UNIVERSITY OF UTAH

J. Willard Marriott Library

Digital Preservation Program: Organizational Policy Framework (06/07/2010)

SECTION A ........................................................................................................................................... 2 - 5
Purpose
Mandate
Objectives
Scope
Attributes and Responsibilities
Challenges and Incentives
Operating Principles
Roles and Responsibilities
Cooperation and Collaboration
Selection and Acquisition
Access and Use

SECTION B ........................................................................................................................................... 6 - 12
Policies and Procedures
Roles and Responsibilities
Collection Owner Responsibilities
Collection Manager Responsibilities
Digital Preservation Program Responsibilities
Quality Creation and Benchmarking
Selection and Acquisition Policies and Procedures
Transfer Requirements and Deposit Guidelines
SIP Requirements
Deposit Agreement Requirements and Responsibilities
Access and Use Policies
Digital Archive Operations
Platform Requirements and Procedures

SECTION C ............................................................................................................................................ 13- 14
Implementing the Framework
Publicizing and Promulgating the Policy and Plan

APPENDICES .......................................................................................................................................15 - 25
A: Definitions
B: Digital Preservation Decision Checklist
C: Supported Formats Table
D: De-accession Criteria
F: Detailed Functional Model of the OAIS Reference Model
G: References
SECTION A

PURPOSE
The J. Willard Marriott Library (hereafter, Library), in keeping with its mission, serves as a trusted caretaker of the Library’s collections of enduring value\(^1\), including those in digital format. The Digital Preservation Policy Framework supports this mission and is the highest level digital preservation policy document at the Library. The framework makes explicit the Library’s commitment to preserving its digital collections through a comprehensive digital preservation program for both born-analog and born-digital collections. The framework reflects the goals defined in the Library’s SMART goals 2008-2009 and contains references to other relevant Library policies and procedures. The audience for the framework includes Library employees, digital content contributors, donors, and users.

MANDATE
Although many programs and projects both within and outside the Library make objects available to users online, digital preservation implies more than making an object available in a digital format. Digital preservation has been defined by the American Library Association (ALA) as “policies, strategies, and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time.”

The mandate for digital preservation at the Library is linked to institutional responsibility, legal obligations, scholarly commitment, contractual obligations and grants, and membership services (such as Utah Academic Library Consortium (UALC), Greater Western Library Alliance, Mountain West Digital Library (MWDL), etc). Special Collections, Information Technology, University Archives and the Institutional Repository all have missions, whether explicit or implicit, to collect, preserve, and provide access to the historical collections and institutional and scholarly records they hold. In some cases analog preservation will not suffice and the digital preservation of such objects can be inferred.

Additionally, the Library receives grant funding to ensure that specific collections are digitized and made available to online users and the sustainability and long-term accessibility of those collections is often required.

The Library also provides services for outside institutions that need items digitized and made available online. As part of these services, the long-term preservation of selected materials has been written into many formal agreements.

\(^1\) The J. Willard Marriott Library has defined enduring value as unique materials concerning Utah life and/or history as well as materials created by University faculty or Utah residents that fit the Library’s collection mission.
OBJECTIVES
The overall mission of the digital preservation program is to preserve and sustain long-term accessibility to all digital collections created or collected throughout the Library by maintaining a comprehensive digital preservation program. Additionally, it should be noted that in order to manage digital collections over time, the program must include the accessibility of the software and other discovery tools associated with those collections.

Within the overall mission, we have the following objectives:
• Enable uninterrupted (not necessarily instant) access to digital content over time as technology for digital content evolves.
• Collaborate with campus partners, and regional and national institutions to make the best use of resources and avoid duplication of effort.
• Comply with and contribute to the development of the standards and best practices of the digital preservation community.

SCOPE
The Library has primary responsibility for preservation of:
• Digital library resources of enduring value
• Digital resources from outside sources that the Library has contracted to preserve for long-term access

Program limitations: This program’s top priority will not be to preserve objects that are already commercially available elsewhere or that are preserved with a trusted digital repository, except in the case of a future digital preservation strategy (such as the LOCKSS model). The program will assess candidates for digital preservation within budget limitations as well as explicit criteria specified by the Library’s Digital Collections Policy and checklist (url forthcoming).

