

The Application of Microcomputer-Based Local Systems with the MARC AMC Format

FREDERICK L. HONHART

THIS ARTICLE WILL CONSIDER the effects of the new MARC AMC format when applied to a microcomputer system for archival operations. Much of the basis for this article is the result of research at Michigan State University (MSU) in the University Archives and Historical Collections and the Applications Programming Division of the MSU Computer Laboratory. At this time (January 1988) over four years have been spent developing a MARC AMC format microcomputer system for use by archival and manuscript repositories and other appropriate applications. The initial version of the system was offered for sale in September 1986, but research continues to expand and enhance its capabilities. Because this is the only microcomputer-based MARC AMC format system currently in existence, discussions of various issues, of necessity, will be specific to this system.

Development of this system was a direct result of the creation and acceptance of the MARC AMC format. While the total impact of the AMC format on archival operations remains to be seen, already it has become a major factor in the operation or future operations of archives nationally, if not internationally. In contrast to other attempts to develop automated formats for archives this one had widespread acceptance well before it was even officially released in October 1984 by the Library of Congress. One reason for the immediate and widespread support of the new format (as opposed to the previous manuscripts format) was its development over a period of years by archivists for

Frederick L. Honhart is Director, University Archives and Historical Collections, Michigan State University, East Lansing, Michigan.

archival applications. Progress reports were made periodically, so the profession was well prepared for its appearance. Another reason for its acceptance is that it is supported by the Library of Congress; this means a developer or a user of an automated system designed to utilize the format can be confident it will continue to be supported in the future.

The decision by the two major library online cataloging utilities, the Online Computer Library Center (OCLC), and the Research Libraries Group's Research Library Information Network (RLIN), to implement the new AMC format in their respective systems furthered acceptance and use of the format. It is now potentially available to hundreds of archives in the country because of their membership in one of these utilities. As described in other articles in the issue, by entering MARC AMC format records into OCLC or RLIN or other online cataloging utilities, archives have now enabled researchers to search dozens and eventually hundreds or more repositories for materials pertaining to a particular area of research. The net result is that the MARC AMC format has become the de facto standard for the description of archival and manuscript records in a machine-readable format.

One alternative to using OCLC or RLIN or a similar system is to use a local system. Information created on a local system, using the MARC AMC format, potentially could be transferred to another system's database. A local system could make the AMC format available to any archival repository, not just those affiliated with one of the utilities, and it could serve as a supplemental system for those repositories having access to RLIN or OCLC.

Developing a microcomputer-based local system is possible because of a series of events in the microcomputer field. The entry of IBM into the manufacture of microcomputers profoundly affected the industry; IBM brought credibility and rationality to an industry that had been short on both. Until their entry into the field it seemed as if every microcomputer manufacturer used a different operating system. This meant that a program that would operate on machine *A* could not operate on machine *B*. IBM's use of the Microsoft Disk Operating System (MS-DOS) and the availability of MS-DOS for use by other microcomputer manufacturers—making them IBM compatible—meant that a program written for the IBM microcomputer would also operate on other microcomputers that used MS-DOS.

At that point, the potential existed to at least consider using microcomputers for archival work. Yet there was one hurdle still to be met. Initially, the storage of information on microcomputers was limited to the storage capacity of a floppy disk or generally 360 kilobytes. This was

Microcomputer-Based Local Systems

an impractical size for all but the smallest and most limited databases. The introduction of the hard disk drive and a storage capacity of 10,000 kilobytes, or 10 megabytes, finally provided sufficient information storage capacity. Archives could now consider developing a local system for archival description and management.

For archival use a local, stand-alone, microcomputer-based system should incorporate several basic criteria—the most important of which is it should support the MARC AMC format. Failure to incorporate the format in a local system means the information cannot be exchanged with other systems using the AMC format. The system should be totally self-contained, independent of any other software or hardware requirements. The needs of archival and manuscript institutions, which are not members of OCLC, RLIN, or other bibliographic utilities, could be addressed by such a system. In order to best serve the user and the needs of the profession, a local system should be able to create MARC AMC format records for use by other MARC automated systems—that is, the records can be entered into other systems' databases for use by researchers as well as other archivists.

The development of the Michigan State University Microcomputer System, known as MicroMARC:amc, was first contemplated in early 1983. At that time a proposal was submitted by the director of the University Archives and Historical Collections for a modest university research grant to explore automated archival systems in existence or under development. The proposal was funded, and the project director visited several institutions around the country that were involved with automation. At the same time discussions were held with the Applications Programming Office of the MSU Computer Laboratory regarding the feasibility of developing a microcomputer system that could utilize the MARC AMC format. It was decided that creating such a system seemed possible, and a grant proposal was submitted to the National Historical Publications and Records Commission. The proposal was funded, and this grant, along with the financial support from the university, made the development of the system possible.

