UNT Libraries: TRAC Conformance Document

Date: October 2015

Version: 1.0

Contributors:

- **Mark Phillips** Assistant Dean for Digital Libraries
- **Hannah Tarver** Department Head, Digital Projects Unit
- **Ana Krahmer** Supervisor, Digital Newspaper Unit
- **Daniel Alemneh** Supervisor, Digital Curation Unit
- **Laura Waugh** Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Table of Contents

Introduction
   Background
   TRAC or ISO 16363
   How We Expect this Document to be Used

Section A: Organizational Infrastructure

Section B: Digital Object Management

Section C: Technologies, Technical Infrastructure, and Security

Appendices
Introduction
The UNT Libraries: TRAC Conformance Document is designed to supplement and provide extended reference to the UNT Libraries’ TRAC Audit Checklist (Appendix A), which outlines the requirements of a Trusted Digital Repository. The self-assessment of the UNT Libraries and its Digital Collections encompasses an evaluation of its associated policies, procedures, workflows, modelling, and technical infrastructure in the TRAC audit process.

Background
Activities related to Web archiving, digitization, and digital object management started at the University of North Texas (UNT) Libraries in the early 2000s. Those initiatives have grown over time, resulting in a large online collection and a number of departments largely or solely responsible for activities to support the Digital Collections.

The Digital Collections
The UNT Libraries host and actively facilitate collection of digital materials for the Digital Collections, which comprise three public-facing interfaces: the UNT Digital Library, The Portal to Texas History, and The Gateway to Oklahoma History. The Digital Library (http://digital.library.unt.edu) contains materials owned or licensed by the Libraries or university entities, and items created by the UNT Extended Community (faculty, staff, students, and administrators) including scholarly materials. The Portal (http://texashistory.unt.edu) is collaborative and contains materials from collections owned by more than 300 partner institutions across the state of Texas. The Gateway (http://gateway.okhistory.org) contains materials owned by the Oklahoma Historical Society, primarily newspapers and photographs.

Administratively, as many of the processes as possible are uniform across collections and material types. All items in the Digital Collections reside in a single infrastructure, built in-house from open source components. Archival files and technical/administrative metadata are stored in the Coda repository; Web derivatives and descriptive metadata are put in the Aubrey access system. Automated processes package, verify, and deposit all digital files for these systems.

Descriptive metadata is also standardized for the Digital Collections. All items are described using the same twenty-one locally-modified Dublin Core metadata fields. This metadata is stored in a local (UNTL) format, but is also available as normalized Dublin Core.

The UNT Libraries supports the Digital Collections to uphold a commitment to their various designated communities. As a program within UNT, the materials support research and scholarship for persons directly associated with the university, but also make the scholarship of UNT available to the wider public. Additionally, materials support the services offered by partner institutions and provide worldwide access to their submitted content.
At the time of this writing, more than 1.4 million items (nearly 163 million files) are archived in the Digital Collections; roughly 4% are not yet visible, awaiting metadata completion. The UNT Libraries provide the widest access possible, based on collection requirements and licensing and restrictions. A majority of items are completely unrestricted and can be viewed (or played) by any member of the public. Around 1% of visible items are restricted in various ways, generally to use by current UNT community members.

UNT Libraries Staff

Development and maintenance of the Digital Collections infrastructure and content is managed within the UNT Libraries by the Digital Libraries Division, with support from the Facilities and Systems Division (see organizational chart below). Summary of the groups directly involved and their primary related activities:

- **User Interfaces Unit** - Involved in the design of Aubrey interfaces for public access and metadata editing
- **Digital Curation** - Manages the processing of born-digital materials for inclusion in the Digital Collections
- **Digital Projects Lab** - Digitizes a variety of materials on flatbed and duplex scanners and creates descriptive metadata for the collections
- **Digital Newspapers** - Handles digitization, processing, and description of newspapers in physical, microfilm, and born-digital formats
- **Software Development** - Develops and maintains components of the archival (Coda) and access (Aubrey) infrastructure
- **Technology and Computer Operations (TACO)** - Provides hardware and software support for Digital Libraries activities
- **External Relations** - Liaises with external partners to collect content and to promote the Digital Collections to user communities

This document references specific responsibilities and activities of these groups when relevant.
TRAC or ISO 16363

Several organizations worked together to develop the Trusted Repository Audit Checklist (TRAC) process -- including OCLC, the U.S. National Archives and Records Commission (NARA), and the Center for Research Libraries -- as a way of evaluating various aspects of repositories to determine if they meet expectations for long-term planning and preservation of their materials. The TRAC documentation is built around the Open Archival Information System (OAIS) reference model that expresses a framework for a repository to ingest, store, and share materials. Formal TRAC audits are administered by the Center for Research Libraries (CRL), though they have also released various tools for organizations that want to perform a self-audit.

Additionally, the International Standards Organization (ISO) started developing a similar metric for determining trustworthiness that could be adopted as an international standard. ISO16363 reached a semi-final “Magenta Book” stage in 2011 and essentially supersedes the TRAC process, though auditing bodies must be certified using guidelines of ISO 16919:2014 (Requirements for Bodies providing Audit and Certification) and it may be some time before the process is fully articulated or available to interested repositories.
The UNT Evaluation
For our process, we chose to use TRAC because more information is readily available, and because other institutions that have completed the TRAC process have posted their documentation, creating a body of examples to use as reference material. Although ISO16363 will likely be considered the “gold standard” going forward, the documentation and work to show TRAC compliance is not mutually exclusive with the ISO standard.

The UNT Libraries’ Digital Libraries Division out of the University of North Texas (UNT) collaborated with the University of Florida to prepare documentation for a TRAC self-audit and to design a process for cross-institutional, peer review of TRAC compliance. Multiple groups have been involved at UNT, particularly in the Digital Libraries Division, in developing the documentation necessary to fulfill the requirements of the TRAC Audit Checklist and TRAC Conformance Document.

How We Expect this Document to be Used
This document was developed to track and evaluate current practices at the UNT Libraries that support the Digital Collections. In this document, we have made a conscious effort to respond each of the sections of the TRAC criteria. The ultimate goal of this document is to provide supportive documentation about the UNT Libraries’ Digital Collections that addresses questions about the policies, processes, technologies, models, and practices of the repository.
Section A: Organizational Infrastructure

A number of policies and documents govern the organizational infrastructure of the UNT Libraries’ Digital Collections and the staff who manage the resources.

University of North Texas Policies (http://policy.unt.edu/)
- Faculty Appointment and Granting of Tenure (15.0): http://policy.unt.edu/policy-by-number/continue/15
- Staff Development/Performance Planning and Review (1.7.3): http://policy.unt.edu/policy-by-number/1

University of North Texas Libraries Policies (https://www.library.unt.edu/policies)
- UNT Libraries’ Digital Preservation Policy Framework (Appendix H)
- UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)
- Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C)
- Locally Created Cataloging and Metadata Records Rights Policy

UNT Libraries Digital Libraries Division Documentation
- Division website: https://www.library.unt.edu/digital-libraries
- Staff directory: https://www.library.unt.edu/staff-directory/department/digital-libraries
- Information about partnering with the Digital Collections: http://www.library.unt.edu/digital-projects-unit/about-partnering

UNT Libraries Digital Libraries Partnership Agreements
- Standard Partnership Agreement (Appendix D)
- Portal-branded Memorandum of Agreement for Digital Rights (Appendix E)
- Digital Library-branded Memorandum of Agreement for Digital Rights (Appendix F)

A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.

The UNT Libraries’ Digital Collections mission statement is documented in the Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C):

In support of the UNT Libraries Mission, the UNT Libraries’ Digital Collections is dedicated to the long-term collection, production, maintenance, delivery, and preservation of a wide range of high-quality digital resources and services for the UNT Community and users throughout the world.
A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.

The UNT Libraries’ Digital Collections are administered by the UNT Libraries’ Digital Libraries Division as a primary service of the UNT Libraries, to support students, faculty, staff, alumni, and administrators of the University of North Texas. The Digital Collections encompass a broad set of digital items, services, technology infrastructure, and standards, administered by professionals whose goals are to provide services to the UNT Extended Community.

The Digital Collections comprise several primary discovery interfaces: The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History. These interfaces exist to provide access to the continuously-growing collection of digital resources made available by partner institutions throughout the UNT Libraries, the University of North Texas, and the state of Texas. Through collaboration, these collections are leveraged to provide an unprecedented amount of digital content for research, scholarship, and learning to the broadest number of users possible.

To ensure access and preservation services for such a large and diverse group of partnering organizations that provide content to the Digital Collections, a multi-faceted approach is in place to provide the organizational rights, succession plans, and formal arrangements for long-term sustainability of the resources entrusted to our repository. Two documents approved by the UNT Office of General Counsel are utilized in this work: a Standard Partnership Agreement and a Memorandum of Agreement for Digital Rights. Additionally, for certain partnerships or collections, an alternative formal agreement, e.g., a Memorandum of Understanding (MOU), is utilized to describe services provided by UNT for digital library collection support and partnership terms.

**Standard Partnership Agreement** (Appendix D): This is a standard contract with content partners interested in contributing resources to the repository. The agreement was accepted by the University of North Texas Legal Council in 2006 and gives the Dean of Libraries the authority to enter into contract with other organizations for digital library services. The agreement was revised in 2013 to stipulate that the metadata created for items in the UNT Libraries’ Digital Collections is in the public domain. This agreement covers the terms and conditions of the access, reuse, rights, and responsibilities of both the UNT Libraries as well as the partner organization. In the event that the UNT Libraries can no longer meet the conditions of the agreement, this agreement states that the content contributed by the partner institution will be returned to the partner.
Memorandum of Agreement for Digital Rights (Appendices E and F): This contract was approved in 2008 by the University of North Texas Legal department and provides the UNT Libraries a non-exclusive license to preserve and provide access to resources owned by the rights owner for works. In the event that the UNT Libraries is no longer able to provide preservation and access services to these resources, they are to be returned to the rights holder.

Memorandum of Understanding (MOU): Large-scale, long-term partnerships are sometimes covered by another form of agreement, a Memorandum of Understanding. These agreements cover the activities of participating institutions involved in digital preservation efforts, are unique to the institutions involved in the respective agreement, and they include provisions about what to do in the event that the UNT Libraries can no longer provide access and preservation services to resources covered by these specific agreements.

Current organizations with which the Digital Collections have MOUs or formal partnerships:

- Government Publishing Office
- National Archives and Records Administration
- Office of the Secretary of State of Texas
- Center for Research Libraries
- Dolph Briscoe Center for American History (Pending)
- Texas State Library and Archives Commission

The documents in the three classes above are archived in both print and electronic formats, with physical issues housed at the UNT Libraries’ Office of External Relations and the digital copies available via the Digital Libraries Division shared network drive in both the “Partnerships” and “Legal” folders. Copies of these agreements are available on request.

While the UNT Libraries is committed to the long-term stewardship of the digital resources entrusted to them and made available by the Digital Collections, there is always the possibility that on an infinite time scale this commitment will shift or change. To ensure viability of the resources to the widest subset of users in the situation where that UNT Libraries can no longer to commit to maintain these resources, the resources will be re-distributed to those partner institutions that originally contributed the resources. Although these partner institutions may not have the local capabilities to provide the same level of services to the digital resources, the UNT Libraries will make an effort to provide a transition plan to the most appropriate organization(s) to maintain the digital library infrastructure available at that time.
A2.1 Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties.

The UNT Libraries established the Digital Libraries Division in 2010 to provide leadership, standardization, stewardship, and technological infrastructure for the UNT Libraries and UNT Extended Community. This division houses the primary units responsible for the day-to-day maintenance and upkeep of the Digital Collections. In close collaboration with the UNT Libraries Facilities and Systems Department and the UNT Libraries Office of External Relations, and with guidance from the various collecting units throughout the UNT Libraries, the Digital Libraries Division provides services in the areas of digitization, born-digital document management, digital preservation, web archiving, metadata creation, user interface design, and preservation of news content. Individuals throughout the UNT Libraries who provide deep subject-specific knowledge in areas such as archives and special collections, music libraries, government documents, collection management, and public services supplement these broad areas of expertise.

The Digital Libraries Division comprises 19 full-time librarians and staff members, and is supplemented by 25-35 part-time student assistants. The UNT Libraries Department of Facilities and Systems fund 6 full-time staff positions who assist with maintenance and upkeep of technical and network infrastructure required by the UNT Libraries’ Digital Collections. The UNT Libraries Office of External Relations works closely with the Digital Libraries Division and assists with two full-time librarian positions that spend large portions of their time working with management of external partnerships, assistance in securing external funding opportunities, and responding to end-user feedback for The Portal to Texas History. Finally, dozens of subject experts throughout the UNT Libraries assist with a variety of services necessary for a large-scale digital library operation.

The overall operation of the Digital Collections is the responsibility of the Assistant Dean for Digital Libraries. The person in this position serves as a member of the UNT Libraries’ administrative council (Dean’s Council) and works with other division and office heads to identify priorities at a high level and to secure technical and personnel infrastructure. The Assistant Dean for Digital Libraries works with stakeholders throughout the UNT Libraries to ensure that the Digital Collections infrastructure and services meet their needs. A collaborative effort between the Digital Libraries Division and Office of External Relations is in place to interact with external partners to integrate their feedback into the development of services and infrastructure.

Librarian and staff positions have established mechanisms outlining job responsibilities and expectations, as well as metrics to ensure that employees are meeting those expectations. Position descriptions for all positions in the Digital Libraries Division are available upon request.
Librarians: Full-time librarians employed by the University of North Texas hold faculty status and are part of a formal library Personnel Affairs Committee (PAC) peer-review process, in conformance with UNT Policy 15.0.1, “Faculty Appointments and Granting of Tenure” (https://policy.unt.edu/policy-by-chapter). UNT Libraries’ PAC documentation outlining internal metrics and processes is maintained and updated on a local SharePoint website and is available upon request.

At the beginning of each fiscal year, every librarian develops a performance agreement with the direction and approval of their immediate supervisor, the Assistant Dean of their division, the Associate Dean, and the Dean of Libraries. The performance agreement is based on three areas: primary assignment, research, and service. An annual performance review evaluates performance in all areas, with input from the supervisor, Assistant Dean for the Division, Associate Dean, Dean of Libraries, and the UNT Libraries PAC.

Staff: UNT Libraries staff also undergo an annual review process, completed through two UNT Human Resources forms: UPO-31 (Planning Guide and Performance Rating Form), and UPO-35 (Performance Agreement and Review). Annual reviews of all UNT Libraries staff take place between April 1st and May 31st of each year, in accordance with UNT Policy 1.7.3 Staff Development/Performance Planning and Review (http://policy.unt.edu/policy-by-number/1). Each year, the staff member’s immediate supervisor evaluates the staff member’s effectiveness in meeting the expectations of the position (UPO-31) and outlines goals for the upcoming year (UPO-35).

A2.2 Repository has the appropriate number of staff to support all functions and services.

The UNT Libraries’ Digital Libraries Division is staffed with 15 full-time positions that directly support the UNT Libraries’ Digital Collections on a daily basis. These positions cover a range of roles and are divided into several sub-groups that work together to provide services to the rest of the UNT Libraries and to the various partners of the Digital Collections. These functional areas include:

Digital Projects Lab: The Digital Projects Lab is responsible for digitization and metadata services for the Digital Collections. This lab is supported by four full-time positions, entailing three librarians and one staff member, as well as a number of student assistants. The digitization and metadata workforce can grow as projects warrant, with a maintenance number of twenty to twenty-five student assistants employed year-round. In addition to digitization activities, the Lab supports digital curation projects for born-digital resources.

Digital Newspapers Team: The Digital Newspapers Unit is the home of the Texas Digital Newspaper Program and associated newspapers collections in the Digital Collections. They have historically been involved in the Texas, Oklahoma, and New Mexico National Digital Newspaper Programs as either the lead institution (Texas) or the technology partner (Oklahoma
and New Mexico). This unit is also responsible for processing newspaper content for The Gateway to Oklahoma History. There are four full-time staff members and four to eight part-time student assistants involved in managing newspaper projects.

**Digital Curation Team:** The Digital Curation Team takes a lead role in projects that comprise primarily born-digital content. Major collections managed by this team include the UNT Scholarly Works Repository, the UNT Theses and Dissertations Collection, and the UNT Data Repository. Primary work areas include interfacing with a variety of local UNT units including the Toulouse Graduate School, UNT Press, UNT Libraries Scholarly Publishing Services, and the UNT Libraries Collection Management Division. This team also works with external partnerships through the Texas Secretary of State for projects such as the Texas Register and the Texas Laws and Resolutions Archive. The Digital Curation Team consists of two full-time librarians and several part-time student assistants.

**Software Development Team:** The Software Development Team is responsible for planning, building, customizing, and maintaining software associated directly with the Digital Collections, including the Aubrey access system, the Coda preservation repository, and the various other systems and tools required to operate the digital library infrastructure. Three programmers make up this team.

**User Interfaces Unit:** The User Interfaces Unit works with stakeholders throughout the UNT Libraries, as well as external partners, to help develop Web interfaces for the Digital Collections. These include the front-end interfaces to The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History. Additionally, they are responsible for Web interfaces for the metadata editing environment and other tools utilized by the digital library team to produce and maintain the Digital Collections. There are two librarians, one project manager, and one interface programmer in this unit.

These units and teams work collaboratively with librarians and staff throughout the UNT Libraries to provide digital library services in support of their local collection development policies and missions. Additionally, these units work with the UNT Libraries Facilities and Systems Department on storage, computing, and network infrastructure, and with the UNT Libraries Office of External Relations, which interfaces with external partners contributing to the Digital Collections.

A2.3 Repository has an active professional development program in place that provides staff with skills and expertise development opportunities.

The UNT Libraries offers a range of training opportunities that directly and indirectly benefit the operations of the Digital Collections and develop the skills and expertise of its staff. These opportunities fall into several key areas including:

- On-the-job training for skill development
- Travel budget to attend conferences and workshops
- Library-wide professional development training and opportunities
- Division-wide professional development opportunities in group writing projects
- UNT Libraries’ Green Light to Greatness internal grant program
- On-site training sessions either in-person or via Web-enabled technologies

Training and Professional Development

Within the UNT Libraries structures, librarians are required to complete annual Performance Agreements at the beginning of each fiscal year (September 1) that cover goals for the librarians’ Primary Assignment, Research and Scholarship, and Service to library and university. In these documents, librarians negotiate with their supervisors for portions of time to complete research and scholarship within their specific research interest areas. The primary assignment offers the supervisor and librarian the opportunity to outline training and professional development opportunities for the upcoming year. At the end of the year, the librarian annotates this with information about how his/her activities of the previous year have fulfilled the performance agreement.

Classified staff at UNT annually complete a performance planning document called a UPO-31. This document outlines their required job duties and areas of professional development deemed necessary for job success in the following year. At the end of the year these documents are used as the instrument for annual review and assessment. If there are specific training needs identified either by the staff member or by their supervisor, these included in this document for the upcoming year.

Travel Funding

The UNT Libraries offers librarians and staff funding by providing a travel allotment that librarians use to participate in conferences, professional development, and training. At this time each librarian is budgeted roughly $1,000 per year for these travel expenses; research time and activities are negotiated annually. Each staff member of the library has a small allotment of $50 per year for professional development. The Digital Libraries Division supplements this travel and
training funding with additional funding from such sources as Indirect Cost Accounts (IDC) from grants received by the division. Each year, members of the Digital Libraries Division propose training and associated travel to the Assistant Dean for Digital Libraries as part of their annual performance agreements. The Assistant Dean for Digital Libraries provides oversight to the process and tries to take advantage of opportunities such as combined training and travel.

In-House and On-the-Job Training

The Digital Libraries Division offers training for the activities required of the librarians, staff, and student assistants. This is provided in a variety of means, including one-on-one training, group training led by another member of the Digital Libraries Division, or via documentation available on the Division's internal wiki. When a new service, workflow, or technology is investigated, documentation that can be used to provide cross-training and can contribute to future documentation is developed as part of the onboarding process.

If a supervisor within the division identifies, or is alerted to, the need for training in an area where sufficient expertise is not internally present, the Digital Libraries Division works with other divisions in the library to investigate opportunities to train other employees, and an appropriate trainer is then contracted.

The Digital Libraries Division offers training related to the Digital Collections to partners across the University of North Texas, to external partners in an on-demand capacity, and to other divisions in the UNT Libraries. Training sessions are varied but include topics such as: digitization basics, managing digital projects and workflows, Web archiving, digital preservation, and the creation of metadata according to the UNT Libraries' metadata documentation and procedures.

See Appendix G for a listing of travel and professional development for the Digital Libraries Division from 2013 to 2015.

A3.1 Repository has defined its designated community/communities and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation requirements will be met.

The designated community of the UNT Libraries’ Digital Collections comprises a broad audience, including:

- The UNT Extended Community (UNT students, faculty, staff, alumni, and administrators)
- Contributing partners and their constituents
- The larger academic community and researchers
The general public throughout the world

The Digital Collections provide primary services of long-term preservation of the content held (both bit-level preservation and format migration) and support for an array of uses of that content, including:

- Persistence of object address (OAIS “Reference”)
- Discovery
- Access

The UNT Libraries’ Digital Preservation Policy Framework (Appendix H) formalizes the UNT Libraries’ continued commitment to the long-term stewardship for, preservation of, and sustainable access to its diverse and extensive range of digital assets. This policy outlines the mandate and reasons for UNT Libraries to preserve digital materials, categories of commitment, levels of preservation, a plan for maintaining preservation, and a glossary of relevant terms.

For further explanation of specific elements and mechanisms of preservation at the UNT Libraries, see sections B and C of this document, particularly B1.4 (the ingest process), B1.5 (physical control over digital objects), B1.7 (demonstrated preservation responsibility for objects), B4.1 (documented preservation strategies), 4.2 (archival storage and migration), C2.1 (appropriate hardware technology), and C2.2 (appropriate software technologies).

**A3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolve.**

The policies and procedures below address the mandatory responsibilities cited by OAIS 3.1 and include the following documentation.

- Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C)
- UNT Libraries’ Digital Preservation Policy Framework (Appendix H)
- UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)
- UNT Libraries’ Policy on Creation and Licensing of Metadata
- UNT Libraries’ Open Source Software Policy (Appendix Q)
- UNT Libraries’ OAIS Information Package Specifications (Appendix M)
- UNT Libraries Archival Storage Replacement Fund (Appendix O)
- Data Loss Escalation Procedures (Appendix R)
- UNT Digital Libraries: Risk Analysis and Management Strategy Plan (Appendix S)
- Preferred File Formats (Appendix L)
- Timetable for Documentation Review (Appendix I)
The UNT Libraries policies will be reviewed and updated as needed with a full review every two years to assure timely revisions as technology progresses and preservation strategies mature (see Appendix I).

**A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed.**

The UNT Libraries’ Digital Collections plan to secure the long-term preservation and access to the resources collected, created, or acquired for the UNT Extended Community and the partners who contribute content to the repository. Consequently, prior to initiating a project that results in building access to a digital collection of materials, contributing partners and digital rights holders are required to enter into formal agreements with UNT Libraries, to extend rights for storing digital files and providing access to the designated community.

These partnerships take a variety of forms, most commonly the form of a Standard Partnership Agreement (Appendix D) or a Memorandum of Agreement for Digital Rights (Appendices E and F), but may be through Memoranda of Understanding (MOUs) and similar formal agreements and documentation. These documents are approved by the UNT Office of General Council. Signed copies of these agreements are stored on a local network drive and are available on request.

In the event that a formal agreement results in extending rights that the contributing partner inappropriately assigned (e.g., through a misunderstanding regarding the content rights holder) or other rights dispute, the UNT Libraries will work with the various parties involved to remedy the situation. The UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J) provides general guidelines regarding some types of removal requests, including requests for redaction and rights challenges. In some situations, a new Partnership Agreement or MOU may designate appropriate rights, or changes to the descriptive metadata could accurately reflect information. For the most extreme circumstances, in which an agreement cannot be reached, the archived digital objects would not be removed from the digital preservation repository (Coda), however, a note can be attached to the archived objects and the digital resources can be removed from the access system (Aubrey).

Also see the UNT Libraries’ Digital Preservation Policy Framework (Appendix H) for more information about the preservation mandates, expectations, and guidelines followed by the UNT Libraries; section A1.2 for a description of partnership agreements; and section A3.1 regarding long-term preservation.
A3.4 Repository is committed to formal, periodic review and assessment to ensure responsiveness to technological developments and evolving requirements.

The Digital Libraries Division, under the auspices of the UNT Libraries, is committed to a formal self-review of the Digital Collections operations and infrastructure every five years of its operation. The first such review was conducted in FY2014-2015 in partnership with the University of Florida. The goal of this partnership was to each complete a local self-review and audit according to the TRAC requirements (see Appendix A) and provide a peer-review to the other institution of their self-review. The findings of these reviews, along with all documentation that is not security sensitive, will be made available online by each institution.

Additional periodic evaluation, review and refinement of documentation and policies relevant to the Digital Collections are included in Appendix I as a general timetable of documentation review and certification. Also see sections C1.7 and C2.2 regarding hardware and technology assessments and updates.

A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time.

The UNT Libraries Digital Libraries Division solicits and actively responds to feedback requests from a number of sources, including Feedback Forms attached to all primary access systems and the UNT Libraries’ Ask Us service. In addition, the Digital Libraries Division welcomes and responds to e-mail and phone communication submitted directly to repository staff. The UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J) describes the policies governing feedback and the interaction with external users in relation to materials in the Digital Collections (http://www.library.unt.edu/policies/).

The UNT Libraries is committed to assessment of its services and the Digital Collections have been included in a variety of assessment activities that include studies of The Portal to Texas History and the UNT Digital Libraries:

- Measuring the Impact of Preserving Digital Assets
  http://digital.library.unt.edu/ark:/67531/metadc93306/

- Applying User-Centered Design Principles to Redesign the Interface to the Portal to Texas History: The IOGENE Project
  http://digital.library.unt.edu/ark:/67531/metadc28324/

- UNT Digital Libraries Values Study
  http://digital.library.unt.edu/ark:/67531/metadc307537/
The Portal to Texas History formally solicits feedback from the Portal’s contributing partners through interactions with the Portal to Texas History Project Development Librarian in the Office of External Relations to better understand how the The Portal to Texas History and the greater UNT Libraries are meeting partners needs.

While user- and producer-feedback is often solicited, collected, and addressed--and has been throughout the history of the Digital Collections--formal studies have only been conducted in an ad-hoc manner. To ensure a more systematic process, from 2015 to 2016, the Digital Libraries Division will develop a plan and set of procedures for collecting feedback from both users of the Digital Collections, and from partners who fill the role of producers for digital resources included in the Digital Collections. This plan will be available via the UNT Libraries website.

**A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content.**

The UNT Libraries make available documentation related to digital library infrastructure, policies, and procedures. These documents affect the preservation strategies of the Digital Collections and may have potential effects on preserving and providing access to digital content. The UNT Libraries’ Digital Preservation Policy Framework (Appendix H), Preferred File Formats document (Appendix L), and UNT Libraries Archival Storage Replacement Fund (Appendix O) are reviewed and modified as needed to keep current with changes in technology and associated developments in the preservation world. Previous versions of these documents are stored on the Digital Libraries Division shared network drive in the *Documentation* folder for the purpose of retaining document versioning.

The Digital Collections document major events relevant to the lifecycle of objects, aggregations of objects, and the system itself in the Digital Libraries Division’s Major Event Log. This application allows us to document events, discuss potential effects of these events on the underlying system or digital objects, and refer to the events in a persistent way.

Changes to the underlying digital repository architecture and software are codified in a locally-maintained version-control system and associated change requests, logs, and specifications are available through the Digital Projects Unit Trac system. The Software Development Unit in the Digital Libraries Division is in the process of migrating code, documentation, and versioning history to a public GitHub account ([https://github.com/unt-libraries/](https://github.com/unt-libraries/)). The Digital Projects Unit makes use of a local installation of MediaWiki to document projects, workflows, and processes used day-to-day in that unit. This platform offers versioning for all content added or edited.
The UNT Libraries performs Web archiving on UNT’s Web presence twice per year and deposits these as part of the University Archive in the UNT Libraries Special Collections Department. These Web archives contain versions of public documentation about and related to the UNT Libraries’ Digital Collections and their change over time. These are accessible through the UNT Digital Library interface.