Program priorities:
• Unique materials in danger of obsolescence in analog form and identified as “critical need” for digital preservation
• Unique materials in digital form in danger of obsolescence or loss.
• Digital collections earmarked by our patrons as requiring long-term access

Timeframe: Our policy, procedures, current and needed technical infrastructure, refined selection criteria, and resources framework will be completed by the end of

---

2 There may be cases in which the program will archive an object that is also available within another trusted repository in order to retain the integrity of the collection. An example of this would be if a faculty member deposits her research with the Library’s Institutional Repository (USpace) but also deposits or publishes some research elsewhere. Another example would include a case in which the Library digitizes a rare book from its collection and Google Books digitizes the same book at a later date. Although that content may be duplicated, the Library would have a responsibility to retain its copy of the digitally preserved book because Google Books has made no claim to be a Trusted Digital Repository and therefore its content is not guaranteed to be available in perpetuity.
2010. At that point, we will assess the overall timeframe for an operational, sustainable, comprehensive digital preservation program.

**ATTRIBUTES and RESPONSIBILITIES**

This framework follows digital preservation standards as defined in OCLC’s Trusted Digital Repositories: Attributes and Responsibilities. Accordingly, the attributes of a trusted digital repository are:

- **Open Archival Information System (OAIS) compliance**
- **Administrative responsibility**
  - Accept responsibility for the long-term maintenance of digital resources on behalf of its depositors and for the benefit of current and future users.
- **Organizational viability**
  - Establish an organizational system that supports not only long-term viability of the repository, but also the digital information for which it has responsibility.
- **Financial sustainability**
  - Demonstrate fiscal responsibility and sustainability.
- **Technological and procedural suitability**
  - Develop policies, practices, and performance that can be audited and measured.
- **Systems security**
  - Ensure the ongoing management, access, and security of materials deposited within it.
- **Procedural accountability**
  - Dependably carry out its long-term responsibilities to depositors and users openly and explicitly.

**CHALLENGES and INCENTIVES**

- **Budget limitations.** We must always live within our financial means. Realistically, we will not be able to preserve everything, making our selection criteria for preservation all the more imperative.
- **Keeping up with technological change in terms of hardware, software, new formats, etc.** A key question here deals with emulation vs. migration of formats.
- **Creating and following submission standards**
- **Meeting the education needs of staff involved with (but not explicitly responsible for) digital preservation.**

**OPERATING PRINCIPLES**

The Library will strive to:

- Comply with OAIS and other digital preservation standards and practices
- Ensure that content remains readable and understandable
- Participate in the development and adoption of digital preservation community standards, practice and solutions
• Develop a reliable, scalable, sustainable, and auditable digital preservation repository
• Manage the hardware, software, and storage media in accordance with environmental standards, quality control specifications, and security requirements

ROLES and RESPONSIBILITIES
The Library accepts responsibility for preserving its digital assets. The Digital Library Council evaluates high-level policy documents and reviews programmatic plans and progress. The Associate Director for Information Technology and the Associate Director for Scholarly Resources and Collections provide input and guidance to the work being done by the Digital Preservation Archivist to manage the digital preservation program and the lifecycle of digital objects of enduring value within the Library. The Head of Digital Technologies, Head of University Archives and Records Management, and the Institutional Repository Coordinator also contribute to the program at various levels.

COOPERATION and COLLABORATION
The Library acknowledges that its digital preservation goals will likely exceed available resources and therefore not be able to guarantee the safety of all digital assets. Therefore, collaboration and partnerships with regional and/or like-minded organizations will be required to ensure the program’s success and to properly prioritize which assets will be addressed and in what order. These may include working with state and regional cultural heritage organizations. Such collaborations may require formal agreements that make explicit the roles and responsibilities of each member in any collaborative.

SELECTION and ACQUISITION
The Library Digital Collection Policy sets the policies and criteria for creating and acquiring digital collections and the Digital Preservation Decision Checklist (Appendix B) guides collection owners regarding preserving digital content of enduring value. The Decision Checklist also reflects criteria for deposit.

ACCESS and USE
Stakeholders of the Library’s digital preservation program include traditional users such as Library departments, patrons, and faculty, and newer stakeholders such as the University and cultural heritage organizations that have deposited archival masters with the Library for long-term preservation. Restrictions to use of collections are defined by the collection holder and vary from collection to collection. Archival files that are open to the public are becoming instantly accessible as we create links between access and archival versions of digital objects. Archival files that have access restrictions require a slightly longer access process that will be detailed later in this document.
POLICIES and PROCEDURES

Roles and Responsibilities
There are several individuals responsible for the digital content connected with the Library’s Digital Archive throughout the content’s lifecycle. Main roles and responsibilities are divided between the collection owner who is submitting materials to the archive (whether they be physical or born digital), collection managers who digitize physical materials, and the archive staff.

