January 1980

Archives, Automation, and National Networking: Is There a Future?

Karen Benedict
Nationwide Insurance Companies

Follow this and additional works at: http://digitalcommons.kennesaw.edu/georgia_archive

Part of the Archival Science Commons

Recommended Citation
Available at: http://digitalcommons.kennesaw.edu/georgia_archive/vol8/iss2/2

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Georgia Archive by an authorized administrator of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
ARCHIVES, AUTOMATION AND NATIONAL NETWORKING:
IS THERE A FUTURE?

Karen Benedict

In the July 1976 issue of American Archivist, Michael E. Carroll discussed the UNESCO Intergovernmental Conference on the Planning of National Documentation, Library, and Archives Infrastructures held in Paris in 1974. UNESCO proposed the creation of a National Information System (NATIS) in the United States designed to provide users with access to all of the relevant bibliographic information on a given subject through documentation, library, and archives services. NATIS would meet international descriptive bibliographic standards and would be compatible with an international system similar to, but broader in scope than, the current World Science Information System (UNISIST).

The concept of an international network of all types of information services on a broad range of subjects is exciting, but is as far from fruition now as it was when UNESCO made its proposal for NATIS. The prospects for an international group of librarians, archivists, and information specialists reaching agreement on a set of descriptive bibliographic standards for all printed matter, nonprint media, manuscripts, and archival records; a standard format for recording that bibliographic data; and a universal system of subject classification for retrieving that information do not appear good. Within the United States alone, librarians and archivists cannot agree upon standards for the description of manuscripts and archival records, and archivists cannot even agree among themselves on standards and formats for description of manuscripts and records.
The library profession has had success in establishing national and international bibliographic standards for cataloging of print and nonprint media. In 1908 the library associations of Great Britain and the United States established the Anglo-American Code (also known as the Joint Code) in an effort to create an accepted cataloging standard throughout the English-speaking world. Through the years librarians continued to revise and amend the code to improve its usefulness and to adapt to the proliferation and dynamic nature of information generated in a high technology society. The end product of this effort was the 1967 Anglo-American Cataloging Rules (AACR) and the 1979 Anglo-American Cataloging Rules II (AACR II). The International Federation of Library Associations (IFLA) through its International Office for Universal Bibliographic Control has established International Standards for Bibliographic Description (ISBD) for monographs and serials.

In general archivists believe that the type of cataloging which librarians practice, with its subject matter orientation, is inappropriate for archival records, and archivists have rejected the descriptive standards for cataloging manuscript materials developed in AACR and AACR II. Nor have archivists created their own code for bibliographic description accepted by the entire profession, in spite of early efforts like Margaret Cross Norton's 1938 Catalog Rules: Series for Archival Material. Without any established standards for description of archival and manuscript collections, each institution has had carte blanche to go its own way and to devise its own descriptive information for collections. Until the archival profession sets standards for description, or cooperates with the library profession's efforts to do so, very little can be done to create a national information network.

Just as librarians are ahead of archivists in standardization, so have they had more success with cooperative and computerized networking ventures. Because most of their materials are duplicated elsewhere,
libraries readily discern financial benefits from cooperation. Networking allows libraries to distribute the cost of equipment, data bases, and technical computer staff among several institutions; and standardized bibliographic description has facilitated computerization of a number of operations, especially cataloging, interlibrary loan transactions, and acquisitions.4

Holdings in archives and manuscript repositories, on the other hand, consist primarily of unique items. Therefore most archivists do not see the same sort of financial gains accruing to their institutions from cooperation, thus eliminating the main incentive for cooperation and networking. Nevertheless there are good reasons for archives and manuscript repositories to cooperate and to form networks. Knowledge of the holdings of other institutions can prevent duplication of effort and unnecessary competition for collections in a subject or geographical area. Networking would also enable institutions to direct prospective donors to the appropriate repository for their materials. The greatest benefit of networking, however, would be in reference services and the major advantage would be for the user.

At present researchers must depend upon the Hamer Guide, the National Union Catalog of Manuscript Collections,5 and the work of fellow scholars to locate manuscript and archival collections for their work. None of these sources is exhaustive. The profession has not yet been able to marshal sufficient cooperation among institutions to create a comprehensive guide to institutional holdings on a national level. A combination of the lack of national standards for description of holdings, the absence of substantial financial incentive, and the lack of commitment to provide better reference service has kept archives and manuscript repositories from making meaningful efforts to cooperate and to create networks.

