Microcomputer Applications in Archives: A Study in Progress

by RICHARD M. KESNER and DON HURST[1]

North American archivists have recently followed the example set by their colleagues in Great Britain in turning to the computer as an administrative tool and as an assistant in the management of information. Due to the costs involved and the expertise required for the proper handling of computers and machine-readable records, it is not surprising that both the Canadian and U.S. federal governments have taken the first steps in establishing automatic data processing (hereafter ADP) archives.[2] Even so, these on-going ADP programs are still in nascent stages of development.[3] In the area of automated archival

---

1 This essay draws upon material presented previously in workshops before the staff of the Public Archives of Canada and students attending its 1980 Archives Course (17 March 1980) and the Society of American Archivists (3 October 1980) as well as in a paper presented before the Society of Georgia Archivists (21 November 1980). The authors would like to acknowledge the generous support of the Research Programs Division of the National Endowment for the Humanities, East Tennessee State University, and the University of Colorado at Boulder in providing us with both the means and the release time to pursue this research project. The authors alone take full responsibility for any errors in fact or judgement in what follows.

2 See, for example, Charles M. Dollar, "Computers, the National Archives, and Researchers," *Prologue* 8, No. 1 (1976): 29-34; United States, National Archives and Records Service, *Catalog of Machine-Readable Records in the National Archives of the United States* (Washington, D.C., 1977), passim; Meyer H. Fishbein (ed.), *The National Archives and Statistical Research* (Athens, Ohio, 1973), passim; Everett O. Alldredge, "Inventory Magnetic-Media Records," *American Archivist* 35, Nos. 3/4 (1972): 337-45; and Jack Dennis, "The Relation of Social Science Data Archives to Libraries and Wider Information Networks," *Proceedings of the Conference on Interlibrary Communications and Information Networks*, Joseph Becker (ed.), (Chicago, 1971), pp. 117-120.

3 This position is certainly born out by the studies conducted by Carroll, DeWhitt, Geda, and others. M.E. Carroll, "Public Archives of Canada Survey of EDP Installations," *ADPA* 1, No. 2 (1974): 15-25; Ben DeWhitt, "Archival Use of Computers in the United States and Canada," *American Archivist* 42, No. 2 (1979): 152-7; Carolyn L. Geda, "Social Science Data Archives," *American Archivist* 42, No. 2 (1979): 158-66; Inter-University Consortium for Political and Social Research, the Bentley Library, and the University of Michigan School of Library Science, *Proceedings of a Conference on Archival Management of Machine-Readable Records, Held at the Bentley Library, the University of Michigan, February, 1979* (Chicago, 1980), passim.

management systems, there are even greater deficiencies and challenges.[4] However, activity on both sides of the Atlantic since 1975 suggests that the profession is about to enter the "cybernetic age."[5] Archivists are demonstrating a growing awareness of the potential benefits of automation in improving their services to their patrons and in eliminating some of the mechanical and repetitive aspects of archival administration.

Indeed, a poll of Society of American Archivist (SAA) members taken in late December 1979 clearly demonstrated a shift within the profession regarding the role and potential significance of the computer. When asked to list the five most pressing problems that they as archivists anticipated in the next five years, forty-five per cent of those polled indicated technological change as the major challenge in the years ahead.[6] Only scarcity of resources, the perennial problem plaguing archival institutions, gathered more votes than technology. This dramatic shift of interest towards automation has also manifested itself in the offerings of professional workshops and meetings. Each year since 1975, for example, has experienced a growth in the number of seminars and sessions at the annual meetings of both the Association of Canadian Archivists (ACA) and the Society of American Archivists devoted — at least in part — to automated indexing and retrieval or machine-readable records.

Perhaps the single most dramatic indication of this trend is reflected in the diversity and quality of professional publications pertaining to this emergent sub-field. In 1979, the *American Archivist* devoted an entire issue to data archives and computer-based finding aids. The journal has also published individual essays in these and related fields on a more frequent basis.[7] In addition,

---

4 H. Thomas Hickerson, Joan Winters, and Venetia Beale, *SPINDEX II at Cornell University and a Review of Archival Automation in the United States* (Ithaca, New York, 1976), pp. 23-34. MARC Development Office, *Manuscript: A MARC Format* (Washington, D.C., 1973). OCLC has recently established a "Manuscript Task Force," chaired by Helen Slotkin of MIT. This body is examining ways to improve and expand the MARC manuscript format so as to make it more useful to archivists. The latest reports from the task force indicate that some progress has been made in this regard. See Helen M. Slotkin, "Report of OCLC Manuscript Task Force," (xerox copy), 15 April 1980.

5 Frank G. Burke, "Report on a Survey of Automation Activities in Archives and Manuscript Repositories in the United States and Canada," *American Archivist* 31, No. 2 (1968): 208-210. Burke's article suggests that the archival profession was by the late 1960s nearing an active period of inquiry into the role of automation in archives. Unfortunately, neither his evidence nor subsequent professional activities support this proposition. The development of SPINDEX, a project with which Mr. Burke was intimately involved, was a notable exception to this state of affairs.

6 *SAA Newsletter* (March, 1980): 1-2. The membership of the SAA indicated that they were concerned with "keeping current with technological change, the impact of computers on archival theory and practice, especially in the area of finding aids, and new information formats [e.g. machine-readable records and computer-output microfilm] require retooling of archivists and archival operations."

