Nancy Olson

volume 8
number 2

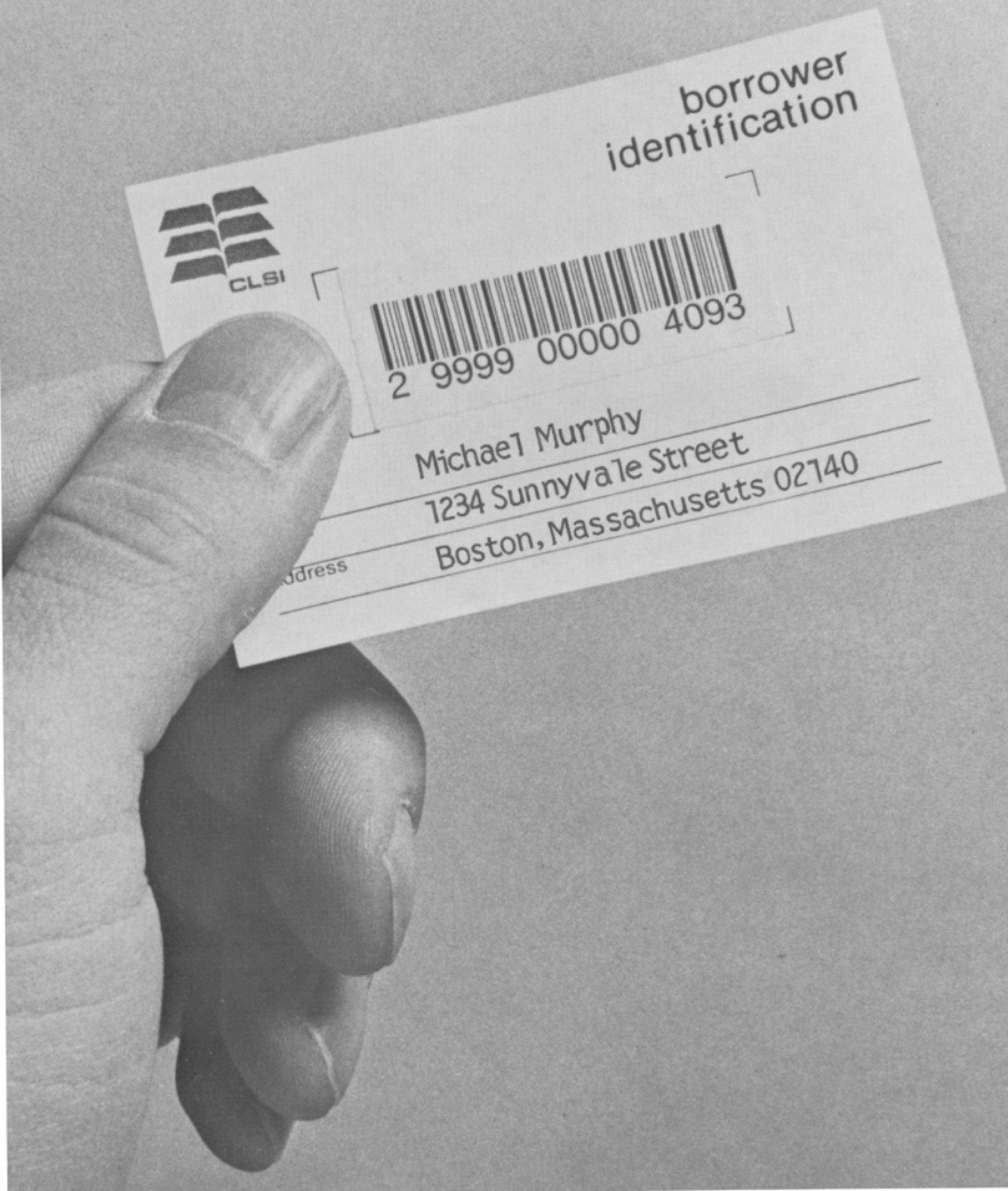
ISSN 0022-2240

journal of library automation

june, 1975

jlauay 8 (2) 81-168 (1975); j libr automat

We're CLSI. Here's our card.



JOURNAL OF LIBRARY AUTOMATION

Volume 8, Number 2: June 1975

CONTENTS

JLAUAY 8 (2) 81-168 (1975); J Libr Automat

Manuscripts of articles should be addressed to Mrs. Susan K. Martin, Editor, *Journal of Library Automation,* The General Library, University of California, Berkeley, CA 94720. Technical communications and news items should be addressed to Don L. Bosseau, *JOLA Technical Communications,* The Library, Emory University, Atlanta, Georgia. Copies of books submitted for review should be addressed to Peter Simmons, School of Librarianship, University of British Columbia, Vancouver, British Columbia.

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

The *Journal of Library Automation* is a perquisite of membership in the Information Science and Automation Division, American Library Association. Subscription price, $7.50, included in your membership dues. Nonmembers may subscribe for $15.00 per year.

Circulation and Production: American Library Association, 50 E. Huron St., Chicago, IL 60611. Advertising: Leona Swiech, Advertising Traffic Coordinator, ALA Headquarters. Please allow six weeks for changes of address.

Publication of an article or advertisement in the *Journal of Library Automation* does not constitute official endorsement by ISAD or ALA.

Abstracted in *Chemical Abstracts, 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.

Facing Our Technology

The technology of information display played a role in shaping the nature of the message long before this age of McLuhan. According to Hellmut Lehmann-Haupt, Gutenberg's "powerful creative dream was the reproduction in print of the great medieval liturgical manuscript in all its colorful beauty."* It was necessary for Gutenberg and those following to transcend his original dream so that printing could develop into the medium of wide-scale information dissemination that made our current civilization possible.

When the computer was first used for information display by persons then called documentalists, its ability to rearrange data was immediately recognized. Hans Peter Luhn developed the KWIC display technique for giving multiple access to titles by arranging entries alphabetically under each significant word. Mortimer Taube and others developed a variety of techniques involving precoordinate and postcoordinate indexing. A number of systems using roles and links have been developed. The PRECIS system, fathered by Derek Austin, has carried many of these techniques into the subject index to the *British National Bibliography*.

With the exception of the understaffed, rather inconclusive work of the Technical Processes Research Office under Richard Angell at the Library of Congress, there has been little attention paid by mainstream American librarianship to the appropriateness of the intellectual content of our records in light of the emerging technologies. The SPIRES/BALLOTS effort at Stanford and the LEADERMART effort at Lehigh both addressed themselves to using the power of the computer to improve search access to library catalogs. Both, however, ignored the nature of the content of the records they manipulated.

Our computerized catalog systems are still largely card catalog facsimile reproduction systems. While the specialized information retrieval services and many special libraries have been developing successful new methods of indexing and cataloging, the mainstream of American librarianship has steadfastly dodged the issues.

Our current forms of subject access to library collections have proven inadequate. It is inappropriate for our expensive automation efforts to continue to perpetuate the already bankrupt subject approach now used in our

* Hellmut Lehmann-Haupt, *Gutenberg and the Master of the Playing Card* (New Haven: Yale Univ. Pr., 1966), p.47.

new computer systems. We have great new tools in our grasp if we have the courage to use them.

We have left behind our second generation computer systems and developed far more powerful software techniques using the capabilities of third generation devices. We must now face up to the intellectual challenge of providing new forms of access to library collections and cease being facsimile reproducers.

BRIAN AVENEY
UNIVERSITY OF CALIFORNIA

MEDLARS II: A Third Generation Bibliographic Production System

Robert V. KATTER and Karl M. PEARSON, Jr.: Information Retrieval and Library Systems Projects, System Development Corporation, Santa Monica, California.

MEDLARS II, the replacement for the MEDLARS system used by the National Library of Medicine over the past decade, incorporates a number of major advances in the state of the art for massive information retrieval systems: on-line access to a number of very large bibliographic files, an efficient throughput figure, validation and mapping of inputs against authority files, and modularity and parametric programming to provide the flexibility needed to support future system enhancements. Major MEDLARS II subsystems described are: (1) specification maintenance, (2) input and release, (3) file maintenance, (4) retrieval, (5) publication production, and (6) management reporting. The MEDLARS II retrieval subsystem (ELHILL) incorporated the results of the experimental AIM-TWX service.

INTRODUCTION

The purpose of this article is to give an overview of MEDLARS II as an example of a major bibliographic processing system—one that supports on-line access to a number of very large files, has efficient throughput, and is operated on a single large-scale computer (an IBM 370/158) concurrently with other tasks and with minimal assistance from computer technicians. MEDLARS II represents a major gain in efficiency over MEDLARS I, and it extends the functions performed by that system and by the MEDLINE (formerly AIM-TWX) on-line retrieval service.

MEDLARS I was developed during the early 1960s by the National Library of Medicine (NLM) and its contractor, the General Electric Company. MEDLARS I was a batch retrieval system and ran on a Honeywell 800-200 computer. From its inception, the system was one of the largest and most successful library automation projects, and it served NLM well over its ten-year life span in supporting the rapidly growing publication production and reference searching functions for medical literature that could not have been sustained, and in some ways could not even have been attempted, without computer assistance. The original system continued to be modified and extended by the NLM programming staff, but by the end of the 1960s it was clear that a major redesign would be necessary in order to make use of new computer and information processing technology, to continue supporting the increasing workload, and to incorporate on-line retrieval. The history, principles, and detailed descriptions of the operating system have been documented elsewhere.[1]

NLM investigated several on-line data processing packages and, in 1969, settled on System Development Corporation's ORBIT® as the basis for the pilot AIM-TWX on-line searching service for *Abridged Index Medicus*. The ensuing success of on-line searching of a portion of the MEDLARS data base, as evidenced by the rapid increase in usage, was underwritten by the care NLM devoted to training medical librarians in using the service as well as by the depth of understanding on the part of the searchers themselves regarding the content and indexing structure of the data base. The sound design of the information retrieval command language and functions of ORBIT also played an important part. Experience with AIM-TWX identified the need for an expanded on-line system from the NLM.[2-5] By the end of 1972, AIM-TWX had become MEDLINE and was using an adapted version of the retrieval program (now called ELHILL II) to support retrieval on a much larger base. A history and description of the development of MEDLINE appeared in *Science*.[6]

As the success of the on-line retrieval service became obvious, NLM decided to contract with System Development Corporation (SDC) to assist in designing and implementing MEDLARS II. The effort was begun in fall 1971 and was completed by fall 1974.

MAJOR FEATURES OF MEDLARS II

MEDLARS II incorporates the set of functions performed by MEDLARS I, namely:

- *Input Processing*, which encompasses error detection, automatic validation of record elements such as author names and descriptors (subject headings) against authority files, and mapping of additional information from the authority files and cross-references into the input record. Records are released to the data base only when they meet prescribed standards for completeness and correctness.
- *File Generation and Maintenance*, both of system data base files and indexes thereto and of machine-readable output files for distribution outside NLM.
- *Search and Retrieval*, both in response to individual queries and to obtain contents for SDI and published products such as *Index Medicus*.
- *Publication Production*, which includes the sorting of bibliographic records into desired sequences; the addition of print headings, indexes, and cross-references; and the attachment of photocomposition

instructions indicating type size, font, and page-positioning of each entry element.

- *Management Information Reporting,* which consists of the production of processing run activity reports and of summary reports of system processing activities aggregated over various desired time periods.

To this array of capabilities of MEDLARS I, MEDLARS II adds five ingredients which represent a decade of growth in the state of the art:

- *Processing speed is increased, and processing cost per unit is reduced.* The new system's ability to handle large volumes of many kinds of work loads in a given period of time is significantly greater. More units of processing can be accomplished for the same amount of resources. This is due largely to the fact that the new program software is designed to take advantage of the larger, faster, more powerful computers available today.
- *Standards for data are higher.* Bibliographic records contain fewer errors because a higher degree of precision and error detection is embodied in the new system. This higher standard is made possible by the accumulated operational experience of NLM, together with more stringent data-checking capabilities offered by newer hardware and software.
- *Retrieval service responsiveness is increased.* The incorporation of interactive searching and retrieval, utilizing random-access storage, is a vast improvement over the batch-mode linear searching capability of MEDLARS I. All kinds of MEDLARS II bibliographic records—citations, authority records, and catalog entries—can be retrieved via a wide variety of access points. Search formulations may be as simple or complex as necessary, to fit both the user's and the data base's idiosyncrasies. User aids, such as saving search formulations for later use and automatically incorporating hierarchies of descriptors, are available where proven to be desirable through operational experience with AIM-TWX and MEDLINE.
- *The potential for creating new products is enhanced.* The new system's data base configurations and processing activities are controlled by tables of specifications stored on random-access disks, rather than being "hard-coded" into the various programs that make up the system. Modifying and expanding these tables is far easier than changing hard-coded programs. Therefore, it becomes not only feasible but also attractive to proceed with projects to amplify the scope and variety of data bases, retrieval services, and published products that can be offered.
- *A baseline testbed is provided for future developments.* Large strides have been made by library and information science and technology in the past decade. The feasibility of investigating and of exploiting innovative concepts often depends on the availability of a *baseline* system of sufficient generality, flexibility, and functional modularity

to permit new ideas to be converted to realities within the limits of available resources. The new system, MEDLARS II, represents a significant step in achieving such a baseline.

MEDLARS II, as was its predecessor, is a pioneering system for the support of information services of increasing sophistication. It will be ready to support the new methods of library and information science needed to cope with the phenomenal burgeoning of the world's scientific and technical literature in the next decade as predicted by Georges Anderla.[7] In addition to providing a major resource for a national on-line bibliographic network, MEDLARS II will be able to provide the interrecord linking within files via cross-citations and the large and complex record structures needed to incorporate the abstracts, summaries, and condensations that may, in the near future, supplant full-text items for many uses.

The inherent flexibility and functional modularity of MEDLARS II allows the system to support the sharing of library resources and processing of workloads on a network basis. New files with different record formats can be specified, designed, and easily implemented without interfering with the mainstream of activities. It is easy to install selected components of the system at other sites. For example, MEDLARS II system subsets have been installed on the computers at the Australian National Library at Canberra.

MEDLARS II SYSTEM DESCRIPTION: GENERAL CHARACTERISTICS

MEDLARS II is installed on an IBM 370/158 at the NLM in Bethesda, Maryland, and operates under VS2 without modification. For ease of maintenance and modification, the programs are written in PL/1. A few routines are programmed in BAL for the sake of efficiency or to perform special functions. With two exceptions (the Input Subsystem and the Retrieval Subsystem), all programs fit into computer core memory partitions of 256K bytes or less. The Retrieval Subsystem requires less than 300K bytes, and the Input Subsystem less than 500K bytes.

All programs are designed to run together in any desired combination without mutual interference in a multiprocessing environment. Interlocks are provided to prevent one program from trying to use a file that another program is updating. The flow of work through the system is routed through a series of files that perform "holding" functions between processing modules. A module early in the processing chain deposits the results of its processing in such a file as it completes its work; those results can be read by the next module in the chain on a convenient schedule. This process of interfacing modules through files is greatly facilitated by the use of a standardized record structure (called a "unit record" in this system) that is common to most of the files in the system.

Data files associated with each task, process, and product are identified by tasknames, publication and publication section names, processing indicators, and names of authorizing individuals that are input by order decks (control cards placed in the system input file). For users of the on-line facilities, there is a *user-Id*-keyed system for controlling access by individuals to data files and to operating commands and options.

MAIN SUBSYSTEMS AND DATA PATHS

The system contains eighteen program modules and several subordinate routines that are grouped into four main processing subsystems and two support subsystems. Figure 1 shows these subsystems and the information flow paths among them. The dashed lines connect the Specification Maintenance Subsystem (described later), and the dotted lines connect the other support subsystem, the Management Information Analysis and Report Generator (MIARG). The solid lines depict the main flows of processed data through the four main processing subsystems: Input (and Release); Select, Generate, and Maintain; Retrieval; and Publication Production.

The main bibliographic data processing flow, charted by solid arrows, starts at the left-hand center of the figure, with data being entered via key-to-tape devices located at NLM in Bethesda and at a nearby commercial keyboarding firm. Keyboard operators work from journal articles and data-input forms on which indexers, catalogers, or vocabulary analysts have recorded additional information. Each form results in one record being entered into the system. The information on the input tapes is processed by the Input Subsystem's Unit Record Input Control (URIC) module, which scans each record to perform editing and validation checks on data values in the record. URIC uses name and vocabulary authority files to perform any required reference mapping of content descriptors and authors' names. Detected errors are printed out on a proof listing on which clerks mark corrections for keyboarding and resubmission to the Input Subsystem. A major design goal of MEDLARS II was to catch most errors early, when the cost of correction is relatively small. After the Input Edit Loop (shown in the figure) has operated one or more times to correct the records, batches are released by orders given to the Batch Release (BATREL) module, which transfers the records to a holding file in preparation for further processing.

The records can then be processed by the Select, Generate, and Maintain Subsystem (the second block in the diagram of the major flow-path). Within this component, selection orders in the form of control cards input to the Selection and Formatting (CELLFORM) module cause the selection and transfer to output files of the records appropriate to each particular retrieval file, as a preliminary to the main step. As part of this preliminary operation, CELLFORM can be directed to reformat records by transferring only a specified part of each record, while discarding the rest. In the main step for the subsystem, the file generation and maintenance module (REFILL/EMPTY) is run for each retrieval file (a data base with an assigned name, such as the CATALOG FILE, the CURRENT

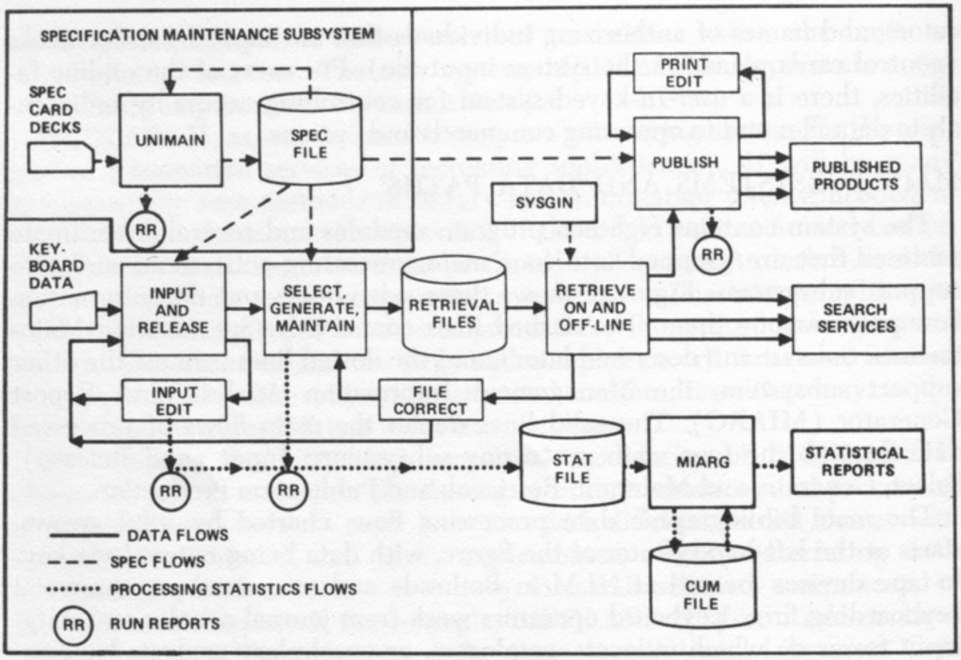


Fig. 1. MEDLARS II Main Components and Flows.

CITATION FILE, SDILINE, or MEDLINE) to be updated.

The Retrieval Subsystem (the third block in the main flow path of Figure 1) has separate programs for retrieving data base records in on-line and off-line mode. These two programs are much alike, except that ELHILL III (the on-line retrieval program) must support a large number of searching users at the same time, allocating its work load to each, while OFFHILL, since it is not operating interactively, processes each search to completion before starting another one. Consequently, the very large, time-consuming production searches involved in retrieving the contents for some publications are run by OFFHILL. A special connecting file between the two programs allows on-line ELHILL III users to formulate searches that are stored for execution by OFFHILL during low-workload periods or for running against files that are not currently available.

The procedure for correcting or updating a unit record in a retrieval file utilizes the Retrieval Subsystem to identify and make a copy of the record to be corrected. These copies are returned to the Input Subsystem, where they are corrected, revalidated, and released back to the appropriate data base using the same procedures followed for new records. Despite the delay inherent in the correction process, this method was chosen over the method of updating a record in place as a means of guaranteeing the integrity of the data base.

The retrieval programs can be regenerated by means of the SYSGIN program to allow them to handle a newly defined data base. Thus, new and different retrieval files can be added to the system with no reprogramming effort.

The main outputs from the Retrieval Subsystem can take either of two major routes. A number of Search Services channels connecting on-line interactive terminals may be operating, each delivering information retrieval products from various data bases. As many as fifty such terminals may be running simultaneously on the system, depending on other load factors on the computer. The channel to the Publication Subsystem consists of a number of kinds of holding files that accept the output of records identified by the Retrieval Subsystem searches. These constitute the various published products to be processed in the last main step: the production of formal publications and informal reports.

The Publication Subsystem contains a Publication Structure Processing module and a Publication Format Processing (photocomposition) module. In the first stage, Publication Structure Processing examines each record and assigns it a classification code which is then used by the subsystem to guide the processing of that record's contents. Data elements required to make up the content of a publication entry are selected from the unit record and converted to their "print" versions. Entries are then sorted into the sequences required for the publication and may be duplicated in cases where the element contents are required at more than one place in the publication. Headings, indexes, and cross-references are added, and a processing identification code is attached to each processed element. The correctly structured contents of the publication are then generated as the Sorted Publication Contents File for passing to the Publication Format Processing module.

In photocomposition processing, the Publication Format Processing module reads the identification code attached to each content element it draws in sequence from the Sorted Publication Content File. It consults the proper specification instructions, indicated by the code, and then selects specified type size and font; positions the element as required on the page; inserts interline leading and interelement spacing as required; attaches page numbers and running heads; and outputs the results in a form that is independent of any particular type of photocomposing device. At this point in the process, a particular Device Dependent Composer Module (DEVCOMP) is operated to convert the output to the unique format required for that device.* This arrangement allows MEDLARS II to utilize a new or different photocomposition device simply by creating an additional DEVCOMP module.

SUPPORTING SUBSYSTEMS

The Specification Maintenance Subsystem, depicted in the upper left

* The Publication Format Processing Subsystem contains a DEVCOMP for the Photon 901 and a DEVCOMP for line-printer output.

portion of Figure 1, maintains tables of parameters used for automatic guidance and control of the processing steps performed by the other subsystems. Specifications in MEDLARS II have two major roles:

1. To provide a *control* mechanism whereby system outputs (publications, data files, and reports) can be created, modified, or deleted without making program changes.
2. To standardize and unify the *definition* and *processing* of data elements and formats throughout the system, so that standard data-handling routines can be incorporated in programs as needed and so that the impact of changes in data definitions and structures can be identified, reduced, and localized.

The specifications concept, also known as table-driven programming or parametric programming, provides a degree of flexibility and system consistency far greater than can be achieved when all processing instructions are embedded in programs alone. In effect, the specifications provide limited and identifiable points at which minor changes can be made to the system; these points are located where such changes are most likely to occur and where their impact on existing system software is essentially nil. The use of specifications may also reduce program size, since one generalized routine may perform a set of similar operations on a number of different records or data elements.

Most importantly, specifications provide a device whereby library personnel desiring a modification in the content or format of system outputs are able to make the change without requiring assistance from computer programmers or affecting other portions of the system. By changing or adding a specification, the library staff member directly responsible for the production of a particular publication can personally cause a software change to be incorporated in the system to produce the desired alteration to, or addition of, a publication.

Along the bottom of Figure 1, the dotted lines connect the output of run status information from several major processing modules to the Management Information Analysis and Report Generator Subsystem. As indicated in the figure, run reports (RR) are produced at the end of the run of a module or subsystem. In some cases, the same run information is also routed to the statistics file to be held for processing by MIARG, which stores the results of such processing in the cumulative file. On demand, MIARG accesses the cumulative file to find specified data, format them, and create the desired management information report, typically containing processing activities of a subsystem summarized over a desired interval of time.

SYSTEM DEVELOPMENT NOTES

SDC developed MEDLARS II at its own Santa Monica headquarters, using a small team of analysts and programmers. At its largest, this team numbered sixteen, including the project head, the technical director, and the chief programmer. Development of the photocomposition module was subcontracted to Operating Systems, Inc. (OSI); they employed three programmers on the task. These programmers worked closely with the SDC team, occupying an adjacent office during program design and implementation. One full-time-equivalent systems analyst was employed during the design phase and again for the user manual preparation and training tasks. Several specialists, both within and without SDC, contributed to the system design, each being employed for a relatively short time.

Despite the costs for travel and communications, the continent-wide separation of NLM and SDC had some real benefits. At NLM expense, a remote job entry terminal was installed in an office adjacent to SDC's programmers, and NLM leased a line to its computer in Bethesda. Because it is penny-wise and pound foolish to economize on the machines needed by programmers to do their work, we also installed two keypunches, two 2741 on-line terminals, and a teletype for programmer use. These arrangements provided for full SDC control over its own staff, kept NLM and SDC staff from interfering with each other, reduced to a minimum the number of meetings and conferences that might otherwise have wasted manpower, and gave the programmers a considerable degree of control over the computer resources required to meet their needs. The significance of this control is underlined by the fact that the module with the severest production scheduling problems was the one that, because of its size, could not be run during the day throughout most of the development period, requiring programmers to work at unusual hours.

