identical to what was installed at Illinois. In the future only the source changes in standard IBM IEBUPDTE form will be supplied to replication sites. This will better enable these sites to integrate the new version into theirs.

OTHER SITES

The University of Missouri is likely to be the third replicant of the system, since it has just selected WLN as the basis for its online catalog system. Installation is planned before the end of 1981. The National Library of New Zealand has also indicated that it intends to purchase the system. The Southeastern Library Network (SOLINET) has obtained the system in order to convert it to a Burroughs facility. While this is a software license, it is not a replication. The resulting system, however, would be available from WLN for installation on Burroughs equipment.

WLN has not implemented data sharing with Australia, but is testing the loading of Illinois data into its bibliographic file. WLN libraries should see Illinois records on a regular basis by late summer of 1981. Similar arrangements will be made with Missouri and SOLINET.

Shared development has gotten off to a start with the National Library of Australia having done the work necessary to add the IBM 3270 type of terminal to those that can support cataloging input and edit on WLN. Illinois will be undertaking the development of enhancements to make the system easier to use as a public online catalog, in addition to other possible areas of concern. WLN, of course, continues its in-house de-

velopment, which has recently seen the implementation of a new batch retrospective-conversion subsystem, and added COM catalog options and online authority verification during input/edit.

While not the only bibliographic system to be successfully replicated, the WLN Computer System is becoming the most systematically replicated main-frame facility, with a broad range of future possibilities, including that of a truly turnkey system. WLN's experience indicates that, if a system is designed for ease of maintenance at perhaps some sacrifice of efficiency, it will be readily transportable and allow others to obtain the benefits of a highly sophisticated bibliographic capability without the ever-increasing cost of original development and, more importantly, without having to support the ongoing maintenance of a unique system.

A General Planning Methodology for Automation

Richard W. MEYER, Beth Ann REULAND, Francisco M. DIAZ, and Frances COLBURN: Clemson University, Clemson, South Carolina.

INTRODUCTION

A workable planning methodology is the logical starting place for the successful implementation of automation in libraries. An automation plan may develop on the basis of an informal arrangement or from the efforts of one individual, but just as often, automation plans are developed by committees. An automation planning committee must determine and execute some kind of planning methodology and is more likely to be successful if it starts with clear guidelines, good leadership, and a thoroughly proven approach.

As a summary review of the literature will bear out, many libraries have developed their own planning techniques in-house. Some of these, which are addressed to the issues of cataloging rule changes and public-access catalogs, have been very well thought out.[1] However, these techniques are generally not directed to planning for

library-wide automation, and are usually designed to meet the specific needs of an individual library. Although the pattern for these studies is often similar, they do not seem to be based upon any general automation design methodology. Neither, in addition, does there seem to be a general methodology available through any external library agency. The Office of Library Management Studies of the Association of Research Libraries has developed a number of programs designed to assist libraries with their planning efforts, some of which appear to be useful in automation development.[2] But for many libraries, these programs may be too broad, too time-consuming or too expensive. As an alternative, some libraries will need to look elsewhere for a general automation planning methodology. This problem was addressed by the administration of the Clemson library, and was resolved in a unique way.

BACKGROUND

The Robert Muldrow Cooper Library of Clemson University has the responsibility of acquiring, preserving, and making available for use the many materials needed by faculty and students in their research and instructional efforts. At a typical land-grant institution like Clemson, the amount of scholarly publishing and the pressure to develop research proposals has risen sharply in recent years. The increased needs of users working with an expanding and diversified collection have resulted in a doubling of circulation activity, and have required the growth of library staff by 70 percent over the last decade. Furthermore, acquisition, processing, and access problems are compounded by the high inflation rate of materials, particularly serial publications, and manpower costs.

Even though user demands heavily burdened the traditional manual systems, the extent of library automation at Clemson had been limited to a batch circulation system, a simple serials-listing capability, and the use of bibliographic utilities. Although it had been generally accepted for some time that the acquisitions and fund-control functions at Clemson were in need of automation, no concrete approach to develop-

ing a system had been established. In addition, there was some concern that the development of an automated acquisitions system shouldn't be initiated without a clear understanding of how such an effort would affect the rest of the functions in the library. With this in mind, and as an initial part of planning, the library administration decided to implement a programmed study to determine specific needs and problems of the whole library at Clemson and to determine the attendant costs and benefits of their resolution. Since developing the methodology for this kind of study effort in-house has been shown by experience elsewhere to be both expensive and time-consuming, a planning methodology was sought which could be brought in from outside the library and applied in a timely fashion. The International Business Machines Corporation (IBM), through their local marketing representative, volunteered to supply that methodology by means of an Education Industry Application Transfer Team (ATT) study. In order to implement the study, a team was organized consisting of representatives from the library, from the university's Division of Administrative Programming Services (DAPS), and from the IBM Corporation. The purpose, approach, and results of that study constitute the rest of this paper.

PURPOSE

The Application Transfer Team methodology was implemented to fulfill a fourfold purpose.

• First, it was necessary to act on the recognized need for a library-wide automation plan with something tangible that library and university administrators could use in the decision-making process.

• Second, basic objectives and implementation estimates were required to provide groundwork to the development of systems specifications and evaluation.

• Third, the planning process needed to provide a forum for meaningful participation by a number of library staff and users.

• Fourth, the planning needed to be accomplished rather quickly.

The ATT met all these requirements. Although the ATT study technique is generalized for work on any problem in the educa-

tion arena, it seems particularly well suited to the library environment because it is oriented toward developing applications that solve production problems.

The Application Transfer Team methodology was developed by the IBM Corporation for customer use. The ATT methodology evolved from IBM's Business System Planning Function, which has been operational since the early 1970s. Although the methodology has been used several times in the academic environment, this is the first time, to our knowledge, that it has been used in a library operation. The strength of the ATT is that it helps members of a team with diverse backgrounds to understand the environment under study. Its final goal was "to improve operational productivity, provide better service to students, and provide information which can enhance management planning and decision making."[3] Put to work, the methodology is straightforward and effective. From beginning to end, the ATT process took Clemson slightly more than three months elapsed time. Total work time (including all report writing) for library staff was approximately one thousand man hours.

As the initial step with the ATT methodology, it was necessary to engage a sponsor and to select a team. For this study, the sponsor chosen was the dean of graduate studies, who reported directly to the vice-president for academic affairs. In turn, the director of computing and the director of the Division of Administrative Programming Services (DAPS) reported to the dean of graduate studies. Although it was not critical that the sponsor be intimately involved in the project, his level of authority within the university administration would help to secure acceptance of the study's recommendation. The sponsor also provided cogent advice along the way, based upon his understanding of institutional resources, and he served as a communication link with other university administrative offices.

The study team was chosen by the library administration with the intention of getting diverse involvement and expertise. Library staff included the associate director, the head of circulation, the serials cataloger, and a reference librarian. Although only the associate director brought significant

experience in library automation development, the head of circulation contributed substantial practical experience with automation systems. The cataloger offered specifics of bibliographic problems, cataloging rule changes, and serials control issues, and the reference librarian contributed a comprehensive knowledge of information-retrieval concerns.

Outside staff included the director of DAPS, who furnished details on the Clemson computing environment, and an IBM marketing representative, who provided appropriate help with hardware capabilities, the ATT methodology, and legwork. In addition, Clemson was also able to engage the help of a representative of IBM's Education Industry Division to guide the ATT efforts on the basis of his experience in the use of the methodology. From time to time, other IBM and DAPS staff were involved in assisting with interviews and report writing. The associate director served as team chair in order to act as spokesperson, to coordinate team effort, and to edit the final report.

METHODOLOGY

The Application Transfer Team methodology is applied in six phases. IBM recommends that these phases be conducted sequentially, and that they last from five to sixteen weeks, depending on the size of the problem. Throughout the process, verbal reviews were conducted by the team with the sponsor and with the library staff.

The first phase involved an organizational session. Following the introduction of team members, the IBM Education Industry Division representative presented an overview of the methodology and explained the mechanics of the ATT study process. The team then established the scope of the study by choosing an application area on which to focus and by determining the general objectives of the final system to be implemented. Since part of the purpose of the project was to develop a plan for library-wide automation, it was quickly recognized by the team that the application area should be an integrated library information system. However, the IBM representative suggested that this scope was too broad for the study and that one functional area such as

acquisitions be chosen, with other functions reserved for subsequent ATT studies. Given time constraints, a compromise arrangement was made in which serials control was determined as the scope. Since serials control is a single functional area, but encompasses nearly all bibliographic issues, it served as a microcosm of overall library operations. Therefore, it was generally accepted that a plan that effectively accommodated serials would constitute an integrated system plan. The organizational phase continued by determining who to interview during the data-collections phase and by setting up an interview schedule. This phase was concluded by developing an outline of the final report and by assigning writing responsibilities to individual team members.

The data-gathering effort constituted phase two. This involved structured interviews of representative staff of each unit of the library who were involved in routine interactions with any phase of serials control at Clemson. Interviews were conducted with staff from acquisitions, cataloging, circulation, reference units, and branch libraries as well as the university business office, students, and faculty.

Following an outline in the ATT, each person interviewed was asked for specific details of his work with serial publications regarding (1) interfaces (or points of interaction), (2) concerns or needs, (3) suggested improvements, (4) expected values or benefits of improvements, (5) work volume, and (6) cycles. Data gathered in each of these interview sessions were immediately documented in a letter to the interviewees. These letters were reviewed by those interviewed for corrections and added detail.

Data from completed and documented interviews were consolidated during the third phase of the study into a matrix of each of the six questions plotted against operational areas of the library, graphically designating areas of the greatest concern to the largest part of the library. This composite was analyzed to separate problems that could be reasonably handled by an integrated automation system from those that needed the attention of administrative policy and direction. Functions for automation consideration were then examined in a

"blue sky" session of the committee to envision what system would accommodate the specifications for serials control and access that each library unit and serials user required. From this session a synthesis emerged of the architecture for an integrated system.[4] This architecture included a description of the basic relationships of functional modules of the system, a list of the various files needed to contain system information, and a list of data elements required for bibliographic holdings, acquisition, and patron records in the system database.

Phase four called for the translation of the architecture and general system requirements into modules on basic access, acquisition or processing functions, and into the individual programs needed to execute each module. The team divided into two parts. The IBM and DAPS personnel, with the associate director, listed the modules and programs and formulated descriptions of each. Part of the description effort involved drafting approximate flowcharts of each program. Using algorithms developed by IBM, these descriptions were used to assign estimates of person hours required to create the necessary modules. In order to determine the overall cost of system development the person-hour figures were converted to dollars using an average hourly cost for Clemson DAPS personnel.

Committee members not involved in program/module design formed a group to evaluate anticipated benefits defined in the interviews, to collect data from library staff to support these expectations, and to assign a value to them. Benefits from reduced file maintenance, processing, and tracking time were valued as person hours saved by the new system. Additional improvements were projected for the system's capability for better fund control, more complete and immediate on-order, claiming, and in-process information, and statistical collection development/use data. These benefits were assigned the value of estimated duplicate and inappropriate material acquired under the present system. A value was not assigned to user benefits. Faculty and student satisfaction is intangible, and variable from case to case. Enhanced user service was recognized as a substantial benefit of the proposed system, but was not quantified. The cost factors determined in phase four were consolidated with derived benefit values to form a cost/benefit analysis, which constituted phase five.

In the sixth and final phase an implementation plan was formulated. This plan, along with recommended target dates, was presented orally to library staff and university administration. In addition, the entire process, recommendations, and plan of action were documented in a written report.[5]

RESULTS

Within the ATT report were a description of the current library environment, objectives and description of the proposed system, implementation considerations, a cost/benefits analysis, and recommendations for a plan of action. Although care was taken to "walk through" the function of each module of the described system, the report was not intended to provide detailed computer program specifications ready to be coded by a programmer. It described a useful and powerful integrated serials system in sufficient detail to be a working tool in the hands of a knowledgeable systems analyst to match (or revise) already available systems and programs to the library's specifications. The report itself also served as an effective communication link with the university administration, setting out library concerns and giving rational solutions to the pervasive problem of serials control and, in the long term, to an integrated library information system.

The timing of the ATT study was fortunate for the Clemson library. The university was on the eve of an accreditation self-study. As often happens with the examination of any organization, a host of related, but unacknowledged, problems surfaced in the course of the ATT study. During the interviews, staff members felt free to bring up matters of unclear policies, misunderstood hierarchical arrangements, and staffing inadequacies throughout the library. The number and importance of non-

automation concerns was significant enough that an administrative report was written to articulate these problems to the university administration.[6] It is interesting to note also that, while in every instance the team received enthusiastic cooperation from all those interviewed, there was fear among some staff members that any automation project would necessarily cut staff positions. Once this worry was identified, the study team was able to allay those fears by explaining the study's purpose.

One of the greatest contributions of the ATT study has been the direction it has given the library for future goals and priorities. By focusing on the problems of serials control, the team evaluated a microcosm of library problems. Investigating these problems in the environment of more limited budgets, possible future closing or freezing of the card catalog, and increased user demands for services has helped the library develop a course of action, a resolve of mission, and a direction for future growth.

The staff of DAPS and the library are conducting a review of existing software and systems potentially appropriate for a comprehensive serials control system. The ATT study was the tool successfully used to elicit university support for library automation. The university has given its approval, and supplied funding, to proceed with the determination of available systems and with the development of a request for quotation.

REFERENCES

1. For example: University of Rochester, River Campus Libraries, Task Force on Access Systems, *Report* (Rochester, N.Y.: Univ. of Rochester, 1980), University of California, Berkeley, General Library, Committee on Bibliographic Control, *Future of the General Library Catalogs of the University of California at Berkeley* (Berkeley: Univ. of California, 1977); Pennsylvania State University Libraries, Systems Development Department, *Remote Catalog Access System: General System Specifications* (University Park: Pennsylvania State Univ., 1977).

2. Association of Research Libraries, Office of Management Studies, *Annual Report, 1979* (Washington, D.C.: The Association, 1979).

3. International Business Machines Corporation, *Application Transfer Teams: Application Description* (White Plains, N.Y.: The Corporation, 1977), p.1; International Business Machines Corporation, *Application Transfer Teams: Realizing Your Computing Systems' Potential* (White Plains, N.Y.: The Corporation, 1977).

