# Table of Contents

1. Executive Summary .................................................. 2  
2. Key Findings .............................................................. 3  
3. Recommended Components of a Preservation Plan ............ 4  
4. Open Archival Information System (OAIS) Compliance ....... 6  
5. Archival Information System (AIP) ............................... 7  
6. Storage ........................................................................ 9  
7. Management of Preservation Files/Sustainability Model .. 10  
8. Current Assessment ...................................................... 11  
9. Open Source Preservation Tools .................................... 12  
10. Resources for Further Research ................................... 13  
11. Works Cited ............................................................... 14
1. Executive Summary

The Council of the University of Arkansas College and Research Libraries (CUACRL) Digital Initiatives Task Force (Task Force) was charged with developing and sharing a *Preservation Plan* for unique materials at CUACRL’s member institutions. After reviewing the preservation literature, the Task Force found that progress in digital preservation has been made since approximately 2000. Despite those efforts, the Task Force did not locate a digital preservation plan to use as a guide or template that meets the needs of current CUACRL’s member institutions. There are several frameworks available from the University of Minnesota Libraries and Ohio State University, etc. but not a specific preservation plan detailing the steps on how to achieve digital preservation.

Digital preservation is the current focus of ongoing research at several organizations and institutions. CUACRL’s preservation plan suggests guidelines for member institutions to use when creating a preservation plan. The Task Force understands that implementation will vary based on each member institution’s commitments, resources, and needs.
2. Key Findings

The Task Force, while attempting to develop the preservation guidelines, identified two key findings. The first key finding is the massive production of digital collections without the creation and implementation of preservation and maintenance plans. In the article *Bridging the Gap: Taking Practical Steps toward Managing Born-Digital Collections in Manuscript Repositories*, Ben Goldman states that the “age of digital production creates wide-ranging issues that special collections and archives must confront and manage” (14). Goldman emphasizes, in reference to preservation, that the most critical issue is the “active acquisition of digital-born materials without formal plans for ongoing management and preservation of these materials” (12). This critical issue is referred to as the metaproblem in digital collections and in manuscript repositories. The second key finding is the different meanings attached to the concept of digital preservation. For the remainder of this document, the Task Force interprets digital preservation as follows: “Digital preservation combines 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” (ALCTS Preservation and Reformatting Section).
3. Recommended Components of a Preservation Plan

Below are possible components that each CUACRL member institutions may consider including in a preservation plan.

  a. Creating a mission statement

“Digital Preservation starts at the point of accession and continues throughout the processing of files to end” (The Preservation of Digitized Reproductions). From the Task Force’s interpretation, digital preservation continues through the process of file storage to user access. The Task Force decided that each institution may consider creating a mission statement that identifies the materials for inclusion in the digital preservation plan. The plan should reference the specific institution’s collection goals, research topics, and sustainability factors. It may also include an administrative responsibility statement that lists any mandates and authority statements developed to sustain the program. Member institutions may consider stating, perhaps in the same document, the types of materials and formats the institutions will accept for digitization. For example Harvard Libraries lists their eligible materials as “Items that are out of copyright or which the library has the right to reproduce, items needed for long-term teaching and research use, items from special collections and non-circulating general collections, and requests for can be submitted for one or more items” (Digitizing Library Materials for Teaching and Learning).

  b. Identifying the sustainability factors

The Sustainability of Digital Formats Planning for Library of Congress Collections, located on the Library of Congress website, states the following sustainability factors:
disclosure, adoption, transparency, self-documentation, external dependencies, impacts of patents, and technical protection mechanisms.

c. **Establishing acceptable file formats**

Each member institution may consider clearly stipulating the file formats accepted for preservation of digitized and born digital files. The Library of Congress’ *Format Descriptions* website offers additional guidelines on acceptable file formats on the *Sustainability of Digital Formats Planning for Library of Congress Collections* website.

d. **The Big Picture – Preservation Plan Steps**

i. Accession (analog or born-digital)

ii. Digitization (creating the content for the digital collection) or analyzing born-digital files

iii. AIPs creation (based on OAIS guidelines-create the Archival Information Packages

iv. Storage (move AIPs to the servers assigned for preservation purposes)

v. User’s access (CONTENTdm, IR, etc.) or storage in dark archives

vi. Monitor preservation files
4. Open Archival Information System (OAIS) Compliance

“The Open Archival Information System (OAIS) is a high-level model depicting the combination of tasks carried out by humans and technology for the purpose of digital archiving. The OAIS model encompasses the people necessary, such as Producers, Managers and Consumers, the functional components of preservation, and the information objects produced” (Lavoie 71). Usually, one person is not performing all of these tasks; it is a joint effort.

5. Archival Information System (AIP)

One main component of OAIS is creating the Archival Information Package (AIP).

An AIP is defined as follows: “An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS” (OAIS 9).