The development of MEDLARS II has been successful. Although the final portions of the system were installed several months beyond the initially targeted completion date, the bulk of the software was provided as needed. Several versions of the retrieval software, each containing new capabilities, were produced during the development period. The contract with SDC was on a fixed-price basis, and the system was delivered within budget.

There are a number of reasons for the success of the project. Members of NLM and its technical staff devoted a great deal of effort to assisting, and at times working with, the SDC team in preparing functional requirements, implementation planning, and system installation. This close cooperation ensured that the programmers building the system received all needed guidance without undue delay. As one mark of achievement for this cooperative effort, the functional requirements, once published, stood throughout the implementation phase virtually without change.

Another factor, at least for the initial stage of the project, was NLM's employment of a consultant (from the RAND Corporation) to work with the SDC team and with NLM to ensure effective interorganization coordination. This consultant coordinated the efforts of both parties to the contract and helped in establishing a close working relationship. He assured the NLM staff that the design approach taken by SDC was viable and effective, while at the same time assisting the NLM staff in establishing their system requirements and developing firm control over the project.

Using the proven ORBIT® retrieval system (in its ELHILL II incarnation) as a major subsystem of MEDLARS II contributed greatly to the development of the total system in a relatively short time and with a relatively low level of effort, since the knottiest problems of file design, user interface, and retrieval mechanisms were presolved. In fact, embedding the retrieval package in a system which included data input, file maintenance, and data output functions revealed some unsuspected capabilities for data management in a bibliographic citation-retrieval system such as ORBIT. It may be that library automation software has more data management applicability than is generally recognized. The emphasis in library automation on providing a human-oriented retrieval interface certainly should receive more attention on the part of data management system designers.

Perhaps the most basic factor for the success of the project was the fact that development of the system was wholly result-oriented. Considerations that would detract from getting the job done were not allowed to impact production. Ideas for additions and elaborations to system capabilities were usually compiled and saved for implementation after the initial system was installed and operating successfully. System production difficulties were identified and dealt with expeditiously by mutual agreement between NLM and SDC. There was agreement on all sides as to what was necessary to contribute to get MEDLARS II operational. And finally, in the fall of 1972 the NLM and SDC project leaders jointly lit a candle to the success of the system in the Cathedral at Reims.

ACKNOWLEDGMENTS

The MEDLARS II Development Project was chartered by Dr. Martin Cummings, director of NLM. The project was under the overall direction of Davis B. McCarn. The library operations work of the NLM staff was directed jointly by Dr. Joseph Leiter and his deputy, William Caldwell, and included, among many others, the efforts of Betty Sawyers, Emily Wiggins, Lillian Kozuma, Dr. Clifford Bachrach, and Dr. Norman Shumway. The computer programming design consulting and signoff work for NLM were under the direction of John Cox, with main contributions coming from Tuck Arnold, Joe Hutchins, and Dave Kenton. Don N. Cohen of RAND, operating as consultant to the director of NLM, filled the important role of maintaining close coordination between NLM and SDC.

The SDC MEDLARS II project team was managed by Dr. Robert Katter, with John Moran as technical programming director for a one-year period during the project. William Schoene was lead programmer and had responsibility for the Specification Subsystem. Programmer responsibilities were as follows: Input Subsystem: Dave Londe, with contributions from Michael Adams and Matty Park. File Generation and Maintenance: Marty Vago. Retrieval Subsystem: Robert Burket, with contributions from Ray

Baker and Donald Blankenship (consultant). Publication Structure Subsystem: Jack Newsbaum. Publication Format (photocomposition) Subsystem: Dorothy Ringer, Jeffrey Trefftz, and Ronald Case (Operating Systems, Inc.). Management Information Subsystem: Ray Baker. Library Systems Analysis work for the SDC team was done by Karl Pearson and Don Black. The original software design concept was the joint product of Robert Burket, Donald Blankenship, and Dave Kenton. Robert Patrick provided design review consultation for the project.

Questions concerning MEDLARS II should be directed to Mr. Davis B. McCarn, National Library of Medicine, 8600 Rockville Pike, Bethesda, Maryland 20014. Questions concerning ORBIT® and information system development may be directed to the authors.

The research upon which this publication is based was performed pursuant to Contract No. NIH 71-4728 with the National Library of Medicine, National Institutes of Health, Department of Health, Education, and Welfare.

REFERENCES

1. National Library of Medicine, *The MEDLARS Story of the National Library of Medicine* (Washington, D.C.: 1963); National Library of Medicine, *MEDLARS: 1963-1967* (PHS Pub. No. 1823 [Washington, D.C.: 1968]); National Library of Medicine, *The Principles of MEDLARS* (Washington, D.C.: [1970]).
2. R. V. Katter and D. B. McCarn, "AIM-TWX—An Experimental On-Line Bibliographic Retrieval System," in D. E. Walker, ed., *Interactive Bibliographic Search: The User/Computer Interface* (Montvale, N.J.: AFIPS Press, 1971), p.121-41.
3. W. Moll, "AIM-TWX Service at the University of Virginia: A Review and Evaluation," *Medical Library Association Bulletin* 59:458-62 (July 1971).
4. F. W. Lancaster, *Evaluation of On-Line Searching in MEDLARS (AIM-TWX) by Biomedical Practitioners* (Champaign, Ill.: Univ. of Illinois Graduate School of Library Science, 1972).
5. C. A. Cuadra, "On-Line Systems: Promises and Pitfalls," *Journal of the American Society for Information Science* 22:107-14 (March-April 1971).
6. D. B. McCarn and J. Leiter, "On-Line Services in Medicine and Beyond," *Science* 181:318-24 (July 27, 1973).
7. G. Anderla, *Information in 1985* (Paris: Organization for Economic Co-operation and Development, 1973).

Library Support Through Automation: The California State University and Colleges Plan for Library Automation

John C. KOUNTZ: Associate for Library Automation, Office of the Chancellor, The California State University and Colleges, Los Angeles, California.

The nineteen campuses of the California State University and Colleges (CSUC) are spread across a distance equal to that between Maine and North Carolina. The extreme distances between the individual campuses and the enormity of the student population served, one quarter million FTE, lead to a variety of problem types and levels for library automation. It is within this context that a systemwide approach using standardized equipment, procedures, and techniques for library automation is being developed and coordinated and implemented for the California State University and Colleges. As a result of over two years' planning, specifications, schedules, and budgets have evolved for a ten-year program leading to a total system of library automation for the nineteen libraries. In this article, software, hardware, procedural, and operational components of the CSUC systemwide approach to library automation are reviewed with specific emphasis on the time periods involved, the phasing of implementation, and finally, the costs and benefits anticipated from this project.

INTRODUCTION

The approach to automated support for the libraries of the California State University and Colleges centers on resource sharing. Classically libraries have recognized and attempted to enhance their ability to provide the breadth and depth of materials required by their clients through a variety of formal and informal resource sharing arrangements. Typical of such arrangements are the American Library Association's Interlibrary Loan Procedures, reciprocal borrower privileges, the establishment of pools or banks of specialized materials which can be used mutually by several libraries, and the development of special collections by subject area or material type. Many of these and similar approaches reflect the library's acknowledgment that each library cannot purchase, process, and provide to its clientele all the appropriate materials available in the marketplace.

More recently, budget constraints, inflation, and the increasing number of available titles have further reduced the share of published materials the individual library can and should procure and provide to its users.

Thus, fully applying the practice and techniques of resource sharing can be hypothesized as permitting the maintenance, at a minimum, of existing levels of user service for a fixed budget. While successful resource sharing will dramatically improve user services, such success is contingent upon the integration of policy objectives, a systemwide recognition of the benefits to be derived, and the intelligent application of automation.

The areas defining resource sharing are shown in Figure 1. Note that policy, organizational, and procedural areas differ from the area supported by automation. Note, too, that while automation supports resource sharing, it cannot of itself effect resource sharing. The intent of automation in the libraries, then, is to support resource sharing which in turn is anticipated to avoid additional costs and to improve user service.

SYSTEMWIDE ACCESS

SYSTEM INTEGRATION

STRUCTURED COLLECTION BUILDING

SYSTEMWIDE RESOURCE VISIBILITY AND CONTROL;
THE COMPONENT SUPPORTED BY AUTOMATION

Fig. 1. Library Resource Sharing.

That automation can provide this support feasibly, effectively, and economically for specific functions has been amply demonstrated by numerous libraries across the United States and is clearly documented in the recent *Survey of Library Automation Systems* performed for CSUC by Inforonics, Inc.[1] The survey also showed that in library automation unconstrained development leads invariably to unstructured systems. For these reasons, and to preclude an untoward risk, CSUC has selected the inventory or, in library terms, the circulation control and interlibrary loan function for its library automation objective. Of equal importance is the fact that this specific function was early recognized and presented by CSUC to the National Commission on Libraries and Information Science in November 1972 and has not changed significantly since that time in spite of delays caused by procedural difficulties.[2]

However, before entering into technical considerations relating to the automation of specific activities, let us review the circulation control and interlibrary loan functions as they relate to resource sharing: first, as these functions are currently performed, and then, to set the automation of these specific functions in focus, how computer processing can support and facilitate circulation control and interlibrary loan. Finally, all the subprograms prerequisite to automating the circulation control and interlibrary loan function will be described and presented in terms of their status, schedule, and budget and their relation to the overall project.

CURRENT RESOURCE SHARING

With the exception of the libraries at CSU Northridge, Fullerton, Los Angeles, and Sacramento where machine-readable files exist, each CSUC library presently maintains a separate manual local inventory file (their local shelflist and in-circulation files). Because these files are large and bulky, they do not physically lend themselves to easy replication or transportation, therefore access to the information in them concerning the existence, availability, and disposition of specific items in a particular library's collection is only possible at that library. In short, as shown in Figure 2, current sharing procedures require that to find out if a certain book is available, library A must interrogate each library (B, C, D, etc.) sequentially until the book is either found and released to library A, or it is determined that no library in the system has the requested book. At that time library A may attempt to procure the requested item. The current proce-

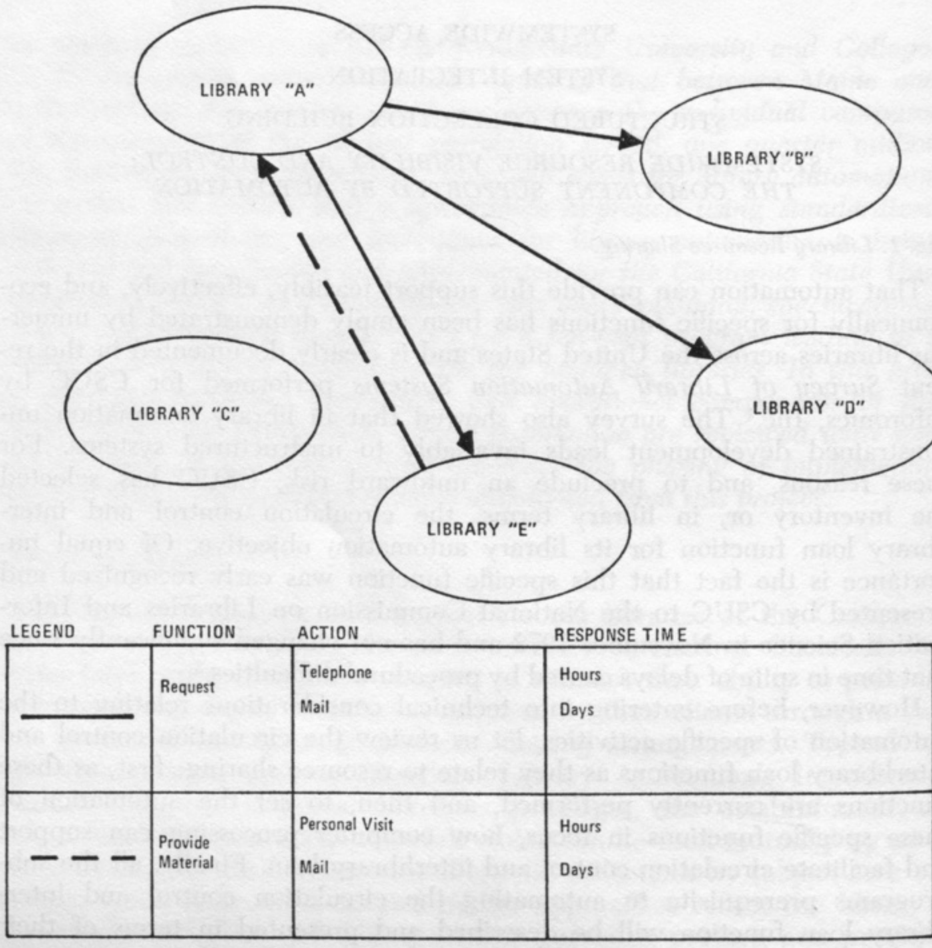


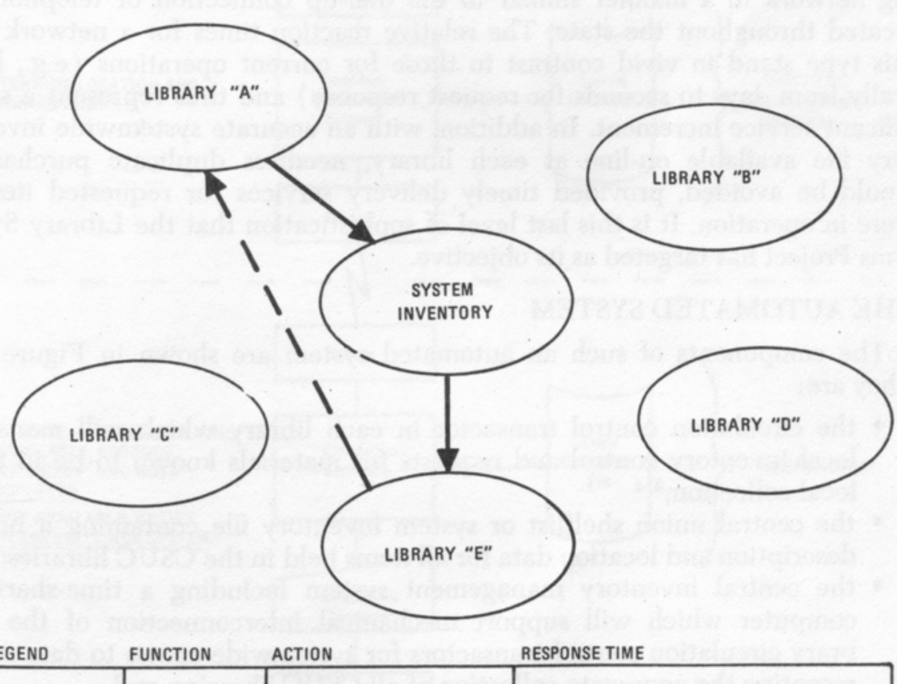
Fig. 2. Library Resource Sharing—Current Operations.

dure is labor intensive, time consuming, and subject to error because it is dependent on manual search of manually maintained files. As a result, the typical library reaction to this complicated operation has been to buy "all" items at the outset and thus avoid the loan procedure if possible. However, as we are all aware, this mode of operation is no longer economically realistic or viable.

AUTOMATED SUPPORT FOR RESOURCE SHARING

The key to resource sharing and the foundation upon which all subsequent library automation activities will depend is a system inventory or union shelflist master file for the entire system (Figure 3). This centrally processed and maintained master file will contain a brief description and the location of every book in the CSUC libraries.

At the most rudimentary level this master file would permit the timely



LEGEND	FUNCTION	ACTION	RESPONSE TIME
———	Request	Mechanically scan system inventory; local circulation, and; create pick-ticket.	Seconds
— — —	Provide material	Courier	Hours

Fig. 3. Library Resource Sharing—Computer Supported Operations.

publication of indexes and cumulative supplements representing all items held by the CSUC libraries. Such indexes or finding lists would be distributed to each library and would literally allow a librarian at any site in the CSUC to determine if a specific book were held by the system and which library had it. Subsequent to such a determination, the librarian would have to use traditional communications to complete the circulation or interlibrary loan process.

At a moderately sophisticated level, this master file could be linked mechanically to cathode ray tube (CRT) terminals in the libraries to allow computer assisted search. As in the first instance, subsequent processing by the libraries would follow traditional procedures.

At a still more sophisticated level, intelligent terminals capable of handling the local in-circulation processing could be connected together through the system master file by means of a centrally maintained switching network in a manner similar to the dial-up connection of telephones located throughout the state. The relative reaction times for a network of this type stand in vivid contrast to those for current operations (e.g., literally from days to seconds for request response) and thus represent a significant service increment. In addition, with an accurate systemwide inventory file available on-line at each library, needless duplicate purchases would be avoided, provided timely delivery services for requested items were in operation. It is this last level of sophistication that the Library Systems Project has targeted as its objective.

THE AUTOMATED SYSTEM

The components of such an automated system are shown in Figure 4. They are:

- the circulation control transactor in each library which will manage local inventory control and requests for materials known to be in the local collection;[3, 4]
- the central union shelflist or system inventory file containing a brief description and location data for all items held in the CSUC libraries;
- the central inventory management system including a time-sharing computer which will support mechanical interconnection of the library circulation control transactors for systemwide access to data representing the aggregate collection of all CSUC libraries; and
- the communications facilities over which availability messages will be transmitted.

The status of these components at this point in time is:

- Circulation control transactors which due to procedural difficulties in procurement are now the object of competitive bid for the fourth time.
- Central union shelflist or system inventory file is being developed in phases. Conversion has been completed for periodical titles and items

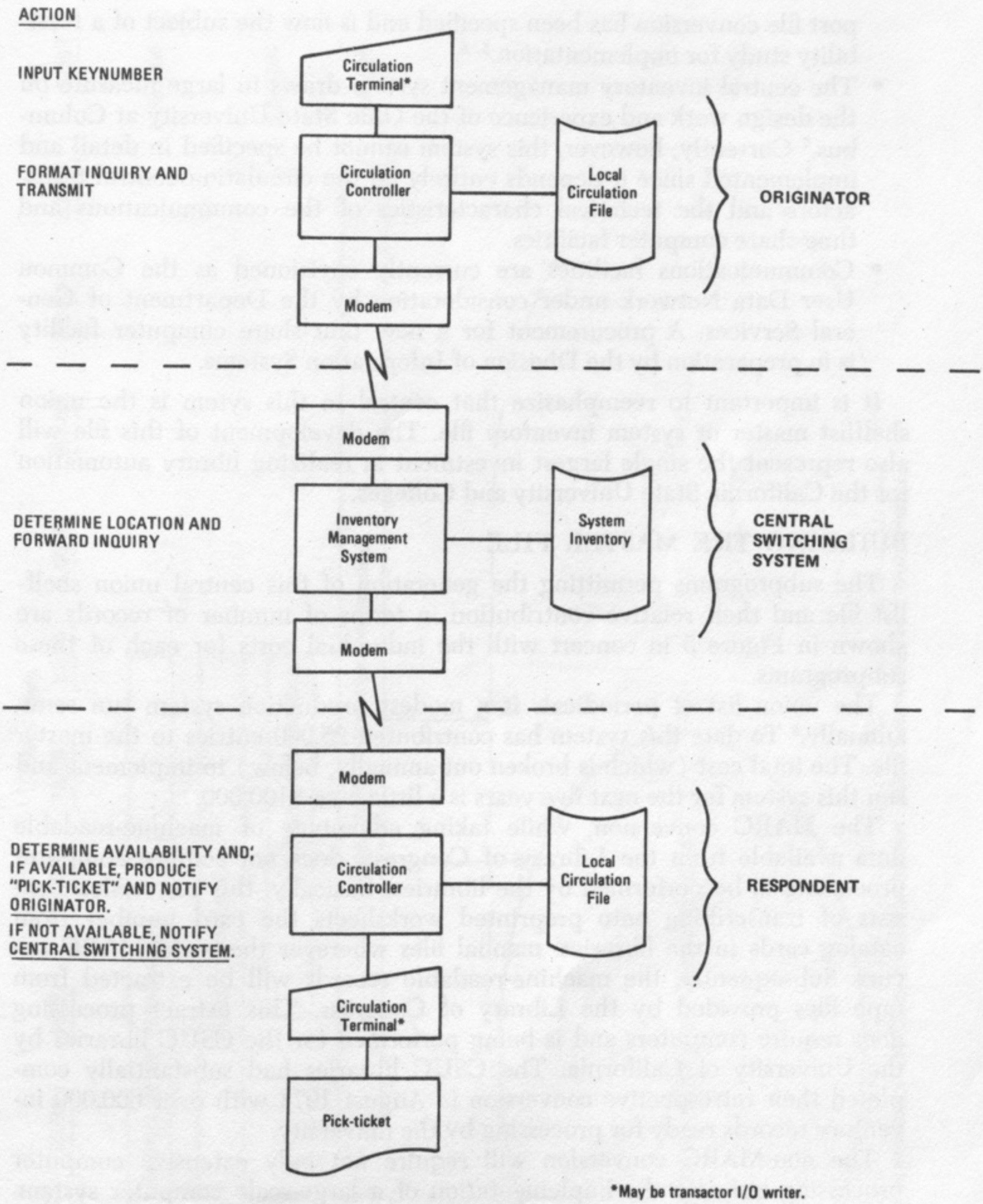


Fig. 4. Automated System Components.

for which Library of Congress MARC (machine-readable) records are available. For the other materials held by the CSUC libraries, a file conversion technique has been designed, and a computer system to support file conversion has been specified and is now the subject of a feasibility study for implementation.[5, 6]

- The central inventory management system draws in large measure on the design work and experience of the Ohio State University at Columbus.[7] Currently, however, this system cannot be specified in detail and implemented since it depends entirely on the circulation control transactors and the technical characteristics of the communications and time-share computer facilities.
- Communications facilities are currently envisioned as the Common User Data Network under consideration by the Department of General Services. A procurement for a new time-share computer facility is in preparation by the Division of Information Systems.

It is important to reemphasize that central to this sytem is the union shelflist master or system inventory file. The development of this file will also represent the single largest investment in realizing library automation for the California State University and Colleges.

BUILDING THE MASTER FILE

The subprograms permitting the generation of this central union shelflist file and their relative contribution in terms of number of records are shown in Figure 5 in concert with the individual costs for each of these subprograms.

The union list of periodicals is a modest production system run semiannually.[8] To date this system has contributed 25,000 entries to the master file. The total cost (which is broken out annually, below) to implement and run this system for the next five years is a little over $100,000.

The MARC conversion, while taking advantage of machine-readable data available from the Library of Congress, does not require computer processing to be performed by the libraries. Basically, this conversion consists of transcribing onto preprinted worksheets the card number from catalog cards in the libraries' manual files wherever the word MARC occurs. Subsequently, the machine-readable records will be extracted from tape files provided by the Library of Congress. This extract processing *does* require computers and is being performed for the CSUC libraries by the University of California. The CSUC libraries had substantially completed their retrospective conversion in August 1974 with over 600,000 inventory records ready for processing by the university.