7 The *American Archivist* 42, No. 2 (1979) includes articles by Ben DeWhitt and Carolyn L. Geda on data archives, and Alan Calmes and David Bearman on automated indexing and retrieval systems. Other important articles pertaining to computers and archives that have recently appeared in the *American Archivist* include: Charles M. Dollar, "Appraising Machine-Readable Records," *American Archivist* 41, No. 4 (1978): 423-30; Richard H. Lytle, "Intellectual Access to Archives: I. Provenance and Content Indexing Methods of Subject Retrieval," *American Archivist* 43, No. 1 (1980): 64-75; and Richard H. Lytle, "Intellectual Access to Archives: II. Report of an Experiment Employing Provenance and Content Indexing Methods of Subject Retrieval, *American Archivist* 43, No. 2 (1980): 191-208.

the Society of American Archivists has released a series of separate publications, including an annotated bibliography, a basic manual, and a volume of symposium proceedings.[8] The bibliography itself documents the growing research commitment among archivists regarding computers and machine-readable records.[9] Finally, one must not lose sight of equally important efforts by our colleagues abroad. In particular, it is noteworthy that the International Council on Archives (ICA) established a "working party on the implications of automatic data processing for archival management" in 1972.[10] This working party has served as an international forum for the discussion of archival problems involving automated records and techniques. ICA also publishes a bilingual (French/English) journal, *ADPA* (Automatic Data Processing in Archives).[11]

However, as all of these publications demonstrate, current efforts to develop computerized systems for archives focus upon the services of large, so-called "main-frame" computers and the partnering services of a well staffed computer programming/operating facility. This approach raises the obvious question of how these new systems might be employed in more modest settings, such as county or municipal archives, smaller universities and colleges, and historical societies? How is this large and certainly interested professional constituency to be brought into more direct and productive contact with automation? And more importantly, how are these archives and their staffs to obtain the training and experience in dealing with computers that so many appear to desire? It is the purpose of this essay to suggest an alternative course of research and development to the archivist's reliance on large, expensive, and often unavailable main-frame computers. The key to the future of archival automation lies in recent information processing trends that emphasize in-house systems, so-called "distributed processing systems," and in the arrival of the microcomputer as a powerful administrative and data management tool.

According to the advocates of distributed processing, the reliance on main-frame computers places the user at the mercy of his or her resident computer centre or a computer service bureau. They point out that the way to avoid the low-priority service usually accorded libraries and archives in such a setting is

---

8 Richard M. Kesner, *Automation, Machine-Readable Records, and Archival Administration: An Annotated Bibliography* (Chicago, 1980); H. Thomas Hickerson, *Archives and Manuscripts: Automation* (Chicago, 1980); and ICPSR, the Bentley Library, and the University of Michigan School of Library Science, *Proceedings of a Conference on Archival Management of Machine-Readable Records, Held at the Bentley Library, the University of Michigan, February, 1979* (Chicago, 1980).

9 Richard M. Kesner, pp. 4-10.

10 The ICA Working Party met in Spoleto, Italy, 23-25 May 1972. Meyer Fishbein of the National Archives and Records Service served (and still serves) as the U.S. representative on that body. For the policies and objectives of the ICA Working Party, see *ADPA* 1, No. 1 (1972): 1-15 (minutes); Appendix I (resolutions).

11 *ADPA* is published irregularly as issues become ready for press. The subscription rate is therefore calculated on three issues (currently U.S. $9) rather than upon the calendar year. U.S. archivists may obtain subscriptions to ADPA through Meyer H. Fishbein, Director, Military Archives Division, NARS, Washington, D.C. 20408 and Canadian archivists may contact Harold Naugler, Machine-Readable Archives Division, Public Archives of Canada, 395 Wellington Street, Ottawa, Ont. K1A 0N3 for more information.

to establish an in-house computer centre to meet the immediate needs of the institution.[12] Recent innovations in microcomputer technology have led to a situation where the affordable microcomputers can do many if not all of the tasks previously assigned to main-frames. Like its main-frame counterparts, the microcomputer has the ability to index, manipulate and retrieve data. It can generate forms, keep track of department finances, and manage user services and instruction. As a result, the microcomputer is rapidly becoming a common tool in businesses, government offices, and records management operations.[13] Unlike a main-frame computer system, however, the microcomputer is relatively inexpensive to purchase and maintain. Even so, only a few archives have actually considered microcomputer applications, and among these, the Smithsonian Institution Archives stands out as the first agency of its kind to take positive steps in this direction. The Smithsonian has just obtained a microcomputer to use in conjunction with their main-frame SELGEM system as a means of enhancing their ability to produce quality, computer-generated publications.[14] In its initial stages at least, the Smithsonian project will continue to rely on a large computer for the storage and manipulation of its collection description data base. The Institute's staff has not as yet committed the Archives to design and implement microcomputer programme packages (software) for use in other archival settings.

During the fall 1979 business meetings of the SAA Task Force on Automated Records and Techniques, the paucity of research and development in this area arose as a point of discussion. Following that gathering, the authors of this essay initiated a preliminary inquiry into the feasibility of employing microcomputers to address the management, and indexing and retrieval needs of small and medium-size archives. This in turn led to a successful grant application to the National Endowment for the Humanities that proposed:

---

12 The obvious advantage of distributed processing to an archives lies in the fact that the system would be housed within the archives itself and would be operated by archives personnel. Thus, the system would be utilized to serve the needs of the archives first and foremost. While the Public Archives of Canada might need a main-frame or minicomputer, the microcomputer can serve the information processing needs of most other Canadian archives. The literature pertaining to distributed processing varies widely. *The Office, Computer Decisions* and *Infosystems* are three monthly publications with regular features concerning DP. For an introduction to DP, see P.J. Down and F.E. Taylor, *Why Distributed Computing* (Manchester, 1976); and Judson Breslin and C. Bradley Tashenberg, *Distributed Processing Systems* (New York, 1978).

13 Computers Plus, Inc., for example, markets FMS-80, a file management system that is not only useful to records managers but may also prove applicable in various archival settings. *FMS-80, File Management System Users Manual* (Alexandria, VA, 1978, 1979). Apple Computers, Radio Shack, Commodore Pet, and other microcomputer manufacturers have all developed file management systems that operate on their machines.