The last of these obstacles may be the crucial one. The archival profession has placed far more
emphasis upon administrative concerns than upon the need to provide information services. As James M. O'Toole pointed out in a 1975 address to the American Society for Information Science, "Archivists and manuscript curators . . . persist in handling similar problems in vastly different ways and in the fussy habit of holding crucial information in their own heads and confiding it to no one." For computer technology and the attendant possibilities for the creation of national finding aids and institutional networks to receive the support necessary to achieve significant results, a major shift in the focus of the profession to an emphasis upon the information function will be required.

The tendency to stress administrative control at the expense of greater intellectual control of collections to the detriment of the researcher is rooted in the history of the archival profession in the United States. Men like Theodore Schellenberg and Ernst Posner adopted the cardinal principles of provenance and original order from European archival practice, while developing the American practice of arranging records to follow the organization and function of the agencies which created them. They believed that arrangement should reflect the process by which the records came into existence.

Schellenberg, Solon Buck, and others devised the term "record group" to define the main unit of arrangement for the records of administrative units at the bureau level of government. These founders of the profession established that record groups be arranged in either organization arrangement, reflecting the hierarchical structure of the organization, or in functional arrangement, reflecting the interrelationship of function of several agencies and offices. The organization of record subgroups was based either on the organization or the function of sections within the administrative unit or upon the physical characteristics of the records themselves. Series within subgroups reflected the particular filing system of the
administrative unit. The series were made up of individual file units—volumes, folders, or individual documents—arranged sequentially as they were established by the creating body, based on their relationship to the organization, function, chronological period, place, or subject. Thus the arrangement of archival records as established by the National Archives was a well thought-out system based upon scholarship and the European experience.

The guidelines which were created for the description of the records, however, were designed for the purely practical task of maintaining control at the National Archives. The device used for description of archival records was the inventory, an initial brief list of record units. Katherine E. Brand of the Library of Congress designed a similar tool, the register, as the basic finding aid for manuscript collections. Neither the inventory nor the register describes the piece-by-piece contents or arrangement of a record group or collection. The register indicates the size, inclusive dates, and basic scope and content of a collection. The inventory contains the same sort of brief information for the record group, its subgroups, and series. Inventory description at the National Archives rarely, if ever, goes beyond the series level.

The decision not to implement description beyond the series level was pragmatic, the result of insufficient funds and staff to support the work. It did not reflect any reasoned conclusion that item-level description was inappropriate or unnecessary for archival records. Early archivists assumed that the inventory and register were preliminary tools to insure the institution's basic control over its holdings and that when staff and budget increased the collections would receive additional attention. However, time has shown that staff and budget never increase sufficiently to allow an institution to rehandle records that have received initial attention.
The unfortunate consequence is that, without any theoretical analysis of the ramifications of the failure to gain complete intellectual control over collections, item-level description has been eliminated as a viable practice for archives. Rare is the institution which has a staff large enough to prepare calendars and other detailed finding aids for its holdings. Moreover, the sad truth is that these sorts of detailed guides are looked upon by much of the profession as "unprofessional," the fussy work of little institutions run by dedicated ladies with time on their hands. What began as the accidental consequence of limited resources has been raised to a canon by the profession.11

Archivists must make a more reasoned decision about the level of description which all institutions should set as the standard practice. Archivists must also agree upon a uniform format for collection description before it will be possible to create a regional or national computer network. That format should cover the type of information which must be provided for each collection or record group, the measurements to be applied to them, the amount of detailed description expected, and the order in which the information is to be recorded.

In spite of the great obstacle of not having uniform standards for description, archivists have made some progress in creating networks and sharing information. The Library of Congress has taken the lead by launching projects like the National Union Catalog of Manuscript Collections (NUCMC) and Selective Permuted Indexing (SPINDEX) which have encouraged cooperation and have utilized computer technology.

