INTRODUCTION

Many words have been written on the subject of the Machine-Readable Cataloging (MARC) Program: the events that led to the pilot project, the development of the format, the operational Distribution Service, the influence of MARC on standardization, and the impetus it gave to library automation projects and to the creation of networks here and abroad. This article serves to gather together all aspects of the national and international MARC system.

Much of what follows has been fairly well documented in many published reports, journal articles, etc., and therefore this article relies heavily on that material. A Bibliography based on the article’s main headings has been included for those readers who wish to explore any aspect of MARC in greater depth.

BACKGROUND, 1961–1965

The Library of Congress’s (LC) investigations of the possibility of using automated techniques for its internal operations began in the late 1950s. As a result of the interest generated by these investigations, the Librarian of Congress requested a grant from the Council on Library Resources (CLR) for a study to determine the feasibility of applying automated techniques to the operations of LC. The study, published in 1963,[1] recommended that a group be established to design and implement the procedures required to automate the cataloging, searching, indexing, and document retrieval functions. A systems staff was assembled in the Office of the Librarian to proceed with the recommendations. During this same period, motivated by the increasing attention given to automation in libraries, CLR awarded a contract for a study of the possible methods of converting the data on LC cards to machine-readable form for the purpose of printing bibliographical products by computer. The report of the study[2] was the subject of a conference held at LC in January 1965, under the sponsorship of LC, the Committee on Automation of the Association of Research Libraries (ARL), and CLR. Participants included representatives from universities, research agencies, government agencies, and private industry. The conferees concluded that:

1. Availability of machine-readable catalog records produced and distributed by LC would help those libraries that have automated systems.
2. The machine-readable record should include all the data presently available on LC’s printed card, plus additional information to produce a multipurpose record.
3. Agreement by a broad segment of the library community on the elements to be included in the record was most desirable, and the design of the record at LC was probably the best means of achieving standardization.

Three LC staff members were assigned to the task of analyzing cataloging data from a machine processing point of view. A report,[3] issued in June 1965, was reviewed by over 150 LC staff members and their comments were issued as a supplement to the report. Comments were also elicited from many representatives of the library community.

The report was discussed at a conference supported by CLR and held at LC in November 1965. The enthusiasm generated at this meeting again demonstrated the desirability of LC becoming the distribution source of machine-readable records. As a result, LC sought funds from CLR to conduct a pilot project and in December 1965 received a grant to test the feasibility and utility of the distribution of cataloging data in machine-readable form from LC to user libraries. The project was named MARC, for machine-readable cataloging.

MARC PILOT PROJECT, 1966–1968

Planning for the pilot project began early in January 1966. CLR funds were used for contractual support: 1) to develop procedures and programs for the conversion, file maintenance, and distribution of MARC data, as well as programs to use the data at the participating libraries; and 2) to assist in the evaluation of the project. Participants in the pilot project were invited by LC, and from the 40 libraries that responded affirmatively, 16 were chosen.
Selection was based on type of library (special, government, state, university, public, and school), geographic location, availability of personnel, equipment, funds, proposed use of MARC data, and willingness to evaluate the utility of the data and prepare written reports.

The participants selected were: Argonne National Laboratory; Georgia Institute of Technology; Harvard University; Indiana University; Montgomery County Public Schools; Nassau (County) Library System; National Agricultural Library; Redstone Scientific Information Center; Rice University; University of California Institute of Library Research, Los Angeles; University of Chicago; University of Florida; University of Missouri; University of Toronto; Washington State Library; and Yale University. By February 1966, the LC staff assigned to the project was prepared to call a conference of the participants for the purpose of describing: 1) the concepts, objectives, schedules, functions, and requirements of the experiment; 2) the operation at LC; 3) the MARC format; 4) the materials to be sent weekly to participants; and 5) the content of the reports expected from the participants. The February meeting was considered the official opening of the pilot project.

LC set for itself the herculean task of completing within eight months the design of procedures and computer programs required both for the LC operations and for the participants. The MARC I format had to be stabilized by April 1966 to begin programming. Consequently, the time spent in analysis and design was severely limited and the form of material was restricted to book materials.

The first distribution was to begin in September 1966. Actually, the first test tape was mailed in October and the weekly service began in November.

The months following the initial distribution were hectic for all. The computer programs provided to participating libraries were not, in all cases, error free and had to be modified; the participants were busily engaged in writing their own tailor-made programs; and LC was learning better production methods "on the job" to reduce the large number of incorrectly edited and inadequately verified records. The load on the telephone lines between the participants' locations and Washington, District of Columbia, was heavy from November 1966 through the next several months.

*It was thought that for the project to provide the most useful information, different types of libraries and geographical areas should be represented. The first (types of libraries) was a valid decision; the second (geographical areas) is questionable.

The pilot project format, later in the life of the project, was called MARC I to differentiate it from the distribution format (MARC II), which was designed later.

The pilot project had originally been scheduled to end in June 1967. This termination date, however, seemed unrealistic. Setup time had taken longer than anticipated at the participating libraries and, with June as a cutoff date, there could be little operational experience. LC stated at the midwinter meeting of the American Library Association (ALA) in January 1967 that the pilot would be extended. During the next six-month period, it became obvious that MARC served a useful purpose and, encouraged by the enthusiasm of the library community, LC announced at the June 1967 ALA/MARC meeting that a full-scale operational MARC Distribution Service was in the planning stages. The pilot project would continue through June 1968 while the operational system was being planned and implemented. The pilot service was extended to four additional participants—California State Library, Illinois State Library, Cornell University, and the State University of New York Biomedical Communications Network.

The participating libraries experimented with card catalog and book catalog production, current awareness listings, filing arrangement by computer, etc. Not all participants were successful in mounting an operational system but, nonetheless, all cooperated to the fullest in reporting back to LC the pros and cons of each aspect of the project. In particular, the need for timely receipt of data and quality of records was stressed.

Part of the rationale for the pilot project was the test of a machine format under operational conditions in order to design a next generation format based on the experience gained. In addition to the substantive evaluation of the MARC I format provided by the participating libraries, an important event occurred during the project which exerted a strong influence on the design of MARC II. The interest expressed by the British National Bibliography (BNB) in mounting a UK/MARC Pilot Project and the many visits from foreign librarians directed thinking toward a standard communications format suitable for interchanging bibliographic data, not only from one organization (LC) to many, but also among organizations, perhaps crossing national boundaries.