Finally, the Digital Libraries Division participates in the greater scholarly discussion of digital preservation and access, and describes the process and workflows of the repository and its sub-systems in its writings. All of these publications are made available via the UNT Scholarly Works repository in compliance with the Open Access, Self-Archiving, and Long-Term Digital Stewardship for University of North Texas Scholarly Works Policy, available here: http://openaccess.unt.edu/unt-open-access-policy

A3.7 Repository commits to transparency and accountability in all actions supporting the operation and management of the repository, especially those that affect the preservation of digital content over time.

The UNT Libraries is committed to transparency and accountability in all of its endeavors. The UNT Libraries distributes a regular newsletter to partners and individuals who are interested in the activities of The Portal to Texas History and the UNT Libraries. These newsletters highlight the activities of the UNT Libraries, collection news and events, development efforts, and other information. The UNT Libraries places documentation related to aspects of the repository including technology, policy and procedures, papers and presentations, and accountability (e.g., trusted repository documentation) on its public website. The UNT Libraries is making strides to make the source-code for the repository itself available to the general public for reuse, study, and scrutiny via the UNT Libraries’ GitHub site with an appropriate open-source license (currently a BSD 3-clause license). Making this code available is important to the transparency of repository operations and is covered by the UNT Libraries Open Source Software Policy (Appendix Q, http://www.library.unt.edu/policies/).

As a public university and entity of the State of Texas, the University of North Texas is subject to the Texas Public Information Act, set forth in Chapter 552 of the Government Code. The UNT Policy Manual (https://policy.unt.edu/policy-by-chapter) and specifically the UNT Administration Policy 10.6, “Request for Information Made Under the Texas Public Information Act” (https://policy.unt.edu/policydesc/request-information-made-under-texas-public-information-act-10-6) address the policies and procedures of UNT.
A3.8 Repository commits to defining, collecting, tracking, and providing, on demand, its information integrity measurements.

The UNT Libraries recognize that the integrity of digital resources deposited into the repository stands as the cornerstone concept on which all additional services and operations are built. As such, a number of workflows, processes and systems are in place that verify, log, and store information about the integrity of digital objects throughout their lifecycle in the UNT Libraries’ Digital Collections.

Workflows track the integrity of digital objects from the time that they are deposited into the system in the format of a Submission Information Package (SIP). Internally, the UNT Libraries uses the MD5 message-digest algorithm throughout the repository infrastructure to verify that the content is the same as when it was originally deposited. As outlined in the UNT Libraries OAIS Information Package Specifications (Appendix M) the UNT Libraries uses the BagIt Specification in conjunction with the UNT Libraries Archival Information Package METS Profile for the encapsulation of fixity and other preservation metadata with the object in the systems Archival Information Package.

Whenever a digital object is transferred between storage nodes (network hop) or between logical disks within a computer, a full validation of the digital object is performed to verify that no data was lost. Special care is taken within the conversion code used to convert a digital object from a SIP to an Archival Information Package (AIP) so that fixity values are reused when present and never re-calculated and then stored. This helps to ensure that fixity values for a given file persist across instances of information packages. Before a digital object is accessioned into the Coda repository, it is once again validated as part of the ingest process.

The UNT Libraries performs continuous integrity checks on all content in the Coda repository with the Validate service (also see section C1.5 for a description of the “Validate” module). The Validate service currently operates continuously at a rate that ensures that each digital object in the repository is validated for integrity at least once every 365 days. Additionally, the Coda repository provides curators the ability to order on-demand integrity checks of items within the repository. When ordered, this integrity check is pushed to the top of the Validate service work queue and is scheduled as the next item to verify.

The Digital Collections systems record all actions that occur on objects, including ingest, fixity checks, replication, migration, and other events that act on the content in any way, from the time it enters the repository. These events (including events that happen as a result of continuous or on-demand integrity checks through the Validate service) are logged using the PREMIS Event model and stored in the PREMIS Event Service, a component of the Coda repository. These logged PREMIS Event Entities are associated with the relevant AIPs in the repository and can be viewed as part of the history of the digital object in the Coda repository user interface.
See Appendix N for screenshots of Coda repository tracking and activities.

**A3.9 Repository commits to a regular schedule of self-assessment and certification and, if certified, commits to notifying certifying bodies of operational changes that will change or nullify its certification status.**

The UNT Libraries' Digital Libraries Division will ensure that appropriate policies are viewed and updated as needed with a full review of documentation associated with this conformance document is completed as stated in the Timetable for Document Review (Appendix I) to assure timely revisions as technology progresses, collection focus changes, and digital library collections mature.

This TRAC Conformance Document and other associated documentation contained in the Appendices will be published and linked from the UNT Libraries' website upon completion in Fall 2015. This TRAC Conformance Document stands as the self-assessment piece in the overarching goal of recognizing the UNT Libraries' Digital Collections as a Trusted Digital Repository.

Following the self-assessment, the University of Florida will complete a peer-review of this document during the FY2015-2016 year; the outcome of that external assessment will be available on the UNT Libraries' website for public reference.

As the landscape of external certification in the area of Trusted Digital Repositories matures the UNT Libraries’ will continue to investigate the opportunities of a formal, externally-certified audit.

**A4.1 Repository has short- and long-term business planning processes in place to sustain the repository over time.**

The UNT Libraries’ Digital Collections serve a broad range of end users including the UNT Extended Community, participating partners, and the general public. As such, the long-term sustainability of the repository is of critical importance. The UNT Libraries is part of the formal budget process of the University of North Texas with the majority of its funding coming from a per-credit-hour library use fee.

The Digital Collections are referred to both explicitly and implicitly in the two most recent strategic planning documents for the UNT Libraries, including the goals for the diversification and increase of funding pools for these resources. Several targets and goals are identified specifically for the Digital Collections in each document:

UNT Libraries Strategic Plan 2011-2015: A Plan to Advance the Strategic Research Value of the University of North Texas Libraries
In 2010, the UNT Libraries established the Cathy Nelson Hartman Portal to Texas History Endowment that directly supports The Portal to Texas History. This endowment is to be considered fully funded when it reaches two million dollars and there is an active campaign to secure this funding. Disbursements from this endowment are used to support scholarship, technology, and infrastructure for The Portal to Texas History.

The UNT Libraries established a UNT Libraries’ Archival Storage Replacement Fund (Appendix O) in 2013 to support the ongoing support of the storage used by the Digital Collections. This fund is used to help support the planned refresh costs of storage infrastructure associated with the Digital Collections. See section C 1.7 for a description of the five-year technology rotation.

**A4.2 Repository has in place processes to review and adjust business plans at least annually.**

A discussion of the budget and any related changes to the process and operations of the UNT Libraries’ Digital Collections is taken into account annually by the UNT Libraries as part of the yearly budget allocation by the Dean of Libraries with input from the Dean’s Council. As changes are required throughout the year, these requests are brought to the UNT Libraries’ Dean’s Council and finally the Dean of Libraries for necessary funding approval and allocation. The requests, allotments, and overall budgeting is usually presented in the support for the Digital Libraries Division and for projects originating in other departments that have a digital component.

**A4.3 Repository’s financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.**

The UNT Libraries’ financial procedures are subject to audit by University of North Texas System Internal Audit (http://www.untsystem.edu/internalaudit/purpose.htm) and any outside auditors contracted by the UNT System Board of Regents. All documented activities are subject to the Texas Public Information Act (http://www.statutes.legis.state.tx.us/Docs/GV/htm/GV.552.htm) and may be reviewed with appropriate requests made to the University of North Texas System Office of General Counsel (http://www.untsystem.edu/generalcounsel/pub-information.htm).
Financial reports for the University of North Texas are submitted annually in compliance with the TEX.GOV'T CODE ANN §2101.011 (https://policy.unt.edu/policydesc/financial-reporting-2-0-1) and in accordance with the requirements established by the Texas Comptroller of Public Accounts. Financial reports are considered for audit by the State Auditor as part of the audit of the State of Texas Comprehensive Annual Financial Report (CAFR) (http://www.texastransparency.org/State_Finance/Budget_Finance/Reports/Comprehensive_Annual_Financial/).

A4.4 Repository has ongoing commitment to analyze and report on risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).

UNT’s Office of Risk Management http://riskmanagement.unt.edu/ is responsible for assessing risk for all UNT units, including the Libraries.

The UNT Libraries is subject to the processes and procedures for auditing including asset labeling, licenses, and liabilities as specified by the University of North Texas, The UNT System and the UNT Board of Regents. Financial Reports for the University of North Texas are publicly available: http://www.untsystem.edu/financialreports.htm

The UNT System Office of Finance is responsible for investment and debt management. All expenditures, including those for assets and licenses, are managed and monitored by various offices within the Office of Finance. Liabilities are also managed and monitored by the Office of Finance. The UNT Libraries, as part of the University of North Texas System, are recipients of the services of all the UNT System offices and subject to management and monitoring by them. The Libraries are responsible for reviewing all of its revenues and expenditures using online financial reporting tools available to all UNT units. The Libraries and all UNT units are expected to be stewards of the Universities’ money and, as such, to responsibly use the funds budgeted for them. The Libraries are committed to regularly assessing future needs for technology infrastructure to include the storage and equipment needed for managing and providing access to digital resources. The Libraries are responsible for maintaining reserves adequate to those future needs.

The Facilities and Systems Division in the UNT Libraries manages external licenses for software used by the Digital Collections as part of its technical infrastructure. Additionally this group maintains documentation related to the technology lifecycle and refreshment periods of various pieces of infrastructure such as servers, storage and networking tools used by the Digital Collections.
In 2015 the Digital Libraries Division developed a document titled *UNT Digital Libraries: Risk Analysis and Management Strategy Plan* (Appendix S) to help identify perceived and potential threats, and planned or implemented responses that related to the Digital Collections’ funding, staffing, technology, and infrastructure including physical and natural risks.

**A4.5 Repository commits to monitoring for and bridging gaps in funding.**

The Digital Collections within the UNT Libraries are operated by the Digital Libraries Division for the UNT Extended Community including students, staff, faculty, alumni, administrators. Additionally external partners, their constituents, and the general public rely on services provided by the Digital Collections in a variety of ways. Because of this, it is imperative that the Digital Collections are able to continue operation through a gap in funding.

The Assistant Dean for Digital Libraries is responsible for monitoring high-level projects, infrastructure developments and purchases, personnel, and costs associated with running the repository. This information is reported to the UNT Libraries Administration, including any changes in the expected costs for operating the Digital Collections above allotted funding levels defined during the yearly budgeting process, or when new projects and initiatives occur.

The Digital Collections is positioned in a number of ways to bridge gaps if they occur:

**Infrastructure planning** - The Digital Collections operates on technical infrastructure that has been deployed in stages through a five-year period so that the entirety of the technical infrastructure will never need to be replaced at the same time. This allows the technology costs associated with replacement to be spread across fiscal years and for more reasonable planning.

**UNT Libraries’ Archival Storage Replacement Fund** (Appendix O) - This fund was designed as a way of creating a reserve that is used for the replacement of the Digital Collections technical infrastructure. Having this fund in place allows for the allocation of funds in preparation of technology purchases several years in the future.

**Cathy N. Hartman Portal to Texas History Endowment** - The disbursements of this endowment fund can be used to support a wide range of activities for The Portal to Texas History including technology infrastructure purchases. Because of the shared infrastructure design of the Digital Collections, purchases for The Portal to Texas History can also be leveraged by other portions of the infrastructure and will support the greater Digital Collections.

**Local funds allocated for use by the Digital Libraries Division** - The Digital Libraries Division has two local departmental accounts that can be used for a wide range of needs including the salaries of staff and librarians, travel, training, infrastructure and technology purchases. In the event of a funding gap these accounts would be available to help bridge the gap as necessary.
UNT Libraries budget - The UNT Libraries budget can be utilized to help bridge any identified gaps. The UNT Libraries maintain a reserve in its budget to deal with unexpected expenses and shortfalls in funding for important services and programs, to be allocated at the discretion of the Libraries Administration. This fund in conjunction with the overall Libraries budget could be used to help bridge any gaps in funding that occur for the Digital Collections.

UNT Campus and System Funding - The Digital Collections are a major service of the UNT Libraries to the UNT campus and as such, if there were a catastrophic gap in funding that the previously-mentioned options could not address, the UNT Libraries could approach both the UNT Denton Campus and, if needed, the UNT System for resources. Although this is considered a last resort in the list of possibilities, it could be leveraged in extreme circumstances.

A5.1 If repository manages, preserves, and/or provides access to digital materials on behalf of another organization, it has and maintains appropriate contracts or deposit agreements.

Deposit of content in the Digital Collections is governed by a Standard Partnership Agreement (Appendix D), Memorandum of Agreement for Digital Rights (Appendices E and F), or Memorandum of Understanding (MOU) between the University of North Texas and another organization.

Information for organizations to become partners of The Portal to Texas History or the UNT Digital Library can be found at: http://www.library.unt.edu/digital-projects-unit/about-partnering
This information explains the process for becoming a partner, models of partnership for working with the UNT Libraries, funding options for partners to contribute materials, responsibilities of partners and UNT for project completion, and the documents and agreements required to begin the partnership.

The documents in the three classes above are archived in both print and electronic formats, with physical issues housed at the UNT Libraries’ Office of External Relations and the digital copies available via the Digital Libraries Division shared network drive in both the “Partnerships” and “Legal” folders. Copies of these agreements are available on request.

Also see section A1.2 for more information about the agreement types and standing MOUs.
A5.2 Repository contracts or deposit agreements must specify and transfer all necessary preservation rights, and those rights transferred must be documented.

As part of the Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, or Memorandum of Understanding (MOU), the UNT Libraries secures necessary rights to preserve digital resources into the future.

The documents in these three classes are archived in both print and electronic formats, with physical issues housed at the UNT Libraries’ Office of External Relations and the digital copies available via the Digital Libraries Division shared network drive in both the “Partnerships” and “Legal” folders. Copies of these agreements are available on request.

A5.3 Repository has specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties.

As part of the Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, and Memoranda of Understanding (MOUs), the UNT Libraries has agreements with depositors of content to cover the acquisition, maintenance, access, and withdrawal of items in the UNT Libraries’ Digital Collections. All signed agreements between partner institutions, individual rights holders, or other organizations in the form of a formal agreement or MOU are shared between the institution and are permanently retained by the UNT Libraries in both paper and electronic formats.

See also section A1.2 for further information about the documents listed above and standing MOUs, and section A5.1 for information about deposit agreements and about partnering with the Digital Collections.

A5.4 Repository tracks and manages intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.

The Standard Partnership Agreement (Appendix D), Memorandum of Agreement for Digital Rights (Appendices E and F), and Memoranda of Understanding (MOUs) specify the responsibilities of the UNT Libraries in tracking and managing intellectual property rights for resources deposited in the Digital Collections. All items not in the Public Domain or under other Open Access License (such as Creative Commons) are covered by copyright.

The UNT Libraries' Digital Collections Usage and Feedback Policy (Appendix J) defines the levels of access that users have to resources in the Digital Collections that include: Public
Access, UNT Community, UNT Community Strict, and Physical Premises. These access levels are used to grant or restrict access to end users of resources in the Digital Collections, according to the guidelines and definitions in the policy.

As of 2014 the UNT Libraries and partners of The Portal to Texas History make available all bibliographic or descriptive metadata as Public Domain content under a CC0 license. This treatment of descriptive metadata was incorporated into the UNT Libraries Standard Partnership Agreement in 2014 and these agreements were re-signed by partner institutions. The UNT Libraries’ Locally Created Cataloging and Metadata Records Rights Policy is available at: http://www.library.unt.edu/policies/

See also section A1.2 for further information about and definitions of agreement versions.

A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights.

The UNT Libraries make every effort to ensure that it has appropriate rights to ingest and provide access to content. In cases where the rights are unclear (e.g., when staff members encounter copyright information that, relative to the work in hand, is ambiguous or contradictory), the UNT Libraries archives, but does not provide access, to the item. The UNT Libraries work with the contributing partners for items that have identified ambiguous or contradictory information to clarify rights associated with the object.

The UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J) covers procedures for responding to complaints and for addressing challenges to access rights. In addition, issues related to personally-identifiable information and requests for item redaction are covered by this policy.

The UNT Libraries do have a mechanism in place to ensure that partner institutions do in fact have the rights to the resources they deposit into the Digital Collections. These rights statements are explicitly defined in the Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, or a Memorandum of Understanding (MOU) signed by the University of North Texas and other institutions.

While specific verification of rights is not handled by the UNT Libraries, if there are concerns or questions identified by staff for the Digital Collections or other curators, they are submitted to the partner organization responsible for the items in question. These concerns could be for either resources that seem to be beyond the scope of rights that might be held by the partner, or for items that contain obvious personally-identifiable information. When identified, the UNT Libraries works with the depositing partner to define and implement a workable remedy to the questions.
If a resource already online in the Digital Collections has associated questions related to copyright or personally-identifiable information, the UNT Libraries encourage people with information about the item to contact utilize the feedback mechanism of in the access interface for the Digital Collections to elaborate on the permissions associated with the resource. The process for working with these requests are discussed in the UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J).
Section B: Digital Object Management

The UNT Libraries adopted the OAIS Reference Model as the guiding model for the design and implementation of the UNT Libraries’ Digital Collections and its underlying infrastructure and policies.

Additionally, there are various policies and documented standards that support the digital object management activities at UNT:

- UNT Libraries policies (http://www.library.unt.edu/policies):
  - UNT Libraries Digital Preservation Policy Framework (Appendix H)
  - Collection Development Policy for the UNT Libraries' Digital Collections (Appendix C)
  - UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)
- UNT Libraries METS profile registered at the Library of Congress (http://www.loc.gov/standards/mets/profiles/00000045.xml)
- OAIS Reference Model documentation:
  - Formal Statement of Conformance to ISO 14721:2012 (Appendix B)
  - UNT Libraries OAIS Information Packaging Specification (Appendix M)
  - UNT Libraries Digital Workflow Diagram (Appendix T)
  - UNT Libraries Ingest Workflow Diagram (Appendix U)
  - UNT Libraries SIP-to-AIP Conversion Workflow (Appendix V)
  - UNT Libraries AIP-to-DIP/ACP Conversion Workflow (Appendix W)
- UNT Digital Libraries standards and public documentation:
  - Digitization Standards (Appendix K) - http://www.library.unt.edu/digital-projects-unit/standards
- Standard forms for collecting rights/provenance information about items:
  - Standard Partnership Agreement (Appendix D)
  - Portal Memorandum of Agreement for Digital Rights (Appendix E)
  - Digital Library Memorandum of Agreement for Digital Rights (Appendix F)
○ Project- or partner-specific agreements or Memoranda of Understanding are available upon request

B1.1 Repository identifies properties it will preserve for each class of digital object.

The UNT Libraries are committed to preserving the intellectual content and, in most cases, the appearance and layout of materials deposited in the UNT Libraries’ Digital Collections. For materials that have a standardized sequence (e.g., books, manuscripts, documentation, audio, and video records), the UNT Libraries’ Digital Collections store and preserve metadata detailing the sequence of files for the digital object.

The UNT Libraries maintains a set of standards for the digitization of image- and text-based materials, moving images, and audio (see Appendix K). These guidelines dictate the minimum standards accepted for digitized analog materials and include information such as resolution, bit-depth, file format, and file compression.

While the UNT Libraries do not have an explicit list of accepted file types, the “Digital File Formats” document (Appendix L) includes a list of commonly-accepted and preferred formats and standardized normalization procedures to provide a framework. Born-digital resources are acquired in the richest (most original) file formats available; most proprietary files undergo format normalization to convert them into a more standardized format as identified in Digital File Formats (Appendix L). In all cases, the master copies of resources are retained as part of the Archival Information Package (AIP).

Recognizing the complex issues of file formats and their preservation needs, the UNT Libraries are committed to bit-level preservation and format migration of materials created according to our documented specifications as technology, standards, and best practices in the digital library community change. More information can be found in the UNT Libraries Digital Preservation Policy Framework (Appendix H).

B1.2 Repository clearly specifies the information that needs to be associated with digital material at the time of its deposit (i.e., SIP).

The Submission Information Package (SIP) used by the UNT Libraries (UNTL-SIP) encapsulates information about the digital object and includes provenance information about the depositor of the digital resource. The specification for UNTL-SIPs was developed in 2009 and has changed very little over the subsequent years. As of 2015, there have been over 1.3 million submission packages created and ingested into the Digital Collections and processed into both Archival Information Packages (AIPs) and Dissemination Information Packages (DIPs).
For full documentation of the UNTL-SIP, see the UNT Libraries’ OAIS Information Package Specification (Appendix M).

**B1.3 Repository has mechanisms to authenticate the source of all materials.**

The UNT Libraries' Digital Collections consist of digital content created, collected, or contributed by the UNT Extended Community or by contributing content partners. Material contributed by partners external to the UNT Libraries are covered by one of the following: Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, or standard Memorandum of Understanding (MOU) signed by the University of North Texas and other institutions. Appendices D-F display blank copies of the Standard Partnership Agreement and the Memorandum of Agreement for Digital Rights for The Portal to Texas History and the Digital Library, respectively. Specific Memoranda of Understanding or signed agreements are available upon request.

The UNT Libraries maintain two pieces of metadata necessary to authenticate the source of materials from a given partner. First, information about the depositing entity is documented as part of the BagIt Specification bag-info.txt document (see Appendix M) and can be used to track a digital object to its original partner and the project or workflow that was responsible for the collection or creation of that resource. Additional information identifying the partner institution responsible for the digital resource in the Digital Collections is recorded in the descriptive metadata record. This information appears in the “Partner” or “Institution” field in the UNTL Metadata Schema. Because changes to descriptive metadata records are versioned and maintained over time, it is possible to identify if the partner institution value was ever changed or if it remains the same as when originally deposited.

The Digital Libraries Division has standard workflows for processing and managing born-digital and digitized physical materials when received from external sources or when digitized by the Digital Projects Unit, Digital Curation Unit, or the Digital Newspaper Unit. These workflows are documented on an internal wiki used by the Division, and is available upon request.

**B1.4 Repository’s ingest process verifies each submitted object (i.e., SIP) for completeness and correctness as specified in B1.2.**

The UNT Libraries ingest process conducts the following tests on each submitted item:

- Container: internal tests to ensure BagIt conformance.
- Digital Signatures (MD5 checksums): checksum verification of all deposited files against fixity values submitted with the SIP.
- Digital Object Well-formedness: verifies that automated packaging and conversion into UNT Libraries’ Archival Information Package (AIP) will be successful at an object level.
- Metadata: internal tests to ensure UNTL conformance and completeness as defined by the UNTL XML Schema - [http://digital2.library.unt.edu/untl.xsd](http://digital2.library.unt.edu/untl.xsd)
• File Characterization: JHOVE used for well-formedness and validation of individual files upon ingest. Output of this process is stored as part of the AIP. File characterization is distinct from digital object well-formedness in that it occurs at the file-level rather than the object-level.

Note: These processes do not detect problems originating in capture (e.g., missing pages, audio tracks, etc.), nor do they detect readability or other problems that cannot be algorithmically detected.

Detailed information about the requirements related to packaging of content for ingest into the system is documented in the UNT Libraries’ OAIS Information Packaging Specification (Appendix M).

**B1.5 Repository obtains sufficient physical control over the digital objects to preserve them**

As a function of the design of the UNT Libraries’ Digital Collections’ infrastructure, all preserved digital content resides locally on UNT’s servers. In support of this infrastructure, the UNT Libraries maintain the following:

1. **Standard Partnership Agreements and Memoranda of Agreement for Digital Rights (Appendices D-F):**
   - The UNT Libraries obtain signed copies of all Standard Partnership Agreements and Memoranda of Agreement for Digital Rights before acquiring and ingesting materials into the Digital Collections. These agreements are filed with the UNT Libraries Office of External Relations.

2. **Workflow Documents:**
   - The UNT Libraries Digital Workflow Diagram (Appendix T) documents ingest procedures for content collected and digitized by the UNT Libraries Digital Projects Unit and for content contributed via the Digital Projects Unit from external partners.
   - The UNT Libraries Ingest Workflow Diagram (Appendix U) documents the steps of digital object creation, from the time objects are ingested until they are accessible to end-users.

3. **Digital Collection Policy:**
   - The Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C), available at [http://www.library.unt.edu/policies/](http://www.library.unt.edu/policies/), specifically states that items added to the collection will be permanently maintained on local storage through the UNT Libraries.

4. **Records of preservation events:**
   - As digital objects are ingested, or when preservation operations are conducted, a record of these events and the date and time of their occurrence is documented in the log files. These records use the PREMIS XML container and are logged within the Coda PREMIS Event Service, operated by the UNT Libraries Digital
When logged with the PREMIS Event Service, records of each event are referenced by the unique identifier used within the Coda Repository.

**B1.6 Repository provides producer/depositor with appropriate responses at predefined points during the ingest processes**

The Digital Libraries Division notifies partners after the successful ingest of items into the system, or if there are issues with the deposited content that cannot be remedied through standard workflows. In addition to this notification of completion, partners can query the administrative system or Application Programming Interfaces (API) associated with their content to understand what has been successfully ingested. If the partner needs to track their resources during the ingest workflow before it is completely ingested, they can contact the Digital Project Unit staff member for information regarding the ingest process of their content.

**B1.7 Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e., SIPS)**

After the ingest process is complete for a given item, the item is registered in the preservation repository (Coda) and the appropriate access system (Aubrey). The presence of the resource in these two systems represents formal acceptance of preservation responsibility.

**B1.8 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition)**

The UNT Libraries package digital content according to the UNT Libraries OAIS Information Packaging Specification (Appendix M), which defines the structure and required elements necessary for the creation of a UNT Libraries Submission Information Package (UNTL-SIP).

A UNTL-SIP contains metadata about the digital object being packaged, its provenance, and fixity information about files that constitute the digital object (OAIS Content Information). Metadata is stored using the BagIt specification. Relevant metadata describing the content being packaged, the person responsible for packaging the item, original identifiers, and file names, and dates of packaging (which signifies the start of the ingest) are included in the bag-info.txt file in the bag. Fixity information in the form of MD5 hashes is stored in the manifest-md5.txt file in the root of the bag.

It is standard practice to extract preservation metadata for each of the files that constitute a digital object during the SIP-to-AIP conversion process. Preservation metadata in this context encompasses information about the file -- including creation date, file format information, and fixity information -- which is serialized as a PREMIS Object Entity using the PREMIS XML.
format, one entity per file. During this process additional packaging information is created for the candidate digital object and subsequently stored in a METS record conforming to the UNTL METS Archival Information Package Profile available on the METS Profile section of the Library of Congress’ Website: http://www.loc.gov/standards/mets/profiles/00000045.xml.

**B2.1 Repository has an identifiable, written definition for each AIP or class of information preserved by the repository.**

The Archival Information Package (AIP) used by the UNT Libraries is formally described in the UNT Libraries’ OAIS Information Package Specification (Appendix M).

The UNTL-AIP serializes the structure of the digital object with a METS record conforming to the UNTL METS Archival Information Package Profile (http://www.loc.gov/standards/mets/profiles/00000045.xml) and further encapsulates the files associated with the digital object within a bag as specified by the BagIt 0.97 Specification.

**B2.2 Repository has a definition of each AIP (or class) that is adequate to fit long-term preservation needs.**

The UNT Libraries uses a unified Archival Information Package format for all classes of digital objects deposited in the Coda repository. This UNTL-AIP serializes the structure of the digital object with a METS record conforming to the UNTL METS Archival Information Package Profile (http://www.loc.gov/standards/mets/profiles/00000045.xml) and further encapsulates the files associated with the digital object within a bag as specified by the BagIt 0.97 Specification.

The Archival Information Package (AIP) used by the UNT Libraries is formally described in the UNT Libraries’ OAIS Information Package Specification (Appendix M).

**B2.3 Repository has a description of how AIPs are constructed from SIPs.**

The UNT Libraries maintains documentation that defines the conversion process for UNTL-SIPS into Archival Information Packages (AIPs). Documentation on the UNT Libraries SIP-to-AIP Conversion Workflow (Appendix V) is codified in the ingest process workflow in a step called makeAIP.py as a set of Python scripts that are run on the content on one of several ingest servers in use by the Digital Projects Unit.
B2.4 Repository can demonstrate that all submitted objects (i.e., SIPs) are either accepted as whole or part of an eventual archival object (i.e., AIP), or otherwise disposed of in a recorded fashion.