What are the responsibilities of collection owners?
- Intellectual property rights: Ensure all proper permissions associated with the deposited content are fully established. This includes the content’s subsequent preservation treatment, e.g. copying.

- Metadata: Submit appropriate descriptive, administrative, structural and possible preservation metadata as required by Library documentation. If collection is submitted to the archive in digital form, technical (and possibly preservation) metadata should be submitted by collection owner at time of deposit. (See http://mwdl.org/index.php/about for current Mountain West Digital Library metadata guidelines).

- Agreement: Sign and maintain a formal Agreement with the Library specifying current materials being deposited and current contact information. This must be completed before the digital collection is created.

What are the responsibilities of the collection manager?
- Reliability: Carry out all digitization processes according to formal Agreement between collection owners and the Library.


What are the responsibilities of the Digital Preservation Program?
1. Insure digital stewardship for all objects.

Harvard University’s digital preservation program defines digital stewardship as “the management of digital objects over the long term through careful digital asset management practices.” Collection managers and Digital Archive staff must work together to manage stored digital objects throughout all phases of the objects' life cycle. The phases are:

a. Assessment phase: Collection manager performs a curatorial assessment of materials intended for the Archive. Assessment includes filling out the New Collection Deposit Form, which specifies such things as initial
format, archival format, access considerations, copyright restrictions, etc.

b. Acquisition and creation phase: Collection manager selects digital formats and defines technical specifications and workflow processes for creation of objects and related metadata. For objects that the Archive will be digitizing, this will include a workflow for digitizing according to archival specifications and metadata creation. For born-digital objects submitted to the Archive, this will include a workflow for possible migration to accepted Archive format and initial checksum verification.

c. Deposit phase: Digital Archive validates each package of digital objects and related metadata that is submitted.

d. Archive and preservation phase: Archive staff will perform yearly fixity checks to ensure the usability of digital objects over time. This includes periodic reports to collection managers about their objects and their refreshment and possible migration to new formats.

2. Reliability: Provide services as agreed to in all Agreements with collection owners.

3. Documentation: Maintain current documentation of supported formats and disseminate the preservation action plan for each supported format. (See Appendix C for supported formats).

4. Financial: Determine costs of long-term preservation and services and disseminate them to Library Administration and collection owners.

5. Preservation: Provide data preservation treatments that are as lossless as required given the Library’s resources and current knowledge.

6. Sustainability: Professionally manage the Program in a way that is administratively, financially, and technically viable long-term.

DIGITAL ASSETS

Quality Creation and Benchmarking

The Library’s Digital Archive is committed to providing long-term storage to all deposited content by applying best practices for data management and digital preservation while also acknowledging the complexities involved in preserving digital information. The Archive commits to preserving content in the form it is originally deposited if deposited in an acceptable Archive format (See Appendix C). The Archive will preserve the content, structure and functionality of the files through migration to newer formats or other preservation strategies, where feasible. The Archive will provide basic services including secure storage, backup, management, and fixity-checks.

At the outset, the Archive will provide preservation support for specific file formats only. We have determined these by applying a set of evaluation criteria including:
prevalence of the format in the marketplace, availability of tools for migration and availability of local resources to take specific preservation actions. Over time, our ability to provide full preservation support for more formats is likely to grow as additional tools and techniques are developed and adequate staff and resources are allocated to fully support the service offered.

This service is currently provided only for formats that are both publicly documented and widely used, giving us a high degree of confidence in our preservation commitment, making it more likely that tools will exist or be developed to undertake preservation actions, and that those actions will result in an understood and controlled migration. The content may also be normalized (transformed to another stable format) to provide additional assurance that functionality is preserved. Finally, if possible, the content will be preserved as originally deposited to ensure the original bit stream is always available. TIFF is an example of a supported format, as its specifications are publicly available and it is well supported and widely deployed.