The philosophy behind MARC II was the design of one format structure (the physical representation on a machine-readable medium) capable of containing bibliographic information for all forms of material (books, serials, maps, music, journal articles, etc.) and related records (name and subject reference records, etc.). The structure, or "empty container," the content designators (tags, indicators, and subfield codes) used to explicitly
MARC DISTRIBUTION SERVICE, 1968–

From June 1967 through June 1968, LC staff, in addition to concentrating on the MARC II format design and the expanded character set, were engaged in the redesign of the procedures and programs for the operational MARC Distribution Service. The pilot project was officially terminated in June 1968, and from July through March 1969 LC shifted over to a period of testing the new procedures and programs.

During this practice period, a proposal was written by LC to the Information Science and Automation Division of ALA recommending that a series of workshops (which became known as the MARC Institutes) be held throughout the country for the purpose of briefing library staffs on the MARC II format, LC procedures, and the uses of MARC data by libraries during the pilot project. The first institute was held in Seattle, Washington, in July 1968. This series continued for several years and was attended by over 2000 individuals.

The first Subscriber’s Guide to the MARC Distribution Service [which later became Books: A MARC Format] was published in August 1968 so that institutions planning to subscribe to MARC would have the necessary information for programming. The Guide was followed by a test tape in the fall of 1968 to provide the means of checking user programs.

In March 1969 the operational system was launched covering all English-language monographs cataloged by LC including titles acquired through the National Program for Acquisitions and Cataloging (NPAC). The distribution cycle was weekly, and approximately 1000 records appeared on each MARC tape.

Before the actual implementation of the distribution service, a great deal of interest was generated in MARC procedures. Consequently, in 1969, LC issued, through ALA, the first edition of the MARC Manuals. This publication contained the Data Preparation Manual: MARC Editors, the Transcription Manual: MARC Typists, the Subscriber’s Guide to the MARC Distribution Service, and Computer and Magnetic Tape Unit Usability Study (an analysis to determine which computers and peripheral devices could be used to process MARC tapes).

Because of hardware limitations and time and funding constraints, the new MARC system was designed as a batch tape system composed of four subsystems: input, file maintenance, retrieval, and output. The programs within each subsystem were generalized, data independent where possible, table driven where the characteristics of the data were known (i.e., validation of tags across

---

Identify or to additionally characterize the data elements, and the content, the data itself (author’s names, titles, etc.), are the three components of the format. It was recognized that under ideal conditions the universe of material would be studied at one time for a more coordinated approach to the assignment of content designators. However, those responsible knew the magnitude of such a task, the time required, and the need for specialists to be involved. Consequently, to make progress in the near term, it was decided to handle one form of material at a time, beginning with books.

LC staff held many meetings discussing the book format contents and content designators. The intent was to provide a multipurpose record “rich” enough in detail to allow inclusion or exclusion of data elements based on the user’s needs. The resultant MARC II format for books reflects the consensus of a large number of librarians and systems personnel.

Another significant outcome of the pilot project was the design of an extended character set for roman alphabet languages. At the onset of the project, a character set was specified based on the work in progress by the Library Typewriter Keyboard Committee of the Resources and Technical Services Division of ALA. Once the pilot was in operation, LC, in consultation with the National Agricultural Library and the National Library of Medicine, turned its attention to the development of an extended character set to cover all the major roman alphabet languages as well as the romanized forms of nonroman alphabets.

In December 1967 a meeting funded by CLR was held at LC for the purpose of discussing the MARC II format and the proposed character set. This meeting was particularly significant since the participants, based on shared experiences in the pilot project, were setting the framework for an operational MARC system that would have wide implications for the entire library community and the future of automated library systems.

At the conclusion of the pilot project in June 1968, LC had distributed approximately 50,000 machine-readable records for English language book materials. All concerned were a little wiser than at the onset of the experience. LC had learned a great deal about the procedures and the funds needed to convert cataloging data to machine-readable form. The participants realized the complexities inherent in applying the computer as a tool for library operations and the requirement for management support, time for implementation, and funds. All organizations were aware of the close cooperation required between computer and library personnel for a successful project.

A project report was published in 1968 describing the experiences of LC and participating organizations, the expanded character set, and the MARC II format.
forms of material), and parameterized to the extent possible to allow the specification of unique requirements (e.g., output formatting).

From the beginning of the design phase, LC staff recognized that the system would require updating to use disk as the storage medium and to provide for online correction. In 1971, work was begun on the Multiple USE MARC System (MUMS) to provide online capability and on the redesign of the MARC input and maintenance system.

In accordance with the original plans to specify MARC formats for forms of material other than books, LC published formats for serials and maps in 1970, films in 1971, and manuscripts in 1973. The music and sound recordings format is in draft form at the time this is written and should be published in 1975. In the design of all formats, LC worked with other organizations that had expert knowledge of the material concerned.

The distribution service has been expanded as funding permitted. Distribution of records for films began in 1972, and for serials, maps, and French-language monographs in 1973. The present services represent approximately 105,000 records per year. If the service is expanded to include books in other romance languages and in German late in 1974, the total record count will reach approximately 150,000 annually. As of summer 1974, the database totaled approximately 500,000 records. The present service is available in several forms. These include: 1) all MARC records; 2) books (All) covering records for English-language monographs including titles acquired through NPAC, titles in all languages in the Cataloging in Publication (CIP) Program, and French-language monographs; 3) books (English) as described in books (All) above, but excluding the French-language records; 4) films, covering motion pictures, filmstrips, transparencies, slide sets, and other projected materials, in all languages; 5) maps, covering records for single and multisheet thematic maps, map sets, and maps treated as serials, in roman-alphabet languages; and 6) serials, covering all serials receiving printed card cataloging, including romanized records for titles in nonroman alphabets. Records for books (English) and films can be purchased in annual cumulations. Test tapes for all services are available for experimental purposes.