The non-MARC conversion will require not only extensive computer processing but also the implementation of a large-scale computer system. This system was functionally specified by Library Development and Services in January 1974. As mentioned earlier, the implementation and operation of this system is the subject of a feasibility study now being performed by the CSUC Division of Information Systems. The existing machine-readable files from the four campuses identified earlier will be used by this system in concert with data from the MARC conversion to form

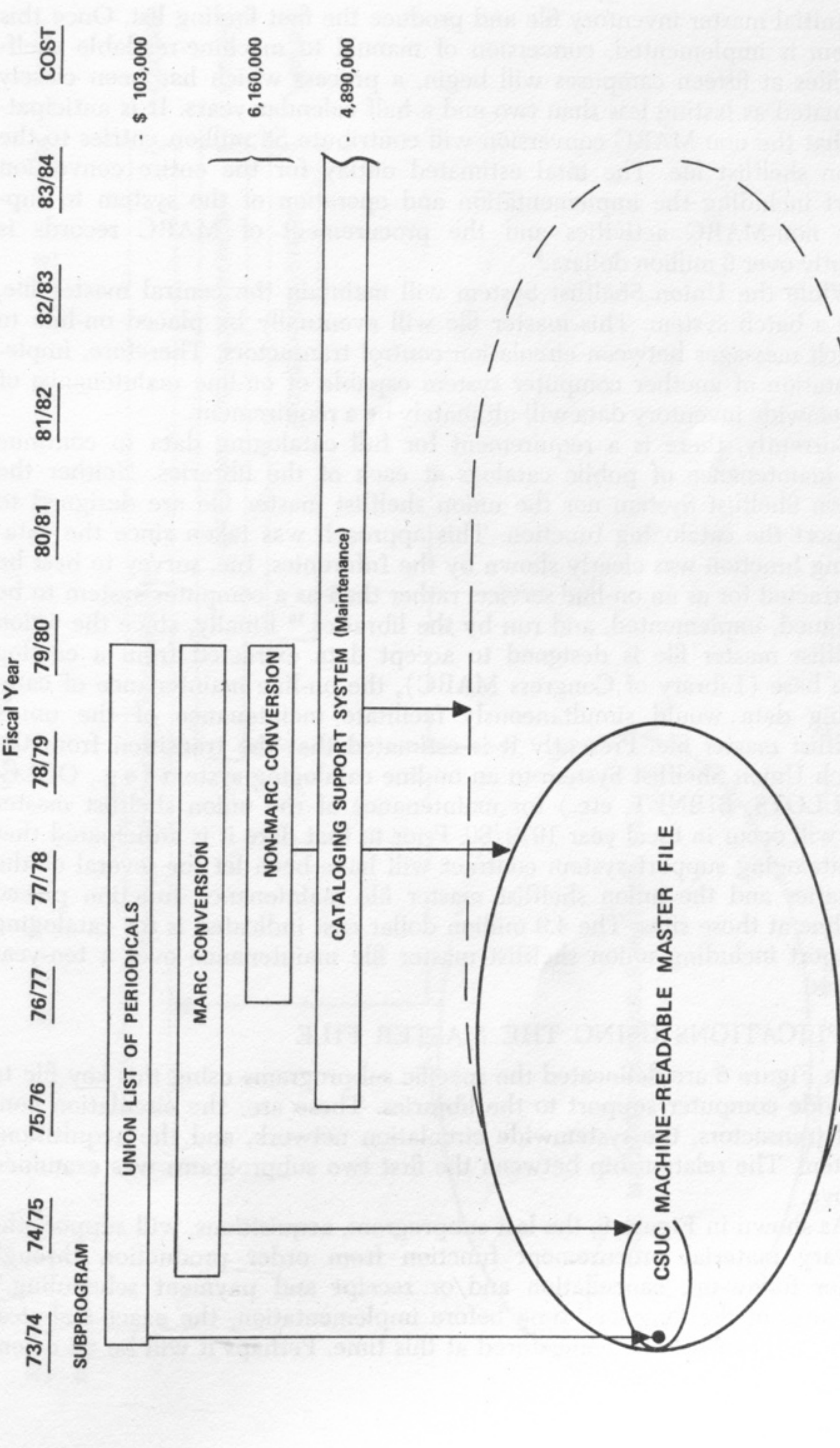


Fig. 5. Specific Subprograms for Machine-Readable Master File Construction and Maintenance.

the initial master inventory file and produce the first finding list. Once this system is implemented, conversion of manual to machine-readable shelflist files at fifteen campuses will begin, a process which has been closely estimated as lasting less than two and a half calendar years. It is anticipated that the non-MARC conversion will contribute 5½ million entries to the union shelflist file. The total estimated outlay for the entire conversion effort including the implementation and operation of the system to support non-MARC activities and the procurement of MARC records is slightly over 6 million dollars.[9]

While the Union Shelflist System will maintain the central master file, it is a batch system. This master file will eventually be placed on-line to switch messages between circulation control transactors. Therefore, implementation of another computer system capable of on-line maintenance of systemwide inventory data will ultimately be a requirement.

Currently, there is a requirement for full cataloging data to continue the maintenance of public catalogs at each of the libraries. Neither the Union Shelflist System nor the union shelflist master file are designed to support the cataloging function. This approach was taken since the cataloging function was clearly shown by the Inforonics, Inc. survey to best be contracted for as an on-line service, rather than as a computer system to be designed, implemented, and run by the libraries.[10] Finally, since the union shelflist master file is designed to accept data extracted from a catalog data base (Library of Congress MARC), the on-line maintenance of cataloging data would simultaneously facilitate maintenance of the union shelflist master file. Presently it is estimated that the transition from the batch Union Shelflist System to an on-line cataloging system (e.g., OCLC, BALLOTS, BIBNET, etc.) for maintenance of the union shelflist master file will occur in fiscal year 1979/80. Prior to that date it is anticipated that a cataloging support system contract will have been let for several of the libraries and the union shelflist master file maintenance function placed on-line at those sites. The 4.9 million dollar cost indicated is for cataloging support including union shelflist master file maintenance over a ten-year period.

APPLICATIONS USING THE MASTER FILE

In Figure 6 are delineated the specific subprograms using this key file to provide computer support to the libraries. These are: the circulation control transactors, the systemwide circulation network, and the acquisitions system. The relationship between the first two subprograms was examined above.

As shown in Figure 6, the last subprogram, acquisitions, will support the library material procurement function from order production through order follow-up, cancellation and/or receipt and payment scheduling.[11] Because of the long lead-time before implementation, the exact technical approach can only be conjectured at this time. Perhaps it will be an exten-

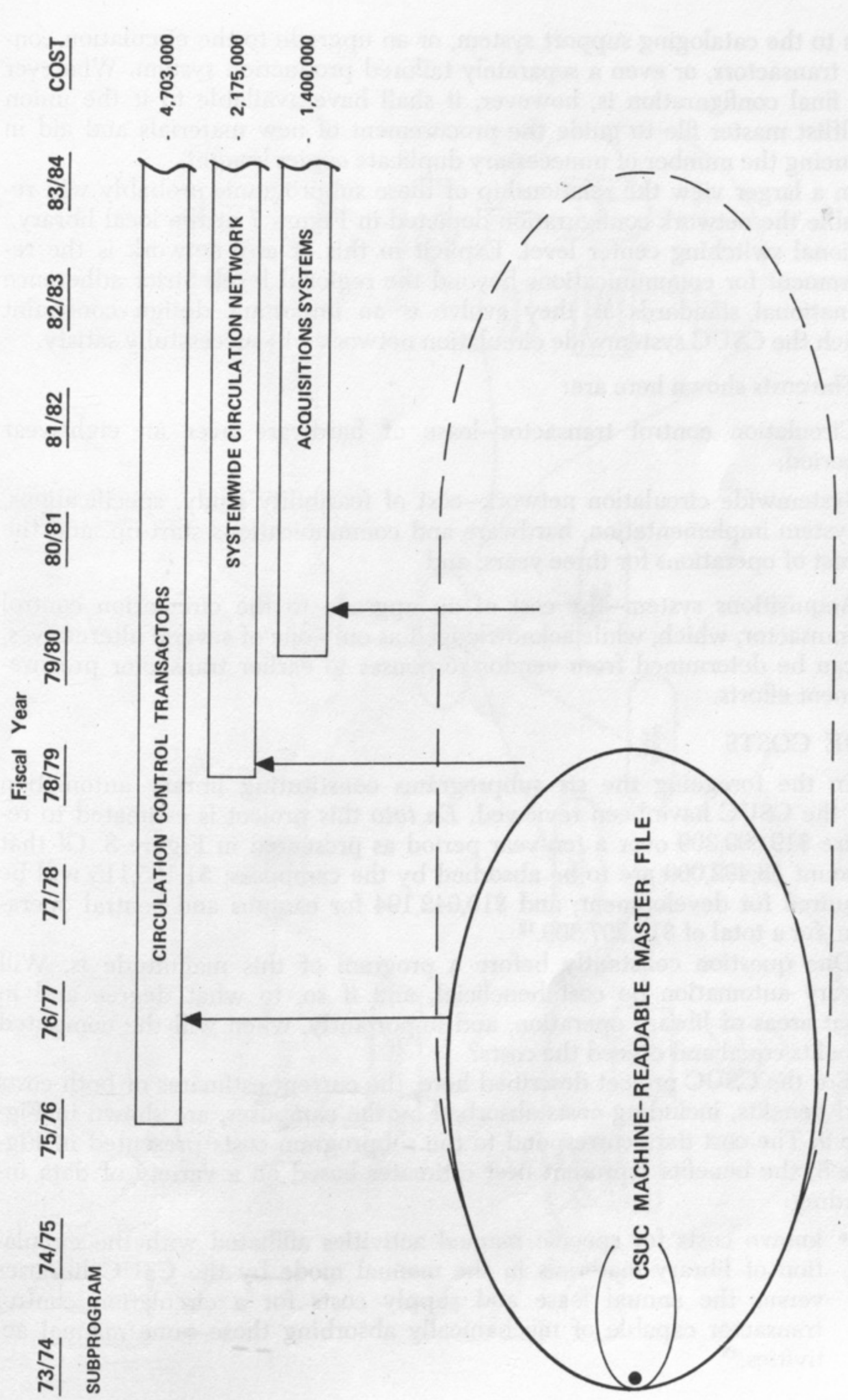


Fig. 6. Machine-Readable Master File Use by Specific Subprograms.

sion to the cataloging support system, or an upgrade to the circulation control transactors, or even a separately tailored production system. Whatever the final configuration is, however, it shall have available to it the union shelflist master file to guide the procurement of new materials and aid in reducing the number of unnecessary duplicate copies bought.

In a larger view the relationship of these subprograms probably will resemble the network configuration depicted in Figure 7 at the local library/regional switching center level. Explicit in this or any network is the requirement for communications beyond the regional level. Strict adherence to national standards as they evolve is an important design constraint which the CSUC systemwide circulation network will successfully satisfy.

The costs shown here are:

Circulation control transactor—lease of hardware over an eight-year period;

Systemwide circulation network—cost of feasibility study, specifications, system implementation, hardware and communications start-up, and the cost of operations for three years; and

Acquisitions system—the cost of an upgrade to the circulation control transactor, which, while acknowledged as only one of several alternatives, can be determined from vendor responses to earlier transactor procurement efforts.

THE COSTS

In the foregoing the six subprograms constituting library automation for the CSUC have been reviewed. *En toto* this project is estimated to require $19,699,309 over a ten-year period as presented in Figure 8. Of that amount $8,492,000 are to be absorbed by the campuses; $1,165,115 will be required for development; and $10,042,194 for campus and central operation, for a total of $11,207,309.[12]

One question constantly before a program of this magnitude is, Will library automation be cost beneficial, and if so, to what degree and in what areas of library operation, and importantly, when will the computed benefits equal and exceed the costs?

For the CSUC project described here, the current estimates of both costs and benefits, including costs absorbed by the campuses, are shown in Figure 9. The cost data correspond to the subprogram costs presented in Figure 8; the benefits represent best estimates based on a variety of data including:

- known costs for specific manual activities affiliated with the circulation of library materials in the manual mode by the CSUC libraries versus the annual lease and supply costs for a circulation control transactor capable of mechanically absorbing those same manual activities;[13]

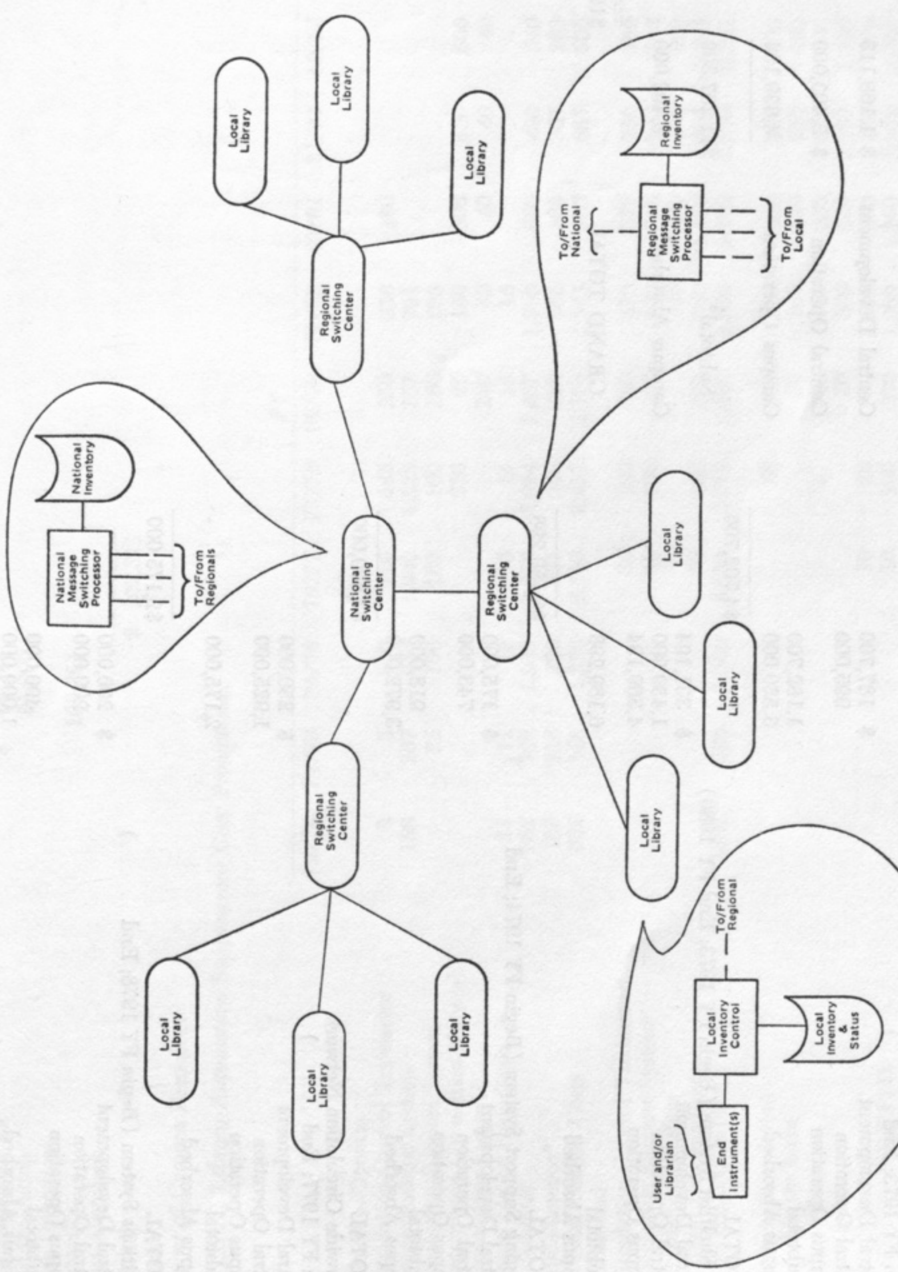


Fig. 7. Network Configuration.

Program	Program Cost Totals (Where end dates are blank, costs are projected through FY 1983/84)	
A. Circulation Control Transactor (Begin FY 1973; End ......)		
Central Development	$ 187,700	
Central Operation	995,000	
Campus Operation		
Subtotal	1,182,700	
Campus Absorbed	3,520,000	
TOTAL		$4,702,700
B. Union Shelflist System (Begin FY 1973; End FY 1980)		
Central Development	$ 331,104	
Central Operation	1,459,000	
Campus Operation	4,369,194	
Subtotal	6,159,298	
Campus Absorbed		
TOTAL		$6,159,298
C. Cataloging Support System (Begin FY 1974; End ......)		
Central Development	$ 175,000	
Central Operation	743,000	
Campus Operation		
Subtotal	918,000	
Campus Absorbed	3,972,000	
TOTAL		$4,890,000
D. Systemwide Circulation Network (Begin FY 1977; End ......)		
Central Development	$ 250,000	
Central Operation	1,925,000	
Campus Operation		
Subtotal	2,175,000	
Campus Absorbed		
TOTAL		$2,175,000
E. Acquisitions System (Begin FY 1978; End ......)		
Central Development	$ 200,000	
Central Operation	200,000	
Campus Operation		
Subtotal	400,000	
Campus Absorbed	1,000,000	
TOTAL		$1,400,000
F. Union List of Periodicals (Begin FY 1973; End FY 1978)		
Central Development	$ 21,311	
Central Operation	81,000	
Campus Operation		
Subtotal	102,311	
Campus Absorbed		
TOTAL		$ 102,311

Aggregate Costs By Cost Center		
Central Development	$ 1,165,115	
Central Operation	$ 5,403,000	
Campus Operation	4,639,194	
Subtotal*	$11,207,309	
Campus Absorbed†	8,492,000	
GRAND TOTAL		$19,699,309

* General Fund
† Redeployment of Campus Resources

Fig. 8. The CSUC Library Automation Subprogram Cost Projections.

	FY 73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84
COSTS											
Central Develop/Operate											
1. Circulation Control Transactor	3	5	150	120	145	250	330	180			
2. Union Shelflist System	158	267	724	1,940	1,525	757	791				
3. Cataloging Support System		25	150	160	160	160	263				
4. Systemwide Circulation Network					250	25	150	300	450	500	500
5. Acquisitions System						200	40	40	40	40	40
6. Union List of Periodicals	21	11	12	13	14	15	16				
7. *Subtotal*	*182*	*308*	*1,036*	*2,233*	*2,094*	*1,407*	*1,590*	*520*	*490*	*540*	*540*
8. Miscellaneous*	180	218	220	307	358	369	355	366	367	368	369
9. Total Central Costs	362	526	1,256	2,540	2,452	1,776	1,945	886	857	908	909
Campus Absorbed											
10. Circulation Control Transactor†				30	100	200	370	570	750	750	750
11. Cataloging Support System†				40	160	320	480	743	743	743	743
12. Acquisitions System							50	100	200	300	350
13. Total Campus Costs				70	260	520	900	1,413	1,693	1,793	1,843
14. TOTAL COSTS	362	526	1,306	2,610	2,712	2,296	2,829	2,299	2,550	2,701	2,752
BENEFITS											
15. Campus Circulation					50	100	250	400	500	500	500
16. Interlibrary Loan						50	100	250	250	250	250
17. Acquisitions						50	150	350	550	600	650
18. Processing						50	300	500	1,100	1,100	1,100
19. Cataloging				10	50	100	200	300	400	400	400
20. TOTAL BENEFITS				10	100	350	1,000	1,800	2,800	2,850	2,900

* Includes Couriers, Administrative Burden, OE, and for FY 1974/75, $68,000 for SFSU Reclassification
† Automated Equipment or Services

Fig. 9. Annual Cost/Benefit Detail for FY 1973 through FY 1983 (In Thousands of Dollars).

- known costs of processing new materials based on the CSUC Technical Processing Cost Study versus estimated costs of performing specific manual operations by computer including automated on-line file search and file maintenance to facilitate cataloging;[14, 15]
- known costs to execute interlibrary loan requests based on an interlibrary loan study recently completed for the libraries versus the estimated costs of performing the same function using a computer based network;[16] and
- costs of specific manual functions to acquire materials versus the annual costs of an upgrade to the circulation control transactor to absorb those specific functions and the provision of systemwide collection visibility to reduce the number of items procured which may be little used.[17]

It was extremely important to understand that automation of libraries *improves* service to patrons and that the improvement results in greater use being made of the collections. In other words, as experience elsewhere has amply indicated, because of the efficiencies provided through automation, even greater usage of the collections and demand by patrons will result. However, most improvement is measurable. Two typical examples of such measurement are:

- the reduction of time to get materials to users, which could be measured as: time from patron's request for a specific item to his receipt of the item; and
- the support of statewide visibility of library resources which could be measured as: an increase in the percentage of materials at each library about which location and availability can be made known to all the system's patrons.

Increased collection usage is a workload factor recognized by the staffing formulas of the CSUC Board of Trustees library development program and can be predicted based on experience in automated libraries across the country. The dollar estimates of benefits, however, as shown in Figure 9, are conservative, taking into account only workload estimates based on the latest FTE projections and labor rates for fiscal year 1972/73.

These cost and benefit projections are presented in another form in Figure 10. Barring further delays in the CSUC program, the benefit curve is expected to cross the cost line in fiscal year 1980/81 depending, of course, on the future of clerical and professional salaries and fringe benefits. It also must be recognized and acknowledged that the projected benefits are available only after all related automated programs are fully operational.

Finally, it is emphasized that this brief overview and especially the costs and benefits represented by the project are based solely on a CSUC developed and operated program. Thus, although staff of the CSUC meet periodically with their counterparts in the University of California and the

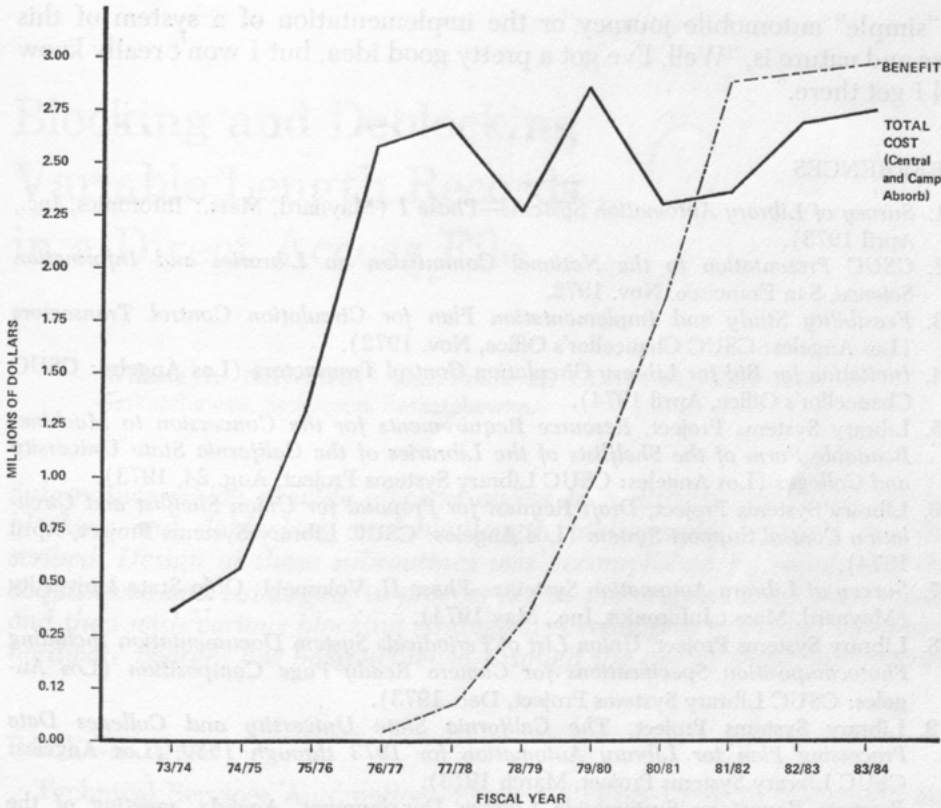


Fig. 10. Library Automation Cost/Benefit Curves Representing FY 1973/74—FY 1983/84 (In Millions of Dollars).

California State Library on cooperation matters, and while the CSUC is fully supportive of significant and meaningful efforts to cooperate with the other segments in library matters, it would be impossible to project costs and benefits based on the potential implied in envisioned joint efforts with other state entities.

CONCLUSION

Clearly, at the outset of a large program occupying a time-frame of almost ten years, there are incredible difficulties in accurately determining all costs or projecting all benefits. Of equal importance is the impact of such a system upon existing library operations and procedures; impact potentially extending unto the tradition which is the library itself. Finally, there is the world of supporting technology which is by no means static and cannot be overlooked.

It is much like the old story about the time it takes to drive a car from Los Angeles to New York City. There are so many variables that it would be almost impossible to carefully assess and summarize each and every potential variable. Thus, the best one can say with regard to either a "simple" automobile journey or the implementation of a system of this size and nature is, "Well, I've got a pretty good idea, but I won't really know 'til I get there."

REFERENCES

1. *Survey of Library Automation Systems—Phase I* (Maynard, Mass.: Inforonics, Inc., April 1973).
2. *CSUC Presentation to the National Commission on Libraries and Information Science*, San Francisco, Nov. 1972.
3. *Feasibility Study and Implementation Plan for Circulation Control Transactors* (Los Angeles: CSUC Chancellor's Office, Nov. 1972).
4. *Invitation for Bid for Library Circulation Control Transactors* (Los Angeles: CSUC Chancellor's Office, April 1974).
5. Library Systems Project, *Resource Requirements for the Conversion to Machine-Readable Form of the Shelflists of the Libraries of the California State University and Colleges* (Los Angeles: CSUC Library Systems Project, Aug. 24, 1973).
6. Library Systems Project, *Draft Request for Proposal for Union Shelflist and Circulation Control Support System* (Los Angeles: CSUC Library Systems Project, April 1974).
7. *Survey of Library Automation Systems—Phase II,* Volume 1, Ohio State University (Maynard, Mass.: Inforonics, Inc., May 1974).
8. Library Systems Project, *Union List of Periodicals System Documentation Including Photocomposition Specifications for Camera Ready Page Composition* (Los Angeles: CSUC Library Systems Project, Dec. 1973).
9. Library Systems Project, *The California State University and Colleges Data Processing Plan for Library Automation for 1973 through 1980* (Los Angeles: CSUC Library Systems Project, March 1973).
   Progress Report on Systemwide Library Development, Agenda, meeting of the Board of Trustees of the California State University and Colleges (Item 6) May 23, 1973.
10. *Survey of Library Automation Systems—Phase II,* Volumes 4 (University of Massachusetts), 5 (Ohio College Library Center), 2 (Stanford University—BALLOTS-SPIRES), and 6 (Syracuse University), (Maynard, Mass.: Inforonics, Inc., May 1974).
11. *Position Paper on Acquisitions,* Application Standard Sub-Committee of the Library Technical Advisory Committee, Los Angeles, Jan. 1974.
12. *Library Automation Plan for the California State University and Colleges; Sub-Programs, Including Resource Requirements and Schedule: 1973/74–1983/84* (Los Angeles, CSUC Library Development and Services—Library Systems Project, Revised Sept. 1974).
13. *Final Report on the Time/Methods Study Project for Library Circulation Control* (Sacramento, Calif.: California State University, March 15, 1974).
14. *Report on a Cost Study of Specific Technical Processing Activities of the California State University and Colleges Libraries* (Los Angeles: CSUC Chancellor's Office, Feb. 1973). (ED 073779).
15. Dan Du Bois, *Library Labor Cost Accounting System* (Los Angeles: CSUC Chancellor's Office, 1972). (ED 075063).
16. *Report on an Interlibrary Request and Loan Study—1973* (Los Angeles: CSUC Chancellors' Office, 1974).
17. *Library Cooperation: A Systems Approach to Interinstitutional Resource Utilization.* Program Review Branch, Audits Division. California State Department of Finance. June 1973.

