

# Automation in North American College and University Archives: A Survey

LEON J. STOUT *and* DONALD A. BAIRD

**Abstract:** The authors summarize survey results covering six topics: responsibility for machine-readable records, campus administrative databases, automated control systems, use of word processing, entry of holdings information in library network databases, and the place of archives in library automation planning. Almost one-fifth of the 979 institutions surveyed were active in one or more of these areas. Findings include the need for more emphasis on appraisal and preservation in the development of machine-readable archives, the probable increase in the use of microcomputers and database management systems for internal control, and the probable development of closer integration of archives and libraries as library automation begins to be more accommodating to archival needs. Educational issues remain highly important. While archivists may not need to be expert programmers, a greater grasp of computer capabilities and the environment of computer professionals is needed both to preserve the machine-readable record and to utilize automated techniques in physical and intellectual control.

*About the authors:* Leon J. Stout has been head of the Penn State Room and Associate Librarian in Special Collections at The Pennsylvania State University since 1974. He received his B.A. and M.A. degrees in history from Pennsylvania State University in 1969 and 1972, and the M.L.S. degree from the University of Pittsburgh in 1974. He is a member of the Society of American Archivists, Mid-Atlantic Regional Archives Conference, Pennsylvania Historical Association, and Society for Ch'ing Studies. He has served on the SAA Automated Records and Techniques Task Force since 1977 and has been a frequent speaker at SAA and MARAC meetings. Stout has most recently served as consultant to the Pennsylvania State Historical Records Advisory Board and prepared the state's historical record assessment report.

Donald A. Baird is University Archivist and Adjunct Professor of History, Simon Fraser University. He previously served as University Librarian at Simon Fraser and Assistant Librarian at the University of Alberta. He received his B.A. from the University of British Columbia in 1950 and M.S. from Columbia University in 1951. He has served as president of the Association of British Columbia Archivists (1982-83) and is a member of the Association of Canadian Archivists, the Society of Archivists, and the Society of American Archivists. He has served on the SAA Automated Records and Techniques Task Force since 1977 and is the author of *The English Novel, 1578-1956* (1959).

The authors wish to acknowledge the financial assistance of The Pennsylvania State University Libraries, Simon Fraser University Library, and the Society of American Archivists in conducting this survey.

ONE OF TODAY'S MORE COMMONPLACE TRUISMS is that an information revolution has swept through large organizations. The direct stimulus for this revolution has been the growing use of computers to store, process, and make available information of all kinds. Perhaps more than any other sector of modern society, colleges and universities are witnessing a great surge in the use of computers. Instructional, research, and administrative uses of computing power are now the most dynamic areas of higher education. While the archives traditionally has trailed other institutional branches in the adoption of new technology, there are, nevertheless, many striking new developments in automation on the college and university archival scene.

In order to ascertain exactly what those developments are, we conducted a survey among North American college and university archives. In particular, we sought information on three general subjects. The first was archival responsibility for machine-readable records and computerized storage media. In addition to this traditional approach, however, we also examined the relationship of archives to the growing area of on-line databases, or automated management support systems, in colleges and universities. The second subject of the survey was the use of automation in the intellectual and physical control of holdings by means of internal automated control systems and the use of word-processing and other similar systems for the preparation of finding aids. Third was the sharing of holdings information in library or archival databases and the consideration of archives in the planning of library automation systems on campus. This

topic is of particular interest in view of recent developments concerning the MARC Archival and Manuscripts Control (AMC) format.

The survey was undertaken as a project of the Automated Records and Techniques Task Force of the Society of American Archivists. The task force has previously undertaken similar surveys of state archives and business archives.<sup>1</sup> Planning is also under way for a survey of automation in religious archives. These surveys to determine the extent and nature of automated support systems, machine-readable records, and automated finding aids are part of the work of the task force in establishing a baseline of trends and activities for future planning and study. In particular, they should help the task force determine the profession's need for education in automation and enable the task force to plan program sessions, workshops, and other educational activities to meet these needs.