The UNT Libraries’ ingest process is designed to validate each SIP received to ensure that it aligns with the UNT Libraries’ Submission Information Package (SIP) specifications before converting from SIP to AIP. Each UNTL-SIP is converted to a corresponding UNTL-AIP, with a one-to-one relationship.

If a SIP fails to meet these specifications, it is quarantined. These anomalies generate documentation in log files and information about what caused the SIP to fail during the ingest process. The person responsible for ingesting items to the repository is responsible for identifying the cause of the issue via available log files and enacting necessary actions to remedy the resource prior to the batch being considered “complete.” Items that fail the ingest process are immediately dealt with so that they do not remain in an un-ingested state for any lengthy period of time.

The only time a SIP would be disposed of completely without conversion to an AIP is if the SIP was submitted as part of a testing run. These test ingests are carried out in a section of the repository infrastructure that does not result in the acquisition of the test resource into the Coda repository. Appendix M, “UNT Libraries OAIS Information Package Specification,” further details this process.

B2.5 Repository has and uses a naming convention that generates visible, persistent, unique identifiers for all archived objects (i.e., AIPs).

The UNT Libraries make use of Archival Resource Keys (ARKs) as part of its persistent identifier strategy. Current and previously used identifier designations include:

- The Name Assigning Authority Number (NAAN) for the UNT Libraries is 67531, i.e., ark:/67531/uniquename.

- Unique names are assigned by a locally-developed identifier service, referred to at UNT as a “number server,” which mints identifiers under five current namespaces: metapth, metadc, metarkv, metatest, and coda. (See Table 1.)

- A historic and now-deprecated prefix, metacrs, is also present in the Aubrey access system, but no new identifiers are minted under this.
Upon verification that a SIP is valid and meets the requirements for a successful conversion from SIP to AIP, an external-facing ARK identifier is assigned based on the namespace associated with the ingest workspace.

In all situations, a digital object retains the locally-supplied identifier (via the submitted SIP folder name and bag-info.txt file), and is assigned an external-facing ARK identifier and internal-facing Coda ARK identifier, which are stored with the object and made discoverable via the Coda repository system. (See section C1.3 for a description of how internal and external identifiers are assigned during the ingest process to track items.)

<table>
<thead>
<tr>
<th>Namespace</th>
<th>Status</th>
<th>Internal/External Facing</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>metapth</td>
<td>Operational</td>
<td>External-Facing</td>
<td>Used for The Portal to Texas History</td>
</tr>
<tr>
<td>metadc</td>
<td>Operational</td>
<td>External-Facing</td>
<td>Used for the UNT Digital Library and The Gateway to Oklahoma History.</td>
</tr>
<tr>
<td>metarkv</td>
<td>Operational</td>
<td>External-Facing</td>
<td>Used for items that are “archive only” and do not rely on Aubrey for access. (Web Archives)</td>
</tr>
<tr>
<td>metacrs</td>
<td>Deprecated</td>
<td>External-Facing</td>
<td>Historically used for Congressional Research Service (CRS) Reports; discontinued use in 2007 in favor of metadc.</td>
</tr>
<tr>
<td>metatest</td>
<td>Operational</td>
<td>External-Facing</td>
<td>Testing namespace.</td>
</tr>
<tr>
<td>coda</td>
<td>Operational</td>
<td>Internal-Facing</td>
<td>Namespace used for the Coda repository system.</td>
</tr>
</tbody>
</table>

Table 1: Namespaces employed by the UNT Libraries for minting ARK identifiers
B2.6 If unique identifiers are associated with SIPs before ingest, the repository preserves the identifiers in a way that maintains a persistent association with the resultant archived object (e.g., AIP).

When digital objects are deposited into the UNT Libraries’ Digital Collections, the ingest process retains existing unique and other meaningful numbers associated with the SIPs. Since materials in the Digital Collections come from a number of sources, item identifiers are based on a wide range of policies, based on the partnering organizations or departments that contribute them. The UNT Libraries refer to these as “locally supplied identifiers” or “Local Control Numbers” and notate identifiers both in the descriptive metadata record and the bag-info.txt file in the BagIt bag used as the container for the UNT Libraries SIP and AIP.

To ensure that all items have identifiers unique to the Digital Collections, the UNT Libraries assign ARK identifiers for AIPs in the Coda repository and separate, connected ARKs for DIPs in the Aubrey access system. (See sections B2.5 and C1.3 for a further discussion of ARK identifiers in the Digital Collections.) Additionally, because the local control numbers are retained in both systems, the original identifiers can be used to locate items in both the Coda and Aubrey systems, though this method is less specific and may be less successful than the assigned ARKs.

B2.7 Repository demonstrates that it has access to necessary tools and resources to establish authoritative semantic or technical context of the digital objects it contains (i.e., access to appropriate international Representation Information and format registries).

Although the UNT Libraries will accept any file types for submission to the Digital Collections and are committed to bit-level preservation of all files, the Digital File Formats document (Appendix L) outlines three categories of responsibility:

- **Preferred File Formats** - Most common and generally most archival (least compressed) versions of various media types. The UNT Libraries are committed to bitstream preservation as well as functionality and content preservation.

- **Recognized File Formats** - Less preferred or less archival file types (also encompasses all file formats not listed elsewhere in the list). The UNT Libraries are committed to bit-level preservation without expectations of preserving long-term functionality.

- **Standard Normalizations** - For some proprietary formats (e.g., Microsoft Office file types), the UNT Libraries has standardized normalization paths to create additional derivative files (usually PDF and .jpg) in order to increase the likelihood of content and functionality preservation.
Based on the wide adoption of all file formats in the Digital File Formats list, the UNT Libraries expect ongoing support within the digital community and sufficient warning should any of them become less accepted. In the event that any of the identified format standards should become deprecated or obsolete, the UNT Libraries would work with the digital community to ensure that repository content and associated tools and software (for validation, rendering purposes, etc.) can be migrated safely to the new and broadly-accepted format standards.

When the Digital Projects Unit encounters file formats with which it has no previous experience, it utilizes a variety of resources to assist in the identification of the file formats and to gain necessary information related to the preservation and display of that file format, including:

- Sustainability of Digital Formats Planning for Library of Congress Collections
  http://www.digitalpreservation.gov/formats/

- Wikipedia - List of file formats

- UK National Archives - PRONOM
  https://www.nationalarchives.gov.uk/PRONOM/

- Archiveteam’s File Format Wiki
  http://fileformats.archiveteam.org/

**B2.8 Repository records/registers Representation Information (including formats) ingested.**

Currently the ingest workflow uses two tools to identify and report characteristics of files added to the Digital Collections: the JSTOR/Harvard Object Validation Environment (http://jhove.sourceforge.net/) and the standard Unix file command (https://en.wikipedia.org/wiki/File_(command)). Each file is processed with these tools and the output for each file is stored in the UNTL-METS-AIP record created for each AIP in the repository. This extracted information is additionally integrated into the PREMS Object Entity XML block also stored for each file in the UNTL-METS-AIP record and into the METS fileSec element for each file in the digital object. Information extracted with these tools includes the file format and mime type among other information. Currently the ingest workflow for the UNT Libraries’ Digital Collections does not formally record format information in local registries nor does it explicitly relate identified formats to other available format registries like PRONOM.

The Digital Projects Unit will investigate the integration of a format identification tool that supplies mappings between identified formats and international format registries. At this time staff members are evaluating two tools, FIDO (https://github.com/openpreserve/fido) and DROID.
(http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/), for integration into the existing file characterization workflows present for content that is ingested into the Digital Collections. The Digital Projects Unit will integrate this new format registry information into the PREMIS Object Entity XML block that is stored for each file inside the UNTL-METS-AIP record. It is expected that this work will be completed by December 2015 with subsequent integration into the ingest workflow in place no later than April 2016.

B2.9 Repository acquires preservation metadata (i.e., PDI) for its associated Content Information.

The ingest process for the Digital Collections involves the extraction and generation of a significant amount of preservation metadata and documentation important to the lifecycle of digital objects maintained by the UNT Libraries. This preservation metadata is recorded in two primary ways:

1. Embedded metadata is stored within the UNTL-AIP as a part of the UNTL-METS-AIP that conforms with the UNTL METS Archival Information Package Profile (http://www.loc.gov/standards/mets/profiles/00000045.xml)
2. Referenced metadata uses persistent links to associate digital objects with preservation events recorded by an external databases such as the PREMIS Event Service or the Digital Libraries’ Major Event Log.

The UNT Libraries follow suggested practices for recording PREMIS preservation metadata in METS records and locally makes use of the PREMIS Data Dictionary for Preservation Metadata V2 including the Version 2 XML scheme (http://www.loc.gov/standards/premis/v2/). Both embedded and referenced preservation metadata can be viewed in the Coda user interface, as part of the history of the digital object.

B2.10 Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability.

The Digital Projects Unit has a number of checks in place to help ensure that the quality of digital objects added to the Digital Collections is consistent with internal standards and local processes. Examples of these include:

File Formats Document (Appendix L): establishes levels of responsibility that the Digital Projects Unit accepts for various file formats to ensure the future usability of items in the Digital Collections.
Digitization Standards (Appendix K): provides guidance in digitization workflows to further ensure the quality of digital objects added to the Digital Collections. Digitized objects added to the Digital Collections must meet these minimum digital object standards.

Input Guidelines for Descriptive Metadata (https://www.library.unt.edu/digital-projects-unit/input-guidelines-descriptive-metadata): outlines the formatting and content standards expected for metadata records in the Digital Collections. Metadata records that do not meet these standards are modified through consultation with the originating contributing partner, or by taking responsibility within the department to improve records for consistency and compliance.

The UNT Libraries provides several methods (e.g., e-mail links and feedback forms in the user interface) for people to report problems with the Digital Collections interfaces including the usability and utility of the digital objects. Contact via any of these methods results in a response from a member of the Digital Projects Unit or User Interfaces Unit.

Additionally, the UNT Libraries maintains an active program of user experience research, including usability testing, to continuously improve access to materials in the repository.

In the event that resources added to the Digital Collections are no longer directly usable by its designated community, the UNT Libraries will work with a combination of local discipline experts at the UNT Libraries, across the UNT Community, and the greater digital preservation community to establish a usable form of the digital object or provide additional information to the designated community of users for the resources.

B2.11 Repository verifies each AIP for completeness and correctness at the point it is generated.

The UNT Libraries has a single unified workflow for the conversion of Submission Information Packages (SIPs) to Archival Information Packages (AIPs) and finally to a Dissemination Information Package (DIP) that is locally referred to as an Access Content Package (ACP). In order for a digital object to be accessioned into the Digital Collections it must successfully be converted from SIP to AIP to DIP/ACP. This conversion code is implemented in a series of Python scripts that verify the SIP, extract preservation information, and subsequently create Web derivatives for viewing. AIPs are added to the Coda repository and DIPs/ACPs are added to the Aubrey access system.

If a digital object is malformed and fails the conversion workflow, the item is moved to a quarantine area. Information logged during the failed conversion is stored with the invalid digital object where it can be investigated by the operator of the ingest process. All quarantined items are examined by the operator and either fixed and moved forward or removed from the workflow before the process is considered complete.
The ingest workflow used by the UNT Libraries can be tested or demonstrated on demand when new file formats or processing instructions are created by using a test_dropbox workflow. Digital objects processed using these test_dropbox workflows are complete, valid AIPs and DIPs/ACPs but they are not retained in the repository and are considered transient.

Also see Appendix M: UNT Libraries OAIS Information Package Specification for further information about the workflow for conversion of SIPs to AIPs and to DIPs/ACPs.

B2.12 Repository provides an independent mechanism for audit of the integrity of the repository collection/content.

In addition to documentation provided in B2.1 through B2.6, in the Fall of 2014 the Digital Projects Unit performed a full content audit verifying that each item accessible in the Aubrey access system has an associated AIP present in the Coda repository. UNT Libraries has a goal that this full audit of items will be conducted once every two years as an external validation of the workflows and processes for moving content through the ingest process and into the Coda repository and the Aubrey access system. (See also Appendix I: Timetable for Documentation Review.)

B2.13 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (AIP creation).

Once an item is added to the Coda repository, all subsequent events pertinent to the lifecycle of the digital resource are logged as PREMIS Event Entities using the UNT Libraries PREMIS Event Service. At present, the events that are captured include: ingest, fixity, replication, and migration. These PREMIS Event Entities and their associated PREMIS Agent Entities are associated with the digital objects they describe using its internal facing ARK identifier.

B3.1 Repository has documented preservation strategies.

The UNT Libraries has a Digital Preservation Policy Framework (Appendix H) that describes the expected process for identification and remittance strategy for preservation issues that may arise over time. Because of the relative stability of the preferred and recognized formats (see Appendix L) ingested thus far by the Digital Projects Unit, this process has only been tested in the abstract.
B3.2 Repository has mechanisms in place for monitoring and notification when Representation Information (including formats) approaches obsolescence or is no longer viable.

The UNT Libraries’ Digital Libraries Division is actively involved in and contributes to the wider conversation in the national and international digital preservation community and thus works to stay informed of changes to file formats that directly affect the UNT Libraries and the Digital Collections.

While most of the preferred formats specified by the Digital Projects Unit (see Appendix L) have been relatively stable for the past decade, there is always the possibility that these formats will begin to shift in their wide acceptability. The UNT Libraries rely on two methods to avoid undesirable effects of this shift:

- If one or all of the format standards currently identified as preferred should become deprecated or obsolete, there should be sufficient warning (due to their wide adoption) to ensure that repository content and associated tools and software -- for validation, rendering purposes, etc. -- can be migrated safely to the new standards.
- On a biennial basis, the Digital Projects Unit’s list of preferred formats for deposit (see Appendix L) will be verified, and changes or additions to this list will be discussed as part of the review process (see Appendix H: UNT Libraries Digital Preservation Policy Framework). Additionally, the list of preferred formats will be adjusted to include formats identified as preservation concerns by the UNT Libraries’ and the Digital Collection’s designated community.

See also section B2.7 for information on preferred primary content formats.

B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities.

The UNT Libraries Digital Preservation Policy Framework (Appendix H) states that the Digital Libraries Division will review and revise the policy once every two years. Information changed to provide continued relevance or to address other changes will be shared with UNT Libraries’ stakeholders for their acceptance. If changes to the policy require an amendment to existing partnership agreements or Memoranda of Understanding (MOUs) with external entities, all partners will be contacted in order to solicit input on the proposed changes.
B3.4 Repository can provide evidence of the effectiveness of its preservation planning.

The UNT Libraries can demonstrate that any object ingested into the Digital Collections is available and valid by showing the availability of the DIP/ACP via the Aubrey access system and the digital object’s presence as an AIP in the Coda repository. Furthermore, for any particular item, the UNT Libraries can display all PREMIS Event Entities associated with the object, acquire the Archival Information Package (AIP) for the object from Coda, and demonstrate that it is viable as an object by validating the package and by reprocessing it with the most current AIP-to-DIP/ACP conversion process.

In addition to the above on-demand demonstrations, the UNT Libraries conducted a full repository audit in 2014 to ensure that all items represented in the Aubrey access system had a valid registered AIP in the Coda repository. From the 750,000 items in that audit, 513 were never deposited into Coda from the legacy infrastructure, all but 41 of the 513 were located on previous storage arrays for the Digital Projects Unit, and the remaining 41 items were acquired from the original partner and ingested into the system. This audit is to be carried out once every two years (see Appendix I) and will next be conducted in Fall 2016.

B4.1 Repository employs documented preservation strategies

The UNT Libraries employs a number of policies that govern the overall digital preservation strategy of the Digital Collections and related services, including:

- Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C)
- Digital Preservation Policy Framework (Appendix H)
- UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)

Together these policies provide the framework for operations and describe the expected process for activities such as the identification and remittance strategies for preservation issues that may arise over time.

In addition to reliance on standard formats, and the format normalization of non-standard formats pre-ingest (see Appendix L), the UNT Libraries employ the following preservation strategies:

- Bit-stream copying
- Periodic refreshing of storage infrastructure
- Replication to multiple storage systems in multiple locations
- Encapsulation
- Normalization
Additionally, the UNT Libraries’ digital preservation infrastructure actively creates and logs preservation metadata about every file and digital object ingested into the Digital Collections. These logs are in the form of encapsulated PREMIS Object Entities in the UNTL-AIP-METS record and furthermore as PREMIS Event Entities stored in the PREMIS Event Service and the Digital Projects Major Event Log connected via persistent links to the digital objects. This ongoing commitment to the creation and management of preservation metadata about the digital objects in the repository demonstrates the implementation of workflows and local practices informed by the policy documents.

B4.2 Repository implements/responds to strategies for archival object (i.e., AIP) storage and migration.

The UNT Libraries store AIPs in the Coda repository using a standard POSIX file system. Each AIP is stored in a separate directory that includes all files and metadata associated with the digital object upon ingest. Each AIP contains technical and administrative metadata (e.g., MD5 checksums). Each object is described in an associated UNTL-AIP-METS record, also stored as part of the digital object in that directory (see sections B1.2 and B1.8 for more information about SIP and AIP packaging).

The UNT Libraries maintain uniformity of AIPs throughout the repository as an important strategy for storing and providing access to content in the repository, and to allow for easier migration of content should the need arise.

While objects added to the UNT Libraries' Digital Collections are not limited to a specified list of formats (see B2.7 for a discussion of accepted formats), the Digital Projects Unit maintains a list of preferred formats (see Appendix L) that it recommends to partners and local content providers interested in depositing items. Currently there have been no documented issues with those formats labeled “preferred.” The Digital Libraries Division also has a pre-ingest normalization strategy to convert non-preferred formats to preferred formats (see Appendix L), primarily used for specific collections including the UNT Scholarly Works Repository, UNT Data Repository, UNT Theses and Dissertations, and the Texas Digital Newspaper Program Collection.

In the event that a full migration away from an existing preferred format is required, the Coda repository has mechanisms to enable a format change. The workflow for this process includes exporting the original AIP from the Coda repository, carrying out the necessary transformations to the affected file formats, verifying a satisfactory transformation, and re-ingesting the newly-migrated digital object into the repository. The Coda repository would retain both the original version and the newly-migrated version of the digital object.
B4.3 Repository preserves the Content Information of archival objects (i.e., AIPs).

The Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C) and the UNT Libraries Digital Preservation Policy Framework (Appendix H) state that items once added to the collection will be considered permanent and all efforts will be made to keep the resources in perpetuity. The strategy to version resources in the Coda registry allows for revision of objects to be ingested into the system while still maintaining the originally-deposited copy of the object. This allows the UNT Libraries to ensure that resources, once added to the Coda repository, are not deleted.

The UNT Libraries can demonstrate the history of events associated with any of the AIPs or groups of AIPs in the Coda repository during their lifecycle within the system.

Also see sections B2.7 for the UNT Libraries’ level of responsibility for various file types and B3.4 for information on the UNT Libraries’ preservation planning.

B4.4 Repository actively monitors integrity of archival objects (i.e., AIPs).

The ingest process used by the UNT Libraries performs integrity checks on items throughout the process of adding them to the Coda repository and Aubrey access system. Once ingested into Coda, fixity checks are carried out when an item is replicated to a secondary or tertiary storage location, or during the migration from one storage infrastructure to another. There is a continuous fixity verification process that runs on the primary and secondary Coda repository instances that checks the validity of each object based on its conformance to the BagIt specification. All fixity checks are logged with the UNT Libraries PREMIS Event Service; the events are associated with individual digital resources by referencing the assigned internal-facing ARK identifiers. The continuous fixity verification service is scheduled to validate all digital objects in the Coda repository during a 365-day period.

B4.5 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Archival Storage).

Once an item is added to the Coda repository, all subsequent events relevant to the lifecycle of the digital resource are logged as PREMIS events modeled as PREMIS Event Entities and stored in the UNT Libraries PREMIS Event Service. At present, captured events include: ingest, fixity, replication, and migration. These PREMIS Event Entities and associated PREMIS Agent Entities are connected to the digital objects they describe using internal-facing ARK identifiers.
At a system-wide level (as opposed to object-level), the Digital Libraries' Major Event Log (http://digital2.library.unt.edu/major-event-log) creates PREMIS Event Entities to record events that affect multiple digital objects. Examples of these events include: dates associated with storage infrastructure migration or a change in specifications (such as the METS profile used by the UNT Libraries to serialize its Archival Information Packages).

B5.1 Repository articulates minimum metadata requirements to enable the designated community(ies) to discover and identify material of interest.

All resources added to the UNT Libraries' Digital Collections are discoverable via the three primary discovery systems built on top of the Aubrey access system: The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History. The Aubrey access system requires that each item be deposited with a UNTL metadata record that is valid against the UNTL XML schema at http://digital2.library.unt.edu/untl.xsd.

The University of North Texas Metadata Guidelines (http://www.library.unt.edu/digital-projects-unit/metadata) act as documentation for the input rules and guidelines for the UNTL metadata format.

The UNTL metadata format includes an inherent notion of a minimally-viable record that represents the most basic information required in order to classify a record as “complete.” When a record is indexed, the system calculates a completeness metric based on the existence of values in the eight required fields; this value is available to curators and metadata editors via the Aubrey Edit system. Further documentation about what constitutes a minimally-viable record is available on the University of North Texas Metadata Guidelines web pages http://www.library.unt.edu/digital-projects-unit/minimally-viable-records. The Digital Projects Unit supplements efforts made by contributing partners to elevate descriptive records for all resources in the Digital Collections to at least the level of a minimal record. The current number of record at or below the minimally-viable threshold is available via the Aubrey Edit system.

The UNT Libraries have a locally-created Python module used throughout its systems for processing these UNTL metadata records, called pyuntl. This package is available through the UNT Libraries GitHub account under a BSD Open Source License (https://github.com/unt-libraries/pyuntl/).

The UNT Libraries has a goal to continuously improve the descriptive metadata in the UNT Libraries' Digital Collections in order to allow the widest possible discovery and use of these resources. Active analysis of metadata records and their quality are presented as features in the Aubrey Edit system for metadata editors to use in identifying metadata records that need improvement over time. An example of this type of improvement is the use of the Extended Date Time Format (EDTF) (http://www.loc.gov/standards/datetime/) for date representations in
the descriptive metadata records. When a metadata record is indexed, a field in the index denotes if all dates in the record are valid EDTF date strings; if they are not valid, a metadata editor can quickly identify these records and remedy the situation by adjusting the date string so that it conforms to the EDTF specification.

Descriptive metadata in the Aubrey system is versioned during each save, so it is possible to track back a given change in a metadata record, what changed, and who was responsible for editing the record. These versioned records are also useful for understanding how the metadata in the repository changes over time, and for making decisions based on a knowledge of these changes in aggregate.

**B5.2 Repository captures or creates minimum descriptive metadata and ensures that it is associated with the archived object (i.e., AIP).**

The UNT Libraries work with content contributors to identify the easiest method for creating the required metadata to facilitate discovery and retrieval of their digital resources. Responsibility for metadata creation may rest solely with content contributors or with the UNT Libraries, or may be shared.

When descriptive metadata will be completed after ingest, the Digital Projects Unit staff members add collection, partner, resource type, and format information. During the digitization process, items are tracked using identifiers submitted by the partner institution as folder and file names for each object. Upon upload, the folder name (unique identifier) for the item is automatically generated as a local identifier as as a temporary title for the item in each individual template metadata record. This brief metadata allows the most basic access to the resource until additional metadata can be entered and the item can be made publicly available.

The UNTL metadata format includes an inherent notion of a minimally-viable record that represents the most basic information required in order to classify a record as “complete.” When a record is indexed, the system calculates a completeness metric based on the existence of values in the eight required fields. (Also see section B5.1 for further documentation about what constitutes a minimally-viable record.) The Digital Projects Unit supplements efforts made by contributing partners to elevate all resources in the Digital Collections to at least the level of a minimal record. The current number of record at or below the minimally-viable threshold is available via the Aubrey Edit system.

Items that are left in the system for extended periods of time (over 3 years) without complete metadata records are typically adopted by the UNT Libraries as “orphaned records”; staff members complete these records when resources are available.
B5.3 Repository can demonstrate that referential integrity is created between all archived objects (i.e., AIPs) and associated descriptive information.

The UNT Libraries uses the Archival Resource Key (ARK) specification as a component of its persistent and unique identifier implementation. A public-facing ARK identifier is created during the ingest process and digital objects are accessed using this ARK identifier in the Aubrey access systems. A separate, internal-facing ARK identifier is created on ingest into the Coda repository system. Linkage is maintained through the AIP in the Coda repository and the DIP in the Aubrey system by using these identifiers.

The Aubrey system maintains the descriptive metadata for each of the items in the Digital Collections. Items discovered using metadata or characteristics of the resource itself (such as full text) can be located in the Coda repository with the public-facing ARK identifier from the Aubrey system.

This linkage between the repository and access systems can be demonstrated on demand if requested.

Also see section B2.5 for further information about the ARK identifiers in place for the Digital Collections, section B2.6 for more information about unique identifiers assigned to SIPs prior to ingest, section B2.13 for information about PREMIS Event Entities associated with objects according to ARK identifiers, section B4.4 for further information about events (such as fixity checks) tied to digital objects according to ARK identifiers, and section C1.3 for a description of when different ARKs are assigned during the ingest process.

B5.4 Repository can demonstrate that referential integrity is maintained between all archived objects (i.e., AIPs) and associated descriptive information.

The UNT Libraries maintain referential integrity between all archived objects (AIPs) and associated descriptive metadata using the public-facing Archival Resource Key (ARK) assigned to each digital object on ingest into the system.

The ARK identifier for a digital object is not directly editable via the Aubrey Edit interface as a safeguard against accidental modification of this value. Furthermore, all changes to the metadata records via this system are versioned on edit, which allows staff members to investigate items if a resource’s referential integrity appears to have been intentionally or unintentionally modified.

The maintained linkage between the systems can be demonstrated on demand if requested.
This linkage between the repository and access systems can be demonstrated on demand if requested.

Also see section B2.5 for further information about the ARK identifiers in place for the Digital Collections, section B2.6 for more information about unique identifiers assigned to SIPs prior to ingest, section B2.13 for information about PREMIS Event Entities associated with objects according to ARK identifiers, section B4.4 for further information about events (such as fixity checks) tied to digital objects according to ARK identifiers, and section B5.3 for further information about the establishment of referential integrity.

**B6.1 Repository documents and communicates to its designated community(ies) what access and delivery options are available.**

The UNT Libraries publicly documents various policies important to the access, delivery, and long-term maintenance and preservation of digital resources including:

- Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C)
- UNT Libraries’ Digital Preservation Policy Framework (Appendix H)
- UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)

**B6.2 Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/depositors.**

All resources added to the UNT Libraries' Digital Collections are discoverable via the three primary discovery systems built on top of the Aubrey access system: The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History. The Aubrey system provides access to the DIPs/ACPs to end users.

If users require access to high-resolution versions of resources and permission has been granted by the content contributor or partner, the Digital Projects Unit has the ability to provide the high-resolution files to the requestor on behalf of the original partner.

Access and reuse policies for the Aubrey access systems are addressed in the UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J).

Usage information is calculated for the Aubrey access system at the item level and then aggregated to the collection and partner levels on a nightly basis. The unit of measurement is a “Use,” comprising all interactions with a digital object by a single IP address for a thirty-minute
period of time. These use statistics are available to contributing partners and the general public as a user-interface feature of the Aubrey system.

The Coda repository restricts access by IP address to those authorized to retrieve master files. Usage and access information for the Coda repository are logged as standard Web server logs and are processed on an as-needed basis with commodity statistics tools and packages.

**B6.3 Repository ensures that agreements applicable to access conditions are adhered to.**

Users may view or use items in the Digital Collections through the primary interfaces for the Aubrey access system. Access policies associated with these systems are governed by the Collection Development Policy for the UNT Libraries' Digital Collections (Appendix C), and the UNT Libraries' Digital Collections Usage and Feedback Policy (Appendix J). Further refinement of these general policies may be present in a formal Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, or other Memorandum of Understanding (MOU) between UNT and a partner.

Access to the Coda repository is restricted to IP-based usage from Digital Curators and associated staff in the Digital Libraries Division and the Office of External Relations.

