

Digital Preservation at the UO Libraries : A Snapshot

by Carol Hixson, with the assistance of Corey Harper

August 2004

Statement of need:

We are at risk of losing key parts of our digital content *in the near term* unless we take immediate steps to fill in some of the gaps relating to our policies, technology, and resources. If the content is worth creating or licensing, it is also worth making sure that it is going to be available and usable five, ten, or two hundred years from now. It is crucial to begin now before we lose any more content. We have already experienced some loss of digital content. For instance, the original sheet music index was created on a platform that became obsolete and was not migrated before the means to read and access the data in it was lost. We are in danger of losing some content of digital images created in Special Collections in the recent past that is stored on CDs where there is already some identified deterioration of the media. There are many other examples of lost or at-risk content that could be cited throughout the Libraries.

Definition of digital preservation:

Preserving digital content entails far more than making backup copies and storing them in disparate locations. Digital preservation is a series of managed activities necessary for ensuring both the long-term maintenance of the files and continued accessibility of their contents. Designed to extend the usable life of machine-readable files and protect them from media failure, physical loss, and hardware and software obsolescence, these activities include:

- ensuring the long-term maintenance of a bitstream (the zeros and ones):
 - backing up files and keeping a copy at an offsite location
 - running checks to track the deterioration of storage media, files or bitstreams
- providing continued accessibility of the contents:
 - *viability* - making sure that information is intact and readable from the storage media
 - *renderability* - making sure that information is viewable by humans and able to be processed by computers
 - *understandability* – making sure that information is able to be interpreted by humans

Attributes of a Trusted Digital Repository (TDR):

A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources for a designated community, now and in the future. Its attributes are:

- Administrative responsibility
 - make explicit commitment to the development of a TDR that complies with prevailing standards, policies, and practices that can be audited and measured
- Organizational viability
 - demonstrate that its organizational mission, legal status, and operations are sufficient to support a TDR
- Financial sustainability
 - establish and can sustain a sound financial base for the TDR
- Technological and procedural suitability
 - avoid technological obsolescence of hardware and software needed to create, store, and access digital content
- System security
 - ensure the reliable operation of the TDR by adopting appropriate security measures
- Procedural accountability
 - establish a systematic approach and supporting techniques to provide comprehensive documentation of all relevant policies, procedures, and practices
- OAIS compliance (see below)

OAIS overview

The OAIS (Open Archival Information System) Reference Model was developed by NASA's Consultative Committee for Space Data Systems to "create a consensus on what is required for an archive to provide permanent, or indefinite long-term, preservation of digital information." It is quickly becoming the accepted conceptual framework for an archival digital preservation system and has been used by OCLC and RLG to develop the framework for Trusted Digital Repositories (TDR). The model attempts to identify the responsibilities and components of an archival system:

- the roles of people and institutions that interact in an archive. These appear in the model as: Producer, Administration, and Consumer.
- the digital objects (information packages) that are managed by an OAIS system. The model identifies three distinct types of information packages: Submission Information Package (SIP), Dissemination Information Package (DIP), and Archival Information Package (AIP). A key concept is that the formats for ingesting and accessing information do not necessarily have to be identical to the format used for archival storage.
- the major functions (things that happen or that should be done) of a fully-functioning archive. The six higher-level functions, which appear in the model as Ingest, Data Management, Archival Storage, Access, Preservation Management, and Administration, represent thirty-three lower-level functions. The diagram of the model intentionally doesn't include a seventh high-level function, Common Services, because its functions are assumed to be provided by any modern information technology system.

Next steps for UO Libraries:

- Convene Digital Collections Gatekeepers and set up monthly meetings
- Working with Digital Collections Gatekeepers and other stakeholders, conduct a gap analysis of our digital collections strategy and infrastructure (using worksheets from workshop to identify issues, problems, etc.)
- Conduct a series of workshops for library staff about the issues of digital preservation
- Inventory all digital content being created or licensed by the library (extent of collection, file formats, current backup procedures, hardware and software used, storage medium, etc.)
- Revise or draft digital collections mission statement for institutional review
- Rewrite or draft key policies related to digital collections (MSU guidelines, Access to Digital Collections, etc.)
- Begin working with GWLA and Orbis/Cascade institutions to share approaches and developments
- Explore possibility of exchanging backup copies of digital content with one or more sister institutions
- Working with Digital Collections Gatekeepers, develop checklist for digital collections approval (following up on work of the Digital Library Initiative)

Essential Background Information:

Digital Preservation Management Tutorial by Cornell University Libraries

<http://www.library.cornell.edu/iris/tutorial/dpm/index.html>

OAIS Reference Model by NASA's Consultative Committee for Space Data Systems

<http://ssdoo.gsfc.nasa.gov/nost/wwwclassic/documents/pdf/CCSDS-650.0-B-1.pdf>

Trusted Digital Repositories : Attributes and Responsibilities : An OCLC-RLG Report

<http://www.rlg.org/longterm/repositories.pdf>