

**Proceedings of the 5th Annual
Federal Depository Library Conference
April 15 - 18, 1996**

**Preservation and Archival Issues for
Electronic Records: the Center for
Electronic Records of the National Archives
and Records Administration**

Fynnette Eaton, Theodore J. Hull National Archives and Records Administration
College Park, MD

Access to government information is essential in a democratic nation. It depends, to a large degree, on good record-keeping by government agencies, on the preservation of important records, and on public access to those records. The information in records has considerable value in documenting the vast scope and great diversity of Federal Government activities and functions. Federal records often are the best available sources of information for various kinds of research, including research in history, sociology, economics, and the natural sciences.

The Center for Electronic Records is the organization within the National Archives that appraises, accessions, preserves, and provides access to Federal records in a format designed for computer processing. The Center maintains electronic records with continuing value created by the Congress, the courts, the Executive Office of the President, numerous Presidential commissions, and over 100 bureaus, departments, and other components of executive branch agencies and their contractors. Originally, these records are created or received by agencies of the Federal Government. They may concern virtually any area or subject in which the government is involved. They may be from any type of computer application such as data processing, word processing, computer modeling, or geographic information systems.

Archival Services and Electronic Records Preservation

Identifying permanently valuable records for retention by the National Archives involves cooperation between the Archives and the various agencies of the Federal Government. Through the process of appraisal, the Center identifies and selects the electronic records it judges to have enduring value.

When appraising electronic records, archivists evaluate the contents of the records for their evidential, legal, and informational value and for their long-term research potential. Some of the factors in the decision include whether the information is in its original form, whether the records can be linked with other information, how the records relate to past and present research trends, which Federal agency created them and why, how the information was used, and what impact the records had on Federal programs and policy. Generally, only a small percentage of all electronic records are evaluated as having enduring value.

When the National Archives receives a transfer of electronic records, two copies of the original records are made onto 3480-class tape cartridge, the current preservation media. The Technical Services Branch of the Center for Electronic Records administers the preservation program for electronic records. Regulations require Federal agencies to submit their records to the National Archives in a hardware and software independent format, on either 9-track magnetic tape or 3480-class tape cartridge, encoded in ASCII or EBCDIC and blocked no higher than 30,000 bytes. Recently, the National Archives has announced the acceptance of data transferred on CD-ROM media, accepting only those records conforming to a select set of 'standard' formats. In addition, any transfer of electronic records to the National Archives must be accompanied by adequate and proper documentation about the contents of the records. Electronic records transferred to the National Archives on CD-ROM will be copied to archival media, namely 3480-class tape cartridge, and the original CD-ROM will be used for reference, when appropriate.

The preservation program of the Center for Electronic Records provides valuable services to both Federal agencies and the general public. The Technical Services Branch has implemented a pro-active program to ensure the long-term accessibility of the records, as the preservation media for electronic records is fragile and deteriorates relatively rapidly compared to other forms of archival records. The branch annually conducts an analysis of a statistical sample of the tapes and cartridges to ensure that the environmental conditions are not adversely affecting the media. The environmental conditions at the National Archives for the storage of electronic records conform to current standards. The preservation program also involves rewinding those records preserved on open reel magnetic tape at a constant tension at normal tape speed every time they are placed on a tape drive. Finally, copies of permanent records are made every ten years to new media, or more frequently when necessary, to prevent the physical loss of data or the technological obsolescence of the medium.

Recently, the Technical Services Branch developed and implemented a new system for the preservation of electronic records. Prior to moving to the new National Archives facility in College Park, Maryland, the Center conducted the majority of its preservation activity on mainframe computing facilities at another government agency. The system now operational at the College Park facility is the Archival Preservation System (APS). The APS allows for the preservation of electronic records received on a variety of media, such as CD-ROM, onto the current archival media. Alternatively, the APS will eventually allow for researchers to request copies of files not only on the two available media (9-track magnetic tape and 3480-class tape cartridge), but on a variety of output media. This development should enhance access to electronic records for those researchers who do not have access to mainframe computing resources.

The Center's Holdings

In general, the electronic records transferred to the National Archives are themselves evidence of evolution in the application of computer technology in support of program administration, planning, record-keeping, and research within agencies of the Federal Government. A few of the data files were originally created as early as World War II and reflect punch-card technology in use since the 1880s; an even smaller number contain information from the 19th century that has been converted to an electronic format. However, most of the electronic records in the Center's holdings have been created since the 1960s. Numbering well over 20,000 unique files of electronic records, measuring more than 325 gigabytes in volume, the scope of the Center's holdings is quite diverse--as diverse as the activities and interests of the Federal Government itself.