Example AIP


The following paragraphs provide a brief explanation of components of an AIP. This is not a prescriptive system; not all components will be used for every package. Each institution may decide which components to include.

a. Unique ID, link to Descriptive Record – Every collection is assigned a unique identifier within the repository.
b. Preservation Description Information
   a. Documentation – The provenance of the collection, which includes deeds of gifts, transfer correspondence, etc.
   b. Logs – Records of system actions such as migration, renaming, accruals, deaccessions, etc.
   c. Manifests – Checksums, file characterization data. Documentation of the accuracy of the changes implemented to the collection.
   c. Preservation Files (locked) – The original files unchanged as created if from analog or as received if from born-digital resources.
   d. Processed Access Copies – The finished files as published or posted in the collection.
      a. Near line files – Files that were or are open for consideration in the collection but not included as of now.
      b. Online files – Files in the collections as processed or modified.
6. Storage

Although storage is not preservation, an institution cannot maintain a file that is safe and free from decay if the file was not initially saved and stored. In reference to storage, there are several theories as to how many copies and what formats to store. The Trustworthy Repositories Audit and Certification Criteria and Checklist (TRAC) indicate that regardless of the storage decision “The repository system must be able to identify the number of copies of all stored digital objects, and the location of each object and their copies. This applies to what are intended to be identical copies, not versions of objects or copies” (44). The following are a few storage recommendations from other sources:

a. The SAA 3-3-3 Redundancy policy recommends having “3 copies, stored in 3 different ways, and in 3 different places” (Colati, Jessica B.; Colati, Greg 52).

b. The LOCKSS project recommends that Lots of Copies Keep Stuff Safe, this project is an agreement among various institutions to store copies of another institution’s material to ensure preservation.

c. Peter Krogh, from the American Society of Media Photographers, recommends that each institution “Have at least 3 copies of your data, Keep these backups on 2 different media, and Store 1 backup offsite” (Krogh, Peter).
7. Management of Preservation Files/Sustainability Model

This section discusses the issue of volatility. Assuming that CUACRL member institutions have created copies of digital collections, in various media, and stored in at least one on-site and one off-site location, the question to answer is how often to audit the files for potential obsolescence. A file format migration is in order if member institutions are aware that a particular file format might become useless in the near future. However, as a policy, each institution should decide how often to inspect the collections in order to accurately render the files to users. Some institutions recommend auditing the files every 5 years. Even in the cases in which obsolescence is not a concern, conducting periodic fixity checks more frequently is also recommended. Anti-virus protection is another aspect of preservation that should be considered when conducting an audit of file formats volatility.

An important aspect of sustainability is the financial commitment by the hosting institution. The institution needs to commit to financial support of long-term file storage and management by staff. The Task Force recognizes that each institution has different needs and capabilities, and urges each institution to create a digital preservation plan specific to the institution that will be completing the work.
8. Current Assessment

There are several tools available to assess your institution’s status within digital preservation. The *National Digital Stewardship Alliance* (NDSA) offers a model of reviewing the technical aspects of various levels of digital preservation. This information is located on the *Levels of Digital Preservation* website; see the works cited section, page 15, for further details.

Another assessment tool is the adaptation of Cornell University’s maturity model for management of digital assets. This assessment is a more comprehensive plan that includes the stages of digital preservation and a self-assessment. For more information about this resource, please see the works cited section under AIDA, page 14.
9. Open Source Preservation Tools

a. Transferring and Processing files
   a. Teracopy – “TeraCopy can be used to move or copy files from one location to another. An advantage of this tool is that the created date does not change (often the created date changes to the date a file is copied). The program also uses CRC checks to verify that the files were copied correctly. (Windows Only)” (Kussmann, Carol R.). Page 15
   b. Data Accessioner - Transfers files from one location to another, creates checksums and basic preservation metadata, page 14
   c. BitCurator- Software used to ingest and to analyze groups of digital files, page 16

b. Bulk processing
   a. Bulk Rename Utility – Renaming files in bulk, page 14
   b. Removing Empty Directories (RED), page 16

c. Checksums
   a. DROID (Digital Record Object Identification), page 16
   b. Karen’s Directory Printer, page 16
   c. Fixity - Regular fixity (checksum) checks, page 14

d. File duplication detection
   a. HashMyFiles, page 15

e. Metadata
   a. DataAccessioner, page 14
   b. NARA – File analyzer and metadata harvester, page 14

f. Other
   a. Disk Inventory X, page 14
b. Windows Directory Statistics, see page 16

10. Resources for Further Research

a. *You’ve Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*, article. The access information is located on page 15 under Online Computer Library Center, Inc.

b. *Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house*, article. The access information is located on page 15 under Online Computer Library Center, Inc.

c. Preserving (digital) Objects With Restricted Resources (POWRR). The access information is provided on page 15 of this document under POWRR.

d. PREMIS - Preservation Metadata Maintenance Activity. The access information is provided on page 15 of this document under the Library of Congress, PREMIS.

e. DACS - Describing Archives A Content Standard. The access information is provided on page 15 of this document under Society of American Archivists.
11. Works Cited


<http://library.duke.edu/rubenstein/uarchives/about/data-accessioner>.


<http://hcl.harvard.edu/collections/digital_collections/digitization_program.cfm#teaching>.