In 1973, LC suggested to the Resources and Technical Services Division/Reference and Adult Services Division/Information Science and Automation Division (ALA RTSD/RASD/ISAD) Committee on Representation in Machine-Readable Form of Bibliographic Information (MARBI) that a MARC advisory committee be formed to work with LC regarding changes to the various MARC formats. The MARBI committee decided it would be the MARC advisory committee, and LC was asked to prepare a paper proposing how such a committee would operate in relationship to the MARC Development Office (MDO). The LC proposals and recommendations were adopted by the MARBI committee during its meeting at the 1974 ALA midwinter meeting.

The term “change” includes additions, modifications and deletions of content data (in both fixed and variable fields), and content designators made to the format as well as additions, modifications, and deletions made to the tape labels. The changes under consideration fall into five categories: 1) changes resulting from a change in cataloging rules or systems, 2) changes made to satisfy a requirement of LC, 3) changes made to satisfy a subscriber’s request, 4) changes to support international standardization, and 5) changes made to expand the MARC program to include additional services.

Guidelines have been established to handle the above categories of changes and a time period set for the MARBI committee, working through the MARC users, to react to a proposed change. LC agreed to notify MARC subscribers two months prior to including any change in the MARC Distribution Service.

LC plans to continue to expand MARC until all of LC cataloging (approximately 250,000 titles annually) is encompassed within the MARC system. In certain nonroman alphabets this extension may mean romanization. Equally as important as the bibliographic records is the availability of name and subject authority records. The procedures to input subject records and to maintain the files have been operational for some time. The design of similar procedures for name records, as well as distribution procedures for name and subject records, began in 1974, and it is hoped that in 1975 the library community will have these records for processing in their local systems.

Although the primary advantage of the distribution service is considered to be the cost savings resulting from centralized cataloging and from centralized editing and transcription of machine-readable records, another byproduct of MARC is often overlooked. It is impossible to estimate the resources (people and time) saved by national and international organizations implementing automated systems through the use of MARC publications. The LC language and country codes and character sets have been widely adopted, and the various MARC editorial and transcription manuals adapted to local needs. Thus,

Since LC redistributes all changed records and, in addition, each CIP record is updated at least once to a full MARC record, the actual number of records distributed is significantly larger than the number of titles cataloged and included within the scope of MARC at any time.
through the efforts of a dedicated LC staff, significant cost savings have been effected throughout the world.

**RETROSPECTIVE CONVERSION, 1968–**

With the assurance of an ongoing MARC Distribution Service, libraries throughout the country began to discuss and, in some instances, to plan the conversion of their existing catalogs. Such uncoordinated conversion activities would result in costly duplication of titles in machine-readable form. Any future amalgamation of the multiple databases into a single unified national database would entail a formidable task of human editing to make the name and subject headings assigned by many libraries consistent with authorized forms.

Since LC was also interested in the conversion of its retrospective records, it appeared timely to investigate the feasibility of a large-scale centralized conversion of retrospective cataloging records and their distribution to the library community.

In 1968 a proposal was submitted to CLR for a feasibility study to be conducted by LC. Recognizing the far-reaching significance of a centralized conversion effort, CLR responded immediately by providing funds. The responsibility for the conduct of the study which was dubbed *Retrospective Conversion* (RECON) was assigned to a working task force. An advisory committee composed of members of the library profession was appointed to provide guidance for the study. Both groups included representatives from LC as well as from other organizations.

A report published in 1969 described the work performed during the feasibility study and stated the major conclusions reached and recommendations made by the task force. The task force examined in detail: 1) the hardware and software required for a large-scale conversion; 2) existing LC files to select the one most suitable for conversion; 3) the rationale for setting priorities for conversion, and how best to accomplish the job; and 4) the costs of hardware, software, and manpower for a conversion project.

As a result of its investigations, the task force arrived at the following conclusions:

1. The MARC Distribution Service should be expanded to cover all languages and all forms of material as rapidly as resources and technology allow. There should be no conversion of any category of retrospective records until that category is being currently converted.

2. Conversion of some portion of retrospective records to machine-readable form should be an early goal for library automation efforts.

3. Conversion for a national bibliographic database requires standardization of bibliographic content and machine format. Standards for conversion of retrospective records should be the same as those for current records.

4. The highest priority for retrospective conversion should be given to records most likely to be useful to the largest number of libraries. As nearly as possible, subsequent priorities should be determined by the same criteria.

5. Large-scale conversion should be accomplished as a centralized project. Decentralized conversion would be more costly and unlikely to satisfy requirements for standardization. The project should be under the direction of the Library of Congress.

Based on these conclusions, the task force recommended that conversion should take place in reverse chronological order, by language, using the LC Card Division record set, which is arranged in sequence by date. Each record would be compared against the LC Official Catalog for updating purposes.

The order of conversion should be: 1) English-language monograph records issued from 1960 to date, 2) Romance- and German-language monograph records issued from 1960 (this conversion would not begin until MARC had been expanded to include these languages), and 3) English-language monograph records issued from 1898 to 1959. Every effort should be made to convert items 1) and 2) within four years.

An initial method of conversion was prescribed and it was proposed that further studies be initiated to determine: 1) the best means of obtaining standard records for items not represented in the LC Official Catalog to build a complete national bibliographic data store, and 2) the feasibility of establishing a national union catalog by recording holdings of libraries in the data store.

The task force further recommended that an implementation committee be established to investigate sources of funds for a detailed system design for a large-scale conversion effort and the initiation of a pilot project to test the proposed conversion system.

The implementation committee was never formed. However, LC took the initiative and proposals were submitted to CLR and the U.S. Office of Education (OE) to augment the resources committed by LC with additional funds required to support both a pilot project and the continuation of the activities of the working task force. Both CLR and OE responded positively and in August 1969 the RECON Pilot Project was initiated, with the additional funds designated for the studies by the task force and the travel expenses of both the task force and the advisory committee.
The pilot project conducted by LC covered five major areas:

1. Conversion techniques postulated in the RECON feasibility report were tested in an operational environment by converting English-language monographs cataloged in 1968 and 1969 but not included in the MARC Distribution Service. (Many of the 1968 titles were formerly distributed to the participants of the MARC Pilot Project in the MARC I format.)

