
Reeling Them In: Accessioning the Electronic Records of the United States Government

by Mark Conrad¹
Archivist
National Archives and Records Administration

The Center for Electronic Records is the unit of the National Archives and Records Administration (NARA) that is responsible for, among other things, appraising the electronic records of federal agencies; accessioning those electronic records identified as permanently valuable; preserving the records once they have been acquired; and providing reference services for the records. This paper will offer an overview of the procedures used in carrying out these responsibilities. Particular emphasis will be placed on the steps involved in accessioning electronic records.

First, we need a definition of records. Records, as defined in 44 U.S.C. 3301,

..include all books, papers, maps, photographs, machine readable materials, or other documentary materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of data in them. Library and museum material made or acquired and preserved solely for reference or exhibition purposes, extra copies of documents preserved only for the convenience of reference, and stocks of publications and of processed documents are not included.

There are several ways in which the Center for Electronic Records may have its first contact with an agency concerning a dataset. Three of the most common ways are:

First, an agency may contact NARA to schedule the final disposition of records. Under 44 U.S.C. 33, no federal agency may dispose of federal records without the approval of the Archivist of the United States. Agencies usually obtain approval for the final disposition of their records through the use of a Standard Form 115, Request for Records Disposition Authority. This form is often referred to as a

“schedule”, “records schedule”, or “SF 115”. A records schedule contains a description of the records to be disposed of and the proposed disposition. If a records schedule contains electronic records, the Center is contacted.

Second, the agency may contact the Center with a direct offer of unscheduled records. This sometimes happens if the agency has produced a dataset that they consider important to documenting the mission of the agency.

Third, the Center for Electronic Records may initiate contact with the agency in an effort to acquire particular datasets with lasting value. If the Center staff learns that an agency is producing important datasets, we contact that agency in an effort to get them to schedule those records. Unlike textual records which can sit for thirty to fifty years or more without significant deterioration, it is important to acquire electronic records soon after their creation to ensure their preservation.

While there are a number of different ways that the Center for Electronic Records can become involved with the acquisition of electronic records, once the process begins, the same steps are usually followed. The records are first appraised to determine if they have sufficient value to warrant their acquisition and preservation.

It would not be practical, possible, or desirable for the Center for Electronic Records to acquire a copy of every electronic record produced by every agency of the United States Government. Appraisal is the process of determining which records will be retained and preserved and which records will be discarded.

Records are appraised as permanent if they have high legal, evidential or informational value. Records have high legal value if they are required to preserve legal rights of the government or individuals affected by the government. Records have high evidential value if they document how a government agency carries out its mission, develops policy, or makes key decisions. Records have high informational value if they contain data that might be valuable to a researcher for reasons

other than the reason the federal agency gathered the information in the first place.

I do not intend to offer a complete explanation of the appraisal process in this paper. The appraisal of electronic records could be the subject of a paper by itself.

The appraisal archivist must weigh a number of factors related to the electronic records at hand in order to determine whether they should be preserved. Some of the questions to be answered are:

- Where did the data come from?
- How was it used?
- Did this dataset have a major impact on federal policy?
- Is the dataset unique?
- Is the electronic format the most desirable format for keeping the information?
- Is the data available in a software/hardware independent format?
- Is there adequate documentation of the dataset to make the data accessible to a researcher?
- Is this dataset likely to be used by researchers?

Once the appraisal archivist has reached a decision and that decision has gone through an extensive review process, the Center for Electronic Records will begin the negotiations with the agency to accession those datasets that have lasting value. Accessioning is the process of transferring legal and physical custody of the records from the agency to the National Archives.

Standard Form 258, Request to Transfer, Approval, and Receipt of Records to the National Archives of the United States (SF 258), is the document used for transferring title of the records to the National Archives. The agency sends the datasets, documentation for the datasets, and the SF 258 to the Center for Electronic Records to begin the accessioning process.

The Center for Electronic Records requires the agency to transfer the electronic records in a prescribed form. All records must be transferred in a software/hardware independent format. Records must be transferred on 7 or 9 track, 1/2 inch open-reel tape at 1600 or 6250 bpi, or on 3480 class cartridges. Records must be in ASCII or EBCDIC with no internal control characters and blocked no higher than 30,000 bytes. These specifications are

used to ensure that the Center will be able to transfer the electronic records to new media as current media become obsolete.

When the tapes arrive, they are checked for readability. If they contain data errors that cannot be corrected by cleaning the tape, the agency is contacted for replacement tapes. If the tapes can be read with no problems, tape maps and dumps of a limited number of records from each dataset are printed. The accessioning archivist compares the tape map with the documentation to verify that the tape described in the documentation is, in fact, the tape she/he is looking at.

