CHARLES R. SAGE
Coordinator of Automated Library Services
Libraries of the State Universities of Iowa
Iowa State University
Ames, Iowa

Utilization of the MARC II Format for Serials in an Inter-University Environment

This paper deals with a relatively small project undertaken by the libraries of the three state universities of Iowa that standardized their handling of machine-readable serial records using the MARC II format. The first section deals exclusively with programming techniques and conventions employed in the MARC II format. The second section describes procedures, generalization, and compromises that permitted the development of a generalized packaged program to serve three academic libraries. The third section outlines the early work environment with special emphasis placed on the library/data processing center relationship. Current environment is also described and projections are made about the next phase this project will enter. Conclusions drawn from this project bear on future handling of computer applications in libraries.

Programmed System

The initial impetus for developing a machine-readable central serials file incorporating the MARC II format came from a consideration of *Serials: A MARC Format,*[1] a working document by the Information Systems Office of the Library of Congress. At that time, August 1969, the three state university libraries of Iowa were exploring the feasibility of compiling and producing a union list of serials. Each university library was operating independently in recording active titles and corresponding holdings. The University of Iowa and Iowa State University maintained their files on magnetic tape. The University of Northern Iowa maintained a 3 by 5 inch card file. The three systems were

dissimilar except that the two computerized systems were written in 1401 Autocoder and operated on IBM 360s.

We felt adoption of the MARC format would accomplish two objectives: (1) it would serve as an experiment and exercise to expose and familiarize our personnel with the MARC II communications format; and (2) it would provide adoption of a standard format that will conform to national standards and fit into a library computerized technical processing system of the future.

Input to the current system is in two forms. Original data are punched into 80 column IBM cards and input in conventional batch mode. An input card format was designed specifically for this function. (See figure 1b.) The bulk of serials information at the University of Iowa and Iowa State University was information already contained on magnetic tape. Special conversion programs were written, and tape-to-tape conversions were processed.

The master files are maintained in the MARC II communications format. Minor modifications were made relative to variable and fixed length records; however, the working format is the communications format rather than an abbreviation or subset[2] of the communications format.

At the present level of development three output formats are produced. They are: (1) card image of the input alphabetized by title and identified by a local system number (figure 1c); (2) an image of the masterfile in the MARC II format which does not include the "Record Directory" and the "Leader" (figure 2a); (3) selected variable fields in a formal printout for reduction and reproduction (figure 2b).

To adopt the MARC II format in its entirety as an internal systems format did present programming difficulties. In spite of these minor hardships, we find the description adequate for our current needs, and anticipate very few problems when we further develop the capabilities of this system.

Because of the nature and characteristics of our hardware and software, we actually use a fixed length record rather than a variable length record. The IBM 360 and 370 series machines and COBOL-F or USASI COBOL do not lend themselves readily to variable length records. We modified all records to 2048 characters with the variable differential padded with blanks. The variable MARC II record within the 2048-character record is terminated with an ASC II extended 8-bit "End of Record" (1D16) character for recognition purposes. It should also be pointed out that having programmed in USASI COBOL very little use is made of the "Record Directory." Manipulation of data is exclusively controlled by "Field Terminators" (8-bit, 1E16); however, the "Record Directory" is actively maintained and updated.

The use of special characters unique to the MARC II format compounds the difficulty of maintaining data in the EBCDIC form. We allowed three