14 SELGEM (*Self-Ge*nerating *M*aster) includes a series of information indexing and retrieval programmes originally developed by the Smithsonian Institution for the better management of its museum collections. It was subsequently adapted to the needs of the Institution's Archives. At this time, over two hundred museums and archives employ SELGEM in the management of their holdings. See A.L. Bain, "Computer Applications to Archives and Manuscripts at the Smithsonian Institution Archives," *ADPA* 2, No. 3 (1978): 13-21. Mr. Bain also prepared a report on the Smithsonian Archives' consideration of microcomputer and word processing systems. A more recent version of this paper was presented before an SAA workshop on microcomputers. See A.L. Bain, "SIA Systems Review," (draft report, xerox copy), October 1979.

(1) to consider the feasibility of microcomputer applications in archives in some detail,
(2) to develop prototype systems for the accessioning and description of archival collections, for the registry of researchers, for the management of archival personnel, budgets, and supplies, and for the automated searching of collection guides as a function of reference services, and
(3) to share findings with colleagues in the hope that others will build upon our example.[15]

From the outset, we found ourselves without the benefit of earlier research in this field. We therefore viewed the production of an introductory essay concerning the use of microcomputers in an archival setting as essential both in fulfilling our grant obligations and in laying the ground work for ourselves and others. This paper seeks to acquaint the reader with microcomputers and introduce MARS (Microcomputer Archives and Records Management Systems). While MARS is in its early stages of development, our plans, objectives and preliminary findings will, it is hoped, stimulate others to take up the challenge and design alternative automated indexing and retrieval systems of their own.[16]

MICROCOMPUTERS AS AN ARCHIVAL TOOL

Main-frame oriented, automated archival systems are often too costly and require considerably more equipment and specially trained personnel than are available to most archives. Certainly, many of the great U.S. research institutions will continue to use SPINDEX and its sister programming packages. But this approach to archival automation is inaccessible to the majority. Microcomputers offer a practical alternative in that they are affordable, portable, flexible, and with a little training and effort may be run by archivists themselves without recourse to computer specialists. Furthermore, microcomputers lend themselves more readily to interactive networking through telephone links. Simply put, it is the contention of this paper that the microcomputer is not only a practical solution to immediate information management and administrative problems confronting archives, but that it is also the key to many of our ultimate objectives as an international community of archivists and information managers.

A microcomputer is a device that can carry out many of the mathematical, analytical, and text manipulation functions of a large, main-frame computer. The salient difference between the two is that the central processing unit or brain (hereafter CPU) of the microcomputer is made up of small silicon chips

---

15 On 11 June 1980 the National Endowment for the Humanities awarded the Archives of Appalachia a grant to carry out this project and to share the results of this study with the greater archival community. During the period leading up to the grant award, Don Hurst moved from Head of Systems at the Sherrod Library, East Tennessee State University to Head of Systems, University of Colorado at Boulder Libraries, but along with Richard M. Kesner, Director of the Archives of Appalachia (ETSU), he remained as chief researcher on the project.

16 In this regard, we have already been contacted by the staff of the Institute Archives at the Massachusetts Institute of Technology concerning our work and encourage other interested parties to share information with us.

called microprocessors. These microprocessors are large-scale integrated circuits (LSI's) that often include over one thousand microscopic "gates" each through which electric current passes. By opening and closing these gates, microcomputers can store information in binary form, like that of any other computer. The significance of the microprocessor resides in its remarkable miniaturization. This feat greatly reduces construction costs while enhancing machine efficiency.[17] As a result of technological innovations over the past ten years, microcomputers are now quite powerful but cost a fraction of what one might pay for a main-frame configuration.[18] The low costs of microcomputers argues strongly in favour of their adoption by archives to meet their ADP needs.

Due to the competitive market within which most microcomputer manufacturers find themselves, the buyer of a microcomputer system has, like the purchaser of stereo equipment, a wide field from which to choose. One can often mix components from different manufacturers. Indeed, the growing adaptability of microcomputer "interface" units is making such a practice a more common occurrence.[19] The fierce competition and low production costs faced by microcomputer makers have led, as with pocket calculators, to a decline in prices. Furthermore, those archives connected with educational institutions will benefit from the growing desire among manufacturers to strengthen their standing in the educational market. As a result, discounts of from ten to twenty per cent on the purchase of equipment are not uncommon.

The plethora of microcomputer lines makes it difficult if not impossible to discuss the "ideal" computer system.[20] There are, however, elements that are common to most of these hardware (equipment) packages. First among these is the CPU or brain of the system. The CPU includes a memory unit of from 4K bytes to 64K bytes and a "scratch pad" where it manipulates data and carries out calculations.[21] The greater the memory capacity of the CPU the more it will cost. The microcomputer's brain is connected to a keyboard which acts

---

17 For a brief introduction to microprocessor and microcomputer technology, see Laurence Altman, *Microprocessors* (New York, 1975); Michael Hordeski, *Illustrated Dictionary of Microcomputer Terminology* (Blue Ridge Summit, Pa., 1978); Infotech International, *Infotech State of the Art Report: Microprocessors: I Analysis and Bibliography* and *Infotech State of the Art: Microprocessors: 2 Invited Papers* (Maidenhead, Berkshire, U.K., 1977); and Anne Carol, *Microcomputer Design* (Englewood Cliffs, N.J., 1978). Two of particularly high quality in this field are: John D. Lenk, *Handbook of Microprocessors, Microcomputers and Minicomputers* (Englewood Cliffs, N.J., 1979); and Charles J. Sippl and Fred Dahl, *Computer Power for the Small Business* (Englewood Cliffs, N.J., 1979).