Blocking and Deblocking Variable Length Records in a Direct Access File

William L. NEWMAN* and Frank H. COXFORD: University of Saskatchewan, Saskatoon, Saskatchewan.

Subroutines which provide efficient access to a directly organized file by blocking and deblocking variable length bibliographic records are described. Design of these subroutines was accomplished by simulating the acquisitions and cataloging in-process file in the original unblocked mode, and then with various blocking strategies. A reduction in overall computer hardware charges has been achieved through the use of these subroutines.

BACKGROUND

Technical Services Automation–Phase I (TESA-1) is an integrated acquisitions and cataloging system at the library of the University of Saskatchewan, Saskatoon, which has been operational since November 1970.[1-4] Since the system provides on-line inquiry and efficient duplicate order checking, the in-process file is organized using IBM's basic direct access method (OS/BDAM). The file retains a stable size of approximately 20,000 logical records, each consisting of fixed length fields, a record directory, and variable length bibliographic fields.

When TESA-1 was in the design stage, a study was done on the lengths of MARC records. The results indicated that after the MARC record length was modified to reflect the corresponding TESA-1 record format, the average record length is 460 characters, and 89 percent of the records would contain 600 characters or less. The modification of the MARC record length consisted of subtracting the lengths of fields not used in TESA-1 (only fields needed for local processing requirements are transferred from MARC to TESA-1), subtracting 10 characters per record directory entry (each TESA-1 directory entry occupies two bytes), and adjusting the fixed field length to that of the TESA-1 format (154 characters). The study also indicated that the system must have the capability of handling records with lengths exceeding 2,000 characters.

* Now at the National Library of Canada, Ottawa, Ontario.

Two direct access methods are available under IBM's Operating System. In COBOL, only fixed length records may be used with the basic indexed sequential access method (BISAM).[5] BDAM supports the use of variable length records; however, records cannot be blocked in order to conserve disc space, nor, more importantly, can record lengths be changed during update operations.[6] Of course, the ability to change the length of a record was required in TESA-1. For example, the record length is usually increased at the cataloging stage when a MARC record is located and transferred to the in-process file, or when original cataloging is performed.

The BDAM file structure implemented in the original TESA-1 system incorporated the capability of having as many 600-character physical records as were necessary to store a complete logical record. TESA-1 could, therefore, handle bibliographic records of any length. Note that when randomly retrieving or storing a logical record, only one I/O operation was required in 89 percent of the cases.

The in-process file was designed with a capacity of 20,000 logical records at an 80 percent file packing density. This meant that with each logical record requiring on the average 1.15 physical locations, and with nine physical locations available per track on an IBM 2314 disc pack, the space required for the in-process file was 3,190 tracks, or room for 28,710 physical records. That is, space required in tracks is approximately

$$\frac{100}{80} \times \frac{1.15 \times 20{,}000}{9}$$

DISC STORAGE—A PROBLEM

Each item in TESA-1 is identified by an item number, the first six characters of which are numeric, the seventh an alphabetic check digit, and the eighth a numeric supplement digit. The key-to-address conversion algorithm employed originally in TESA-1 is called the prime modulo technique.[7] The first six characters of the item number, the supplement digit, and the physical record number (the first physical record of a logical record has physical record number one, the second two, etc.) were concatenated to form an eight-digit number which was input to the prime modulo algorithm. As reported by Mitchell and Burgess, this technique tends to build long strings of contiguously occupied physical record locations.[8] When the in-process file approached the 20,000 logical record mark in December 1971, reorganizations were being required twice monthly because the limit of twenty-three tracks of contiguously occupied locations was being exceeded frequently. Additional disc storage space was not available. Under the University of Saskatchewan charging policy, disc I/Os cost $1.25 per 1,000 I/Os. A reorganization involves placing the in-process file on a back-up unit, then removing completely processed items during replacement. Since approximately 30,000 physical locations existed in the in-process file, the charge for disc I/Os alone in the reorganization was over $75 and the total cost of the run over $100.

In order to provide relief from the constant file reorganizations, the random number key-to-address technique suggested by Mitchell and Burgess was then implemented. Satisfactory results were obtained for approximately six months; then, in the summer of 1972, due to increased ordering from a special budget allocation and decreased cataloging because of vacations, the in-process file approached 22,000 logical records. Again, frequent reorganizations were required (at least weekly). As a temporary solution, the size of the file was increased and the number of tracks searched for an empty location for an addition was extended from twenty-three to thirty.

TOWARD SOLUTION OF THE STORAGE PROBLEM

Even though the number of logical records was later reduced to under 20,000, research toward the development of a more efficient and economical means of directly storing and retrieving in-process records was started. It was found that the most likely means of improvement would be the construction of subroutines which would perform all the in-process file I/O, and would block and deblock in-process file records. The main problem in designing such subroutines is the determination of how many tracks to search for sufficient space when adding a record, or for the record itself when changing a record. If BDAM is not extended by the user, records cannot be blocked. The format of user records on a track which can be accessed by BDAM is shown in Figure 1.

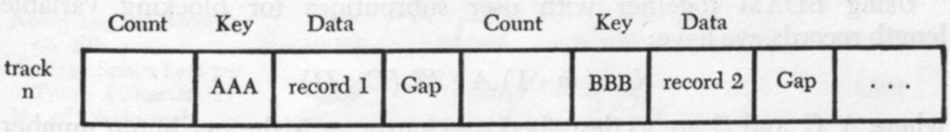


Fig. 1. Formatted Records on an IBM Disc Pack.

If record 2 in Figure 1 were to be updated, the user program would supply the track number relative to the beginning of the file, that is, *n*, and the record key, that is, BBB, to BDAM. The disc hardware searches a user specified number of tracks, starting at track *n*, for the record with key BBB. Whether or not the record is located and regardless of the number of tracks searched, the user is charged for only one I/O. By blocking records, the user loses this feature, so that an I/O charge occurs for each track searched. However, with proper design, the user can employ variable length records, change their lengths, and, if the block is set to the maximum allowable for a particular device, the overhead space required for each physical record is saved. This overhead is considerable; for instance, on an IBM 2314 disc pack, 146 character positions are required for the gap and count areas between each physical record on the 7,294-character track, and on an IBM 3330 disc pack, 191 character positions of overhead are required between each physical record on the 13,030-character track.

I/O CALCULATIONS

The number of in-process file I/Os per month in TESA-1 may be computed as follows:

$$N = N_1 + N_2 + N_3$$

where N is the total number of I/Os per month; N_1 is the total number of I/Os occurring in the update program; N_2 the number of I/Os incurred in reading the in-process file sequentially (accounts, claims, and reorganization programs); and N_3 is the number of I/Os incurred in reading the in-process file directly in all other programs.

Using BDAM and the original TESA-1 file structure:

$$N_1 = 1.15\,[2\,(A + C + D)]$$

where 1.15 is the average number of physical records read or written per logical record; A is the number of additions to the in-process file; C the number of changes; and D the number of deletions.

$$N_2 = 28{,}710 n_2$$

where 28,710 is the number of physical records on the in-process file, and n_2 is the number of times per month that the in-process file is read sequentially.

$$N_3 = 1.15 n_3$$

where n_3 is the number of times that a logical record is accessed per month.

Using BDAM together with user subroutines for blocking variable length records, we have:

$$N_1 = (m + 1)\,A + 2X\,(C + D)$$

where A, C, and D are as described previously, m is the maximum number of tracks which will be searched to verify that the record does not already exist when making an addition, and X is the average number of tracks searched to locate a particular record.

$$N_2 = Y n_2$$

where Y is the number of physical records on the in-process file, and n_2 is as described previously.

$$N_3 = X n_3$$

where X and n_3 are as previously described.

IN-PROCESS FILE I/O SIMULATION

A program was written to simulate the in-process file with blocked records. Variables in the simulation were the number of tracks allocated to the file, and the number of tracks specified to search for a record (changes) or for space (additions). The latter variable is called the *restricted search parameter*. The average number of transactions applied to the in-process file in a week, the probability distribution of record lengths, the probability of record length increases due to changes, and the initial value for number of records were all determined through analysis of live data. Y was fixed at 3,000 tracks (physical records) on an IBM 2314 disc pack. During the simulation runs, transactions representing weekly activity were applied to the table representing the in-process file until the restricted search parameter was exceeded. The length of each addition was obtained using a random number generator and the probability distribution of record lengths. In a similar manner, record length increases due to changes were computed. Simulation runs were executed with the restricted search parameter set at 3, 4, and 5 with the results summarized in Tables 1–3. These results indicate that during the first week of updating, the required space is usually found on the first track searched. Note also that on the actual in-process file one record existed with a length greater than 2,500 characters.

The results of a simulation of the original in-process file I/O method are recorded in Table 4. Since the number of transactions was slightly overestimated, this method did not last even a week without requiring a reorganization.

The aim of the simulations was to arrive at the restricted search parameter which would allow for reorganization monthly, and yet still require

Table 1. Restricted Search of 3 Tracks on a 3,000-Track File (IBM 2314, Blocked Records)

		File Creation	*End of 1st Week*	*End of 2d Week*	*End of 3d Week*	*Restricted Search Exceeded*
No. of Logical Records on File		19,196	20,046	20,896	21,746	21,876
Average Space Left per Track (Characters)		4,224	4,010	3,794	3,579	3,550
No. of Tracks with:						
0–5 Records		941	794	688	588	574
6–10 Records		2,008	2,127	2,193	2,236	2,243
11–15 Records		51	79	119	176	183
No. of Logical Records with Record Length						
Under 501 Characters		13,015	13,616	14,210	14,804	14,887
501–1,500 Characters		6,143	6,380	6,620	6,856	6,894
1,501–2,500 Characters		37	49	65	85	94
Over 2,500 Characters		1	1	1	1	1
No. of Tracks Having Space Left of:						
Under 1,001 Characters		10	26	47	85	92
1,001–3,000 Characters		310	460	620	768	781
3,001–5,000 Characters		2,027	2,009	1,935	1,851	1,839
Over 5,000 Characters		653	505	398	296	288
Displacement of Record From Home Track During:						
Additions	0	19,196	849	843	838	127
	1	0	1	7	12	3
	2	0	0	0	0	0
Changes	0	0	3,841	3,814	3,797	0
	1	0	9	36	53	0
	2	0	0	0	0	0

Table 2. Restricted Search of 4 Tracks on a 3,000-Track File (IBM 2314, Blocked Records)

	File Creation	*End of 1st Week*	*End of 2d Week*	*End of 3d Week*	*End of 4th Week*	*End of 5th Week*	*Restricted Search Exceeded*
No. of Logical Records on File	19,196	20,046	20,896	21,746	22,596	23,446	24,247
Average Space Left per Track (Characters)	4,224	4,011	3,798	3,581	3,371	3,159	3,007
No. of Tracks with:							
0–5 Records	941	806	700	596	505	417	352
6–10 Records	2,008	2,110	2,173	2,226	2,249	2,252	2,229
11–15 Records	51	84	127	178	246	331	419
No. of Logical Records with Record Length							
Under 501 Characters	13,015	13,859	14,175	14,760	15,346	15,948	16,504
501–1,500 Characters	6,143	6,400	6,646	6,897	7,128	7,349	7,608
1,501–2,500 Characters	37	56	74	88	102	118	134
Over 2,500 Characters	1	1	1	1	1	1	1
No. of Tracks Having Space Left of:							
Under 1,001 Characters	10	19	40	70	115	180	232
1,001–3,000 Characters	310	474	635	811	951	1,085	1,183
3,001–5,000 Characters	2,027	2,006	1,919	1,814	1,705	1,556	1,432
Over 5,000 Characters	653	501	406	305	229	179	153
Displacement of Record From Home Track During:							
Additions 0	19,196	848	845	843	839	823	763
1	0	2	5	5	11	25	35
2	0	0	0	2	0	2	2
3	0	0	0	0	0	0	0
Changes 0	0	3,829	3,828	3,821	3,805	3,708	0
1	0	11	22	23	45	136	0
2	0	0	0	6	0	6	0
3	0	0	0	0	0	0	0

fewer I/Os than the original method. Setting the restricted search parameter at four would easily handle the reorganization requirement (see Table 2); and, as illustrated below, considerably fewer I/Os would be required than with the original BDAM method. From Table 2 it is also evident that when the restricted search is exceeded, the average space left per track is still substantial. Further simulation could be done to explore the tradeoff of file size versus the restricted search parameter. However, the 3,000-track file was available at the time, and in the interest of rapid response for on-line inquiry, it was decided to retain this file size.

I. Original BDAM method (IBM 2314).

$$N_1 = 1.15\,[2\,(3{,}400 + 15{,}400)] = 43{,}240$$

$$N_2 = 6\,(28{,}710) = 172{,}260$$

$$N_3 = 1.15\,(30{,}000) = 34{,}500$$

$$N \;= 250{,}000$$

Table 3. Restricted Search of 5 Tracks on a 3,000-Track File (IBM 2314, Blocked Records)

	File Creation	End of 1st Week	End of 2d Week	End of 3d Week	End of 4th Week	End of 5th Week	End of 6th Week	End of 7th Week	Restricted Search Exceeded
No. of Logical Records on File	19,196	20,046	20,896	21,746	22,596	23,446	24,296	25,146	25,996
Average Space Left per Track (Characters)	4,224	4,015	3,802	3,583	3,367	3,145	2,928	2,716	2,528
No. of Tracks with:									
0–5 Records	941	807	689	579	494	406	342	275	222
6–10 Records	2,008	2,114	2,188	2,243	2,261	2,265	2,221	2,176	2,110
11–15 Records	51	79	123	178	245	329	437	549	668
No. of Logical Records with Record Length:									
Under 501 Characters	13,015	13,600	14,186	14,779	15,358	15,903	16,467	17,068	17,664
501–1,500 Characters	6,143	6,398	6,650	6,885	7,140	7,424	7,694	7,930	8,158
1,501–2,500 Characters	37	47	59	81	97	118	134	147	173
Over 2,500 Characters	1	1	1	1	1	1	1	1	1
No. of Tracks Having Space Left of:									
Under 1,001 Characters	10	22	44	81	127	196	282	380	467
1,001–3,000 Characters	310	448	614	768	929	1,091	1,220	1,303	1,378
3,001–5,000 Characters	2,027	2,006	1,930	1,825	1,692	1,514	1,353	1,213	1,082
Over 5,000 Characters	653	524	412	326	252	199	145	104	73
Displacement of Record from Home Track During:									
Additions 0	19,196	848	842	837	837	822	811	810	770
1	0	2	8	13	12	25	36	34	258
2	0	0	0	0	1	3	2	6	51
3	0	0	0	0	0	0	1	0	10
4	0	0	0	0	0	0	0	0	0
Changes 0	0	3,843	3,818	3,770	3,787	3,726	3,657	3,652	3,040
1	0	7	32	78	55	109	172	173	258
2	0	0	0	2	8	15	14	25	51
3	0	0	0	0	0	0	7	0	10
4	0	0	0	0	0	0	0	0	0

II. BDAM and blocking subroutines with restricted search of 4 tracks on a 3,000-track file (IBM 2314).

$$N_1 = (4+1)\ 3{,}400 + 2\ (1.002)\ (15{,}400) = 47{,}862$$

$$N_2 = 3\ (3{,}000) = 9{,}000$$

$$N_3 = 1.002\ (30{,}000) = 30{,}060$$

$$N\ = 86{,}922$$

Table 4. Original In-Process File I/O Method on a 3,000-Track File (IBM 2314, Unblocked Records)

	File Creation	Restricted Search Exceeded
No. of Logical Records on File	19,196	20,034
No. of Physical Records on File	22,798	23,833
No. of Tracks with:		
0–3 Records	33	21
4–6 Records	719	504
7–9 Records	2,248	2,475
No. of Logical Records with Record Length:		
Under 501 Characters	13,015	13,593
501–1,500 Characters	6,143	6,392
1,501–2,500 Characters	37	48
Over 2,500 Characters		
Displacement of Physical Record From Home Track		
0 Tracks	21,279	21,832
1 Track	1,182	1,344
2 Tracks	282	367
3–5 Tracks	48	204
6–10 Tracks	5	62
11–15 Tracks	2	17
16–20 Tracks	0	5
21–25 Tracks	0	1
26–30 Tracks	0	1

Soon after this study was completed, the University of Saskatchewan Computation Centre installed IBM 3330 disc drives. This had the following effect on the calculations.

I. Original BDAM method.

Since this method is independent of storage device as long as the same number of physical record locations are assigned to the file:

$$N = 250{,}000$$

II. BDAM and blocking subroutines with restricted search of 2 tracks on a 1,425-track file (IBM 3330).

$$N_1 = (2 + 1)\ 3{,}400 + 2\ (1)\ (15{,}400) = 41{,}000$$

$$N_2 = 3\ (1{,}425) = 4{,}275$$

$$N_3 = 30{,}000$$

$$N\ = 75{,}275$$

Note that virtually no track overflows occur in this case if the file is reorganized monthly.

A SOLUTION TO THE STORAGE PROBLEM

Two assembler subroutines accomplish all of the in-process file I/O. One subroutine reads and writes the file sequentially and the other reads, writes, and rewrites (adds and changes) records directly.

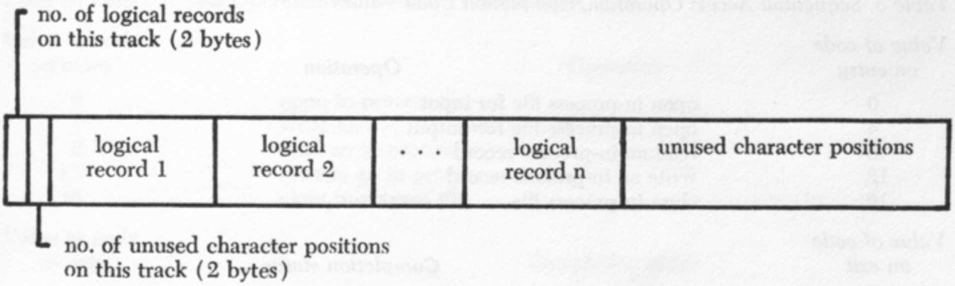


Fig. 2. In-Process File Block (or Track).

The in-process file is organized using the BDAM technique called relative block addressing. Each block is the size of a track on the external storage device. Figure 2 gives a schematic layout of the format of a block or track used by the subroutines.

Each in-process logical record contains 154 characters of fixed length information, two characters of which are a count of the number of variable characters present. The location of the next logical record is thus obtained by incrementing the current record position by 154 plus the number of variable characters in the current record.

SEQUENTIAL ACCESS SUBROUTINE

Two parameters are passed between this subroutine and the calling program. The first is a code specifying which in-process I/O operation is to be performed. On exit from the subroutine, this code reflects the completion status of the operation. The second parameter is the address of an area into which or from which the in-process logical record is to be moved.

Table 5 gives the meanings of values of the operation/completion code described above.

When reading the in-process file, one logical record per call is passed to the calling program. The subroutine stores the location of the current record and the number of records already retrieved from the current track in order to determine when another track should be read.

When writing in-process records, which must be supplied to the subroutine in relative track address order, one record at a time is accepted by the subroutine and placed in position after the previous record. The space left and number of records on the track are then updated accordingly. This process continues until the relative track address is greater than that of the track being loaded, or until the track becomes full. In either case the track is then written. If a record cannot be placed on a track which is within the range of the restricted search, completion code eight is returned to the calling program. Of course, if file size and average record size are accurately predicted, this condition should never occur. Tracks which do not contain any records are initialized for space left and number of rec-

Table 5. Sequential Access Operation/Completion Code Values

Value of code on entry	*Operation*
0	open in-process file for input
4	open in-process file for output
8	read an in-process record
12	write an in-process record
16	close in-process file

Value of code on exit	*Completion status*
0	successful completion
4	end of file
8	record could not be written within the range of the restricted search

ords, and then are written. When all records have been written, the calling program sets the extent of the file by supplying a maximum relative track address. Tracks are initialized by the subroutine until this address is reached; then the file is closed.

DIRECT ACCESS SUBROUTINE

The two parameters passed between the direct access subroutine and the calling program have the same functions as the sequential access parameter. Table 6 gives the meanings of values of the direct access operation/completion code.

The first step in each of the read, write, and rewrite operations is a check of the address of the track which is currently in memory. If the address is the same as that of the required track, an I/O operation is saved.

Write (addition) and rewrite (change) operations have special design features. The first step in a write operation is a search to verify that the record does not already exist on the file. During this search, which begins at the record's home track and extends over the range of the restricted search, the first track located that has enough space to accommodate the record is stored. If the record does not already exist on the file, it is added to the stored track, and then the track is rewritten. If space for the record cannot be found, the record is stored in an overflow track.

In a rewrite operation, the updated record will be either shorter, the same length, or longer than the corresponding record on the track. In the first two cases, the updated record overlays the corresponding record on the track. In the first case, records following the overlayed record are shifted to contiguous locations. After replacing the record, the track is rewritten. In the third case, a check must be made to determine whether or not there is room on the track for the enlarged record. If so, records are shifted to make room for the updated record. If space on the track is not available, tracks are searched sequentially within the range of the restricted search in an attempt to locate a track with the required space. Upon locating such

Table 6. Direct Access Operation/Completion Code Values

Value of code on entry	*Operation*
0	open in-process file
4	write an in-process record (addition)
8	read an in-process record
12	rewrite an in-process record
16	close in-process file

Value of code on exit	*Completion status*
0	successful completion
1	record to be written already exists
2	space cannot be found during write (addition)
3	record to be read is not on the file
4	space cannot be found during rewrite (change-length expanded)
5	record to be rewritten (changed) is not on the file

a track, the record is removed from the original track, the original track is rewritten, the updated record is added to the new track, and then the new track is rewritten. If space cannot be located, a similar procedure is followed to add the record to an overflow track.

CONCLUSION

In addition to reducing sequential access I/O costs and increasing storage capacity, the use of the blocking/deblocking subroutines permits a dynamic reorganization of the in-process file. Whenever records are added to the overflow area, a switch file on disc is set to indicate that the in-process file requires reorganizing. After the in-process file update program has completed execution, the reorganization switch is tested. If the test is positive, the in-process file reorganization is automatically executed.

Due to the low cost of reading the in-process file sequentially with the subroutines, the design of several TESA-1 programs has been changed. For example, before the subroutines were implemented, the program which produced invoice clearance reports maintained a file of item numbers which were applicable to this program. The in-process records corresponding to these item numbers were read directly. Now, since only 1,425 I/Os are required to read the entire in-process file, the invoice clearance report program reads the in-process file sequentially, and the maintenance of the file of approximately 3,000 item numbers applicable to invoice clearance reports is completely eliminated. This change has reduced the cost of producing these reports by more than one-third.

The costs of acquiring and cataloging items through TESA-1 have been calculated using accepted techniques, and documented.[9,10] Even though TESA-1 costs were more than competitive with other reported costs prior to the implementation of these subroutines, a 40 percent reduction in computer charges has been achieved through their use.[11]

REFERENCES

1. E. J. Buchinski, W. L. Newman, B. A. F. Burns, and F. H. Coxford, "Technical Services Automation—Phase I," Paper in preparation.
2. E. J. Buchinski and W. L. Newman, "The Cataloguing System at the University of Saskatchewan, Saskatoon: A Subsystem of TESA-1 (Technical Services Automation —Phase I)," in *Automation in Libraries;* papers presented at the C.A.C.U.L. Workshop on Library Automation, Vancouver, June 19–20, 1971 (Ottawa: Canadian Association of College and University Libraries, 1971), p.3.1–3.23.
3. B. A. F. Burns and P. Pridmore, "TESA-1/Cataloguing: A Second Look," in *Library Automation in the Nation;* papers presented at the C.A.C.U.L. Workshop on Library Automation, Regina, June 10–11, 1972. (Ottawa?: Canadian Association of College and University Libraries, 1973), p.10.1–10.57.
4. W. L. Newman and E. J. Buchinski. "Entry/Title Compression Code Access to Machine Readable Bibliographic Files," *Journal of Library Automation* 4:72–85 (June 1971).
5. *IBM OS Full American National Standard COBOL Compiler and Library. Version 2 Programmer's Guide.* GC28-6399-2 (3d ed.; New York: IBM, 1972).
6. *IBM OS Full American National Standard COBOL.* GC28-6396-3 (4th ed.; New York: IBM, 1972).
7. *Introduction to IBM Direct-Access Storage Devices and Organization Methods* (New York: IBM, 1973).
8. P. C. Mitchell and T. K. Burgess, "Methods of Randomization of Large File with High Volatility," *Journal of Library Automation* 3:79–86 (March 1970).
9. Buchinski, Newman, Burns, and Coxford, "Technical Services."
10. B. A. F. Burns and P. Pridmore, "TESA-1 Cataloguing," p.10.35–10.43.
11. Buchinski, Newman, Burns, and Coxford, "Technical Services."

Automated Serials Control: National and International Considerations

Mary Kay DANIELS: Library of Congress

The proceedings of the October 1974 ISAD institute on automated serials control are summarized.

INTRODUCTION

On October 11 and 12, 1974, in Atlanta, Georgia, an institute on "Automated Serials Control: National & International Considerations" was cosponsored by the Information Science and Automation Division of the American Library Association and the American Society for Information Science. Moderated by S. Michael Malinconico of the New York Public Library, the institute brought together speakers who touched on both the current cataloging and automation concerns in the area of serials handling.