### Review and Methodology

Several surveys of college and university archives have been made since 1949, but only the most recent, made by Nicholas C. Burckel and J. Frank Cook in 1982, even mentions computers. The Burckel-Cook survey included computers as one of the choices of primary finding aids, and an average of 3 percent of the repositories reported using them. All were classified, according to the survey's system of categorization, as large, public universities.<sup>2</sup> This result seems to confirm the commonly held view that there is very little automation activity in college and university archives. Further results from the Burckel-Cook survey, showing that a typical repository

<sup>1</sup>Ben DeWhitt, "Archival Users of Computers in the United States and Canada," *American Archivist* 42 (April 1979): 152-57; and Richard M. Kesner, "Automated Records and Techniques in Business Archives: A Survey Report," *American Archivist* 46 (Winter 1983): 92-95.

<sup>2</sup>Nicholas C. Burckel and J. Frank Cook, "A Profile of College and University Archives in the United States," *American Archivist* 45 (Fall 1982): 425.

has a budget of \$21,000 (median of \$7,500), would seem to indicate that few if any repositories could afford an automated system regardless of how anxious they might be to install one.<sup>3</sup> Nevertheless, the very limited information concerning automation activities does not give a full picture, nor could a sample of only 10 percent provide as much data as a study covering all college and university archives could for future planning and evaluation. Furthermore, the archival literature has provided little substantial information on automation in college and university archives. Hence, this survey was undertaken to fill the void.

Our wish to retrieve as much information as possible from the population of college and university archives had to be tempered by the knowledge that our resources were not unlimited and that there might be only a small amount of activity to be discovered. Therefore, a two-level survey was chosen for the project. In late 1981, a survey containing six questions was mailed to 939 U.S. and 40 Canadian repositories identified in the *SAA Directory of College and University Archives in the United States and Canada* (1980). The six questions each contained explanations and qualifiers, but in essence they asked repositories about

- 1) responsibility for machine-readable records;
- 2) access to, or responsibility for, institutional automated management support systems;
- 3) automated control systems for archives;
- 4) automated means of preparing finding aids;
- 5) entry of bibliographic records of archival holdings into automated library or archival databases; and

6) consideration of archives in planning or operation of local automation systems in college libraries.

Archival units were asked to respond by returning a postage-paid card if they could answer that they were active in one or more of the six areas. By mid-1982, responses had arrived from 167 U.S. and 22 Canadian archives (a 19.3 percent total rate of response) that were active in one or more of the areas of automation. The rate of response alone was surprising and was very encouraging.

A second-level survey was then sent to each of the 189 institutions that had responded in the first round. To lighten the burden on the archivists, these questionnaires were modularized. That is, in-depth questions relating to each of the six basic question areas were placed on separate, color-coded forms, and only those forms corresponding to a positive response on the first-level survey were sent to each institution. Thus, if an institution had reported involvement in only the first and second areas, it would receive in-depth survey forms relating only to those areas. Responses to the second-level survey were received and tabulated by mid-1983 from 96 U.S. and 12 Canadian institutions (a 57.1 percent total rate of response, representing about 11 percent of the original first-level population.) Responses to the two questionnaires are tabulated in Table 1.

### **Machine-Readable Records and Management Information Systems**

The area of machine-readable records (MRR) and automated management support systems, or management information systems (MIS), is probably the most problematic aspect of overall operations for archives today. In his 1978 study of

<sup>3</sup>Ibid.