Considerable attention was given not only to the design of the system but also to the manner and development of a structure that would be utilized in creating the system. A substantial contribution to this part of the project was made by the consultant systems analyst. An advisory board of knowledgeable archivists reviewed the development of the system at appropriate times. Over a year was spent on the analysis and design of the system before any programs were written by the Applications Programming Office. A modular design was chosen for

the system. MicroMARC:amc is currently composed of four basic modules: Editor, for creating AMC records; Search, which does online searching of selected fields; Report, which will access the entire database; and the MARC I/O record module which inputs and outputs USMARC:AMC Format Records. The system was initially tested at the University Archives and Historical Collections at Michigan State University. Once a module appeared to meet the desired goals and was error free, it was sent to seven archival repositories for further testing (beta testing). The same testing procedure was followed for the documentation for the system.

It is interesting to look back at the initial assumptions that were made regarding the development of the system and see what changes took place and what problems occurred. Initially, the intention was to use an existing, commercial database management system for the storage of information. Upon further investigation, the decision was made to write a special database management program for the system. Such a database management system would not be a compromise, as any off the shelf system would be, and it would be more efficient, practical, and amenable to future refinements of the system. In addition, the system would not be dependent on another software manufacturer, nor would there be any need to negotiate proprietary rights to the software. A more general change has been the evolution of MicroMARC:amc into a more complex and sophisticated system than was originally envisioned—as we determined how we wanted the system to perform, adjustments were made to the original design. Those changes included allowing the user to create defaults for specified fields; to design a nonlinear form—to generate, for example, a catalog card—as well as linear output of selected information from the database; and to search the database to identify linked records.

Another factor that must be considered in any microcomputer-based system is the technological evolution/revolution that is taking place. Since beginning the development of the MicroMARC:amc system there have been significant changes in microcomputer technology, and new changes continue to take place. The basic IBM PC/XT has become more powerful; the current PC/AT configuration has up to 30 megabytes of disk storage, 1.2 megabyte floppy disks, and faster clock speed. Similar advances have been made by other manufacturers of microcomputers. New generations of microcomputers using the 386 chip are now available. In comparison to their predecessors, these machines have taken a quantum leap in operating speed and storage capacity. The result is that while the speed and capacity of microcomputers has

Microcomputer-Based Local Systems

increased, their effective cost has decreased. Consequently, the practical availability of microcomputers has increased; repositories which in the past, because of cost, could not consider using an automated system are now in a position to do so.

There are important implications in the new developments for both software manufacturers and users. Because MS-DOS and the IBM PC/XT and compatible microcomputers have set a widely adhered to standard, it is essential for most microcomputer programs to function within such equipment/operating system configurations. Designers of microcomputer software also must be prepared to take advantage of the constant evolution in the technology so that their software will operate on future generations of microcomputers. Obviously, it is of considerable concern to users of particular programs that when they upgrade their microcomputer equipment, their software is not made obsolete.

Use of a microcomputer-based local system brings many opportunities, as well as some drawbacks, for archival repositories. The most obvious benefit is that the system is local—it is dependent only on the user of the system. A one-time expenditure for equipment and software results in potentially lower costs for operating the system. There are no monthly fixed or use charges as is the case with OCLC, RLIN, and other similar bibliographic utilities. This means that many archival repositories that could not afford to participate in established systems will now have a financially viable alternative with a local system that uses the MARC AMC format. In the future, monetary considerations should prevent only the most limited repositories from using some kind of automated system. As will be discussed, there are some excellent reasons for repositories that are part of an online bibliographic utility to consider using a local system as well.

A microcomputer system does have some inherent limitations when compared to a system that operates on a mainframe computer. In general, the speed of processing information is slower, although system design, and the size of the database, can alleviate some of the problem. Newer microcomputers now have an increased clock speed, for example, which results in faster processing of information. However, it must be realized that some functions—for example, the time it takes to read a database—will simply take longer to perform because they are being done on a microcomputer as opposed to a mainframe computer.

Another limitation is the data storage capacity. The IBM PC/XT and compatible microcomputers have a basic 10 megabyte disk capacity (which can be upgraded). Newer, enhanced machines now have disk capacity of 20 or 40 megabytes and ones of 130 megabytes or more should

be available in the near future. Add-on hard disk drives and tape backup devices further expand data storage capacity. However, because the information being stored is normally only for one repository and in one format, as opposed to the contents of an online bibliographic utility, there is not the same need for mainframe storage and processing capacity. The new generation of 80386-chip-based microcomputers promises to significantly increase storage capacity, memory, and operating speed, making current microcomputer limitations even less of an issue in the future operation of a local system.

