
An Archivist's Challenges: Adapting to Changing Technology and Management Techniques

by Donald Fisher Harrison¹
National Archives and Records Administration
Washington, D.C., United States of America

Introduction

Over twenty years ago, the National Archives of the United States embraced the concept that automated records were actually records which could be considered permanent within the meaning of the Federal Records Act and set about collecting them. Since then it has confronted problems incident to finding these automated records, acquiring them, preserving them and making them available to the public. Previous papers have discussed access to public automated records in the normal sense; that is, the ability of the researcher to get at them. In this paper I wish, however, to discuss the National Archives' acquisition process as a form of access.

This paper addresses three threats to the acquisition of machine-readable records: the threat of an onslaught of hardware and software incompatibility, the threat of discontinuity within textual records series brought about by end-users with microcomputers and the threat brought about by new management techniques from the Paperwork Reduction Act of 1980. Archivists ought to view these threats as challenges. When overcome, the challenges will have presented the Archives with the opportunity to create a better collection of automated records.

Software and hardware dependency

The first challenge to the National Archives is well publicized and needs no significant introduction in this treatise. The Archivist of the United States, confronted with the research community's complaint that valuable data were being created by Federal agencies without any consideration for their preservation or dissemination to the public, established in the

¹Presented at IASSIST/IFDO International Conference May 1985, Amsterdam.

** The views expressed in this paper do not necessarily correspond with those of the National Archives and Records Administration. I wish to acknowledge that some of the material has been developed out of long standing collaboration with fellow archivist Dr. W. Jon Hedgesheimer, but any mistakes in concept or fact are mine.

1960's the forerunner of today's Machine-readable Branch. This branch was given the task of inventorying Federal data bases and deciding how best they should be preserved for posterity. We accessioned a number of machine-readable data files created in the 1960's. Some of these files were dependent on other outside factors and could not be read on their own. Three examples of software and hardware dependency illustrate our initial problems.

The first example came early in our organizational being. We received over thirty-five machine-readable data systems from the Office of the Secretary of Defense and the Office of the Joint Chiefs of Staff. These systems were encoded in a data base management system called the National Information Processing System (NIPS). They caused serious problems in access and handling and a considerable backlog in the accessioning workload.

NIPS was devised for generalized file handling using languages designed to support user requirements in six components. It afforded any data center the capability of reporting long and involved statistical manipulations on extremely short notice to a variety of users. However, the software was compatible only with IBM computers.

The presence of NIPS files suggested serious difficulties in providing a uniform reference service to researchers and brought up the whole question of software dependent files. To retain the files in NIPS would constitute a precedent. Since researchers by and large preferred to use their own utility software, transportable files would afford a range of options that encoded files would not. Last but not least, maintaining large inventories of software would add to the preservation costs and require more shelf space. For all these reasons we decided to decode the files. It appears now, with hindsight, that, despite the fact that these files were unique and

very valuable, we should have insisted that the material be transportable before being accepted by the National Archives.

The second example was the National Archives' accessioning of a microfilm series of records containing pictures of captured North Vietnamese documents. These were filmed in Saigon during the war by the Combined Document Exploitation Center on 94 oversize (13-inch) rolls of 35mm microfilm, each roll 1000 feet long. The documents were on one side of each frame, with digital bar codes on the other side to provide indexing and control information.

Soon after we received the microfilm we discovered to our chagrin that the material was hardware dependent in a system known as "File Search." Four configurations of this machine had been manufactured and sold to Federal agencies in the 1960's. The last model (generation four) had a small computer in it. It could therefore provide a printout by reading the bar code on the film strip, transferring it to magnetic tape, which in turn could be manipulated and dumped on to paper. The machines cost \$250,000 new and were used only by military and intelligence agencies, as far as is known.

It was only after this information was made available to us that we discovered that other file systems were known to exist in this environment and were equally unreadable without any machines in existence to retrieve data. These included some important files in the Navy Sea Systems Command (in Arlington, Virginia) and the Navy Oceanographic Command (in Bay St. Louis, Mississippi), including the Defense Intelligence Agency. Recently we have discovered the existence of an intact File Search model in salvage channels. We have requested that it be turned over to the National Archives, and we think we have the technical expertise to restore the model to operating condition.