Table 1--Survey Response

	Level 1			Level 2		
	U.S.	Can.	Total	U.S.(%)	Can.(%)	Total(%)
Machine-Readable Records	69	12	81	37 (54%)	5 (42%)	42 (52%)
Automated Mgt. Support Systems	9	4	13	2 (22)	1 (25)	3 (23)
Automated Control Systems	33	8	41	14 (42)	7 (88)	21 (51)
Finding Aids Preparation	40	7	47	29 (60)	5 (71)	29 (62)
Holdings in Library Network	89	2	91	38 (43)	1 (50)	39 (43)
Consideration in Library Planning	102	14	116	57 (55)	7 (50)	64 (51)
Total Responses	167	22	189	96 (58)	12 (55)	108 (57)
% of total population	(18)	(55)	(19)			

state archives, Ben DeWhitt found that fewer than one-third of the states and provinces had begun to schedule MRR and that only a handful had actually accessioned tapes or other storage media.<sup>4</sup> Similarly, in his survey of business archives in 1982 Richard Kesner found that, of the 23 companies that kept MRR, in only one was the archivist responsible for the preservation of data files.<sup>5</sup> Maris A. Vinovskis, in his article on preservation of MRR in colleges and universities, noted that in most institutions no systematic effort was being made to save administrative MRR. He concluded that "we are unintentionally erasing our past even though as academics most of us have a strong personal commitment to the value of the preservation of information and the needs of scholarly analysis."<sup>6</sup>

The area of automated management support systems is even more troublesome for archives. These integrated, on-line database systems are becoming more commonplace in colleges and universities as administrators try to take advantage of technological advances to make their activities more cost-effective or responsive to student demand. Pennsylvania State University, for example, has undertaken an ambitious Administrative Information System project to integrate its various databases for records of students, alumni, accounting, facilities, business services, budget, and planning and to make them available on-line at all twenty-two campuses. A number of issues for archivists arise here: What version of a constantly changing file should be preserved? How can the appearance of a data file at the time an important decision was made

be recreated for the researcher? Other concerns in the area of MRR include the availability of documentation and software, the obsolescence of hardware, and general access to computer resources.

In the area of responsibility for MRR, the survey revealed that a surprisingly large number of institutions deals with this issue, and that the typical institution is active in this area for about seven years. A significant proportion of these archives have records management functions combined in the same unit (54.7 percent); and in almost two-thirds of the repositories, it was the archivist or the records manager who had initiated work with MRR. It is important to note that work with MRR is not usually considered to be a specific assignment: in 92.1 percent of the responding archives, the activity was judged to fall within the normal responsibilities of the unit.

Nevertheless, storage appears to be the primary characteristic of the activity. While two-thirds of the repositories had accessioned records (primarily, tapes and punched card decks in equal proportions), only about one-fourth reported any use of these records. In all cases the use was administrative and averaged only 4.1 uses in the preceding year. Access to the data files was primarily through traditional archival descriptive practices (30.4 percent), with half as many relying on existing documentation, library cataloging, or newly created record format definitions. More than half had no documentation at all. When queried about the lack of use, archivists cited restrictions on access to records, lack of means of access, and lack of user awareness of availability of the records.

<sup>4</sup>DeWhitt, p. 155.

<sup>5</sup>Kesner, p. 94.

<sup>6</sup>Maris A. Vinovskis, "Are We Erasing Our Past? Research Problems and Opportunities with Machine-Readable College and University Records," in Carolyn L. Geda, Erik W. Austin, and Francis X. Blouin, Jr., eds., *Archivists and Machine-Readable Records* (Chicago: Society of American Archivists, 1980), p. 44.

Even though we should be pleased to cite even storage as a primary activity at this point, the repositories have far to go in providing storage in an appropriate fashion. In general the archives did not make special arrangements for these records. Only two repositories indicated the presence of special vaults or storage cabinets, while two others had racks to store tapes. Nearly two-thirds took no special measures for preservation. Six institutions cited the creation of backup or security copies, two reported periodic recopying, and only one carried out tests for errors.

These findings suggest that archivists must be made more aware of problems in the preservation of MRR. Ben DeWhitt, in his survey of state archives, noted that much could be done to preserve MRR without actually having a computer or extensive training.<sup>7</sup> Yet, a minimal understanding of the preservation needs of these relatively fragile records is important and must be further emphasized in our training efforts. Generating increased use, however, is more problematic. In part, this requires making researchers more aware of the possible uses of machine-readable college and university records. Of course, this may not prove to be easy for, as Vinovskis points out, most researchers in higher education performing quantitative analysis rely on cross-institutional survey data.<sup>8</sup> Increased use of MRR also requires changes in attitude on the part of administrators: they must be convinced of the validity of permitting secondary research use of their machine-readable records.