2. Procedures and programs to implement format recognition were developed.

3. Problems associated with the conversion of records in other languages and cataloged according to other conventions and cataloging rules were analyzed for their effect on a conversion effort.

4. The state of the art of input devices for a large conversion effort was monitored. This included the testing of two devices in a production mode.

5. Microfilming techniques and their associated costs were studied to determine a best method for copying the sections of the LC Card Division record set. Microfilming was suggested because the record set is a ‘‘high use’’ file that cannot be withdrawn in whole or in part for any substantial period of time.

The pilot project continued for two years (August 1969–August 1971) and is fully documented in a report published in 1972. The findings of the project are briefly summarized below.

1. Format recognition applied to unedited records is a practical computer technique, eliminating the need for human editing prior to keying. The costs for keying and proofing for format recognition remain approximately the same as for processing records fully edited by humans. However, applying format recognition in place of human editing does effect a cost reduction of approximately 12% in manpower costs.

2. The magnetic tape selectric typewriter, currently in use, was found to be as cost efficient as any other off-line device given the requirements for easy accommodation of variable length records and an expanded character set. Cathode ray tube terminals were also investigated for use in an on-line mode for corrections. Only one device was available with the expanded character set. The results of testing direct-read optical character readers (OCR) to convert LC cards to machine-readable form were negative; no device could perform adequately.

3. Filming all cards in a given series of the Card Division record set against a form and reproducing hard copy proved to be the most efficient method of producing source documents. The desired subset of records would then be selected for conversion.

4. Processing of older catalog records and those in foreign languages is more complex than processing current English-language records for books, and therefore conversion costs will be higher.

5. The lowest RECON unit cost that can be anticipated for a record processed by format recognition is approximately $3.06.

Approximately 58,000 records were converted during the pilot project. The 1969 titles were distributed free of charge to MARC subscribers in 1971 and the 1968 titles were made available through sale.

The most significant technical achievement of the pilot project was the development and implementation of the format recognition process. Automatic identification of the data elements of machine-readable records had been tested in other organizations, namely the Institute of Library Research at the University of California, Berkeley, and the Bodleian Library, Oxford. However, LC was the first to attempt to assign the content designators (i.e., tags, indicators, and subfield codes) and the fixed fields for a full MARC II record by program. The cataloging source data are transcribed from left to right and from top to bottom. The data are input as fields, which are detected by the program because each field ends with a carriage return, and each field continuation is indicated by a carriage return and tab. The title paragraph is the only field where the program is given an additional clue in the form of a delimiter (a special character used to separate units of data) separating the title, edition, and imprint. The format recognition program searches first for the collation statement, which is easily located and always present. Once the collation statement is found and identified, the fields preceding and following collation are analyzed and identified separately using a variety of clues, such as first characters present and punctuation. The first pass of the program provides gross identification (in most instances only a partial tag is provided) and the remainder of the process is a reexamination of each field to provide final and complete tags, indicators, and subfield codes. After each variable field is fully identified, the data are scanned for information required to set the fixed fields. Tables of

---

5Format recognition is a technique for the automatic assignment of tags, indicators, subfield codes, and fixed fields to machine-readable bibliographic records.

6Projected costs were based on the MARC Pilot and Distribution Service experience because RECON was an experiment rather than a production effort. Time was expanded on testing devices, alternative methods, etc., and consequently there was no basis for determining accurate cost projections.

7The source document is either an LC printed card or a MARC worksheet.
keywords consisting of items such as U.S. cities, foreign cities, geographical names, meetings, and honorary titles associated with personal names are used for matching purposes to aid in the identification of a variable field, or a data element within a field, as well as providing, in some instances, the information needed to fill in a fixed field.

Format recognition is currently used for the processing of English- and French-language monographs. Additional keyword lists will be built for other languages as the MARC service is expanded. The program has been operational since May 1971. RECON records were processed to test the technique in a production environment until January 1972. At that time, satisfied with the results, LC proceeded to use the technique for the processing of current MARC records.

Concurrent with the RECON Pilot Project, the task force considered certain basic issues of retrospective conversion that were of national scope. The studies, with conclusions and recommendations, were reported in a publication in 1973. Underlying all the work performed was the reaffirmation of the importance of a coordinated retrospective conversion activity to keep duplication of effort to a minimum and to achieve a high degree of compatibility among records.

The special studies involved four areas that are described along with the findings of the task force.

1. There have been many attempts made and a great deal of effort expended to determine a subset (hereafter referred to as level) of the MARC format that would still allow a library using it to be part of any future national network. A level of a MARC record can be distinguished by differences in: 1) the data content of a record, and 2) the extent to which its contents are separately identified. Upon agreement as to the use to be made of machine-readable cataloging records, two or more parties can define a level of the MARC format to satisfy that use. The task force was concerned with the creation of machine-readable catalogs for national use, i.e., either the recording of holdings to a national union catalog or the distribution of cataloging information in machine-readable form to library networks, systems, or individual institutions. The task force concluded that level of MARC records can be defined for each of these functions. Since the union catalog function is well defined, a level of less completeness than the full MARC could be determined. However, the record for the distribution function must satisfy the needs of diverse installations and applications and, consequently, must be the fullest MARC record.

2. A large pool of machine-readable bibliographic records exists. Since MARC is still limited in its coverage, there is the potential of using this pool of records to accelerate the building of a national database and to reduce the duplication of input with a resultant cost savings. The task force undertook to investigate representative machine-readable databases and to assess their potential for this purpose. The study indicated that the per record cost of converting non-MARC records to MARC, comparing them with the records in the LC Official Catalog (to ensure consistency and completeness), and updating their contents, approached the per record LC/MARC cost. However, if the record was similar to an LC/MARC record in terms of fullness of catalog entry, format, and character set, i.e., a high potential record, then comparing only its access points with the LC Official Catalog record and updating those access points when required, would result in a conversion cost estimated to be in the order of one-half of present LC/MARC conversion costs.