LC, MARC, and CONSER

Josephine Pulsifer, group leader within the MARC Development Office of the Library of Congress (LC) in charge of the development of the Library of Congress' automation program for serials, opened the first session with a discussion of the current method of handling of serials at the Library of Congress and the anticipated effect of the CONversion of SERials (CONSER) project on LC operations.

The Serial Record Division, which is part of the Processing Department of LC, has the responsibility of manually searching and accessioning over 1.4 million serial pieces each year. If a piece cannot be matched against an entry in the serial record file, it is forwarded to the selection officer of LC for consideration. If the selection officer determines that the item should be added to the collections, it is then sent to the Cataloging Section of the Serial Record Division for full descriptive cataloging. Approximately 12,000 to 13,000 titles, in all languages, are processed by this section each year.

Once the descriptive cataloging is completed, a worksheet containing these data is forwarded to the National Serials Data Program (NSDP), a program which is now under the administrative control of the Processing Department, for the addition of a key title and International Standard Serial Number (ISSN). Subject headings and an LC classification number are then added by the Subject Cataloging Division, and the LC call number is completed by the Shelflisting Section of the Subject Cataloging Division. Dewey numbers are also added for all English and French language serials. At this point, the serial is forwarded to the Serials Division, while the worksheet is sent to the MARC Development Office (MDO) for entry into MARC. Ms. Pulsifer noted, however, that since this data input function is now in production, it will be transferred from MDO to the Serial Record Division.

Following MARC editing for content designators (tags, indicators, and subfield codes) and for editorial details, the serial cataloging data are keyed on a magnetic tape selectric typewriter (MTST) containing the 175 characters of the standard ALA character set. The typewriter cassette is converted to a machine-readable computer tape, and computer processing is performed to create the bibliographic record and a diagnostic listing. The diagnostic listing is used by the MARC editors to verify the record against the data on the original worksheet, and any corrections required are forwarded to the MTST operators for keying.

Twice a week, tapes are sent to the Card Division, which is responsible for the conversion of MARC records into the communications format and for their distribution to subscribers once every four weeks. The *MARC Distribution Service—Serials* includes all serials cataloged since January 1973, with the exception of nonroman alphabet titles cataloged between January and September 1973. The data base currently contains approximately 11,600 records, over 10,300 of which had been distributed to MARC subscribers as of the end of September 1974. Fifty-six percent of the data base consists of English language materials, although fifty-six different languages have been included in the service. One-third of the data base consists of U.S. imprints, 5 percent are West German, 4.6 percent are British, and 4.3 percent are Indian. Ms. Pulsifer noted that although nonroman alphabet serials are input into the MARC data base in romanized form, they are never printed in this form on catalog cards or book catalogs produced by LC, but are set in hot type by the Government Printing Office.

During the CONSER project, it is anticipated that the Library of Congress will input all newly cataloged serials on-line through a CRT device hooked up to the Ohio College Library Center (OCLC) in Columbus. OCLC would than return these records to LC in LC's internal processing format for use in generating tapes for the *MARC Distribution Service—Serials* and for internal LC use.

It is anticipated that CONSER operations at the Library of Congress will occur in two stages: (1) input into the on-line OCLC data base will be done after the completion of the entire LC cataloging process; and than later, (2) input into the on-line OCLC data base will be done following the LC descriptive cataloging process but before the subject cataloging process, thus allowing partial records containing descriptive information, ISSN, and key title to be distributed as soon as they are available. During this second phase, subject heading, LC classification and Dewey classification information would be added to the record at a later stage, resulting in a second distribution of the record.

When the second stage of the planned approach is operational, two additional situations may occur: (1) a CONSER record in the OCLC data base may be used as the base record on which the Library of Congress would build its record, and (2) LC will undertake to postedit, or authenticate, all non-LC records in the data base against the LC Official Catalog. The only exception to this authentication process would be the Canadian names, which would be referred to the National Library of Canada for handling. It is anticipated that these authenticated records would be distributed as uniquely identified records in the *MARC Distribution Service —Serials*. It is also hoped that input of *New Serial Titles* reports can begin via the CONSER system early in fiscal 1976.

At the end of the planned two-year CONSER project, it is anticipated that participating libraries will be able to input cataloging records online to the Library of Congress. This would require on-line access to the serials data base and hopefully to LC's authority file. As a part of this build-up, all new names with references will be input into machine-readable form beginning in 1975. It is expected that these name authority records will be made available to libraries through a MARC distribution service.

The last major topic of Ms. Pulsifer's talk concentrated on the cooperative nature of the serial format that has been developed. *Serials: A MARC Format* is an outgrowth of the original Serials Data Program which began in 1967 and the MARC Pilot Project, which began in late 1966.[1] The structure of the format conforms to the standards set by the American National Standards Institute (ANSI), later adopted by the International Organization for Standardization (ISO). Data elements which occur both in *Books: A MARC Format* and *Serials: A MARC Format* are identified by the same content designators.[2]

The development of the serials format was complicated by the existence of two schools of thought concerning the treatment of serials: (1) serials should be treated as much as possible like monographs, or (2) serials should be treated as a world unto themselves. The resulting format is a compromise between these two positions. The first edition of *Serials: A MARC Format*, issued in 1970, and "Addendum no. 1," issued in 1971, reflected the actual experience that LC personnel encountered in converting to machine-readable form serials held in the Main Reading Room collection.[3] The second edition of the format, issued in 1974, includes only a few changes, notably the use of the LC card number instead of the ISSN as control number and the addition of key title required by the National Serials Data Program.

As a result of the impending CONSER project, several changes and additions in the MARC serials format have been proposed in order to accommodate the needs of the participants. At this time, these changes have been submitted to the RTSD/ISAD/RASD Representation in Machine-Readable Form of Bibliographic Information (MARBI) Committee of the American Library Association for review. Each of the changes will be considered by the committee separately, and any change not accepted by the committee will not be included in the communications format used to distribute MARC serial records, although it may be included in the internal processing format used within the Library of Congress. Ms. Pulsifer did not review the proposed changes in detail, but noted that information about them could be obtained from the MARC Development Office upon request. In addition, as any or all of the changes are approved by MARBI, they will be made generally available, probably in the form of an addendum to the second edition of *Serials: A MARC Format.*

Although no firm deadline for the start-up of the CONSER project had been settled, the Library of Congress was working on the computer program changes that will be required within LC, as well as on the preparation of complete editing and cataloging manuals. As the terminals to the OCLC data base are already in place at LC, it is anticipated that once the go-ahead sign is given, LC will immediately begin to input via the CONSER project.

CATALOGING CONSIDERATIONS

Joseph Howard, chief of the Serial Record Division at the Library of Congress, and formerly chief of the Descriptive Cataloging Division, discussed the cataloging rules for serials and their implications for automated processing. As Mr. Howard noted, serials are a knotty problem. And, like people, they are born, may marry, divorce or stay single, and eventually die. They also have the facility to change names and to propagate at will.

In defining the specific problems in cataloging serials, Mr. Howard noted that there are three major areas: (1) choice of entry (What is the main entry of the item going to be?); (2) form of entry (Once the main entry has been chosen, in what form will the entry be given?); and (3) descriptive information. To date, these areas of serials cataloging have been influenced primarily by five publications: (1) the *Statement of Principles adopted at the International Conference on Cataloging Principles (Paris Principles)* of 1961; (2) the *ALA Cataloging Rules for Author and Title Entries* of 1949; (3) the *Anglo-American Cataloging Rules (AACR)* of 1967, which replaced the *ALA Rules* as the American standard in 1967; (4) the *Guidelines for ISDS* of 1973, which emphasize the International Standard Serial Number and key title; and (5) the *International Standard*

Bibliographic Description for Serials, or *ISBD(S),* of 1974, which deals only with descriptive information.[4-8]

The *Paris Principles* of 1961 provide that the choice of main entry for a serial should be: (1) a corporate body for those serials whose titles consist of a generic term preceded or followed by the name of a corporate body, providing the serial includes some account of the activities of that corporate body, or (2) a title for those serials which are known conventionally or primarily by title. The *Principles* also provide that if a serial is issued successively under different titles, an entry should be made for each title, with an indication given of at least the immediately preceding and succeeding title.

Since, as Mr. Howard noted, this guidance is somewhat vague, the authors of the North American text of the *AACR* departed from these principles. The authors of the *AACR* felt the name of the corporate body responsible for a serial was too powerful a criterion to be nullified if no activities of the corporate body were included in the serial. They also felt that the rule of entering a serial under title if it were primarily or conventionally known by title was too vague to be applied with any consistency. Consequently, although cooperation on a national and international level is becoming increasingly imperative in a world of rising costs and continued information expansion, bibliographic control is complicated by the differences in cataloging practices.

Recent efforts directed at developing serial cataloging codes for international use have resulted in the development of the *ISBD(S)* and the *Guidelines for ISDS.* To date, some of the differences among these codes have been reconciled, but some are still pending. For example, the rules for recording of title under *ISDS, ISBD(S),* and *AACR* have recently been brought into agreement by two changes to the *AACR:* rule 162B has been dropped, with the result that title statement is no longer truncated, as it was in the past. The ISDS distinctive title concept is now being practiced in that an author statement is being added to titles which consist only of generic terms.

Insofar as choice of entry is concerned, however, a conflict between the *Guidelines for ISDS* and the *AACR* still exists. The problem is caused by the fact that ISSNs are assigned on the basis of title rather than on the basis of bibliographic entity and that a new entry may be made under *AACR* because of a change in issuing body and/or title, while a new entry may be made under the *Guidelines* only for a change in title. This problem could be overcome by using the ISBD(S) distinctive title as the main entry for all serials and using added entries for issuing bodies; or, serials consisting of generic terms plus issuing body could be entered under corporate body and all other serials could be entered under title with added

* This matter was a topic of discussion in an open forum at the 1975 ALA Midwinter Meeting in Chicago.

entries given for other issuing bodies.* Mr. Howard emphasized that the use of either of these alternatives implies a change to rule 6, not a dropping of rule 6. He also emphasized that the choice of entry really doesn't matter as long as the rules developed can be interpreted in the same way by everyone. It is a problem which must be tackled with international, national, and local concerns in mind. This idea of cooperation, though, is predicated on the assumption that it is more worthwhile to catalog new materials than to continue recataloging and reconciling differences among older materials.

Like choice of entry, the form of entry is another area of major concern. Under the ALA Rules many corporate bodies were entered under place. Under the AACR some corporate bodies are entered under place, but most are not. Consequently, when the AACR was adopted in 1967, the policy of superimposition was implemented at LC. In essence, superimposition meant that newly cataloged materials were cataloged under the AACR only if the corporate heading form had not been established previously under the ALA Rules. Adopted to avoid large-scale changing of headings in the LC catalogs, this policy created as many problems as it solved. LC is now considering whether to drop superimposition or not. This problem is currently being reviewed in light of possible future changes in the choice of entry, the automation of the authority file, and the requirements imposed by the CONSER project.

Mr. Howard also discussed a third problem area: the successive entry rules of the AACR code versus the latest entry rules of the ALA code. Under the ALA code the cataloger with a new serial in hand must relate it to former titles of the same serial by recataloging the entire history of the serial and treating it as one unit record. Under the AACR code a new title is always treated as a new unit record, although links are provided to at least the previous title and the title immediately following.

In reviewing the present status of the ISBD(S) code, Mr. Howard noted that it has been published, and a copy of it has been sent to the Cataloging Code Revision Committee of ALA, which has approved it in principle. If the ISBD(S) were also to be approved in principle by the Joint Steering Committee at the 1975 ALA Midwinter Meeting, the second edition of the *AACR* will reflect the ISBD(S) in the redraft of chapter 7. Paul Winkler, principal descriptive cataloger at the Library of Congress, would be responsible for this redraft since he has been asked by the Joint Steering Committee to act as one of the five editors of the second edition of the *AACR*. The Joint Steering Committee has indicated, however, that it does not wish to consider publishing changes to the existing edition of the *AACR* by itself, but would rather publish a new edition.

From a long-range point of view, adoption of the ISBD(S) by the Library of Congress could resolve many of the cataloging problems. However, from a short-range point of view, LC is committed to AACR for descriptive cataloging, to ISDS for conforming to the role of the U.S. arm of the international group, and to CONSER. With cooperation, though, it may be possible to catalog an item only once in the U.S. and possibly only once in the world.

INTERNATIONAL STANDARDS

Lucia Rather, assistant chief of the MARC Development Office of the Library of Congress, spoke about past and present international efforts to develop standards for serials cataloging, machine-readable record formats, and character sets. In the area of cataloging, two standards for serials exist: (1) the *International Standard Bibliographic Description for Serials,* and (2) the *Guidelines for ISDS.*

The history of *ISBD(S)* began with the work on an international standard for monographs in 1970, although it wasn't until 1971 that the International Federation of Library Associations (IFLA) began to work on a similar standard for serials. Throughout 1972 and into 1973, several meetings were held and three working papers were developed for consideration scriptive cataloging, to ISDS for conforming to the role of the U.S. arm by the serial cataloging community. In its current form, the ISBD(S) is largely based on the ISBD(M), that is, the *International Standard Bibliographic Description for Monographs.*[9] Like ISBD(M), the ISBD(S) puts forth the elements in a prescribed order which, regardless of language, is intelligible to any cataloger through the use of standard punctuation. Unlike ISBD(M), however, the ISBD(S) relates not only to bibliographic description, but also to identification.

Under ISBD(S) the title statement must be a distinctive title, and should include the name of the issuing body as part of the title. The cataloging treatment of serials under ISBD(S) would, therefore, be a compromise between handling a serial like a monograph as opposed to handling a serial as a unique type of material.

Ms. Rather next turned to international developments in the area of automation. Regarding the format structure of a machine-readable record, three basic requirements are included in the 1972 ISO 2709 standard: (1) a twenty-four-character leader; (2) a record directory pointing to a variable length field; and (3) a series of variable length fields. Currently, though, a working committee is reviewing a group of changes to the standard that have been proposed by various interest groups. Some of the changes suggested are: (1) lengthening the record directory to include a notation as to which field has been updated if the record is being issued as a change record and information about field sequencing if the same field tag is repeated more than once in a record; (2) allowing tags to include alphabetics rather than being limited to numerics; (3) moving the position of the variable-length fixed fields from after the record directories to before the record directories; (4) allowing the interchange of bibliographic records in EBCDIC as well as ASCII; and (5) the abolition of field separators, which are redundant since the length of each field is given in the record directory. Ms. Rather noted that only the first two proposals will probably be adopted in some form by the committee.* However, it should be recognized that these proposed changes may not necessarily have any effect on the U.S. standard, as it has also been proposed that changes made to ISO 2709 should not cause existing standards based on the old ISO 2709 to be changed.

Internationally, serials projects are widespread. In addition to the MARC serials project in the United States, there is a Canadian MARC serials effort, two British serials efforts (one as part of the British National Bibliography and one as part of the Birmingham Library Project), a French effort, a German Union List of Serials, and a Scandinavian Union List of Serials which is under development. The MARC serials tapes are also distributed to the University of Sydney in Australia and to the Biblioteca Nazionale Centrale in Rome, Italy.

While some of the national formats that have been and are being developed are compatible with each other, others are not. Consequently, since 1973 the IFLA Working Group on Content Designators has met four times in an effort to develop a multimedia format which can be used on an international level of exchange. Under this concept, the national formats would remain intact for national use, but would be translated to an international format for international exchange. The main problems of the committee have been threefold: (1) formats tend to follow cataloging rules, yet the same rules are not used on an international basis; (2) there is a lack of agreement as to how complex the format should be—for example, the U.S. MARC has a 245-title field containing three subdivisions, the French format has a 245-title field which contains thirteen subdivisions, while the German format divides the title into more than thirteen areas which are different from the French subdivisions; and (3) there is a lack of agreement as to the purpose of the format—whether it is a bibliographic tool or a tool for information retrieval and control.

Another major interest being studied on an international level is the character set. Currently, there exists the ALA standard character set (175 characters), used for roman alphabet languages and for romanized forms of nonroman alphabets. This character set has become widely available on line printer and photocomposition devices and on CRT terminals used for input, display, and, in some instances, printing. It is used in the United States and Britain, and an expanded version has been proposed for use in France.

At the 1972 ISO meeting, it was proposed to develop extended character sets not only for the roman alphabet, but also for Cyrillic, Greek, and African alphabets. Presently, drafts for the extended roman, Cyrillic, and Greek character sets are almost completed. The draft for the African set,

* At a meeting in Stockholm in October 1974 all proposals except the second were rejected.

however, is far from completion, mainly due to a large variety of symbols and characters used in the numerous African languages. This effort has been complicated by the differences in the alphabets used by newly emerging African countries as opposed to the notations previously defined by missionaries. One additional character set, a mathematical character set, has also been added to the list of extended character sets being developed.

Ms. Rather noted that when the results of this standards work is completed, it will eventually be incorporated into the MARC effort at the Library of Congress. In addition, it should be realized that the implementation of an extended roman character set will have implications for modification of the ALA character set.

ISDS AND NSDP

Joseph Price, acting director of the National Serials Data Program, summarized the history of the National Serials Data Program. In 1964 the Task Force for Scientific and Technical Journals (COSATI) created a special subcommittee to look into matters such as (1) the definition of a serial, (2) the estimate of a serials population, (3) the procedures for starting an inventory, and (4) benefits to be derived from such an inventory. In 1965, a contract was awarded to the Information Dynamics Corporation to study the feasibility of creating a national inventory. By 1966, LC was approached for its views on the proposed national inventory. In 1967, the National Library of Medicine, the National Agricultural Library, and the Library of Congress agreed to work on a cooperative serials data base program. The concept included the creation of a machine-readable data base covering all subject fields and a description and location of where the serial material could be found. LC was to act as the executive agency, and the Information Systems Office as the director. The Joint Committee on the Union List of Serials was to act in an advisory capacity.

To determine the elements required in the data base to be developed, the Nelson Associates undertook a user survey. As a format for machine-readable records, the MARC format for serials became available as a working document in 1969 and was a timely and helpful augmentation to the NSDP effort. A pilot project was begun in September of 1969 at the Library of Congress under the direction of the Association of Research Libraries (ARL), based on funds solicited by the U.S. National Libraries Task Force on Automation and Other Cooperative Services and provided by the National Agricultural Library. By December 1969, approximately 7,000 records had been coded and verified.

In 1972, NSDP was established as an operational, on-going program. The program follows international standards for bibliographic description and interchange, principally those of the ISSN (ISO 3297) and *Format for Bibliographic Information Interchange on Magnetic Tape* (ISO 2709). The ISSN is a unique eight-digit number applied to a serial publication and intended to serve as an unambiguous identifier of that serial. The

NSDP serves as the U.S. representative to the International Serials Data System (ISDS), a UNESCO-sponsored developing network of national and regional centers responsible for systematic registration of serials. As a member of ISDS, the NSDP ascribes to the ISDS *Guidelines,* which are rules for description of serials developed for the most part to be in conformance with the developing *International Standard Bibliographic Description for Serials* rules.

As the ISDS *Guidelines* were developed, conflicts with the *Anglo-American Cataloging Rules* came to light. Chief among these conflicts were rules for choice and form of entry and for determining entry changes. This is a critical point since a new title is always given a new ISSN. Some relief has been obtained by the recent dropping of *AACR* 162B and subsequent rule interpretations which allow the placing of the author in the title whenever the title consists of a generic word, e.g., "journal." In addition to the ISSN and rules for its assignment, the ISDS code also designates a number of elements that should always be included in a machine-readable record describing a serial. The operating details for using ISDS are described in the *Guidelines for ISDS,* which are available from NSDP upon request.

One of the primary concerns of the NSDP office and of the ISDS is the distribution of the ISSNs which have been assigned. One mode of distribution is the *MARC Distribution Service—Serials;* another is the ISDS *Bulletin* which is issued bimonthly, and which lists serial titles by key title and ISSN, as well as giving links by publisher, place, country code, and abbreviated key title. The base file of serials under development is the responsibility of the International Center of the ISDS in Paris. To date, the Bowker and Ulrich publications dealing with serials have included ISSNs which have been added to the publications automatically. Although some problems have since been created because the titles in question do not always conform to the *ISDS* code, the benefits for identification had been considered to be of greater significance than the possible problems. It is hoped that ISSNs will become readily available on all serials and will be used by librarians, subscription agencies, and abstracting services. The aim is to see the ISSN employed worldwide as a unique identifying number.

CONSER PROJECT

Lois Upham, acting coordinator of the Cooperative CONversion of SERials project gave a summary of the background of CONSER, its current status, and its future plans. Ms. Upham's paper was published in its entirety in the November 29, 1974, *LC Information Bulletin* and is not further summarized here.

IMPACT OF NATIONAL DEVELOPMENT AND AUTOMATION ON LIBRARY SERVICES

Paul Fasana, chief of preparation services for research libraries at the

New York Public Library, spoke from the point of view of a technical services librarian, noting their apprehensions, frustrations, and concern about automation projects in general and about the CONversion of SERials (CONSER) project in particular. Mr. Fasana's paper was published in its entirety in the December 1974 issue of *JOLA,* and is not summarized here.

CANADIAN MARC

Edwin Buchinski, head, Canadian MARC Office, alleviated some of the apprehensions of those present by explaining that the Canadian effort is being promoted by the National Library of Canada as one facet of its commitment to create a national information network. The Canadian MARC serials format is the second format to be designed by the National Library of Canada, and the Canadian MARC serial records will be available in early 1975.

In reviewing the history of the Canadian MARC effort, Mr. Buchinski noted that the current program began with the introduction of the National Library Act of September 1, 1969. Two task groups, one on cataloging standards and one on machine-readable cataloging, were appointed in 1970. The Canadian Task Group on Cataloguing Standards paid particular attention to the bilingual needs of Canada. In its final report issued in 1972, the Canadian Task Group on Cataloguing Standards recommended that all headings be in English and French and that bilingual publications be given a full cataloging treatment in both English and French. The task group also recommended the adoption of the *Anglo-American Cataloging Rules* as the standard cataloging code with the exception of rules 98 and 99.

The MARC task group undertook to investigate the requirements for a Canadian MARC format. Its final recommendations included a format that would be highly compatible with the Library of Congress MARC II format. It also recommended in the final report, issued in 1972, that the National Library of Canada carry out further developmental work on the format.

The Canadian MARC Office has now developed a Canadian MARC format for serials which is compatible with the Canadian MARC format for monographs and the Canadian cataloging standards and which is adaptable to international standards. The Canadian MARC Office worked in close cooperation with the staff of the MARC Development Office of the Library of Congress and the participants of the CONSER project to assure that the Canadian serials format would be compatible with other North American efforts. The *Canadian MARC Communication Format: Serials* manual will be available from Information Canada, Ottawa, in February 1975.[10]

In comparing the Canadian and Library of Congress MARC formats for serials, Mr. Buchinski noted that both are magnetic tape formats, written in ASCII, based on the *Anglo-American Cataloging Rules,* and are patterned on the MARC II format for monographs. Differences between the two formats are mainly due to bilingual requirements of Canadian users.

If the publication being cataloged is a bilingual work, two Canadian MARC records are issued for the same work, one in English and one in French. The same control number is used for both records except for the English/French language designator which can be used to readily distinguish one from the other. The access points (main entries, added entries, and subject headings) are given in French for the French language record and in English for the English language record. LC subject headings or Canadian headings are used for English subject access; Laval headings are used for French language subject approach.*

If the publication being cataloged is a unilingual work, only one Canadian MARC record is issued, with subject headings, and main and added entries in both English and French if available and applicable. The choice in using the French or English form of a heading in the 1xx and 7xx fields of a record for a unilingual publication is determined by the language of the work. For Romance language publications, the French form of the heading is used; the English form of the heading is used for all other publications. The equivalent form of heading is carried in the 9xx field of the unilingual record and a special linking field, 990, is provided to indicate which fields have equivalent forms. Subject headings are all carried in 6xx fields; the second indicator differentiates the source of the heading and therefore its language.

Other differences in the Canadian format for serials as opposed to the LC MARC format for serials include the following: (1) record link code information may be carried in the record (position 19 of the leader); (2) historical notes on the work may appear in local field 9xx; (3) authority and cross-reference information may be given in 9xx fields for the main entry, added entries or other access points in the record; and (4) the use of subfield codes may vary, such as the inclusion of additional values in the 245, 246, and 247 title fields. The subfield codes used by Canada in tags 245, 246, and 247 will not affect the CONSER format, as it is anticipated that these will be internally translated by OCLC. In addition, the content of the record may vary due to a different interpretation of the *AACR*. Mr. Buchinski also noted that with the implementation of desuperimposition by the Library of Congress, the uniformity of content should be promoted.

At present, the Canadian MARC Office is planning a monthly record distribution program. It is anticipated that the service will include approximately 5,000 records each year, all of which will be made available to OCLC for the CONSER project. In addition to the Canadian MARC effort, a national International Serials Data System Centre was established

* Canadian subject headings are listed in *A List of Canadian Subject Headings* issued by the Canadian Library Association. Laval headings are found in *Répertoire des vedettes-matières,* 7e edition hors commerce, issued by the Bibliothèque de l'Université Laval.

in Canada in 1973. The internal Canadian MARC format used at the National Library is designed to accommodate the ISDS, and a program is being developed to translate records in the internal format to the ISDS standard format. Canada is presently negotiating an exchange of tapes with the Library of Congress, Great Britain, and France. The aim is to establish a universal bibliographic control and a national record service which would be authoritative, timely, and in a standard form.

