Creating a Flexible Preservation Infrastructure for Electronic Records

Karen Estlund and Heather Briston

University of Oregon
kestlund@uoregon.edu, hbriston@uoregon.edu

Abstract. As universities begin to address their first significant collections of electronic records, the needs of the collections often outstrip the resources and support available. This poster will illustrate the steps taken to transition and preserve a presidential electronic records collection into an university archives with limited systems support and preparation for future preservation needs. The infrastructure created was designed to quickly ingest at-risk records and allow for file migration and system evolution as future technologies are implemented.

Keywords: Digital Preservation, Digital Libraries, Preservation Planning, Institutional Archives, Migration

1 Introduction

The transition of a presidency is always a momentous occasion and that is certainly the case in a public university. In summer 2009, the University of Oregon president retired after fifteen years. Although previous presidents had sporadically created electronic records, this was the first presidency to create significant amounts of electronic records along with their traditional paper records. Because of the level of decisions documented and the breadth of topics covered, the records of the presidency are the most important collected by the University Archives and Libraries. Under Oregon Administrative Rule 166-475, the Oregon University System records retention schedule, [1] the bulk of records created in the office are deemed as permanent and must be transferred to the archives when no longer active. This collection of records created a sense of urgency to collect the records before any loss, a need to integrate discovery with the accompanying paper materials, and a need for an infrastructure in the Libraries to handle the records. Based on the needs for this new collection, an opportunity arose to create a new organizational collaboration between the university historian and archivist, the digital collections coordinator, and library systems personnel, creating new working standards for this type of collection.

2 Preservation Planning

Ideally, a repository system based on the Open Archival Information Standard (OAIS) [2] would have been implemented for this collection. The two digital asset management systems used by the University Libraries, CONTENTdm and DSpace, did not fully meet the needs of the collection, which needed to be described on a collection rather than item level and did not require a web presence. The presidential electronic records contain over 6,000 files and additional embedded files. In addition, this collection was unique in the need to fully integrate retrieval of the paper documents along with the electronic records and not support separate systems for each format of electronic files (e.g. images vs. email). In order to facilitate a system that could quickly be constructed for secure ingest, we attempted to follow the principles of the OAIS model with manual controls in an infrastructure where selections for automations and migration to an OAIS compliant repository can easily be added. We used the PLATTER documentation for planning for a trusted repository [3] to help guide the decisions for these tasks, migration schedules, institutional support, and access conditions.

3 Administration

The president's office was accustomed to providing paper records to University Archives; however, a separate conversation was needed for the electronic records. Although they are ubiquitous today, outside of the university archives, most offices, even the president's does not think about the long term preservation and access of electronic records. There is a level of trepidation when it comes to the transfer of electronic records to the archives because of the sensitive nature of materials. While most documents created by the university are public records and subject to the Oregon Public Records law [4], there are numerous exemptions from disclosure, as well as other state and federal laws that require some documents or information to be kept confidential. Many records creators at the university are concerned about the ease of inadvertent disclosure of electronic records, especially outside of their context.

Once transfer was agreed upon, we worked with the Executive Assistant to the President to prepare the electronic records. She went through the documents and e-mail collections and filed many messages, discarded junk mail, as well as flagged confidential or otherwise sensitive items prior to our ingest of the records. In a collaborative meeting with campus IT and the president's office, we paved the way to insure everyone was comfortable with the records transfer, security and access. In this situation, the total transfer was less that 4GB, and for expediency was transferred via a DVD. In the immediate future, the Libraries will have a Web Dav Share in place, so that files can be moved and ingested without transfer via media and set the stage for an OAIS compliant system to be implemented for ingest.

4 Ingest

The preparation conducted by the former Executive Assistant to the President allowed for all permanent electronic materials to be transferred in one batch including: the official university records from the president and his assistant and the president's personal records. The first step in turning these personally managed collections into a library archives collection required an inventory of the file types found among the records. During the evaluation of the files, we also took steps to prepare the files for transfer including changing file names to standard forms without special characters or spaces and ensuring that proper file extension were applied to all files.

