TECHNICAL COMMUNICATIONS

ANNOUNCEMENTS

New COLA Chairman

Brian Aveney, of the Richard Abel Co., has been elected Chairman of the COLA Discussion Group, effective January 1974.

Prior to his present position with the Design Group at Richard Abel, Mr. Aveney was head of the Systems Office at the University of Pennsylvania libraries.

The COLA Discussion Group traditionally meets on the Sunday afternoon preceding each ALA conference. Meetings are open, and all are invited to attend.

And A Book Review Editor

A member of the University of British Columbia Graduate School of Library Science faculty, Peter Simmons, has been appointed Book Review Editor of the *Journal of Library Automation.* Mr. Simmons is the author of the "Library Automation" chapter in the *Annual Review of Information Science and Technology,* volume 8, the most recent of his publications.

Authors and publishers are requested to send relevant literature to Mr. Simmons at the Graduate School of Library Science, University of British Columbia, Vancouver, British Columbia, for review.

Missing Issues?

The rapid publication sequence of the 1972 and 1973 volumes of the *Journal of Library Automation* has created problems for some ISAD members and subscribers. If your address changed during 1973, or if your ALA membership suffered any quirk, you are especially likely to have missed one or more of the issues due you. If this is the case, please write to the Membership and Subscription Records Department of the American Library Association, 50 E. Huron St., Chicago, IL 60611. Indicate which issues you are missing, and every attempt will be made to forward them to you as quickly as possible.

New ERIC Clearinghouse

Stanford University's School of Education has been awarded a one-year contract by the National Institute of Education (NIE) to operate the newly-formed ERIC Clearinghouse on Information Resources under the direction of Dr. Richard Clark.

The new Clearinghouse will be part of the Stanford Center for Research and Development in Teaching.

The Clearinghouse on Information Resources is the result of a merger of two previous Clearinghouses—the one on Media and Technology formerly located at the Stanford Center for Research and Development in Teaching, and the one on Library and Information Sciences formerly located at the American Society for Information Science in Washington, D.C.

The new Clearinghouse is responsible for collecting information concerning print and nonprint learning resources, including those traditionally provided by school and community libraries and those provided by the growing number of technology-based media centers.

The Clearinghouse collects and processes noncopyright documents on the management, operation, and use of libraries, the technology to improve their operation, and the education, training, and professional activities of librarians and information specialists.

In addition, the Clearinghouse is collecting material on educational media such as television, computers, films, radio, and microforms, as well as techniques which are an outgrowth of technology—systems analysis, individualized instruction, and microteaching.

LIBRARY AUTOMATION ACTIVITIES—INTERNATIONAL

Computerized system at the James Cook University of North Queensland library.

The system design phase of an integrated acquisitions/cataloging system for the library at the James Cook University of North Queensland has been completed by a firm of computer consultants, Ian Oliver and Associates, and programming has commenced.

History

The system, known as CATALIST, is a batch system to be operated on the university's central computer, a PDP-10. It will be programmed in FORTRAN and MACRO, the assembly language of the PDP-10.

Description

The system will cover all aspects of cataloging/acquisitions procedures for all library material apart from serials including:

(a) production of orders, followups, reports
(b) budget control
(c) fund accounting
(d) routing slips
(e) accessions lists
(f) in-process and catalog supplements (author/title and added entry) and subject catalog supplement shelflist and supplement
(g) catalogs (author/title and subject)
(h) union catalog cards.

Some features of the system include the maintenance of average book price in all subject areas. These are continually updated by the system to reflect the current fluctuations in the trade.

This information will be used together with machine-based arrival predictions to control the budget and fund allocations. MARC data will be used as much as possible, with records for individual items being supplied from external sources on request.

The In-Process Catalogues, which will contain items on order, items arrived, and items cataloged since the previous edition of the catalog, will contain added entries for all material where such information is available.

The catalogs will be produced on COM. Roll film will be used for public catalogs and fiche for in-house use. Data for the National Union Catalogue will be submitted on minimally-formatted computer-produced cards.

For further information contact Ms. C. E. Kenchington, Systems Librarian, Post Office, James Cook University of North Queensland, Australia 4811.