18 When used in reference to a computer, the term "power" or "powerful" usually refers to the device's ability to manipulate large quantities of data at high speeds. Most microcomputers run off of standard 110 volt wall current. Microcomputer specialists predict that within ten years, we will witness the development of a small computer with the power of the currently popular IBM 370 main-frame. See Robert N. Noyce, "From Relays to MPU's," *Computer* 9, No. 12 (1976): 26-29 and Mark Shepard, Jr., "Distributed Computing Power: A Key to Productivity," *Computing* 10/11 (1977): 66-74.

19 See, for example, Jonathan A. Titus, *TRS-80 Interfacing* (Indianapolis, 1979); Howard Anderson, "The IBM Grand Strategy," *Mini-Micro Systems* 10/10 (1977): 20-22; and A. Nichols, Joseph C. Nichols and Peter R. Rony, Z-80 *Microprocessor: Programing and Interface*, Books I and II (Indianapolis, 1979).

20 For a useful introduction to microcomputer hardware, see Richard Didday, *Home Computers: 2^{10} Questions and Answers, Vol I: Hardware* (Forest Grove, Or., 1977).

21 Microcomputer memory capacity is usually expressed in terms of thousands of "bytes".

as the primary communication device between the user and the CPU. A television monitor usually serves as the link between the CPU and the operator, flashing the message typed into the CPU through the keyboard and relating the CPU's response. A hard-copy (i.e. paper) terminal can be employed in place of the television monitor for those desiring a permanent record of these transactions. In most instances, the archivist user will need to enter large quantities of data, such as finding aids and accession records, into the computer and will use either a cassette tape recorder or a disk drive for this purpose. Though more expensive, the use of a disk drive and floppy diskettes for data storage is preferable because they are so much faster and more efficient. Since archives may find it useful to retain hard-copy versions of searches and finding aids carried out by the microcomputer, the staff will in all likelihood desire the addition of a printer as well as a television monitor. Thus, from the perspective of maximizing the utility of a microcomputer system within an archives, such a configuration would most probably include: a CPU, a keyboard, a television monitor, one or more disk drives, and a printer.[22]

The availability of a wide range of computer systems means that the cost of a microcomputer configuration may vary considerably. In 1981, for example, CPU's may cost a few hundred to well over two thousand dollars. A television monitor may cost as little as fifty dollars or as much as six hundred dollars depending upon its size and whether or not the user requires a color-graphics capacity. Disk drives and printers also vary according to their design and performance abilities. Generally speaking, however, microcomputer systems fall into three basic price categories.[23] The first of these includes all systems ranging up to five thousand dollars in cost. The currently popular Radio Shack TRS-80, the Apple II Plus, and the Atari, all fall into this category. The second group, ranging upwards to ten thousand dollars, encompasses many word processing systems, such as those marketed by Wang, IBM and CMC. Finally there are microcomputer systems that cost as much as one hundred thousand dollars. Configurations in this category offer high-power graphics and design capabilities used primarily in the graphics, publications and engineering industries.

From the outset, we have sought to impress upon our readers the economic arguments in favor of the use of microcomputers in archives. In conducting our own study, we have therefore limited ourselves to the systems in category one. We chose an Apple II Plus system with 32K bytes memory, communicator card, dual disk drive, black and white television monitor, and a small, seven-inch paper, silent-type printer. The entire package cost us slightly less

---

Thus, 48K = 48,000 bytes and 16K = 16,000 bytes. Each byte contains 8 "bits" (i.e. binary pulses of either "1's" or "0's"). A single character, either numeric or alphabetic, requires one byte of storage if it is to be stored by the computer. A large archives finding aid might require 30K to 40K of system storage if it is to be saved in its entirety.

22 Microcomputer systems also require "interface" units that serve as connectors between the CPU and the systems various input/output devices. We found a communication board with interface also essential for communicating over telephone lines with other microcomputers as well as with main-frame systems.

23 A detailed breakdown of system costs may be found in the following: Richard Didday, pp. 200-203; and Charles J. Suppl and Fred Dahl, pp. 90-105.

than three thousand two hundred dollars. A comparable, though not quite as powerful, Radio Shack system would have cost slightly less and a more versatile Atari system slightly more. The important fact to bear in mind is that there is no ideal or perfect system. The configuration that meets the needs of the archives under consideration is the "ideal" microcomputer for that setting, and our readers should therefore evaluate their own needs most carefully before purchasing a system. Whatever you decide, it is unlikely that your system will cost your organization more than four thousand dollars. When compared with the cost of a main-frame computer, this is a small amount indeed. The savings realized by automating will quickly recoup this modest investment.

Beyond the cost argument, there are other reasons why the microcomputer offers archivists an attractive alternative to main-frame systems. First among these is the ease with which one can learn to operate and program microcomputers. Most systems operate in BASIC, a high-level programming language that is structured to approximate the English language. BASIC allows the user to communicate with the computer with relative ease and to design programs resembling the conversational structure of the English language.[24] In addition, due to the popular, non-professional nature of a large segment of the microcomputer market, the shelves of most local bookstores contain a variety of helpful material on the programming of microcomputers.[25] While these readings are not particularly conversant in those areas of greatest interest to archivists, they do provide an adequate starting point and may be supplemented thereafter by the more specialized volumes found in most well-stocked university libraries.