The third example entailed the 1960 Decennial Census, offered to the National Archives by the Census Bureau in the mid 1970's. These records were created by a UNIVAC II-A computer, of a generation that had been effectively phased out of use in the Federal government after the tapes had been created. It has been reported that once the tapes became available for transfer to the National Archives, only two such machines existed to read them, one in Japan and one in the Smithsonian Institution.² Eventually a reasonable approach was agreed to by the National Archives and the Census Bureau, to convert the data into a compatible format, making them available for preservation in our vaults.

These three examples are illustrative of the long term problems created by hardware and software dependency of records created in the 1960's, when computers were maintained in relative isolation from each other. It was a period in which data managers were concerned with the creation and the use of computer products and were by and large ignorant of the long term value of these products as Federal records. It can be said that the letter of the law — the fact that the tapes were handed over to the National Archives — was carried out. The fact that the tapes were unreadable because of software and hardware dependency was a new problem that had never been faced with paper acquisitions. For their part, agencies were understandably reluctant to disperse funds solely for the benefit of depositing these records in the National Archives. Thus reason has had to prevail in our dealings on transfer of the tapes, and no one solution can be applied in all cases.

Small computers and office automation

The second challenge to the smooth flow of records into the archives stems ironically from the very machine meant to facilitate administrative operations in the modern office. For several years, most federal agencies have been extending the advantages of their word processing pools by placing terminals at the hands of management officials, giving fingertip control to their own records creation. Office automation (AO), more aptly termed "the paperless office", is based on a series of compatible, menu-driven programs utilizing a centralized data base for common shared-use data and unique smaller databases for individual users. These systems have the ability to transfer data and information between data bases through a network or a distributed system.

The advantages of such a system are obvious. Federal managers frequently need information suddenly and immediately, and often the demands for this information come after the staff has left for the day or the weekend. Managers would like the ability to search for the data or reports they need through an indexed automated bibliographic/numeric data base, access and use the appropriate software to perform simple to moderately complex analyses of this data (e.g. forecasts, correlations, etc.), use graphics to illustrate their results, access word processing/office automation tools to produce a memo in the appropriate format, and finally send this report/memo electronically to the recipient's office, all without the necessity of using the phone, typewriter, or staff that are not available.

Keeping all this in the system can cause an archival "log jam". The designers and the users of paperless office systems are frequently ignorant of the paper systems they are replacing and the archival need for intellectual continuity. Outside contractors compound the problem. In

²Committee on the Records of Government, *Report*, Washington, DC, March 1985, pp. 86-87.

the absence of any other information the hardware and software dependency problem has reemerged in the small computer world, and agencies are finding that transportability cannot cross the boundaries between offices. Software now provides end users with ultimate fingertip access. This allows handcrafted programming and instant manipulative gratification. The same person who creates data on the system can now dispose of it with equal ease. By closing the gap between the user and the machine, the system eliminates the apparent need for the data middleman, to say nothing of the records manager who, under other circumstances, looked after standardized formats, ensured traditional records disposition practices and provided for a continuity of records series in the agency.

Thus the danger inherent in OA is that the practice concentrates on the information as it is used immediately after creation without making a record of actions taken. It is said to parallel the dangers of telephone use when first introduced. With that instrument, managers needed go through no intermediate device for communication. Telephones assured privacy of communication and were sheltered from the public record. The comparison with OA is evident. Just as managers could converse at the push of a telephone button, so they do now with electronic mail. Further, if one of the parties is absent, there need be no callback, because the mail has already been delivered electronically.

Like the telephone, the OA challenge is to find a way to record the communication. With the small computer, software must be devised to ask the user for a determination of the ultimate value of the information before it is ever keyed into the system. This software has been integrated into the planning for OA systems in most Federal agencies. Whether or not it solves the problem in practice remains to be seen.