4. International Business Machines Corporation, *Business Systems Planning: Information Systems Planning Guide* (White Plains, N.Y.: The Corporation, 1975), p.49.

5. Richard W. Meyer and others, *Total Integrated Library Information System: A Report on the General Design Phase* (Syracuse, N.Y.: ERIC Clearinghouse on Information Resources, 1980), ED 191 446.

6. Richard W. Meyer, *Cooper Library: Status and Agenda. A Report on FY 1979–80* (Clemson, S.C.: Clemson Univ., 1980).

Reports and Working Papers

Inclusion of Nonroman Character Sets

The following document was prepared by staff of the Library of Congress as a working paper for discussions on incorporating the techniques described into the MARC communications format.

The document defines the principles for inclusion of nonroman alphabet character sets in the MARC communications format and the procedural changes needed to allow implementation of the principles. This technique was agreed upon at the MARBI Committee meeting on February 2, 1981.

Any questions on the description of the inclusion of nonroman character sets in the MARC communications format should be addressed to: Library of Congress, Processing Services, Attention: Mrs. Margaret Patterson, Washington, DC 20540.

1. INTRODUCTION

The cataloging rules followed by American libraries favor recording the title page data in the original script when possible. This helps those who consult catalogs to read the most essential information about the book. (Reading his or her name in romanized form is just as difficult for someone who knows Arabic as reading your name when it's written in Arabic.) The new cataloging rules also specify that names and titles in notes be given in their original script, AACR2 1.7A.3. Technological advances have made it possible to provide many, if not all, nonroman alphabets in machine-readable cataloging records. OCLC and RLIN are in the process of enhancing their systems so they can handle some nonroman writing systems. The Library of Congress has entered into a cooperative agreement with RLIN for the development and use of an augmented RLIN system for East Asian (i.e., Chinese, Japanese, and Korean) bibliographic data. Although the Library itself will not be creating and distributing MARC records with nonroman characters in the near term, the goal of this proposal is to define how these data can be included now so others can do so soon.

The technique known as an escape sequence announces that the codes which follow will represent letters in a specific different alphabet instead of the roman letters the codes would otherwise stand for.

2. PRINCIPLES

The following principles will govern inclusion of other alphabets in MARC records. Note that these deal only with the MARC communications format record, not the details of its processing—keying, sorting, display, etc.—by any bibliographic agency or utility. These principles are a slightly revised version of ones reviewed and approved in principle by the MARBI Character Set Committee in 1976. The earlier version was also distributed that year as working paper N77 of ISO TC46/SC4/WGI.

(1) Standard character sets should be used when available.
(2) Standard escape sequences should be used when available.
(3) Escape sequences should be used only when needed.
(4) Escape sequences are locking within a subfield but revert at any delimiter or field or record terminator code.

Example: (For demonstration purposes only, *EC* represents escape to Cyrillic and *EA* escape to ASCII)

245 10$a*EC*Russian title proper :$b*EC*Russian Subtitle. **F**

not

245   10$a*EC*Russian   title   proper :*EA*$b*EC*Russian subtitle. *EA***F**

and not

245 10$a*EC*Russian title proper :$bRussian subtitle.**F**

(5) Records which contain an escape sequence will also contain a special field which specifies what unusual character sets are present.

3. IMPLEMENTATION

The following will be done to realize these principles.

- The ALA character set will be redefined—see table 1.
- A new character sets present field will be defined.
- Details of application such as distribution, filing indicator values, etc., will be defined.

3.1 Discussion—ALA Character Set

A character set is a list of characters with the code used to represent each one. Using this definition, the ALA character set as given in appendixes III.B and III.C of *MARC Formats for Bibliographic Data* actually consists of eight character sets.

(1) ASCII and ALA diacritics and special characters with their eight-bit code.
(2) Superscript zero to nine, plus, minus, open and close parentheses with their eight-bit code.

(3) Subscript zero to nine, plus, minus, open and close parentheses with their eight-bit code.
(4) Greek lowercase alpha, beta, and gamma with their eight-bit code.
(5–8) The same characters with their six-bit codes.

The six-bit character sets are used to distribute MARC records on seven-track tapes. There are very few subscribers. It is unlikely that a method can be devised for distribution of nonroman character sets records on such tapes. The present seven-track subscribers should be asked if they know of any way to do so. If they do not, the alternatives are to cease distribution of seven-track tapes entirely or limit them to those records containing only roman alphabet characters—those without a character sets present field. In the latter case, they should pay proportionately less for their subscription.

The present four eight-bit character sets and their escape sequences do not conform to present standards. The present standards did not exist when the character sets were being defined. To avoid creating and distributing records containing both standard and nonstandard character sets and stan-

Table 1. Proposed Revised ALA Character Set

dard and nonstandard escape sequences, the ALA character set should be redefined. This change will be much less traumatic than it sounds. No new characters will be added; only the codes used to represent subscript, superscript, and Greek characters will be changed. These characters were found in the title field of 8.59 out of 1.1 million records. If, as seems plausible, most or all MARC subscribers translate tapes into their own character set codes as a first step and for communication translate from their own codes into the ALA character set as the last step before distribution, only these two programs would need to be changed.

The proposed redefined ALA character set is shown in table 1. On it, columns two through seven are the American standard code for information interchange (ASCII) which is a recognized standard with a registered escape sequence. Columns ten through fifteen are the ALA extension of ASCII with special characters and the three Greek letters in columns ten and eleven, superscripts in column twelve, subscripts in thirteen, and diacritics in columns fourteen and fifteen. (It should be noted that six ASCII codes will not occur in MARC records: codes 5/14 circumflex, 5/15 underline, 6/0 grave, and 7/14 tilde are redundant with the codes for these diacritics in columns fourteen and fifteen; codes 7/11 left brace and 7/13 right brace never occur because these characters do not occur in bibliographic data. No change in this practice is proposed. It is the fact that these last two codes are used in some nonroman alphabet standard character sets that makes nonroman six-bit codes impossible.) The ALA extension of ASCII is not an official standard now; it does not have an escape sequence yet.

In addition to the ALA extension of ASCII, there is a draft international standard extended Latin alphabet character set for bibliographic use—ISO DIS 5426 (table 2). While both sets are identical in purpose, they differ in the characters they contain and the codes used to represent them. The ABACUS group has agreed that ISO 5426 be used for international distribution of MARC records among the bibliographic agencies they represent once it is an ap-

proved international standard, cf. LC *Information Bulletin,* November 16, 1979, p. 475. The Library will, however, continue to use the ALA extension for U.S. distribution. Some of the characters only on the ISO set could be added to the ALA extension without affecting existing records. An ANSI Z39 subcommittee has been established to consider this possibility. While some changes may be desirable to the ALA character repertoire, it is important that this issue not delay the separate matter of providing for inclusion of nonroman alphabets in MARC.

3.2 Discussion—Escape Sequence

For purposes of this discussion, escape sequences are defined as a combination of three characters. (See table 3.) The first is an escape character, hex 1/11. The second character specifies which codes are having different characters assigned to them, those in columns 2–7 or those in columns 10–15. The third character defines what characters are being assigned to these codes, e.g., Cyrillic, Greek, etc. This is a greatly simplified explanation of the escape sequence standards, ISO 2022 and ANSI X3.41. (Both are in the process of revision.) These standards provide for two types of escape sequences: public ones which reference registered character sets, and private ones for unregistered character sets. While the meaning of the latter is governed by an agreement between the sender and the receiver, they are in conformity with the standard. Until the ALA extension of ASCII has a registered escape sequence, such a "private" escape sequence could be defined for it in the character set appendix and used.

The second character of an escape sequence which changes the meaning of the codes in columns 2–7 contains either an open parenthesis, hex 2/8, or a less than sign, hex 2/12. The second character of an escape sequence which changes the meaning of the codes in columns 10–15 contains either a close parenthesis, hex 2/9, or an equal sign, hex 2/13.

The third character of escape sequences for certain registered character sets has been defined as follows:

Table 2. Extended Latin Alphabet Character Set

Bits b7 b6 b5	b4	b3	b2	b1	Row \ Column	0 (0 0 0)	1 (0 0 1)	2 (0 1 0)	3 (0 1 1)	4 (1 0 0)	5 (1 0 1)	6 (1 1 0)	7 (1 1 1)
	0	0	0	0	0				´	?	˒		
	0	0	0	1	1			¡	˙	ˋ	¿	Æ	æ
	0	0	1	0	2			··	˛	´	ʋ	Đ	đ
	0	0	1	1	3			£		^	ᴗ		ð
	0	1	0	0	4			$		~	˚		
	0	1	0	1	5			¥		¯	˘		l
	0	1	1	0	6			＋	‡	˘	.	IJ	ij
	0	1	1	1	7			§	·	•	··		
	1	0	0	0	8			ˈ	ˮ	··	_	Ł	ł
	1	0	0	1	9			‘	’	··	=	Ø	ø
	1	0	1	0	10			“	”	˚	ı	Œ	œ
	1	0	1	1	11			«	»	ˊ	ˆ		ß
	1	1	0	0	12			♭	#	˙		Þ	þ
	1	1	0	1	13			©	´	˝	⌐		
	1	1	1	0	14			Ⓟ	˝	˴	⌐		
	1	1	1	1	15			Ⓡ	¿	ˇ	�‿		

The sixteen codes in column three can be used to designate sixteen different "private" character sets. In MARC records, ASCII and Russian would be assigned to columns 2–7, while Greek and the extended Cyrillic (and the ALA extension of ASCII) would be assigned to columns 10–15.

Escape sequences would be given where needed in data fields. If necessary, it is permissible to embed escape sequences within a word. For example, a Latin diacritic might be needed with an extended Cyrillic letter to represent a letter in one of the non-Slavic languages of Central Asia which uses the Cyrillic alphabet.

In addition to escape sequences for non-roman alphabets described above in which one code stands for one letter, the escape standards also define escape sequence procedures for changing to multiple byte character sets. Because the ideographic writing

Table 3. Escape Sequence Character Set

4321 BITS		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0000	0			SP	ø	ю	п	Ю	П					г	±	Г	Ђ
0001	1			!	1	а	я	А	Я					ђ	е	Ђ	θ
0010	2			"	2	б	р	Б	Р					ѓ	v	Ѓ	V
0011	3			#	3	ц	с	Ц	С					є	ж	Є	Ж
0100	4			¤	4	д	т	Д	Т					ё		Ё	
0101	5			%	5	е	у	Е	У					s		S	
0110	6			&	6	ф	ж	Ф	Ж					i		I	
0111	7			'	7	г	в	Г	В					ї		Ї	
1000	8			(	8	х	ь	Х	Ь					j		J	
1001	9			)	9	и	ы	И	Ы					љ		Љ	
1010	10			•	:	й	з	Й	З					њ		Њ	
1011	11			+	;	к	ш	К	Ш					ћ	[	Ћ	
1100	12			,	<	л	э	Л	Э					ќ		Ќ	
1101	13			—	=	м	щ	М	Щ					ў	]	Ў	
1110	14			.	>	н	ч	Н	Ч					џ		Џ	
1111	15			/	?	о		О						ъ		Ъ	

GOST 13052-67 Russian   ISO DIS 5427 Extended Cyrillic

systems of East Asia use thousands of different characters, it will be necessary to use two or three bytes/codes to identify a single specific character uniquely. The Japanese Industrial Standard character set, JIS 6226, uses two bytes per character, and it has been submitted to ISO to obtain a registered escape sequence. The first volume of the Chinese Character Code for Information Interchange, CCCII, has been issued; the second is expected in December. It uses three bytes per character. In all probability the LC/RLIN East Asian cooperative project will adopt either these character sets and their escape sequences or machine reversible adaptations of them. The need to expand East Asian character sets constantly to provide for infrequently used characters poses problems whose solutions cannot be predicted at this time.

3.3 Discussion—Character Sets Present Field

As specified in the sixth principle, there is need for a special field which specifies what character sets are present whenever a set other than ASCII and the ALA extension of ASCII are present in a record. The proposed field will use tag 066 and be defined as follows:

066 Character Sets Present

This field specifies what character sets are present in the other than ASCII and the ALA extension of ASCII. The field is not repeatable. Both indicators are unused and will contain blanks.

$a This subfield will contain all but the first character of the escape sequence to the default character set in columns 2–7 whenever the default character set is not ASCII. This is not likely to occur in records created in the United States. Since there can only be one default character set, the subfield is not repeatable.

$b This subfield will contain all but the first character of the escape sequence to the default character set in columns 10–15 whenever the default character set is not the ALA extension of ASCII. This is not likely to occur in records created in the United States. Since there can be only one default extension character set, this subfield is not repeatable.

$c This subfield will contain all but the first character (or all but the first if a longer escape sequence is used) of every escape sequence found in the record. If the same escape sequence occurs more than once, it will be given only once in this subfield. The subfield is repeatable. This subfield does not identify the default character sets.

Example: bb$c)W A record containing the ISO extended Cyrillic character set.

bb$c)W$c)X A record containing both the ISO Greek and extended Cyrillic character sets.

3.4 Discussion—Other Details

When a field has an indicator to specify the number of leading characters to be ignored in filing and the text of the field begins with an escape sequence, the length of the escape sequence will not be included in the character count.

When fields contain escape sequences to languages written from right to left, the field will still be given in its logical order. For example, the first letter of a Hebrew title would be the eighth character in a field (following the indicators, a delimiter, a subfield code, and a three-character escape sequence). The first letter would *not* appear just before the end of field character and proceed backwards to the beginning of the field.

A convention exists in descriptive cataloging fields that subfield content designation generally serves as a substitute for a space. An escape sequence can occur within a word, after a subfield code, or between two words not at a subfield boundary. For simplicity, the convention that an escape sequence does not replace a space should be adopted. One other convention is also advocated: when a space, subfield code, or punctuation mark (except open quote, parenthesis or bracket) is adjacent to an escape sequence, the escape sequence will come last.

Wayne Davison of RLIN raised the following issue. After the Library of Congress has prepared and distributed an entirely romanized cataloging record for a Russian book, a library with access to automated Cyrillic input and display capability will create a record for the same book with the title in the vernacular. (Since AACR2 says to give the title in the original script "wherever practicable," the library could be said to be obligated to do so.) In such an event the local record could have all the authoritative Library of Congress access points. To keep this record current when the Library of Congress record is revised and redistributed, it would be necessary to carry the LC control number in the local record. Most automated systems are hypersensitive to the presence of two records with the same control number. The two records can be easily distinguished: in the Library of Congress record, the modified record byte in field 008 will be set to "o" and it will not have any 066, character sets present field.