Within a year or two, most microcomputers will have the capacity to operate with PASCAL as well as BASIC.[26] PASCAL affords the programmer greater flexibility in terms of text manipulation. This is particularly important to archivists concerned with design problems related to the description, indexing and retrieval of manuscript materials. Though one may employ BASIC to carry out the same tasks that PASCAL is capable of executing, it requires

---

24 The literature treating BASIC is quite substantial. A few suggested readings include: Bob Albrecht, LeRoy Finkel and Jerald R. Brown, *BASIC for Home Computers: A Self-Teaching Guide* (New York, 1978); Radio Shack, *LEVEL II BASIC Reference Manual* (Fort Worth, Tex., 1978); John G. Kemeny and Thomas E. Kurtz, *Basic Programming*, 2nd edition (New York, 1971); and William Baden, Jr. *How to Program Microcomputers* (Indianapolis; 1977).

25 See, for example, Richard Didday, *Home Computers: 2^{10} Questions and Answers, Vol. II: Software* (Forest Grove, Or., 1977); Christine Doerr, *Microcomputers and the Three R's: A Guide for Teachers* (Rochelle Park, N.J., 1979); Nancy B. Stern, *Flowcharting: A Tool for Understanding Computer Logic* (New York, 1975); and Mitchell Waite and Michael Pardee, *Microcomputer Primer* (Indianapolis, 1976).

26 At this time PASCAL is not as easily transportable as BASIC. Even among those manufacturers, such as APPLE, who market PASCAL cards, the added costs and memory requirements do not make it particularly attractive. But within a few years, PASCAL's time will come and it may ultimately eclipse BASIC as the universal microcomputer language. For an introduction to PASCAL, see I.R. Wilson and A.M. Addyman, *A Practical Introduction to PASCAL* (New York, 1979); Jim Welsh and John Elder, *Introduction to PASCAL* (Englewood Cliffs, N.J., 1979); and Michael G. Schneider, Steven W. Weingart, and David M. Perlman, *An Introduction to Programming and Problem Solving with PASCAL* (New York, 1978). This last volume is perhaps the best single book currently available on the subject. The sections on programme development are excellent.

greater programming ingenuity and longer, more involved programs. As PASCAL becomes as universal as BASIC, microcomputers will become more flexible and powerful archival tools. With their wider availability, the cost of equipping a microcomputer with PASCAL capabilities will also diminish.

A final attraction of the microcomputer as a tool for archivists relates to the often discussed notion of a national or international archival information network. Initially, SPINDEX promised to serve as an automated link between archives but the package shows little sign of developing in that direction. Nor does SPINDEX or any of its sister systems, with the possible exception of SELGEM, hold out the hope of an interactive capability within the foreseeable future. Microcomputers by contrast are interactive, highly flexible and have such a similarity in design to suggest that networking is within their information management capacities. We do not as yet have microcomputer software that compares with SPINDEX or SELGEM, but the machines have the ability to provide links through the use of telephone lines between archives. If the programming packages are developed to meet the needs of archives and manuscript depositories, the microcomputer may very well provide our profession with the means of communicating and sharing information on a large scale. The evolution of MARS and its implementation as a prototype system for archival administration and information management serves as a useful first step in the realization of these objectives.

THE MARS PROJECT

The Microcomputer Archives and Records Management Systems (MARS) at the Archives of Appalachia grew out of a desire to explore the possibilities of employing microcomputers in archival settings. From the outset, the project staff developed systems that are highly portable and largely free from hardware dependence. In selecting a computer with which to carry out this project, we restricted ourselves to a system that did not exceed four thousand dollars in cost. We chose the Apple II Plus because it appeared at the time to meet our financial and research criteria. In the early stages of our study, the staff designed a series of working models but our ultimate objective is the creation of a so-called "turn-key" software (computer programme) package that other archives may use with ease. As a result, MARS will perform a variety of standard archival operations without any need for significant restructuring by a particular user. Certain changes in the package will prove necessary in transferring it from one type of microcomputer to another and formats may require alteration to bring them in conformity with differing archival practices. MARS will nevertheless serve as both a guide post and model to others which they may adapt to meet their own specific needs.

Portability (the ability of the software to work in a number of different archival settings) remains a key element in MARS development. For this reason, the staff chose the use of BASIC over that of PASCAL as our programming language. PASCAL does lend itself to the types of design problems we face, particularly in the area of text manipulation. However, only a limited number of microcomputers currently have PASCAL capabilities and this requires additional memory capacity (from 16K bytes to 32K bytes), and hence

great hardware costs. To keep the costs of equipment within our self-imposed limits and to insure the greatest degree of portability, we have therefore decided to use high-level BASIC as our programming language. We will be able to do the same sorts of things with BASIC that we had hoped to do with PASCAL and can upgrade our systems to the latter as it becomes more universal and less expensive.

While BASIC is common to most microcomputers, each manufacturer adds his own enhancement that make BASIC software hardware-dependent. System commands, for example, such as PRINT, READ, AND RUN, are common to most versions of BASIC but different brands of BASIC (eg. BASIC Level II or Applesoft) respond differently in terms of the way the CPU will interact (interface) with system input/output devices, such as disk drives and printers. A program written in BASIC Level II will therefore require certain modifications if it is to be run on an APPLE computer. This problem, by the way, is also faced by those transferring a software package from one mainframe computer to another. Fortunately, manufacturer enhancements do not constitute a major obstacle to systems portability. Through the development of a special "scanner" routine, the recipient hardware configuration can edit the incoming software to make it compatible with the computer system. Some of our readers may weary of these technical considerations, but these concerns are central to any appreciation of the types of problems confronting archivists engaged in this line of research and are therefore one of the overriding interests of the MARS research staff. Most manufacturers and users of microcomputers are not overly worried about the portability of software packages, except perhaps within the same brand-name system. Our project staff, on the other hand, assumes that archivists will require systems with the greatest degree of flexibility if they are to address the varying needs of archives. We have structured our research and development efforts accordingly.