The UNT Libraries can demonstrate how the different levels of access restriction are assigned to digital objects in the Aubrey access system and in the Coda repository, based upon agreements between content contributors and the UNT Libraries.

**B6.4 Repository has documented and implemented access policies (authorization rules, authentication requirements) consistent with deposit agreements for stored objects.**

Access definitions and policies are defined in the UNT Libraries' Digital Collections Usage and Feedback Policy (Appendix J).

Access restrictions for the UNT Libraries' Digital Collections are exercised by a combination of IP-based and Local Directory Authorization Protocol (LDAP) authentication that uses the UNT Active Directory system and the UNT EUID authentication mechanism.

The UNT Libraries can demonstrate the different levels of access restriction implemented on digital objects in the Aubrey access system and in the Coda repository, per established deposit agreements. Additionally, this allows for a count of objects that have each kind of restriction, as well a demonstration of what happens when an unauthorized or unauthenticated user attempts to access restricted items.
See section C3.3 for further information about staff roles and access privileges.

**B6.5 Repository access management system fully implements access policy.**

UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J) clearly outlines access and usage policies which govern the Digital Collections, including files in both the access system (Aubrey) and the archival repository (Coda). The UNT Libraries can provide examples of items under restricted access or demonstrate processes noted in the policy as needed.

Also see section B6.3 for information about agreements that govern access conditions and section B6.4 for information about how access restrictions are implemented when required by agreements.

**B6.6 Repository logs all access management failures, and staff review inappropriate “access denial” incidents.**

The Digital Collections provides unrestricted access to the majority of digital objects in the Aubrey access system. Because of this, only 1% of the items in the utilize access management features of the Aubrey access system.

The Digital Projects Unit is putting a mechanism into production for the Aubrey access system that will allow staff to monitor http server logs for inappropriate “access denial” incidents. Additionally, this logging mechanism will help to identify general access management failures that are not currently logged, as well as events that are logged but that are not currently aggregated in a usable or actionable way. This feature is expected to be in operation by December 2015.

**B6.7 Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is completed in relation to the request.**

The UNT Libraries has standard documentation that defines the AIP-to-DIP/ACP Conversion Workflow (Appendix W). This documentation identifies the process and mechanisms that a user or administrator goes through to acquire a complete DIP.

Once an AIP has been successfully transformed into a DIP/ACP, it is made available to the designated user via one of the three end user interfaces built on top of the Aubrey access system. Requests to one of these three user interfaces are processed and fulfilled by the Aubrey access system.
The UNT Libraries can demonstrate that the process of creating a DIP/ACP from an input AIP is completed in a consistent manner as defined in the AIP-to-DIP/ACP Conversion Workflow (Appendix W) documentation.

The UNT Libraries can demonstrate that the requests made by the Designated Community are fulfilled by the Aubrey access system with investigation of server log files and by real-time inspection of requests to the system.

**B6.8 Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is correct in relation to the request.**

The UNT Libraries can demonstrate that the process of creating a DIP/ACP from an input AIP is completed in a consistent manner, as defined in the AIP-to-DIP/ACP Conversion Workflow (Appendix W) documentation.

The UNT Libraries can additionally demonstrate that the digital object delivered as part of the Aubrey access system is the same item requested by the user who made the original request.

**B6.9 Repository demonstrates that all access requests result in a response of acceptance or rejection.**

The Aubrey access system that provides end-user access to the designated community for the Digital Collections processes each user request, which either succeeds or fails. Both success and failure of requests, as well as the amount of time elapsed during the request, is captured in log files created by the http server infrastructure that serves the Aubrey access system. These server logs are permanently retained (see Appendix J: UNT Libraries’ Digital Collections Usage and Feedback Policy). Information about successful requests -- including the elapsed time and the amount of time a user interacts with the Aubrey access system -- is also captured as part of the Google Analytics client-side logging code present in the different discovery systems built on top of Aubrey. This Google Analytics data is covered by the UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J) and is used to improve the overall services offered by the UNT Libraries.

The Digital Projects Unit is putting an additional logging mechanism into production for the Aubrey access system that will allow for the further monitoring of important events that result in the subsequent success or failure of a request to the Aubrey system. This logging mechanism is expected to be in operation by December 2015.
B6.10 Repository enables the dissemination of authentic copies of the original or objects traceable to originals.

The UNT Libraries maintains the authenticity of the deposited content throughout each digital resource’s lifecycle as represented in these documents:

- Ingest Workflow Diagram (Appendix U)
- UNT Libraries OAIS Information Package Specification (Appendix M)
- UNT Libraries SIP-to-AIP Conversion Workflow (Appendix V)
- UNT Libraries AIP-to-DIP/ACP Conversion Workflow (Appendix W)

Also see section B2.5 for a discussion of the process that the UNT Libraries use to mint and assign unique ARK identifiers for resources.

Policies that inform discussions related to authenticity of copies of digital object in the Digital Collections include:

- Collection Development Policy for the UNT Libraries’ Digital Collections (Appendix C)
- UNT Libraries’ Digital Preservation Policy Framework (Appendix H)
- UNT Libraries’ Digital Collections Usage and Feedback Policy (Appendix J)

The UNT Libraries can demonstrate the authenticity of DIP/ACPs created during the process of converting an AIP to a DIP/ACP in a repeatable fashion. This conversion -- as documented in UNT Libraries AIP-to-DIP/ACP Conversion Workflow (Appendix W) -- is expected to change over time as new technologies are introduced that influence the delivery of content via the Web.
Section C: Technologies, Technical Infrastructure, and Security

The Digital Collections are operated by the UNT Libraries, a part of the University of North Texas (UNT) and the UNT System. As a subsidiary of the UNT Libraries, there are a number of standard policies and procedures with which the Digital Collections comply, and structures which guide the development and deployment of Information Technology and Web-based technologies.

UNT System Policies and Regulations

- UNT System Administrative Policies site: http://www.untsystem.edu/policies-sysadmin.htm
  - Chapter 08 Information Technology - 08.100 Information Security Policy: http://www.untsystem.edu/pdfs/policies-admin/08.100/08.100_Information-Security-(00127965x1C146B).pdf
- UNT IT Shared Services Policy Homepage: https://itss.untsystem.edu/it-policies/policies-home
  - Chapter 06 Information Technology - 06.1000 Information Security: http://itss.untsystem.edu/sites/default/files/pdf/06.1000%20Information%20Security%20(00084028%29.PDF
- UNT IT Shared Services IT Standards: https://itss.untsystem.edu/it-policies/it-standards

University of North Texas Policies

- UNT Policy Homepage: http://policy.unt.edu/
  - Computer Use Policy (3.10): http://policy.unt.edu/policy/3-10
  - Information Resources Security Policy (3.6): http://policy.unt.edu/policy/3-6
  - Network Connections Policy (3.11): http://policy.unt.edu/policy/3-11
  - Web Publishing Policy (3.9): http://policy.unt.edu/policy/3-9
  - Web Accessibility Policy (5.1): http://policy.unt.edu/policydesc/web-accessibility-policy-5-1
While the Digital Collections are directly managed by the Digital Libraries Division, responsibility for the underlying technical infrastructure is shared with the UNT Libraries Facilities and Systems Division, requiring significant coordination. Within the Facilities and System Division, the Technology and Computer Operations (TACO) group is the primary partner in relation to computing, storage, and technology infrastructure.

**UNT Libraries Policies**

- UNT Libraries Policies Homepage: [http://www.library.unt.edu/policies/](http://www.library.unt.edu/policies/)

**C1.1 Repository functions on well-supported operating systems and other core infrastructural software.**

The Digital Libraries Division's Software Development Team works in close collaboration with the UNT Libraries Technology and Computer Operations (TACO) group to follow industry standard procedures for proper support of all hardware, software, and operating system aspects of repository systems.

**Repository Servers**

The primary server operating system is RedHat Enterprise Linux 6.x, and is kept on current long-term support releases with support from the vendor via a UNT campus site license. All production servers are either managed under RedHat Satellite or are being migrated to Satellite management, with an expected migration completion date of December 2016.

When possible, vmWare server clusters are used to increase uptime and reduce the possibility of service interruptions and data corruption.

The Digital Libraries Division and TACO collaborate on the documentation of servers and associated software involved in the infrastructure for the UNT Libraries' Digital Collections. This documentation includes server names, information about firewall settings, both blacklist and whitelists of IP addresses, operating system versions, standard software installed on each machine, and additional software required on a per-machine basis. This documentation is shared between the two units via an internal departmental wiki that is not available to the public.

**Digital Library Systems**
The core of the Digital Collections is built upon commonly-available open-source components that have been combined in a way that supports the needs of the designated community and provides a platform that allows for significant growth in a variety of key areas including the addition of partners, collections, content, and the increase in end-user access.

The majority of components for the Digital Collections are built using the Python programming language with the current production version being the Python 2.7 release. These components include Web-based applications and ingest workflows. Work is underway to migrate existing Python 2.7 code to the Python 3.4 release with a schedule of completion of April 2016. After migrating to Python 3.4, the new production environment will be transitioned to the more modern version of Python on the Python 3 branch.

Web-based services -- including the Coda repository, Aubrey access system, and the Aubrey Edit system -- are provided using the Django Web Framework that has been used successfully since 2008 to build services for the UNT Libraries’ Digital Collections. Additional components include: Apache Solr, an enterprise search server, used to provide full-text and metadata indexing for all items ingested into the repository; MySQL, used for relational database management system related activities in the digital library infrastructure; and an Apache Web server, used to provide content to end users.

The UNT Libraries adopted an Open Source Software Policy (Appendix Q) that helps to establish a mechanism for the UNT Libraries to contribute custom code and local configurations of standard software to the greater community. As existing software is reviewed, it is added to the UNT Librarie’s GitHub repository (https://github.com/unt-libraries/) using an Open Source License (BSD 3 Clause). Newly-created code is added to this site as it is available. Contributing this code is intended to allow others to learn from the work carried out at the UNT Libraries and also to provide a mechanism for the open vetting of the infrastructure used by the Digital Collections.

C1.2 Repository ensures that it has adequate hardware and software support for backup functionality sufficient for the repository’s services and for the data held, e.g., metadata associated with access controls, repository main content.

To address and fulfill hardware needs, all storage and tape infrastructure used by the UNT Libraries, and therefore the UNT Libraries’ Digital Collections, are supported on a five-year extended contract. This five-year rotation is described in more detail in Section C1.7. Software is updated with long-term support versions serving as the preferred versions for core infrastructure activities.

The technical infrastructure for the UNT Libraries has two strategies related to the backup of the resources held by the UNT Libraries’ Digital Collections.
Replication:
The rich digital master files held in the Coda repository in the form of an Archival Information Package (AIP) and the content access files stored as static resources in the Aubrey system that form the Dissemination Information Package (DIP) are actively replicated between two physically independent systems.

The Coda repository environment replicates AIP content packages between the primary datastore in the UNT Libraries’ Willis Library server room, and the UNT Discovery Park where the secondary copy of all content is stored. These replications are managed automatically by the Coda repository infrastructure and make use of ResourceSync to communicate information between the primary and secondary datastores to keep them in balance. This replication process includes full fixity checking of the entire AIP package; the resulting replication and fixity event are logged with the PREMIS Event Service in the Coda repository. In the event that one of the instances of the Coda repository becomes unavailable, the unaffected instance would be able to immediately take over the fulfillment of requests for items from the Coda repository.

Replication of the DIP as part of the Aubrey system’s static file storage is accomplished with traditional RSync technologies and replicates the end-user access copies between two physically separate storage systems in the Willis Library’s Server Room. Because of the design of the Aubrey system, there is immediate failover to these replicas if there is an issue with the primary version of the data.

Backup:
A traditional backup infrastructure based on the LTO5 tape backup system enables replication of a number of services that can be sent offsite, on a weekly basis. Nightly incremental backups are run and kept onsite, while weekly differential and full backups are sent to offsite secure storage in case of a disaster.

For the UNT Libraries’ Digital Collections, several sub-components fall into this traditional backup workflow. These sub-component systems include the descriptive metadata system used to power the Aubrey system, the MySQL database used by the Django Web framework, and other software that utilize SQL database and tables. All metadata in this system is versioned per change and is backed up as part of the nightly backup routine. The primary MySQL database used for the Aubrey system is also part of this nightly backup. The Subversion system responsible for locally managing the code repository for all software created or customized by the Digital Libraries Division is also included on the nightly backup schedule.
C1.3 Repository manages the number and location of copies of all digital objects.

The UNT Libraries actively maintain two copies of all resources added to both the Coda repository and the Aubrey access system.

Coda

The Coda repository is replicated to two physically-separate locations: one in the Willis Library Server Room and another in a facility at UNT Discovery Park, the university’s research campus, which is located four miles due north of the main campus.

When a digital object is ingested into the UNT Libraries’ Digital Collections, it is processed through a specific workflow that ensures tracking within the repository:

1. The Submission Information Package (SIP) is verified
2. An Archival Resource Key (ARK) is minted in an external-facing namespace for the SIP, and this ARK is used to manage the item throughout its lifecycle
3. An Archival Information Package (AIP) is created and registered with the Coda Repository
4. An additional ARK is minted in an internal-facing namespace to guarantee uniqueness within Coda and to track versioning
5. An ingest PREMIS event record is logged with the PREMIS Event Service upon successful ingest
6. After ingest, the digital object is replicated from the primary Coda instance to the secondary Coda instance located at the UNT Discovery Park.

If the original digital object needs to be replaced for some reason (e.g., missing images, or a need to redact personal information), a mechanism is in place to version the items in the Coda repository. In this process, previous copies are kept and a new internal-facing ARK is assigned to the most-current version of the files for the digital object. This means that there will be a growing difference in the number of resources in the Coda repository and the number of items in the Aubrey access system.

At any point it is possible to verify the number of digital objects in either the primary or secondary Coda repository.

Aubrey

The Aubrey access system is actively replicated to two identical storage arrays in the Willis Library Server Room, thus providing redundant active access to the Dissemination Information Packages (DIPs) in the Digital Collections. These are locally identified with the original external-facing ARK identifier assigned during the conversion from SIP to AIP in the ingest process. This ARK identifier is used to connect the static access files with the records in the
metadata storage system including descriptive metadata (UNTL metadata format) and the structural and administrative metadata (UNT METS format). If the AIP is needed from the Coda repository, it can be queried and retrieved via this ARK identifier. At any time, it is possible to show the number of digital objects in the Aubrey system.

For each Aubrey ARK, there is one or more Coda ARKs associated with it and this can be demonstrated on-demand. (See also section B2.5 for more information about the ARK namespaces used in Coda and Aubrey.)

**C1.4 Repository has mechanisms in place to ensure any/multiple copies of digital objects are synchronized.**

The UNT Libraries manage the replication of the Coda repository using ResourceSync and a locally-written replication queue process that is built into Coda. At least twice per week, the secondary Coda repository checks for items in the primary repository that are new or for which it does not have a record. When the software encounters new items, it replicates the content to its local instance, validates the BagIt Bag, and then registers a replication PREMIS Event with the primary Coda instance. Replication statistics are reported on a weekly basis to the Digital Libraries Division as part of the *Software Development Unit’s Weekly Update*. At any point a curator or manager may query the Coda repository to see if a digital object has been replicated, when it was replicated, and the final status of the replication (e.g., success or failure).

The Aubrey access system is kept up-to-date across the two content stores with standard ResourceSync technology, executed on a nightly basis. This process replicates additions to the access dataset that have occurred during the previous 24-hour period.

See section C1.3 for more information about the process for tracking and verifying objects in the Coda or Aubrey systems.

**C1.5 Repository has effective mechanisms to detect bit corruption or loss.**

The UNT Libraries verify the integrity of digital objects throughout the ingest process, including the creation of an AIP from a SIP and the subsequent loading of the AIP into the Coda repository. Once in Coda, valuation of a digital object is measured by the fact that the deposited BagIt bag is valid as defined by the BagIt specification ([https://tools.ietf.org/html/draft-kunze-bagit-06](https://tools.ietf.org/html/draft-kunze-bagit-06)). In this process the various BagIt tools used to validate content verify that the files contained within the bag’s `data` directory are present and valid. The UNT Libraries use MD5 Hashes as the cryptographic hash function to record and test against bit corruption or loss.

The Coda system has a module called “Validate” which is responsible for the continuous validation of digital objects in the repository. This module runs independently for both the
primary and secondary instances of the Coda repository and logs the resulting fixity information with the PREMIS Event Service in the appropriate Coda repository instance. The Validate service has reporting functionality to demonstrate the number of digital objects processed in the past day/week/month/year and gives an estimation of the amount of time required to process the entire repository as well as the percentage of the repository that has been processed in a designated period. The Validate module checks every digital object at least once per year.

The following appendices help to illustrate various aspects of validation in the Digital Collections:

Appendix M: UNT Libraries OAIS Information Package Specification
Appendix N: Coda Screenshots showing system dashboards, and views displaying example information and actions
Appendix U: Ingest Workflow Diagram
Appendix V: UNT Libraries SIP-to-AIP Conversion Workflow
Appendix W: UNT Libraries AIP-to-DIP/ACP Conversion Workflow

C1.6 Repository reports to its administration all incidents of data corruption or loss, and steps taken to repair/replace corrupt or lost data.

The Digital Libraries Division partners with the Facilities and Systems Division to ensure transparency in the event of data-loss or other events that result in the loss of access to resources in the Digital Collections.

The Digital Libraries Division’s Data Loss Escalation Procedures Document (Appendix R) discusses the procedures for alerting the UNT Libraries administration, affected partner institutions, and, depending on severity, the greater designated community, if the Digital Collections experience technological issues that cause data loss or restriction of access.

After data loss events are remedied, the events and their remediating steps are tracked in the Digital Libraries Division’s Major Event Log for future reference. This system creates a valid PREMIS event for each occurrence, which can be referenced by documentation as needed.

At this time there has not been any loss of data reported in the Coda repository system.

In 2013 a human-error incident in the Aubrey access system resulted in the accidental deletion of 7TB encompassing 140,000 digital objects. Due to further configuration errors with the automatic backup scripts, these deletions propagated through the replica storage system resulting in the complete loss of the DIPs. A notice to users was added to the access systems to alert end users, and staff followed steps outlined in the Digital Libraries Division’s Data Loss Escalation Procedures Document. 100% of data was recovered by reprocessing AIPs and
creating new DIPs that were subsequently added back into the Aubrey storage infrastructure. The reprocessing and reloading of the 140,000 digital objects took three weeks of machine time using two ingest servers. Changes were then made to better train technicians and correct issues with the automatic backup scripts so that this kind of loss would not happen again.

**C1.7 Repository has defined processes for storage media and/or hardware change (e.g., refreshing, migration).**

The UNT Libraries have a five-year hardware rotation schedule for all physical infrastructure components (servers and storage arrays) that serve the Digital Collections. Every four years the Digital Libraries Division coordinates with TACO to plan for the transition and to outline a working timeline of events. The last refresh of physical hardware for the Coda repository was in 2012, with the next scheduled refresh to occur in 2017. Planning for this event will begin in September 2016.

All server and storage hardware is purchased with a five-year extended support agreement including four-hour Pro Support on servers and Next Business Day Pro Support on storage.

All storage resides on a RAID6 or equivalent redundancy, and each storage array has 2-8 hot spare disks specified, depending on the physical disk count of each array. This is based on the current release of Dell PowerVault MD3260i Storage Arrays and MD3060e expansion cabinets with sixty 4TB drives. As technology changes, the Digital Libraries Division and the Facilities and Systems Division determine hardware needs based upon current technological demands.

As digital objects are migrated from a legacy hardware installation to a new hardware platform, migration events are simultaneously logged with the PREMIS Event Service as part of the primary Coda instance. This log documents the movement of the digital resources from the legacy hardware installation to the new platform. Upon successful migration of all resources to the new storage hardware platform, the Digital Libraries Division’s Major Event Log system logs an event and description for future reference.

Additional documentation related to the hardware and associated software for the server, storage, and network infrastructure is housed with the UNT Libraries Technology and Computer Operations (TACO) group and is available upon request.

**C1.8 Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository’s ability to comply with its mandatory responsibilities.**

The Digital Libraries Division maintains a Major Event Log system ([http://digital2.library.unt.edu/major-event-log/](http://digital2.library.unt.edu/major-event-log/))
that supports a public interface to document changes within the digital library infrastructure for the Digital Collections, including descriptions and dates of changes. This system enables the Digital Libraries Division to log events that may affect the digital objects stored within the Digital Collections and makes these events available in public display, utilizing PREMIS event terminology provided in the PREMIS Data Dictionary. These events are maintained and can be referenced in documentation and consulted as time passes from the date of the original event.

In addition to the Major Event Log, passive mechanisms are in place to document the changes within the underlying digital library infrastructure. Specifically, all code that manages these systems is versioned with either Subversion or Git and changes to the underlying codebase are recorded in the Digital Libraries Division’s Subversion repository or the UNT Libraries’ Github.com public repositories.

Each of the units of the Digital Libraries Division maintains documentation on an internal wiki that includes details of technology implementations, workflows, mitigation strategies, and training materials. These documents are updated regularly and are used to train new personnel.

**C1.9 Repository has a process for testing the effect of critical changes to the system.**

The UNT Libraries locally develops, deploys, and maintains the software infrastructure for digital preservation of and access to the UNT Libraries’ Digital Collections.

The Digital Libraries Division’s Software Development Team coordinates with TACO to isolate software development environments, thus supporting separate development, staging, and production instances of the Coda repository and the Aubrey access systems. Software that is locally developed by the Digital Libraries Division uses a Development, Staging, and Production (DSP) Model. Suggestions for system improvements and identified software bugs are submitted via a locally-maintained TRAC ticketing system (http://trac.edgewall.org/).

When developers complete features or correct software bugs, the software is moved to a staging instance and tested in a semi-production environment, which interacts with a read-only instance of the production environment. Once the software has cleared through test cases, it is deployed to the production environment. Software deployed to the production environment is actively monitored until the software development team lead and developers involved in the upgrade deem that no adverse effects are observable in the infrastructure.

Whenever software interacts with digital objects as part of the ingest workflow or in Coda, the version of the software responsible for the interaction is logged as part of a PREMIS Event registered with the PREMIS Event Service. This allows the Digital Libraries Division to track the effects of software on the digital objects and as needed investigate changes to the underlying software that have had adverse affects on the digital objects themselves. This information is
taken from the Subversion revision number, which is an explicitly-set version number that is incremented whenever changes occur in how the software processes a digital object.

**C1.10 Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment.**

The UNT Libraries operates within the overall organizational principles and policies of the University of North Texas computing infrastructure. As part of this larger service, the UNT Libraries have access to information about new vulnerabilities or threats that are available to the IT Shared Services Security Team.

Servers are managed with RedHat Satellite to enable patch application when security patches are released, as well as allowing an overview of all server resources and their current patch level.

When new security updates are available via RedHat Satellite, they are scheduled for installation by Lib-TACO in consultation with the Digital Libraries Division to ensure a balance between the highest level of server availability and security as possible. When the IT Shared Services Security Team, Lib-TACO, or the Digital Libraries Division identify critical security updates, these updates are quickly scheduled and deployed with priority on all servers that operate the Digital Collections.

**C2.1 Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed.**

All systems managed by the UNT Libraries are monitored by Cacti (http://www.cacti.net/) for general system health and longitudinal data collection, and the Digital Libraries Division also runs Nagios (https://www.nagios.org/) for more detailed health monitoring and alerting. These systems are restricted to authorized users on the Digital Projects Stats Server.

Environmental monitoring in each server room sends alerts to the Network Manager in the event of humidity or temperature alarms, along with other issues like fire or water intrusion.

The Digital Libraries Division and TACO engage in active dialog about the needs associated with the Digital Collections’ technical infrastructure, and this includes regular discussion about networking, storage, computing, and the underlying system software used by the system. Additionally, discussions about networking and computing infrastructure used by the UNT Libraries ensure sufficient capacity required for acquiring new digital content without negatively impacting other users of the UNT Libraries or its systems.
The UNT Libraries has established a five-year replacement timeline for digital library-related systems such as archival storage for the Digital Collections. (See section C1.7 for more details about the Libraries’ hardware replacement schedule.) The Digital Libraries Division and Lib-TACO work together to research new technologies that meet the needs of the digital library infrastructure. Research in this context takes the form of vendor demonstration, product reviews, feedback from colleagues in both digital library and information technology roles, and testing of new systems or architecture. Once adequate research has been conducted the Digital Libraries Division, and Lib-TACO work with the UNT Libraries Administrative Offices to purchase the new servers, storage and networking equipment.

**C2.2 Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed.**

The UNT Libraries has worked with local constituents and partner institutions to identify and develop a set of services to meet local and external (partner) needs. These services include the interfaces and capabilities of the Aubrey access system, and the formats and features related to working with the digital object in the system. With this information about partner needs, the UNT Libraries also regularly work to develop tools and services that meet the widest range of needs through a combination of user interfaces, application programming interfaces (APIs) and documentation.

The current interface for The Portal to Texas History was developed as part of a federally-funded National Leadership Research Grant from the Institute of Museum and Library Services (IMLS) and was designed around a subset of users identifying themselves as genealogists. Lessons learned from these formal user experience sessions and user-centered design processes were also incorporated into the design and layout of the UNT Digital Library and The Gateway to Oklahoma History. As changes are applied to the public facing interfaces, both the public users as well as partner institutions are consulted with the goal of presenting the most useful system possible.

The UNT Libraries is also cognizant that many users are now interested in interfacing with digital library and repository systems using Application Programming Interfaces (API) and as a result continue to add new APIs to public-facing interfaces in addition to library-standard interfaces such as OAI-PMH, SRU, and more widely-adopted technologies, such as Atom Feeds and OpenSearch Interface. These APIs are available for specific subsets of materials designated as collections or belonging to a particular partner, as well as the entire holdings for each of the access interfaces: The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History.
The Digital Libraries Division’s User Interfaces Unit has participated in the Certified User Experience Professional Training developed and delivered by Texas Tech University's Technical Communication and Rhetoric Program in the Department of English. Their skills are used for major user interface changes for all of the Aubrey systems.

C3.1 Repository maintains a systematic analysis of such factors as data, systems, personnel, physical plant, and security needs.

In 2015 the Digital Libraries Division conducted a risk analysis related to data, systems, personnel, physical plant, and security needs and drafted the “UNT Digital Libraries: Risk Analysis and Management Strategy Plan” (Appendix S) as a way of documenting these factors that could directly or indirectly affect the Digital Collections.

This plan describes threats that could disrupt normal operations of the UNT Libraries' Digital Collections and prevent the repository from fulfilling its mission for short or long periods of time. This document details risk-minimization strategies, system infrastructures, operating practices, and business relationships that the UNT Libraries employs to reduce the potential impact of these threats.

The Digital Libraries Division used the framework established by the Canada Based Scholars Portal to guide the creation of this document. The original documentation is available on the Scholar’s Portal documentation site (https://spotdocs.scholarsportal.info/display/OAIS/Risk+Analysis+and+Management+Strategies).

The "UNT Digital Libraries: Risk Analysis and Management Strategy Plan" will be reviewed and revised once every three years by the Digital Libraries Division to update known risks, or to add newly-identified risks. This review schedule is documented in Timetable for Documentation Review (Appendix I).

C3.2 Repository has implemented controls to adequately address each of the defined security needs.

The UNT Libraries follows widely accepted practices regarding system security including firewalls, patching, and limited administrative access. The University of North Texas security procedures and activities are listed in the UNT System Information Security Handbook and the UNT IT Shared Services Policy Handbook.

The UNT Libraries use various software and other measures to monitor and protect Digital Collection components:

- **Server management:**
  - All servers in the UNT Libraries are managed centrally by RedHat Satellite.
○ All systems managed by the UNT Libraries are monitored by Cacti for general system health and longitudinal data collection.
○ The Digital Libraries Division also runs Nagios for more detailed health monitoring and alerting for repository systems.

● **Workstation software:**
  ○ All workstations in the UNT Libraries are updated centrally with Microsoft System Center Configuration Manager and Windows Server Update Services to ensure current software.
  ○ All workstations in the UNT Libraries are monitored with McAfee VirusScan Enterprise & ePolicy Orchestrator. Management infrastructure for this is supplied and monitored by UNT IT Shared Services and is also monitored within the Libraries.
  ○ All workstations in the Libraries also have the most restrictive firewall settings possible while still allowing day-to-day work.
  ○ Library workstations require valid UNT employee credentials and have different access levels based on the user’s account.