A Comparison of OCLC, RLG/RLIN, and WLN

University of Oregon Library

The following comparison of three major bibliographic utilities was prepared by the University of Oregon Library's Cataloging Objectives Committee, Subcommittee on Bibliographic Utilities. Members of the subcommittee were Elaine Kemp, acting assistant university librarian for technical services; Rod Slade, coordinator of the library's computer search service; and Thomas Stave, head documents librarian.

The subcommittee attempted to produce a comparison that was concise and jargon-free for use with the university community in evaluating the bibliographic utilities under consideration. The University Faculty Library Committee was enlisted to review this document in draft form and held three meetings with the subcommittee for that purpose. The document was also shared with library faculty and staff in order to elicit suggestions for revision.

A copy of the draft was sent to each utility with a request for suggestions for correction and/or clarification of the report. Each of the utilities responded promptly, and their recommendations were reviewed by the subcommittee and have been incorporated into the report as it appears here.

In reading this report two considerations should be kept in mind: (1) the information is current as of December 1980, and (2) the efforts at brevity and jargon-free comparison may have resulted in oversimplification in some areas.

This report is one aspect of the six-months-long decision-making process that led the University of Oregon Library to select OCLC, Inc. (now the Online Computer Library Center).

INTRODUCTION

An online bibliographic utility provides computer services to member libraries who, in turn, contribute computer-readable records to a common database. The database is a collection of catalog records input by the members and other sources such as the Library of Congress, the Government Printing Office, and the National Library of Medicine. Use of the database is *online*, meaning that each member library accesses the computer directly and carries out its work in an interactive, conversational manner through a computer terminal located in the library. Communications with the central computer are carried over a leased long-distance telephone line. The bibliographic utility produces two primary products—catalog cards and magnetic tapes of a library's catalog records—and offers many other services for processing and bibliographic control in libraries.

In addition to providing the products and services of a bibliographic utility through the Research Libraries Information Network (RLIN), the Research Libraries Group (RLG) has three other goals: (1) to provide a structure through which common research library problems can be addressed, (2) to provide scholars and others with increasingly sophisticated access to bibliographic and other forms of information, and (3) to promote, develop, and operate cooperative programs in collection development, preservation of library materials,

and shared access to research materials.

The purpose of this report is to provide an overview of considerations in selecting an online bibliographic utility and a comparison of the three utilities being reviewed by the University of Oregon Library. Each consideration is accompanied by a brief definition or explanation, and a summary of each utility's capability in providing the necessary services or products. An attempt has been made to distinguish between currently available services and those that are planned for the future, but technological and organizational changes in the utilities have complicated this task and, in some cases, made it difficult for the subcommittee members to distinguish between operational and projected capabilities.

BASIC CHARACTERISTICS
HISTORY
OCLC

OCLC, Inc., was founded in 1967 by the Ohio College Association as the Ohio College Library Center, to be the first online shared cataloging network. It has since expanded beyond the confines of the State of Ohio and is currently used by nearly 2,400 member libraries in the United States and abroad. In 1977 it adopted its present name.

RLG/RLIN

The Research Libraries Group, Inc., was established in 1974 by four major research libraries. In 1978 it acquired from Stanford University the BALLOTS bibliographic data system, which became the foundation for RLIN (Research Libraries Information Network), RLG's wholly-owned bibliographic utility. Besides being the basis for RLG's cooperative processing activities, RLIN supports its other three programs: Shared Resources, Cooperative Collection Development, and Preservation. RLG presently has 23 owner-members.

WLN

In 1975 the Washington Library Network began testing its online system using as its base a computerized bibliographic database that several Washington libraries had been building since 1972. WLN is a project of the Washington State Library and pres-

ently has over 60 members, primarily in the Northwest.

Membership Configuration
OCLC

OCLC had 2,392 member libraries, in early 1981, including about 1,300 college and university libraries, 330 public libraries, 250 federal libraries, 145 special libraries, 77 law libraries, 71 members of the Association of Research Libraries, 168 medical libraries, 37 state libraries, and at least 48 art and architecture libraries.

RLG/RLIN

In December 1980, there were 23 owner-members (21 university libraries, The New York Public Library, and the American Antiquarian Society), two associate members, two affiliate members, and several museum and three law library special members. Libraries which formerly contracted for BALLOTS cataloging services from Stanford University are still being served by RLIN. These include 52 libraries using RLIN for online cataloging and 136 libraries using RLIN on a search-only basis.

WLN

WLN had 65 members, in early 1981, including 34 college and university libraries, 21 public libraries, two special libraries, three state libraries, five law libraries, and the Pacific Northwest Bibliographic Center.

Governance

Methods of governance are of concern to libraries considering membership inasmuch as they determine to a great extent the responsiveness of the utilities to the needs of their members and the ability of members to participate in setting the direction and priorities for the utility.

OCLC

A 15-member Board of Trustees holds the powers and performs the duties necessary for governance (including filling management vacancies and approving policy and budgets). A Users' Council, elected by the members, participates in the election of trustees and represents the interests of the membership in an advisory capacity. It also

must ratify amendments to the OCLC *Code of Regulations* and *Articles of Incorporation*. Of the 69 delegates to the Council, 44 are from academic libraries. Various advisory groups exist representing the interests of special groups within the membership, including a Research Libraries Advisory Group. Twenty regional networks contract with OCLC to provide services to their members. OCLC libraries in Oregon participate through the OCLC Western Service Center, Claremont, CA, and are served by OCLC's Portland office.

RLG/RLIN

RLG/RLIN operates through a Board of Governors consisting of one representative from each full member institution with the President as chief operating officer. Standing committees for Collection Management, Public Services, Preservation, and Library Technical Systems & Bibliographic Control; and Program committees for East Asia, Art, Law, Theology, and Music are composed of appointees from member institutions and report to the President.

WLN

An 11-member Computer Services Council is elected directly by the online participant libraries. Legal responsibility for WLN resides with the Washington State Library Commission.

Financial Stability

An indicator of a utility's financial stability is its proven ability to generate sufficient revenues to cover expenses with the least recourse to outside funding sources.

Financial stability in a utility is a concern to a library considering membership not only from the standpoint of a utility's mere survival, but because of its implications for future system developments, possible dramatic fee increases should outside funding evaporate, and maintenance of high quality services and products.

OCLC

OCLC, Inc., is a not-for-profit corporation, with tax-exempt status having been granted under section 501(c)(3) of the Internal Revenue Code. It is self-supporting, receiving no government or private subsidies,

and issuing no stock. Its revenues alone support existing operations, expansion, and research and development activities. Revenues result from fees charged member libraries for products and services. OCLC's estimated assets for fiscal year 1980 were over $55 million and its revenues approximately $24 million. Its revenue base is its 2,400 member institutions.

RLG/RLIN

The Research Libraries Group, Inc., is a tax-exempt corporation owned by its 23 owner-member institutions. Revenues result from fees charged members for use of the RLIN database. RLG currently must supplement this income with foundation grants and loans from Stanford University, because of relatively high development costs and relatively low revenues. As of this year, nearly $5.25 million has been received in grants and a $2.2 million loan was obtained, to be repaid by August 1986.

RLG has projected that in 1982–83 ongoing operating costs will be met by fee-generated income. RLG's Board of Governors recently approved a new income/expense structure to take effect September 1, 1981: "operating expenses matched by rates for services; system development matched by grants and loans; program and administration matched by a program partnership fee." This new program partnership fee will be a flat annual rate for full members in the range of $20,000 to $25,000.

A decline in the number of units cataloged by member libraries (due in part to decreased acquisitions budgets), which is the basis for fees charged, forced the Board to institute this new fee. RLG is encouraging member libraries to seek these additional funds from institutional sources outside the libraries' own budgets.

The new financial structure appears to reflect a recognition of the need for outside resources to provide for research and development for at least the immediate future, and at the same time an effort to reconcile income and expense in the areas of operating expenses and program administration. Its revenue base is its membership of 23 institutions. In the past RLG has estimated that financial stability would be reached

when membership reached 35, but it is unclear how the new rate structure will affect that projection.

WLN

The Washington Library Network receives revenues in the form of fees for services and products. As a division of the Washington State Library, it also receives some funding from the State of Washington. WLN has been the recipient of some outside grants, but does not appear to rely heavily upon grant monies to meet ongoing expenses or system development costs. WLN would like to lessen its dependency upon the State of Washington, and has taken the first step by broadening the base of its advisory committee to include out-of-state members. Its revenue base is its membership of approximately 60 libraries. The committee preparing this report does not have information as to the proportion of revenues generated by fees. However, a recent (July 1, 1980) 10% increase in service rates was put into effect for these stated purposes, among others: "to recover the cost of operation of the computer service" and to "allow a modest margin to insure stability."

Track Record in Meeting Past System Development Deadlines

Past success or failure in meeting announced deadlines for system developments may be indicative of future performance in this regard. All three utilities are heavily engaged in research and development and, while we are primarily interested in the features that are presently available, it is also important to try to gauge what each system will look like several years from now. The amount of information available to the committee varied according to the utility, so these columns are not directly comparable, but merely suggestive.

OCLC

OCLC tries not to attach dates to its projections because of early failures to meet announced deadlines. However, its interlibrary loan system was implemented one year early and its searching improvements are claimed to be ahead of schedule. The planned acquisitions subsystem had been

scheduled for completion in Summer 1980, and is currently being tested by a small number of member libraries.

The conversion of OCLC's database to accommodate the new cataloging rules and include new forms of names was completed on schedule in December 1980. The serials union listing capability was also completed on time. (See p. [224])

RLG/RLIN

A study dated August 1978 performed for the University of California listed planned BALLOTS system developments with projected completion dates. This list follows, with actual completion dates or revised projections added:*

Network File System (now called "reconfigured database" by RLIN)
  Projected January 1979
  Revised projection April 1981
Serials cataloging
  Projected January 1979
  Actual completion late 1979
Authority control system, phase 1
  Projected January 1979
  Revised projection Spring 1981
Authority Linking and Control, phase 2
  Projected Fall 1979
  Revised projection Spring 1981
Generalized acquisitions
  Projected Fall 1979
  Revised projection (in two phases) June 1981, October 1981
Serials Control
  Projected 1980
  Revised projection post-1982
Library Management Information System
  Projected 1979
  No projected date, no resources allocated
Book/COM Catalog Interface
  Projected 1980
  Revised projection 1981

WLN

WLN's present online system was one year late, and its acquisitions module was also late. The processing of retrospective conversion tapes which had been three months behind was current by early 1981,

*Since 1978 the RLG Board of Governors has determined the order of priorities for research and development.

with the exception of two special projects.

Large-scale system adjustments to accommodate new cataloging rules were completed on schedule, as was implementation of roll-microfilm catalogs.

Database Size and Components

The size and makeup of the utility's database is of concern to libraries considering membership because those factors have the greatest bearing on the library's likelihood of obtaining a large portion of its cataloging information from the system.

OCLC

Size.
Over 7.1 million bibliographic records (February 1981)
Books: 4.9 million (October 1979)
Serials: 341,000 (October 1979)
Other: 340,000 (October 1979)
Name authority records: 500,000 (est. by 1981)

Formats Available.
Books
Serials
Films (AV)
Maps
Manuscripts
Music recordings
Music scores

Sources of Data.
Member-contributed records
Library of Congress–produced Machine-readable cataloging records (MARC) (1968 to date)
Government Printing Office–produced records (cataloged directly into OCLC by GPO)
CONSER records (*Conversion of Serials*—a project of 15 major libraries to produce machine-readable serials cataloging records). Data are entered directly into OCLC, then authenticated by the Library of Congress and the National Library of Canada.
National Library of Medicine–produced records

Additional sources include the following databases:
Canadian MARC serials
Minnesota Union List of Serials
Pittsburgh Regional Library Center Serials

RLG/RLIN

Size.

Over 3 million bibliographic records (June 1980)

Books: 2.5 million (June 1980)

Serials: 460,000 (June 1980)

Authority records: 1.6 million (early 1981)

Formats Available.

Books

Serials

Films (AV)

Maps

Music recordings

Music scores

Sources of Data.

Member-contributed records

MARC (excluding 1968–1972)

GPO records (to be added Spring 1981)

CONSER records

Cataloging records from Columbia and Yale Universities and University of Minnesota Biomedical Libraries, previously put into machine-readable form, have been added to RLIN. Records from the New York Public Library, Northwestern and Pennsylvania State Universities will be added in the near future.

Additional sources include the Avery Index to Architectural Periodicals.

WLN

Size.

2 million bibliographic records (January 1981)

Authority records: 2.3 million (January 1981)

Holdings records: 2.3 million (December 1980)

Formats Available.

Books

Serials

Films (AV)

Music recordings*

Music scores*

Sources of Data.

Member-contributed records

MARC (1968 to date)

GPO records

CONSER records (except those not yet authenticated by the Library of Congress)

Machine-readable records from the University of Illinois will be added to WLN's

*Awaiting implementation by the Library of Congress.

database on a weekly basis by mid-1981. Records from certain libraries in the Southeastern Library Network (SOLINET) will be added in the future, as part of an arrangement whereby WLN made its computer software package available for use by Illinois and SOLINET.

RESOURCE SHARING
Interlibrary Loan (ILL)

ILL is the process by which library materials are lent and borrowed by libraries in the U.S. and foreign countries. A bibliographic utility provides two tools to aid in this process: an online union catalog used to determine which library owns the needed material, and a message switching system used to communicate among libraries and to carry out the transaction. ILL at the University of Oregon Library is currently accomplished using a large number of printed union catalogs and is communicated by mail or Western Union teletype. A bibliographic utility will not completely replace ILL transactions carried out in this manner. The number of requests for materials from the library collection will probably increase due to the "visibility" gained in the online union catalog.

OCLC

The OCLC database provides the largest online union catalog through a holdings record listed with each catalog entry. The ILL message system transfers records from the database to the lending library in a request form, automatically sends the request to up to five libraries, generates records on the status of each request, and provides statistics on ILL transactions. OCLC ILL transactions are generally faster than traditional methods of interlibrary loan because of the ability to move data directly from the online union catalog to the request form without re-typing and the ability to have requests automatically forwarded if a library is unable to fill the request immediately. OCLC's ILL subsystem has been in operation for a year and participating libraries have reported general satisfaction with its performance.