The formats specified in Appendix C will be re-evaluated at the end of each calendar year to determine if new standards should be adopted. New formats will be evaluated on a periodic basis and when determined to be acceptable standards, those formats will be accepted and old formats migrated to the new standard. In the early days of the program, the Library’s Digital Archive is highlighting functionality over format, meaning it is more important to archive and make accessible the content of a digital item than the actual format it was originally presented in. In the future, it may be possible to preserve both functionality and format.

Selection and Acquisition Policies and Procedures
See the Marriott Library Digital Collection Development Policy for information on the selection and acquisition of digital collections. It is important to note that while many types of digital collections will be supported by the Library for access purposes, only unique collections that are selected for long-term preservation will become the highest priority for long-term management and preservation.

Digital Archive staff and the Digital Library Council will make the final decision regarding which collections will be preserved within the Digital Archive (following the recommendations of collection owners, archivists and subject specialists who are depositing materials into the archive and who will in turn become some of the archive’s main community of users). The Digital Preservation Decision Checklist (Appendix B) will be used as a guide in this process.

As mentioned in Section A, in order to maintain their research value, significant individual collections may be retained in their entirety, even if only a portion of the collection is unique material. This will be done on a case-by-case basis and be decided by collection managers and the Digital Library Council. If space and funding constraints become too difficult, these will be the first collections to be re-evaluated.
It should be noted that the Digital Archive, guided by the Digital Library Council, reserves the right to de-accession collections from long-term preservation on a case-by-case basis, with due observance of institutional and contractual obligations. In cases of de-accession, collections can and should be transferred to another trusted digital repository, if possible. De-accession criteria is detailed in Appendix D.

**Transfer Requirements and Deposit Guidelines**
The diagram below provides a visual representation of the Open Archival Information System (OAIS) model and shows the relation of the various steps in the archival process. The next two pages will go into further detail with regards to the following terms: Submission Information Package (SIP), Archival Information Package (AIP), Dissemination Information Package (DIP). For a more detailed listing of the activities contained in the OAIS model below, please see Appendix F.

![OAIS Functional Entities Diagram](image)

Deposit process: It is our aim to provide public access to as many collections as possible via our Digital Asset Management system (DAM) because we believe that preservation is of little use without minimum level access. Therefore, when a collection is deposited for long-term preservation, the process will depend on whether or not:
1. The collection will be or has been uploaded to our DAM for public access or
2. It is a private collection that needs to be archived but its content not made available to the public for a specified time period

---

3 Currently, our digital asset management is CONTENTdm, which is OCLC’s single software solution that handles the storage, management and delivery of digital collections to the Web.
If the collection IS being uploaded to our DAM
- Our DAM’s Archival File Manager must be enabled and maintained throughout entire upload process.
- Submission Information Package (SIP) needs to be created

If the collection is NOT being uploaded to our DAM
- Submission Information Package (SIP) needs to be created

SIP Requirements
- The SIP consists of:
  - Submitted Digital Preservation Decision Checklist (see Appendix B)
  - Original objects to be digitized or the digital objects to be submitted
  - Associated descriptive (and sometimes technical/preservation) metadata received from collection owners.

All information on the decision checklist and deposit form is informative for digital preservation, but the essential information to document the deposit transaction includes:
- Information that identifies the depositor and a description of the deposit
- Exact listing of the files received (original file name and checksum are good identifiers)
- Date of the deposit

The SIP includes additional metadata, files, or replacement files that were requested or received from the depositor to complete the deposit, when applicable. Persistent identifiers for the submission and the files should be assigned upon arrival or as soon after as possible. The SIP could also include the Metadata Encoding and Transmission Standard (METS), which is the standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language.

The SIP forms the basis of the Archival Information Package (AIP). As the scope of digital content received by the Library expands (e.g., websites, audio and video files), the level and nature of the metadata for new digital content will adjust accordingly.

Deposit agreement requirements and responsibilities
Collection managers can de-accession objects from the Digital Archive as part of a responsible collection management decision, but the Archive will not be a temporary storage solution for digital objects. Temporary storage can be dealt with on a more cost-effective, less curatorially intensive basis by other solutions offered by local or external IT providers.
Access and Use Policies
What are the Library’s access and use policies for objects held in the Digital Archive?

An important part of the Archival Information Package (AIP) is the Dissemination Information Package (DIP). The DIP includes derivative versions of the processed files in acceptable distribution formats, relevant set-up files, and metadata required to read and use the files by patrons of the Digital Archive who will be accessing the content after it has been deposited with the Archive.