NUCMC provided researchers with the most complete national guide to the holdings of manuscript repositories and set the first accepted interinstitutional standard for collection description. Because most archivists support the concept of a union list of manuscript collections, institutions have cooperated fairly well in providing the required information to the

http://digitalcommons.kennesaw.edu/georgia_archive/vol8/iss2/2
Library of Congress; and to date approximately 39,770 collections have been reported. And because NUCMC requires repositories to collect and report certain information about each collection in a particular format, the cooperating institutions have tended to include the same information in the same format in their own institutional guides to collections.

However, NUCMC excludes archival collections which are maintained by their creating agencies. This seriously limits the ability of NUCMC to serve as a stimulus to full interinstitutional cooperation and as a source of information for a national network of collection information for archives and manuscript repositories. Not only is important information about archival holdings not available to researchers and other repositories, but the excluded archival institutions have not accepted the NUCMC format for description of their collections.

The Library of Congress developed SPINDEX in response to the overwhelming task of creating a date, author, and recipient index for the hundreds of thousands of items in the presidential papers microfilm project. This index project initially employed a punch card system of automated data processing to sort information, but in 1964 Library staff transferred the data to computer to complete the indexes. A decision to employ the computer for description of the Manuscript Division's 3,000 collections followed the success of this automated indexing venture.

The index produced for the presidential papers did not provide subject control; therefore the Library decided to create a system which would produce a modified "Key Word in Context" (KWIC) index based on subjects and names gleaned from the container lists which had been produced for the collections. This SPINDEX system employed a fixed-field format using the standard eighty character computer card. Testing proved that the fixed-field format did not provide adequate space for collection description, and in 1966 the Library of
Congress suspended the SPINDEX project.

The National Archives revived and revised the SPINDEX project in 1967 with the help of a $40,000 two-year grant from the Council on Library Resources. The new system, christened SPINDEX II, changed from a card to a tape format to allow for variable-length fields and utilized lower case as well as upper case type for the first time. The system now had the capacity for on-line correction and updating. Nine other repositories joined the National Archives and Records Service (NARS) to test the potential of the system for providing interinstitutional description for archival collections. Most of the allotted project time was spent attempting to produce a standard format which would be acceptable to all of the participating institutions and easily implemented by them. The testing of the proposed indexing system bogged down, and the grant expired before SPINDEX II could be implemented.

The National Archives then assumed full responsibility for the SPINDEX II project. Several of the original participating institutions dropped out of the project and others joined it. In June 1973 the National Archives held a conference of original and subsequent SPINDEX users to evaluate the system. At that conference NARS indicated that, although the system had been successfully used to index the papers of the Continental Congress and the guides to the captured German documents and other institution projects, the production of such detailed indexes to the Archives was not feasible, evidently for financial reasons. When other conference participants expressed concern that NARS abandonment of SPINDEX would endanger the concept of a national data bank, the Archives promised to make SPINDEX II available at a reasonable cost as it developed and to serve as the clearinghouse for information on the system. The Archives refused, however, to commit additional money to the development of an information retrieval system which would be used principally, perhaps exclusively, by other institutions.12
The SPINDEX II experience soured the National Archives on the prospects for indexing its holdings by subject. In his 1973 article "Automation and Archives," Frank Evans argued that it was futile for the Archives to attempt to analyze its entire holdings by item, folder, or series. Therefore the Archives would abandon the notion of information retrieval based on subject indexing in favor of a system of administrative control at the record group level.

This was a severe blow to the prospects for inter-institutional cooperation. Even though it is quite clear that its sole responsibility is its own administrative problems, its size and prominence make the National Archives the leader in the archival field. When the National Archives abandons the development of information retrieval systems with subject indexing capacity, it makes a de facto decision for the rest of the profession.

It was evident from the proceedings of the 1973 SPINDEX users' conference that some smaller institutions were less concerned than the National Archives with administrative control of holdings and more committed to the establishment of a national archival network. Therefore a number of archival institutions have adopted SPINDEX II in spite of its shortcomings. This has not, however, increased the viability of its adoption as a national network program, because individual institutions have had to modify the program to suit their particular needs. The South Carolina Department of Archives and History, for example, has modified the program so that it can supply a personal name index, a chronology, a place name or locality file, a main topic or subject list, and a list of documents by type. The modifications which have been made in the system vary from institution to institution and may inhibit the ability to interface programs.