RLG/RLIN

The RLIN database provides an online

union catalog through a holdings record listed with each catalog entry. Materials not located in the RLIN database may be referred to the Bibliographic Center at Yale University for further manual searching through printed union catalogs. The RLG Message System may be used to create and send ILL requests to other RLG libraries, though this system is not specifically designed as a comprehensive ILL support system. The Shared Resources Program committee has recently formed a task force charged with the responsibility to create a functional specification for an automated interlibrary loan system, and to determine the priority for its implementation. RLG resource sharing policy requires members to give priority to ILL requests from other RLG members, to suspend fees to members, to provide on-site access to users from members' libraries' institutions, and to provide free photocopies of non-circulating materials.

WLN

The WLN database provides an online union catalog through a holdings record listed with each catalog entry. This online union catalog includes the local library call number and, for serials, the specific holdings of the library. The *WLN Resource Directory* is a microfiche listing of the bibliographic and holdings information in the database. WLN offers no message switching system for ILL, though this is their highest priority for future development. In cooperation with Pacific Northwest Bibliographic Center, WLN is planning experiments with a message switching system for interim use until the comprehensive ILL system is developed.

Cooperative Acquisitions

Cooperation in purchasing library materials is done in order to minimize the duplication of expensive purchases and to ensure that important works are easily available to users of the library, whether they are actually owned or not.

OCLC

Member libraries may search the database to determine the holdings of particular items by other member libraries, in order to

avert undesirable duplicative purchases.

RLG/RLIN

Members actively coordinate purchases of certain categories of materials in designated fields in order to avoid extensive duplication and to ensure that at least one copy of every item of research value be acquired by a member institution.

In support of this effort is an automated "cooperative purchase file," containing limited bibliographic information and acquisition decisions of RLG members for all new serials on order and for all expensive items ($500 or more).

Member institutions agree to develop conspectuses reflecting their level of holdings and development in certain fields (subjects, language, and formats). These conspectuses are time-consuming to develop. A survey of holdings in Chinese, Japanese, and Korean languages has been finished by 12 members. Older members have completed Language and Literature, Fine Arts, Philosophy, and Religion. History is expected by March, 1981, to be followed by the hard sciences. Based upon these conspectuses, RLG members will build a system-wide collection development policy. New members are expected to begin work on their conspectuses as soon as possible, but not necessarily immediately after joining RLG.

WLN

Members may search the database to determine the holdings of particular items by other member libraries, in order to avert undesirable duplicative purchases. Libraries may also search the in-process file to determine if items are on order by one of the 23 libraries using WLN's Acquisitions Subsystem.

Support for Collection Development Activities

A bibliographic utility is potentially useful for collection development in that it provides a large file of bibliographic records that may be searched to assist in a) determining the existence of published materials in specified categories (on a particular subject, by a particular author, in a particular series, for example), and b) obtaining cor-

rect bibliographic information about specific items to assist in ordering them. Important features in a utility in this regard are database size and variety of access points (subject, author, series titles, etc.).

OCLC

Useful access points by which the database may be searched include:

- Personal author
- Corporate author
- Title
- Series title
- Variant names (e.g. Clemens or Twain)
- Conference names

The database must be searched using a "search key" (a code based upon a sequence of initial letters in the words to be searched), not real words.

RLG/RLIN

Useful access points by which the database may be searched include:

- Personal author
- Corporate author
- Conference names
- Title
- Series title
- Subject heading or call number range (excluding items cataloged by the Library of Congress)
- Publisher, using a truncated ISBN (International Standard Book Number) [restricted to items cataloged by the Library of Congress]

A search of RLIN is likely to produce multiple records for particular items because an item held by more than one member will be displayed for as many libraries as have cataloged it through the system. It is projected that by April, 1981, RLIN's "reconfigured database" will have solved that problem by attaching holdings information to one unified record. It will also have merged the two bibliographic subfiles (Library of Congress and member cataloging) so that access by subject heading, call number range, and ISBN will be available for the entire database.

WLN

Useful access points by which the database may be searched include:

- Personal author
- Corporate author or corporate author keyword (keyword searching permits the user to search for items using either the full heading: American Society for Information Science; or words from the heading: "society" and "information." This capability is useful when the complete phrase is not known.)
- Title
- Corporate or conference author/title series (keyword)
- Series title or truncated series title
- Subject heading and/or subdivision or truncated subject heading
- Corporate and conference name subject headings (keyword)

Preservation of Library Materials

All bibliographic utilities, because of their function as a union catalog of their members' machine-readable cataloging information, have some usefulness for libraries making decisions about preservation priorities. A library may, for example, choose to give preservation treatment to item A rather than item B because item B is owned by several other libraries in the vicinity, whereas item A appears to be unique. It must be remembered, however, that many older items will not appear at all, because they were cataloged long before the utilities came into existence.

OCLC

Members may search holdings information in the database to determine the relative rarity of an item that is a candidate for preservation treatment.

RLG/RLIN

Members may search holdings information in the database to determine the relative rarity of an item that is a candidate for preservation treatment.

A computerized list of members' micropreservation activities is provided.

Experimental programs are conducted to test new preservation technologies and applications of existing processes.

Preservation microfilming is being done for members by staff at Yale and Princeton.

Funds are provided to members for preservation activities.

These activities are part of RLG's Preservation Program, one of its four major programs.

WLN

Members may search holdings information in the database to determine the relative rarity of an item that is a candidate for preservation treatment.

TECHNICAL PROCESSING
Acquisitions

The steps by which the library purchases books and other materials include:

1. Pre-order searching to determine that a requested item is not already owned by the library or on order.
2. Selecting a dealer likely to be able to supply desired item.
3. Placing the order.
4. Receiving the item.
5. Clearing the order records.
6. Processing the invoice for payment.
7. Maintaining precise accounting of all book funds.
8. Inquiring about the status of items which are not received when expected.
9. Cancelling orders and adjusting accounting records when items are not available.

At the UO most acquisitions forms and files are created and maintained manually. In an automated acquisitions system the placing of the initial order generates an acquisition record for each item, which is updated as the item moves through the cycle outlined above. This eliminates the need for maintaining separate files according to the status of an order.

OCLC

Operational. OCLC has an online name-address directory which presently can be searched while using other OCLC subsystems. This file contains information about publishing, educational, library, and professional organizations and associations. This information will be automatically transferrable to forms being produced online.

Planned. OCLC's Acquisitions Subsystem, which is presently being tested by se-

lected member libraries, is projected to be generally available in Spring 1981.

When operational the Acquisitions Subsystem will permit users to:

Place orders for all types of bibliographic materials (forms generated will be sent directly to supplier with copy to library)
Renew subscriptions
Request publications or price quotations
Create deposit account orders
Send prepaid orders
Cancel orders
Create and adjust fund records
Receive periodic fund reports

RLG/RLIN

Operational. RLIN does not have an operational Acquisitions Subsystem. Stanford University is continuing to use a system developed as part of BALLOTS.

Planned. The RLG Board of Governors has approved functional specifications for an Acquisitions Subsystem to be introduced in two phases. By June 1981, RLIN plans to have a centralized in-process file which will contain records of all new orders, gifts, subscriptions, etc. of members, and will be able to support non-accounting aspects of the acquisitions process. The capability to store and maintain an online book fund accounting system will be achieved in October 1981.

RLIN expects to be able to support all files, processing, and products necessary to establish, coordinate, and monitor materials acquisitions from the point of selection decision, request, order, or receipt through completion of technical processing activity.

WLN

Operational. WLN's Acquisitions Subsystem, which has been operational since May 1978, is comprised of four files:

1. In-process file which supports the majority of acquisitions activities.
2. Standing orders file which has records for subscriptions and other items which are renewed or reordered on a continuing or periodic basis.
3. Name and address file which contains names and addresses of book dealers and other vendors, main libraries, branch libraries, etc.
4. Account status file which provides ca-

pability to maintain up-to-date accounting. Information keyed into the terminal during the day is entered against the accounts nightly and is reflected in the account totals available online the following day.

Records of completed transactions are transferred to a magnetic tape history file and can be used for generating statistical and other reports.

With each step of the order cycle, appropriate forms and reports are generated. Special system reports reflecting the status of the four files may be generated on request. Instructions entered at the time of the initial order provide for automatic generation of notification forms for individuals requesting the specific item being ordered or inquiry notices for materials not received after a specified period.

Planned. Further refinements of the procedures and capabilities of the system.

Cataloging

The creation of a cataloging record involves:

1. Describing an item
2. Assigning headings for names of persons or organizations and titles by which the user might be expected to seek the item in the catalog
3. Assigning a unique call number which will place the item with others of a similar nature, and
4. Assigning subject headings which reflect the content of the item.

Because most libraries collect many of the same materials, the concept of sharing the responsibility for cataloging was developed which makes materials available more quickly at reduced cost. With the establishment of national and international cataloging rules and standards, and the growth of large online computerized databases, it is becoming increasingly feasible to have each item cataloged only once with that cataloging information available for all libraries to use. The Library of Congress catalogs approximately 250,000 titles per year into machine-readable form. This cataloging is available through each of the bibliographic utilities and may be used for the creation of local catalogs. When the Library of Con-

gress has not yet cataloged a specific item, a utility member library may prepare the cataloging according to specified standards and enter its cataloging into the database for use by other member libraries and for its own catalog.

Another aspect of the cataloging activity is the creation of a local database which can be used as the basis of not only the local library catalog, but also of a local circulation, acquisitions, and serials system, as well as for regional union catalogs. In order to provide total access to a library's collection in this machine-readable database, information concerning every item in the library must be entered into the system. This process is called retrospective conversion.

During the retrospective conversion process the library can choose to eliminate existing inconsistencies in the treatment of library materials including reclassifying books so that most materials are retained in one main classification system.

The University of Oregon Library has as a long-term goal completing total retrospective conversion of its collection so that all materials can be searched and located in an online catalog.

OCLC

Operational. OCLC's online Cataloging Subsystem has been operational since 1971. Based on the experience of similar libraries, the University of Oregon Library might expect to find entries in OCLC's database for over 90 percent of the items searched.[*] These cataloging records can be modified online or accepted as is. The local library's symbol is added to indicate that it has used the cataloging record and then presorted, alphabetized catalog cards are ordered. The cards are printed overnight and shipped on a daily basis. Many OCLC libraries print their call number labels by means of a printer attached to their terminal.

Once a cataloging transaction has been completed, it is not possible to retrieve your local modifications online in the OCLC system. The record of your transaction is stored and sent to your library on magnetic tape on a periodic basis. These magnetic archive tapes can be used by a vendor or

*See footnote on page 225.

local computing center to generate a local microform or online catalog, run a circulation system, etc.

It is presently possible to catalog most types of materials in the OCLC system including books, serials, microforms, motion pictures, music, sound recordings, maps, and manuscripts.

Increased emphasis has been placed on quality control and adherence to specified standards in the creation of cataloging records, but there is no official editing of cataloging records by OCLC staff.

In 1979–80 nearly 45 percent of the activity on OCLC's Cataloging Subsystem was related to retrospective conversion. OCLC's large database, extended hours of service, and special pricing schedules for retrospective conversion and reclassification make it attractive for these activities. OCLC charges 60 cents per retrospective conversion record during hours of peak system activity (prime time) and five cents per retrospective conversion record during less busy hours (non-prime time).

Planned. OCLC continues to explore means of improving quality control. After moving their central facility to new quarters in early 1981, OCLC will reconsider the possibility of storing and displaying the number and location of local copies of a title.

RLG/RLIN

Operational. At this time the University of Oregon might expect to find cataloging available for 70 to 90 percent of its ongoing work in RLIN.† A search of RLIN's database retrieves multiple records because each library's records are stored separately. The

library selects the desired record, modifies or accepts it, enters the library's symbol, and orders cards which are printed nightly and sent in presorted, alphabetized batches. No call number labels are produced, and it is not presently possible to print labels from the terminal.

Local library modifications are accessible online. Magnetic tapes or cataloging transactions may be purchased and used to create local online or microform catalogs.

Most materials may be cataloged with RLIN including books, serials, microforms, motion pictures, music, sound recordings, and maps.

Member libraries agree to catalog in conformity with RLIN standards, but there is no formal editing of records by RLIN staff on an ongoing basis. Sample quality checking is the responsibility of a newly-created position of Quality Assurance Specialist.

With only 23 owner-members, RLG must carefully consider the impact on the system of allowing individual members to undertake retrospective conversion projects. Each project must be approved by the Board of Governors, and members are encouraged to seek outside financial support rather than asking RLIN for reduced rates. RLIN has just received a 1.25 million dollar grant including $600,000 to support retrospective conversion projects. RLIN does not charge for retrospective records which are completely recataloged and upgraded with the book in hand. The prices for other levels of retrospective conversion cataloging range from fifty-five cents to $1.85 per record.

Planned. In April 1981, RLIN plans to reformat its database so that there will be

†A wide range of success rates for searching each system are cited in the literature, each dependent on the sample procedures used. The University of Oregon Library had 100 items searched against each database. This sample excluded books with printed Library of Congress card numbers, and included books, serials, microforms, music scores, recordings, documents, and non-book materials. Of this sample OCLC found 96, RLIN found 65, and WLN found 38. The range of figures cited in this report allows for variation between studies cited in the literature, word-of-mouth reports from librarians using these systems, and the University of Oregon Library's own sample. An analysis of this sample is being prepared.

Recent comparisons of searching success are found in the following: *Linking the Bibliographic Utilities: Benefits and Costs*, submitted to the Council on Library Resources . . . by Donald A. Smalley [and others]. Columbus, Ohio, Battelle, 1980; Matthews, Joseph R., "The Four Online Bibliographic Utilities: A Comparison," *Library Technology Reports* 15:6 (November–December 1979), p. 665–838; Tracy, Juan I. and Remmerde, Barbara, "Availability of Machine-Readable Cataloging: Hit Rates for BALLOTS, BNA, OCLC, and WLN for the Eastern Washington University Library," *Library Research* 1:3 (Fall 1979), p. 227–81.

only one copy of each cataloging record. Member libraries' symbols and local cataloging information will be displayed with the appropriate records.