If microcomputer-based local systems have some limitations, they also have some considerable advantages for the user. When used with the MARC AMC format, local systems should provide the user with greater management and control capabilities because the system is designed to meet the specific needs of archival and manuscript repositories. For example, with a local system the user can search the entire database, not just certain fields or subfields, and the system should be able to do this in conjunction with one or more Boolean operators. Finally, a local system will allow the user, not the system, to define which search strategy and fields are to be examined.

In an archival repository, the hierarchical linking of records is an important management and research function because of the methodology used in the arrangement and description of records. The AMC format has provisions, if somewhat limited, for identifying and linking records. A local system should be able to utilize the record-linking capabilities of the format and easily identify hierarchically linked records in the database.

In contrast to large multiuser systems, local microcomputer-based systems most likely will not impose controls over the entry of illegal information into various fields. This is both an opportunity and a problem for the user. The reason for this freedom is to ensure that the user has maximum flexibility for entering records. For the software manufacturer it is also the least costly solution to the problem of creating records that meet the standards of other systems. The lack of system controls over the entry of illegal information, however, does pose potential problems. If the user contemplates transferring AMC records to another AMC system, then care must be taken to ensure that the records created locally meet the same standards as the system in which they will be entered.

A local system utilizing the MARC AMC format has potential advantages even for users of the bibliographic utilities. Records could be written, corrected, and revised, and then, after a certain number of

Microcomputer-Based Local Systems

records had been created, they could be batch-loaded into the utilities database. Another function could be to create from the main database specialized databases that relate to specific functions or topics. These databases could be maintained solely on the local system and hard copy created when needed. The records could be changed or updated when necessary, or the entire database searched at no additional cost.

When using a local system for the creation of MARC AMC records, there are some specific factors that need to be taken into account. A local system most likely will not have any controls over the entry of illegal characters into the database. Quality control of the record becomes a user responsibility. Another consideration is the exchange of records with other systems—one of the basic rationales for using the MARC formats. If exchange is contemplated, either with one of the bibliographic utilities or another local system, then the records should meet the record requirements of the system they are to be entered into. Failure to establish minimum information requirements could result in an inability to transfer the information or in having to reenter and correct information in the database.

Applying in-house standards to AMC records should be given careful consideration by users of local systems. Two potential benefits of using the MARC formats in conjunction with automated systems are greater control over archival materials and greater access to the information contained therein. However, these benefits can only be achieved if the MARC records are entered into the system in a standardized fashion. Ideally, the system itself should have some means of checking the records for uniformity. Generating a printout of subject headings for use as an authority list is one such way to verify entries.

The MARC AMC format is complex, comprehensive, and subject to various interpretations as to the use of some fields. Only through trial and error with the format can a user become familiar with it. A local system permits the user to repeatedly reenter corrections, additions, or deletions of information with few penalties—the only cost is the time required to enter or delete information.

The advent of local, automated, MARC AMC format systems has tremendous potential for researchers as well as for archivists. The creation of the AMC format, and its growing acceptance as the standard for entry of archival records into a machine-readable format, is a major step in the development of a national information research database; the format is employed by OCLC, RLIN, and the National Union Catalog of Manuscript Collections, creating a standardized pool of information for use by researchers. Repositories that do not use one of the large

bibliographic utilities now have the capacity to create their own AMC format records. Information which otherwise would be inaccessible, or accessible on a very limited basis, once entered into and made available through larger networks, is potentially available to any researcher who has access to an online terminal connected to one of the national databases.

The sophistication of a local system in the management and control of archival records also benefits the researcher. Local systems have the potential to perform manipulations of the database that are often not practical, or possible, for large systems with an extensive number of records. Complex searches of the database will provide researchers with information about the archival records and manuscripts—such as indexes, subject lists, and collection descriptions—of a repository that otherwise would either be unavailable or not feasible to create. Local systems can meaningfully increase access to information about a repository's archival records.

Online national databases will significantly affect not only the increase of information for researchers, but may also reduce the cost of doing research. For example, time spent traveling to examine records could be reduced and be more productive. Local systems will be an asset in this respect because materials that otherwise would remain largely unknown now have the potential to be available on a national basis. These systems, using the MARC AMC format, provide both a viable alternative and a supplement to the bibliographic utilities.

Automation and the MARC AMC format will have an enduring effect on the way archival repositories function. Both their internal and external operations will be affected and microcomputer-based local systems using the MARC AMC format will be a significant agent of change.