● **Physical server security:**
  ○ Environmental monitoring is accomplished in each server room sends alerts to the Network Manager in the event of humidity or temperature alarms, and other issues such as fire or water.
  ○ Personnel are only provided physical access to the server room in the main library after a review by the UNT Libraries’ Network Manager.
  ○ Security at the remote server location at the UNT Discovery Park Data Center is controlled by UNT System ITSS and only allows access by appointment, for authorized personnel designated by the UNT Libraries’ Network Manager.

See section C3.3 for additional information about restrictions on making changes within the repository.

### C3.3 Repository staff have delineated roles, responsibilities, and authorizations related to implementing changes within the system.

Authorization to the UNT Libraries’ Digital Collections content and servers is administered at four different levels:

1. Ability to authorize access to repository file system
2. Read/Write Access to full repository
3. Read Access to full repository (no write permissions)
4. Advanced Content Editing (Read/Write Access to Web Interfaces) on Coda/Aubrey Systems

This access hierarchy is intended to minimize or prevent the threat of risks.

<table>
<thead>
<tr>
<th>Person/People</th>
<th>Level of Repository Access</th>
<th>Authorized By</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Administrator</td>
<td>Ability to authorize access to repository file system</td>
<td>Network Administrator</td>
</tr>
<tr>
<td>System Administrators</td>
<td>Read/Write Access to full repository</td>
<td>Network Administrator</td>
</tr>
<tr>
<td>Core Developers</td>
<td>Read/Write Access to full repository</td>
<td>Network Administrator</td>
</tr>
<tr>
<td>Broader Developers (e.g., interface developers; firewall setup)</td>
<td>Read Access to full repository</td>
<td>Core Developers &amp; System Administrator</td>
</tr>
<tr>
<td>Digital Curators/Content Managers</td>
<td>Advanced Content Editing (Read/Write Access to Web Interfaces) on Coda/Aubrey Systems</td>
<td>Core Developers</td>
</tr>
<tr>
<td>Metadata Editors</td>
<td>Advanced Content Editing (Read/Write Access to Web Interfaces) on Coda/Aubrey Systems</td>
<td>Core Developers</td>
</tr>
<tr>
<td>Feedback Reviewers &amp; Reference Personnel</td>
<td>Advanced Content Editing (Read/Write Access to Web Interfaces) on Coda/Aubrey Systems</td>
<td>Core Developers</td>
</tr>
</tbody>
</table>

Requests for access according to one of the four levels defined above come from the Digital Libraries Division. They are reviewed individually by the UNT Libraries’ Network Manager.
C3.4 Repository has suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s).

Disaster Preparedness and Recovery Plans

Note: At this time there is no formal Disaster Recovery Plan that directly covers the UNT Libraries Technology Infrastructure used by the UNT Libraries’ Digital Collections.

In 2015 the Digital Libraries Division developed a Risk Analysis and Management Strategy Plan (Appendix S).

The UNT Libraries will develop and enact a formal disaster recovery and preparedness plan based on the Risk Analysis and Management Strategy Plan during the 2016 fiscal year (Sept. 2015-Aug. 2016).

The Digital Collections use technologies and infrastructure shared at different levels between the Digital Libraries Division, UNT Libraries (managed and operated by the Facilities and Systems Division), the UNT IT Shared Services, and the UNT System. Due to this, a number of policies and procedures are in place that directly or indirectly govern the operation of technology and physical infrastructure in normal situations and in response to physical or manmade disasters. Relevant documents to this section include:

UNT System / IT Shared Services

ITSS Disaster Recover and Business Continuity Planning: https://itss.untsystem.edu/security/disaster-recovery-business-continuity-planning

ITSS Services Catalog: https://itss.untsystem.edu/services

UNT Libraries / Facilities and Systems Division

LIB-TACO Standard Practices and Methodologies
Internal documentation (Available upon Request)

LIB-TACO UNT Libraries Technology Services Catalog
Internal documentation (Available upon Request)

UNT Libraries Digital Libraries Division
Redundancy and off-site storage of data
The UNT Libraries' Digital Collections operates on storage and computing infrastructure managed and operated by the Facilities and Systems Division in collaboration with the Digital Libraries Division.

Repository infrastructure ensures that two copies of all Archival Information Packages (AIPs) are ingested into the system. Each copy is in a physically-separate location with the primary data storage facility residing in the Willis Library Server Room, and the secondary data storage facility at the UNT Discovery Park Data Center. Active monitoring of digital objects ingested into the system is performed by Coda to ensure that digital objects are replicated between the two systems. See section C1.2 for a more detailed description of replication and backup measures.
Appendices

Appendix A: UNT Libraries TRAC Audit Checklist
Appendix B: Formal Statement of Conformance to ISO 14721:2012
Appendix C: Collection Development Policy for the UNT Libraries’ Digital Collections
Appendix D: The Portal Partnership Agreement
Appendix E: Portal Memorandum of Agreement for Digital Rights
Appendix F: Digital Library Memorandum of Agreement for Digital Rights
Appendix G: Travel/Professional Development Document
Appendix H: UNT Libraries’ Digital Preservation Policy Framework
Appendix I: Timetable for Documentation Review
Appendix J: UNT Libraries’ Digital Collections Usage and Feedback Policy
Appendix K: Digital Projects Unit Standards
Appendix L: Preferred File Formats
Appendix M: UNT Libraries OAIS Information Package Specification
Appendix N: CODA Screenshots
Appendix O: UNT Libraries Archival Storage Replacement Fund
Appendix P: UNT Libraries’ Policy on Creation and Licensing of Metadata
Appendix Q: UNT Libraries Open Source Software Policy
Appendix R: Data Loss Escalation Procedures
Appendix S: UNT Digital Libraries: Risk Analysis and Management Strategy Plan
Appendix T: UNT Libraries Digital Workflow Diagram
Appendix U: UNT Libraries Ingest Workflow Diagram
Appendix V: UNT Libraries SIP-to-AIP Conversion Workflow
Appendix W: UNT Libraries AIP-to-DIP/ACP Conversion Workflow
UNT Libraries TRAC Audit Checklist

Date: October 2015

Version: 1.0

Contributors:

- Mark Phillips Assistant Dean for Digital Libraries
- Daniel Alemneh Supervisor, Digital Curation Unit
- Ana Krahmer Supervisor, Digital Newspaper Unit
- Hannah Tarver Department Head, Digital Projects Unit
- Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Example Evidence</th>
<th>Response</th>
<th>Documentation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section A. Organizational Infrastructure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A1. Governance &amp; organizational viability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.</td>
<td>Mission statement for the repository; mission statement for the organizational context in which the repository sits; legal or legislative mandate; regulatory requirements.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.1.1</td>
<td></td>
</tr>
<tr>
<td>A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.</td>
<td>Succession plan(s); escrow plans; explicit and specific statement documenting the intent to ensure continuity of the repository, and the steps taken and to be taken to ensure continuity; formal documents describing exit strategies and contingency plans; depositor agreements.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.1.2</td>
<td></td>
</tr>
<tr>
<td><strong>A2. Organizational structure &amp; staffing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2.1 Repository has identified and established the duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties.</td>
<td>A staffing plan; competency definitions; job description; development plans; plus evidence that the repository review and maintains these documents as requirements evolve.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.2.1</td>
<td></td>
</tr>
<tr>
<td>A2.2 Repository has the appropriate number of staff to support all functions and services.</td>
<td>Organizational charts; definitions of roles and responsibilities; comparison of staffing levels to commitments and estimates of required effort.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.2.2</td>
<td></td>
</tr>
<tr>
<td>A2.3 Repository has an active professional development program in place that provides staff with skills and expertise development opportunities.</td>
<td>Professional development plans and reports; training requirements and training budgets, documentation of training expenditures (amount per staff); performance goals and documentation of staff assignments and achievements, copies of certificates awarded.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.2.3</td>
<td></td>
</tr>
<tr>
<td><strong>A3. Procedural accountability &amp; policy framework (documentation)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3.1 Repository has defined its designated community(ies) and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation service requirements will be met.</td>
<td>Mission statement; written definitions of the designated community(ies); documented policies; service-level agreements.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.1</td>
<td></td>
</tr>
<tr>
<td>A3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolves.</td>
<td>Written documentation in the form of policies, procedures, protocols, rules, manuals, handbooks, and workflows; specification of review cycle for documentation; documentation detailing review, update, and development mechanisms. If documentation is embedded in system logic, functionality should demonstrate the implementation of policies and procedures.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.2</td>
<td></td>
</tr>
<tr>
<td>A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed.</td>
<td>Deposit agreements; records schedule; digital preservation policies; records legislation and policies; service agreements.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.3</td>
<td></td>
</tr>
<tr>
<td>A3.4 Repository is committed to formal, periodic review and assessment to ensure responsiveness to technological developments and evolving requirements.</td>
<td>A self-assessment schedule, timetables for review and certification; results of self-assessment; evidence of implementation of review outcomes.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.4</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time.</td>
<td>A policy that requires a feedback mechanism; a procedure that addresses how the repository seeks, captures, and documents responses to feedback; documentation of workflow for feedback (i.e., how feedback is used and managed); quality assurance records.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.5</td>
<td></td>
</tr>
<tr>
<td>A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content.</td>
<td>Policies, procedures, and results of changes that affect all levels of the repository: objects, aggregations of objects; object-level preservation metadata; repository’s records retention strategy document.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.6</td>
<td></td>
</tr>
<tr>
<td>A3.7 Repository commits to transparency and accountability in all actions supporting the operation and management of the repository, especially those that affect the preservation of digital content over time.</td>
<td>Comprehensive documentation that is readily accessible to stakeholders; unhindered access to content and associated information within repository.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.7</td>
<td></td>
</tr>
<tr>
<td>A3.8 Repository commits to defining, collecting, tracking, and providing, on demand, its information integrity measurements.</td>
<td>An implemented registry system; a definition of the repository’s integrity measurements; documentation of the procedures and mechanisms for integrity measurements; an audit system for collecting, tracking, and presenting integrity measurements; procedures for responding to results of integrity measurements that indicate digital content is at risk; policy and workflow documentation.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.8</td>
<td></td>
</tr>
<tr>
<td>A3.9 Repository commits to a regular schedule of self-assessment and certification and, if certified, commits to notifying certifying bodies of operational changes that will change or nullify its certification status.</td>
<td>Completed, dated audit checklists from self-assessment or objective audit; certificates awarded for certification; presence in a certification register (when available); timetable or budget allocation for future certification.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.3.9</td>
<td></td>
</tr>
<tr>
<td>A4.1 Repository has short- and long-term business planning processes in place to sustain the repository over time.</td>
<td>Operating plans; financial reports; budgets; financial audit reports; annual financial reports; financial forecasts; business plans; audit procedures and calendars; evidence of comparable institutions; exposure of business plan to scenarios.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.4.1</td>
<td></td>
</tr>
<tr>
<td>A4.2 Repository has in place processes to review and adjust business plans at least annually.</td>
<td>Business plans, audit planning (e.g., scope, schedule, process, and requirements) and results; financial forecasts; recent audits and evidence of impact on repository operating procedures.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.4.2</td>
<td></td>
</tr>
<tr>
<td>A4.3 Repository’s financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.</td>
<td>Demonstrated dissemination requirements for business planning and practices; citations to and/or examples of accounting and audit requirements, standards, and practice; evidence of financial audits already taking place.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.4.3</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>A4.4 Repository has ongoing commitment to analyze and report on risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).</td>
<td>Risk management documents that identify perceived and potential threats and planned or implemented responses (a risk register); technology infrastructure investment planning documents; cost benefit analyses; financial investment documents and portfolios; requirements for and examples of licenses, contracts, and asset management; evidence of revision based on risk.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.4.4</td>
<td></td>
</tr>
<tr>
<td>A4.5 Repository commits to monitoring for and bridging gaps in funding.</td>
<td>Fiscal and fiduciary policies, procedures, protocols, requirements; budgets and financial analysis documents; fiscal calendars; business plan(s); any evidence of active monitoring and preparedness.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.4.5</td>
<td></td>
</tr>
<tr>
<td>A5. Contracts, licenses, &amp; liabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5.1 If repository manages, preserves, and/or provides access to digital materials on behalf of another organization, it has and maintains appropriate contracts or deposit agreements.</td>
<td>Deposit agreements; policies on third-party deposit arrangements; contracts; definitions of service levels; Web archiving policies; procedure for reviewing and maintaining agreements, contracts, and licenses.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.5.1</td>
<td></td>
</tr>
<tr>
<td>A5.2 Repository contracts or deposit agreements must specify and transfer all necessary preservation rights, and those rights transferred must be documented.</td>
<td>Contracts, deposit agreements; specification(s) of rights transferred for different types of digital content (if applicable); policy statement on requisite preservation rights.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.5.2</td>
<td></td>
</tr>
<tr>
<td>A5.3 Repository has specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties.</td>
<td>Submission agreements/deposit agreements/deeds of gift; written standard operating procedures.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.5.3</td>
<td></td>
</tr>
<tr>
<td>A5.4 Repository tracks and manages intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.</td>
<td>A policy statement that defines and specifies the repository’s requirements and process for managing intellectual property rights; depositor agreements; samples of agreements and other documents that specify and address intellectual property rights; demonstrable way to monitor intellectual property; results from monitoring.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.5.4</td>
<td></td>
</tr>
<tr>
<td>A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights.</td>
<td>A definition of rights; citations for relevant laws and requirements; policy on responding to challenges; documented track record for responding to challenges in ways that do not inhibit preservation; examples of legal advice sought and received.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document A.5.5</td>
<td></td>
</tr>
<tr>
<td>Metric</td>
<td>Example Evidence</td>
<td>Response</td>
<td>Documentation</td>
<td>Notes</td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>----------</td>
<td>---------------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>B. Digital Object Management</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1.1 Repository identifies properties it will preserve for digital objects.</td>
<td>Evidence: Mission statement; submission agreements/deposit agreements/deeds of gift; workflow and policy documents, including written definition of properties as agreed in the deposit agreement/deed of gift; written processing procedures; documentation of properties to be preserved.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.1</td>
<td></td>
</tr>
<tr>
<td>B1.2 Repository clearly specifies the information that needs to be associated with digital material at the time of its deposit (i.e., SIP).</td>
<td>Evidence: Transfer requirements; producer-archive agreements.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.2</td>
<td></td>
</tr>
<tr>
<td>B1.3 Repository has mechanisms to authenticate the source of all materials.</td>
<td>Evidence: Submission agreements/deposit agreements/deeds of gift; workflow documents; evidence of appropriate technological measures; logs from procedures and authentications.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.3</td>
<td></td>
</tr>
<tr>
<td>B1.4 Repository’s ingest process verifies each submitted object (i.e., SIP) for completeness and correctness as specified in B1.2.</td>
<td>Evidence: Appropriate policy documents and system log files from system performing ingest procedure; formal or informal “acquisitions register” of files received during the transfer and ingest process; workflow, documentation of standard operating procedures, detailed procedures; definition of completeness and correctness, probably incorporated in policy documents.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.4</td>
<td></td>
</tr>
<tr>
<td>B1.5 Repository obtains sufficient physical control over the digital objects to preserve them.</td>
<td>Evidence: Submission agreements/deposit agreements/deeds of gift; workflow documents; system log files from the system performing ingest procedures; logs of files captured during Web harvesting.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.5</td>
<td></td>
</tr>
<tr>
<td>B1.6 Repository provides producer/depositor with appropriate responses at predefined points during the ingest processes.</td>
<td>Evidence: Submission agreements/deposit agreements/deeds of gift; workflow documentation; standard operating procedures; evidence of “reporting back.”</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.6</td>
<td></td>
</tr>
<tr>
<td>B1.7 Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e., SIPs).</td>
<td>Evidence: Submission agreements/deposit agreements/deeds of gift; confirmation receipt sent back to producer.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.7</td>
<td></td>
</tr>
<tr>
<td>B1.8 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition).</td>
<td>Evidence: Written documentation of decisions and/or action taken; preservation metadata logged, stored, and linked to pertinent digital objects.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B1.8</td>
<td></td>
</tr>
<tr>
<td><strong>B2. Ingest: creation of the archival package</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2.1 Repository has an identifiable, written definition for each AIP or class of information preserved by the repository.</td>
<td>Evidence: Documentation identifying each class of AIP and describing how each is implemented within the repository. Implementations may, for example, involve some combination of files, databases, and/or documents.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.1</td>
<td></td>
</tr>
<tr>
<td>B2.2 Repository has a definition of each AIP (or class) that is adequate to fit long-term preservation needs.</td>
<td>Evidence: Documentation that relates the AIP component’s contents to the related preservation needs of the repository, with enough detail for the repository's providers and consumers to be confident that the significant properties of AIPs will be preserved.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.2</td>
<td></td>
</tr>
<tr>
<td>B2.3 Repository has a description of how AIPs are constructed from SIPs.</td>
<td>Evidence: Process description documents; documentation of SIP relationship to AIP; clear documentation of how AIPs are derived from SIPs; documentation of standard/process against which normalization occurs; documentation of normalization outcome and how outcome is different from SIP.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.3</td>
<td></td>
</tr>
<tr>
<td>B2.4 Repository can demonstrate that all submitted objects (i.e., SIPs) are either accepted as whole or part of an eventual archival object (i.e., AIP), or otherwise disposed of in a recorded fashion.</td>
<td>Evidence: System processing files; disposal records; donor or depositor agreements/deeds of gift; provenance tracking system; system log files.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.4</td>
<td></td>
</tr>
<tr>
<td>B2.5 Repository has and uses a naming convention that generates visible, persistent, unique identifiers for all archived objects (i.e., AIPs).</td>
<td>Evidence: Documentation describing naming convention and physical evidence of its application (e.g., logs).</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.5</td>
<td></td>
</tr>
<tr>
<td>B2.6 If unique identifiers are associated with SIPs before ingest, the repository preserves the identifiers in a way that maintains a persistent association with the resultant archived object (e.g., AIP).</td>
<td>Evidence: Workflow documents and evidence of traceability (e.g., SIP identifier embedded in AIP, mapping table of SIP IDs to AIPs).</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.6</td>
<td></td>
</tr>
<tr>
<td>B2.7 Repository demonstrates that it has access to necessary tools and resources to establish authoritative Representation Information of the digital objects it contains.</td>
<td>Evidence: Subscriptions or access to such registries; association of unique identifiers to registries of Representation Information (including format registries); Viewable records in local registries (with persistent links to digital objects); database records that include Representation Information and a persistent link to relevant digital objects.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.7</td>
<td></td>
</tr>
<tr>
<td>B2.8 Repository records/registers Representation Information (including formats) ingested.</td>
<td>Evidence: Viewable records in local format registry (with persistent links to digital objects); local metadata registry(ies); database records that include Representation Information and a persistent link to relevant digital objects.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.8</td>
<td></td>
</tr>
<tr>
<td>B2.9 Repository has documented processes for acquiring preservation metadata (i.e., PDI) for its associated Content Information and acquires preservation metadata in accordance with the documented processes. The repository must maintain viewable documentation on how the repository acquires and manages Preservation Description Information (PDI).</td>
<td>Evidence: Viewable documentation on how the repository acquires and manages Preservation Description Information (PDI).</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.9</td>
<td></td>
</tr>
<tr>
<td>B2.10 Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability.</td>
<td>Evidence: Retention of individuals with the discipline expertise; periodic assembly of designated or outside community members to evaluate and identify additional required metadata.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.10</td>
<td></td>
</tr>
<tr>
<td>B2.11 Repository verifies each AIP for completeness and correctness at the point it is generated.</td>
<td>Evidence: Description of the procedure that verifies completeness and correctness; logs of the procedure.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.11</td>
<td></td>
</tr>
<tr>
<td>B2.12 Repository provides an independent mechanism for audit of the integrity of the repository collection/content.</td>
<td>Evidence: Documentation provided for B2.1 through B2.6; documented agreements negotiated between the producer and the repository (see B 1.1-B1.9); logs of material received and associated action (receipt, action, etc.) dates; logs of periodic checks.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.12</td>
<td></td>
</tr>
<tr>
<td>B2.13 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (AIP creation).</td>
<td>Evidence: Written documentation of decisions and/or action taken; preservation metadata logged, stored, and linked to pertinent digital objects.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document B.2.13</td>
<td></td>
</tr>
</tbody>
</table>
## B3. Preservation planning

**B3.1 Repository has documented preservation strategies.**
- Evidence: Documentation identifying each preservation issue and the strategy for dealing with that issue.
- Conforms See TRAC Conformance Document B.3.1

**B3.2 Repository has mechanisms in place for monitoring and notification when Representation Information (including formats) approaches obsolescence or is no longer viable.**
- Evidence: Subscription to a format registry service; subscription to a technology watch service; percentage of at least one staff member dedicated to monitoring technological obsolescence issues.
- Conforms See TRAC Conformance Document B.3.2

**B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities.**
- Evidence: Preservation planning policies tied to formal or information technology watch(es); preservation planning or processes that are timed to shorter intervals (e.g., not more than five years); proof of frequent preservation planning/policy updates.
- Conforms See TRAC Conformance Document B.3.3

**B3.4 Repository can provide evidence of the effectiveness of its preservation planning.**
- Evidence: Collection of appropriate preservation metadata; proof of usability of randomly selected digital objects held within the system; demonstrable track record for retaining usable digital objects over time.
- Conforms See TRAC Conformance Document B.3.4

---

## B4. Preservation Planning - implementing the planning in storage and maintaining digital objects Archival storage & preservation/maintenance of AIPs

**B4.1 Repository employs documented preservation strategies.**
- Evidence: Documentation of strategies and their appropriateness to repository objects; evidence of application (e.g., in preservation metadata); see B3.3.
- Conforms See TRAC Conformance Document B.4.1

**B4.2 Repository implements/responds to strategies for archival object (i.e., AIP) storage and migration.**
- Evidence: Institutional technology and standards watch; demonstration of objects on which a preservation strategy has been performed; demonstration of appropriate preservation metadata for digital objects.
- Conforms See TRAC Conformance Document B.4.2

**B4.3 Repository preserves the Content Information of archival objects (i.e., AIPs).**
- Evidence: Policy documents specifying treatment of AIPs and whether they may ever be deleted; ability to demonstrate the chain of AIPs for any particular digital object or group of objects ingested; workflow procedure documentation.
- Conforms See TRAC Conformance Document B.4.3

**B4.4 Repository actively monitors integrity of archival objects (i.e., AIPs).**
- Evidence: Logs of fixity checks (e.g., checksums); documentation of how AIPs and Fixity information are kept separate.
- Conforms See TRAC Conformance Document B.4.4

**B4.5 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Archival Storage).**
- Evidence: Written documentation of decisions and/or action taken; preservation metadata logged, stored, and linked to pertinent digital objects.
- Conforms See TRAC Conformance Document B.4.5

---

## B5. Information management

**B5.1 Repository articulates minimum metadata requirements to enable the designated community(ies) to discover and identify material of interest.**
- Evidence: Descriptive metadata.
- Conforms See TRAC Conformance Document B.5.1

**B5.2 Repository captures or creates minimum descriptive metadata and ensures that it is associated with the archived object (i.e., AIP).**
- Evidence: Descriptive metadata; persistent identifier/locator associated with AIP; system documentation and technical architecture; depositor agreements; metadata policy documentation, incorporating details of metadata requirements and a statement describing where responsibility for its procurement falls; process workflow documentation.
- Conforms See TRAC Conformance Document B.5.2
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Evidence</th>
<th>Conforms</th>
<th>See TRAC Conformance Document</th>
</tr>
</thead>
<tbody>
<tr>
<td>B5.3</td>
<td>Repository can demonstrate that referential integrity is created between all archived objects (i.e., AIPs) and associated descriptive information.</td>
<td>Descriptive metadata; persistent identifier/locator associated with AIP; documented relationship between AIP and metadata; system documentation and technical architecture; process workflow documentation.</td>
<td>Conforms</td>
<td>B.5.3</td>
</tr>
<tr>
<td>B5.4</td>
<td>Repository can demonstrate that referential integrity is maintained between all archived objects (i.e., AIPs) and associated descriptive information.</td>
<td>Log detailing ongoing monitoring/checking of referential integrity, especially following repair/modification of AIP; legacy descriptive metadata; persistence of identifier/locator; documented relationship between AIP and metadata; system documentation and technical architecture; process workflow documentation.</td>
<td>Conforms</td>
<td>B.5.4</td>
</tr>
<tr>
<td>B6.1</td>
<td>Repository documents and communicates to its designated community(ies) what access and delivery options are available.</td>
<td>Public versions of access policies; delivery policies; fee policies.</td>
<td>Conforms</td>
<td>B.6.1</td>
</tr>
<tr>
<td>B6.2</td>
<td>Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/depositors.</td>
<td>Access policies; use statements.</td>
<td>Conforms</td>
<td>B.6.2</td>
</tr>
<tr>
<td>B6.3</td>
<td>Repository ensures that agreements applicable to access conditions are adhered to.</td>
<td>Access policies; logs of user access and user denials; access system mechanisms that prevent unauthorized actions (such as save, print, etc.); user compliance agreements.</td>
<td>Conforms</td>
<td>B.6.3</td>
</tr>
<tr>
<td>B6.4</td>
<td>Repository has documented and implemented access policies (authorization rules, authentication requirements) consistent with deposit agreements for stored objects.</td>
<td>Access validation mechanisms within system; documentation of authentication and validation procedures.</td>
<td>Conforms</td>
<td>B.6.4</td>
</tr>
<tr>
<td>B6.5</td>
<td>Repository access management system fully implements access policy.</td>
<td>Logs and audit trails of access requests; information about user capabilities (authentication matrices); explicit tests of some types of access.</td>
<td>Conforms</td>
<td>B.6.5</td>
</tr>
<tr>
<td>B6.6</td>
<td>Repository logs all access management failures, and staff review inappropriate &quot;access denial&quot; incidents.</td>
<td>Access logs; capability of system to use automated analysis/monitoring tools and generate problem/error messages; notes of reviews undertaken or action taken as result of reviews.</td>
<td>Conforms</td>
<td>B.6.6</td>
</tr>
<tr>
<td>B6.7</td>
<td>Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is completed in relation to the request.</td>
<td>System design documents; work instructions (if DIPs involve manual processing); process walkthroughs; logs of orders and DIP production; test accesses to verify delivery of appropriate digital objects.</td>
<td>Conforms</td>
<td>B.6.7</td>
</tr>
<tr>
<td>B6.8</td>
<td>Repository can demonstrate that the process that generates the requested digital object(s) (i.e., DIP) is correct in relation to the request.</td>
<td>System design documents; work instructions (if DIPs involve manual processing); process walkthroughs; logs of orders and DIP production.</td>
<td>Conforms</td>
<td>B.6.8</td>
</tr>
<tr>
<td>B6.9</td>
<td>Repository demonstrates that all access requests result in a response of acceptance or rejection.</td>
<td>System design documents; work instructions (if DIPs involve manual processing); process walkthroughs; logs of orders and DIP production.</td>
<td>Conforms</td>
<td>B.6.9</td>
</tr>
<tr>
<td>B6.10</td>
<td>Repository enables the dissemination of authentic copies of the original or objects traceable to originals.</td>
<td>System design documents; work instructions (if DIPs involve manual processing); process walkthroughs; production of a sample authenticated copy; documentation of community requirements for authentication.</td>
<td>Conforms</td>
<td>B.6.10</td>
</tr>
</tbody>
</table>
## C. Technologies, Technical Infrastructure, & Security