3. The National Union Catalog (NUC) serves two major functions: the provision of cataloging data and a finding list for the locating of a title. With the advent of LC/MARC and the proliferation of library automation projects, it was logical for the task force to consider the implications of a national union catalog in machine-readable form. On-line access to a national union catalog did not appear to be likely in the near future. Attention was consequently turned to the automatic production of the NUC in book form or microform. This mode of production would relieve humans of the drudgery of preparing the catalogs, make the information more rapidly available, and provide the facility for additional access points to the data. The task force selected the register/index as the optimum format for NUC. This type of catalog is composed of a register of complete bibliographic entries arranged in numeric sequence according to a number assigned to each entry when the register is prepared. The register is never cumulated. The indexes are derived from the register entry and contain a brief bibliographic entry and the register number. The indexes are cumulated at predetermined intervals. This format has the advantage of providing several points of access without having to print the full bibliographic record more than once. The task force concluded that the cost of the automated system with its associated advantages would not exceed the cost of the present manual system, and the cost of producing the quinquennial would be significantly reduced. In addition, the database might eventually form the basis for an on-line network of regional bibliographic centers.

4. Experience in the RECON Pilot Project indicated that a large-scale conversion project as projected in the
RECON feasibility study demanded far more resources (staff, space, and funds) than there is hope of obtaining. To convert chronologically by language as recommended in the feasibility study, on a lesser scale, would be too slow to satisfy those libraries requiring machine-readable records for their automation efforts. Therefore, the task force examined alternative strategies, hoping to arrive at a means of more rapidly responding to the needs of libraries, but concluded that there was no perfect solution to the problem, since preferences for languages, dates, and forms of material are dependent on the type and size of library involved. Systematic conversion, i.e., orderly conversion by language and date, has the advantage of allowing the user to predict with reasonable certainty what is in machine-readable form. It has the disadvantage that users may wait for long periods of time before desired records may be available. Conversely, nonsystematic conversion (based on some criterion such as all records in a given bibliography) might have the advantage of more rapidly converting records desired by a larger number of organizations, but it would have the disadvantage that, without querying the database, the user cannot predict which titles are in machine-readable form.

The task force concluded that LC should concentrate on going forward as rapidly as possible to convert all of its current cataloging to machine-readable form and that an agency, established expressly for the purpose of a large-scale retrospective activity, should undertake to convert the retrospective LC records that are most in demand and be responsible for adapting records from libraries other than LC for inclusion in the national database. The task force recommended that since the problem of retrospective conversion is of concern to all of the nation’s libraries, the National Commission on Libraries and Information Science (NCLIS) might be the appropriate agency to determine a course of action and explore sources of funding to implement a national program.

To date, no central agency has been established nor has NCLIS taken any express action concerning retrospective conversion. LC, when resources permit, continues to convert retrospective titles selected from the Card Division as well as MARC Pilot records that were not converted during the RECON Pilot Project. These records are distributed as part of the MARC service.

There is continuing interest in a centralized conversion activity. The concern of the participants in a series of meetings [17] sponsored by CLR is one manifestation of this interest. The purpose of these meetings is to seek a means of exchanging machine-readable records among organizations until such time as a national service may exist to perform this function centrally. The records under consideration are of two types: 1) records converted from LC cataloging copy not included in the MARC Distribution Service (i.e., either items cataloged prior to the MARC service for a particular language or form of material or records not yet within the scope of MARC), and 2) records representing titles not cataloged by LC. The emphasis here is placed on all records, both retrospective and prospective, being converted by organizations other than LC. As progress is made on the automation projects of individual institutions and as regional networks evolve, it is an accepted fact that institutions will be doing local conversion. The excessive cost of duplicate conversion that will result if machine-readable data are not shared, as well as the problems of an inconsistent data base, can only be resolved if a single organization assumes the responsibility and has the resources for building a national data base.

INFLUENCE OF MARC ON STANDARDIZATION

Libraries acknowledged the importance of standardization long before MARC came into being. Machine systems, however, showed the need for an extra measure of conformity over what had been apparent in the past. What appears in the printed record is human readable, not bit configurations representing characters or codes or explicitly identifying content data. Variations in the placement or the formatting of words, paragraphs, and numbers in the printed record are not too damaging to the initiated, but can be intolerable to efficient machine processing.

The interest of libraries in the computer for library operations was increasing in the 1960s. The availability of cataloging data in machine-readable form supplied by LC, the need to input cataloging data locally (data not within the scope of MARC or titles not cataloged by LC), the possibility of sharing these locally generated records, the potential for using computer programs across organizations to reduce the high cost of designing and writing software, and the need for hardware capable of handling large character sets were all factors that put increased emphasis on the establishment of and conformity to standards.

---

\(^{1}\)Initially, these records might be those ordered from the LC Card Division most frequently.

\(^{2}\)The MARC Editorial Division, responsible for the conversion of MARC records, is geared to peak load conditions. When the volume of records diminishes for any period of time, the gap is filled in by converting retrospective titles.
The impetus given to standardization by LC/MARC is doubtless one of its most important results. The establishment of bibliographic standards did not happen chronologically but in many instances overlapped, i.e., the development of a new standard might begin while a proposed standard was in the process of adoption. Consequently, reporting of the standards cannot be made in strict chronological order.

**Format for Information Interchange**

During the operation of the MARC Pilot Project, BNB personnel visited LC to investigate the possibility of a UK/MARC Pilot Project. At a later date, when BNB decided to proceed with its MARC project, its staff worked with the MARC staff on the MARC II format in an effort to satisfy the requirements of both agencies. This cooperation had long-term effects. Both agencies recognized the future implications of an interchange format and the importance of two major publishing countries agreeing on a standard. No only would it be possible to exchange machine-readable records between the United States and the United Kingdom, but the way would be open for other countries to follow the lead and develop their own MARC projects. This possibility was particularly interesting to LC because of NPAC, where LC uses the descriptive cataloging as given by other countries in the program, adding or modifying main and added name entries, subject headings, etc., to the record as necessary. The entire process is a manual operation using printed records. The potential exists for updating machine-readable records from countries in the program and thus reducing the cost of conversion at LC.