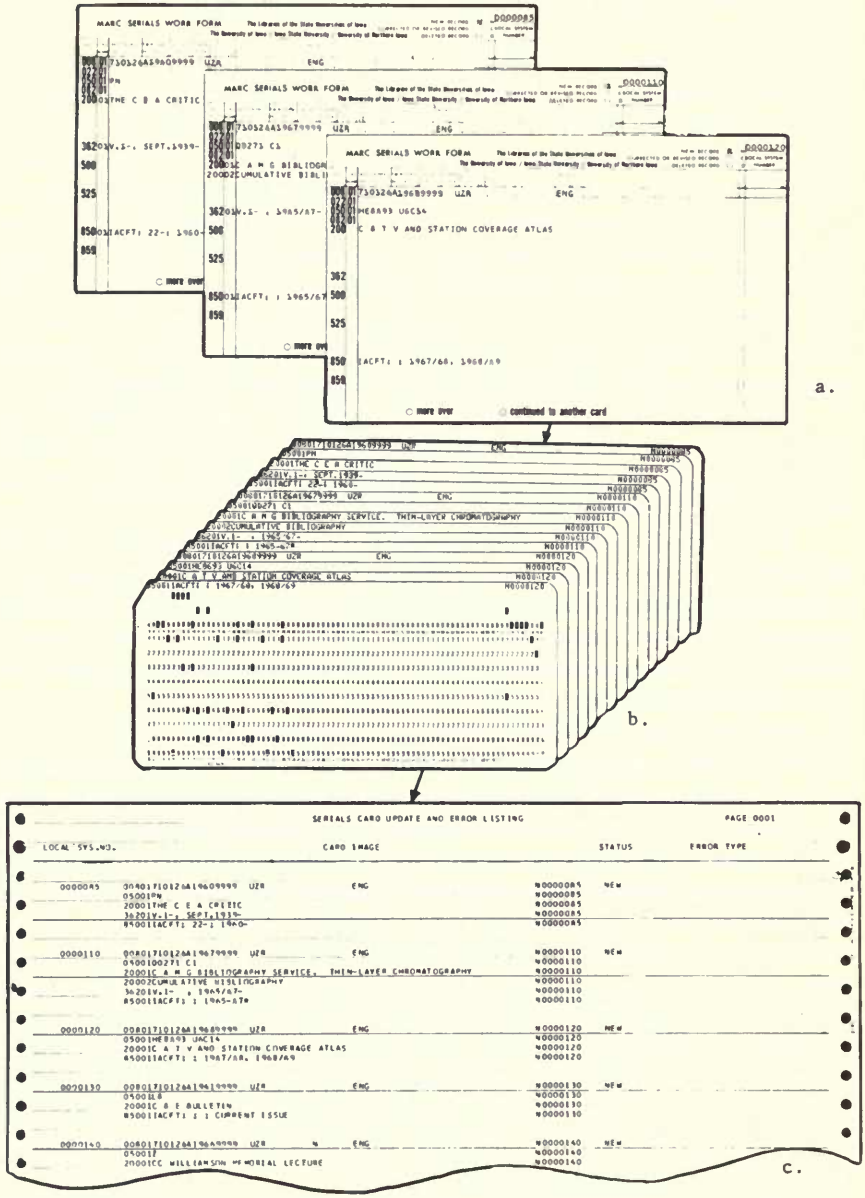
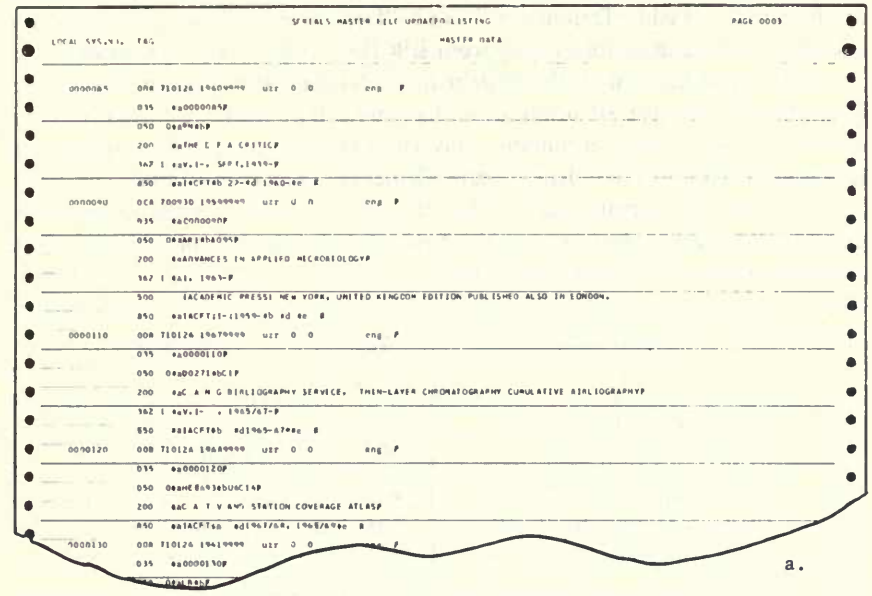
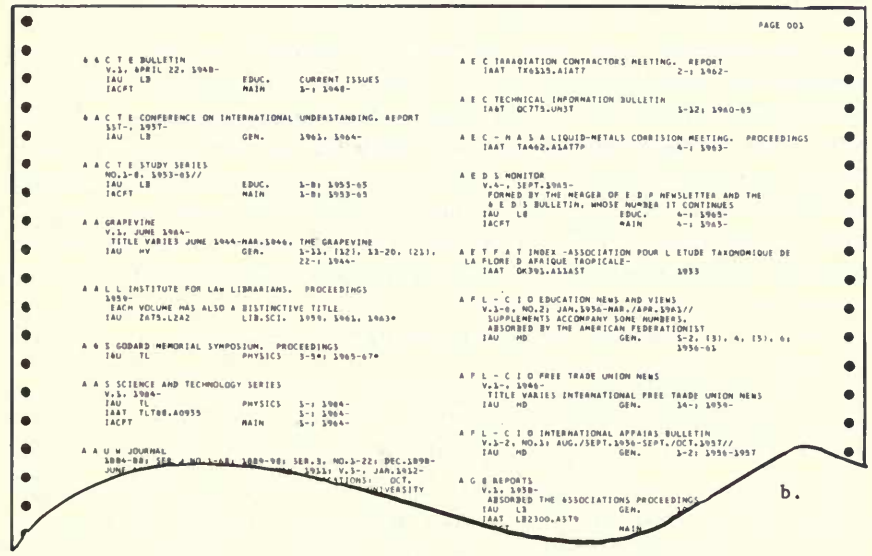


