Proceedings of the 6th Annual
Federal Depository Library Conference

April 14-17, 1997

Planning for Preservation of Digital Information: an Archival Perspective

Kenneth Thibodeau, National Archives and Records Administration
Washington, DC

Archival Preservation of Records Is Different than the Preservation of Publications, Because of Differences in the Basic Nature of Publications and Records

Publications are instruments intended to communicate information to an unspecified audience. That is, while there may be a target audience for a publication, it is not normally addressed to specific individuals. Records, on the other hand, are instruments and by-products of intentional activities. Their primary purpose is not communication itself, but support of the activities in which the records are created, or other subsequent activities connected to the ones in which the records originated. The target of a record is specific. Rather than author and readers, in the case of a record we have author and addressee.

Records are used to communicate within a group where a lot of information is common; therefore, there are a lot of lacunas or ellipses in records. Commonly shared information does not have to be repeated in each item. Thus, some information which is essential to understanding a record is likely not to be in the record itself. Records can only be understood in the context of other records which precede and follow in the course of an activity.

Therefore, archivists are not concerned with the preservation of a record, but with the preservation of a collection of records. In records preservation, the archival bond is as important as the individual record. The archival bond is the specific relationship, established by the records creator, between a record and other records in a ‘fonds,’ which is the totality of records created by an identified records creator. If we break or obscure the archival bond, then the meaning of the record is altered. Subtract the archival bond and you cannot preserve records, only documents.

Archival Preservation of Electronic Records Is Not Like the Preservation of Paper Records

For hard copy records, preservation equals holding on to what you have. For example, we preserve paper records by deacidification of paper, repair of tears, or replacement of originals by copies that are "true to form."

For electronic records, holding on to what you have will eventually mean loss of records, because all of the technology on which the records originally depend eventually will become obsolete, making the records inaccessible. So, while a basic technique for preserving hard copy records is the use of permanent papers, for electronic records, a permanent medium is irrelevant.

All digital media will eventually become obsolete. With new media formats being introduced on a time scale of 18 months, that eventuality tends to be very soon in arriving. Even if we could preserve a digital medium for a long period of time, along with maintaining the required hardware and software in operating condition, it would not be desirable to do so. The ever increasing expense of maintaining obsolete technology would be compounded by the ever decreasing costs of digital storage. Maintaining obsolete technology would also cut off opportunities to benefit from continuing improvements in information technology. These improvements reduce costs for both preservation and access and, typically, enable a greater volume of information to be delivered in a smaller physical volume and more quickly.

Planning to Preserve Electronic Records

In its forthcoming Guide for Managing Electronic Records from an Archival Perspective, the Committee On Electronic Records of the International Council on Archives sets forth, as a principle in the framework for managing electronic records, "The archives must articulate preservation and access requirements to ensure that archival electronic records remain available, accessible, and understandable."[1]

Implicit in this principle is the view that:

> "Preservation and access to archival electronic records are interdependent;
>
> "Available records are physically intact, identified, and readable;
>
> "Accessible records can be selected within search strategies consonant with the way the creator organized the records, and presented in an historically authentic form; and
>
> "Understandable records are records which can be used as historical evidence. This requires identification of the provenance of the records, maintenance of the original order of the records, and the availability of related records and other contextual information."[2]

T**he purpose of preservation is access**, but in the archival realm access means not access to information, but access to authentic records. This creates a basic tension in the preservation of electronic records. Preserving authentic records ordinarily means preserving

them in their original form, without alteration. But, as stated, this is impossible for electronic records.

A **digital preservation program must be dynamic**, both to counteract obsolescence and to take advantage of improvements in information technology and decreasing costs of new technology. A dynamic preservation program is one that will select the best options for preserving the records. But this selection will be made from a set of possibilities that itself will change. The key to making the right decisions is a sound migration strategy. As stated in the 1996 publication, Preserving Digital Information. Report of the Task Force on Archiving Digital Materials, a task force sponsored by the Commission on Preservation and Access and the RLG,

> "Migration is a set of organized tasks designed to achieve the periodic transfer of digital materials from one hardware/software configuration to another, or from one generation of computer technology to a subsequent generation. The purpose of migration is to preserve the integrity of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology."[3]

Migrating electronic records across generations of technology will inevitably involve transforming the records. The greatest challenge in archival preservation of electronic records derives from these inevitable transformations. The greatest challenge is not overcoming media fragility, or rapid obsolescence, or taking advantage of continuing improvements in technology. It is to distinguish the essential attributes of electronic records, which must be preserved, from other attributes, which are merely artifacts of technology, and therefore can be discarded without loss of authenticity or alteration of meaning.[4] There are a few, simple cases where these 'throw-away' attributes are well known. They include the physical media on which electronic records are written and the specific way they are physically inscribed on those media. The National Archives is at the fourth generation of digital media used for preservation. It will change to a fifth medium within a few years.

But beyond this simple level, there has not been much done in identifying the essential characteristics of electronic records.

The ICA Guide, which I cited earlier, recognizes the problem:

> "Over time, it will be necessary to transform the records in order to migrate them from obsolete technology to current forms. Archival preservation requires that such transformations respect the authenticity of the records and that such changes enable the records to be retrieved and understood. Such transformation must be thoroughly documented."[5]

The problem is also recognized in a white paper which was issued on April 10 by the Consultative Committee for Space Data Systems, Reference Model for an Open Archival Information System.[6] While this report is a product of the space science community, it does reflect archival principles and concerns. An archivist from the Center for Electronic Records, Dr. Bruce Ambacher, is a member of the panel of technical experts which produced the white paper. The white paper provides a generic framework for addressing the issue of

preserving and providing access to digital records. But the problem itself remains to be addressed. It is essential to do so in order to develop sound and viable plans for preserving electronic records.

1. International Council on Archives. Committee on Electronic Records. Guide for Managing Electronic Records from an Archival Perspective. 1997 (In press). p. 28.

2. Ibid.

3. The Commission on Preservation and Access. Preserving Digital Information. Report of the Task Force on Archiving of Digital Information. Commissioned by the Commission on Preservation and Access and the Research Libraries Group, Inc. Washington, D.C. 1996. p. ii.

4. Kenneth Thibodeau. "Boundaries and Transformations: an Object-Oriented Strategy for Preservation of Electronic Records." Proceedings of the DLM Forum on Electronic Records, Cooperation Europe-Wide, Brussels, 18-20 December 1996. European Commission. Brussels. 1997 (In press).

5. International Council on Archives. 1997. p. 29.

6. Consultative Committee for Space Data Systems, Reference Model for an Open Archival Information System. Report Concerning Space Data Systems Standards. (CCSDS 650.0W-1) Washington, D.C. National Aeronautics and Space Administration, CCSDS Secretariat. 1997.