Recognizing the worldwide impact of sharing machine-readable records, LC and BNB, in addition to being active in their respective national standards organizations, in 1969 submitted the then proposed American National Standards Institute (ANSI) format and its equivalent, the proposed British Standards Institute (BSI) format to the International Organization for Standardization (ISO). The MARC projects in the United States and Great Britain and the developing ANSI, BSI, and ISO standards created the environment for many countries (including Australia, Canada, Denmark, France, Germany, Italy, Japan, Latin America, Norway, and Sweden) to begin to plan for and implement their own national MARC system. (In addition to national libraries and bibliographies, many organizations based outside of the United States, such as UNISIST, the International Labor Organization, Information Services in Physics, Electrotechnology, Computers and Control, and the International Atomic Energy Agency, have adopted or recommended adoption of the format structure.)

The MARC format *structure*, as a vehicle for information interchange, was adopted in the United States by the three national libraries, ALA, ARL, the Committee on Scientific and Technical Information (COSATI), the Educational Resources Information Center (ERIC), the Association for Scientific Information Dissemination Centers (ASIDIC), and others. The format structure became an ANSI standard in 1971[^18] and an international standard in 1973 [ISO 2709–1973 (E)].[^19]

**International Standard Bibliographic Description**

In 1969 the International Federation of Library Associations (IFLA) convened an International Meeting of Cataloguing Experts. Since the Paris Principles had set the framework for an international cataloging code, it seemed appropriate to call an international meeting to discuss the next step, a standard bibliographic description. A working paper[^20] prepared by Michael Gorman of the BNB, with the support of a United Nations Educational and Scientific Organization (UNESCO) grant, demonstrated the difference in the order of data and in punctuation patterns of various national bibliographies (NUC represented the United States bibliography). A working party was selected to draft an international standard bibliographic description for monographs. An international standard bibliographic description would provide a standard set of descriptive elements in a standard order using standard punctuation to separate the elements. Such a description would facilitate: 1) the comprehension of the description in printed form regardless of language, and 2) the automatic recognition of elements of the description for machine identification. The first published draft of the International Standard Bibliographic Description (ISBD) was issued in 1971, followed by the first standard edition in 1974.[^21] [A recommended ISBD for serials was also published in 1974.[^22]]

**Content Designators**

The ISBD for monographs, the ISO standard format structure for bibliographic interchange, Kaltwasser’s paper[^23] concerning universal bibliographic control, and the ever-increasing number of national MARC formats were all positive steps toward international cooperation. Problems remained, however. One of the significant areas of nonstandardization among national MARC formats was in the assignment of content designators to elements of information in the machine-readable records. Accordingly, IFLA established, under the auspices of the Committees on Cataloguing and Mechanization, an
The Working Group realized that the issues concerning data organization and functions of content designators could be resolved (subsequently agreement on these points was reached in a meeting in Brussels in February 1974) but that the lack of an internationally accepted cataloging code and the dissimilarities in the functions of different bibliographic services were areas over which it had no control.

Encouraged by the fact that it has been possible to work around these differences in the formulation of the ISBD, the Working Group proceeded to adopt the concept of a SUPERMARC as an international system for exchange, leaving the various national systems as they now exist. Each country would have an agency that would translate its own machine-readable records into the format of the SUPERMARC system; likewise, each agency would translate SUPERMARC records into its own national format.

The international machine-readable record was divided into functional blocks (e.g., intellectual responsibility block, descriptive block, subject block). There was unanimous agreement that the descriptive block would follow the ISBD, and it is this block where agreement on content designators and content is most likely possible.

It is recognized that the content and consequently the content designation of certain blocks is dependent on cataloging codes and practices. Therefore, the best that can be hoped for in the near term is agreement on content designation at a level of specificity less than that presently used by national systems. For example, the intellectual responsibility block contains names of persons and organizations associated with the work described. In national systems, each type of name may be further expanded to multiple levels of subtypes and subfields coded in a variety of ways. In the international record, the expansion may be limited to indicating personal, family, corporate, and conference names; and primary, alternate, and secondary responsibility.

The concept is a record composed of two principal parts: 1) the identification of the item (descriptive block modeled after the ISBD), and 2) all other elements making up the bibliographic record, e.g., headings, subjects, and classification numbers. This will allow the user to accept the descriptive block for inclusion in his system and provide the option of using or discarding the other elements of information.

Work on this international standard is still underway. It is hoped that agreement will be reached on a standard set of content designators by the IFLA meeting in 1975.

Other Activities

The expanded roman-alphabet character set designed by LC was adopted by ALA, has been used in information systems throughout the world, and is the basis for the current work in progress in ISO/TC46/SC4/WG1, Character Sets for Documentation and Bibliographic Use. The acceptance of a character set for bibliographic use, indicating marketability to hardware vendors, led to the manufacture of a typewriter sphere, a print train, and several cathode ray tube terminals.

Country of publication codes and language codes originally established by LC for the MARC Pilot Project, and, based on experience, modified for the MARC Distribution Service have also been accepted as the codes for several systems. Although these codes have not been adopted as ANSI or ISO standards, they do have an impact on current standardization activities.

The challenge to the individuals involved in standards setting is great. Much remains to be accomplished. Progress has been made, however, and it appears that ‘International MARC’ may become a reality.

MARC USERS

Use of MARC at the Library of Congress

The main thrust of LC’s bibliographic automation program during the years 1966–1970 was the distribution of MARC data. In terms of LC, MARC up to 1974 covered approximately 40% of LC’s total cataloging effort.

In 1970, sights began to turn inward, and MDO was established in the Processing Department to concentrate on automating the functions concerned with technical processing. An article describing the automation programs of the Processing Department was published in 1972. The program can be considered to have a three-pronged approach: 1) expansion of the MARC Distribution Service, 2) design and implementation of a core system and provide the option of using or discarding the other elements of information.

1 Before June 1970, staff responsible for MARC, RECON, etc., was part of the Information Systems Office (ISO), organizationally situated in the Office of the Librarian. At the same time that MDO was established, ISO was transferred to the Administrative Department.
bibliographic system, and 3) products and services for LC and the library community. It is principally in item 3) that the MARC data base is used at LC at this time. Several of these projects are briefly described below.

Book catalogs

1. The Main Reading Room catalog consists of records included in MARC (books and serials) and records input especially for the project. Catalogs have been produced, via computer printer, arranged by call number, author, title, and subject for use in LC. Plans include the general distribution of the catalogs in this form as well as photocomposed issues at a later date.