Economy and portability remain basic design objectives, but the key element in the actual creation of microcomputer software for archives is, in our view, the standardization of archival procedures. We recognize that some of our colleagues dislike standardization as an unnecessary restriction on the activities of archivists. Admittedly, much of our work does not lend itself to formal structuring. However, a modicum of standardization is essential if the computer is to be employed and such rigour can be imposed without inhibiting the creative aspects of the archivist's job. By structuring the format of various archival procedures, such as accessioning and collection description, we can allow considerable flexibility to enter into the process while feeding information into the computer in a form that it can digest. This may require archives to reduce the number of formats that they follow for a given procedure, but in the long run, this is preferable to a myriad of different formats even if the archives decides not to resort to the computer.

The Archives of Appalachia has standardized its procedures to the extent that each stage of processing, from acquisition to description, is governed by a particular form. We have adopted these practices in anticipation of automating our activities. The use of standardized formats facilitates the transition to machine-readable record keeping and allows for a more economical conversion from the former to the latter. Our accessioning procedures, for example,

are governed by one of two forms — a general form for all collections and a special form for photographic collections. The Archives also employs a single finding aid format. This does not mean that each guide reads the same but rather that the provenance statement, the index, and the box and file folder inventories are always located in the same place within each guide. At times, we have added appendices to finding aids in an effort to accommodate special materials, such as video tape recordings, but we have never deviated from the general format. Thus, we are in a position to transfer the contents of our finding aids to diskette without any significant reworking of the original. The wording of any particular entry will vary widely but the order in which this information enters the computer system remains constant. Standardization along these lines is essential if archivists are to use microcomputers or for that matter main-frame computers with any benefit to their operations. The exact format employed is at the discretion of the individual institution; ours are cited in this essay only as the basis for discussion. Nevertheless, the careful structuring of archival practices with a view towards their eventual conversion to a machine-readable form will considerably improve the chances for success of any plan to automate these procedures.

With our administrative records and finding aids in standardized formats and with our general design objectives in mind, the project staff found themselves in a favourable position to commence with the actual structuring of MARS. We began with accessioning because the process lends itself most readily to machine manipulation. During accessioning, the Archives assigns a unique number to each new collection. As newly arrived materials are added to existing archival holdings, they are listed on the original accession sheet and do not receive a number of their own. We also maintain fixed shelving assignments for our collections though any given collection may be stored in different locations within our facility. If a collection comes to us piecemeal, the various deliveries are usually stored separately until processing begins. As long as all of the shelving locations are noted on the accession sheet, there is no real danger in this arrangement. It avoids back-shifting in the stacks and the wasting of storage space on the anticipated arrival of additions to archives collections. In our view, MARS ought to be able to enter new accessions, assign accession numbers and shelf locations, up-date accession records, and provide regular accountings of the accession status of any given collection as well as of the availability of space within the storage facility as a whole. These objectives have directed us in our design of the MARS sub-routine "ACCESSION."

In terms of MARS user services, we sought two complementary objectives. First, we planned to employ the microcomputer to enter Archives' finding aids onto diskettes. All corrections and editing would occur at this stage. Once completed, the guide could be examined by researchers on the television monitor or in hard-copy form through the use of the system's printer. The MARS sub-routine "ENTER COLLECTION" serves this function while allowing Archives personnel to correct and add to existing guides as the need arises. In addition, this sub-routine will generate listings of all archival holdings by collection name and could ultimately serve as the basis for a program to produce a comprehensive and detailed publication of our holdings. "ENTER COLLECTION" will also provide us with the ideal data base for researcher inquiries. However, we currently face the problem imposed by the size of our

finding aids, ranging in some cases from thirty to forty thousand characters, and the limited capacity of our storage medium (i.e. floppy diskettes). Indications are that diskette technology is about to take a major leap forward.[27] When it does, we will have the ability to carry out complete textual searches of collections down to the folder level.

Until that time, we must content ourselves with a more limited data base that nevertheless allows researchers to survey our holdings. To achieve these ends, we plan to enter a significantly reduced set of vital data elements about each of our collections into the computer system. In this sub-routine each collection guide will include donor name, collection size, span dates, key indexing terms, such as major correspondents and subjects covered in the collection, and a brief abstract. All of these data points are already standardized components of our collection guides and will therefore require little additional work prior to their actual entry onto diskettes. The "QUERY COLLECTION" subroutine will then search this more limited data base for as many as thirty terms linked by "ands" and will identify, and if the user so wishes, print out the capsule descriptions of those collections netted in the search. The researcher may then call up more detailed guide entries through the use of "ENTER COLLECTION."[28]

The final sub-routine of the MARS package is in fact a series of small programs under the general heading of "ADMINISTRATIVE SUBSYSTEMS." These sub-systems include a program for the management of staff work schedules, for the control of supply and equipment inventories, for budgets and financial management, and for researcher registration. In many respects, all of these sub-routines assist in the day-to-day administration of the Archives. The staff schedule program, for example, keeps track of employee hours, sick leave and vacation time, staff work assignments and locations, benefits, and salaries. In a matter of seconds, it can alert an administrator to staff shortages or the status of a particular worker within the Archives. Those institutions with workers under a variety of different funding formulas, such as CETA, work-study or grants, will find considerable use for this program. Similarly, the supply inventory and budget programs are designed to indicate the rate of consumption and current status of various supply categories and financial accounts. The supply program automatically alerts the staff when materials have dropped below certain thresholds.

---

27 See A.J. Kolk, Jr, "Low-Cost Rotating Memories: Status and Future," *Computer* 9, No. 3 (1976): 30-34; Kent Winton and James L. Patterson, "Hard Disk Drives," *Digital Design* 8, No. 12 (1978): 20-28; and William R. Miller, "Floppy Disk Drives," *ibid.*, 30-32.