We created a plan for migration for files in proprietary and unsustainable formats. The native files and the converted file type are stored in the preservation copy of the collection. A majority of the office documents were converted to PDF files, which we were able to utilize batch processing migration tools. Since .pst file format has recently been released as an open standard, we opted to contain the email in the original format and put on the list of file types to monitor for future migration.

5 Archival Storage, Data Management and Access

After taking the multiple personal information management systems and combining them into one record of an individual and his job as president of the university, we created three distinct areas for the collection:

1. The preservation layer consists of the files in their native format and structure delivered by the President's office along with migrated versions of the files. Records that were created in Archivists' Toolkit[1] to describe the files are exported into EAD XML and live on the server with the archival files. The files are backed up in multiple locations and check-sums are run to avoid bit rot. The log of the file types created during the inventorying process are kept to help monitor future migration needs if they arise. Because issues of confidentiality, privacy, and state and federal record laws apply, access to this section is restricted.
2. The archivists' layer consists of the preservation format of the files and is organized and tagged according to the system devised by the university historian and archivist. Records in Archivists' Toolkit are used to describe the content and point to the server location of these files along with the paper records.
3. The public access layer is a redacted copy of the archives of files that have been determined not to breach confidentiality or contradict any laws guiding access. These files are available on a file server that allows for designated public terminals and staff computers to access read-only versions of the files. Future plans include providing access to files online, as risk is assessed, and integrating into the existing UO Office of the President's Digital Collection.[2]

---

[1] Archivists' Toolkit, `http://www.archiviststoolkit.org/`

[2] `http://boundless.uoregon.edu/digcol/uopres/`

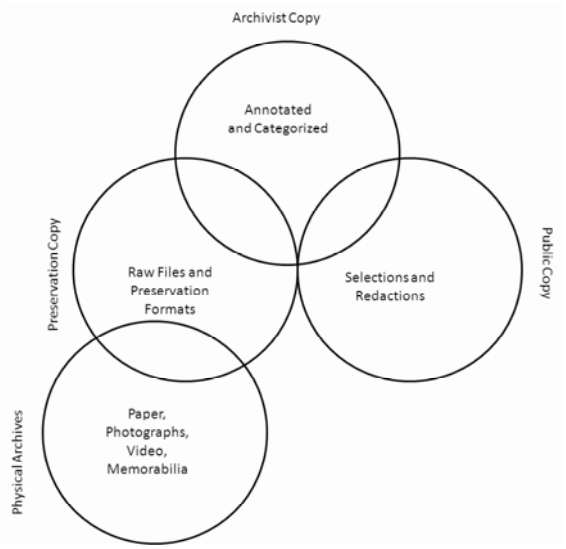

Fig. 1. Preservation and Access of Presidential Records

6 Conclusion

Although it was unrealistic to implement a fully OAIS compliant repository in time to collect these important records, by following the tools and standards provided by the OAIS model and the PLATTER toolkit, we were able to implement a transitional system that meets current needs and can easily be adapted as future technologies are integrated.

References

1. Oregon Administrative Rule, Secretary of State, Archives Division, Oregon University System Records, http://arcweb.sos.state.or.us/rules/OARS_100/OAR_166/166_475.html
2. Reference Model for an Open Archival Information System (OAIS) CCSDS 650.0-B-1 Blue Book (January 2002), http://public.ccsds.org/publications/archive/650x0b1.pdf
3. DigitalPreservationEurope, (April 2008), “DPE Repository Planning Checklist and Guidance DPED3.2", http://www.digitalpreservationeurope.eu/publications/reports/Repository_Planning_Checklist_and_Guidance.pdf
4. Oregon Revised Statutes, Chapter 192, Records, Public Reports and Meetings (Public Records Law), http://www.leg.state.or.us/ors/192.html