TECHNICAL EXCHANGES

Editor's Note: The two following articles, prepared by the Library of Congress and the Council on Library Resources, respectively, have been distributed through various LC publications. Due to the importance of the two documents, however, and to the fact that they may not have reached the entire library community, it seemed therefore appropriate to publish the papers again in Journal of Library Automation.

Sharing Machine-Readable Bibliographic Data: A Progress Report on a Series of Meetings Sponsored by the Council on Library Resources

Beginning in December 1972 and continuing since that date, the Council on Library Resources has convened a series of meetings of representatives of several organizations to discuss the implications of bibliographic data bases being built around the country and the possibilities of sharing these resources.

Although the deliberations are not yet completed, the Council, as well as all participants in the meetings, felt that it was timely to make the progress to date known to the community. Since publication in the

open literature implies a long waiting period between completion of a paper and the actual publication date, it was decided that this paper should be written and distributed as expeditiously as possible. Since the Library of Congress has vehicles for dissemination of information in its MARC Distribution Service, *Information Bulletin,* and *Cataloging Service Bulletin,* LC was asked to assume the responsibility for the preparation of a paper to be distributed via the above mentioned channels as well as sending copies to relevant associations. The institutions participating in the deliberations have been included as an appendix to this paper.

The bibliographic data bases under consideration at individual institutions contain both MARC records from LC as well as records locally encoded and transcribed. These local records represent: (1) titles in languages not yet within the scope of MARC; (2) titles in languages cataloged by LC prior to the onset of the MARC service; (3) titles not cataloged by LC; and (4) titles cataloged by LC and recataloged when the LC record cannot be found locally. The first two categories, in many instances, are being encoded and transcribed by institutions using LC data as the source, i.e., proofsheets, *NUC* records, and catalog cards. These are referred to for the remainder of this paper as LC source data and the third and fourth categories as original cataloging.

All participants agreed that the structure of the format for the interchange of bibliographic data would be MARC but several participants questioned if a subset of LC MARC could not be established for interchange for all transcribing libraries other than LC.[1, 2] Although LC had reported its survey regarding levels of completeness of MARC records and the conclusions reached by the RECON Working Task Force, namely, "To satisfy the needs of diverse installations and applications, records for general distribution should be in the full MARC format," it appeared worthwhile to once more make a survey to see if agreement could be reached on a subset of data elements.[3] The survey included only those institutions participating in the CLR meetings. The result of the survey again demonstrated that considered collectively, institutions need the complete MARC set of data elements. The decision was made that the LC MARC format was to be the basis of the further deliberations of the participants.

Attention was then turned to any additional elements of the format or modifications to present elements that may be required in order to interchange bibliographic data among institutions. All concerned recognized that although networks of libraries, in the true sense, still do not exist today, much has been learned since the development of the MARC format in 1968.

Certain ground rules were established and are given below:

1. The material under consideration is to be limited to monographs.
2. The medium considered for the transmission of data is magnetic tape.
3. Data recorded at one institution and transmitted to another in machine-readable form is not to be retransmitted by the receiving institution as part of the receiving institution's data base to still another institution.[4]
4. Any additions or changes required to the MARC format for "networking" arrangements are not to substantially impact LC procedures.
5. Any additions or changes required to the MARC format for "networking" arrangements are not to substantially affect MARC users.

Long discussions took place concerning modifications to LC source data by a transcribing library and the complexity involved in transmitting information as to which particular data elements were modified. Ground rule 6 was established stating that if any change is made to the bibliographic content of a record copied from an LC source document (other than the LC call number), the transcribing library would be considered the cataloging source, i.e., the machine-readable record would no longer be considered an LC cat-

aloging record. Any errors detected in LC MARC records are to be reported to LC for correction.