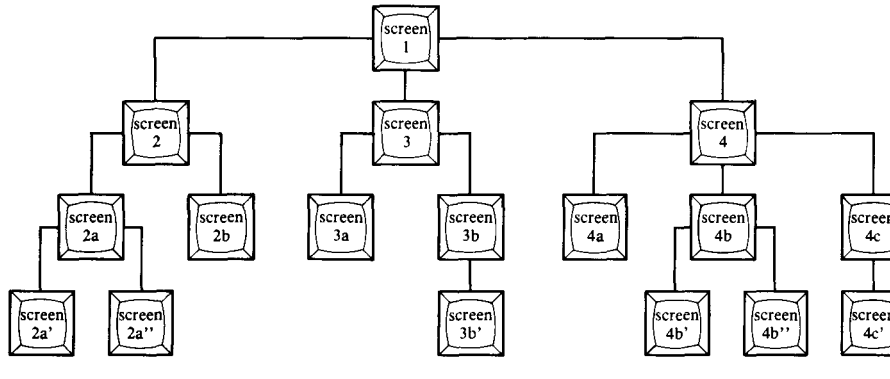
28 For the moment, we are limiting MARS's data base search structure to linked "and" statements, such as "SEARCH: Kentucky AND Mining AND Accidents AND Insurance." As we gain experience, we plan to incorporate more sophisticated boolean search strategies into the software package. In effect, MARS will establish two separate data bases for the Archives. The first will describe our holdings at the collection level ("QUERY COLLECTION") and the second will describe our holdings at the box and folder as well as the collection level ("ENTER COLLECTION"). Both of these sub-routines will have free-text searching and full-text retrieval. The Smithsonian Institution Archives uses a similar arrangement for the storage of information on their collections. See A.L. Bain, "Computer Applications to Archives and Manuscripts at the Smithsonian Institution Archives," *ADPA* 2, No. 3 (1978): 13-21.

The researcher registration program is more complex than the other components of the "ADMINISTRATIVE SUBSYSTEMS" sub-routine. This program records information about the patron including: name, address, institutional affiliation, and research interest. It also records patron use of the archives in terms of date, time, seat location, and collections examined. This information is currently maintained in a paper format and its conversion to a machine-readable format will allow us to improve user services while enhancing physical control over the collection in our care. The information gathered during this process will also provide us with the basis for a quantitative evaluation of the research objectives of our patrons. With this data, we will be in a much better position to train our reference staff to deal with research inquiries. Furthermore, we will always have on hand up-to-date statistics concerning our operations for those times when we are obliged to report upon our activities to funding agencies or the university administration.

In summary, the project staff anticipates the establishment of four sub-routines within MARS, "ACCESSION," "ENTER COLLECTION," "QUERY COLLECTION," and "ADMINISTRATIVE SUBSYSTEMS." MARS is designed for use of a moderately priced microcomputer system and is largely independent of any particular hardware configuration. The software package is written in high-level BASIC but will eventually operate in PASCAL. If the formats employed in MARS are not well suited to the needs of a given archival institution, the package can at least serve as a useful models which others may adapt as they develop microcomputer software of their own.

PROJECT PROGRESS

The project staff has proceeded with MARS according to what is referred to among programmers as a "top-down" approach. This process is perhaps best visualized as a series of logically linked screens. Each screen is hierarchically linked to the screens above and below it:



In this drawing, screen 1 represents the top, directional unit of the program which guides the user to the sub-routine of his choice. Screens 2, 3, and 4 represent the sub-routines and the frames under them signify the further elaboration of their respective programs.

This approach to computer programming, often referred to as a Hierarchical Input Process Output or "HIPO" structural style, allows the software system to evolve as a series of self-contained sub-routines. Each part stands alone and is documented separately.[29] Thus, an error in one sub-routine will not interfere with the operation of the remaining parts of the program. In addition, the logical structure common to the top-down approach and the opportunity it affords to address difficult programming problems individually make it more appealing to non-specialists. Our staff prefers HIPO structuring because it relates more directly to the way the human mind considers problems and because programming errors are more easily corrected. This decision has significantly contributed to the progress of our efforts in developing MARS.

The "top" or initial screen of the MARS pyramid indicates the various sub-routines available to the user and asks him to choose the routine he desires:

(SCREEN 1)
1. ACCESSION
2. ENTER COLLECTION
3. QUERY COLLECTION
4. ADMINISTRATIVE SUBSYSTEMS
ENTER NUMBER (1-4)? _____

In this example, the user is interested in registering a new patron to the archives. Since the program for research registration is located within the "ADMINISTRATIVE SUBSYSTEMS" sub-routine, the user would indicate "4" in the space provided. Upon striking the "RETURN" key on the terminal, the computer will shift to screen number four, indicating the various programs that fall under "ADMINISTRATIVE SUBSYSTEMS", thus:

(SCREEN 4)
ADMINISTRATIVE SUBSYSTEMS
1. WORK SCHEDULES
2. SUPPLY INVENTORIES
3. BUDGET
4. RESEARCHER REGISTRATION
ENTER NUMBER (1-4)? _____

Again, the user would indicate "4" in the space provided and would then press the "RETURN" key. Upon receiving this instruction, the computer will present the user with a formatted screen showing a blank researcher registration form. Once the user has completed the form, MARS will assign the patron a researcher number and store the information on diskette. If the form should be recalled to add further information, such as the collections examined, the completed form will appear on the screen. The system will also take certain general information from the form, like the patron's research topic, and store it separately as part of MARS's overall statistical gathering capabilities.

---

29 Large scale information systems, such as the *New York Times* Data Base and to a certain extent SELGEM, also employ the HIPO structural style in their programming.