Furthermore, archivists will have to take more initiative in seeking out machine-readable data files that have potential research value. Only about half of the institutions had inventories of their

MRR and almost 90 percent did not know the quantity of these records in their institutions. About half the institutions did not include MRR in retention and disposition schedules and three-quarters had no established appraisal standards for these materials. Fewer than half of the repositories without scheduled MRR or appraisal standards intended to develop them. It is interesting to note that three institutions had acquired computerized data files created by faculty members and fourteen more planned to do so. Two archives had acquired MRR from other organizations and fifteen others intended to acquire MRR from local or state government, businesses, or other organizations. These findings suggest that, while extensive training in data processing may not be needed to manage machine-readable archives, more familiarity with quantitative research techniques probably will be needed. Archivists must seek closer relations with social scientists who study higher education as well as with data managers who oversee the records in their active life.

Lack of familiarity with computers in general is probably at the root of the most frequently cited problems in this area—getting started and personnel. Funding and administrative relationships were also cited as problems, but they were listed less frequently. Much of the concern over funding is caused by a lack of awareness of actual costs in establishing a program. This problem is exacerbated by the lack of communication with the data processing community. The barriers are not caused by jargon alone; they also lie in the tendency of data managers and data archivists in computation facilities to regard their environments as self-contained and self-sufficient. After all, they “archive” data, do they not? We

<sup>7</sup>DeWhitt, p. 156.

<sup>8</sup>Vinovskis, p. 34.



need to explain to data professionals both the concept of life cycle of records and the archivist's concern for information of permanent value, regardless of medium. An alliance must be built in the spirit of mutual cooperation.

An even greater need for increasing awareness and communication lies in the area of automated management support systems. With only thirteen institutions reporting the ability to access or have responsibility for these systems, we have a long way to go. Of those responding to the in-depth questions, three repositories could access these databases and two could add data. In most cases, only administrative staff could use this information, and the means of doing so through software were limited. These archives had no apparent plans for expansion.

In both of these areas, the lack of involvement of archivists in systems planning is significant. Only three repositories reported that archivists were involved in planning in terms of records keeping, and only one had an archivist serving on an automation planning committee. If MRR are truly the archives of the future, and if there is any validity in the promise of the paperless office, we as a profession have been sadly lax. These records have a limited lifespan because of their storage media, and the urge to erase data once it has served its primary purpose is strong among data processing staffs. Our survey reveals that many more archivists than we might have suspected have begun to act on automated records, but more education and assistance is needed to improve the performance of these repositories as well as to initiate new programs at other institutions.

### **Automated Control in Archives**

In this segment of the survey, archives were asked to provide information on their use of automated systems that permit administrative and/or intellectual

control over a database of information about holdings. Secondly, repositories were asked to specifically comment on the use of automation in the preparation of finding aids. The results indicate that these archives began activity in this area in the early 1970s, although three-quarters of the automated control systems were brought up in the last six years.

The character of these systems is quite varied. Seven systems were developed by archival staff, eight were developed by campus resource people (computer center or library systems staff), and nine were external systems. Typically, the data input represented two-thirds of the holdings (number of collections and amount of material), with more than a quarter of the repositories having 100 percent of their holdings accounted for in those systems. The hierarchical level of description represented is interesting. Five repositories have both record group/collection descriptions and series descriptions in their systems, four each have one or the other, and another six have either folder-level or document-level control only.

Use is also of considerable interest. Of the sixteen systems that permitted retrieval by subject for users or the preparation of special lists of materials, eleven also made provision for administrative control activities such as maintenance of location or accession registers and disposition schedules, the recording of use, and the generation of statistics. Only two repositories reported no intellectual control activities. Only one-third of the systems provided users with direct access. Surprisingly, only one of these systems is based on a microcomputer.