For most objects, public access will be via our DAM. The DAM allows most objects within the Digital Archive to be seen by the public, but it does not constitute the archive itself. Currently, access to archival files will be through the CONTENTdm’s Archival File Manager component, either via directly available online links to the archival files for open collections, or through special request for closed collections. Collections that are not included in our DAM will be accessible via special request with due recognition of the Government Records Access Management Act (GRAMA) and the Freedom of Information Act (FOIA).

The Library’s Digital Archive does not manage descriptive metadata. Collection owners, along with collection managers, are responsible for making sure that descriptions of their objects are available in the access portal (i.e. our DAM). This requirement exists at the metadata level only; there is no requirement that the general public has instant access to the high-resolution files for any stored objects. In fact, many collections, due to copyright and donor restrictions, will never make their archival files available to the general public without prior consent by the collection owner.

DIGITAL PRESERVATION STRATEGIES
The digital preservation strategies employed by the Digital Archive can be divided between program and collection level.

Program-level

- Monitor collection owners/users to track changes in their service requirements.
- Monitor available technologies for possible upgrades.
- Assess risks for loss of content posed by technology variables such as commonly used proprietary file formats and software applications.

Collection-level

- On a yearly basis, the Digital Archive staff will evaluate the digital content objects to determine what type and degree of format conversion or other preservation actions should be applied.
• Metadata attachment. The Digital Archive staff will determine the appropriate metadata needed for each object type and how it is associated with relevant objects.

TECHNOLOGICAL INFRASTRUCTURE
Digital Archive Operations

Ingest file formats
Pre-ingest files need to be in one of the acceptable formats outlined in Appendix C. Whether this is accomplished by the collection owner or collection manager once the collection has been transferred to the Digital Archive will be on a case-by-case basis.

Security
The current security of objects within the Digital Archive is assured following specifications outlined in Appendix E.

Emergency Planning
The Digital Archive’s current emergency plan is detailed in Appendix E.

Platform Requirements and Procedures
Platform requirements and procedures most likely will change as improved technology becomes available. These changes can be made by the Digital Archive without the consent of the collection owner, but must maintain the current level of service. Current platform requirements and procedures are detailed in Appendix E.
SECTION C

IMPLEMENTING THE FRAMEWORK

Our short-term plan for implementing the Organizational phase of our Digital Preservation Policy includes both ongoing, yearly actions as well as a two-year strategy towards the implementation of a viable digital preservation program.

Ongoing yearly actions

- Conduct format migration appraisal and migrate to new formats as needed. This will be accomplished in-house, until an improved viable, affordable digital preservation solution is available.

- Continue to provide skilled staff and research, with links to a wider network of distributed development activity, researchers, and services for digital preservation.

- Add to our set of central services, standards, and tools for a greater range of distributed digital preservation services.

2010

- Engage in pilot projects with OCLC’s Digital Archive, Ex Libris’ Rosetta, MetaArchive (and other DP solutions) in order to more fully evaluate and understand Digital Preservation solutions currently available.

- Draft and vet the Technological phase of our Digital Preservation Policy.

- Draft and vet the Financial phase of our Digital Preservation Policy.

2011

- Use the Trustworthy Repositories Audit and Certification (TRAC) criteria and checklist to conduct a gap analysis to determine the trustworthiness of our proposed Digital Preservation Program. This will be done by mapping the complete framework to the existing components, technical infrastructure, and available resources of our digital preservation practices. By doing this, we will be able to more fully understand where we are en route to becoming a Trusted Digital Repository and can more fully map out the specific steps that will lead to our goal. The next steps will become additional goals.

- Further explore other Digital Preservation solutions by engaging in pilot projects or by working with one on a more permanent basis.

- Plan for and gain approval for financial considerations related to implementing a viable digital preservation program.
• Implement the technological and financial phase of our Digital Preservation Policy via a viable digital preservation solution. This could extend beyond 2011.

PUBLICIZING and PROMULGATING THE POLICY and PLAN
The first steps in publicizing the organizational framework will be to:

• Create and maintain a website with links to all related policies, forms and appendixes.

• Engage in public relations activities surrounding website to highlight and disseminate our policy and future plans.

• Engage in training/education activities surrounding policy and plan implementation.

• Compile the experience of research undertaken to draft the policy framework for potential conference presentation and/or written articles on the subject.