MARS employs a combination of query and formatted screens. The query screens, as illustrated above, ask questions and give the user a number of optional responses. The formatted screens are employed primarily to imput data into system memory or storage. These screens resemble the paper forms currently in use at the Archives of Appalachia for accessioning, researcher registration, and collection description. A blinking cursor is used to indicate which part of the form is to be completed next. As the user finishes a given section, he presses the "RETURN" key and the cursor automatically skips to the next section of the form. Formatted screens with their accompanying blinking cursors facilitate the rapid and accurate entry of large amounts of data into the system. Once trained, a clerical member of the staff could easily maintain the operation with only limited supervision. Query screens by contrast give MARS a less structured, informal appearance. Research inquiries, for example, are carried out in the query mode, moving from one set of questions, to a structured search strategy, to a listing of collection citations. By moving between query and formatted screens within the HIPO structure, MARS offers visual variety that both staff and patrons appreciate.

The structure, format and sequence of MARS screens is governed by the system's "resident programme" — the actual controlling mechanism within the software package. The resident programme will, for example, instruct the user as to which diskettes are required for a given research inquiry or administrative procedure. It will also inform the user of the order in which the diskettes are to be played. The shifting of diskettes itself is not as cumbersome as the reader might suppose. In the first place, many functions, such as accessioning, do not entail the creation and maintenance of lengthy records. Neither these files nor researcher registration forms are likely to include more than one thousand characters (i.e. letters or numbers) per record, and in the case of accessioning, perhaps five hundred characters is closer to the mark.[30] This means that one can store anywhere from 116 to 232 accession records on a single diskette or as many as 464 accession records on two diskettes mounted for use on a dual disk drive. Thus, current technology affords sufficient storage capacity for most administrative records. The actual shifting of diskettes is no more difficult than changing a phonograph record and the resident program will tell the user when to change diskettes and which ones to select. As mentioned above, diskette technology may shortly come to our assistance with new storage devices of such enhanced capacities that these concerns will be eliminated.[31]

Prototypes for MARS "ACCESSION" and "QUERY COLLECTION" sub-routines are already in operation. The programs of "ADMINISTRATIVE SUBSYSTEMS" are in various stages of development. Progress is slow because each new step brings the project staff up against new problems. For example, we recently devised a means of programming around BASIC's treatment of commas so that we can feed long character strings, such as those employed in collection descriptions, into the computer without a loss of data

---

30 The term "record" is employed among computer users to refer to one discrete set of data within a "file" of data. Thus, one completed accession form, when stored on diskette becomes a "record" within the "file" of accession forms.

31. See note 27.

or cumbersome inputting procedures. Another problem that we are dealing with at present concerns "key board bounce," caused by the rapidity with which the CPU scans the keyboard. There are various ways of dealing with this problem but the most economical is through the use of a "debounce" software package. We expect to eliminate this mechanical source of record error within the next few months.[32]

Within the course of the next year, the project staff plans to complete and test the remaining sub-routines of MARS and to mount the system for testing at the Archives of Appalachia and perhaps the University of Colorado Archives. Beyond this formidable challenge, the staff must find an effective means of evaluating the system from both an in-house and a patron perspective. Much has been written about systems analysis, but the refinement of these procedures for use in an archival setting needs further consideration.[33] Finally, the project staff will need to develop an appropriate format for the duplication and the dissemination of research findings and the MARS software package. At present, Greenwood Press has expressed interest in distributing MARS as a module within a larger set of programs devoted to microcomputer applications in a library/archives environment.

PROSPECTS

The MARS project has already demonstrated the economy and flexibility of these machines when compared to their sister main-frame systems. Microcomputer technology is moving at such a pace that we can anticipate within the next five to seven years the marketing of inexpensive micros with the power of large IBM 370 main-frames. As an operating system, MARS or something like it will provide archives staffs with the ability to automate record accessioning, patron registration, and collection description. We will use microcomputers as word processors to edit archives publications and guides before they go to press and to solicit financial support and process grant applications. The reader may ask, well, is all of this going to save us money? The answer to this question is both yes and no.

The microcomputer is itself an affordable tool which certain archives could employ in place of current arrangements with a centralized computer centre or service bureau. The benefit of such an arrangement would be more in the area of direct control over information processing than in the saving of dollars. Certainly, the use of a microcomputer as a word processor will save archives staff time and hence money and will be providing them with higher quality publications. However, the true worth of a microcomputer or for that matter of any computer in an archival setting ought to be measured not in its labour-saving ability but rather in its capacity to manage information. In today's world, our patrons want greater and faster access to information. Library

---

32 Though there is nothing in the literature that actually addresses this problem, microcomputer manufacturers are aware of key board bounce and produce a pressure-sensitive keyboard with no moving parts that eliminates it. However, the cost of this piece of equipment is currently quite high and the small debounce program, taking from fifty to seventy bytes of memory, appears to work just as well.

33 For an excellent example of systems evaluation in a library setting, see Eva L. Kiewitt, *Evaluating Information Retrieval Systems* (Westport, Conn., 1979), pp. 89-163.

services with their increasing reliance on automated data bases have attempted to keep up with these trends. Within the archival profession, we are at last discussing a common national and international data base for archival and manuscript materials. The microcomputer has a role to play in this new environment. Through the use of sophisticated software packages, archivists can exploit the microcomputer to search their own holdings and to communicate their patron's requests to other repositories. In-house, the microcomputer can afford greater intellectual control over collections; internationally, it can serve as the conduit for information pertaining to the holdings of sister institutions. The promise is there and the capability is there. All we need do as archivists is find the means to exploit it.

Résumé

Cet article se veut une démonstration des possibilités que peut offrir le micro-ordinateur dans un contexte d'archives, aussi bien en ce qui touche le contrôle des dossiers que l'accès aux documents ou la compilation de rapports statistiques. Les auteurs donnent un compte-rendu du progrès de leur projet expérimental MARS aux Archives des Appalaches de la "East Tennessee University".