<table>
<thead>
<tr>
<th>Metric</th>
<th>Example Evidence</th>
<th>Response</th>
<th>Documentation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1.1 Repository functions on well-supported operating systems and other core infrastructural software.</td>
<td>Evidence: Software inventory; system documentation; support contracts; use of strongly community supported software (i.e., Apache).</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.1</td>
<td></td>
</tr>
<tr>
<td>C1.2 Repository ensures that it has adequate hardware and software support for backup functionality sufficient for the repository’s services and for the data held, e.g., metadata associated with access controls, repository main content.</td>
<td>Evidence: Documentation of what is being backed up and how often; audit log/inventory of backups; validation of completed backups; disaster recovery plan—policy and documentation; “fire drills”—testing of backups; support contracts for hardware and software for backup mechanisms.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.2</td>
<td></td>
</tr>
<tr>
<td>C1.3 Repository manages the number and location of copies of all digital objects.</td>
<td>Evidence: random retrieval tests; system test; location register/log of digital objects compared to the expected number and location of copies of particular objects.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.3</td>
<td></td>
</tr>
<tr>
<td>C1.4 Repository has mechanisms in place to ensure any/multiple copies of digital objects are synchronized.</td>
<td>Evidence: Workflows; system analysis of how long it takes for copies to synchronize; procedures/documentation of operating procedures related to updates and copy synchronization; procedures/documentation related to whether changes lead to the creation of new copies and how those copies are propagated and/or linked to previous versions.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.4</td>
<td></td>
</tr>
<tr>
<td>C1.5 Repository has effective mechanisms to detect bit corruption or loss.</td>
<td>Evidence: Documents that specify bit error detection and correction mechanisms used; risk analysis; error reports; threat analyses.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.5</td>
<td></td>
</tr>
<tr>
<td>C1.6 Repository reports to its administration all incidents of data corruption or loss, and steps taken to repair/replace corrupt or lost data.</td>
<td>Evidence: Preservation metadata (e.g., PDI) records; comparison of error logs to reports to administration; escalation procedures related to data loss.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.6</td>
<td></td>
</tr>
<tr>
<td>C1.7 Repository has defined processes for storage media and/or hardware change (e.g., refreshing, migration).</td>
<td>Evidence: Documentation of processes; policies related to hardware support, maintenance, and replacement; documentation of hardware manufacturers’ expected support life cycles.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.7</td>
<td></td>
</tr>
<tr>
<td>C1.8 Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository’s ability to comply with its mandatory responsibilities.</td>
<td>Evidence: Documentation of change management process; comparison of logs of actual system changes to processes versus associated analyses of their impact and criticality.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.8</td>
<td></td>
</tr>
<tr>
<td>C1.9 Repository has a process for testing the effect of critical changes to the system.</td>
<td>Evidence: Documented testing procedures; documentation of results from prior tests and proof of changes made as a result of tests.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.9</td>
<td></td>
</tr>
<tr>
<td>C1.10 Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment.</td>
<td>Evidence: Risk register (list of all patches available and risk documentation analysis); evidence of update processes (e.g., server update manager daemon); documentation related to the update installations.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.1.10</td>
<td></td>
</tr>
<tr>
<td>C2. Appropriate technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requirement</td>
<td>Evidence</td>
<td>Conforms</td>
<td>TRAC Conformance Document</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>C2.1 Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed.</strong></td>
<td>Technology watch; documentation of procedures; designated community profiles; user needs evaluation; hardware inventory.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.2.1</td>
<td></td>
</tr>
<tr>
<td><strong>C2.2 Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed.</strong></td>
<td>Technology watch; documentation of procedures; designated community profiles; user needs evaluation; software inventory.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.2.2</td>
<td></td>
</tr>
<tr>
<td><strong>C3. Security</strong></td>
<td>ISO 17799 certification; documentation describing analysis and risk assessments undertaken and their outputs; logs from environmental recorders; confirmation of successful staff vetting.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.3.1</td>
<td></td>
</tr>
<tr>
<td><strong>C3.2 Repository has implemented controls to adequately address each of the defined security needs.</strong></td>
<td>ISO 17799 certification; system control list; risk, threat, or control analyses; addition of controls based on ongoing risk detection and assessment.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.3.2</td>
<td></td>
</tr>
<tr>
<td><strong>C3.3 Repository staff have delineated roles, responsibilities, and authorizations related to implementing changes within the system.</strong></td>
<td>ISO 17799 certification; organizational chart; system authorization documentation.</td>
<td>Conforms</td>
<td>See TRAC Conformance Document C.3.3</td>
<td></td>
</tr>
<tr>
<td><strong>C3.4 Repository has suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s).</strong></td>
<td>ISO 17799 certification; disaster and recovery plans; information about and proof of at least one off-site copy of preserved information; service continuity plan; documentation linking roles with activities; local geological, geographical, or meteorological data or threat assessments.</td>
<td>Partially-Conforms</td>
<td>See TRAC Conformance Document C.3.4</td>
<td></td>
</tr>
</tbody>
</table>
Formal Statement of Conformance to ISO 14721:2012

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips  Assistant Dean for Digital Libraries
Ana Krahmer  Supervisor, Digital Newspaper Unit
Hannah Tarver  Department Head, Digital Projects Unit
Daniel Alemneh  Supervisor, Digital Curation Unit
Laura Waugh  Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Formal Statement Of Conformance to ISO 14721:2012

This formal statement of conformance to the Open Archival Information System (OAIS) model is based on the official standard (ISO 14721:2012), which is not available online. The text at the Consultative Committee for Space Data Systems (CCSDS) is an earlier version from the official standard, but may prove helpful in the interpretation of this document.

Overview and Criteria for Conformance

Section 1.4 of ISO 14721:2012 defines the requirements of conformance to the OAIS Standard as an archive that can:

- Support the model of information described in Section 2.2
- Fulfill the responsibilities listed in Section 3.1

Criteria for Section 2.2

The information model of Section 2.2 describes information transferred to and from the custody of the OAIS-conforming system in terms of Information Packages, which have particular characteristics.

Each Information Package should contain three parts:

- Content Information
- Preservation Description Information: Provenance (the source), Context (links to other objects), Reference (identifiers for retrieval), and Fixity
- Packaging Information

The system is required to support three types of Information Packages:

- Submission Information Package (SIP)
- Archival Information Package (AIP)
- Dissemination Information Package (DIP)

ISO 14721:2012 also requires that Information Packages be associated with Descriptive Information sufficient to locate them.
Criteria for Section 3.1

The mandatory requirements of Section 3.1 apply to the organization operating the OAIS archive, requiring the OAIS-conforming system to enable the organization to:

- Negotiate for and accept appropriate information from information Producers.
- Obtain sufficient control of the information provided to the level needed to ensure Long Term Preservation.
- Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided, thereby defining its Knowledge Base.
- Ensure that the information to be preserved is Independently Understandable to the Designated Community. In particular, the Designated Community should be able to understand the information without needing special resources such as the assistance of the experts who produced the information.
- Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, including the demise of the Archive, ensuring that it is never deleted unless allowed as part of an approved strategy. There should be no ad-hoc deletions.
- Make the preserved information available to the Designated Community and enable the information to be disseminated as copies of, or as traceable to, the original submitted Data Objects with evidence supporting its Authenticity.

UNT Libraries’ Digital Collections Conformance Statement

The UNT Libraries and the UNT Libraries’ Digital Collections are committed to developing our digital preservation systems and practices in conformance to the OAIS Reference model standard.

Conformance with Section 2.2

The UNT Libraries’ Digital Collections meet the requirements for Information Package types and contents and have established these practices in a number of documents.

Content Information: The UNT Libraries’ Digital Collections support a wide range of activities for their designated community but remain true to their mission of long-term collecting, preserving, and providing access to digital resources that support research, teaching, and scholarship of all types.
As such, the UNT Libraries have a number of policies, specifications, standards, and best practices outlining the requirements of the Content Information they collect and maintain as part of their repository. These requirements articulate both the intellectual content collected (what kinds of data are or are not acquired) as well as the technical specifications in place for this content. Additionally these policies define the levels of support available for content that does not match the minimum specifications at a technical level.

*Related documentation:*
  - The UNT Libraries: OAIS Information Package Specification
  - UNT Libraries Policy: Collection Development Policy for the UNT Libraries’ Digital Collections
  - Digital Libraries Policy: Digital Collections Usage and Feedback Policy

*Preservation Description Information:* The UNT Libraries maintain preservation description information about all digital objects acquired and preserved as part of the UNT Libraries’ Digital Collections. This preservation description information is generated throughout the lifecycle of the resource and is serialized in a number of locations throughout the infrastructure, often in an intentionally redundant fashion.

For each file in a digital object the following preservation metadata is programmatically generated and stored within the AIP created for the object:

- PREMIS Object Entity record for each file
- JHOVE output stream for each file
- UNIX File command output serialization for each file
- METS file-level metadata in the METS File Section for each file

During and after ingest into the Coda repository, PREMIS Events are created and collected using the PREMIS Event Service application and associated with each resource. PREMIS Events captured at present include:

- Ingest
- Fixity Verification
- Replication
- Migration
Packaging Information: The UNT Libraries use the BagIt specification to encapsulate both the Content Information and the Preservation Description Information for each of the digital objects added to the UNT Libraries’ Digital Collections. Additionally each digital object added to the Coda repository must be a valid UNT-AIP as defined by the UNT Libraries OAIS Information Packaging Specification. This requires the use of a METS to encapsulate Content Information and relevant Preservation Description Information in a document that conforms to the UNTL METS Archival Information Package Profile.

Information Package Types (SIP, AIP, DIP): The UNT Libraries: OAIS Information Package Specification acts as the guiding document that specifies the creation of Information Packages in the UNT Libraries’ Digital Collections. Additionally, the METS Archival Information Package Profile is used as part of that specification in the creation of Archival Information Packages (AIPs) for the Coda repository. The UNT Libraries has documentation covering the conversion between a SIP to an AIP and from an AIP to a DIP.

Descriptive Information: Resources collected and ingested into the UNT Libraries’ Digital Collections all have Descriptive Information associated with each object in the form of a metadata record that can be validated against the UNTL XML Schema, and which uses the UNTL Metadata Guidelines for input guidelines during the creation of the metadata. This Descriptive Information is paired with the originating Content Information by reference with the use of Archival Resource Keys (ARKs) that connect the Descriptive Information held in the Aubrey access system and Archival Information Package (AIP) in the Coda repository.
Conformance with Section 3.1

The UNT Libraries as an organization meet the required responsibilities, as outlined in this section.

**Negotiate for and accept appropriate information from Information Producers.**

The UNT Libraries work with Information Producers to acquire needed information for the creation of Submission Information Packages and associated Descriptive Information.

**Obtain sufficient control of the information provided to the level needed to ensure Long Term Preservation.**

The UNT Libraries require contributing partners to sign a Standard Partnership Agreement, Memorandum of Agreement for Digital Rights, or other legal document such as a Memorandum of Understanding that outlines the rights and information needed by the UNT Libraries to preserve and make available items in the UNT Libraries' Digital Collections.

**Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided, thereby defining its Knowledge Base.**

The UNT Libraries' Digital Collections have a broad Designated Community interested in utilizing the digital resources, including:

- The UNT Extended Community (comprising students, faculty, staff, alumni, and administrators)
- Contributing partners and their constituents
- The larger academic community and researchers
- The general public throughout the world

Further information related to the Designated Community of the UNT Libraries' Digital Collections can be found in the following policies at [http://www.library.unt.edu/policies/](http://www.library.unt.edu/policies/).
● UNT Libraries Policy: Collection Development Policy for the UNT Libraries' Digital Collections
● UNT Libraries Policy: Digital Preservation Policy Framework
● UNT Libraries Policy: Digital Collections Usage and Feedback Policy

Ensure that the information to be preserved is Independently Understandable to the Designated Community. In particular, the Designated Community should be able to understand the information without needing special resources such as the assistance of the experts who produced the information.

Resources collected, preserved, and made accessible by the UNT Libraries' Digital Collections are available via access interfaces such as The Portal to Texas History, the UNT Digital Library and the Gateway to Oklahoma History. These systems make available the Dissemination Information Packages (DIPs) generated from the Archival Information Packages (AIPs) preserved in the Coda repository. By actively making these resources available via the Web using common formats, the UNT Libraries' Digital Collections ensure that the information preserved is understandable to their Designated Community.

Additional policy documents that discuss issues related to the understandability and utility of items by the designated community include:

● UNT Libraries Policy: Collection Development Policy for the UNT Libraries' Digital Collections
● UNT Libraries Policy: Digital Preservation Policy Framework
● UNT Libraries Policy: Digital Collections Usage and Feedback Policy

Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, including the demise of the Archive, ensuring that it is never deleted unless allowed as part of an approved strategy. There should be no ad-hoc deletions.

The UNT Libraries have a number of policies, procedures, workflows, and systems in place to ensure that preserved content remains viable through a wide range of contingencies. These are in part discussed by the UNT Libraries Policy: Digital Preservation Policy Framework and the UNT Libraries TRAC Conformance Document in more detail. These documents include information about who would assume custody of digital resources in the event that the UNT Libraries were not able to manage the resources anymore; the majority of the content would
return to the partner institutions, and large portions would be covered by memoranda of understanding (MOUs) with state and federal institutions such as the Texas Secretary of State's Office, the Government Publishing Office, and the National Archives and Records Administration.

Related Documentation:
- UNT Libraries TRAC Conformance Document

Make the preserved information available to the Designated Community and enable the information to be disseminated as copies of, or as traceable to, the original submitted Data Objects with evidence supporting its Authenticity.

The UNT Libraries Digital Collections can provide the full history of the preserved information to the Designated Community when requested. Events important to the lifecycle of the resource are captured as part of the PREMIS Event Service, the Digital Libraries’ Major Event Log system, and in transactional versioning of Descriptive Information and end-user access metadata as part of the Aubrey metadata editing infrastructure.

Additional documentation related to the authenticity of objects and associated metadata is available in the following documents:
- UNT Libraries Policy: Collection Development Policy for the UNT Libraries' Digital Collections
- UNT Libraries Policy: Digital Collections Usage and Feedback Policy
- UNT Libraries TRAC Conformance Document

References
Digital Projects Unit: Digitization Standards -
http://www.library.unt.edu/digital-projects-unit/standards

Digital Projects Unit: Digital File Formats -
http://www.library.unt.edu/digital-projects-unit/digital-file-formats

JHOVE - http://jhove.sourceforge.net/

PREMIS - http://www.loc.gov/standards/premis/

UNT Libraries: OAIS Information Package Specification -
http://www.library.unt.edu/digital-libraries/trusted-digital-repository

UNT Libraries Policy: Collection Development Policy for the UNT Libraries’ Digital Collections -
http://www.library.unt.edu/policies/

UNT Libraries Policy: Digital Preservation Policy Framework -
http://www.library.unt.edu/policies/

UNT Libraries Policy: Digital Collections Usage and Feedback Policy -
http://www.library.unt.edu/policies/

UNT Libraries SIP-to-AIP Conversion Workflow -
http://www.library.unt.edu/digital-libraries/trusted-digital-repository

UNT Libraries AIP to DIP/ACP Conversion Workflow -
http://www.library.unt.edu/digital-libraries/trusted-digital-repository

UNT Libraries TRAC Conformance Document -
http://www.library.unt.edu/digital-libraries/trusted-digital-repository


UNTL METS Archival Information Package Profile -
http://www.loc.gov/standards/mets/profiles/00000045.xml

UNTL XML Schema - http://digital2.library.unt.edu/un tl.xsd
Collection Development Policy for the UNT Libraries’ Digital Collections

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips  Assistant Dean for Digital Libraries
Hannah Tarver  Department Head, Digital Projects Unit
Ana Krahmer  Supervisor, Digital Newspaper Unit
Daniel Alemneh  Supervisor, Digital Curation Unit
Laura Waugh  Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Note: This is an archived version of this document.
An up-to-date version if available at: http://www.library.unt.edu/policies/
Collection Development Policy for the
UNT Libraries’ Digital Collections

1. Scope
This document describes the policies governing the acquisition, curation, and management of materials in the UNT Libraries’ (UNTL) Digital Collections.

2. UNT Libraries’ Digital Collections Mission Statement

In support of the UNT Libraries Mission, the UNT Libraries’ Digital Collections is dedicated to the long-term collection, production, maintenance, delivery, and preservation of a wide range of high-quality digital resources and services for the UNT Community and users throughout the world.

3. Guiding Principles
The UNT Libraries’ Digital Collections are maintained in accordance with the following guiding principles:

- Provide unified digital access to resources of value to the UNT Community
- Make access to these digital collections easy and transparent
- Sustainably manage the digital collections in perpetuity

3.1 UNT Libraries’ Digital Collections Purpose
Under the umbrella of UNTL’s mission, the UNT Libraries’ Digital Collections exist to:

- Acquire, preserve, and provide access to digital resources required by the UNT community to successfully execute the mission of UNT as a major public research university deeply committed to advancing educational excellence and preparing students to become thoughtful, engaged citizens of the world
- Provide an optimum research environment to faculty, staff, students, and community members by enabling open access to significant research collections
- Support open access to and long-term preservation of the scholarly output of UNT to the best of its ability
• Encourage long-term preservation of content contributed by partners, researchers, and institutions
• Provide digital preservation support for non-UNT community organizations in Texas who otherwise do not have access to preservation support

3.2 Audience
The UNT Libraries’ Digital Collections have a broad audience interested in utilizing the digital resources, including:

• The UNT Extended Community (comprising students, faculty, staff, alumni, and administrators)
• Contributing partners and their constituents
• The larger academic community and researchers
• The general public throughout the world

4. Collection Scope
The UNT Libraries’ Digital Collections are comprised of three publicly-facing interfaces: The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History. The digital holdings within these three interfaces are owned by the UNT Libraries or contributing partners.

It should be noted that these resources are different in scope from “Electronic Resources,” which are resources purchased from and managed by third party vendors.

4.1 The Portal to Texas History
The Portal to Texas History includes materials collected and curated by institutions throughout the State of Texas. Most often these resources align in some way to the subject of Texas history, the geography of Texas, or are somehow related to resources useful to users of the Portal. While the majority of the resources in the Portal relate to Texas history, the subject matter is broader and encompasses resource collected and held by institutions and private individuals in Texas.

The Portal to Texas History contains collections from partner institutions across the state. Collections that have received significant support either internally or externally include:

**The Texas Digital Newspaper Program** provides access and preservation to digitized and born-digital newspaper issues from the earliest days of Texas until the present.
The Texas Register is a partnership with the Office of the Texas Secretary of State, Texas Register Section, providing free access to all issues of the Texas Register from Volume 1, No. 1 (January 6, 1976) to the present.

Texas Laws and Resolutions Archive embodies a collection of legislative bills that were filed with the Office of the Texas Secretary of State - Statutory Documents beginning with the 78th Legislative Session. All bills, joint resolutions, and concurrent resolutions passed by the Texas Legislature, including those vetoed by the Governor. In partnership with the Office of the Texas Secretary of State.

Texas State Publications includes digitized Texas documents distributed to the UNT Libraries by the Texas State Library and Archives Commission (TSLAC) as part of the Texas Depository Library Program. In addition this collection contains Web-harvested documents from Texas agencies which fall within the scope of the UNT Libraries Texas State Documents Collection.

4.2 UNT Digital Library
The UNT Digital Library is a centralized repository for the collections held by the libraries, colleges, schools, and departments at the University of North Texas. This system provides access to a wide range of collections from partners around campus. Collections of note include:

UNT Scholarly Works
This collection serves as the UNT institutional repository, housing materials from the UNT community's research, creative, and scholarly activities.

UNT Data Repository
This collection is a central archive for research data outputs of the UNT community. The UNT Data Repository works in conjunction with the UNT Scholarly Works institutional repository to ensure accessibility to the full range of research outputs from UNT.

UNT Theses and Dissertations
This collection houses theses and dissertations created by masters and doctoral students in the degree-seeking process at UNT.

CyberCemetry
This collection is an archive of government websites that have ceased operation. The CyberCemetry features a variety of topics indicative of the broad nature of government information and in particular, websites that cover topics supporting UNT's curriculum, and particular program strengths.

Technical Reports and Image Library (TRAIL)
This collection includes hard-to-find reports published by various government agencies. The publications in this collection contain reports, images, and technical descriptions of research performed for U.S. government agencies prior to 1975 and covering a broad range of topics.

### 4.3 The Gateway to Oklahoma History

The Gateway to Oklahoma History is a portal to resources housed at the Oklahoma Historical Society and is made available in partnership between the Oklahoma Historical Society and the University of North Texas. Collections of note include:

**Oklahoma Digital Newspaper Program**

This collection houses digitized newspapers from throughout the state of Oklahoma.

**Oklahoma Publishing Company Photograph Collection**

The Oklahoma Publishing Company is the parent company several newspapers in Oklahoma. This collection features a wide-variety of photographs taken for stories in these newspapers.

### 5. Selection Criteria

The UNT Libraries’ Digital Collections make available collections that support research and scholarship according to the needs of the UNT Extended Community. For inclusion within the UNT Libraries’ Digital Collections, objects must meet all *Level One* criteria and at least one of *Level Two* criteria.

*Level One* criteria cover objective standards such as copyright, extent, and format issue. Items must meet *all* of the criteria in *Level One* to be eligible for addition into one of the UNT Libraries’ Digital Collections:

Each item must:
- Be in the public domain, or have documented, non-revocable permission granted by the copyright holder, or be allowed by Section 108 (17 U.S.C. § 108)
- Be intended for public viewing and use
- Be complete, such as an entire publication, article, etc., and not a “part” such as an abstract, forward, or title page
- Be in a standard format accessible through current file viewers or have a documented conversion path to move the format into a standard format
- Be intended for permanent storage in the UNT Libraries’ Digital Collections
Level Two criteria cover subjective elements requiring review, assessment, and the professional judgement of UNT Digital Libraries’ librarians and staff in collaboration with subject experts. Items must satisfy at least one criterion from these categories:

- **Utility** - the items support the University of North Texas’ Extended Community goals:
  - Research
  - Scholarship
  - Teaching
- **Demand** - items meet anticipated or demonstrated community demands:
  - Departmental priorities
  - Citation frequency
  - Digital item usage
- **Commitment** - items meet an obligation or requirement established by external means:
  - Institutional archives
  - Legal requirements
  - Contracts or MOUs

### 6. Retention and Evaluation

All objects included in the UNT Libraries’ Digital Collections are intended to be retained permanently upon acceptance into the UNT Libraries’ Digital Collections. The UNT Libraries’ Digital Collections will not be used as a temporary storage facility for digital items.

### 7. Collection Curators

The UNT Libraries provide curators for their digital collections who exercise intellectual and administrative stewardship responsibilities for digital objects stored in the UNT Libraries’ Digital collections. The Digital Collection curator is a role rather than a specific individual, and it is understood that, over time, many different individuals will play this role for a specific collection.

Each Digital Collection Curator will ensure sound stewardship of the collection and will appropriately manage these digital objects over the long-term through standard digital management practices, which include the following:

- **Intellectual property rights**
  - managing the legal rights necessary for all services over time, and
  - where necessary, obtaining legal clearances for the right to
    - make copies for backup purposes,
    - make derivative copies, and
    - distribute publicly
- **Metadata** - follow the UNT Libraries Metadata Guidelines, publicly available from [http://www.library.unt.edu/digital-projects-unit/metadata](http://www.library.unt.edu/digital-projects-unit/metadata), to support the design of
appropriate descriptive, administrative, technical, and structural metadata are created for each object

- Discovery and access - ensure that metadata, or descriptions of objects are publicly available

8. Access and Usage

Some materials may have restrictions on access and/or usage. See the UNT Libraries’ Digital Collections Usage and Feedback Policy, sections 2. Access and 3. Rights/Terms of Use for complete guidelines.

9. Review Cycle

This policy will be reviewed and updated as needed with a full review every two years to assure timely revisions as technology progresses, collection focus changes, and digital library collections mature.
The Portal to Texas History:
Standard Partnership Agreement

This work is licensed under a Creative Commons Attribution 4.0 International License.
This Agreement is entered into as of the ____ day of ____________, 201__ by and between:

**Type Institution Name Here** (“Institution”), notices for whom should be sent to the address shown below or to such address as Institution may inform UNT of in writing:

Name: ______________________________________________________________
Address: ____________________________________________________________
____________________________________________________________________
____________________________________________________________________

and the University of North Texas (“UNT”), with offices at Willis Library, Room 370, 1516 West Highland Street, Denton, TX 76201. Notices for the UNT Libraries’ Portal to Texas HistorySM should be sent to UNT Libraries, Attn: Digital Projects Unit, 1155 Union Circle #305190, Denton, TX 76203-5190.

WHEREAS, UNT has a repository of digital information (“The Portal to Texas HistorySM”) in which participating institutions shall contribute historical materials, and UNT has agreed to store and deliver this information to Public, and

WHEREAS, Institution desires to contribute materials to the Portal to Texas HistorySM,

NOW, THEREFORE, the parties agree as follows:

1) Definition:

“Public” shall mean users via general Internet access.

“Licensed Materials” shall mean all digital files contributed by Institution to be stored and delivered from UNT computers other than public records.

“Metadata” shall mean all information used to search and locate an object such as title, author, subjects, keywords, descriptions, or publisher contributed by Institution, derived from the provided digital files, or created by UNT.

“Public Records” shall mean records that a governmental agency or body is legally required to keep.

“Public Domain” shall mean a work free of copyright.
2) The Institution represents and warrants that:

   a) The Institution has the full right and authority to enter into this Agreement;
   
   b) The Institution has the full right and authority to perform its duties as described
       herein and to authorize UNT to perform duties as described herein in reliance on the
       Institution’s warranty;
   
   c) The Institution has not assigned, pledged or otherwise encumbered the rights in the
       Licensed Materials, corresponding digital reproductions, nor related material created
       for or provided to UNT under this Agreement. The Institution represents and warrants
       that neither the digital reproductions and/or related material created for or provided to
       UNT under this Agreement will interfere with or violate any rights of third parties,
       any copyright, or any other right, and will contain no matter that is libelous or in
       violation of any rights of privacy or that is otherwise contrary to law.
   
   d) The Institution will give UNT prompt written notice of any suit, action, proceeding,
       or complaint brought against it alleging facts, which if proven, would constitute a
       breach of the warranties made by the Institution.

3) UNT will cease to store and/or distribute products if so ordered by any judicial rule or
   order or otherwise at the discretion of UNT. UNT reserves the right to discontinue storage
   and distribution of the Licensed Materials.

4) Institution hereby grants to UNT the perpetual and unlimited rights to: permit UNT to post
   Licensed Materials; permit Public to search, view, and browse the Licensed Materials;
   permit Public to print or download the Licensed Materials for all non-commercial purposes
   allowed by this Agreement; create lower resolution derivative versions of Licensed
   Materials for systematic distribution to Public; and use Licensed Materials in presentations
   about and promotions for the project, and in online exhibits or educational materials, such
   as lesson plans or learning activities.

5) Excluding the use of materials in the Public Domain and except as expressly permitted in
   this Agreement, UNT will not use the Licensed Materials for any substantial or systematic
   reproduction; re-distribution, re-selling or sub-licensing in any manner including in
   connection with fee-for-service use; or a commercial use other than one expressly agreed
   upon by the parties.

6) Any Metadata created or contributed by Institution or derived from Licensed Materials by
   UNT is in the public domain. Institution warrants that Institution has dedicated any
   created, contributed, or derived Metadata to the public domain by waiving all of
   Institution’s rights to the Metadata worldwide under copyright law, including all related
   and neighboring rights. UNT may copy, modify, and distribute Metadata, even for
   commercial purposes without permission.
7) UNT shall post appropriate notices informing users of their obligations under this agreement and of the applicability of United States Copyright Act to the reproduction of the Licensed Materials. UNT shall provide appropriate notices informing users of the name and address of the institution that submitted each item or collection.

8) Except as permitted by this Agreement, neither UNT, nor Public of the UNT Libraries’ Portal to Texas History℠ may modify, adapt, transform, translate or create any derivative work based on any materials included in the Licensed Materials, or otherwise use any such materials in a manner that would infringe the copyright or other proprietary rights therein. UNT shall post appropriate notices informing Public of their obligations under this Agreement and of the applicability of United States Copyright laws to the reproduction of the Licensed Material.

9) In no event shall UNT be liable for incidental, special, or consequential damages suffered by Institution, even if it has previously been advised of the possibility of such damages. Institution’s exclusive remedy for violation of this Agreement shall be the right to terminate this agreement upon 30 day notice and the right to use the licensed materials. Nothing herein shall restrict the right of the Institution from taking such legal action as it considers necessary to prevent the unauthorized use of Licensed Materials by third parties.

10) UNT shall not be liable to Institution for the unauthorized use of Licensed Materials or other violations by Public

11) Institutions may submit digital files that include public records or material in the Public Domain that are subject to the provisions of Texas law, which governs the use of such records. Use restrictions contained in this Agreement shall not apply to any public records or material in the Public Domain.