The next step is for the archivist to manually compare the printout of several records with the documentation to ensure the record layout and codebook match the actual records. If, for example, a record layout indicates that columns sixty through sixty-five should contain the "respondent's date of birth" and the archivist finds "Peoria" in those columns, clearly there is a problem. The archivist consults the agency for a new record layout or new records.

This validation can be very time consuming. The archivist usually validates less than ten records per dataset, but an accession may contain several hundred datasets and individual records may be several thousand bytes long. The Center for Electronic Records is currently developing a system to automate the validation process.

The application will allow a staff member to enter information about the record layout and codes for a particular dataset. The dataset can then be checked by the computer, byte for byte, against the entered description. The dataset descriptions can be saved and used for analysis of other datasets with the same record layout and codes. The application should allow the Center to preserve the links between individual files from a relational database system. Researchers may eventually be able to formulate and execute queries against the data. This will enable them to identify diverse files sharing common attributes that could be used to link the files. This capability would allow researchers to perform analyses that were not contemplated by the agency or agencies that produced the files. The application should be able to produce public use files for those files that contain restricted data.

In addition to validating the datasets, the accessioning archivist must process the documentation that accompanies the data. The archivist must first verify that the documentation is adequate. While a record layout, code book, and technical data about the tapes used for transferring the records may be all that is necessary to validate

the datasets, these files are electronic records of the agency that produced them. The archivist must make sure there is enough information to document who used the data, how they used the data, and what impact the use of these records had on the agency. Sometimes this information has not been recorded in any form. In this case the archivist may have to interview agency personnel to gather this information.

The archivist must take steps to ensure the documentation will be preserved for the long term. Documentation in paper form is placed in acid-free folders and boxes. Preservation photocopies are made of unstable documents such as newspaper clippings and fax pages. Electronic documentation is subject to the same preservation regimen as the electronic records themselves.

Once the accessioning archivist is satisfied that the data matches the documentation and the documentation is adequate, two copies are made of each dataset. These copies are compared with the original records to verify that they are true copies of the files. The original tapes are then returned to the agency.

At this point the accessioning archivist prepares a final accessioning report. This report lists all the datasets received and processed, identifies any problems that remain to be resolved, suggests how the records should be described, identifies any restrictions on the use of the data, and recommends that the SF 258 should be signed by the appropriate parties to complete the transfer. Once the SF 258 has been signed and copies distributed to the appropriate parties, the accession is completed.

Agencies are required to take steps necessary to ensure the preservation of all unscheduled or permanent electronic records in their custody. When an agency transfers their records to the Center for Electronic Records, the Center assumes responsibility for their preservation. Under the provisions of 36 CFR 1234.28, unscheduled or permanent electronic records must, among other things, be stored under tight temperature and humidity controls. Tapes have to be rewound under controlled tension every 3 1/2 years. The agency must test a sample of their tapes for data loss annually. All records must be copied onto new media at least once every ten years.

These requirements serve as an incentive for agencies to schedule all their electronic records and transfer their permanent records to the Center for Electronic Records on a timely basis as well as ensuring against loss of data in the agencies. At the same time they provide for the preservation of the records until the agency does schedule or transfer their records.

Once electronic records have been accessioned, the

Center staff takes steps to provide access to the records. The records are described. The Center forwards descriptions of new accessions to Archival Publications and Accessions Control. This unit maintains a centralized database of descriptions of records accessioned by units of the Office of the National Archives.

The Office of the National Archives is currently developing a new centralized computing system. The Archival Information System (AIS), as presently planned, will include a module for detailed description of electronic records. When AIS is fully operational researchers will be able to query the database interactively to find records related to their research by using one or more sophisticated search paths.

The reference staff maintains the "Partial and Preliminary Title List of Holdings". This list contains the titles and some information about the availability of some of the datasets that the Center has accessioned. This list is available upon request from the reference staff. In addition to the "Title List", the reference staff produces finding aids to particular collections of data that may be of interest to a wide audience. The reference staff has in depth knowledge of the Center's holdings and offers personalized service to the records.

The Center for Electronic Records does not presently provide on-line access to its holdings. At this time a researcher may purchase a copy of a dataset of interest on 1/2 inch open reel tape.

The Center plans to make datasets available on 3480 cartridges this summer. The Center is also investigating the possibilities for making datasets available on media other than 1/2 inch open-reel tape and 3480 cartridges. The Center does plan to make some of its datasets available for on-line access within the next three years.

This has been a brief overview of the procedures used by the Center for Electronic Records in appraising electronic records of the United States Government, accessioning those records determined to have lasting value, preserving those records that have been acquired, and providing reference services for the records. These procedures are constantly being re-examined and revised in an effort to better meet our responsibilities.

1. Paper Presented at IASSIST 92, Madison, Wisconsin.