Fig. 1. MARC II Serials Input



a.



b.

Fig. 2. MARC II Serials Output

exceptions to be maintained in 8-bit ASC II form. Those exceptions are: "End of Record," "Field Terminator," and "Delimeter." These characters are exempted from conversions either from EBCDIC to ASC II or conversely.

All input data enter the system in card form. This form of input does limit the effective use of lower case characters to conform to MARC requirements. To compensate, automatic conversion is made internally of upper case to lower case for "fixed length data elements" character groups. The lower case character, comprising half of the "Subfield Codes" of variable fields, are assigned internally when a variable field is being added or updated. Tags currently processed are given in table 1.

Tag	*Description*	*Tag*	*Description*
001	Control number	269	Vendor name and address
008	Fixed length data	350	Subscription price
022	Standard serial code	362	Dates and volumes description
035	Local system number		
050	LC call number	500	General note
082	Dewey decimal classification	525	Supplement note
090	Local call number	850	Holdings
200	Title	857-9	Local holdings

Table 1. Tags Currently Processed

An exception was made to MARC II format with regard to local holdings. We added this tag so that we could report the holdings of peripheral libraries and reading rooms at each campus. The format is identical to the 850 entry except a local mnemonic is substituted for the NUC location symbol. Multiple peripheral libraries and reading rooms for the same titles are recorded in a linear string of characters separated by "Subfield Codes." Each subsequent location is a prefixed with a "±a" "Subfield Code." Any "Other" Tag entry can be input as long as it does not require more than one subfield code.

Packaged Program Concept for Libraries

The most difficult task in designing and implementing a general purpose programmed system is definition of data. Fortunately this task was provided in more detail than we required with the MARC II record format for serials. Adjustments and minor revisions were made to that format as described in the previous section. However, the ease and convenience of implementing a system cannot be overemphasized when this work is already available. The dialog between the systems analyst and serials librarians centered on the librarians interpreting the MARC II definitions. This dialog was extremely productive

when the conversion programs (local format converted to MARC II format) were designed and written at the University of Iowa and Iowa State University. Identifying and tagging data elements, particularly dates and volumes designations, general notes and various configurations of holdings statements, is a difficult task and would have been next to impossible to delineate without the aforementioned dialog.

The MARC II format was also beneficial in attempting to standardize definitions and notations between our three institutions. Details such as incomplete holdings, ceased publication, "ordered on a selected basis," etc., are handled and noted in a uniform manner. Most conflicts or differences of interpretation of MARC II were handled by telephone conferences. In the period of development only two interinstitutional meetings were necessary to clarify differences.

As with any cooperative programmed or manual system, we had our problems. Two major differences were not solved. The first revolved around the procedure for cataloging monographic series. One institution catalogs each entry separately, the other two do not. Individually this problem is not significant; however, in compiling a union list of serials this quantitative imbalance does distort the relative number or holding of the three universities. A second problem was the use of cross references–Iowa State University and the University of Iowa use cross references quite selectively. The University of Northern Iowa, on the other hand, prefers extensive use of cross references. To compound the differences, the MARC II policies for handling cross references is so complex and difficult to maintain that relative benefits obtained are not warranted. A shortcut method was devised for handling cross references different from the MARC II method, but they are not to be included in union lists.