WLN

Operational. Based on the experience of others, the University of Oregon Library might currently expect to find cataloging records available for 50 to 70 percent of its ongoing work in the WLN database.* Libraries search WLN's database, accept or modify the cataloging records, and order cards and labels which are printed nightly and shipped weekly. (Card sets are not presorted for filing.)

Local cataloging information is accessible online through the library's WLN terminal. Magnetic tapes of a library's cataloging transactions may be purchased to run a local online or microform catalog. WLN also provides microform catalogs on either microfilm or microfiche.

Books, serials, and audio-visual materials, but *not* music, sound recordings, and maps may be cataloged on WLN's system.

Libraries cataloging in WLN must conform to well-defined WLN standards. New cataloging records go through an edit cycle and are reviewed by central WLN staff before being added to the WLN database. Presently this review takes about two weeks. During this period, the cataloging record may not be retrieved online.

The WLN batch Retrospective Conversion Subsystem has been operational since August 1980. Using this system a library enters brief cataloging records which are collected by the system and searched later as a unit through the WLN database. Records for which a match is found are billed at six cents. Records not matched are billed at one cent and may be searched again at a later date. Over 30 WLN libraries are using this capability, which can be made available to non-members under special circumstances.

Planned. WLN is considering dispersing among selected member libraries responsibility for editing member-created cataloging records. WLN will make music cataloging available within the near future.

*See footnote on page 225.

Serials Check-in

Serials are publications issued in successive parts bearing numerical or chronological designations which are intended to be continued indefinitely. They include periodicals; newspapers; annual reports and yearbooks; journals, memoirs, proceedings and transactions of societies; and numbered series. The average research library will have between 15,000 and 20,000 such titles. Precise data must be maintained to enter each issue as received, to discover missing issues, to request replacements for missing issues, to monitor accounting information, to renew or cancel subscriptions, and to maintain binding information.

Serials files contain such information as title, relationship to earlier publications, name and address of publisher, volumes the library owns, call number and location, date, volume, and number of each issue, date each issue was received, subscription dates, price, etc.

At the University of Oregon Library all of this information is maintained in manual files. Once the serials check-in operation is computerized, it is possible to generate a wide variety of serials finding lists, analyses of serials subscriptions by subject, location, department, etc., and to provide current serials information online.

OCLC

Operational. OCLC introduced its Serials Control Subsystem in 1976 and improvements to the system in 1979. Participants create online local data records with information necessary to monitor and control each issue of each serial received by the library. Institutions can check-in currently received issues online.

A recent ancillary to this system is the ability to create and maintain online a cooperative record of serials owned by any group of institutions (a union list of serials).

Planned. OCLC plans to continue upgrading the capabilities of its Serials Control Subsystem as needed.

RLG/RLIN

Operational. None.

Planned. Automated serials check-in is one of several items listed for consideration

after current development activities are released, probably in late 1982. No resources are presently committed to this project.

WLN

Operational. While WLN has no current serials check-in capabilities, it does support maintenance of serials subscriptions in the Acquisitions Subsystem, including automatic renewal and reorder reminders. WLN also produces union lists of serials.

Planned. WLN is investigating existing commercially-created check-in systems to see whether they can purchase an existing system to incorporate into WLN's services.

Management Information

Precise up-to-date information concerning library operations can be very useful in planning improvements in library services and in attaining efficient utilization of available personnel, resources, and materials. Without the computer, the laborious record-keeping necessary to obtain useful management information almost negates the benefits of having the information.

OCLC

Operational. OCLC produces cataloging, interlibrary loan, and serials check-in system use and system performance statistics on a regular basis. Libraries can make local arrangements to create additional analyses of the information stored on subscription archival tapes of their local cataloging activity. OCLC offers semimonthly, monthly, or quarterly accession lists of new materials cataloged by each library. These lists may be in call number or subject sequence. OCLC has produced some special studies for institutions based on their cataloging records.

Planned. When the Acquisitions Subsystem is operational, libraries may choose to receive a cumulative, monthly Fund Activity Report and a periodic, cumulative Fund Commitment Register. These reports will provide institutions with current financial control data.

OCLC plans to continue to develop its ability to provide management information.

RLG/RLIN

Operational. System use statistics are provided in the form of the monthly invoice, which may be used to monitor cataloging and public service activity, and may be broken down into appropriate accounts by pre-planning. Lists in call number order of materials cataloged by a library into RLIN could be produced from local printers attached to the terminal.

Planned. The generation of management information is a future development project; no special management reports are prepared presently. Among the management reports included in the specifications for the Acquisitions Subsystem, projected for implementation by October 1981, are status reports on in-process files, materials awaiting receipt, materials received, and book fund balances.

WLN

Operational. WLN produces aggregate system activity reports monthly, but does not analyze the cataloging activity or subject holdings. WLN's Acquisitions Subsystem can be used to produce acquisitions-related management reports concerning account transactions, account history, standing orders, renewals and reorders, receipts, detailed encumbrances, etc. A microform accession list by title is available. A general-purpose text-editing facility may be used by management to maintain data not derived from WLN operations and to produce formatted reports of this data.

Planned. WLN is developing the capability to store and maintain detailed collection information for each library online, including copy numbers and location symbols for each copy of a title owned by a library. No specific management information plans have been outlined at this point.

PUBLIC SERVICES
Reference Use of the Utility's Terminal

A bibliographic utility has potential for use in library reference services in three major areas:

1. Verification of bibliographic information. The utility's database may be searched for cataloging information

not in the UO Library catalog. A verification search is made to locate a complete catalog description of a specific, known item and is carried out most easily using one of the unique numbers assigned to a publication (Library of Congress Card Number, International Standard Book Number, etc.). If one of these is not known, a combination of author and title words, or a "search key"* based on author and title is used to retrieve the information. Verification places a greater reliance on the quality of bibliographic information in the utility's database than on search techniques used to locate the information.

2. Compilation of subject bibliographies. The utility's database is searched through words in the titles and subject headings in a bibliographic record in order to produce a list of materials on a given subject. This subject query can be modified using the logical relationships AND, OR, and NOT to indicate, respectively, limitations, synonyms, or exclusions in the search. The ability to obtain a *printed* list of references is convenient, if not required.

3. Compilation of author bibliographies. The database is searched to find all material created by a particular individual or corporate body. The size of the utility's database is a major consideration, as is the source of the cataloging found in an author search. Again, a *printed* list is necessary.

OCLC

The OCLC database can be searched in a variety of ways to support reference ser-

*A search key is a code based on a certain number of characters drawn from a particular element in the bibliographic reference. For instance, to find a record for William Manchester's *American Caesar*, an author/title search key using the first four letters of the author's name and the first four letters in the title would be *MANC,AMER*. Various combinations of letters are used to search author names, titles, or author/title combinations. A search key may not necessarily be unique to a given item, and may retrieve other items beside the one desired.

vices, though there is no subject search capability in the system. The following access points may be used in a search:

1. LC Card Number
2. International Standard Book Number (ISBN)
3. International Standard Serial Number (ISSN)
4. CODEN (an abbreviation developed by Chemical Abstracts Service for designating periodical titles)
5. Government Documents Number
6. OCLC Identification Number
7. Personal author (search key, not full words)
8. Corporate author (search key)
9. Performer (search key)
10. Title (search key)
11. Author/Title (search key)
12. Series title (search key)
13. Variant names (search key)
14. Conference names (search key)

Searches may be restricted by year or by type of material, such as books, manuscripts, maps, etc. The logical operators AND, OR, and NOT are not used in OCLC. The OCLC search system is primarily based on search keys and is best utilized to locate a known item. Local printing is available on any OCLC terminal so equipped. There is one standard print format offered.

RLG/RLIN

The following access points may be used in a search of the RLIN database, though not all are currently active in each subfile of the database:

1. LC Card Number
2. ISBN
3. ISSN
4. CODEN
5. Government Documents Number
6. RLIN Identification Number
7. Call number (complete or truncated)
8. Recording Label Number
9. Personal author
10. Corporate authors or conference names (keyword or phrase)
11. Title words
12. Subject headings (keyword or phrase)
13. Music Publisher

Truncation (searching of partial entries) is available to aid in searching incomplete entries and the logical operators AND, OR, and NOT may be used to broaden or restrict a search. Local printers may be attached to the RLIN terminals. A variety of print formats is offered. Plans include unified search access points for all subfiles of the database as of April, 1981.

WLN

The following access points may be used to search the WLN database:

1. LC Card Number
2. ISBN
3. ISSN
4. WLN Identification Number
5. Personal author
6. Corporate authors or conference names
7. Title words
8. Series title (complete or truncated)
9. Corporate or conference author/title series (keyword)
10. Subject headings (complete or truncated)

For a variety of reasons, the WLN search system is the most powerful of the three utilities. Truncation is available and the logical operators AND, OR, and NOT may be applied to broaden or restrict a search. Records may be printed locally in a variety of formats on any WLN terminal so equipped. WLN will also provide printing at the central computer for reference bibliographies. WLN search software may be purchased for local database management applications (see the section on Online Public Catalogs.)

Links to Other Computerized Services

There are presently over 150 reference databases available through commercial computerized reference service vendors. During the last ten to fifteen years, standard bibliographic indexing and abstracting publications such as *Chemical Abstracts, Historical Abstracts* and *Dissertation Abstracts International* have used computerized methods to organize and print references to periodical articles, reports, dissertations, conference papers, etc. The vendor creates a computer searchable version of the reference database and makes it available to libraries for a fee based on their use of the computerized search system. Membership in a bibliographic utility can provide two benefits in the use of other computerized reference services:

1. Discounts on fees through membership in large group contract administered by the utility.
2. Access to the reference vendor's computer through the utility's terminal and communication network.

OCLC

OCLC's Affiliated Online Services program provides access at discounted rates to the information services of Bibliographic Retrieval Service (BRS), Lockheed Information Systems (LIS), and the New York Times Information Bank. OCLC's communications network does not yet permit users to link to the hosts using an OCLC terminal, though this capability is anticipated in the near future.

RLG/RLIN

RLIN does not offer a formal program in this area, though the RLG 40 terminal is compatible with other information retrieval systems.

WLN

WLN does not offer a program in this area, but anticipates offering access to BRS, LIS, and New York Times Information Bank.

Circulation

None of the bibliographic utilities under consideration currently support circulation functions on their computers. However, each system can provide a machine-readable archive tape of our cataloging information to be used in developing a computerized circulation system. In order to keep track of circulation transactions, it is necessary to have complete retrospective conversion of the UO Library catalog. Another important consideration is the transferability of data between the utility's computer and the circulation computer.

OCLC

OCLC anticipates offering support for local circulation systems on their computer

for member libraries and will demonstrate their system in mid-1981. OCLC data has been successfully transferred to many local circulation systems.

RLG/RLIN

RLIN does not anticipate offering local circulation services for member libraries. RLIN data has been successfully transferred to several local circulation systems.

WLN

WLN does not anticipate offering local circulation systems on their computer for member libraries. WLN data has been successfully transferred to local circulation systems and an agreement has been reached with DATAPHASE, a computerized circulation system vendor, to discount purchase of their system by WLN member libraries.

Public Online Catalogs

Again, none of the bibliographic utilities under consideration currently support public online catalogs of an individual library's collection. A public online catalog requires further programming in order to make it easy for the public to locate materials of interest without extensive training; the bibliographic utility's searching procedures are too esoteric to be used by the general public. As in circulation, issues of data transferability and full retrospective conversion of the UO Library's catalog are paramount.

OCLC

OCLC does not currently encourage public access to their database and does not support use of local online catalogs on their computer due to the tremendous demand for computer resources exerted by 2400 member libraries. OCLC and RLG/RLIN are participating in a study of user requirements for a public online catalog. OCLC data has been successfully transferred to several local online catalogs, including Eugene Public Library's circulation and online catalog system, ULISYS.

RLG/RLIN

RLIN anticipates being able to offer public access to their database. They are participating in a study with OCLC of user requirements for such a system, but no date

has been announced for the development of this capability in RLIN. RLIN data has been successfully transferred to a local public online catalog at Northwestern University.

WLN

WLN does not believe that a local online patron accessed catalog should be provided through the WLN computer, even though they anticipate having such a capability within one year. Instead, they encourage libraries to develop local systems for public access to the online computerized catalog and to obtain data from the WLN cataloging system. The University of Illinois is adapting the WLN computer search and database management software to provide a local online catalog and computer-assisted instruction in its use for the public.

Checklist for Cassette Recorders Connected to CRTs

Prepared by Lawrence A. WOODS: Purdue University Libraries, West Lafayette, Indiana, for the Technical Standards for Library Automation Committee, Information Science and Automation Section, Library and Information Technology Association.

INTRODUCTION

A data cassette recorder connected to a printer port is an effective, low-cost method of collecting data in machine-readable form from display terminals such as the OCLC 100/105.

It is important that a *data* recorder be used rather than an audio recorder although the cassette itself can be a good-quality audio tape. It is also important to note that the data recorded on the tape are *not* the same as the data originally transmitted to the display terminal, but are simply a line-by-line image of what appears on the screen. A typical installation will have a minimum of two devices: one attached to the display terminal to collect data, and one attached to a printer or an input device to another computer for playback of the data.

There are more than 150 various data re-

cording devices on the market. This checklist is prescriptive in nature, outlining and describing those features that are necessary or desirable for a typical application. In addition to features, environmental considerations are briefly mentioned along with information for the purchase, lease, or rental of data equipment.

FEATURES

In general, features must be compatible between all devices used for recording and playback in a given application. Some features that are desirable for certain applications are unnecessary or inappropriate for others.

1. *Recording Media*

   The *Phillips cassette* is most widely used and may be interchanged between the recorders of different manufacturers that utilize it. The *cartridge* (either 3M or a vendor proprietary cartridge) is gaining popularity because of its greater storage and transfer rates, but as yet is not widely used.

2. *Code*

   Most print ports on display terminals use ASCII (American Standard Code for Information Exchange) data code. The recorder selected should use the same.

3. *Interfaces*

   The cassette recorder has an "in" plug to accept data. This must be compatible with the print port on the terminal—usually RS232C. The "out" plug on the recorder sends the recorded data to a printer or to a computer. This interface should also be RS232C.