Information resources management

The third challenge to a smooth transfer of records to the archives now comes in the form of an application of new management techniques to the creation and the use of information within the Federal establishment. This new methodology typically accommodates the reality that government must function with less personnel and with individuals of lesser skill and training by altering the way agency missions are carried out. The Paperwork Reduction Act of 1980 was rightly concerned with a problem that had existed for some time in that the Federal government was preoccupied with the physical problems associated with the large volume of paper records created. The authors of the bill reasoned that managers should have been concerned with how the information was being controlled and how it could be shared with the maximum number of sources. Thus the new law espoused intellectual control vis-a-vis physical control, regardless of the medium on which the information had been stored. In order to do this a number of organizational changes have taken place in Federal agencies, each a bit different from the next, in which an "information Czar" has been placed at the highest levels to control access and dissemination of all information, regardless of the medium. This new arrangement has now been entrenched for four years.

A typical arrangement has been established to combine the former functions of "automation, communications, office automation, records management, publications, audiovisual activities and other information activities, services and facilities." An information management plan is usually mandated beginning with a problem analysis, designing a model information system, constructing the "architecture" which produces a program and provides guidance for a budget request. Under this concept, every information system will have centralized management. The

"single manager" concept has been extended to encompass all information, defined as "... all processes by which the user may receive, display or project desired information... (including) voice, text, graphics, audiovisual, video teleconferences, micrographics, files, records management, optical discs and other forms of published information."³

In many ways, the single manager system makes a lot of sense. The information manager is in a unique position to disseminate information within an agency to avoid duplication of effort — or better, to avoid disparate and conflicting data creation. By being organizationally placed at the highest level, the IRM provides information for important decisions and controls a sizeable portion of the agency's budget.

Furthermore, the concept will ease the path of liaison between the agencies and the National Archives. As we began to accession records in machine-readable form in the 1960's, we became increasingly aware of the presence of the data manager as a viable records creator and manager. Between 1961 and 1980, the Machine-readable Branch frequently communicated with the data manager directly when it was not able to get required information any other way. Furthermore, in the first half of this decade, we became more and more concerned with dealing directly with government managers since they were creating (and destroying) information without acknowledging either the Federal Records Act or their agency records administrator. With the advent of the IRM principle, however, the Archives need only deal again with one official, who, if properly briefed on the urgency of the problem, will coordinate the actions of the records manager, the data manager and the end user.

Conclusions

Technology has created new solutions to old problems, but in the process has itself created new problems. The archival community is thus confronted with unique challenges to its traditional role as keeper of the records, which requires our attention. Some measures come to mind as actions to stem the tide.

First, the archivist must keep professional pace with the proliferation of computing technology, not only as it is practiced in Federal agencies in this decade, but also as many writers envision that it might be practiced 25 years from now. Reading the literature is not enough. It requires a shrewd selection of educational services and an on-going dialogue with other archivists. This must include at a minimum the study of software, hardware and storage media as trends develop. An archives must be capable not only of receiving machine-readable records in various modes and written on various media, but also of serving its users with a multiplicity of arrangements.

This leads to the second measure. The archivist must determine far enough ahead in time in what mode and on which medium these new records will appear as candidates for acquisition. To do this, archivists must assert their professional needs to the creators of records throughout the life-cycle of the records. Furthermore, the requirement to deposit tapes and other media in the National Archives should be anticipated and budgeted by Federal agencies.

Third, the archivist must reach end users by some means, to ensure standardization of practices and procedures. It is vitally important to overlay records management practices on the uses and outputs of small computers and of office automation systems. This might include communicating with procurement officers and

³Draft Army Regulations 25-1 and 25-5, 1984.

IRM officials to standardize hardware and software packages which would be interchangeable within and between Federal agencies.

The fourth, and by no means the least important, point is that the IRM developments in Federal agencies, formed as a result of the Paperwork Reduction Act of 1980, must be influenced by direct communication with archival officials.

IRM managers have been imbued with the immediate needs of the agency information program in mind — the here and now concept. There is always the danger that not enough planning will be conducted for the ultimate fate of records. By the way they maintain certain modes of information, IRM officials can influence the disposition, and in turn, the configuration of future holdings of the National Archives.■