Mr. Buchinski predicted that exchange of MARC records with the Library of Congress and CONSER would create an awareness of national differences, but he could see no reason why the formats should not be compatible. For example, current plans call for an authority format to be used in the Canadian MARC system; a draft version of this format has been worked out in cooperation with LC and bears a strong resemblance to the LC version which has been forwarded to MARBI. Within the internal Canadian MARC record format employed at the National Library, the field content will consist of a control number relating to an authority file, rather than the actual authority heading. This decision is expected to reduce keying errors and keying time in the production of Canadian records. Mr. Buchinski anticipated the MARC system would operate without an authority file as of January 1975. He expected implementation of bibliographic data control by an authority file by March 1975.

The possibility of having compatible format standards also seems good. For example, ISBD(M) has eliminated many differences in LC–National Library of Canada cataloging practices. Specifications on changes to LC formats are being made available prior to implementation, and the Canadian MARC Office now has observer status on the MARBI committee. As for content standards, continued and increased cooperation is needed among the Library of Congress, the National Serials Data Program, ISDS Canada, and the National Library of Canada. It is anticipated that the National Library of Canada and the Library of Congress will exchange authority files, which should promote the consistent use of AACR by both libraries.

CANADIAN AUTOMATED SERIALS SYSTEM AND SERVICES

Cynthia (Pugsley) Durance, assistant librarian for planning at the University of Waterloo, and soon to be director of the Cataloguing Branch at the National Library of Canada, focused primarily on the development of a Canadian serials data base, a node of the planned Canadian computer-based national bibliographic network. The National Library of Canada began active planning of this network in 1969, fully cognizant of the difficulties which would arise due to geographical, jurisdictional, and bilingual factors in the country. Throughout the planning for this network, the National Library has actively sought the advice of members of the Canadian library community.

In 1972, as a result of an internal study for the development of the Canadian serials data base, the National Library of Canada contracted with York University to do a study of existing machine-readable serials data bases and supporting systems which the National Library might acquire for use as base files for this system. Richard Anable was the principal investigator. The results of his work, which started with the informal meeting of Canadian and American librarians interested in serials problems, led to the formation of the Ad Hoc Discussion Group on Serial Data Bases and then to the CONSER project. In 1973 a Subgroup on the Union List of Serials of the Canadian Union Catalogue Task Group was formed to "study the problems associated with the development of the Canadian national serials data base and to make recommendations to the Canadian Union Catalogue Task Group for its creation and use in connection with national and regional union lists; and secondly, to act as Canadian liaison with the Ad Hoc Discussion Group on Serial Data Bases."

The subgroup is composed of eleven members of the Canadian library community who have expertise in serials, cataloging, and automated systems. To date the subgroup primarily has been concentrating on the resolution of problems associated with conflicts in serials bibliographic standards, with developing a usable definition of a serial for union list purposes, and with devising a draft standardized holdings notation format.

The building of a Canadian serials data base primarily but not exclusively for union list purposes is now underway. The National Library will be participating in the CONSER project and a program of editing existing manual union lists of serials has been going on for some time. The National Library plans to use the CONSER file as the base file for its serials data base. Data from ISDS, LC MARC, and BNB MARC tapes and from other planned Canadian shared cataloging projects will be input to the system.

During the next three years, work will continue on plans for maintaining a union list of serials network and will progress to decisions on resulting products and services. An operational fully automated system is planned for 1977.

Although decisions on the products and services to be produced from such a system have not been finalized, probable products will include a printed union list of serials, the serials portion of the Canadian national bibliography *Canadiana,* the CAN MARC serials tapes, and the ISDS tapes for the Paris Center.

Other possible products and services from such a machine-readable serials bibliographic and union data base include various types of selected union lists including location lists of titles in SDI services, "critical" lists of the titles held only in one library in the country, lists of key titles and ISSNs, subject lists, etc. A faster and more efficient interlibrary loan location system is also possible. Management information derived from a large union holdings system could also be very useful for such purposes as collection rationalization and development. Whether or not all of these technically possible products and services will ever be produced depends on the need, the cost/benefit and money available, as well as on the willingness of librarians to sacrifice some independence of control in favor of true cooperation.

The National Library of Canada has undertaken the development of an ambitious and complex system. It is still years from completion, but the progress to date has been most encouraging considering the magnitude of the problems to be solved and the multiplicity of the users to be satisfied. The maintenance of a union data base on a national scale depends on input from many sources, and its success is in direct proportion to the excellence and currency of input.

REVIEW OF SYSTEM CONSIDERATIONS

Linda Crismond, assistant university librarian for technical services at the University of Southern California, attempted to bring together much of the theory discussed during the sessions with a review of considerations for approaching serials automation. In looking at the serials problem, Ms. Crismond quoted from the "Survey of Serials" article by Hans Weber, which appeared in the spring 1974 issue of *Library Resources & Technical Services*, that subscription prices for serials rose by 9 to 10 percent in 1973 and by 20 percent in 1974, and totaled 65 to 70 percent of a library's materials budget. At that rate, Mr. Weber speculated that serials would equal 98 percent of a library materials budget by 1978 or 1979.

In looking back to 1965, Ms. Crismond noted that serials was the "hot subject" in terms of automation. To the casual observer, serials processing at that time appeared to be a repetitive, isolated application that would be easy to automate. But, with further analysis, it became obvious that the serial check-in task was not a simple, isolated function as was first thought. Until recently, therefore, the automation of serial processing tended to take a back seat to automation of other library operations, such as acquisitions and book catalogs.

In developing a serials system, consideration must be given both to user needs and to internal processing needs. Specific considerations should include: Is an individual listing or union listing of serials desired? Will the serials file be used primarily as a location tool or as a bibliographic tool? What pieces of information are required in the serials file, such as library locations, holdings, local notes and call number, order information, beginning date of serial, language of publication, place, publisher, identification number, indexes, cross-references, general notes, etc.? What arrangement of the serials is desired on listings, such as entry or title, call number or shelflist number, key-word-in-context (KWIC) or key-word-out-of-context (KWOC), all titles covered by an index or indexes, or by schools or departments? What method of production should be used to generate the listing, such as computer print-out, xeroxed copy, computer output microfilm (COM), or photocomposed copy?

The major internal function that must be considered in detail during automation is the check-in function. To give an idea of the complexity of this function, Ms. Crismond noted that of the twenty-two programs coded for the San Francisco Public Library automated serials system, twenty-one programs were handled by one programmer while one program, the check-in function program, was handled by a second programmer. Under manually operated serials systems, a unit card is usually used to record one or more subscriptions for the same title. Under an automated system, this type of detail can be maintained and accessed by computer programs independently, and can even be used to predict when serials should arrive.

The prediction of serials arrival, however, is not a simple task. As Ms. Crismond noted, approximately 50 percent of the serials issued are of an irregular frequency. Many computer programs, though, can now handle regular irregulars, that is, serials which have an irregularity of frequency which is predictable, such as a publication which appears monthly from September through June but has no issues during July or August. Within the check-in program, the check-in function may be handled with a calendar grid, a frequency statement, a number of issues approach, or a combination of all three. The check-in function as implemented by the user may be handled with a punched card, which could display the first few characters of the serial main entry plus a linking number; or a list, which could be accessed and result in preparation of input; or on-line. One advantage of the on-line system, such as that of the UCLA Biomedical Library, is that a number of back issues can be checked in with ease at one time.

In addition to the check-in function, Ms. Crismond noted that other related functions should be considered for automation: (1) claims letters could be generated based on nonreceipt of materials according to an established frequency schedule; (2) binding information could be stored, specific instructions to the binder could be prepared, and the holdings file statement could be updated when the bound volume was returned; (3) discard slips could be prepared for ephemeral material which has been maintained the required length of time; (4) renewals could be anticipated, orders could be printed on special order forms, and fund number, price, and date of payment information could be recorded when payment was issued; (5) routing of materials, including production of a routing slip with names in a priority sequence, could be printed by computer; and (6) an array of management reports could be created, including statistics on current titles versus titles held, on titles held by branches, etc. To date, some sharing of automated serials systems has occurred. For example, the San Francisco Public Library has given its software to the Jackson Library of Business at Stanford University. Commercially, however, nothing is available as yet, although OCLC does have a system in the design stages.

PANEL DISCUSSION AND OPEN FORUM

Rather than have the wrap-up provided by one speaker, the final session of the institute was devoted to a "panel discussion" which first allowed speakers to respond to the remarks of other speakers and then opened up the discussion to questions from the floor.

Lucia Rather opened the session with comments on the presentation given by Paul Fasana. Ms. Rather noted that Mr. Fasana's charge of poor communication about the CONSER project was a valid one. In fact, that was one of the reasons behind the LC participation in the present institute. It has also been a factor in a current effort to communicate information about developments in the CONSER project through the *LC Information Bulletin* and through *JOLA*.

Ms. Rather also tried to answer some of the specific questions Mr. Fasana raised. Question: Who is in charge and what are their credentials? Answer: The Council on Library Resources (CLR) is the manager of the project for a period of two years. George Parsons, who has been in the field of library automation for a number of years, is the coordinator. Richard Anable, who initiated much of the work on CONSER, is now with CLR and is also associated with the coordinating function. Question: What is the status of the CONSER format? Is it compatible with the MARC format for serials? Answer: There is no CONSER format *per se*. The CONSER group will use the MARC serials format. Changes required in the format to accommodate the needs of the participating libraries have been submitted to MARBI for consideration; if approved, they will be added to the format.

Question: What is the time schedule for implementation of the project? Answer: This depends on when an agreement is reached and a contract is signed with OCLC. If this contract is signed within the next several weeks, the project is scheduled to begin in December.* Question: What are the objectives of CONSER? Answer: As stated in the preliminary document submitted to CLR, the purpose of the project is twofold: (1) to establish a comprehensive serials data base in a timely manner in order to avoid a redundancy of cataloging, and (2) to insure that the resulting data base meets high standards of quality by having the Library of Congress and the National Library of Canada review and authenticate the records that have been input. Question: Is CONSER creating a new standard? Answer: CONSER is trying to build a data base in such a way as to accommodate existing standards for format, description, and identification. It is not trying to create a new standard.

Question: Has CONSER rejected the concept of automated authority control? Answer: No. In looking at the future, LC will begin to distribute authority records in 1975 and thus simultaneously will begin to build its own automated authority file. In reviewing the choices available to CONSER participants, it was felt that it would be better to use a postedit-

* The contract was signed by OCLC and CLR on December 17, 1974. Partial implementation of the project began in early spring 1975, and full implementation should occur in mid-summer 1975.

ing procedure of data against manual files rather than to create a rival authority system or to delay the start of CONSER until an authoritative file were available. Question: Won't the nature of the records created by a cooperative project such as CONSER be disparate? Answer: In point of fact, the file created will be much like NUC or NST in that some records will be LC records and others will be contributing library records. Question: Why should it be necessary to create 200,000 to 300,000 records within two years? Answer: There are over one million serials records currently in machine-readable form. This is an attempt to improve the usability of these records by applying standards to them.

Joseph Howard clarified some apparent misconceptions about the current cataloging standards efforts. The dropping of rule 162B of the *AACR* was prompted by questions raised by CONSER participants as to how conflicts between ISDS, ISBD(S), and AACR could be resolved, but it was not dropped solely for the benefit of CONSER. Insofar as the problem of choice of entry is concerned, no decision has been made as yet as to how to resolve the existing conflict among the standards.

Ruth Tighe, assistant director for field operations for NELINET and chairman of the MARBI committee, was asked to be among the panel participants in order to give an additional perspective to the issues discussed during the session. Ms. Tighe noted that, in some ways, the institute left her with a feeling of disappointment. It was regrettable, she observed, that no member of CLR was present at any of the sessions of the institute. She expressed disappointment that Paul Fasana did not make a stronger case, noting that many of his points were well taken. She felt that the point of view of the speakers was somewhat shortsighted. For example, form and choice of entry will no longer be crucial issues, she observed, as online data becomes retrievable through an increasing variety of access points. Finally, she decried the preoccupation of the panel participants with immediate and pragmatic problems. She felt a better balance could have been achieved if some effort had been made by the panelists to place the discussion in a broader, more philosophical content.

John Byrum, chairman of the Catalog Code Revision Committee, was asked by the panel to comment on current work of the committee. In responding to this request, Mr. Byrum noted that the committee is very interested in receiving reactions from librarians on rule 6 and chapter 7 of the *AACR*, as well as on the *ISBD(S)*. Mr. Byrum noted that no specific revision of rule 6 had been submitted to the committee to date. Instead, after discussions were held on rule 6 within the committee at the 1974 ALA Annual Conference in New York, it was determined that an open meeting on the problem should be held at the 1975 ALA Midwinter Meeting in Chicago. Following this session, the committee plans to meet and draw up whatever revision, if any, should be considered by the Joint Steering Committee. Mr. Byrum emphasized that it is the Joint Steering Committee which has the final responsibility for the *AACR* code.

Mr. Byrum also noted that although the *ISBD(S)* has been approved by the Catalog Code Revision Committee in principle, it has not been endorsed by the Joint Steering Committee to date, but would be considered by that group at the 1975 ALA Midwinter Meeting. If the *ISBD(S)* is endorsed, a redraft of chapter 7 will be assigned to the editors of the North American and British editions of the *AACR*. Mr. Byrum emphasized that such a redraft will be considered in terms of the entire *AACR* and not just a single chapter.

In addition to these specific interests, Mr. Byrum also noted that the Catalog Code Revision Committee is presently discussing whether the *Paris Principles* on which the *AACR* is based should continue to be supported by the committee, or whether other tenets, such as those on which Universal Bibliographic Control is based, should be followed. The status of the whole *AACR* code, and not just rule 6 or chapter 7, could be affected by these discussions.

During the question-and-answer sessions, the topic of locations and/or holdings data in CONSER was considered. Under the present concept, the CONSER effort will include locations (the geographical locations of each serial title), but will not include holdings (information as to which specific issues of a serial title are held by a library). Both Lois Upham and Cynthia (Pugsley) Durance noted that the aim of CONSER is to build a clean bibliographic data base which could then be used for many things. In addition, the problem of standardization of holdings information is under consideration by a CONSER subcommittee headed by Velma Veneziano of Northwestern University. There is no reason that holdings information could not be added at a later date once the technique is formulated and adopted, and there is no reason that location information could not be extracted from the CONSER records as the National Library of Canada is intending to do.

Hank Epstein, from Stanford University, expressed the concern that other on-going serial projects should not be forgotten in the effort to push CONSER ahead. Specifically, he requested that (1) the documentation on the internal functions of CONSER should be made available to groups other than those participating directly in the project, and (2) that nonparticipants should be able to take advantage of the CONSER cataloging data base as it is being built, rather than having to wait until the end of the two-year projected time period for the project. Lucia Rather noted that it was quite probable that the documentation would be made available to all who were interested, but recommended that Mr. Epstein should make known his feelings on the distribution of the data base to the manager of the project, CLR.

Ralph Shoffner, of Richard Abel & Co., in noting a sense of disappointment at the sameness of the statement of the serials problems and solutions as had been given ten years ago, queried the panel and the audience for hard cost data on CONSER and other serial projects. Although no data could be given on CONSER, some cost data were given on two serial projects in California by Linda Crismond: at the San Francisco Public Library, costs average $1,200 per month for updating of the catalog and daily check-in functions of serials; at the University of Southern California Library, costs for catalog maintenance of a serial list, including daily and monthly updates, average $1,500 per year, while costs for printing the list add another $1,500 per year.

In a similar vein, Larry DeVos from Ball State University, stressed the need for cost analysis on the variety of options in the handling of serials. He concurred with Mr. Fasana that cataloging ideas need to be related to the current base in order to maintain consistency. Later, after the alternatives and costs are better understood, will be time enough to experiment.

REFERENCES

1. United States, Library of Congress, MARC Development Office, *Serials: A MARC Format* (2d ed.; Washington, D.C.: 1974). 104p. + Addenda.
2. United States, Library of Congress, MARC Development Office, *Books: A MARC Format* (5th ed.; Washington, D.C.: 1972). 106p. + Addenda.
3. United States, Library of Congress, Information Systems Office, *Serials: A MARC Format* (Preliminary ed.; Washington, D.C.: 1970). 72p. + Addendum.
4. International Conference on Cataloguing Principles, Paris, 1961. *Statement of Principles: Adopted at the International Conference on Cataloging Principles,* Paris, October 1961. (Annotated ed.; London: International Federation of Library Associations, Committee on Cataloging, 1971).
5. American Library Association, Division of Cataloging and Classification, *A.L.A. Cataloging Rules for Author and Title Entries.* (2d ed.; Chicago: American Library Assn., 1949). 265p.
6. *Anglo-American Cataloging Rules.* North American Text. (Chicago: American Library Assn., 1967). 400p.
7. UNISIST International Serials Data System, *Guidelines for ISDS* (Paris: 1973). 58p. + Appendixes.
8. International Federation of Library Associations, Joint Working Group on the International Standard Bibliographic Description for Serials, *ISBD (S)—International Standard Bibliographic Description for Serials* (London: 1974). 36p.
9. International Federation of Library Associations, *ISBD (M), International Standard Bibliographic Description for Monographic Publications* (1st standard ed.; London, 1974). 36p.
10. Ottawa, National Library, Canadian MARC Office, *Canadian MARC Communication Format: Serials,* 1v. (Ottawa: 1975).

TECHNICAL COMMUNICATIONS

ANNOUNCEMENTS

Council on Library Resources Grant to Library of Congress will Fund Expansion of MARC Services

The Council on Library Resources, Inc., and the Library of Congress have announced the issuance of a grant from the council to the library for the expansion of the bibliographic services now provided to the nation's libraries through automated means by the Library of Congress. The grant, in the amount of $106,132, will fund a pilot project and two additional studies at the Library of Congress.

Pilot Project for Certification by the Library of Congress of LC Source Records Converted by Other Organizations—Libraries throughout the country are using LC/MARC data as input to their automated systems. Since the scope of MARC language coverage for bookform material is still limited, libraries in many cases encode and transcribe LC cataloging copy derived from LC printed cards, proofsheets, and entries in the National Union Catalog. Librarians have been aware of the advantages to be gained if a central agency, the Library of Congress, could accept these records in machine-readable form, remove the duplicate records, compare the records with the LC Official Catalog, update them for consistency when required, and redistribute them through the MARC Distribution Service. An important benefit of processing machine-readable records of other institutions at the Library of Congress is the possibility of automatically posting the symbols of the holding library to the National Union Catalog Register of Additional Locations, which is also in machine-readable form.

Since the library has a serious space problem and cannot house the staff required to process large retrospective data files, the project will be concerned only with prospective cataloging data. The Library of Congress and the Council on Library Resources will select participants in the project on the basis of the completeness of the data content of their records and their adherence to MARC encoding. Participants selected will receive the certified records at no cost during the period of the pilot project; other libraries will be able to purchase the records through the MARC Distribution Service apart from the current subscription service.

Of the eighteen months that the pilot project is intended to last, six months will be devoted to the hiring and training of conversion staff, analysis of other organizations' data files, record formats, and design procedures, and the writing of new computer programs or the modification of existing ones. The remaining twelve months will be for the production effort.

The library invites an expression of interest in participation from any organization willing to contribute its records. Interested persons should write to Henriette D. Avram, MARC Development Office, Library of Congress, Washington, DC 20540; the library will respond with further details. The number of participants will be limited by the space available for the operational staff to handle the total volume of records.

Requirements Study—The Library of Congress has for some time been involved with the design and implementation of a Core Bibliographic System. This system takes into account the services needed by outside libraries as well as the technical processing requirements for the Library of Congress. Technical staff have not been available for the full-time analysis needed to define the requirements for a system of this magnitude. The Council on Library

Resources grant will fund a consultant who, together with Library of Congress staff, will define the hardware and communications requirements for the system.

National Union Catalog Format—Since the initiation of the MARC Distribution Service, there have been discussions concerning the possibility of standardizing a subset of the MARC format. Several studies have been conducted and the conclusions have always been the same, i.e., that for the purpose of national distribution, where there are a multitude of users and uses, there is no standard subset. For the purpose of reporting titles in machine-readable form to the National Union Catalog, however, it would be possible to define a less complete record than the present MARC record.

Although the Library of Congress is not prepared at this time to accept such records in machine-readable form for the production of NUC in hard copy or microform or to make these records available in an on-line mode, a record definition now would provide the guidance needed for those libraries that are automating and cannot convert their bibliographic titles to the full MARC format. The resulting NUC record could be used to report locations to the Register of Additional Locations and eventually to report titles for NUC. This work will draw heavily upon the work performed at the National Library of Canada in defining a reporting format for the Canadian National Union Catalog.

CLR and OCLC Sign Agreement on CONSER Project

The Council on Library Resources (CLR) and the Ohio College Library Center (OCLC) have signed an agreement providing for use of OCLC's computer network in the CLR-managed CONSER (CONversion of SERials) project over the next two years. The CONSER project, which should be operational during 1975, is a cooperative effort to convert serials records to a machine-readable format and will result in a relatively comprehensive serials data base.

The file produced by CONSER will be available to the library community through the Library of Congress and the National Library of Canada. The participating institutions will input their serials via CRT terminals on OCLC's on-line shared cataloging system.

Institutions invited as initial participants in the CONSER project include the Library of Congress, the National Library of Canada, the National Library of Medicine, the National Agricultural Library, the State University of New York, the New York State Library, Cornell University, Yale University, the University of Minnesota, and the University of California. The International Serials Data System will participate through the two national centers: the National Serials Data Program within the Library of Congress, and the ISDS/Canada within the National Library of Canada.

A more extensive report on CONSER was made to the Association of Research Libraries meeting in Chicago in January. The March issue of *JOLA* contained an article by Richard Anable entitled "CONSER: An Update."

RESEARCH AND DEVELOPMENT

Computer System to Interact Directly with Library Users

Designing a computer system to interact directly with library users in the solution of individual information problems is the aim of a University of Denver project funded by the Exxon Education Foundation.

A team of three DU faculty members has received a grant of $57,075 to design, implement, and evaluate a system integrating language and resources in the field of sociology.

The project will be conducted at DU's Penrose Library under the direction of Patricia Culkin, assistant professor and technical services coordinator, Ward Shaw, associate professor and associate director of the library, and Thomas Drabek, professor and acting chairman of the sociology department.

Recognizing that information resources, viewed on a community wide basis, are extremely diversified and that language variabilities act as a deterrent to coordination of the resources, the project directors will use computer interface technology to bring users and resources together.

The researcher will interact with the computer system, exploring ways of best expressing his problem and integrating his language with language to which library tools respond. The end result will be an individually tailored guide to resources to direct his research.

The project involves identifying relevant resources, designing the system that will relate resources and language, programming, testing, and evaluating the effectiveness and impact of the system.

The DU project will give primary attention to transferability of concept and design.

Notes from the Office of Science Information Service

Recent advances in retrieval of scientific and technical information (STI) are impressive. Nearly forty STI computer-based abstracting/indexing services are operating. Several commercial firms offer on-line searching across dozens of separate bibliographic data bases. Networking is growing rapidly. Scientific publishers are now turning to photocomposition to reduce rising costs and improve the timeliness and quality of their wares. Also, numerical data banks are proliferating.

While greatly aiding users, these advances also bring an ominous warning: the gap between technological advances and theoretical understanding of how to effect information transfer continues to widen. No one should be surprised at this state of affairs. Technology remains the driving force of research and development. Funds for research and development related to communications equipment and electronics (over $3.2 billion in 1972) are exceeded only by those spent on aircraft and missiles. Consequently, we see continued improvements in hardware and software and systems performance, with decreasing costs and improved capabilities. It is not surprising, therefore, that technological solutions are sought for overcoming barriers to use of information.

But applications of communications technologies occur without a clear understanding of how information transfer occurs. Lacking any systematic theoretical base, information system designers, publishers, librarians, and managers of information/dissemination centers continue to improvise in adapting technological improvements to past practices in meeting vaguely perceived information requirements of users. We are building services and networks without understanding the basic processes involved.

OSIS believes a parallel push is needed—one designed to build a coherent and systematic theory of information transfer. Such a framework would enhance integration of knowledge about human information—using behavior and features of formal information systems, including technological components. Consequently, one of the goals of OSIS is to strengthen information science as a scientific discipline. Results should include refinement of concepts and theoretical relationships and development of theoretical frameworks for understanding information transfer. Research directed to these ends is supported by OSIS under its Research Program, headed by Dr. Edward C. Weiss. Emphasis is on basic research and theoretical studies within the discipline of information science.

Today, the OSIS Research Program represents the only major source of support in the U.S. for development of information science theory. Types of research being supported are illustrated by work in progress at several universities. At the Georgia Institute of Technology, under the direction of Dr. Vladimir Slamecka, OSIS is supporting research on the communication of signs and symbols—the basic elements of messages. Results should aid in development of a theory of information transfer by contributing to understanding of how transmitted symbols convey the desired meaning and how the received meaning affects behavior. Other research is under way at Ohio State University. In this case, Dr. Marshall C. Yovits is attempting to develop a general theory of information transfer based on an integration of information theory and decision-making theory. The purpose is to provide a theoretically-based foundation for the design and evaluation of information systems. At the University of California at Berkeley, Dr. M. E. Maron is analyzing human memory and recall functions in order to formulate more general and efficient theories of information storage and retrieval. Results are expected to be useful for development of future computerized information storage and retrieval capabilities.