4. *Recording Characteristics*

   a. The *number of tracks* can vary from one to four. This is one of the factors that determine the amount of data that can be recorded on a single cassette. Four tracks are recommended.

   b. *Density* also affects the amount of data that can be recorded. Usual densities are 800 or 1,600 bits per inch (bpi).

   c. *Recording mode.* There are several

modes available. *Phase Encoded* (PE) is the best mode for data applications. *Non-Return to Zero* (NRZ) is a popular mode, but has poor error recovery. IBM has a version called NRZI, which improves on NRZ but still is less reliable than Phase Encoded. Other commonly found modes are *Complementary NRZ* and *Ratio Recording.*

   d. *Recording format.* There is a variety of recording formats. To be assured compatibility with the terminal and playback device the format should be either *ANSI* (American National Standards Institute) or *ECMA* (European Computer Manufacturers Association) compatible.

5. *Transmission*

   a. *Duplex.* The recorder should have both full and half duplex available.

   b. *Data transfer rate (baud rate).* Baud rate is usually switch-selectable from 110 to 9600. The recorder must be set at the same speed as the printer port on the terminal. The OCLC 100 and 105 terminals have a printer port baud rate selection switch that may be set at 100, 150, 300, 600, 1200, and a meaningless 1800 baud. Select a recorder that has the fastest compatible setting: 1200 baud is best. Data must be played back at a rate compatible with the receiving device.

6. *Tape Transport Characteristics*

   a. *Read/write speed* is usually a function of the baud rate.

   b. *Nonrecording speeds.* This feature is important for convenience. Fast forward and rewind should be available. One hundred twenty inches per second will rewind a cassette in about thirty seconds.

   c. *Drive mechanism.* Four options are available: capstan, pinch roller, servomechanism, or reel-to-reel. Pinch roller is the most precise but reduces the life of the tape.

7. *Packaging*

This feature can affect the price of the final configuration. If any item is listed as "separate," increase the total price accordingly. Components that can be either internal or separate are: controller, interface, or power supply.

8. *Remote Operation*
Some devices use ASCII control codes to trigger controls automatically. This is a useful feature, but the device *must* have a transparent mode switch, otherwise codes embedded in the data being recorded or sent may trigger undesired operations such as rewind.

9. *Operating Characteristics*
   a. *Rewind, fast forward, initialize, send* and *receive* are all necessary operations and should be switch-controlled.
   b. *Edit, auto program search, string search, skip,* etc., are useful for word-processing operations but are of little use in simple data collection and transmission.
   c. *Read backward* is desirable for sort operations.
   d. *Character mode, line mode,* and *string mode* are useful for printing operations but of little use in data transmission.
   e. *Online/offline* should be switch-selectable.
   f. *Simultaneous read/write* is useful for editing operations.
   g. *Direct block accessing* is useful if there is a need to search for recorded data but is not used in sequential processing.
   h. *Auto reverse* is a useful feature for recording or transmitting more data than can be recorded on one side of a cassette.

ENVIRONMENTAL REQUIREMENTS

1. *Humidity Range*
Humidity range should be 20 percent to 80 percent without condensation. Lower humidity will cause excessive static electricity.

2. *Temperature*
Temperature range should be between ten degrees and forty degrees centigrade.

3. *Power Requirements*
Most recorders require a standard 115-volt alternating current at 47 to 63 Hz. and draw about 60 watts. The circuit should be free from interference such as that caused by florescent lights. A transformer may be required in the outlet to guarantee even power.

4. *Space Requirements*
The recorder usually can be stored on a desk top. It is important that the indicator lights be visible to the terminal operator to monitor its operation.

PURCHASE

1. *Maintenance and Availability*
Ask how many drives the manufacturer has installed to date. This may vary from a few hundred to one hundred thousand or more. Establish a maintenance contract with the company or a local service bureau. It may be necessary to acquire a spare recorder to use as backup.

2. *Price*
Determine ahead of time what features you are actually going to use. Bells and whistles all cost money. A simple reliable recorder can be purchased for around $700. Multiple drive units and other features can run as high as $3,600.

LITA Award, 1980:
Maurice J. Freedman

S. Michael MALINCONICO

This is the third presentation of the LITA Award for Outstanding Achievement. The first two honored individuals whose achievements can be said to have created the discipline we know as library automation. The first award went to Fred Kilgour whose vision, daring, and entrepreneurial and managerial skills changed the way libraries operate almost overnight, and may in the increasingly stringent economic times ahead have helped ensure the economic viability of libraries. The second award went to Henriette Avram, whose untiring efforts on behalf of the MARC formats and their promulgation is only just short of legendary.

This year's winner distinguished himself in a somewhat different manner. His contributions did not lead to the development of new automated systems or services. Rather, his outstanding achievement lies in the creative and pioneering use he made of technology in support of a clear vision of effective library service. His contribution comes from the depth of sensitivity and understanding he brought to the application of technology to library service. Much to our good fortune, he has chosen to share with us through his many writings the insights he has found in his study of the fit between technology and the delivery of effective library service.

This year's winner shares the distinction, with the two previous winners, of being a former president of the division. In fact, he presided over the change from the venerable acronym ISAD to the new name of the division: Library and Information Technology Association (LITA).

It gives me particular pleasure to present this year's award, as it goes not simply to an esteemed colleague but to a valued friend. I first met Maurice (Mitch) Freedman at the first ALA conference I attended—the Midwinter Meeting of 1972. The first session I attended at that conference was a meeting of the Committee on Library Automation (COLA). I had gone to that meeting to report on NYPL's automated cataloging system, which had that month become fully operational with the publication of the book catalogs of the research libraries and of the mid-Manhattan library.

Following the COLA program, Mitch approached me, introduced himself, and inquired about the possibility of using the NYPL system to produce Hennepin County's catalog. The consequences of that afternoon were most salutary both for the Hennepin County Library (HCL) and for me personally. HCL acquired at no cost an automated bibliographic control system, and I gained a friendship that has endured for nearly a decade.

Thus, rather than dwelling on Mitch's professional accomplishments—which are already well known to you—I would prefer to say a few words about the man himself. Perhaps the best way to characterize him is to describe to you his office at

Maurice Freedman (left) receiving 1980 LITA Award presented by S. Michael Malinconico (right).

Columbia University. Prominently displayed on the walls are two enormous posters, one of Bertrand Russell and another of Lenny Bruce. A perhaps odd pair until one realizes that these men had one important attribute in common: neither of them accepted, without incontrovertible proof, truths supported by conventional wisdom alone.

Mitch, like the philosopher and satirist whose images grace the walls of his office, is an iconoclast who insists on more than the endorsement of reigning authority before he will embrace an idea; and he will work tirelessly to change the prevailing wisdom if he finds that it serves to frustrate rather than aid the delivery of the kind and quality of library service to which he feels the patrons of libraries are entitled. Likewise, though he was among the pioneers who helped introduce sophisticated technologies such as automation and micrographics into the operation of libraries, he has always maintained a healthy skepticism, which has prevented him from being seduced by the dry voices of the hollow men who proclaim marvels that are in reality only gilded figures of straw.

Just as Lenny Bruce refused to accept contemporary conventions regarding language and behavior, Mitch Freeman has refused to accept the sanctity of LC subject terminology. He, Sanford Berman, and Joan Marshall have served for more than a decade as LC's conscience, prodding our phlegmatic, de facto national library to action. Just as Bertrand Russell returned to the axioms of Giuseppe Peano in an attempt to secure the foundation of mathematics in formal logic and to

free that discipline of fuzzy thinking, Mitch has returned to the principles articu-
lated by Antonio Panizzi and Seymour Lubetzky, as the tests by which to judge the
claims of the self-assured mountebanks who regale us with newly coined biblio-
graphic wisdom.

In this regard I anxiously await the completion of his doctoral dissertation, in
which he explores the philosophical underpinnings of theories of bibliographic
control (a work that would have proved most useful during the protracted emo-
tional debate that surrounded AACR2).

I expect that it must be particularly gratifying for Mitch to accept his award in
this particular city. Although his physical roots are in the Northeast, I rather think
his intellectual and spiritual roots are here, or more precisely, in the city across the
bay—Berkeley. It was just about twenty years ago that Mitch, after graduating
from Rutgers University, Newark, enrolled as a graduate student in philosophy at
the University of California, Berkeley. While at Berkeley, his sense of social justice
and utter disdain for unsupported dogma—could one expect less of a student of
philosophy?—led him to become active in the free speech movement. Thus, we
find very early in his career a concern for social issues, a concern that reemerged in
his active involvement with the Social Responsibilities Round Table shortly after
joining the library profession.

Before leaving Berkeley, Mitch earned his degree in library science. Thus, he
earned his degree from one of the most prestigious library schools on the west coast,
and now plies his trade as associate professor at one of the most prestigious library
schools on the east coast, the Columbia University School of Library Service. If he
is only moderately successful in conveying to his students his dedication to the
delivery of quality library service, his steadfast conviction that technical services is
in reality the first step in the provision of effective public service, and a respect for
the supremacy of principle over expedience, his graduating classes will constitute a
more lasting and meaningful award than this simple gesture conferred upon him
by his professional colleagues.

News and Announcements

Programmers Discussion Group Meets: PL/I, the MARC Format, and Holdings

Twenty-two computer programmers, analysts, and managers met on June 29 in San Francisco for the formative meeting of the LITA/ISAS Programmers Discussion Group. In an informal and informative hour, the group established ground rules, started a mailing list, planned the topic for Midwinter 1982, and found out more about practices in fifteen library-related installations.

Programming Language Usage

What programming languages are used, and used primarily, at the installations? Nine languages turned up, excluding database management systems (and lumping all "assembly" languages together)—but one language accounted for more than one-half of the responses:

Language	Users	Primary
PL/I	14	13
Assembler/Assembly languages	8	5
COBOL	4	2
Pascal	3	1
BASIC	1	1
C	1	1
MIIS (A MUMPS dialect)	1	1
Fortran	1	0
SNOBOL	1	0

Note: some installations use more than one "primary" language.)

A second round of hands showed only four users with *no* use of PL/I.

MARC Format Usage

These questions are asked on an agency-by-agency basis. One agency made *no* use of the MARC communications format. None of those receiving MARC-format tapes were unable to recreate the format.

Eight of the fifteen agencies made *significant* internal-processing use of the MARC-communications-format structure, including the leader, directory, and character storage patterns; this question was made more explicit to try to narrow the answers. Thus, the MARC *communications* format is used as a *processing* format in a significant number of institutions.

Only three agencies use ASCII internally, most use of MARC takes place within EBCDIC. (All but three agencies were using IBM 360/370 equivalent computers—the parallel is clear.)

Computer Usage

As noted, all but three agencies use IBM equivalents in the mainframe range; three of those use plug-compatible equipment such as Magnuson and Amdahl. The other major computers are CDC, DEC/VAX, and Data General Eclipse systems. Smaller computers in use include DG, DEC 11/70, Datapoint, and IBM Series/1 units.

Home Terminals and Computers

Four of those present currently have home terminals. Three have home computers.

Future Plans for the Discussion Group

The Midwinter 1982 topic will be "Holdings," with some emphasis on dealing with holdings formats in various technical processing systems (such as OCLC, UTLAS, WLN, RLIN). An announcement and mailing list will go to all those on the mailing list, as will an October/November mailing with questions sent to the chair.

Those interested should send their names and addresses to Walt Crawford, RLG, Jordan Quad, Stanford, CA 94305. It is anticipated that papers on the topic may be ready by Midwinter; questions and comments are welcomed. *NOTE:* There will be no set speakers or panelists; this will be a true *discussion* group. The topic for the Philadelphia meeting will be set at Midwinter 1982.—*Walt Crawford, Chair, The Research Libraries Group, Inc.*

CHANNEL 2000

A test of viewdata system called CHAN-NEL 2000 was conducted by OCLC in Columbus, Ohio, during the last quarter of 1980. An outgrowth of the OCLC Research Department's home delivery of library services program, CHANNEL 2000 was developed and tested to investigate technical, business, market, and social issues involved in electronic delivery of information using videotex technology.

Data Collection

Throughout the test, data were collected in three ways. *Transaction logs* were maintained, recording keystrokes of each user during the test, thus allowing future analyses and reconstruction of the test sessions. *Questionnaires* requesting demographic information, life-style, opinion leadership, and attitudes toward CHANNEL 2000 were collected from each user in each household before, during, and after the test. Six focus-group interviews were held and audiotaped to obtain specific user-responses to the information services.

Attitudes toward Library Services

Forty-six percent of the respondents agreed that CHANNEL 2000 saved time in getting books from the library. Responding to other questions, 29 percent felt that they would rather go to a traditional library than order books through CHANNEL 2000, and 38 percent of the users felt that CHANNEL 2000 had no effect on their library attendance.

Forty-one percent of the CHANNEL 2000 test group felt that their knowledge of library services increased as a result of the CHANNEL 2000 test. In addition, 16 percent of the respondents stated that they spent more time reading books than they did before the test.

Eighty-two percent of the respondents felt that public libraries should spend tax dollars on services such as CHANNEL 2000. Although this might suggest that library viewdata services should be tax-based, subsequent focus-group interviews indicated that remote use of these services should be paid for by the individual, whereas on-site use should be "free." Sixty-three percent of the test population stated

that they would probably subscribe and pay for a viewdata library service, if the services were made available to them off-site.

Purchase Intent

Respondents were asked to rank-order the seven CHANNEL 2000 services according to the likelihood that they would pay money to have that service in their home. A mean score was calculated for each CHANNEL 2000 service, and the following table shows rank order of preference.

Rank Order	CHANNEL 2000 Service
1	**Video Encyclopedia** Locate any of 32,000 articles in the new Academic American Encyclopedia via one of three easy look-up indexes
2	**Video Catalog** Browse through the videocard catalog of the Public Libraries of Columbus and Franklin County, and select books to be mailed directly to your home
3	**Home Banking** Pay your bills; check the status of your checking and savings accounts; look up the balance of your VISA credit card; look up your mortgage and installment loans; get current information on BANK ONE interest rates
4	**Public Information** Become aware of public and legislative information in Ohio
5	**Columbus Calendar** Check the monthly calendar of events for local educational and entertainment happenings
6	**Math That Counts!** Teach your children basic mathematics, including counting and simple word problems
7	**Early Reader** Help your children learn to read by reinforcing word relationships

The final report, mailed to all OCLC member libraries, was published as CHANNEL 2000: Description and findings of a viewdata test conducted by OCLC in Columbus, Ohio, October–December 1980. Dublin, Ohio: Research Department, Online Computer Library Center, Inc., 1981. 21p.