12) The warranties contained herein shall survive the termination of this Agreement.

13) This Agreement will be construed according to the laws of the State of Texas.

14) UNT is required by law to provide notice that Chapter 2260 of the Texas Government Code establishes a dispute resolution process for contracts involving goods, services, and certain types of projects. If Chapter 2260 applies to this Agreement, then the statutory dispute resolution process must be used by a party attempting to resolve all of its disputes with UNT arising under this Agreement.
15) General Description of Materials to be added: ______________________________________
___________________________________________________________________________

IN WITNESS WHEREOF, we have signed our names below.

Institution                                      University of North Texas
___________________________________
Name of authorized representative (print)        Martin Halbert
___________________________________
Dean, UNT Libraries

___________________________________
Title

___________________________________
Signature                                     Signature

___________________________________
Date signed                                    Date signed
The Portal to Texas History:
Memorandum of Agreement for Digital Rights

This work is licensed under a Creative Commons Attribution 4.0 International License.
Memorandum of Agreement for Digital Rights

I/we, _________________________________, affirm that I/we own the property described on this paper or attached inventory, and unconditionally give the University of North Texas (UNT), its successors and assigns, permission to digitize, host and publish online, the following items in perpetuity:

_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________

I/we understand that materials hosted by the Portal to Texas History are freely accessible for use in research, teaching (including distribution to classes), and private study. For these uses people may reproduce materials (print, make copies, or download) from the Portal to Texas History without prior permission, on the condition that they properly cite and attribute the source of materials on all copies.

I/we understand that reproduction of Portal to Texas History materials for publication, broadcast, or commercial use requires my/our written permission, unless we have assigned these rights to the UNT Libraries.

Optional: I/we unconditionally assign all rights to this material to the UNT Libraries, including all intellectual property rights. ____Yes  ____No

I/we would like more information about donating these materials to the UNT Libraries in the future.  ____Yes  ____No

_______________________________________________________________________
Signed            Date
Address_______________________________________________
_______________________________________________________________________

Phone_________________________________________________
UNT Libraries: Memorandum of Agreement for Digital Rights

This work is licensed under a Creative Commons Attribution 4.0 International License.
Memorandum of Agreement
for Digital Rights

I/we, _________________________________, affirm that I/we own the property described on this paper or attached inventory, and unconditionally give the University of North Texas (UNT), its successors and assigns, permission to digitize, host and publish online, the following items in perpetuity:

_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________

I/we understand that materials hosted by the UNT Digital Library are freely accessible for use in research, teaching (including distribution to classes), and private study. For these uses people may reproduce materials (print, make copies, or download) from the UNT Digital Library without prior permission, on the condition that they properly cite and attribute the source of materials on all copies.

I/we understand that reproduction of this title for publication, broadcast, or commercial use requires my/our written permission, unless we have assigned these rights to the UNT Libraries.

Optional: I/we unconditionally assign all rights to this material to the UNT Libraries, including all intellectual property rights. ___Yes    ___No

_______________________________________________________________________
Signed            Date
Address_______________________________________________

_______________________________________________________________________

Phone_________________________________________________
Travel and Professional Development Documentation

Date: October 2015

Version: 1.0

Contributors:

Daniel Alemneh  Supervisor, Digital Curation Unit
Mark Phillips  Assistant Dean for Digital Libraries
Ana Krahmer  Supervisor, Digital Newspaper Unit
Hannah Tarver  Department Head, Digital Projects Unit
Laura Waugh  Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
### Table-1: Year-2015

<table>
<thead>
<tr>
<th>Year</th>
<th>No</th>
<th>Event or Conference</th>
<th>Attendee from Digi</th>
</tr>
</thead>
</table>
Mark Phillips |
| 2    | 2  | PASIG - 2015  
March 11-13, 2015  
San Diego, CA | Mark Phillips |
| 3    | 3  | Dspace Training  
March 31, 2015  
Houston, TX | Daniel Alemneh |
| 4    | 4  | Texas Digital Humanities Conference  
April 10-111, 2015  
Arlington, TX | Mark Phillips  
Marcia McIntosh |
| 5    | 5  | University Forum on Teaching & Learning (UFTL),  
April 10, 2015  
Denton, TX | Daniel Alemneh |
| 6    | 6  | UNT Research Symposium on African Studies  
April 11, 2015 | Daniel Alemneh |
<table>
<thead>
<tr>
<th></th>
<th>Event Title</th>
<th>Date</th>
<th>Location</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>IFLA International News Media Conference</td>
<td>April 14-15, 2015</td>
<td>Stockholm, Sweden</td>
<td>Ana Krahmer</td>
</tr>
<tr>
<td>8</td>
<td>TLA-2015</td>
<td>April 14-17, 2015</td>
<td>Austin, TX</td>
<td>Hannah Tarver, Ana Krahmer</td>
</tr>
<tr>
<td>9</td>
<td>IIPC General Assembly 2015</td>
<td>April 21- May 1, 2015</td>
<td>San Francisco, CA</td>
<td>Mark Phillips, Lauren Ko, Brenda Reyes</td>
</tr>
<tr>
<td>10</td>
<td>TCDL-2015</td>
<td>April 27-28, 2015</td>
<td>Austin, TX</td>
<td>Daniel Alemneh, Laura Waugh, Hannah Tarver, Marcia McIntosh, Derek Rankins</td>
</tr>
<tr>
<td>11</td>
<td>Digital Initiatives Symposium</td>
<td>April 29, 2015</td>
<td>San Diego, CA</td>
<td>Marcia McIntosh</td>
</tr>
<tr>
<td>12</td>
<td>DrupalCon</td>
<td>May 10-15, 2015</td>
<td>Los Angeles, CA</td>
<td>Will Hicks</td>
</tr>
<tr>
<td>13</td>
<td>Dodging the Memory Hole II: An Action Assembly</td>
<td>May 11-12, 2015</td>
<td>Charlotte, NC</td>
<td>Ana Krahmer</td>
</tr>
<tr>
<td>No.</td>
<td>Event Name</td>
<td>Dates</td>
<td>Location</td>
<td>Organizer(s)</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>14</td>
<td>Open Access Symposium-2015</td>
<td>May 18-19, 2015</td>
<td>Dallas, TX</td>
<td>Mark Phillips, Daniel Alemneh, Laura Waugh, Dianne Jansing</td>
</tr>
<tr>
<td>16</td>
<td>Columbia University Web Archiving Summit</td>
<td>June 4-5, 2015</td>
<td>New York City, NY</td>
<td>Mark Phillips</td>
</tr>
<tr>
<td>17</td>
<td>Open Repositories Conference</td>
<td>June 8-11, 2015</td>
<td>Indianapolis, IN</td>
<td>Laura Waugh</td>
</tr>
<tr>
<td>18</td>
<td>DC-2015</td>
<td>September 2-3, 2015</td>
<td>Sao Paulo, Brazil</td>
<td>Hannah Tarver</td>
</tr>
<tr>
<td>19</td>
<td>USETDA-2015</td>
<td>September 29 -October 1, 2015</td>
<td>Austin, TX</td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td>20</td>
<td>DLF Forum</td>
<td>October 26-28, 2015</td>
<td>Vancouver, BC, Canada</td>
<td>Marcia McIntosh</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Year</th>
<th>No</th>
<th>Conference</th>
<th>Attendee from Digi</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>1</td>
<td><strong>IFLA International News Media Conference</strong></td>
<td>Ana Krahmer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>February 4-5, 2014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Salt Lake City, Utah</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><strong>Southwestern Association of Law Libraries (SWALL) Annual Conference</strong></td>
<td>Laura Waugh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>March 20-22, 2014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Austin, TX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td><strong>Code4Lib 2014</strong></td>
<td>Will Hicks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>March 24-27, 2014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Raleigh, NC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td><strong>TLA-2014</strong></td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>April 8-11, 2014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>San Antonio, TX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td><strong>TCDL-2014</strong></td>
<td>Daniel Alemneh, Laura Waugh, Hannah Tarver, Jeremy Moore, Derek Rankins, Brenda Reyes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>April 28-29, 2014</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Austin, TX</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td><strong>Open Access Symposium-2014</strong></td>
<td>Daniel Alemneh, Laura Waugh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 19-20, 2014</td>
<td></td>
</tr>
<tr>
<td>Event</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dallas, TX</td>
<td>Hannah Tarver</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| NASIG 2014 | May 1-4, 2014  
Fort Worth, TX | Hannah Tarver |
| Usability Week | June 23-27, 2014  
San Francisco, CA | Will Hicks |
| Cross Timbers Library Collaborative (CTLC) Annual Conference | July 25, 2014  
Dallas, TX | Laura Waugh  
Mark Phillips  
Hannah Tarver  
Daniel Alemneh  
Sarah Lynn Fisher  
Skye Limon |
| American Association of Colleges of Pharmacy (AACP) Annual Conference | July 26-30, 2014  
Grapevine, TX | Laura Waugh |
| IFLA World Library & Information Congress | August 16-22, 2014  
Lyon, France | Ana Krahmer |
| Beyond NDNP Pre-Conference Meeting | September 15, 2014 | Mark Phillips  
Ana Krahmer |
| National Digital Newspaper Program Annual Meeting | September 16-18, 2014 | Ana Krahmer  
Sarah Lynn Fisher |
<table>
<thead>
<tr>
<th></th>
<th>Conference</th>
<th>Dates</th>
<th>Location</th>
<th>Presenter(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>USETDA-2014</td>
<td>Sept. 24-26, 2014</td>
<td>Orlando, FL</td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td>15</td>
<td>DC-2014</td>
<td>Oct. 8 – 11, 2014</td>
<td>Austin, TX</td>
<td>Hannah Tarver</td>
</tr>
<tr>
<td>16</td>
<td>ASIS&amp;T-2014</td>
<td>Oct. 31-Nov. 5, 2014</td>
<td>Seattle, WA</td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td>17</td>
<td>Dodging the Memory Hole</td>
<td>Nov. 10-11, 2014</td>
<td>Columbia, MO</td>
<td>Mark Phillips, Ana Krahmer</td>
</tr>
<tr>
<td>18</td>
<td>University Partnerships Conference</td>
<td>Nov. 12-15, 2014</td>
<td>Denton, TX</td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td>Year</td>
<td>No</td>
<td>Conference</td>
<td>Attendee from Digi</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>1</td>
<td><strong>Code4Lib 2013</strong></td>
<td>Will Hicks, Jason Thomale</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb 11 - 14, 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chicago, IL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td><strong>iConference-2013</strong></td>
<td>Daniel Alemneh, Laura Waugh</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fort Worth, TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td><strong>TxETDA-2013</strong></td>
<td>Daniel Alemneh</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feb. 28-March 1, 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>College Station, TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td><strong>TLA Annual Conference</strong></td>
<td>Hannah Tarver</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>April 24-27, 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fort Worth, TX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td><strong>TCDL-2013</strong></td>
<td>Daniel Alemneh, Laura Waugh, Hannah Tarver, William</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 7-8, 2013</td>
<td>Hicks, Jeremy Moore, Andrew Weidner, Jerrell Jones,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Austin, TX</td>
<td>Brenda Reyes, Robert J. Wilson III</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Event</td>
<td>Date</td>
<td>Location</td>
<td>Speakers</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>6</td>
<td><strong>DrupalCon</strong></td>
<td>May 20-24, 2013</td>
<td>Portland, OR</td>
<td>Will Hicks</td>
</tr>
<tr>
<td>7</td>
<td><strong>Open Access Symposium-2013</strong></td>
<td>May 31, 2013</td>
<td>Dallas, TX</td>
<td>Daniel Alemneh, Laura Waugh, Mark Phillips, Hannah Tarver</td>
</tr>
<tr>
<td>8</td>
<td><strong>Preservation of Photographic Materials Workshop</strong></td>
<td>June 19, 2013</td>
<td>Denton, TX</td>
<td>Hannah Tarver, Jeremy Moore</td>
</tr>
<tr>
<td>9</td>
<td><strong>Open Repositories Annual Conference</strong></td>
<td>July 8-12, 2013</td>
<td>Charlottetown, PEI, Canada</td>
<td>Mark Phillips</td>
</tr>
<tr>
<td>10</td>
<td><strong>Digital Directions: Fundamentals of Creating and Managing Digital Collections</strong></td>
<td>July 21-23, 2013</td>
<td>Ann Arbor, MI</td>
<td>Hannah Tarver</td>
</tr>
<tr>
<td>11</td>
<td><strong>TLA Annual Assembly</strong></td>
<td>July 7-10, 2013</td>
<td>Austin, TX</td>
<td>Daniel Alemneh</td>
</tr>
<tr>
<td>12</td>
<td><strong>Cross Timbers Library Collaborative (CTLC) Annual Conference</strong></td>
<td>August 16, 2013</td>
<td>Denton, TX</td>
<td>Laura Waugh, Daniel Alemneh, Mark Phillips, Sarah Lynn Fisher</td>
</tr>
<tr>
<td></td>
<td>Event</td>
<td>Date</td>
<td>Location</td>
<td>Organizer(s)</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>13</td>
<td>DC-2013</td>
<td>Sept. 2 – 6, 2013</td>
<td>Lisbon Portugal</td>
<td>Andrew Wiedner</td>
</tr>
<tr>
<td>14</td>
<td>Digital Frontiers Annual Conference</td>
<td>Sept. 19-20, 2013</td>
<td>Denton, TX</td>
<td>Hannah Tarver</td>
</tr>
<tr>
<td>15</td>
<td>ETD-2013</td>
<td>Sept. 23-26, 2013</td>
<td>Hong Kong</td>
<td>Laura Waugh</td>
</tr>
<tr>
<td>16</td>
<td>ASIS&amp;T-2013</td>
<td>Nov. 1-3, 2013</td>
<td>Montreal, Canada</td>
<td>Daniel Alemneh</td>
</tr>
</tbody>
</table>
UNT Libraries’ Digital Preservation
Policy Framework

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Note: This is an archived version of this document.
An up-to-date version if available at: http://www.library.unt.edu/policies/
UNT Libraries’ Digital Preservation Policy Framework

1. Scope

This policy addresses preservation of digital collections and resources for which the University of North Texas Libraries (UNTL) is the primary custodian. Although this policy only addresses digital collections and resources for which UNTL is the primary custodian, UNTL, to the best of its abilities, has responsibility for informing, consulting, and as appropriate coordinating with other units of the University of North Texas to ensure that UNT faculty, staff, and students will have adequate ongoing access to administrative, scholarly, and other digital resources created at UNT outside of the UNT Libraries. Further, UNTL personnel will also work externally through consortia (e.g., the Texas Digital Library (TDL), and Cross Timbers Library Collaborative (CTLC)), licensing agreements, etc. to ensure that UNT faculty, staff, and students will have adequate continuing access to all currently available digital resources locally owned and managed by the UNTL. UNTL, however, cannot guarantee preservation for materials that it does not own and manage.

2. Purpose

This document formalizes the UNTL’s continued commitment to the long-term stewardship for, preservation of, and sustainable access to its diverse and extensive range of digital assets. In alignment with the UNTL mission to create, acquire, organize, disseminate, and preserve digital content, this policy makes explicit UNTL’s long-term commitment to the University of North Texas (UNT) community as its trusted digital repository. The UNTL’s digital stewardship efforts contribute to UNT’s mission to:
- Build a world-class faculty
- Develop academic programs that define UNT as a leader in scholarship and artistic endeavors
- Improve the quality of the teaching and learning environment
- Enhance and better serve the student body
- Create a more diverse university community
- Help build Texas’ future by ensuring access to this corpus of information over time

3. Objectives

The primary purpose of digital stewardship and preservation is to collect and maintain the intellectual and cultural heritage important to UNT, while at the same time making sure that it is accessible and held in trust for future use. The objectives in this statement define a framework to:
• Identify, through systematic selection, digital assets to be preserved across new generations of technology.
• Maintain continuous access to reliable data at bit-stream level, the digital assets encoded in the bit streams, as well as access to the intended contextual and intellectual meaning of the digital assets.
• Include in the scope of the program materials that originated in digital form and those that were converted to digital form.
• Protect UNTL’s digital investments through a fully-implemented digital preservation program.
• Demonstrate organizational commitment through the identification of sustainable strategies.
• Develop a cost-effective program through means such as system-wide integration, shared responsibilities, and automating human-intensive efforts, when possible.
• Comply with prevailing community standards for digital preservation and access.
• Seek, expand, and develop digital preservation methods that are appropriate for UNT.

4. Mandate

UNTL’s mandate for digital preservation is five-fold:

• **Scholarship**: As an institution of higher education, UNT is obligated to support core functions such as scholarship, teaching, and learning. As more resources and services associated with these functions become digital, UNTL’s responsibilities must expand to include the identification, stewardship, and preservation of designated digital content.

• **Institutional records**: UNT has charged UNTL with maintaining the University Archives by collecting and preserving university records that best document the history of UNT, including those in electronic format.

• **Legal obligations**: UNT has mandated responsibilities to preserve and maintain access to certain digital collections, as well as responsibilities as a federal depository library. Some legal obligations derived from Federal and State laws require us to maintain files in an archival fashion.

• **Organizational commitment**: UNTL’s commitment to digital preservation is explicitly cited in the UNTL’s current strategic plan, which calls for UNTL to:
  - develop and implement a cross-divisional plan for supporting curation, storage and dissemination of library-created or library managed digital content
  - build a robust, reliable, secure technical infrastructure base including both human and technology resources.

• **Consortia and contractual commitments**: UNTL maintains commitments to partner institutions of the UNT Libraries’ Digital Collections, including The Portal to Texas History, as well as contractual agreements to assume or share in the responsibilities for preserving designated digital content.
5. Challenges

Recognized challenges in implementing an effective digital preservation program include, but are not limited to:

- **Rapid growth**: Technology that enables the preservation and dissemination of a variety of formats changes rapidly. As different types of materials are submitted (e.g., data sets, complex digital objects), monitoring different needs of the materials (storage size, metadata, etc.) and maintaining procedures and policies based on these needs is necessary.

- **Sustainability**: A sustainable digital preservation model should be developed that will respond to technological and staffing changes as needed, without under- or overestimating the needs imposed by these changes. The need for good cost models and affordable programs is widely acknowledged, yet still not fully addressed on a wider public scale. UNTL requires sufficient funding for operations and major improvements for digital asset management, as well as designated library funding to sustain ongoing preservation efforts. Further, there are administrative complexities in ensuring cost-effective and timely action to implement preservation strategies. The scale of funding is based on the level of commitment, therefore the program should reflect reasonable expectations of requisite resources, i.e., UNTL should not promise more than can be delivered.

- **Management**: Moving from well-managed digital collections to preserved collections requires institutional effort, partnership development, and a financial commitment. UNTL should provide a thoughtful balance between access and preservation, while being mindful of preservation’s core role in maintaining access.

- **Partnerships**: UNTL must work with creators and providers of crucial content to employ appropriate maintenance prior to deposit that will facilitate future preservation.

- **Flexibility**: The digital preservation plan must continually review its abilities to respond to the evolving technological capabilities and changing user expectations without jeopardizing the ongoing care of the digital content.

- **Expertise**: UNTL must commit to continually updating staff expertise, where appropriate, as technologies change.

- **Rights**: There may be intellectual property and other rights-based constraints on providing access that impact digital preservation efforts.

- **Privacy**: As UNTL strives to make increasingly more digital collections openly available to the world, personal information might appear within these collections that violates the privacy of the people whom this information represents.
6. Principles

6.1 Guiding principles

UNTL will use consistent criteria for selection and preservation as with other resources in the libraries. Materials selected for digital stewardship and preservation carry with them UNTL’s commitment to maintain the materials for as long as needed or desired.

- The Libraries are committed to the long term preservation of selected content.
- Digital preservation is an integral part of UNTL’s process.
- Processes, policies, and the institutional commitment are transparently documented.
- Levels of preservation and time commitments are determined by selectors and curators, in consultation with technical experts.
- UNTL will participate in the development of digital preservation community standards, practices, and solutions.

6.2 Operating principles

The Libraries will strive to:

- Develop a scalable, reliable, sustainable, and auditable digital preservation infrastructure.
- Manage the hardware, software, and storage media components of the digital preservation function in accordance with environmental standards, quality control specifications, and security requirements.
- Comply with the Open Archival Information System (OAIS) reference model and other appropriate digital preservation standards and practices.
- Ensure that the digital archive is as interoperable as possible by utilizing open source options whenever feasible.
- Ensure the integrity of the data within the digital preservation infrastructure.
- Secure metadata (e.g., administrative, descriptive, preservation, provenance, rights, and technical information) necessary for access to and the use of the digital assets.
- Comply with copyright, intellectual property rights, and/or other legal rights related to copying, storage, modification, and use of digital resources.

6.3 Standards

- UNT is best served when distributed and disparate systems conform to standards and best practices that make communication between these storage systems possible.
- To utilize the OAIS Reference Model as the basis for developing and implementing strategies and tools for long-term digital information preservation and access.
7. Categories of Commitment

UNTl’s levels of commitment as outlined below recognize that developing solutions for “born digital” materials informs solutions for the other categories; it does not imply that these assets are inherently more valuable or important than any of the other categories and/or our traditional, analog materials.

- **Born-digital materials:** Rigorous effort will be made to ensure preservation in perpetuity of materials selected for preservation, both library resources and institutional records.
- **Digitized materials (no available analog):** Every reasonable step will be taken to preserve materials without a print analog, when re-digitizing is not possible or analog versions are located elsewhere. Also included are digitized materials that have annotations or other value-added features making them difficult or impossible to recreate.
- **Digitized materials (available analog):** Reasonable measures will be taken to extend the life of the digital objects with a readily available print analog. However, the cost of re-digitizing as needed will be weighed against the cost of preserving the existing digital objects.
- **Commercially available digital resources:** UNTL has responsibility for working externally through consortia, licensing agreements, etc. to assure that one party or parties provides the necessary infrastructure to provide for preservation activities so that UNT faculty, staff, and students will have adequate ongoing access to commercially available digital resources. If the resources are external to UNTL, there needs to be an articulated exit strategy in the event of the cessation of the consortia or licensing agreements. Particular emphasis should be given to resources which exist in digital form only.
- **Other items and materials:** No preservation steps will be taken for materials requested for short-term use, such as materials scanned for E-reserve and document delivery, or for content that does not meet the requirements of the Collection Development Policy for the UNT Libraries’ Digital Collections

8. Levels of Preservation

UNTl adheres to the National Digital Stewardship Alliance (NDSA) Levels of Digital Preservation developed by NDSA, as a tiered set of recommendations on how organizations should begin to build or enhance their digital preservation activities. to ensure secure, long-term integrity to its digital collections. Based on the current digital preservation standards, UNTL will apply the NDSA levels as a way of identifying gaps and improving its digital preservation infrastructure.
### 8.1 NDSA Levels of Preservation

<table>
<thead>
<tr>
<th>Table 1: Version 1 of the Levels of Digital Preservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage and Geographic Location</td>
</tr>
<tr>
<td>- Two complete copies that are not collocated</td>
</tr>
<tr>
<td>- For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system</td>
</tr>
<tr>
<td>Level 1 (Protect your data)</td>
</tr>
<tr>
<td>Level 2 (Know your data)</td>
</tr>
<tr>
<td>- At least three complete copies</td>
</tr>
<tr>
<td>- At least one copy in a different geographic location</td>
</tr>
<tr>
<td>- Document your storage system(s) and storage media and what you need to use them</td>
</tr>
<tr>
<td>Level 3 (Monitor your data)</td>
</tr>
<tr>
<td>- At least one copy in a geographic location with a different disaster threat</td>
</tr>
<tr>
<td>- Obsolescence monitoring process for your storage system(s) and media</td>
</tr>
<tr>
<td>Level 4 (Repair your data)</td>
</tr>
<tr>
<td>- At least three copies in geographic locations with different disaster threats</td>
</tr>
<tr>
<td>- Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>File Fixity and Data Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Check file fixity on ingest if it has been provided with the content</td>
</tr>
<tr>
<td>- Create fixity info if it wasn’t provided with the content</td>
</tr>
<tr>
<td>- Check fixity on all ingests</td>
</tr>
<tr>
<td>- Use write-blockers when working with original media</td>
</tr>
<tr>
<td>- Virus-check high risk content</td>
</tr>
<tr>
<td>Level 1 (Protect your data)</td>
</tr>
<tr>
<td>Level 2 (Know your data)</td>
</tr>
<tr>
<td>- Check fixity of content at fixed intervals</td>
</tr>
<tr>
<td>- Maintain logs of fixity info; supply audit on demand</td>
</tr>
<tr>
<td>- Ability to detect corrupt data</td>
</tr>
<tr>
<td>- Virus-check all content</td>
</tr>
<tr>
<td>Level 3 (Monitor your data)</td>
</tr>
<tr>
<td>- Check fixity of all content in response to specific events or activities</td>
</tr>
<tr>
<td>- Ability to replace/repair corrupted data</td>
</tr>
<tr>
<td>- Ensure no one person has write access to all copies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Information Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Identify who has read, write, move and delete authorization to individual files</td>
</tr>
<tr>
<td>- Restrict who has those authorizations to individual files</td>
</tr>
<tr>
<td>- Document access restrictions for content</td>
</tr>
<tr>
<td>Level 1 (Protect your data)</td>
</tr>
<tr>
<td>Level 2 (Know your data)</td>
</tr>
<tr>
<td>- Maintain logs of who performed what actions on files, including deletions and preservation actions</td>
</tr>
<tr>
<td>Level 3 (Monitor your data)</td>
</tr>
<tr>
<td>- Perform audit of logs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Inventory of content and its storage location</td>
</tr>
<tr>
<td>- Ensure backup and non-collocation of inventory</td>
</tr>
<tr>
<td>- Store administrative metadata</td>
</tr>
<tr>
<td>- Store transformative metadata and log events</td>
</tr>
<tr>
<td>- Store standard technical and descriptive metadata</td>
</tr>
<tr>
<td>Level 1 (Protect your data)</td>
</tr>
<tr>
<td>Level 2 (Know your data)</td>
</tr>
<tr>
<td>- Storage standard preservation metadata</td>
</tr>
<tr>
<td>Level 3 (Monitor your data)</td>
</tr>
<tr>
<td>- Store standard preservation metadata</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>File Formats</th>
</tr>
</thead>
<tbody>
<tr>
<td>- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs</td>
</tr>
<tr>
<td>- Inventory of file formats in use</td>
</tr>
<tr>
<td>- Monitor file format obsolescence issues</td>
</tr>
<tr>
<td>- Perform format migrations, emulation and similar activities as needed</td>
</tr>
</tbody>
</table>

---

National Digital Stewardship Alliance (NDSA) (2012). NDSA Levels of Preservation
Retrieved from
[http://www.digitalpreservation.gov/ndsa/activities/levels.html](http://www.digitalpreservation.gov/ndsa/activities/levels.html)
9. Roles and Responsibilities

UNTL has identified the following stakeholder categories for the digital preservation program. The terminology is adapted from the OAIS Reference Model (CCSDS 650.0-M-2 (2012)).

- **Producer:** is the role played by those persons or client systems that provide the information to be preserved. Producers include faculty, students, staff, alumni, collectors, creators of content, publishers, and others. Producers will be responsible for complying with established deposit requirements and working with the management of the digital archive to ensure a successful transfer. (expanded OAIS definition)
- **Management:** is the role played by those who set overall OAIS policy as one component in a broader policy domain, for example as part of a larger organization. UNTL’s Deans’ Council will be responsible for setting digital preservation policies and integrating them into broader organizational contexts. (expanded OAIS definition)
- **Administrators:** encompass content stewards (designated staff responsible for selection and for ongoing curation of specific collections), digital preservation specialists, and working teams (see appendix for list). Administrators will be responsible for the establishment of the digital preservation program and for day-to-day management of the digital archives. [Note: OAIS uses Administration Functional Entity: The OAIS functional entity that contains the services and functions needed to control the operation of the other OAIS functional entities on a day-to-day basis.]
- **Co-operating Archives:** includes those Archival Institutions that have Designated Communities with related interests. They may order and ingest data from each other. At a minimum, co-operating Archives must agree to support at least one common Submission Information Package (SIP) and Dissemination Information Package (DIP) for inter-Archive requests. At UNTL we think of this group as collaborators. Examples include: Texas Digital Library (TDL).
- **Consumer:** represents the role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISes, as well as internal OAIS persons or systems.
- **User Groups / Client Groups:** include the various types of clients who use the UNT Libraries’ Digital Collections

10. Collaboration/Cooperation

UNTL acknowledges digital preservation as a shared community responsibility, and as such has long-standing and emerging partnerships with similarly committed organizations (e.g., TDL and MetaArchive) and is committed to collaborating with other institutions, in addition to units internal to UNT in order to:
• Advance the development of the UNTL digital preservation program.
• Share lessons learned with other digital preservation programs.
• Extend the breadth of our available expertise.
• Extend the digital content that is available within a broad information community to UNTL users through cooperative efforts.