In addition, OSIS plans to sponsor conferences to identify trends in basic information science research, promising directions for new research, and to encourage the integration of emerging theoretical formulations. Proceedings will be made available to interested individuals. Meanwhile, OSIS is seeking ideas and proposals for research on basic information science problems. Please communicate with Dr. Edward C. Weiss, Program Director, Research Program, Office of Science Information Service, National Science Foundation, 1800 G Street, NW, Washington, DC 20550.

COMMERCIAL SERVICES/ INFORMATION RETRIEVAL AND DATA BASES

Smithsonian Makes Data Base Available On-Line Through System Development Corporation

System Development Corporation and the Smithsonian Science Information Exchange in Washington, D.C. announced that, effective March 4, 1975, the SSIE file of prepublication information covering research in progress will be available for on-line searching through the facilities of the widely used SDC Search Service. The new SDC/SSIE service is designed to give quick, easy, and economical access, from remote computer terminals, to a file of over 125,000 summaries of ongoing research projects in the life and physical sciences.

The SSIE data base is composed of descriptions of individual research projects which include names of the supporting organization, performing organization, and investigators; title; period of performance; and, in most cases, a 200-word technical summary of the work to be performed. The exchange receives input from over 1,300 funding organizations including federal, state, and local government agencies; colleges and universities; nonprofit foundations and associations; and some foreign groups. Projects are indexed by a staff of scientists trained in each of the major disciplines and then stored for retrieval in SSIE's own computerized data base.

SDC notes that the SSIE file content will complement the research-oriented bibliographic data bases it already accesses. These include SCISEARCH (life sciences), CHEMCON (chemistry), POLLUTION, MATRIX (communications and environment, CAIN (agriculture), COMPENDEX (engineering), GEOREF (geosciences), NTIS (U.S. Government research reports), and others.

SSIE is a nonprofit corporation of the Smithsonian Institution whose purpose is to assist in the planning and performance of research through the timely exchange of information on research in progress. SDC specializes in the development and operation of computer systems for government organizations, academic institutions, and commercial firms.

For further information contact: SDC Search Service Action Desk, 2500 Colorado Ave., Santa Monica, CA 90406; (213) 393-9411, extensions 7277 or 7313, or Jan Goldstein, SSIE, room 300, 1730 M St., NW, Washington, DC 20036; (202) 381-5511.

University of Washington World Treaty Center Completes Computerization of Treaty Details

Details of approximately 22,500 world treaties signed between 1920 and 1970 are now available as a result of a major effort to computerize the information.

Completion of the massive world treaty

information system and publication of treaty index volumes was announced recently by the National Science Foundation (NSF), which granted the University of Washington's (Seattle) Treaty Research Center $300,000 to help fund the project. It is expected to be widely used by teaching and research faculty in law, political science, and international studies as well as by government officials and agencies, reference librarians, and students.

Developed by the center under the direction of Dr. Peter H. Rohn of the university's political science department, the treaty information system consists of a computerized data bank at the university, a five-volume *World Treaty Index*, and a companion *Treaty Profiles* volume to be available later this year.

The *World Treaty Index* is believed to be the only computerized index of its kind. It is compiled in machine-readable form to provide computerized access to detailed information on 11,500 treaties in the United Nations Treaty Series, 5,000 treaties in the League of Nations Treaty Series, and 6,000 treaties in separate national treaty collections. The *Treaty Profiles* volume is a study of world treaty patterns and analyzes the significance of treaty trends in the context of current research in international law and world politics.

The *World Treaty Index* volumes were published by the American Bibliographical Center-Clio Press, Santa Barbara, California.

Chemical Information Systems Linked

Computer-based chemical information systems in West Germany and the United States will be linked under an agreement just concluded between the American Chemical Society and Internationale Dokumentationsgesellschaft fuer Chemie m.b.H. (IDC). The long-term agreement, which went into effect January 1, grants IDC the right to use and the exclusive right to distribute in West Germany and West Berlin the publications, services, and computer-readable files of the society's Chemical Abstracts Service Division. IDC in turn will contribute a share of the cost of producing the Chemical Abstracts Service data base equal to the percentage of documents abstracted by CAS that originate in West Germany—presently about 7 percent. IDC also will be represented on the CAS advisory and editorial boards and will consult with the ACS Board of Directors Committee on decisions affecting CAS publications and services.

IDC takes over marketing responsibility for CAS publications and services in West Germany from Gesellschaft Deutscher Chemiker, which has performed that function since January 1970 under a cooperative agreement with ACS. GDCh's Chemie Information und Dokumentation Division in Berlin also is being transferred to IDC. The new arrangement is in keeping with a program being developed by the West German federal government to encourage and support a coordinated network of information systems to serve the West German scientific and technical community.

IDC, which was established in 1967 by a number of major European chemical companies to operate a centralized computer-based information system to serve the needs of the chemical industry, has developed an extensive data base and search system, including the GREMAS system for coding and searching information on the structures, reactions, and properties of organic compounds. IDC has been working in cooperation with CAS since the 1960s on the development of automated structure-handling techniques. The German organization has developed computer programs for deriving GREMAS codes for chemical structures from the structure representations used in the CAS Chemical Registry System and is working on techniques for identifying reaction and properties data in computer-readable *Chemical Abstracts* subject index files.

With implementation of the agreement with ACS, Chemical Abstracts Service files will provide a substantial part of the input for the IDC data base. The IDC system will make possible the first large-scale substructure searching of the contents of the CAS Registry files. It also will represent the first comprehensive automated search system for the CA subject index content.

Editec: Computer-Based Research Service

Editec, the Electric Library, a computer-based information center operating out of Chicago, has announced extensive information retrieval services. The two-year-old corporation serves clients ranging from large corporations and municipalities in the country to scholars and researchers.

Using a minicomputer, Electric Library research assistants tap more than 25 million different documents in twenty-nine data bases. Each base contains information that has already been screened for significance. Once the computer has selected the maximum number of articles on a subject, the bibliography is delivered with complete citations: name of author, source, code data, and in most cases a 300-word abstract that covers relevant concepts and statistics.

Major subject areas include business management and financial news, financial statistics, national and international industry and marketing, government publications, Library of Congress titles, science and technology, education, urban affairs, medicine, psychology, agriculture, life sciences, chemistry, and general information.

The number of bases grows monthly and there has been a 100 percent increase in the available data base information in the last six months. Data bases are updated daily, with current information included in less than six weeks in most bases. Information from the Electric Library may be ordered by phone, by mail, by facsimile, computer terminal, or TWX-TELEX facilities. Information is returned by mail or electronic reception equipment.

Normal turnaround time is one week if request and information are delivered by mail. One-day service is available if delivery time is shortened by use of electronic reception equipment.

The Electric Library offers a contract for 100 searches and will also do individual bibliographic searches. Further information may be obtained by contacting Editec, 53 W. Jackson, Chicago, IL 60604; (312) 427-6760.

Chemical Industry Notes—Data Base

The American Chemical Society now is offering a computer-readable version of *Chemical Industry Notes,* its weekly digest of and index to business and economic news concerning the chemical industry. The number of publications monitored by the service also was increased from 52 to 73 with the issue of December 9, 1974, to provide expanded coverage of energy, the environment, government regulatory activities, and the pharmaceutical, agrochemical, paint, and chemical specialties industries.

Published jointly by the Society's Chemical Abstracts Service Division and its Division of Public Affairs and Communication, *Chemical Industry Notes* is a weekly compilation of brief, informative extracts of news items on production, pricing, sales, plant expansion, new products and processes, corporate and labor activities and executive changes in the chemically-based industries, and government actions affecting the industries. Each issue covers 900 to 1,000 items selected from key businesses, chemical news, and industry trade publications around the world.

The computer-readable version of the service, which is derived from data on tapes used to compose the printed service, contains the full text of the printed extracts. Key words, phrases, and names, which are highlighted and compiled into an index in the printed version, are specifically identified in the machine-readable service.

The computer-readable *Chemical Industry Notes,* which is offered under a lease arrangement, is priced at $2,500 per volume (52 weekly issues) plus handling charges. Computer-readable files for back issues beginning with Volume 3, No. 1 (December 3, 1973), also are available. A sample issue of the computer-readable service may be obtained for evaluation.

For additional information contact Marketing Department, Chemical Abstracts

Service, Ohio State University, Columbus, OH 43210.

POTPOURRI

Center for Communications Management Announces a Mileage Directory to Evaluate Network Line Charges

CCMI announces the availability of a unique mileage directory to assist in private line network evaluation and planning. This 400+ page reference manual, *Mileage Directory: Primary Rate Centers,* incorporates over 125,000 mileage calculations used by the communications common carriers in their private line ratemaking. Published in hardcover format, the volume includes a separate table of airline rate miles between each of the 370 designated AT&T "high density" rate centers, all commercial Western Union Telegraph Company high density rate centers and all high density international boundary (transiting) points. Additionally, these rate mileages are applicable to most of the private line services offered by the specialized communications common carriers.

The basis for the calculation of rate mileage between two cities is frequently an imaginary point (rate center) within a geopolitical boundary established by the communications common carrier. Consequently the "airline mileage" between two rate centers frequently differs (by as much as 10–15 percent) from the anticipated airline mileage between the true geographic center of the cities. In order to accurately determine the rate mileage between two or more points, therefore, it has been necessary to first obtain from the carrier's tariff the appropriate rate center coordinate points (referred to as V and H coordinates), then execute a time-consuming complex formula: e.g.,

$$\sqrt{\frac{(V\epsilon - V_{\mu})^2 + (H\epsilon - H_{\mu})^2}{10}}$$

CCMI's *Mileage Directory: Primary Rate Centers* provides direct look-up of precalculated rate mileages between 125,000+ city pairs. Its use will greatly simplify the network planning and optimization efforts of those system planners and users who are today confronted with frequent rate restructurings by the established carriers and the plethora of alternative services emerging from the specialized communications common carrier. This reference manual is immediately available from CCMI at $125.00 per copy. For additional information, please contact Minor S. Huffman, Center for Communications Management, Inc., P.O. Box 324, Ramsey, NJ 07446; (201) 825-3311.

Checklist of Computer Contract Clauses

An EDUCOM task force chaired by James Poage, EDUCOM institutional representative from Princeton University, has compiled a checklist of terms and clauses used in contracting for the purchase or lease of computing software packages and the development of custom software. Published in January 1975, *Contracting for Computing Volume II: A Checklist of Terms and Clauses for Use in Contracting with Vendors for Software Packages and Custom Software,* by James Poage and Carolyn Landis supplements an earlier checklist of clauses used in contracting for the purchase or lease of computer hardware and operating system software. Published under the title *Contracting for Computing* in September 1973, the earlier checklist is now in a fourth printing.

Increasing numbers of colleges and universities have begun to write individual contracts for obtaining computing resources rather than relying on vendor contracts. *Contracting for Computing* and the sample contract clauses that it contains should prove to be as useful an aid to contract officers as its predecessor has been.

The checklist includes a discussion of 31 types of clauses that ought to be included in most contracts for computer software packages and 29 types of clauses useful in contracting for the development of custom software. More than 285 clauses taken from actual contracts drawn between universities and vendors are included as examples of the clauses identified in the discussion portion of the checklist. Appendixes to the checklist are organized in catalog fashion to make it easier for the user to find clauses of a particular type. In all cases, sources of sample clauses are identified.

One copy of the checklist has been sent to each EDUCOM member institution. Additional copies are available and may be ordered from the EDUCOM Office, P.O. Box 364, Princeton, NJ 08540. The price is $7.50 for EDUCOM member institutions and $15.00 for nonmembers.

Should States Control University Computing?

Should public colleges and universities operate individual computers or should states control computing for higher education?

This question, asked as frequently by university administrators as by state officials, is the focus of the study conducted by Dr. Charles J. Mosmann for EDUCOM with support from the Exxon Education Foundation. The results of that study have now been released in *Statewide Computing Systems: Coordinating Academic Computer Planning* published by Marcel Dekker, Inc., New York. This hardcover book is a timely analysis of statewide coordination of computing for higher education, a phenomenon which is becoming more widespread each year. This movement will have a significant impact on campus decisions and on the way in which resources are decided upon and supplied to colleges and universities in the future. Dr. Mosmann's study concludes: "Only a cooperative effort will do. Only government's insistence will insure results and only real involvement by campus representatives will guarantee a resource sharing scheme which has promise of educational benefit and to which the campus users may commit their trust . . . Multicampus computing may become the model for the future development of multicampus education, research, administration and service."

Dr. Mosmann's study began with a questionnaire circulated early in 1973 to all state chancellors of higher education and to officials in state governments. Responses were received from each state and from each of the Canadian provinces. Additional information was obtained through interviews at a number of state capitals and universities.

In October 1973, a conference of representatives from education and state government considered results of the survey and the preliminary conclusions posited by Dr. Mosmann. *Statewide Computing Systems: Coordinating Academic Computer Planning* reports on the outcome of the survey and the conference, gives a number of case studies, and presents Dr. Mosmann's conclusions. Also included in the volume are papers contributed by state and university officials who participated in the conference, descriptions of various state plans, and the results of the survey of the states and provinces.

Specific aspects of the study were also discussed at two EDUCOM conferences, fall 1973 and spring 1974. Proceedings of these conferences and the study report are available from EDUCOM Publications, P.O. Box 364, Princeton, NJ 08540. Faculty and staff from EDUCOM member institutions are entitled to special prices on each of these reports. The prices are: **Statewide Computing Systems:* $9.50 members, $13.75 nonmembers; *Facts and Futures,* Proceedings of the fall 1973 conference, $5.00 members, $9.00 nonmembers; *Computing and the Decision Makers:* Proceedings of the 1974 EDUCOM conference, $6.00 members, $9.00 nonmembers.

STANDARDS

American National Standard Committee Z39 X/C 34 on Code Identification of Serial Articles, Draft Code Proposal February 25, 1975

* Review copies of *Statewide Computing Systems* are available on request.

FOREWORD

(This foreword is not part of the proposed American National Standard Code Identification of Serial Articles.)

This standard was prepared by Subcommittee 34 of the American National Standards Institute (ANSI) Committee Z39 on Standardization in the Field of Library Work, Documentation, and Related Publishing Practices. Standards Committee Z39 is organized under the procedures of the ANSI and is sponsored by the Council of National Library Associations.

Various centers have independently accumulated masses of data on journal articles in machine-readable files. Each center has organized its data base but sometimes without consideration for procedures used elsewhere. With the advent of networks for bibliographic information interchange in this area, standardization of journal article identification becomes a prerequisite for successful selection and communication of that information between centers. Otherwise, a specialized method is needed to acquire and process data from each node in a network. Such special methods are to be avoided since they are costly to create and operate, involve a multitude of terms and treatments, lengthen response times, and impair overall efficiency. Further, in the melding of data from different bases, it is necessary to be able to eliminate duplicate references or data element content by machine effort.

While it is possible to build an identification code that would uniquely designate an article, fashioning of a brief code which will provide for identification of a high percentage of serial articles seems justified. Any code so devised would normally be used in conjunction with other existing information (author, title) when uncertainty might arise.

Possible uses of the code identification described in this standard include: (a) exchange of machine-readable data bases; (b) means to eliminate duplication in a machine-readable data base; (c) aid in operation of automated document supply systems; (d) query of files for specific records and for verification; and (e) as a bibliographic strip identification in serials = Bibident.

Suggestions for improvement of this Standard are welcome, and should be sent to the American National Standards Institute, 1430 Broadway, New York, NY 10018.

Subcommittee 34—On Code Identification of Serial Articles, which prepared this standard, had the following members: **Pauline Atherton,** School of Information Studies, Syracuse University; **Anita De Vivo,** American Psychological Association; **Constantine Gillespie,** National Institutes of Health Library; **Lawrence Livingston,** Council on Library Resources; **Harold Oatfield,** Chairman, Formerly, Pfizer Medical Research Laboratories; **Elizabeth Sawyers,** State University of New York at Stony Brook; **Robert Tannehill, Jr.,** Chemical Abstracts Service; **Samuel Waters,** National Agricultural Library; and *Consultant:* **George A. Parsons,** Council on Library Resources. Subcommittee 34 was organized in March 1972.

AMERICAN NATIONAL STANDARD IDENTIFICATION CODE OF SERIAL ARTICLES

1. *Purpose*

This standard provides guidelines for the generation of a code for unique identification of each informational item, e.g., article, published in the serial literature in order to achieve effective information interchange between machine-readable files. It is applicable primarily to automated operations making use of the Z39.2 Standard Format for Bibliographic Information Interchange on Magnetic Tape. For that reason, it is kept brief in the interest of economic computer usage. The code identification is not a replacement for a standard bibliographic reference citation.

2. *Scope*

2.1 The standard is intended to provide means for accurate identification of each informational item published in a serial, regardless of the discipline involved or its subject matter. The standard provides for all variations and combinations of bibliographic data required for identification, such as date, volume, issue, and pagination. It does not include identification of authors or titles of articles. While this standard is primarily designed to identify a specific serial article, the Code may also be used for abstracts, editorials, letters, communications and advertising when these items are to be treated separately.

2.2 This standard is predicated upon the mandatory use of ANSI Standard Z39.9-1971 Identification Number for Serial Publications and ANSI Standard Z39.2 Format for Bibliographic Information Interchange on Magnetic Tape.[1] It cannot be applied to a serial lacking an assigned International Standard Serial Number (ISSN).

2.3 The Code's fixed fields will result in a unique identifier most of the time. Total uniqueness is not considered a practical goal, and thus is not specified. In rare instances manual review will be needed to distinguish between items that carry duplicate codes.

3. *Definitions*

Article. Any separate item of textual information appearing in a serial issue that has, or requires, an identity unique from all other items, whether or not it comprises original material. The term excludes title pages, tables of contents, indexes, and other front and back matter of the serial that go to complete the mechanics of its publication.

Beginning Page Designator. Alphameric designation of the first page of an article.

Data Element. The smallest discrete unit of information or data that can be defined. In machine systems an information unit which is uniquely identified either by fixed field location or by specific tagging in the case of variable length data.

End Designator. A special character used only to mark the end of a coded complete entry.

First Order Designator. Data (numeric or otherwise) supplied by a publisher which provides the main or primary sequential identification for the individual volume or issue of a serial title. Usually this is the volume number but it may be any first order designation that divides the serial into identifiable units, e.g., issue number where no volume number is given.

International Standard Serial Number (ISSN). ISSN is an 8-character code (7 numeric digits plus an 8th check character) that provides unique, permanent, concise, unambiguous identification of a serial publication. ISSN assignment is controlled under central authority of the International Serials Data System, International Center, Paris, France.

Issue. A distinct sequential portion of a serial sent to subscribers by a publisher.

Issue Date. The date printed on each separate serial issue within a publication sequence. "Issue date" does not include coverage dates, dates of meetings, or dates of publication. In the absence of such "issue date," the date of publication should be used.

Issue Identification Data. See Second Order Designator.

Page Fraction. An indicator that specifies location of a particular article on a page when more than one article begins on that same page.

Piece. See Issue.

Second Order Designator. Those data (numeric or otherwise), e.g., issue number as in publications which do not employ consecutive pagination, needed to distinguish among documents that carry identical first order designators. Second order designators divide a first order designator into logical subunits. It identifies the secondary or further division of the publication sequence of a title.

Serial.[2] A publication in successive parts bearing numerical or chronological designation and intended to be continued indefinitely at regular or irregular intervals. Serials include periodicals; newspapers; annual reports, yearbooks; journals, memoirs, proceedings, transactions and similar publications of societies; and numbered monographic series.

Volume Identification Data. See First Order Designator.

4. *Constructions of the Code for Identification of Serial Articles*

The serial article identification code consists of five data elements in a 26-character fixed field format plus a variable length field for explanatory text. The elements identified after the code initials SCISA are the *serial title* (represented by ISSN), *issue date, first order designator, page* on which an article begins, and *end designator*. A distinction is made between data that do not exist and data that may be available, but are not coded or are supplied in a different form. When data are missing, their places are filled with zeros in the designated fixed field. When any element is present but is not being coded, or is being supplied in a different form, its fixed position will be filled with blanks, a flag will be placed in the End Designator data element, and the minimum information needed to complete the unique identification will be inserted in clear text into the variable field provided. The flag for this code will be the dollar ($) sign.

5. *Data Elements*

Field	*Elements*	*Characteristics*
1	International Standard Serial Number	*Fixed* field, nine alphameric characters.
2	Issue Date	*Fixed* field, eight numeric or alphameric characters, zero filled.**
3	First Order Designator	*Fixed* field, four alphameric characters, right justified, zero filled.**
4	Beginning Page Designator	*Fixed* field, four alphameric characters, right justified, zero filled.**
5	End Designator	*Fixed* field, one special character.
6	Explanatory Text*	*Variable* length field, alphameric characters.

5.1 *Code Initials.* The letters SCISA (Standard Code Identification for Serial Articles) and an asterisk shall begin the Code when it is displayed, but are not input into its computer base.

5.2 *International Standard Serial Number,* as assigned to the serial in the International Serial Data System, is an eight-digit number, with a modulus eleven (11) check digit in the rightmost (low-order) position. In the instance in which the modulus eleven (11) check digit would be 10, an X is used instead. The letters ISSN are not input to a computer base but the hyphen separating the two groups of characters is. When displayed, the ISSN always appears preceded by ISSN and is written in two groups of four digits, separated by a hyphen. Examples:

ISSN 1234-5679	ISSN 2357-948X

5.3 *Issue Date.* Eight Arabic numerals, appearing four, two and two as follows: YYYYMMDD for year, month, and day. The first four from left represent the year: for example, 1974, 1897, 1156, etc. The fifth and sixth positions represent the months of the year, January through December, from 01 to 12. The last two digits, 01 to 31, number the days of the month.

Unit	*Symbol*	*Contents*
Year	YYYY	Full designation in Arabic numerals
Month	MM	01 to 12 Jan.-Dec.
Day	DD	01 to 31

Example: 19010312
for March 12, 1901.

If full information for the date is not present on the piece, the appropriate position in this data element is zero filled.*** If any date appears in a form that will not

* Optional

** This field will be blank filled when explanatory text must be provided in field 6.

*** Systems which utilize less information than the full date element (YYYYMMDD) should supply available data in the appropriate positions and fill the remainder of the data element with blanks.

be, or cannot be accommodated in the field provided for it, fill that field with blanks. If data are to be furnished in a different form, fill that field with blanks, place a flag character ($) in field 5 End Designator, and enter a minimum explanatory note in field 6. When more than one month is used to date a piece, as in some bimonthly and quarterly serials, the MM positions of the date element will be blank-filled, and the specified months entered appropriately in field 6. In certain cases use of issue number or other second order designator will suffice. See examples 4, 7 and 8, p.160.

5.4 *First Order Designator*. Four characters, numerals and/or alphabetic, right justified, zero filled. If the first order designator is missing from the piece, this data element will be zero filled (see example 8, page 160). *It is recommended that Roman numerals be converted to Arabic for this field only*. Optionally, the Roman numerals can be given in the Explanatory Text.

5.5 *Beginning Page Designator*. Four characters, Arabic or Roman numerals and/or alphabetic characters, right justified, zero filled. If this designator is greater than four (4) characters, fill the field with blanks, place a flag character ($) in field 5, and enter the data in the Explanatory Text field 6. If pagination does not exist in the piece, this data element will be zero filled.

5.6 *End Designator*. The end of this code will be marked by the character (#). It will appear in field 5 unless a flag has been entered there indicating that the Explanatory Text field 6 has been used to clarify one or more of the preceding fields. In these cases the flag ($) character will be inserted in this field 5, and an explanation in clear text will follow in field 6, with the End Designator character (#) being placed at the end of the Explanatory Text field 6.

5.7 *Explanatory Text*. This portion of the Identification Code is a variable length, alphameric field, containing the MINIMUM TEXT needed to explain any exceptions described above, or to provide additional information deemed appropriate for clarity. This field will be used principally in printed lists to aid users, and will not normally be searched in automated systems. The data in the variable field will be preceded by the related fixed field number and a blank, e.g., "ø2b̸Christmas," and the field will always end with the End Designator (#). The ($) is also used to indicate provision of additional information about a serial article that is not encoded in the pertinent fixed field, or whenever inclusion of further information is needed for clarity. In some cases more than one flag ($) and explanation may be offered in clear text, differentiated by using the relevant field designators.

In addition to the inclusion of date, first order designator, and page data that will not fit into the fixed fields, the two most common types of data that will be entered in this field are Page Fraction and Second Order Designator data (Issue Identification Data). The different data elements entered in this field are always represented by the character "$" preceding each appropriate fixed field number.

5.7.1 *Issue Date*. When issue date information is in a form not provided by the piece it will always be preceded by "$ø2b̸," followed by the data.