While various institutions experiment with SPINDEX II, work is going forward on new automated systems for archival use. SPINDEX III, developed by Frank Burke,
creator of SPINDEX II, is designed to meet the needs of the National Historic Publications and Records Commission in the production of its Directory of Archives and Manuscript Repositories and subsequent projects. The archives of the University of Illinois in Urbana has created the Programmed Annual Report and Digital Information Matrix (PARADIGM) system for administrative control of its holdings at the collection level. Like the National Archives, the University of Illinois archives has rejected subject indexing.

After conducting a $70,000 study, the National Archives has developed the A-1 system to meet its requirement for administrative control of records. NARS selected A-1, a computer-assisted system for text editing, rather than a system designed to retrieve information by subject because the latter necessitated the development of a thesaurus. "A dictionary of terms would have to be developed and applied systematically to all series description . . . ," the Archives' Alan Calmes explained after the decision was made, and "indexing would require that an archivist identify appropriate index terms for each series description. This would slow down the decision-making process during series description writing." The analysts recommended that subject retrieval receive serious attention only after the problems of administrative control were solved. Thus the National Archives administration does not appear to have revised its thinking over the years.

This is the state of automation in the archival profession today. In spite of the quality and quantity of effort that has gone into research and experimentation in the automation field, archivists are as far away from readiness to participate in a national information network as they were in 1976. As a profession archivists have learned the lesson that experimentation with computer technology is a costly business, and that if we deal only with tangible, dollars and cents, benefits it may be more expensive than the results warrant. What we have not done is to analyze realistically what the profession wishes to achieve through automation.
Are we seeking faster and cheaper methods to achieve administrative control over holdings? Are we looking for a reasonable means to provide greater depth of intellectual control over holdings? Do we want to provide users with more information about institutional holdings? Do we want to provide subject access to collections? Do we need more information about the contents of collections to achieve these goals?

As archivists we must clearly define our objectives before we can accurately assess whether automation will deliver sufficient benefits to warrant the expense involved. Once we have established our professional priorities, whether they be administration or a commitment to information and reference services, then we will be in a better position to determine whether we wish to join with other information service professions in a cooperative effort to create national access to information on a scale never before possible.

NOTES

1 Mr. Carroll is chief of the Machine Readable Archives Division of the Archives of Canada. He is also a member of the Society of American Archivists' Committee on Data Archives and Machine Readable Records and the International Council on Archives' Committee on Automation.

1975), pp. 303-56.

A more thorough discussion of the development of archival arrangement and description practices and the influence of library techniques on their development can be found in Richard C. Berner, "Arrangement and Description: Some Historical Observations," American Archivist 41 (April 1978): 169-81. Archivists reject AACR and AACR II standards for cataloging manuscript materials because they are based too closely upon those established for published materials and do not allow sufficient flexibility to deal with the uniqueness and variety of manuscript materials.


This definition of record groups, subgroups, and series is taken from T. R. Schellenberg, Modern Archives: Principles and Techniques (Chicago: University of Chicago Press, Midway reprint, 1975), pp. 181-86. Even these basic concepts of filing units do not have a single standard definition within the profession. See, for example, the definitions in Frank B. Evans et al., "A Basic Glossary for Archivists, Manuscript Curators, and Records Managers," American Archivist 37 (July 1974): 415-33; and the definitions in the glossary of Kenneth W. Duckett, Modern Manuscripts: A Practical Manual for Their
Management, Care, and Use (Nashville: American Association for State and Local History, 1975) which are based on Evans's work.

8For a more complete discussion of the development of the American archival system and its principles and practices of classification and description, see Schellenberg, Modern Archives, and Berner, "Arrangement and Description."


10Schellenberg, Modern Archives, pp. 208-10.


14Cornell University was critical of SPINDEX II's failure to provide a subject-authority for the system. The Minnesota Historical Society was disturbed that the lack of a thesaurus for the system inhibited the
growth of computer usage by the profession. See Hickerson's discussion of SPINDEX in SPINDEX II at Cornell University.

15Duckett, Modern Manuscripts, pp. 157-58. The lists provide citations by record group, series, box or volume, folder or page, and item number.


17For more information on PARADIGM, see Hickerson, SPINDEX II at Cornell University.