A subcommittee was formed to study what MARC format additions and modifications were required. The subcommittee met on several occasions and made the following proposals to the parent committee:

1. Fixed field position 39 and variable field 040, cataloging source, should be expanded to include information defining the cataloging library, i.e., the library responsible for the cataloging of the item, and the transcribing library, i.e., the library actually doing the input keying of the cataloging data.
2. LC should include the LC card number in field 010 as well as in field 001. When the LC card number is known by an agency transcribing cataloging data, field 001 should contain that agency's control number and field 010 should contain the LC card number.
3. Variable field 050 should not be used for any call number other than the LC call number. Transcribing agencies should always put the LC call number in this field if known.
4. A new variable field 059, contributed classification, should be defined to allow agencies other than LC to record classification numbers such as LC classification, Dewey, National Agricultural Library classification, etc., with indicators assigned to provide the information as to what classification system was recorded and whether the cataloging or transcribing agency provided this data.
5. Variable field 090, local call number, should follow the same indicator system as defined in field 059. (090 contains the actual call number used by either the cataloging or transcribing library while 059 would contain additional classification numbers assigned by the cataloging or transcribing library.)
6. LC would assume the responsibility of distributing any agreed upon additions or modifications as either an addendum to or a new edition of *Books: A MARC Format.*

Discussions following the presentation of these proposals indicated concern regarding three principal areas:

1. The modifications of any data element in an LC source document other than the addition of a local call number dictated that the institution performing the modification of the record assume the position of the cataloging source. This resulted in the possibility that a large number of records would undergo minor changes and consequently the knowledge that the record was actually an LC record would be lost. This loss was considered a critical problem.
2. The creation of a MARC record implied that each fixed field and all content designators should be present if applicable for any one record. During the LC RECON project, it was recognized that certain fixed fields could not be coded explicitly because the basic premise in the RECON effort was the encoding of existing cataloging records without inspecting the book. Consequently, the value of certain fixed fields, such as indicating the presence or absence of an index in the work, could not be known. Participants felt that a "fill" character was needed to describe to the recipient of machine-readable cataloging data that a particular fixed field, tag, indicator, or subfield code could not be properly encoded due to uncertainty. The "fill" character will be a character in the present library character set but one not used for any purpose up to this time.
3. Although networking is not clearly defined at this time, participants felt that the MARC format should have the capability to include location symbols to satisfy any future requirement to transmit this information in order to expedite the sharing of library resources.

Majority opinion indicated there was a

need to guarantee the recognition of an LC source record, that a "fill" character could serve a useful function, and that a method of transmitting location symbols was required. Three position papers were written on the topics outlined above giving the rationale for the requirement and describing a proposed methodology for implementation. These papers were reviewed at a meeting of the participants and are presently undergoing modification taking into account recommendations made.

The revised papers are to be distributed prior to the next meeting in January 1974. Following this meeting, another paper will be prepared for publication which will include a definitive account of the modifications and additions recommended for the MARC format as well as describing the rationale for the additions and modifications. At that time the proposals will be submitted to the library community for its review and acceptance.

If the additions and changes are approved by the MARBI[5] Committee of the American Library Association, LC will proceed to amend or rewrite the publication *Books: A MARC Format*. However, the points elaborated below deserve emphasis toward the understanding of the issues described in this paper.

1. The meetings were concerned with a national exchange of data, not international.
2. The additions and modifications recommended for the MARC format, with one exception, affect organizations other than the Library of Congress exchanging machine-readable cataloging data. Except for distributing records with the LC card number in field 010 as well as 001, the MARC format at LC will remain intact.
3. LC will investigate the use of the fill character in its own records, both retrospective and current, and for records representing all types of materials.

Henriette D. Avram
MARC Development Office
Library of Congress

REFERENCES

1. The MARC format has been adopted as both a national and international format by ANSI and ISO respectively.
2. Subset in this context includes both the data content of the record (fixed and variable fields) and content designators (tags, indicators, and subfield codes).
3. RECON Working Task Force, "Levels of Machine-Readable Records," in its *National Aspects of Creating and Using MARC/RECON Records* (Washington, D.C.: Library of Congress, 1973), p.4–6.
4. This rule did not extend to a subscriber to the LC MARC service duplicating an LC tape for another institution. One can readily see the chaos that would result if institution A sent its records to institutions B and C, B then selected all or part of A's records for inclusion in its data base, and then transmitted its records to A and C. The result of the multitransmission of the same records, modified or not, would create useless duplication and confusion.
5. RTSD/ISAD/RASD Representation in Machine-Readable Form of Bibliographic Information Committee.