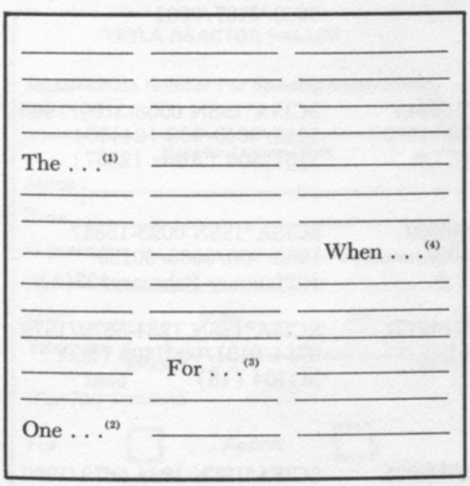
5.7.2. *First Order Designator*. This data element includes the first order designator and associated modifier for a particular publication. The element will always be preceded by "$ø3b̸," followed by the data enclosed in parentheses. Captions used with the first order designator (e.g., "Vol.," "Part," "Section") are not required. Those captions associated with the modifier should always be included, but in abbreviated form. Examples:

	Record as:
Number 3, Supplement a	$ø3b̸(3, Suppl. A)
Supplement to issue 8	$ø3b̸(8, Suppl.)
Spring Issue	$ø3b̸(Spring)
Issue 7, Part II Supplement	$ø3b̸(7, Pt. II, Suppl.)

5.7.3. *Page Data*. The data include

page information that cannot be formatted into the fixed field, or are supplied as supplementary information. In field 6 the entry will always be preceded by "$ø4b̸" and the data enclosed in parentheses.

5.7.4. *Page Fraction Identifier.* This information will always be preceded by "$ø4b̸" followed by the data in parentheses. When more than one article begins on a given page, the article being designated must be uniquely identified by a page fraction identifier. This specificity will be achieved by beginning numbering with the first article in the upper, left-hand corner of the page. That article is numbered "1," and the remaining articles are then numbered sequentially "2" through "n," by proceeding to number down each column and to the right. Example:



When alternatives exist for locating an article on the page, the most direct identifier should be used, i.e., an item number rather than column or page fraction designation.

The format for the *Page Fraction Identifier* is that of this example:

$ø4b̸(3)

6. *Representation of the Identification Code in Printed Materials*

The Code when displayed will begin with its initials SCISA and an asterisk. Its fixed field data elements will be separated by a slash (/). In this way the acronym ISSN, which customarily appears in the display for that element, will be recognized here to be just one of the component parts of the Code. The Code initials will be separated from the letters ISSN by an asterisk.

The *ISSN* will be displayed as ISSN 1234-5679. The *date* will be in its recommended form, e.g., 19730726. The *Explanatory Text* will be written in the same format as it appears in the machine version. The blank (b̸) character following the field number in the explanatory text need not be displayed. The display of the end character is not required. As an example:

SCISA*ISSN 0002-9769/19730726/
0004/0285$03 (5)$04 (09)

7. *Examples*

The following examples illustrate specific data elements and use of the full Code.

Data Element	*Data on Piece*	*Machine-Readable Format*			*Display Format*
		Fixed Field ø1	*Field ø5*	*Variable ø6*	
ISSN	ISSN 1234-5679	1234-5679			ISSN 1234-5679
		Fixed Field ø2			
Date	July 26, 1973	1973ø726			/19730726/
	1973	1973øøøø			/19730000/
	June 1973	1973ø6øø			/19730600/
	Spring 1972	1972b̸b̸b̸b̸	$	ø2b̸Spring	/1972/$02Spring
		Fixed Field ø3			
Volume	No. 5348	5348			/5348/
Identification	Volume 138	ø138			/0138/

Data Element	*Data on Piece*	*Machine-Readable Format*		*Display Format*
Data	Volume 3	ØØØ3		/0003/
	Part 4A	ØØ4A		/004A/
	Volume XL	ØØXL		/00XL/
	Volume 16, Part A	ØØ16 *Fixed Field Ø4*	$ Ø3b̸(PtA)	/0016/$03(PtA)
Beginning Page	123	Ø123		/0123/
Designator	146A	146A		/146A/
	iv	ØØiv		/00iv/

FULL CODE

Data on Piece	*Machine-Readable Format*	*Display Format*
1. Advances in Chemistry Series 89. Isotope Effects in Chemical Processes. 1969, page 65, ISSN 0065-2393	ØØ65-23931969ØØØØØØ 89ØØ65#	SCISA°ISSN 0065-2393/1969 0000/0089/0065
2. Annales Academiae Scientiarum Fennicae, Series A, IV. Biologica. 187, 1972, page 1 ISSN 0066-1988	ØØ66-19881972ØØØØØ 187ØØØ1#	SCISA°ISSN 0066-1988/1972 0000/0187/0001
3. Biological Abstracts, Volume 49, No. 24, December 15, 1968, page 11572, Abstract Number 126867 ISSN 0006-3169	ØØØ6-316919681215ØØ49 b̸b̸b̸b̸$Ø3b̸(24)$Ø4b̸11572 $Ø4b̸(Abstr 126867)#	SCISA°ISSN 0006-3169/1968 1215/0049/$03 (24)$04 11572$06 (Abstr 12867)
4. Kinetics and Catalysis. Volume 3 Number 1, January-February 1962, page 13 ISSN 0023-1584	0023-15841962b̸b̸ØØØØØØ 3ØØ13$Ø3b̸(1)#$02b̸January-February $03b̸(1)#	SCISA°ISSN 0023-1584/ 1962 00/0003/0013$ $02January-February$03(1)
5. The New York Times. Vol. CXXI, No. 41,810, Section M, page 27 New York, Friday, July 14, 1972, item number 15 °ISSN 1234-5679	1234-5679/19720714Ø121 ØØ27$Ø3b̸(Sect M) $Ø4b̸(15)#	SCISA°ISSN 1234-5679/1972 0714/0121/0027$03 (Sect M)$04 (15)
6. Osaka Shiritsu Daigaku Iyaka Zasshi. Vol. 9, No. 12 (Supplement 9), December 1960, page 5721 °ISSN 1234-5679	1234-56791960Ø12ØØØØØ95 721$Ø3b̸(12 Suppl 9)#	SCISA°ISSN 1234-5679/1960 1200/0009/5721$03 (12 Suppl 9)
7. Sulphur Institute Journal. Volume 8, Numbers 1&2 Winter-Spring, 1972, page 10, item number 2. ISSN 0039-4904	ØØ39-49Ø41972b̸b̸b̸b̸ØØØ 8ØØ1Ø$Ø2b̸Winter-Spring $Ø4b̸(2)#	SCISA°ISSN 0039-4904/ 1972 /0008/0010$02 Winter-Spring$04 (2)
8. Library of Congress Catalog Motion Pictures and Filmstrips July-September 1972, page 89 item nr 21 ISSN 0041-7807	ØØ41-78Ø71972b̸b̸ØØØØØØØØ 89$Ø2b̸July-September $Ø4b̸(21)#	SCISA°ISSN 0041-7807/ 1972 00/0000/0089$02 July-September$04 (21)

° ISSN not assigned—sample number

REFERENCES

1. The corresponding international standards are ISO 3297-1974 International Standard Serial Numbering; ISO 2709-1973 Format for Bibliographic Information Interchange on Magnetic Tape.
2. American National Standards Institute. American National Standard for the Abbreviation of Titles of Periodicals. N.Y.C. 1970. 11p. (Z39. 5-1974)

TESLA Reactor Ballot

Photocopy this form to respond to standards proposals. (For details see previous issues of *JOLA*-TC, vol.7, No.3 and 4.) Return ballots to John C. Kountz, Associate for Library Automation, Office of the Chancellor, The California State University and Colleges, 5670 Wilshire Blvd., Suite 900, Los Angeles, CA 90036.

TESLA REACTOR BALLOT

Identification Number For Standing Requirement ______________________

Reactor Information

Name ______________________

Title ______________________

Organization ______________________

Address ______________________

City __________ **State** ________ **Zip** ________

Telephone (____) ____________ - ______
AREA EXT.

Need
(For This Standard)

For ☐ **Against** ☐

Specification
(As Presented in This Requirement)

For ☐ **Against** ☐

Can You Participate In The Development Of This Standard ______________________

Yes ☐ **No** ☐

Reason For Position: (Use Format Of Proposal. Additional Pages Can Be Used If Required)

INPUT

To the Editor:

I have followed the recent controversy arising from the rapid development of CONSER with mixed emotions. On the one hand, the positive benefits of an active catalyzing agent in the area of automated library services is most welcome; on the other hand, many of the concerns voiced by Paul Fasana (*JOLA*, Dec. 1974) transcend CONSER and are cogent observations which must be considered in national planning.

In this light, it seems worth special notice of the article on CONSER in the *L.C. Information Bulletin* of February 14, 1975. This discussion by Larry Livingston of the Council on Library Resources is the most frank and open discussion of the problems encountered in developing an automated library service I have seen in print. ARL, the council, and the *L.C. Information Bulletin* are all to be commended on such an excellent formal communication of current events which previously were common knowledge only in a very limited circle of the library community. If the recent controversy has stimulated such good communication, it will have justified any doubts and misgivings it may have generated.

One can only hope that other general-circulation library journals will provide sufficient space for discussion of such issues of major interest to all libraries.

Brett Butler
Butler Associates

BOOK REVIEWS

Research Libraries and Technology. By Herman H. Fussler. The University of Chicago Studies in Library Science. Chicago: The University of Chicago Press, 1973. 91p. $5.95.

The purpose of this book is "to relate technological potential and limitations to at least some of the major current trends, problems, and objectives of the library." The author concentrates on developments expected to occur in the next five to ten years and focuses "primarily upon the problems of the large, university, research-oriented library." The book was originally a report to the Sloan Foundation on the state of the art in library technology, the foundation having established a program in educational technology. The source of information in the book is largely fifty-nine published documents listed in the reference section. In addition, a score of senior librarians and investigators read and commented on a draft of the report. Personal correspondence with directors of various library computer application projects produced further information for the author. The main value of the book, however, is not in the information gathered and reported but rather the evaluation of that information by the author, a man of outstanding experience and wisdom.

The first chapter reviews major studies of computer technology related to libraries. The second chapter examines data on library cost trends and rates of library growth together with other observations. It is much to the author's credit that in this chapter the tables for the most part contain annual averages of compound percentage growth rates rather than average annual percentage increase most often used, but which always yields too high a figure. Chapter three examines the basic problems of bibliographic control, chapter four looks at three methods for extending resource availability, and chapters five and six discuss computerized library operations and present brief descriptions of eight major computer applications.

The last chapter contains the author's conclusions which are many, valuable, and highly compressed. In general, he finds that "although the probability appears high that most large libraries will make increasing use, directly or indirectly, of computer-based data processing, the mechanisms and structures for such utilization are much less clear"—a sage observation that will be appreciated by many developers of library computerization who find themselves using new concepts and producing new techniques that they did not anticipate but a few years back.

The book is an accurate, cautious, and an important evaluation of the state of the art of library computerization with an estimate of some of the events to occur in the next five years.

Frederick G. Kilgour
Ohio College Library Center

Handbook of Data Processing for Libraries, by Robert M. Hayes and Joseph Becker. Second Edition. Los Angeles: Melville Publishing Co., 1974. 688p. $22.50.

There are few books in our field that deserve a second edition and even fewer that sell in sufficient quantities to receive one. Obviously this work has been well received, for here, after a lapse of only three years, we are presented with a thinner (by 200 pages), only slightly more expensive (by a mere $2.50), revised edition of the most thorough compendium of facts in its field.

Despite the revision, however, it remains very much the book that the first edition was. The same number of chapters appears, each with the same title as its predecessor, and the majority of the information contained therein remains much as before. It has been updated in many places, to be sure, but any reader of the first edition will feel quite at home with the second. Many of the complaints made by Veaner in his review in *JOLA* (Vol. 4, No.4) have been corrected. For example, Veaner noted that much of the tabulated data was either in error or inapplicable to library use. Most of these tabulated data have been removed from the book. This should offer somewhat longer life for the second edition than for the first, since the lists of machines, complete with model numbers and prices, were made obsolete almost immediately upon (or even before) publication.

While this complaint of Veaner's has been answered, however, another has not. "One wishes for more specifics drawn from the real world," said Veaner. This volume offers, in reply, a section on existing automated systems; yet in its efforts to be specific it becomes both less timely than a journal article (of OCLC the authors say, "The technical processing subsystem is due for implementation in 1973.") and less comprehensive than a complete survey. The latter fault is especially evident in what purports to be a survey of Canadian library automation, which lists an Ontario project that ended in 1964 yet ignores completely the University of Saskatchewan, an early and active MARC user and one of the most active libraries in Canada in its use of the computer, and omits the Université Laval, whose on-line serials check in has been widely reported since its implementation in 1968.

To be sure, this volume is more up to date (the latest citation is 1973) and even more willing to attempt to put current projects into perspective: in a section on Computer Usage Trends and System Configurations, mini-computers are explored in some depth. Yet one is still left with a feeling that the work is both too complete in some ways and too superficial in others. Perhaps it is worth repeating Veaner once more, who wrote in 1971, "For the next edition, some consideration might be given to a two-volume work: the first volume for the administrator, and the second containing much more technical detail for the practitioner. If the two volume pattern is followed, a loose-leaf format with regular updating would be most helpful for the second half."

Peter Simmons
University of British Columbia

Radials Bulletin: Research and Development—Information and Library Science. Edited by John Bailey. No. 1- . London: The Library Association, 1974- . Published three times a year. Annual subscription £5.

This publication is a register of ongoing research in library and information science being carried out in the U.K. It has developed out of the similar register which appeared from 1968 to 1973 in the Library Association Year Book, though it does not confine itself to members of the Association. The information given about each project comprises its title; the name(s) of the research worker(s); the institution under whose auspices the work is carried out; the degree to be awarded, if it is academic research; the duration of the project; source of financial support; a brief description of the research; and a note of any publication(s) associated with the work. The entries are arranged under broad subject headings in order of the Classification of Library and Information Science, and each entry is given an accession number indicating the year and issue in which it appeared. A chain-procedure type of index (similar to the one in *LISA*) provides alphabetical access by subject, and a name index of research workers and responsible corporate bodies, a very useful feature of which is the inclusion of the address. Future issues will be cumulative in briefly mentioning all current projects, though perhaps referring to another issue for full details.

The information included in the entries was mostly supplied by research workers in answer to a questionnaire and is sufficient to give an adequate impression of the project, sometimes in detail. Virtues particularly desirable in a bulletin of ongoing research are that it be comprehensive and up to date. Looking, by way of example, at how research presently is being carried on at City University, London, is represented in the *Bulletin,* one sees that work several years ago complete (no longer current) on information transfer among chemists and biologists is reported, while work presently on-going such as of Janet Pope on relevance and Peter Noerr on graph structures in relational indexing is not reported. Also it is somewhat confusing that so few of the research projects cited in the *Newsletter* of the British Library (No.1, September 1974) are represented in the *Bulletin*. It is ungenerous, of course, to generalize from isolated instances. But the instances exist.

As to physical layout, the *Bulletin* suffers from the same minuscule print used in *LISA*, which (though clear in itself) is not easily read by readers of advancing age or deteriorating vision. It would be useful too, if the typography could highlight the title of each project, which at present is not distinguished from the main body of the entry. The first issue contains 163 entries on eleven pages, and one wonders what the rate of increase will be in future issues, in view of the intention to cumulate.

Geoffrey Pendrill
Elaine Svenonius
The University of Western Ontario

The Art of the Librarian; A Collection of Original Papers from the University of Newcastle upon Tyne. Edited by Alan Jeffreys. Newcastle upon Tyne: Oriel Press, 1973. 190p.

This collection of papers has an ostensible cause and an underlying philosophy. Its publication is in honor of Joan Gladstone on her retirement as head of the Order and Accession Department of the University Library; consequently the essays have to do with collections and collection building (however tenuously), and all the contributors have a Newcastle connection. At the same time, Newcastle has long been a pioneer in the field of library automation—UK MEDLARS, the Newcastle file handling system, and two major books, all antedating even MARC I. But like so many pioneers, Newcastle has also been very much a "loner," concerned with its own explorations to the point of exclusion of cooperation with others. Perhaps as a result of these two influences, the papers vary greatly in substance and interest, from a typically high level argument from Maurice Line in favor of locally responsive university library collections backed by a national depository, to a personal reminiscence of libraries that somewhat disjointedly leads to a commendatory paragraph like something from a letter of recommendation (though nonetheless sincere). Among the papers in between are a lucid essay on the value of nonbook media in libraries and their effect on book budgets; a painstakingly detailed record of the physical measurement of books and space on shelves; a predictable description of the Newcastle computerized ordering system, and its apparently improved Manchester nephew; an interesting but bare analysis of user success in on-line MEDLARS searches; a rather whimsical account of the application of the Applepie system to the library of a historic grammar school; a frankly speculative note about the value of computerized amalgamation of book subscription lists; a surprisingly reactionary view of professional and subject knowledge in libraries, which seems to place librarians (even, or perhaps *especially*, computerized librarians) in the category of plumbers' mates; and a charming if nearly irrelevant summary of the travels of seventeenth-century Benedictine librarians.

If the art of the librarian is to be a magpie, a serendipitous gatherer of treasures or trifles whose value will lie in the judgment of the scholar (and at one time this was the librarian's art), then this collection is aptly named; otherwise, for this reviewer at least, it remains a haphazard if interesting collection of partial views.

David Batty
McGill University

INSTRUCTIONS TO AUTHORS

Scope. The *Journal of Library Automation* publishes scholarly papers and technical reports containing findings not published previously in the following fields: research and development in library automation, including inter-library communications; research in information science and educational technology directly related to library activities; and the history and teaching of these subjects. Each paper is accepted with the understanding that it is to be published exclusively in the *Journal of Library Automation,* unless some other specific arrangement is made in advance.

Mailing address. Manuscripts should be mailed to Susan K. Martin, *Journal of Library Automation,* General Library, University of California, Berkeley, CA 94720.

Organization of papers. Reference to earlier issues of the *Journal of Library Automation* will be helpful in preparation. Whenever possible, material should be presented in classical form: 1) introduction, which should set the scene by referring to pertinent related literature and showing how findings in the paper differ from those already in the literature; 2) methods and materials, presented in sufficient detail that others can adapt or adopt them, but omitting information which is generally known or which can be covered by citation of another paper; 3) results, or new data produced; 4) discussion, containing an examination of results with respect to previous findings, or quality, or suggested further investigation, etc.; 5) conclusion, containing evaluation of the significance of the findings or development.

If the paper is a technical report, section 2) is often a detailed description of equipment and procedures of a new system; section 3) reports operation of the system in terms of time to install, cost to install and operate, quality of product, performance, compatibility, flexibility, simplicity, etc.

Papers are edited to improve the effectiveness of communication between author and readers, particularly to remove ambiguities. When extensive editing is necessary, a paper will be returned to the author for correction and approval before type is set.

An indicative abstract of 100 words or less must accompany the manuscript on a separate sheet. It should include purpose, method, results and conclusions, and any specialized information. It should be intelligible without reference to the article, while assuming that the reader has a general knowledge of the subject. Clarity is essential, but complete sentences are not necessary.

Preparation of copy. Manuscripts should be prepared as neatly as possible in the form used by *Journal of Library Automation.* Papers should be submitted in two copies and typed in double space.

The title page shall carry authors' names in full and their affiliations.

References should be typed in double space on a separate sheet. They should be listed in order of citation in the text and numbered consecutively. Citations should be made in the following form:

1. Lois L. McCune and Stephen R. Salmon, "Bibliography of Library Automation," *ALA Bulletin* 61:674–94 (June 1967).
2. ALA Library Administration Division, comp., *Library Statistics of Colleges and Universities, 1965–66 Institutional Data* (Chicago: ALA, 1967), p. 3.
3. Robert R. Freeman and Pauline Atherton, *AUDACIOUS; An Experiment with an On-Line, Interactive Reference Retrieval System Using the Universal Decimal Classification as the Index Language in the Field of Nuclear Science* (New York: American Institute of Physics, UDC Project, 25 April 1968), p. 36. AIP/UDC-7; PB-178 374.

Financial support, and credit for materials, technical assistance and advice, may be cited in a section headed "Acknowledgments," which should appear at the end of the text.

Illustrations. Illustrations must not be larger than 8½ by 11 inches; a larger illustration should be photographically reduced on a glossy print. Line cuts will usually be reduced for printing to 5 inches or less in width, or 7½ inches or less in length. The artist should be informed of the anticipated reduction before a drawing is made. Illustrations should be done in india ink or black typewriter ribbon on white paper. Illustrations should be numbered lightly with soft pencil on the back, and numbered legends typed on a separate sheet.

Tables. Each table should be typed on a separate sheet, numbered, given a title, and cited in the text. Each column should have a heading. Vertical rules may not be used. Tables wider than 76 characters when set must be run endways.

Chargeable costs. Charges will be made whenever special typography (e.g., tables and formulas), corrections in galley proof, etc., cause composition costs to exceed standard text-page composition costs by more than 10%. In the event that an author makes any changes from copy after galley proof has been submitted to the printer, the entire cost of changes will be charged to the author.

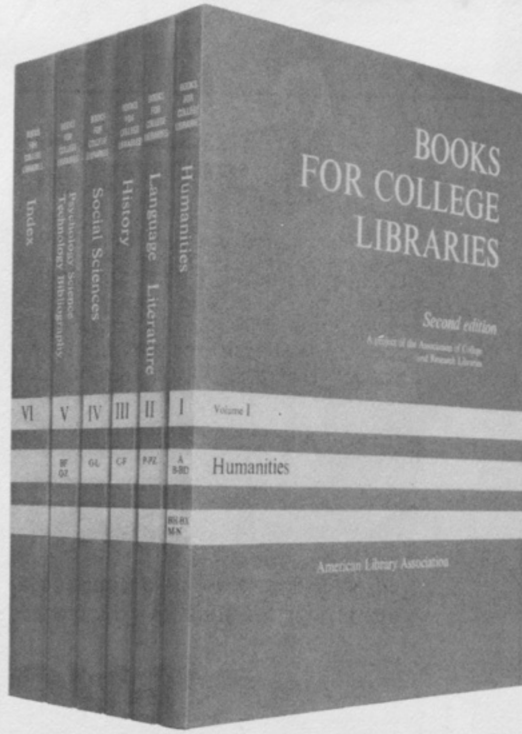
Page proof will not ordinarily be sent to an author.

From reviews of the first edition:

"We cannot think of a single library, from the largest to the smallest, that will not profit from its constant use."
—*A.B. Bookman's Weekly*

"The quality of the titles selected in every class is very high... the scope of the list is engaging... The compilers and their consultants are to be congratulated."
—*Library Journal*

$65.00 ISBN 0-8389-0178-6



6 volumes · almost 40,000 titles · over 2,000 pages

BOOKS FOR COLLEGE LIBRARIES

SECOND EDITION

Compiled by the Association of College and Research Libraries, a division of ALA, with the support of the Council on Library Resources.

Six handsome paperbound volumes:

1. Humanities, 2. Language and Literature, 3. History, 4. Social Sciences, 5. Psychology, Science, Technology, and Bibliography, 6. Author, Title, and Subject Indexes

For this Second Edition, nearly 40,000 titles—four-fifths of those recommended by the 1959 ALA standards as the minimum for a four-year undergraduate library—have been selected and evaluated by subject specialists. Selections have been made from titles in the first edition and from works published between 1964 and 1972. An even greater breadth of sources has been used than before and even more stringent criteria have been applied to selections for this core collection.

Cataloging and classification information has been expanded. Titles are arranged according to Library of Congress classification and are entered in the main catalog only once. Each entry is a virtually complete MARC record: author, title, edition, imprint, collation, subject and other added entries, classification number, LC card number, and ISBN.

AMERICAN LIBRARY ASSOCIATION
Order Department, 50 East Huron Street, Chicago 60611

Lost Horizon . . . Homer's Odyssey . . . or The History of Civilization Vol.II.

Any book can be requested, located and delivered in minutes with a requisition system utilizing Electrowriter handwritten communications equipment.

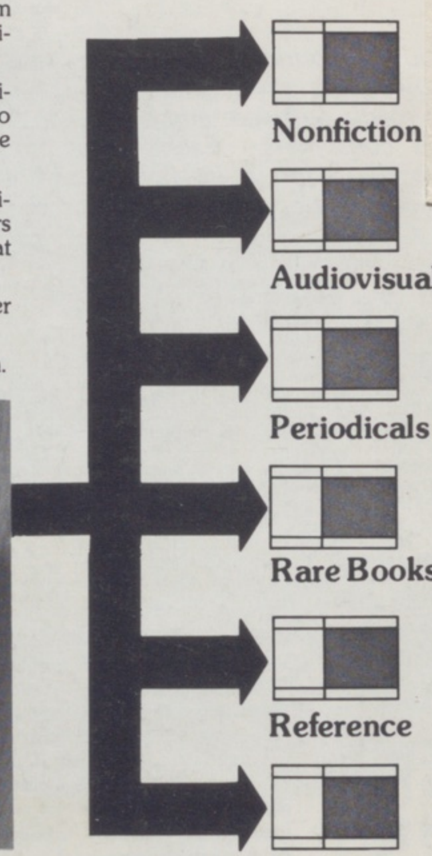
With an Electrowriter system, written requisitions are transmitted from the front desk to the appropriate department for immediate action. Instantly. Accurately.

An Electrowriter system is well suited for library use because it eliminates verbal errors and provides a written, permanent record at all locations for tight control.

And all for as low as a dollar per day per station.

To find out more, fill in and mail the coupon.

Front Desk



VICTOR GRAPHIC SYSTEMS INC

Scanatron Facsimile/Electrowriter Systems
3900 North Rockwell Street, Chicago, Illinois 60618
Subsidiary of Victor Comptometer Corporation

Yes, please send me more information on Electrowriter requisition systems for libraries.

Name ____________________

Title __________ Library __________

Address ____________________

City ________ State ________ Zip ________