NOTIS Software Available

At the 1981 ALA Annual Conference in San Francisco, the Northwestern Univer-

sity Library announced the availability of version 3.2 of the NOTIS computer system. Intended for medium and large research libraries or groups of libraries, NOTIS provides comprehensive online integrated-processing capabilities for cataloging, acquisitions, and serials control. Patron access by author and title has been in operation for more than a year, and version 3.2 adds subject-access capability as well as other new features. An improved circulation module and other enhancements are under development for future release.

Although NOTIS, which runs on standard IBM or IBM-compatible hardware, has been in use by the National Library of Venezuela for several years, Northwestern only recently decided to actively market the software, and provided a demonstration at the ALA Conference. A contract has been signed with the University of Florida, and several other installations are expected within a few months.

Further information on NOTIS may be obtained from the Northwestern University Library, 1935 Sheridan Rd., Evanston, IL 60201.

Bibliographic Access & Control System

The Washington University School of Medicine Library announces its computer-based online catalog/library control system known as the Bibliographic Access & Control System (BACS). The system is now in operation and utilizes MARC cataloging records obtained from OCLC since 1975, serials records from PHILSOM serials control network, and machine-readable patron records. Features of interest in the system are:

1. Patron access by author, title, subject, call number, or combination of keywords. The public-access feature has been in operation since May 1981.

Online instructions support system use, minimizing staff intervention. User survey indicates a high degree of satisfaction with the system.

2. Low cost public access terminal with a specially designed overlay board.

3. Barcode-based circulation system featuring the usual functions, including recalls for high demand items, overdue notices, suspension of circulation privileges, etc.

4. Cataloging records loaded from OCLC MARC records by tape and from a microcomputer interface at the OCLC printer port. Authority control available on three levels: (a) controlled authority, i.e., MeSH or LC, (b) library-specific assigned authority, and (c) word list available to user.

5. Full cataloging functions online, including editing, deleting, and entering records.

6. Serials control from PHILSOM system. PHILSOM is an online distributed computer network that currently controls serials for sixteen medical school libraries. PHILSOM features rapid online check-in, claims, fiscal control, union lists, and management reports.

7. Five possible displays of the basic bibliographic record, varying from a brief record for the public access terminal to complete information for cataloging and reference staff.

8. Two levels of documentation available online.

The software is available to interested libraries, bibliographic utilities, or commercial firms. Contact: Washington University School of Medicine Library, 4580 Scott, St. Louis, MO 63110; (314) 454-3711.

Book Reviews

The Future of the Printed Word: The Impact and Implications of the New Communications Technology. Edited by Philip Hills. Westport, Conn.: Greenwood, 1980. 172p. $25. LC: 80-1716. ISBN: 0-313-22693-8 (lib. bdg.).

The character of this volume is as much that of a topical journal or annual review as that of a monograph. A dozen authors have contributed thirteen chapters, all but one prepared especially for this publication. Ten of the chapters are by British authors, two by Americans, and one by European Community personnel located in Luxembourg. An amusing *Punch* satire about BOOK (Built-in Orderly Organized Knowledge) is reprinted as an unnumbered fourteenth chapter.

In an excellent opening essay, John M. Strawhorn notes: "In this book, the expression *printed word* is construed very broadly, to include words in any kind of display: paper, microforms, CRT's, plasma panels and so on." His essay is a terse but pointed review of the organization of information transfer, some current trends, factors affecting acceptance of new technologies, and some broad projections for the future.

Provocative essays by Maurice B. Line and P. J. Hills, editor of the volume, explore the printed word from the points of view of a bookperson and an educator. In one of the most elegant metaphors to appear in information science literature, Line suggests: "The printed butterfly will emerge from its electronic chrysalis, but it will also return again to it in due time. The vast majority of documents will thus be stored in electronic (chrysalis) form, but the majority of those used at any given time will be in their printed (butterfly) form."

Two incisive and thorough chapters on official information by Patricia Wright systematically explore the use of old and new technologies for forms, leaflets, and signs.

Wright makes acute and useful observations on how technology can hinder or help gathering and dispersion of governmental information.

The Graphic Information Research Unit of the Royal College of Art has done excellent work in recent years in exploring how various display options affect comprehension. Linda Reynolds provides a good essay, "Designing for the New Communications Technology," based on that research.

The review of prospects for electronic journal publishing by Donald W. King is a good overview, especially for beginners. A chapter on Euronet DIANE describes problems in creating an online database capability in the European political environment. Chapters on printing technologies, microforms, and videodiscs cover all major alternatives but suffer from brevity. Two brief but competent speculative essays, which add little, complete the volume.

The work lacks a general index, but the organization of chapters makes this a minor flaw. Use of presumably common British acronyms without explanation, especially in credits and citations, is an irritant for non-U.K. readers.

The work would make an excellent supplementary text for a course on the history of the book. Practitioners in publishing or library and information science will find much of interest.—*Brian Aveney*.

Turnkey Automated Circulation Systems: Aids to Libraries in the Market Place. Edited by Judith Bernstein. Chicago. American Library Assn., 1980. 332p. $10.50.

When my library entered the marketplace for an automated circulation system, I searched the literature for aids. Had I found this book at that time I would have been disappointed. What I would expect from a 332-page book with a subtitle, "Aids to Libraries in the Market Place," would be numerous examples of what had been done

before. I would expect samples of the analyses that other libraries had done to justify entering the marketplace, samples of the RFPs that had been sent to vendors, and samples of the contracts that had been signed. I would like to see a case study (or two) of the complete process of procurement. Admittedly, this expectation is somewhat of an ideal, but these are "aids" that we searched for and that other libraries now ask from us.

What does this book provide? An editorial introduction gives a sense of the difficulties of the marketplace and the frustrations encountered in it. A two-page bibliography gives a reasonable selection of readings to provide a background for decision making. A discussion titled "Hiring a Consultant—Why and How," is a very useful enumeration of details to be considered in the decision to hire a consultant and in the agreement with a consultant. A model request for proposal is a good synthesis of the details to be included in almost every library's RFP and thus provides a starting point for the library new to the marketplace. All of this is what I consider to be the substance of this book, and it ends at page 40. The remaining 292 pages are devoted to the "profiles" of individual libraries which have installed automated circulation systems. The profiles are intended to assist in the identification of libraries to be contacted for further information, but provide little useful information by themselves.

My primary objection to this book is the misleading nature of the citation. One expects more than three hundred pages of "aids" and finds a directory with a forty-page preface. But for the librarian new to the marketplace it may be worth the price.—*Alan E. Hagyard, Yale University Library, New Haven, Connecticut.*

Archives and the Computer, by Michael Cook. London: Butterworths, 1980. 152p. $29.95. LC: 80-41286. ISBN: 0-408-10734-0.

Michael Cook recognizes the special predicament of the archivist whose job consists of trying to satisfy three contradictory needs: (1) the need to arrange and describe archives by their provenance, (2) the need to store them most efficiently by shape and size, and (3) the need to access them to answer inquiries that are mostly subject-oriented. The solution to these conflicting requirements may come from the computer. As Cook says, "The speed and variety of computerized lists and indexes derived from a single data base could solve this problem by producing finding aids in all possible sorts of order."

In a very handsomely produced, sturdily bound book, *Archives and the Computer,* Michael Cook, archivist of the University of Liverpool, reports on various computer systems serving the needs of the archivists. His book starts with a general discussion on the nature of automated systems and their relation to manual ones. This is followed by the description of a select group of archives systems—some still in use, others put to their well-deserved rest after a few years' use. He covers records management systems (i.e., the area of handling current records) and archives management systems (i.e., the handling of noncurrent documents). In the final chapter Cook moves the discussion away from computer processing of traditional, familiar forms of archival material, focusing instead on processing archives that are themselves machine-readable data files. How does the archivist accomplish all of the necessary tasks if the archives are not readable by the human eye? How does he appraise, arrange, describe, and access them?

I like Mr. Cook's cautious and sober attitude. Talking about system design, he remarks, "At this stage decisions will be made which will be irrevocable in practical terms, and may cause much trouble later." About implementation and testing, "computer systems should help people to work more effectively in a more interesting environment; if they fail in this, or appear to fail, there is something wrong, and it would perhaps be better not to introduce the change."

The records management systems he describes are used by British county and city record offices. An interesting feature in one of them, a system called ARMS, is a printout that tabulates for each class of documents the number of requests in a year, per year stored. This printout could be very helpful in modifying established retention periods on the basis of experience.

The following archives systems are described: PROSPEC (adopted by the Public Record Office of London), NARS A-1 (used by the National Archives of the USA), SPINDEX (first used by the National Archives and the National Historical Publications and Records Commission), SELGEM (used by the Archives of the Smithsonian Institution), STAIRS (an IBM system, used, among others, by the House of Lords Record Office in London), PARADIGM (developed and used at the University of Illinois), MISTRAL (used by the National Archives of Ivory Coast), and ARCAIC (used and abandoned by the East Sussex Record Office). Of all these systems, I found the description of SELGEM the most educational. Besides listing the fields making up a computer record, Cook shows an example of an actual record as it appears in the master list, and as it appears in the printed guide to the archives. He also includes an actual segment of the name/subject index.

Although there is a brief mention about the choice between networking versus isolated, separate systems, the book does not speculate about the possibility of a network of many institutions building a common database. Nor does the author discuss the much debated and very timely question of whether archivists could possibly agree on a uniform computer record for the description of manuscripts and archives, similar to the way in which librarians have agreed on using the MARC formats for the description of their materials.

A glossary of technical terms, a "select directory" of archival systems, and a "select bibliography" are useful additions to the main text.

This book is more recommended to the archivist looking for a computer system than for the systems analyst who wants to learn how archives are processed.— *Suzanna Lengyel, Yale University Library, New Haven, Connecticut.*

The Library and Information Manager's Guide to Online Services. Edited by Ryan E. Hoover. White Plains, N.Y.: Knowledge Industry Publications, 1980. 270p. $29.50 hardcover, $24.50 softcover. LC: 80-21602. ISBN: 0-914236-60-1 (hardcover);

0-914236-52-0 (softcover).

Hoover and seven colleagues provide an overview of the main issues and techniques involved in starting and managing an online retrieval service. The emphasis is on a library setting—the implicitly broader focus conveyed by the title is not matched by any specific coverage of, for example, the online search activity of the for-profit information brokers, where funding, staffing, publicizing, and the search process itself are handled differently than in libraries.

The three large, general search services (Lockheed, SDC, and BRS) are used throughout for the descriptions and search examples, and their bibliographic databases inevitably receive the most attention. There is a noticeable slant toward the two agencies with which several of the contributors are or were affiliated—the University of Utah (which doesn't detract from the book's objectivity) and SDC (which does).

The chapters are of uneven quality and scope. Most of the obvious areas are covered—the available search systems and databases; equipment needs; search techniques; managing an online service in a library; training searchers; promoting service; and measurement and evaluation. Taken as a whole, the book is a good state-of-the-art report, even though it is already becoming outdated in terms of industry facts. The numerous charts and tables serve to flesh out the text, but do we really need six photographs of terminals (two of them showing the same searcher at the same terminal, the only difference being that in one there is an onlooker) to illustrate that "some searchers prefer to have the user present"?

Brief chapters on the growing network of online user groups, and on the future of online services (largely derived from Lancaster) end the text, and the book has a serviceable bibliography, glossary, and index.

Six years ago I reviewed one of the first KIPI publications—it was in typescript, comb-bound, a little more than one hundred pages, and it cost $24.50. This is a much better production and, considering inflation since 1975, it represents vastly better value for money. It should serve as a useful handbook for those of us in the field, as well as those just starting, for another

year or two.—*Peter Watson, California State University, Chico.*

Basics of Online Searching, by Charles T. Meadow and Pauline Atherton Cochrane. New York: Wiley, 1981. 245p. $15.95. LC: 80-23050. ISBN: 0-417-05283-3.

The use of online information retrieval services is becoming widespread throughout the information community, whether in traditional libraries or in business, industry, or government offices. The need for trained searchers is evident by looking at the job advertisements and at the quantity of training programs being offered around the country. The programs presented by the Machine-Assisted Reference Section (MARS) of the Reference and Adult Services Division of ALA are always packed. The librarians attending ALA Annual Conferences seem to be hungry for any information available about online information retrieval services. This text fills an obvious need for the professional who attended library school before course offerings in online information retrieval were available. Although online information retrieval is now being taught in most library and information science curriculums, there have been only a few attempts at providing a textbook for beginning students, and none of those has been very successful since the Lancaster and Fayen *Information Retrieval Online* in 1973.

Basics of Online Searching is a text intended "to teach the principles of interactive bibliographic searching . . . to those with little or no prior experience. The major intended audiences are students, working information specialists and librarians, and end users, the people for whom all this searching is done." Because the authors have done an excellent job of targeting their audience and sticking to that target, this text will be useful at the introductory level. The authors cover the elements of interactive searching including the reference interview, Boolean logic, search strategy development, telecommunications and equipment, basic database structure, selective dissemination of information, and how to get help from search-service vendors.

The text is relatively free of jargon and does a good job of defining in context new terms as they appear. The authors begin with basic definitions and a brief overview of the process of interactive searching. The reference interview and search strategy development is covered adequately, first with an introduction and then in a later chapter providing more detailed information. Telecommunications and computer equipment are covered in enough detail for the novice. The next five chapters cover search language, databases, various types of text searching, and how to get on and off the computer. This section of the book uses examples that show the different approaches to the same process on three different systems—BRS, ORBIT, and DIALOG. The authors do not lose sight of their intent to demonstrate the principles of online searching. There is a brief chapter on selective dissemination of information (SDI) and cross-file searching. The chapter explains how SDI is used and gives examples of constructing and saving a search for SDI on each of the three systems. The last chapter of the book, "Search Strategy," is especially good. There seemed to be something beyond the basic elementary information of the preceeding chapters. The authors clearly demonstrate concept development and search strategy formulation.

The authors do an excellent job of integrating the discussion of the three major search service vendors, Lockheed's DIALOG, System Development's ORBIT, and Bibliographic Retrieval Services, Inc. Examples are used from each of the services with a discussion of the differences. The book does clarify the similarity of the services by showing how each function can be accomplished on each system. Searchers using only one system now might use this text to see how easily their knowledge could be transferred to another system.