Among the types of holdings or subject areas represented in the Center's holdings are the following:

- a. Agricultural Data (including the Department of Agriculture's livestock and crop production estimates, the periodic Census of Agriculture, 1949-79, and household food consumption surveys).
- b. Attitudinal Data (including surveys about equal opportunity, crime, violence; surveys sponsored by the U.S. Information Agency (USIA), and the "American Soldier" surveys of soldiers during World War II).
- c. Demographic Data (including data from the Bureau of the Census, U.S. Department of Commerce).
- d. Economic and Financial Statistics (including income, labor, securities, tax, trade, and transportation statistics).
- e. Education Data (including data illustrating the variety of education programs of the U.S. Federal Government).
- f. Environmental Data (including records created on behalf of the Presidential Commissions on the Accident at Three Mile Island and the Coal Industry, the Nuclear Regulatory Commission's Radiation Exposure Information Reporting System, and the National Oceanic and Atmospheric Administration, National Ocean Survey's Nautical Chart Data Base).
- g. Health and Social Services Data (including data incorporating both biomedical and sociological information and efforts to measure the effectiveness of a variety of social programs).
- h. International Data (including import-export statistics and USIA-sponsored surveys).
- i. Military Data (including "American Soldier" surveys, Prisoners of War records for World War II and Korean Conflict, casualty records for Korea and Vietnam Conflicts, and a large collection of datasets resulting from the use of computers for military operations,

management, and research dating from the 1960s, especially during combat in Southeast Asia).

j. Scientific and Technological Data (including registers and surveys of scientific and technical personnel, and data from the National Ocean Survey).

Electronic Records Reference Services

The reference service staff of the Center provides information about the Center's holdings and assists researchers in finding those records that are most useful for their projects. A number of finding aids describe the holdings, including the Title List: A Preliminary and Partial Listing of the Data Files in the National Archives and Records Administration.

Persons seeking information about electronic records should write to the Reference Staff, Center for Electronic Records, The National Archives at College Park, 8601 Adelphi Road, College Park, MD 20740-6001, or telephone the reference staff at 301-713-6645. The Center also uses electronic mail to communicate with researchers; the Internet E-mail address is cer@nara.gov. The Title List and other descriptive materials about the Center's holdings are also available online on the National Archives' gopher/World Wide Web site known as CLIO.

To access the NARA gopher via the Internet, point your gopher client at <gopher.nara.gov>, port 70 (the default):

`gopher://gopher.nara.gov:70`

Researchers can access the Web site with an HTTP client. The URL to use is:

`http://www.nara.gov`

These materials are also available from an FTP site: Use <ftp.cu.nih.gov>, directory `nara_electronic`. A list of the available files on the FTP site is in the file `read.me`.

Copies of records in the Center's custody are available for purchase by researchers in a variety of media, with varying encoding. Photocopies of documentation for electronic records can be obtained either with the data or separately. In some instances, the documentation is available in electronic format or microform in addition to paper. Researchers can also visit the Center's research room to review and self-copy documentation. Currently, the Center for Electronic Records provides researchers with copies of electronic records on either 9-track magnetic tape (1600 or 6250 b.p.i.), or on 3480-class tape cartridge (37,871 b.p.i.). For current information on available media and the cost-recovery fees for reproductions, please request a copy of the Center's descriptive material on services and fees.

A natural outgrowth of the National Archives' improved computer network capabilities and access to information in electronic form in the future may be to provide direct remote access to selected records in the custody of the Center for Electronic Records. The Center is currently planning to build on the capabilities of its electronic records validation system

(AERIC), to improve access to information about the records, or their metadata, and eventually to some of the records themselves.

Some records are natural candidates for remote access and retrieval of specific records. In addition to casualty files from the Korean and Vietnam conflicts, other data files for which such capabilities are appropriate are data files such as the quarterly Federal Assistance Awards Data System and annual Defense Contract Action Data System. They have, respectively, records on Federal disbursements of monies to specific geographic areas by program and contracts of over \$25,000 let by the Department of Defense. However, other large statistical databases, especially survey or sample data, do not necessarily lend themselves to remote access because the records need to be used as full files for analytical purposes or researchers need to identify and extract specific subsets of full files. The Center is exploring using FTP to provide on-line accessible copies of some files, depending on the size of file and other factors. To address the remote access needs of researchers to large, complex data files is a current topic of discussion for future enhancements to the reference services offered by the Center for Electronic Records.

This paper was adapted from the text of National Archives General Information Leaflet 37, "Information About Electronic Records in the National Archives for Prospective Researchers" (rev. 1995) and Theodore J. Hull, "Reference Services and Electronic Records: The Impact of Changing Methods of Communication and Access," *Reference Services Review*, (Summer 1995; vol. 23, No. 2), pp. 73-78..