APPENDIX 1

List of Organizations Participating in the CLR Sponsored Meetings

Library of Congress
National Agricultural Library
National Library of Medicine
National Serials Data Program
New England Library Information Network
New York Public Library
The Ohio College Library Center
Stanford University Libraries
University of Chicago Libraries
Washington State Library
University of Western Ontario Library

A Composite Effort to Build an On-Line National Serials Data Base (A Paper for Presentation at the ARL Midwinter Meeting, Chicago, 19 January 1974)

An urgent requirement exists for a concerted effort to create a comprehensive national serials data base in machine-read-

able form. Neither the National Serials Data Program nor the MARC Serials Distribution Service, at their current rate of data base building, will solve the problem quickly enough. Because of the absence of a sufficient effort at the national level, several concerted efforts by other groups are under way to construct serials data bases. These institutions have been holding in abeyance the development of their automated serials systems, some for several years, waiting for sufficient development at the national level to provide a base and guidance for the development of their individual and regional systems. This has not been forthcoming, and local pressures from their users, their administrators, and their own developing systems are forcing these librarians to act without waiting for the national effort. These efforts are exemplified by the work of one group of librarians, described below.

What has now come to be known as the "Ad Hoc Discussion Group on Serials" had its beginnings in an informal meeting during the American Library Association's Conference in Las Vegas last June. You will also hear this discussion group referred to as the "Toronto Group." This is because its prime mover has been Richard Anable of York University, Toronto, and because the first formal meeting occurred in that city. The expenses of the Toronto and subsequent meetings have been borne by the Council on Library Resources, and Council staff have been involved in each meeting. A fuller exposition of the origins, purposes, and plans of the Toronto group has been written by Mr. Anable for the *Journal of Library Automation*. It appeared in the December 1973 issue.

Quoting from Anable: "At the meeting [in Las Vegas] there was a great deal of concern expressed about:

1. The lack of communication among the generators of machine-readable serials files.
2. The incompatibility of format and/or bibliographic data among existing files.
3. The apparent confusion about the existing and proposed bibliographic description and format 'standards'." End of quote.

The Toronto Group agreed that something could and should be done about these problems. If nothing else, better communications among those libraries and systems creating machine-readable files would allow each to enhance its own systems development by taking advantage of what others were doing.

As the discussions progressed, several points of consensus emerged. Among them were:

1. The MARC Serials Distribution Service of the Library of Congress and the National Serials Data Program together were not building a national serials data base in machine-readable form fast enough to satisfy the requirements of developing library systems. This systems development was, in several places, at the point where it could no longer wait on serials data base development at the national level as long as progress remained at the current rate.
2. The MARC serials format developed at LC offered the only hope for machine format capability. Every system represented planned to use it. For the purpose of building a composite data base outside LC, the MARC serials format would probably require minor modification, principally by extension. These extensions could and should be added on so as to do no violence to software already developed to handle MARC serials.
3. There existed some difference between the LC MARC serials format and that used by the National Serials Data Program. These differences arose from several circumstances. For example, the MARC serials format predated the International Serials Data System (ISDS), the National Serials Data Program, and the key title concept. When these three came along, the requirement existed that the NSDP abide by the conventions of the ISDS. Since the key title

is not yet a cataloging title, but is the title to which the International Standard Serial Number is assigned, it is natural that the approach to serial record creation by NSDP should be different from that of a library cataloging serials by conventional methods. A working group under the auspices of the IFLA Cataloguing Secretariat has devised an International Standard Bibliographic Description for Serials. The working group's recommendations are to be distributed for trial, discussion, and recommendation for change in February. When the ISBD(S) is accepted into cataloging practice, some of the differences in MARC usage and NSDP procedure will disappear. Others will still remain and they *must* be reconciled. We cannot continue with two serial records, both of which claim to be national in purpose but which are incompatible with each other. A good exposition of the differences in these serials records from the point of view of the MARC Development Office is in an article by Mrs. Josephine Pulsifer in the December 1973 issue of the *Journal of Library Automation.*

4. Major Canadian libraries are active in cooperative work on serials and these two national efforts should be coordinated.

Several other circumstances bear on the problem. For example, the National Serials Data Program is a national commitment of the three national libraries. In addition to the funding from the three national libraries, there are excellent chances that the NSDP will receive funds from other sources to expedite its activities. The NSDP is responsible for the ISSN and key title and for relationships with the International Serials Data System. Ultimately, the ISSN and key title will be of great importance to serials handling in all libraries. For all of these reasons, it is imperative that the activities of the NSDP be channeled into the comprehensive data base building effort described in this paper.