Problems with the text do not abound, but there are some that should be brought to the attention of the reader. There is a slight problem with the format of the examples. The reviewer found herself searching for the completion of a paragraph of text on a few occasions. The examples are very good and clear; they are simply not separated from the text adequately for easy reading. There were a couple of instances of unnecessary redundancy. There were two separate

discussions, one on truncation and one on searching word fragments, which could have been improved by integration into one section. There was a repetition of "steps in the presearch interview and the online search" in chapter 3 and then again in chapter 12. This is almost a page of steps, which are very good, but a simple reference back to the earlier list would have sufficed. But the biggest problem with the text in the eyes of this reviewer is that of omission. There was no discussion of citation searching, evaluation of search results, and no mention of the various training options available for the novice searcher. This reviewer would like to have seen more information on where to go next as guidance to the novice. The one hundred pages of appendixes seem unnecessary and will soon be out of date. Library school teachers planning to use this as a text would do well to request free, up-to-date materials rather than relying upon the documents in the appendix, which are more than a year old at the time of this writing. Most every book on this topic has made the same mistake of reprinting search-service and database-producer literature.

Overall, however, the authors have succeeded very capably in their intended endeavor "to teach principles, rather than the detailed mechanics of any particular search system." There is a place in the literature for this very basic text, which is well written, uses clear examples, and teaches in an understated way. For those people who are afraid of automation, afraid to touch a computer terminal, and are insecure about their ability to do online searching, this book will relieve most of those fears and insecurities. The authors acknowledge their desire to give simple instructions and offer a chapter called "Assistance" for people who need more help. Novices might assume they could read this book, purchase a terminal, get a password and system manual, and begin searching. As a matter of fact one could do this, but the results would likely be a discredit to the search-service vendor because of a lack of system-specific training on the part of the searcher. Most people, like this reviewer, can conceptualize a new process, but would feel more comfortable with some type of formal hands-on

training—even for half a day. There are too many little things that can be an impediment to success.

The reviewer would heartily recommend this book to inexperienced searchers and library school students but would warn the experienced searchers that there is nothing new for them.—*Carolyn M. Gray, Western Illinois University, Macomb.*

Quick*Search Cross-System Database Search Guides. San Jose, Calif.: California Library Authority for Systems and Services, 1980. 21 charts. $75 (CLASS members), $95 (nonmembers). ISBN: 0-938098-00-4.

The CLASS On-Line Reference Service (COLRS) is a cooperative program for public, academic, and special libraries offering training and consultation on almost any aspect of online reference searching through the major commercial vendors of databases. This service is a part of CLASS, the California Library Authority for Systems and Services, and acts as a contact point for searchers and the database industry through vendor-training sessions, database training, and the coordination of large group contracts with DIALOG Information Services and Bibliographic Retrieval Services (BRS). This close relationship to the online industry gives CLASS a unique position from which to supply information on databases from a multiple search-system perspective. The publication of the *Quick*Search Cross-System Database Search Guides* is a natural outgrowth of the COLRS program in training and consulting.

The twenty-one charts in *Quick*Search* show the formats used to search for information in a specific database across the two or three vendors offering the database commercially. The databases were selected as the most commonly searched through the major commercial search services: Bibliographic Retrieval Services, DIALOG Information Services, and System Development Corporation Search Service (SDC). Eight databases in the sciences, eight in the social sciences, and five multidisciplinary files are included in the complete set. Two subsets of the science and multidisciplinary files, and the social science and multidisciplinary files are available for $60 for CLASS members

and $80 for nonmembers. The eight science databases are BIOSIS, CAB Abstracts, COMPENDEX, ENERGYLINE, EN-VIROLINE, Food Service & Technology Abstracts, INSPEC, and Oceanic Abstracts. The social science files are ABI/INFORM, ERIC, Exceptional Child Education Resources, Library and Information Science Abstracts, Management Contents, Psychological Abstracts, Social SCI-SEARCH, and U.S. Political Science Documents. The multidisciplinary databases are Conference Papers Index, Comprehensive Dissertation Index, NTIS, PAIS International, and SSIE Current Research.

The stated purpose of the *Quick*Search Guides* is to aid the experienced searcher who must use databases from more than one search service by showing the formats for each vendor of a database side by side for comparison. Because most searchers tend to use a database on only one system, the *Guides* are really more appropriate to an organization where several searchers may be using the same database through different systems and a "universal" quick-reference chart is needed. Because each *Guide* covers only one database, the level of detail shown is much greater than in the simple-command comparison charts previously published.

The *Guides* are arranged to show particular features of the databases as they are used on the different search systems. The file label used to access the database and those fields that are searched when a term is entered with no restriction (the basic index) are shown at the top of each chart. The fields used in subject searching follow and show the field codes used to restrict subject searches, along with the format used online to enter search terms. The typical fields illustrated are title, subject descriptor, identifier, abstract, and category or section code. These fields vary according to database, but include the majority of subject access points used in the file. The balance of the chart is used to illustrate the field codes and formats used to retrieve information from other access points in the database such as author, journal source, language, publication date, document type, report numbers, or update code. These alternate access points vary widely by database, but

each chart provides information on limiting searches by date, language, or update code at a minimum. The *Guides* supply a useful amount of information for the experienced searcher needing a prompt on a form of entry for the fields available in a database, but a good understanding of the search system is required to use them properly.

Given the close contact CLASS has with the database producers and online vendors, it is somewhat surprising to find inaccuracies and some misinterpretation in some of the *Guides*. In the preface, for instance, the editor states, "In many BRS files, UJ and UN are paragraph labels used in addition to DE, MJ, and MN. They are used to indicate major (UJ) or minor (UN) single word descriptors, similar to the DF in DIALOG and IW in ORBIT." It is true that DF is used in DIALOG to indicate a single-word descriptor, but in ORBIT the code is IT. In BRS, UJ and UN mean the term so restricted is an "unbound" part of a multiword descriptor—not a single-word descriptor (see BRS/ERIC database guide, p.14). The use of IW in ORBIT retrieves "unbound" words from the IT field. The most trouble in the charts appears to be in the ORBIT sections. The basic index is misrepresented in several files and the IW field is only irregularly listed, even when it is present in the SDC version of the database. Suggestions on the use of SENSEARCH and STRINGSEARCH are not consistently illustrated for fields that cannot be directly restricted in some databases on ORBIT, such as abstract or supplementary index terms. Many times the suggested search entry would not restrict retrieval to the field indicated on the chart. These inaccuracies would probably not doom an experienced searcher to failure in using a database, but they are annoying and do little to inspire absolute confidence in the information presented.

CLASS is to be complimented on the graphic representations in *Quick*Search* and the heavy stock used for the *Guides* (the paper will probably outlive the information printed on it). Addenda are planned for those databases changed or reloaded since the preparation of *Quick*Search* in October 1980, and a second edition is already under consideration. The *Quick*Search*

Guides are not meant as a replacement for vendor or database documentation and, in fact, are simply repackaged versions of the basic file descriptions available from the online vendors. Considering the price of this publication, organizations would do well to consider investing instead in detailed user guides and updates for their searchers in order to provide the most accurate and current information on databases on a specific system.—*Rod Slade, University of Oregon Library, Eugene.*

Viewdata and Videotext, 1980–81: A Worldwide Report. Transcript of Viewdata '80, First World Conference on Viewdata, Videotex, and Teletext, London, March 26–28, 1980. White Plains, N.Y.: Knowledge Industry Publications, 1980. 623p. $75 softcover. LC: 80-18234. ISBN: 0-914236-77-6.

Videotex 81. Proceedings of Videotex '81 International Conference and Exhibition, May 20–22, 1981, Toronto, Canada. Northwood Hills, Middlesex, U.K.: Online Conferences Ltd., 1981. 470p. $85 softcover.

Viewdata '80 and Videotex '81 were two state-of-the-art conferences for the emerging videotex field. Videotex is the generic name for mass-market, consumer-oriented information retrieval systems of low cost and relative ease of use. Videotex, as a technology, is divided into teletext systems and viewdata systems. Teletext systems sequentially broadcast information using a portion of the television signal. Subscribers, using a special decoder, can select individual pages from the several hundred offered. Viewdata systems, on the other hand, are quite like online information systems except for their use of a television as a display device, their simplicity, and their broader range of transactions and information.

These conference proceedings will be of interest to a limited audience. They are not for the complete beginner. Nor will they provide hours of entertaining reading. Neither meets academic publication criteria; many of the papers are fluff, outlines, or sales pitches. Both proceedings have their share, unfortunately large, of uninformative articles.

But if you are seriously interested in vid-eotex's technology, uses, and social implications, then by all means at least skim the 1981 conference papers. The proceedings *do* describe the state of the art. Moreover, the two proceedings, taken together, show some of the changes in the videotex field in the last year . . . and not only in the spelling of "videotex."

As state of the art, the Viewdata '80 conference proceedings are already superseded. Most of the material has been adequately covered by now in other publications at a much lower cost. There are two exceptions to this, both worth noting. The proceedings has several excellent articles on the Japanese Captain system, the best published on that system. Of additional interest is a report on Control Data Corporation's (CDC) market test of their PLATO educational system. Their report suggests a large consumer market for high-quality educational services even at a relatively high price.

The Videotex '81 conference proceedings are, of course, more current. There are four major topics of interest in the proceedings. Firstly, there are several good presentations on videotex services, such as electronic publishing, retailing, and banking. There is an excellent discussion on what videotex means to newspapers, both in opportunities and threats.

Secondly, and particularly recommended, is a paper by Tydeman and Zwimpfer of the Institute for the Future. The paper outlines some of the social changes and problems that may result from large-scale videotex implementation.

Thirdly, there are updates on the existing videotex technologies and efforts from the French, Japanese, Canadian, and British groups. The British are perhaps the most interesting since they have a year of operational experience with their viewdata system, Prestel. They state that most usage was from the business community, and their reports suggest that services are shifting to attract that market. If this is the case, it is a significant change from the original consumer orientation. There is also a good article on a Prestel information provider's first year. Of additional interest is that Prestel-compatible databases and systems are being constructed in Britain. Thus, people will be

able to access different systems using the same protocol.

Finally, there are numerous fascinating papers on American efforts. The Americans, in contrast to the British, seem very unsettled; there is still a multiplicity of designs. (AT&T's decision on a modified Telidon standard, not reported in the proceedings but a major event of the conference, may ameliorate that.) The papers indicate overall that the "classic" definitions of viewdata and teletext will crumble or will be supplemented in the face of 100-channel, two-way cable systems. Several papers document how these new cable capabilities will provide channels for large amounts of information to be delivered by teletext, viewdata, or hybrid systems. A paper by Simon notes that cable will not only provide large audiences for information services but will also eliminate some of the traditionally defined viewdata functions. For example, people will not buy commodity prices from a viewdata service if that same information

is available on a cable channel at a lower price.

Unfortunately, there are some topics missing from the 1981 conference proceedings. Consumer-oriented educational services are mentioned little. System-performance or human-factor considerations are rarely analyzed. There is much discussion of what services should be offered, but there is little discussion of how those services should be offered. No presentation is made on how to design very large databases for ease of use.

Particularly distressing is the relative omission of the word "quality" from the American papers in both proceedings. One cannot expect every home to be wired to access the entire Library of Congress. Nonetheless, one can hope that videotex will not become merely a medium for used-car advertising.—*Mark S. Ackerman, Department of Computer and Information Science, Ohio State University and OCLC, Inc., Columbus.*

CALL FOR PARTICIPATION

American Society for Information Science 45th Annual Meeting Columbus, Ohio Oct. 17–21, 1982

INFORMATION INTERACTION

The theme for ASIS—82 is based upon the premise that information is derived from, and is the basis for, interaction among organizations as well as individuals. You are invited to participate on topics which include - but are not limited to - the following:

- **Information Systems**—Business, community, home, industry, library, professional, publishing

- **Information System Design for Effective Interaction**—Human factors, database design, standards, information analysis and reduction, computer and network configurations, software developments

- **Organizational Interaction**—Electronic messaging, teleconferencing, word processing, information management in organizations

- **Educational Interaction**—Programmed instruction, computer-based education, use of multi-media (microforms, videodisc, video tape)

- **Communication Technologies**—Transmission media (teletext, viewdata, videotex), future equipment

- **Graphic Communications**—Aids to human comprehension of information, computer graphics (animation, art, photo-composition, simulation), micrographics

FOR FURTHER INFORMATION CONTACT:

John Fried— Room LA
ASIS—82 Technical Program Chairman
Battelle Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

STANDING ORDERS:
THE ULTIMATE TEST OF YOUR BOOKSELLER.

Certainly at Blackwell North America we are proud of all the services we offer as a leading bookseller. But none of them is more demanding of our people and inter-active computer system than our standing order service for books in series.

Kathleen Olesen, at right, and her staff maintain our computer-based history of over 20,000 active series. It is a proven system, and goes back farther than any other—a full ten years. Here's what it means to our standing order customers.

First, oversights are virtually eliminated. As new titles are reviewed by our staff, those which are part of a series are immediately routed to Standing Orders for history verification.

Second, as a major book buyer, we combine efficiency and clout.

We order in volume from all publishers, and sell at competitive prices. Unlike periodical vendors, we do not drop ship books in series. You benefit from timely series shipments with a single invoice.

Third, duplication is eliminated. Our computer system guarantees that you will not receive series volumes on approval if you have a standing order for the series. This simplifies life, and means higher efficiency for us both.

Fourth, our monthly microfiche updates 20,000 series records and assists in record keeping and ordering. Also available on request is a library-specific hard copy standing order record that can be tailored for main and branch libraries.

But the most important test we pass is this: We understand the needs of serials librarians, and attend to them personally. Call your Regional Sales Manager or nearest distribution center toll free and have them put Kathleen Olesen to the test today.

BLACKWELL
Blackwell North America, Inc.

10300 S.W. Allen Blvd.   1001 Fries Mill Road
Beaverton, Oregon 97005 Blackwood, New Jersey 08012
Telephone (800) 547-6426 Telephone (800) 257-7341

OFFICES IN: OXFORD, ENGLAND; BEAVERTON, OREGON; BLACKWOOD, NEW JERSEY; NOVATO, CALIFORNIA; LONDON, ONTARIO, CANADA; HOUSTON, TEXAS; BOSTON, MASSACHUSETTS; ATLANTA, GEORGIA; CHAMPAIGN, ILLINOIS; CANBERRA, AUSTRALIA AND FREIBURG, WEST GERMANY.