About 40 percent of the archives surveyed in these two areas are preparing finding aids using computers independent of an automated control system. Of the 40 percent, about a third use stand-alone

word-processing equipment and the remainder use a text-processing program from the college or university's mainframe computer. This activity is of fairly recent origin; on the average, archives started this procedure four years ago. Almost all of this activity results in finding aids produced in page form, with only two institutions producing Computer-Output Microforms (COM). Only about a third of the collections have been described in this fashion in the average repository.

The interest generated by the new MARC AMC format has somewhat eclipsed the longstanding fascination of the profession with these independent systems. For a number of years, archivists bemoaned the indifference and inappropriateness of library automation to archival concerns. A great deal of experimentation with individual systems has occurred throughout the profession and, it would appear from the survey, at college and university archives in particular. The relative availability of large mainframe computers and research applications staffs at computer centers has fostered much of this development at low cost. When asked to identify the problems in these areas of the survey, archivists most often cited getting started. Funding and personnel were considered to be less difficult, and attitudes were judged to be hardly a problem at all.

The future plans reported by these archives were almost uniformly to expand their efforts for more sophisticated access, indexes, and control tools and also to increase the percentage of collection coverage for these systems. Archives producing finding aids outside of an automated control system intended to either move on to such a system or to integrate data into an on-line library system.

Trends in archival automation that cannot be clarified from these results are the impact of the microcomputer explo-

sion and the implementation of the MARC AMC format. Even as recently as a year ago, there was considerable tension on many campuses over the conflict between proliferating microcomputers and declining resources for central computing. The rapid increase in purchases of computers, concerns for computer literacy, and increasing enrollments in computer science classes, however, have considerably reduced this tension. It is too soon to assess the impact of these developments for archives. But the installation of more microcomputer-based systems, either using independent database management software or in communication with campus mainframe computers for large utility applications such as sorting, seems likely.

### Library Systems Activity

These sections elicited responses from three-quarters of the participating college and university archives. Respondents were asked to comment on the entry of information on their holdings into library or archival network databases and the extent to which they are being considered in local automated library systems. Almost half of the responding U.S. archives were involved in both aspects of this area, while only 15 percent of the Canadian respondents have entered records into network databases.

Typically, network database activity began six years ago and 85 percent of the repositories are entering information into Online College Library Center (OCLC). The average number of records created is 182.3, representing almost a quarter of the collections, with only five repositories entering more than half of their materials. While most archives were interested in the production of catalog cards, two-thirds expected eventually to share their data for reference purposes, and several discussed using their automated bibliographic records for developing their own



on-line systems in the future.

In more than half of the archives, this activity was initiated by library administrators or catalog department heads. Terminal location for input in three-quarters of the institutions was in the library's technical processing area. About half of the respondents indicated archives staff prepared data input sheets and, in a third of the cases, also performed terminal input.

Turning to consideration in library automation planning, we found almost 90 percent of the archives to be administratively part of the library. Not surprisingly, two-thirds of these repositories now have, or are planning to have, archival holdings information represented in their library's database. About one-fifth anticipate, or are using, an automated circulation system to control the occasional lending of material. Two institutions mention microcomputer applications for finding aids and inventory control for the future. The initiator of the planning was as likely to be the library director as the archivist; and, in most cases, planning was a joint effort between archival staff and library technical operations and systems staff.

These sections of the survey received the most extensive comments. Those relating to entry of data into network databases almost uniformly suggested that coverage be expanded to input more of the collections. Involvement in library planning received, by and large, favorable comment. Archivists in general felt that integrating the basic finding aids for their holdings into the rest of the library's catalog data would be important for public service, collection development and management, and efficient technical processing in a research library of any quality.