2. The Science Reading Room file is also composed of MARC records (books and serials) in addition to records input specifically for the project. Computer-printed catalogs have been produced for use in LC.

3. The catalog of Films and Other Materials for Projection for the last quarter of 1972 and the first two quarters of 1973 was produced from the MARC films database in 1974. This catalog is issued in two sections in one volume: 1) title main entries and added author entries, and 2) subject added entries. ‘See’ and ‘see also’ references are included.

All machine-produced book catalogs and other bibliographic listings are arranged by means of a program called Library Sort Key Edit (LIBSKED). This program was designed to implement the new LC filing rules[27] formulated by the Technical Processes Research Office of the Processing Department.

Printed cards

MARC records are input to the Card Division’s automated system to compose and print LC cards, complete with overprinted headings, for use in the LC catalogs. On-demand production of cards for film and map titles in response to orders from subscribing libraries or for cards that are out of stock is operational, but on-demand production for all orders (approximately 40,000 per day) for which there is a MARC record cannot be implemented until quantity limitations imposed by the present capabilities of the offset press and accompanying cutting and collation equipment are resolved. Consequently typeset cards are still being prepared for MARC titles through the Government Printing Office.

MARC retriever

The MARC database is being used more and more frequently in LC as a bibliographic reference to supplement the card catalogs. The machine-readable records provide access via fields not used as headings in the card catalogs, and the richness of the MARC format offers new possibilities for the retrieval of information.

The MARC Retriever was originally seen as a research tool to promote the use of the MARC data base in new and unexplored ways. This experimental mode is still of interest in LC but, in addition, the system is used operationally for many purposes in LC as well as to provide special searches to outside organizations. The Retriever allows the user to search the contents of any variable field, subfield, or fixed field, as well as any specific tag, indicator, or subfield code value.

Recent searches include: 1) monthly lists of books on Africa, Mainland China, Latin America, Eastern Europe; 2) books in translation; 3) children’s books in translation; 4) serials in French published in France, Belgium, and Canada; 5) railroad maps; 6) maps of East Africa, Kenya, Tanzania, and Uganda; 7) CIP records in history, LC class schedules E and F; 8) Festschriften; 9) atlases published in 1972; and 10) books about Africa containing statistical information.

The records retrieved are arranged and printed according to the user’s specifications. The sort may be on author, title, subject, classification, date, etc. Output may be printed in card or list form, with the amount of information displayed (short form or full record printout) a user option.

Other uses in LC

The LC Division for the Blind and Physically Handicapped selects pertinent records from the MARC files for inclusion in its system for the production of bibliographic tools, e.g., printed cards and book catalogs for such items as talking and braille books.

As the scope of MARC expands and as the ongoing projects such as the automation of the PIF become operational, more and more use will be made of MARC at LC, both for technical processing and reference.

MARC is now searchable on-line by LC card number within LC. Access to the file will be expanded to include author/title and title only search keys. Preliminary catalogers will be able to search the MARC file as well as the PIF to determine whether the item in hand is already in the collection or requires cataloging; the Card Division will be able to search for titles for which printed

---

This includes automation of the Order Division functions and the Process Information File (PIF) which are presently in various stages of operation, implementation, and design. The Core Bibliographic System includes the use of LC name and subject authority records to produce book catalogs as well as linking them to MARC records as a cataloging aid. Several ongoing operational projects, such as the building and maintenance of the subject heading file, are modules of this activity.
cards were ordered but for which the LC card number was not given; and the Catalog Publication Division will be able to search for items reported by NUC libraries to ascertain if the reporting record is a title already covered by LC cataloging so that holdings can be posted to the proper record.

Plans include making the MARC file accessible by the Reference Department and expanding searching capabilities, as well as the eventual two-way linking between the authority records (name and subject) and each bibliographic record where the name or subject or both were used. This linking will enhance the present book catalog production system as well as provide the means to use the MARC files as a cataloging aid.

**Use of MARC by Subscribers**

At present (March 1974) there are 74 organizations subscribing to the MARC services. The 74 subscribers purchase the following subscriptions:

<table>
<thead>
<tr>
<th>Service</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARC, Complete</td>
<td>13</td>
</tr>
<tr>
<td>MARC, Books (All)</td>
<td>28</td>
</tr>
<tr>
<td>MARC, Books (English)</td>
<td>19</td>
</tr>
<tr>
<td>MARC, Films</td>
<td>4</td>
</tr>
<tr>
<td>MARC, Serials</td>
<td>12</td>
</tr>
<tr>
<td>MARC, Maps</td>
<td>3</td>
</tr>
</tbody>
</table>

This list can be further subdivided within the category MARC, Complete. For example, there are 13 complete subscriptions and three map subscriptions, meaning that 16 map subscriptions are actually in force. Dividing the subscriptions in this manner results in the following:

<table>
<thead>
<tr>
<th>Category</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Books, Foreign</td>
<td>41</td>
</tr>
<tr>
<td>Books, English</td>
<td>70</td>
</tr>
<tr>
<td>Films</td>
<td>17</td>
</tr>
<tr>
<td>Serials</td>
<td>25</td>
</tr>
<tr>
<td>Maps</td>
<td>16</td>
</tr>
</tbody>
</table>

Since an organization may order more than one type of service, the actual number of individual subscriptions as given above total 169. The subscribers are made up of 18 commercial enterprises, 23 organizations outside the United States (one of these has been previously counted with the commercial enterprises), and the remainder divided among large research, college and university, state, special, and national libraries and library centers.

Institutions are difficult to identify accurately since any one institution may perform the functions of another type, e.g., a university library may act as a center, serving, in addition to itself, several other organizations. Therefore, it is not possible to derive an accurate tally of subscribers by type of library. The subscribers discussed above are considered primary users. All of the commercial vendors and library centers as well as some of the organizations in the other categories of users provide services to other institutions. These latter institutions are considered secondary users of MARC data and, according to LC’s best estimates, total approximately 1500.