When it was realized at the Council on Library Resources that the Toronto Group was serious and that a data base building effort would result, it was obvious that this had enormous significance for the Library of Congress and other library systems because the result would be a *de facto* national serials data base. Accordingly, a paper was prepared and sent to LC, urging that an effort be made in Washington to coordinate the efforts of the MARC Serials Distribution Service, the National Serials Data Program, and this external effort. In addition, it was felt that LC should take a hard look at its own several serials processing flows and attempt to reconcile them better with each other and with the external effort. To do this, LC was urged to do a brief study of LC serials systems, using LC staff and one person from CLR. LC agreed and the study is now very nearly complete. The written guidance given the study group members was quite specific. They were to study all serials flow at LC and make their recommendations based on what LC should be doing, rather than being constrained by what LC is doing. The overall objectives of the study were to aim for the creation of serials records as near the source as possible and one-time conversion of each record to machine-readable form to serve multiple uses. Specifically to be examined were the serials processing flows of the Copyright Office, the Order Division, the Serial Record Division, New Serials Titles, and the National Serials Data Program.

While all of this was going forward, the Toronto Group had some more meetings. OCLC was tentatively selected as the site for the data base building effort. It is understood by everyone that this is a temporary solution; eventually a national-level effort must be mounted which will provide a post-edit capability to bring the composite data base up to nationally acceptable standards. A permanent update capability is also required. This permanent activity, hopefully, will be based at the Library of Congress. OCLC was chosen as the interim site for several reasons, but especially for its proven capability to produce network software and support

which will work. Within a very short time OCLC will have on-line serials cataloging and input capability which will extend to some two hundred libraries. No other system is nearly so far advanced. The Toronto Group has assured itself that the data record OCLC intends to use is adequate and is now working on the conventions required to insure consistency in input and content, to include some recommendations for minor additions to the MARC serials format.

During their deliberations, the Toronto Group realized that, to be effective, their efforts needed formal sponsorship, and discussions to this end were begun. Initially, several agencies were considered to be candidates for this management role. Various considerations quickly narrowed the list down to the Library of Congress, the Association of Research Libraries, and the Council on Library Resources, and representatives of these three met to discuss the matter further. During the discussions, CLR was asked to assume the interim management responsibility until a permanent arrangement could be worked out. CLR was selected because, as an operating foundation under the tax laws, it can act expeditiously in matters of this kind. CLR can also deal with all kinds of libraries and has no vested interest in any particular course of action.

Meanwhile, certain institutions in the Toronto Group had indicated that they were ready to pledge $10,000 among themselves for the specific purpose of hiring Mr. Anable as a consultant to continue his coordinating activities. The group asked CLR to act as agent to collect and disburse these funds.

CLR is ready to assume the initial responsibility for the management of this cooperative data base building effort, if that is the will of the leadership in the library community. CLR is prepared to commit one staff member full time to the project who is well versed in the machine handling of MARC serials records. This is Mr. George Parsons, and other staff members will assist as appropriate. Mr. Anable has agreed to act as a consultant to help coordinate these activities. CLR would aim for the most complete, accurate, and consistent serial record in the LC MARC serials format which can be had under the circumstances. During the effort, CLR will act as the point of contact between OCLC and the participating libraries, assisting in negotiating contracts and other agreements as required. The composite data base will be made available to all other libraries at the least possible cost for copying. Initially at least, the costs of this effort will have to be shared by the participating libraries, since no additional funds are presently available. The goal is to build 100,000 serial records the first year, another 100,000 the second year, and design and implement the permanent mechanism the third year, while file-building continues.