We have been able to exchange bibliographic data between institutions with relative ease. Even though each institution utilizes an autonomous local systems number scheme, exchange, extraction, and reassignment of local systems numbers is easier and more economical than rekeying the desired bibliographic data. Collation and compilation for a union list of serials is available on request. A manual collation and matching is required to generate a union list. However, this is not a large effort (a full-time person for one-half month for a total serials file of 60,000) compared to manually compiling the same file or maintaining a centralized union file of serials.

Program maintenance is handled centrally and does not impose any problems. All program changes are made to all programs and specialized changes for any institution are not done except in special cases such as systems timers, accounting routines, file identification and names, etc. Tables

(locations, symbols, etc.) are a composite of all three institutions and maintained internally in the programs. Job control, tape labels, etc., are handled locally according to the prescribed procedures of the particular data processing center.

Hardware, Software, and Administrative Critique of Project

From the onset of this project, communication with the data processing centers was difficult and we were not really effective in convincing them that we wanted to shed the "Invented Here" syndrome. Our local data processing managers tend to disregard any programming effort that is not created within or supplied by the mainframe manufacturer. However, we mutually agreed this was an experiment and the implication was that there would be no repeat performances.

When the project started we had IBM 360s at all three data processing centers. Two centers programmed exclusively in COBOL and the third programmed in PL/1. Within the IBM repertoire of software, two centers operated under OS (Operating System), one center operated under DOS (Disk Operating System). With those types of hardware and software configurations, fitting our system to each installation was quite easy. Basically the libraries were responsible for programs and the data processing centers were responsible for operations. We jointly documented the system at each data processing center to the local requirements of each center. Since we started operation of our systems, two data processing centers have replaced their IBM 360s with IBM 370s. This change has had no effect on ongoing operations.

The development of the system to its present point has been technically quite easy. Originally our concept was to develop a simple but identical operational system at each institution and then expand applications by delegating modules of expansion to each data processing center. Upon completion and testing of each module, they were to be added to the system as if it were a single system rather than three independent systems. For a number of reasons which will not be detailed here this is neither logistically or administratively workable.

The obvious alternate solution is to continue development at a central source. However, because of budget restraints, this programmed system is in limbo.

Originally plans were to continue development of the serial system to the point that a central serials catalog at each institution would be replaced by a machine-readable file. The basic information to be contained in this file would include full bibliographic data, holdings, minimum of cross references, and added entries of previous changed titles. The system would be interactive

for instantaneous updating and have full use of an expanded upper and lower case character set. With this systems enlargement no longer a reality, scrutiny of the current serials machine-readable file is now being made. The only purpose of this file now is to provide patrons at infrequent intervals (semi-annual and/or annual) with a consolidated book catalog of serials held by each library. The relative benefits of providing such a serials list for patrons versus cost is seriously under question.

The identification and adoption of common standards was made extremely easy through the use of MARC II. Well-selected standards impose a minimum of control and maximum of common advantage. There are even greater benefits to be obtained if we can use generalized packaged programs. Unfortunately our data processing centers have not reached a degree of maturity where developmental schedules and accurate budget projections are demanded or deemed necessary. Also, the application of computers is not widespread enough in libraries that libraries and data processing centers can compare notes with others in similar circumstances and benefit materially from these exchanges.

Penetration of this market will come from two directions once a viable market is realized. Consortia, cooperatives, etc., will offer technical processing services utilizing high speed transmission lines forming regional and national networks. The software industry will offer for purchase or lease standardized packaged programs similar to the hundreds of programs now offered in other fields of endeavor such as Informatics's MARK IV, Applied Data Research's Autoflow, etc.[3] Viability of this market will come to pass only when it will simply be cheaper for libraries to do it this way than through their local data processing centers or by themselves.

REFERENCES

1. U.S. Library of Congress. Information Systems Office. *Serials: A MARC Format*. Washington, D.C., Library of Congress, August 1969-

2. Avram Henriette D. "The MARC Project of the Library of Congress," *Drexel Library Quarterly*, 4:308-09, Oct. 1968.

3. International Computer Programs. *ICP Quarterly*. Indianapolis, 1973.