There are difficulties inherent in any presentation of MARC uses. This difficulty is in part created by the tendency of library functions to overlap. For example, the search/select function is not an end unto itself but instead may be a component of a system to print catalog cards or of an SDI system. There are also various levels of the use of MARC data, depending on the mix of manual and automated processing. MARC records are, in some cases, simply used as a source of bibliographic information with the rest of the processing purely manual, or MARC records become an integral part of an automated system. Some systems are wholly dependent on MARC; in others, MARC is some percentage of the total database. There are systems that are totally batch oriented, others are partially batch oriented and partially on-line, and still others operate principally in an on-line mode. These various differences in systems do not affect the function performed, e.g., looking up a title to find where it is located, regardless of how the system is designed or how the files are organized.

The following discussion of MARC uses attempts to avoid the problems described by addressing MARC uses in general terms and not probing into individual systems.

The majority of uses of MARC data fall into the technical processing area. This is not surprising, since before the reference function can be served adequately, items must be put under bibliographic control. Furthermore, it is in technical processing that the professional librarian needs to be relieved of performing repetitive tasks or supervising others in the performance of such tasks. However, the reference function is somewhat satisfied by SDI services, LC’s MARC Retriever, and the searching of index files of author, title, and subject. Current MARC tapes are used for selection purposes by printing out the records and routing them to persons responsible for selecting materials for acquisition, or by comparing the tapes against stored profiles and printing the records that match. Printed indexes are also prepared from the MARC files (weekly or cumulative) and made available for selection.

To the extent that MARC records are available for items to be ordered, manual bibliographic searching is

---

* The last MARC Survey was conducted in the spring of 1972 and the results published later in that year. The subscriber figures cited in this section have been updated to reflect the current situation.
eliminated, and the MARC records selected are used to produce the orders and associated documents and the records are added to the institution’s in-process or master file.

MARC records are used in the production of catalog cards and book form catalogs. Book catalogs have appeared in dictionary form, register/index form, or as separate author, title, and subject catalogs (or any combination thereof). The level of sophistication of catalog production varies greatly. The production forms include computer line printing, photocomposition, and microform.

A bibliography in microform, called Books in English, is produced by merging BNB/MARC records with relevant LC/MARC records. A catalog of English-language acquisitions of a university library and affiliated institutions is also produced by selecting appropriate records from these two databases.

Shared cataloging systems with access to LC/MARC and locally input cataloging records by LC card number and/or search keys are also in operation. Member libraries search the files for a specific record. If the record is in the file and if the cataloging data is acceptable to the institution, printed cards in the format and quantity desired by the institution are sorted (prearranged in a filing sequence) and prepared off-line for the selected record. If the record is not available, the member library inputs the locally cataloged record that then becomes available to all other participants and the printed cards are prepared from the locally input record. In some systems, records can also be output on tape for any library desiring its cataloging data in machine-readable form. The on-line posting of the locations of the requesting library and the use of this file for interlibrary loan is also in operation.

The production of catalog cards remains the most popular operation, although in many instances card production is not the raison d’être of the system. The production of book preparation materials (bookcards, pockets, and spine labels), as well as circulation cards, is a useful by-product of some systems.

In one system the MARC record is used as a bibliographic source, and agency subject terms, call numbers, etc., are added. The entire MARC file is maintained as a composite database, and a subset of the file relevant to the specialization of the organization is selected by such criteria as LC class numbers or subject headings, added to the records in the system not included in MARC, and both types made available on-line to participating libraries. A full free text search is provided with a modified Key-Word-in-Context (KWIC) display for titles, series names, conference names, corporate names, publishers’ names, and subject headings. Additional on-line searching capabilities include other names, LC call numbers, LC card numbers, international standard book numbers, Dewey Decimal numbers, local library card numbers, and local library location indicators. This sophisticated searching capability is in addition to the production of several library products, e.g., current awareness listings and book catalogs.

**CONCLUSION**

MARC is an assemblage of formats, publications, procedures, people, standards, codes, programs, systems, equipment, etc., that has evolved over the years, stimulating the development of library automation and information networks.

Any discussion about MARC in particular involves discussing library automation in general. Progress in the field is evident. There is a definite advance from the automation of discrete functions toward the automation of modules leading to integrated systems. Emphasis is being placed on the implementation of on-line systems and the use of one system to provide services to many other organizations.

Library automation activities in 1974 appear to fall into the following categories:

1. Institutions are automating their technical processing functions and, as a by-product, providing services, e.g., cards and listings, to other organizations.
2. Institutions are automating a particular aspect of their service, e.g., book catalog production, and as a by-product their procedures and programs are used to produce a like product for another organization.
3. Institutions are creating automated systems designed principally as a bibliographic center to provide services, e.g., printed cards and union catalogs, for other organizations that do not intend to automate.

Libraries have passed through an era of much talk but few results into the 1970s where the automation of library operations is no longer a promise but a demonstrated success. Faced with economic realities, the availability of bibliographic information in machine-readable form and the possibility of sharing these resources through computer technology, there is an increasing awareness that standardization, nationally and internationally, is the sine qua non of the information system.

The benefits that accrue to a library and its clients from the establishment of and the conformity to standards are many. Products from different sources will mesh. Records from different libraries will be interchanged. Machine systems will be more easily developed and shared. Union
Machine-Readable Cataloging (MARC) Program

catalogs will be possible without costly editing for consistency, thus facilitating interlibrary loan. Cost of local changes to catalog records will be minimized. It will be advantageous for vendors to manufacture hardware to handle the requirements of libraries. The process of ordering, cataloging, etc., will be more uniform. Therefore, less searching and bibliographic verification will be necessary and duplication of effort will be avoided. Networking will be facilitated. Various databases will be accessible through the use of standard protocol. Service to the user will be improved, and that is really what MARC is all about.

REFERENCES


**BIBLIOGRAPHY**

**Background, 1961–1965**


**Marc Pilot Project, 1966–1968**


MARC Distribution Service, 1968–


Retrospective Conversion, 1968–


Irvine, R. A note on the conversion of existing catalogue records to MARC format. Program April 1973, 7 (2), 96–100.


Influence of MARC on Standardization


**MARC Users**


Corbett, L.; German, J.A. AMCOM project stage 2; A computer aided integrated system using BNB MARC literature tapes. *Program January 1972*, 6 (1), 1–35.


Machine-Readable Cataloging (MARC) Program