As the project gets under way, it will work like this: a set of detailed written guidelines for establishing the record and creating the input will be promulgated, and agreement to abide by them will be a prerequisite to participation. Selected libraries with known excellence in serial records will be asked to participate; others may request participation. Those selected who already have or can arrange for terminals on the OCLC system will participate on line. This is the preferred method, but it may be possible to permit record creation off line, such records to be added to the data base in a batch mode. It is very difficult to merge serial files from different sources in this way, so an attempt will be made to find a large serials data base in machine-readable form for use as a starting point. This file would be read into the OCLC system. A participating library wishing to enter a record would first search to see whether it existed in the initial data base. If a record is found, it would be updated insofar as this is possible, within the standards chosen for the system. It may be further updated by other participants, still within the system standards, but at some point update on a record in the system will reach a point of diminishing returns and the record will remain static until a post-edit at the national level can be performed. These records will be for use as their recipients see fit, but their prime purpose is to support the development of automated serials sys-

tems while eliminating duplication of effort.

Details of how to flag these records in the OCLC data base as they are being created by this effort will be worked out, as will be the relationship between this effort and the rest of OCLC activities. CLR will, from time to time, report progress to the community.

It would be the hope of CLR that the Toronto Group will continue to assist in the technical and detailed aspects of the project. In addition, and after consultation with the appropriate people, an advisory group will be appointed to advise CLR in this effort.

Lawrence Livingston
Council on Library Resources

INPUT

To the Editor:

Re: file conversion using optical scanning at Berkeley and the University of Minnesota discussed by Stephen Silberstein, JOLA *Technical Communications, December 1973.*

It is rewarding to find someone who has actually read in detail one's published work (Grosch, A. N. "Computer-Based Subject Authority Files at the University of Minnesota Libraries"). I generally agree with Mr. Silberstein's observations regarding the use of optical scanning for library file conversion. However, several points were raised by Mr. Silberstein on which I feel further comment is needed.

Perhaps in my article I should have cautioned the reader that when developing procedure and programs for the CDC 915 page reader, there is a great variance in these machines depending upon:

1. How early a serial number unit, i.e., vintage of machine,
2. What version of the software system GRASP is being used,
3. What degree of machine maintenance is performed out, and
4. What kinds of other customers are using the scanner.

It was our misfortune to have a CDC 915 page reader that had many peculiarities about it which could or would not be resolved by a maintenance engineer. In addition it was not heavily used and what use it did receive was mostly nonrepetitive conversion jobs dealing mostly with mailing address file creation and freight billing.

In our initial testing we tried to use various stock bond paper and had various reading difficulties. In talking with others who had used this particular machine we found that the choice of paper stock was critical on this scanner. I might add that we did not actually use $400 worth of paper on this as I sold half of the stock we had ordered to another user locally who was going to use this device.

It might be worth mentioning that we had a failure of a potentially large conversion project reported to us. This project tried to use this equipment but could not create a suitable input format because of a specific uncorrected peculiarity of not being able to read lines of greater than six inches without repeated rejects. We were aware of this from our experience which is why we kept our line short using the ⌐¬ to terminate reading of the line at the last character position. Also our input was double spaced, not single spaced as you seem to infer in your comments.

With this particular device we also found that the format recognition line was easily lost, necessitating greater time spent in re-running the job. Therefore, even though this was a great commission of sin on our part according to Mr. Silberstein, I then must plead guilty to using expedient methods to turn a bad situation into an acceptable one. I might also point out that this solution had been employed at various times by some past users we contacted. In fact, I have later found out that occasionally such a technique has been resorted to in one of our other local user installations on a much newer machine.

I do not wish to imply that our conversion achieved maximum through-put but that in any case it was a cost effective way to proceed. With a small file conversion such as this one which is to be done on

a one-shot basis, it seemed foolish to me to spend much time optimizing, but rather to find a way that worked as our difficulties were encountered. If this had to be a continuing job we would have had to get a better maintained scanner and invested more time and money into the project. I take the view that we wish to couple modest human costs with modest projects and reserve for greater projects of a continuing nature more optimized procedures.

I agree that file cleansing is undoubtedly the most costly operation but I cannot say by just what amount since my responsibilities did not include such work. This was later performed by our Technical Services Department.

Our general point in writing about this project was to convey our broad experiences using this technique on a subject authority system as we had not seen such use reported in the literature previously. I would hope your comments and mine here serve to illustrate that one's systems problems must be solved in light of the conditions and not always according to what we term the best theory or practice. To this end I hope others will profit from both of our comments.

Audrey N. Grosch
University of Minnesota Libraries