There were concerns, however. Other archivists mentioned the sacrifices made in the attention to cataloging details and

perfection of these bibliographic records. Others were concerned with the lack of priority attached to archival conversion by library staff, were pessimistic about the benefits of the effort, or questioned whether an acceptable cost-benefit ratio could really be achieved in this effort. Several mentioned dissatisfaction with OCLC and its lack of real utility to archives. One respondent did not "believe, at all, in the principle of adding every scrap to OCLC," and one questioned the place of archival and manuscript materials in automated library systems in general.

As mentioned above, perhaps the most significant development in this area is the new MARC AMC format. It has been implemented in the RLIN system of the Research Libraries Group, Inc., and several member institutions are adding data on their holdings. Other format implementations include a project at the Chicago Historical Society and one to bring up the format for Penn State's LIAS integrated library system. It also appears that OCLC, the other major bibliographic utility besides RLIN, will implement the new format. These developments will, for the first time, put archival information securely in the library automation environment. It seems unlikely, however, that they will replace the automated control systems described previously. If anything, archives will continue to rely on parallel systems for internal controls and access at the folder level for most record series and collections, while the MARC records in library databases serve to bring in new users and provide researchers with upgraded access to interinstitutional data.

### **What Have We Learned?**

The survey of automation in North American college and university archives provided a surprising picture of variety and breadth with perhaps a longer history

of automation development than we might have expected. It would be too easy to claim that we have entered a new era in archival automation but, for many archivists, this is a point of decision and redefinition. The microcomputer revolution and the birth of the MARC AMC format certainly open new possibilities for automated activities. Similarly the continued growth of MRR and MIS in colleges and universities is presenting new challenges to the archivist. The relationship of these archives to their library hosts is changing as well and will put new pressures on college and university archives.

A major aspect of professional concern for the Society of American Archivists and its Task Force on Automated Records and Techniques has been education and training. Findings in this survey demonstrate a rather low level of computer training on the part of these archivists. Of the three areas of the survey, staff with computer training were most likely to be found in archives with automated control systems; but even there few claimed formal coursework. It has been frequently asserted in the literature that computer training is not actually necessary—that archivists will not be programming their own system.<sup>9</sup> Nevertheless, the hands-on familiarity gained through a programming course, as well as workshop experiences geared toward specific applications, are invaluable. It is true that one need not be an accomplished programmer to tend MRR or to use a microcomputer. But one must be able to communicate with computer professionals to achieve cooperation for a machine-readable archives program or to understand the limitations of commercial software packages for archival adapta-

tion. Even those archivists who will be moving toward creation of MARC records for their data will enter a new world of tags and field definitions, and some grasp of how systems work will be needed. There can be no doubt that training and continuing education in automation must continue to be a major activity for the profession.

At the same time, however, concerns over availability of equipment and funding may become less pressing. Actual costs of computer equipment are declining, and the availability of computer resources on campus is expanding at an unprecedented rate. Similarly, the attention paid to archives by library automation will now rise with the advent of the AMC format, although archivists will have to move quickly to ensure that library administrators understand that “getting the archives into the database” is not the complete solution. Parallel internal control systems must be developed in microcomputer applications or by connection to the campus mainframe, or the needs must be met through modules of the library system—newly planned or modified—to handle archival information needs.

Overall, library automation may enable the archives to become more than a stepchild of the library. The trade-offs, requiring more rigorous description and acceptance of standardization, are certainly significant, but ultimately these should benefit both the archives and its users. Still the most troublesome area for college and university archives remains machine-readable records. The relationships with data processing professionals probably will be resolved. Appraisal remains the greater challenge. The profession must do more to study the generic forms of these records in institutions of

---

<sup>9</sup>For example, David Bearman, “Automated Access to Archival Information: Assessing Systems,” *American Archivist* 42 (April 1979): 181.

higher education and to determine their information content and their usefulness in research. One aspect of the library environment that seems to strongly influence college and university archives is a tendency to be driven more by informational values than evidentiary values. If this is true, it will be a particularly appropriate characteristic to retain in the automated records area as we continue to develop our partnership with automation.