Generally, in working, cooperating, and collaborating with others, UNTL will strive to:

• Understand the goals, objectives, and needs of the communities of creators and the communities of consumers of its digital resources.
• Identify appropriate partners and stakeholders to contribute to national and international efforts in digital preservation.
• Help develop national and international strategies and initiatives that enable the distribution of collection, description, service delivery, digitization, and preservation activity.
• Work actively with creators of digital materials to encourage and promote standards and practices.

11. Access and Use Criteria

UNTL acquires, manages, and preserves digital resources so that they remain accessible to its constituents over the long term. Certain limitations may be placed on access due to legal, donor, and/or other restrictions; generally, to the extent that it is possible, UNTL endeavors to make its digital resources accessible to all users. Additional restrictions may apply to usage of items. See the UNT Libraries’ Digital Collections Usage and Feedback Policy, sections 2. Access and 3. Rights/Terms of Use for complete guidelines.

12. Implementation

Implementation of this policy framework is contingent upon the infrastructure (technical and human resources) provided by UNT and UNTL, the availability of cost-effective solutions, the adoption of standards, and the evolution of best practices and procedures.

13. Review Cycle

This policy will be reviewed and updated as needed with a full review every two years to assure timely revisions as technology progresses and preservation strategies and experience mature.
Appendix 1: Glossary

**Access:** The process for the retrieval of data and information from storage media, through the use of catalogs, indexes, and/or other tools.

**Acquire:** To take physical and legal custody of data and information.

**Analog:** Data and information in a format that must be digitized to make it digitally accessible.

**Aubrey:** Framework developed by the UNT Libraries to provide end user access to digital collections in for the University of North Texas Libraries.

**Bit-Level Preservation:** Minimum digital preservation standard; the goal is to maintain the integrity of the original bit-stream of a digital object. It is accomplished by maintaining backup copies (onsite and/or offsite), the periodic refreshing of those copies to new storage media, and conducting fixity checking.

**Born-Digital:** Data and information created and maintained in a digital format.

**Coda:** Archival management system created by the UNT Libraries for registering, storing, replicating, and verifying fixity of preservation files.

**Collection:** A group of materials assembled by a person and/or organization, with one or more unifying characteristic.

**Curation:** Activities related to managing data and information throughout its lifecycle, ensuring that data are properly appraised, selected, and securely stored, while appropriately maintaining logical and physical integrity and authenticity. Further, that data is made and remains accessible and viable in subsequent technology environments.

**Data Sets:** Collections of data. The data formats include, but are not limited to, flat file tabular data, relational databases, text corpora, qualitative data in field notes, scholarly editions and thematic research collections.

**Digital Assets:** Digital objects (e.g., text, image, audio-visual files) owned or managed by an institution (or person).

**Digital Collections:** The collective digital library interfaces operated by the UNT Digital Libraries comprising The Portal to Texas History, the UNT Digital Library, and The Gateway to
Oklahoma History. When used in the capitalized form it is meant to refer to the UNT Libraries’ Digital Collections.

**Digital Object:** An entity in which one or more content files and their corresponding metadata are united, physically and/or logically, through the use of a digital wrapper.

**Digital Preservation:** Comprehensive set of managed activities that are necessary to provide continued access to digital objects, beyond the limits of media failure or technology change. At minimum it should include bit-level preservation.

**Digital Rights Agreement:** A legal document that provides the UNT Libraries a non-exclusive license to preserve and provide access to resources owned by the rights owner for digital content in the UNT Libraries’ Digital Collections.

**Digital Wrapper:** A technology that encapsulates administrative, technical, preservation, descriptive, or structural metadata and possibly content files into a specified format for serialization or transmission. An example of this is the Metadata Encoding & Transmission Standard (METS).

**Digitized Materials:** Analog materials that have been transformed into digital form, especially for storage, access and use in a computer environment.

**EUID:** “Enterprise UserID” is the local terminology for the unique identifier given to students, faculty and staff at the University of North Texas as part of the authentication framework in use by the university.

**Institutional Records:** Data or information in a fixed form, regardless of medium, that is created or received in the course of institutional activities and maintained as evidence of that activity for future reference.

**Object Fixity:** The quality of a digital object to be stable and resist change.

**On-Campus:** Physical UNT locations in Denton including the main campus and Discovery Park.

**Partners:** Person(s) or organization(s) that contribute materials to the UNT Libraries’ Digital Collections, and who are usually the content owners or rights holders for the materials.

**Partnership Agreement:** A legal document that defines the relationship between the UNT Libraries and a contributing partner as well as the various rights each retains to the digital content in the UNT Libraries’ Digital Collections.

**Provenance:** Information regarding the origins, custody, and ownership of an item or collection. (Richard Pearce-Moses, A Glossary of Archival and Records Terminology, Society of American Archivists 2005)

**Reformatting:** The process of creating a copy with a format or structure different from the original, for preservation and/or access; this may be accomplished via, transcription, xerography, microfilming, and/or digitization.

**Scholarship:** Use of intellectual resources, such as those managed by the UNT Libraries, for research, teaching, and learning.

**Stewardship:** The responsibility for overseeing and protecting materials and assets.

**Sustainable Access:** The process of providing long-term resources (fiscal, human, and technological) necessary to maintain access to information and digital objects in a repository.

**Trusted Digital Repository:** A repository whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. “Trustworthiness” should be quantifiable via an assessment tool applied to a repository conceptual model (e.g. TRAC / ISO 16363 for OAIS Reference Model).

**University of North Texas Libraries (UNTL):** The library managing the digital collections discussed throughout the Digital Preservation policy documentation.

**UNT Community:** Individuals and entities currently affiliated with the University of North Texas that have a valid/active EUID.

**UNT Digital Libraries:** Administrative unit in the UNT Libraries that manages digital and web-based systems.

**UNT Extended Community:** UNT students, faculty, staff, alumni, and administrators.
Appendix 2: Sources Consulted

The following resources were consulted in the development of this policy framework:

Cornell University Library Digital Preservation Policy Framework (December 2004)
Ohio State University Digital Preservation Policy Framework (August 2013)
Reference Model for an Open Archival Information System (OAIS), Recommended Practice, CCSDS 650.0-M-2 (Magenta Book) Issue 2, June 2012
Society of American Archivists A Glossary of Archival and Records Terminology, 2005
Timetable for Document Review

Date: October 2015

Version: 1.0

Contributors:

Daniel Alemneh Supervisor, Digital Curation Unit
Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Introduction

In order to stay relevant, technical documents must be periodically evaluated to ensure that they meet the needs of the organization or project. Due to this practice, certain UNT Digital Libraries' documents are subject to internal review and approval by various Libraries' units as well as University and external review bodies. This review timetable document compiles major documents that require periodic reviews and gives timetable and the framework over the documents review procedures.

<table>
<thead>
<tr>
<th>No</th>
<th>Document</th>
<th>Review frequency</th>
<th>Year Adopted</th>
<th>Year scheduled for review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TRAC/ISO16393 Conformance Document</td>
<td>Every three years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>Formal Self Audit</td>
<td>Every five years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>Full Audit of Resources in Repository</td>
<td>Every two years</td>
<td>2014</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Digital Preservation Policy Framework</td>
<td>Every two years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td>Digital Collection Development Policy</td>
<td>Every two years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>Digital Library Feedback Policy</td>
<td>Every two years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>7</td>
<td>Preferred File Formats</td>
<td>Every two years</td>
<td>2015</td>
<td>X</td>
</tr>
<tr>
<td>8</td>
<td>UNT Digital Libraries: Risk Analysis and Management Strategy Plan</td>
<td>Every two years</td>
<td>2015</td>
<td>X</td>
</tr>
</tbody>
</table>
UNT Libraries’ Digital Collections Usage and Feedback Policy

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Note: This is an archived version of this document.
An up-to-date version if available at: http://www.library.unt.edu/policies/
UNT Libraries’ Digital Collections Usage and Feedback Policy

1. Scope

This document describes the policies governing the interaction of external users with materials in the UNT Libraries’ (UNTL) Digital Collections.

2. Access

UNTL acquires, manages, and preserves digital resources so that they remain accessible to its constituents over the long term. Although the UNTL supports open access and endeavors to make its digital resources accessible to all users, some items may have certain limitations on access due to legal, donor, and/or other restrictions.

The designation of access for a resource does not imply that the resource is in the Public Domain as many of the items in the UNT Libraries’ Digital Collections are still covered under copyright. It is up to the user to investigate the usage limitations for items they are interested in reusing for the UNT Libraries’ Digital Collections.

2.1 Categories of Access

The four categories of access currently utilized by the UNT Libraries’ Digital Collections include:

Public Access: Resource is available via the Web to the public around the world with no limitations on who can access the resource. (All public domain items and Federal Government Documents fall into this category.)

UNT Community: Resource is available to anyone on the UNT Denton Campus, or by a current UNT Community member with use of a valid EUID when off campus.

UNT Community Strict: Resource is available only to current UNT Community members who are required to authenticate using their EUID before each use, no matter their physical location.

Physical Premises: Resources do not require authentication, but are only available in the physical buildings of the UNT Libraries, on computers connected to the Local Area Network (note WiFi cannot connect to these resources).
2.2 Embargoes

In some cases, materials are included in the UNT Libraries' Digital Collections under embargo, which is essentially a pre-release state. While the embargo is in effect, the digital resource is not available to any users, regardless of location or affiliation. When the embargo date passes, the digital resource automatically becomes available under the stipulated access level. Generally, materials become publicly accessible when the embargo expires, however, it is possible for items to have restricted access and be available to only the specified user group at the end of the embargo period.

Whenever possible, contact information will be included so that users can request a copy directly from the creator(s) while the embargo is in effect. However, we cannot guarantee that contact information is correct, or that a copy will be provided.

3. Rights/Terms of Use

All materials in the UNT Libraries’ Digital Collections are subject to specific terms of use.

3.1 Copyright Statement

Items in the UNT Libraries’ Digital Collections are protected by copyright, with all rights reserved, unless otherwise indicated. The contents of the UNT Libraries’ Digital Collections are made publicly available by our contributing partners for use in research, teaching, and private study.

Although the nature of archival and manuscript collections sometimes makes it difficult to determine the copyright status of an item, it is the end user’s responsibility to use it according to all applicable terms. Please contact the contributing partner for additional information regarding copyright status of a particular digital image, text, data set, or sound or video recording. The UNT Digital Libraries will make contact information for contributing partners publicly available.

3.2 Conditions of Use

By using the digital images, texts, data sets, audio and video recordings, and other materials in the UNT Libraries’ Digital Collections, you, as user, agree to follow these conditions of use:

- UNT Libraries does not have authority to grant or deny permissions to use images or content from resources that are in the public domain or items covered under a Creative Commons License.
- Responsibility for any use of these materials rests exclusively with the user.
- Some materials in these collections may be protected by the U.S. Copyright Law (Title 17, U.S.C.). In addition, the reproduction of some materials may be restricted by terms of use.
gift or purchase agreements, donor restrictions, privacy and publicity rights, licensing and trademarks. Transmission or reproduction of materials protected by copyright beyond that allowed by copyright law requires the written permission of the copyright owners.

- When using items from the UNT Libraries’ Digital Collections, users should give proper credit to the contributing partner and the UNT Libraries’ Digital Collections system from which they retrieved the item.

4. User Privacy

4.1 Surveys and Assessments

Information and data obtained by the UNT Libraries in support of assessment of services, collections, resources, etc., or in support of research related to depository services, are considered confidential and will not be shared except in aggregations or with the express permission of participants, to protect privacy.

4.2 Logging

The UNT Libraries’ Digital Collections access systems (i.e., The Portal to Texas History, the UNT Digital Library, and The Gateway to Oklahoma History) frequently track or “log” the actions performed by users of those systems using mechanisms common to Web server technologies.

Information from transaction logs are aggregated for reporting on types of use and use of materials. For this purpose, information regarding individual identities are removed.

These transaction log files are retained as a research dataset by the UNT Libraries for use in research, system modelling, or the identification of errors. If a valid research request for these log files is accepted by the UNT Libraries, only an IP anonymized version of the files is provided.

4.3 Authentication

The UNT Libraries’ Digital Collections makes use of the Lightweight Directory Access Protocol (LDAP) as a mechanism to allow individuals who are affiliated with UNT to authenticate into the UNT Digital Library.

4.4 Google Analytics

The UNT Libraries’ Digital Collections uses Google Analytics, a Web analytics service provided by Google, Inc. (“Google”) to help understand how the site is being used and to improve our interface and services. Google Analytics operates through the use of a “cookie,” which is a text
file placed on a user’s computer that contains information about their use of the UNT Libraries’ Digital Collections. The information stored in the cookie (including IP addresses) will be transmitted to and stored by Google. The UNT Libraries’ Digital Collections uses this information for analytical and feature-improvement-related purposes only. We do not transmit any information to other third parties. For more information on Google Analytics and Google’s privacy policy, please visit http://www.google.com/intl/en/privacy.html.

A user can choose to opt out of this tracing by using one of the following methods:
- Turn off cookies in the browser preferences setting
- Install the Google Analytics opt-out browser add-on http://tools.google.com/dlpage/gaoptout

4.5 Personally identifiable information in items in the repository
See Redaction Policy below.

5. Redaction Policy

When UNT Libraries is notified of personally-identifying information within a digital object, steps will be taken depending on the type of information in question.

Objects containing specific types of personally-identifying information will be immediately suppressed from public view, and the information will be edited out of the section of the document where it occurs, e.g.: social security numbers and drivers’ license numbers.

For requests to redact information that is not the above type of information, UNT Libraries will direct the person requesting redaction to the owners of the original content, as specified in the Partnership Agreement, Digital Rights Agreement, MOU, or other contract. In such cases, UNT has been granted permission within one of these agreements to digitize, host, and make openly accessible this information, and it is by permission of the rights owner that the request will be honored.

In the event that the rights owners grants permission to redact this information the UNT Libraries will suppress the item from public view while a new derivative of the item is processed which has the information redacted. Upon ingest of the new derivative into the system, the revised item will be made available again. UNT Libraries redacts the minimal amount of data to maintain the integrity of the original object.

As a result, there is some content that UNT Libraries does not redact, such as:
6. Intellectual Property or Copyright Challenge Policy

If a digital object is identified as being online without sufficient rights provided to the UNT Libraries by the rights holder, the UNT Libraries will work with the rights holder to remedy the situation. The expected steps involved in the process include:

- UNTL is alerted to a copyright or intellectual property challenge
- UNTL suppresses the record and access to the digital object during the review period
- A review of the resource, the partnership or rights agreements associated with the item or collection, and the issues brought up by the rights holder is conducted by UNTL staff
- UNTL communicates with the rights holder the findings of the internal review, and if in error works with the rights holder to hopefully restore access to the resource. If an agreement is not made, then the resource remains suppressed from public access.

7. User Feedback

The UNT Digital Libraries receives a variety of comments, questions, and requests through various means, including:

- Forms submitted from The Portal to Texas History, UNT Digital Library, or Gateway to Oklahoma History websites
- Forms submitted through the UNT Libraries website, such as the “Ask Us” form
- Direct contact via e-mail or phone calls with staff members

7.1 Responses

UNT Digital Libraries staff members answer feedback questions as received, within a reasonable amount of time, dependent on staff availability. In some cases, UNT Digital Libraries staff members may forward queries or comments to more appropriate subject experts instead.

Staff members will not answer questions that appear to be automated spam, questions that do not include an e-mail address for reply, or questions that contain inciting comments rather than relevant questions.
All requests for permission to use materials from the UNT Libraries' Digital Collections or to acquire high-resolution copies of image files must be directed to the contributing partner (see Rights/Terms of Use above).

7.2 Examples of Services Not Offered by Digital Libraries Staff

The UNT Digital Libraries do not:

- Provide access to embargoed or restricted materials to persons outside the designated community
- Provide appraisals for materials
- Provide research assistance including genealogy research, or other services that do not directly pertain to the maintenance or usage of the UNT Libraries' Digital Collections
- Provide technical support or guarantee the functionality of sites that are not operated or maintained by the UNT Digital Libraries
- Directly sell copies (digital or physical) of materials in the UNT Libraries' Digital Collections
  - In some specific instances, the UNT Digital Libraries provide links to third-party services that sell a particular item, however, UNT does not maintain responsibility for those sites
- Buy physical or digital items from individuals for inclusion in the UNT Libraries' Digital Collections
- Provide legal advice in any form

8. Review Cycle

This policy will be reviewed and updated as needed with a full review every two years to assure timely revisions as technology progresses, expectations related to access, and digital library collections mature.
Digital Projects Unit Standards

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Daniel Alemneh Supervisor, Digital Curation Unit
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Digital Projects: Standards

**Scanning Standards by Type of Material**

- Text (Books)
- Documents (Handwritten, Ledgers, Receipts)
- Photographs (Prints)
- Photographs (Negatives and Slides)
- Large Format (Posters and Maps) Above 17 x 24 inches (A2)
- Newspapers
- Moving Image
- Audio

Digital image collections incorporated into the UNT Libraries’ permanent collections must adhere to the minimum requirements given below. Depending on the specific objectives of the projects, higher resolutions may be desirable for some materials. Please note that this document is not intended to address all technical issues and does not describe the details of operational procedures. See project-specific guidelines for further information.

### Text (Books)

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (ppi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>B&amp;W Text Only</td>
<td>1-bit</td>
<td>Bitonal</td>
<td>600ppi</td>
<td>100%</td>
<td>Tiff (uncompressed or lossless compression, no LZW)</td>
</tr>
<tr>
<td>B&amp;W Text with Illustrations</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>400ppi</td>
<td>100%</td>
<td>Tiff (uncompressed)</td>
</tr>
<tr>
<td>Text w/ Continuous-Tone Images &amp; Photos (Color)</td>
<td>24-bit</td>
<td>RGB</td>
<td>400ppi</td>
<td>100%</td>
<td>Tiff (uncompressed)</td>
</tr>
</tbody>
</table>

### Documents (Handwritten, Ledgers, Receipts)

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (ppi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>B&amp;W Text Only*</td>
<td>1-bit</td>
<td>Bitonal</td>
<td>600ppi</td>
<td>100%</td>
<td>Tiff (uncompressed or lossless compression no LZW)</td>
</tr>
<tr>
<td>Handwritten Documents, Ledgers</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>400ppi</td>
<td>100%</td>
<td>Tiff (uncompressed)</td>
</tr>
<tr>
<td>Correspondence, Letters</td>
<td>24-bit</td>
<td>RGB</td>
<td>400ppi</td>
<td>100%</td>
<td>Tiff (uncompressed)</td>
</tr>
</tbody>
</table>

* We recommend scanning handwritten documents in grayscale or color.
Photographs (Prints)

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (ppi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>B&amp;W Photographs</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>600ppi</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
<tr>
<td>Color Photographs</td>
<td>24-bit</td>
<td>RGB</td>
<td>600ppi</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
</tbody>
</table>

- As of April, 2009, we scan both sides of a photograph even if one side is blank.

Photographs (Negatives and Slides)

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (ppi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>B&amp;W Negatives &amp; Slides</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>5000 pixels across the longest side*</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
<tr>
<td>Color Negatives &amp; Slides</td>
<td>24-bit</td>
<td>RGB</td>
<td>5000 pixels across the longest side*</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
</tbody>
</table>

- For copy negatives, scan the entire negative. Do not crop to the original image. Do not crop out written notes.

Large Format (Posters and Maps) Above 17 x 24 inches (A2)

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (ppi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>B&amp;W Maps/Posters</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>5000 - 10000 pixels across the longest side</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
<tr>
<td>Color Maps/Posters</td>
<td>24-bit</td>
<td>RGB</td>
<td>5000 - 10000 pixels across the longest side</td>
<td>100% (1:1)</td>
<td>Tiff (uncompressed)</td>
</tr>
</tbody>
</table>

Newspapers

Newspapers should be scanned and digitally preserved in accordance with the national standards (http://www.loc.gov/ndnp/guidelines/) set by the Library of Congress for the National Digital Newspaper Program.

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Bit Depth</th>
<th>Color Space</th>
<th>Resolution (dpi)</th>
<th>Scale</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microfilm</td>
<td>8-bit</td>
<td>Grayscale</td>
<td>400*</td>
<td>100% (1:1)</td>
<td>Tiff 6.0 (uncompressed)</td>
</tr>
<tr>
<td>Physical Pages</td>
<td>24-bit</td>
<td>RGB</td>
<td>400</td>
<td>100% (1:1)</td>
<td>Tiff 6.0 (uncompressed)</td>
</tr>
</tbody>
</table>

- Monitor equipment performance by capturing a standards-based target film strip or color image at the start of each scanning session.
- Split dual images into individual newspaper images as necessary.
- Deskew images with more than 3% skew.
- Crop page image files to the edge of the newspaper, retaining the original edge and up to a quarter inch beyond.
- Produce images that have exactly the same dimensions, spatial resolution, skew, and cropping as the images used for OCR.

*For microfilm, scan at 8-bit grayscale with a resolution of 400 dpi, if possible; otherwise between 300 and 400 dpi (relative to the size of the original newspaper).
## Moving Image (Film and Video)

<table>
<thead>
<tr>
<th>Original Format</th>
<th>File Format</th>
<th>Codec</th>
<th>Audio</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 mm film</td>
<td>QuickTime (.mov)</td>
<td>Uncompressed YUV 10-bit 4:2:2</td>
<td>48 kHz, 16-bit</td>
<td>Standard Definition (720x480)</td>
</tr>
<tr>
<td>VHS, SVHS, Betamax, Betacam, Umatic, DVCAM, DVC Pro, Mini DV, Laser Disc</td>
<td>QuickTime (.mov)</td>
<td>DV50</td>
<td>48 kHz, 16-bit</td>
<td>Standard Definition (720x480)</td>
</tr>
</tbody>
</table>

## Audio

<table>
<thead>
<tr>
<th>Original Format</th>
<th>Bit Depth</th>
<th>Resolution</th>
<th>File Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassette</td>
<td>16-bit</td>
<td>44.1 kHz</td>
<td>Waveform Audio File Format (WAVE) .wav</td>
</tr>
<tr>
<td>Reel-to-Reel, Analog Disk</td>
<td>24-bit</td>
<td>48 kHz to 96 kHz</td>
<td>Waveform Audio File Format (WAVE) .wav</td>
</tr>
</tbody>
</table>
Preferred File Formats

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Note: This is an archived version of this document.
An up-to-date version if available at:
http://www.library.unt.edu/digital-projects-unit/digital-file-formats
Prefered File Formats

Introduction

This document provides a general overview of the file types produced by the activities of the Digital Libraries Division, or accepted for inclusion in the UNT Libraries' Digital Collections.

Preferred File Formats

These are commonly-accepted file formats for which we intend to maintain bitstream preservation in addition to preserving the functionality and content of the files.

<table>
<thead>
<tr>
<th>MimETYPE</th>
<th>Common Accepted Extensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>image/tiff</td>
<td>.tif, .tiff</td>
</tr>
<tr>
<td>image/jpeg</td>
<td>.jpg, jpeg</td>
</tr>
<tr>
<td>application/pdf</td>
<td>.pdf</td>
</tr>
<tr>
<td>text/plain</td>
<td>.txt</td>
</tr>
<tr>
<td>text/html</td>
<td>.htm, .html</td>
</tr>
<tr>
<td>application/x-troff-msvideo</td>
<td>.avi</td>
</tr>
<tr>
<td>video/avi</td>
<td></td>
</tr>
<tr>
<td>video/msvideo</td>
<td></td>
</tr>
<tr>
<td>video/avs-video</td>
<td></td>
</tr>
<tr>
<td>video/quicktime</td>
<td>.mov</td>
</tr>
<tr>
<td>audio/wav</td>
<td>.wav, .wave</td>
</tr>
<tr>
<td>audio/x-wav</td>
<td></td>
</tr>
<tr>
<td>audio/aiff</td>
<td>.aiff, .aif</td>
</tr>
<tr>
<td>audio/x-aiff</td>
<td></td>
</tr>
<tr>
<td>application/octet-stream</td>
<td>.arc, .arc.gz</td>
</tr>
<tr>
<td>application/warc</td>
<td>.warc, .warc.gz</td>
</tr>
<tr>
<td>image/png</td>
<td>.png</td>
</tr>
</tbody>
</table>
Recognized File Formats

These file formats are less preferred. Although we still preserve content and files at the bit level, functionality may be lost over time.

<table>
<thead>
<tr>
<th>Mimetype</th>
<th>Common Accepted Extensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion Picture Entertainment Group</td>
<td>video/mpeg, audio/mpeg</td>
</tr>
<tr>
<td>Flash Video Format</td>
<td>video/x-flv, .flv</td>
</tr>
<tr>
<td>MPEG Group4</td>
<td>video/mp4, .mp4, .m4v,</td>
</tr>
<tr>
<td>Windows Media</td>
<td>video/x-ms-wmv, .wmv</td>
</tr>
<tr>
<td>MPEG Group4</td>
<td>audio/mpeg3, audio/x-mpeg-3, video/mpeg, video/x-mpeg</td>
</tr>
</tbody>
</table>

Standard Normalization Paths

For some born-digital file types, we create a set of derivatives to display the contents in a more standardized way. Although we retain all file types and derivatives, they may not all be visible or accessible from the public interface.

<table>
<thead>
<tr>
<th>Original File</th>
<th>Intermediary File(s)</th>
<th>Display/Public Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word document (.doc, .docx)</td>
<td>.pdf</td>
<td>.jpg</td>
</tr>
<tr>
<td>Portable Document Format (.pdf)</td>
<td>&gt;&gt;</td>
<td>.jpg</td>
</tr>
<tr>
<td>Powerpoint (.ppt, .pptx)</td>
<td>.pdf</td>
<td>.jpg</td>
</tr>
<tr>
<td>Spreadsheet (.xls, .xlsx)</td>
<td>&gt;&gt;</td>
<td>.csv</td>
</tr>
<tr>
<td>OpenDocument file: presentation</td>
<td>.pdf, .ppt</td>
<td>.jpg</td>
</tr>
<tr>
<td>text document</td>
<td>.pdf, .doc</td>
<td></td>
</tr>
</tbody>
</table>
UNT Libraries OAIS Information Package Specification

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Hannah Tarver Department Head, Digital Projects Unit
Ana Krahmer Supervisor, Digital Newspaper Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Introduction

The UNT Libraries operate a digital repository system built around the functional components designated by the Open Archival Information Systems (OAIS) Reference Model\(^1\). As such, this requires a formal definition of the specification of the Submission Information Package (SIP), Archival Information Package (AIP), and Dissemination Information Package (DIP) accepted and processed by the UNT Libraries Digital Library Infrastructure.

Because the SIP, AIP, and DIP are closely related in the UNT Libraries' Digital Collections infrastructure, this document is used to describe all three packages. When all three of the formats are discussed together they are referred to as Information Packages (IPs).

Historically the UNT Libraries have referred to the Dissemination Information Packages (DIPs) for the Digital Collections as Access Content Packages (ACP). In the document below there are references to ACPs that can and should be equated to DIPs when mapping these concepts to the OAIS Reference Model.

Components

There are a number of core components and technologies used by the UNT Libraries to package, transfer, verify, and track Information Packages throughout the system. This section enumerates these components and discusses how each is used for packaging and transferring Information Packages throughout the repository.

BagIt

The UNT Libraries' Information Package (IP) formats make use of the BagIt\(^2\) bag container as a way of encapsulating content submitted and stored by the repository; the container functions as a mechanism for checking for the well-formedness and validity of a resource housed within the bag. Many tools have been built in most popular programming languages that allow for individuals to build, change, and validate BagIt bags. In addition, the BagIt format has a mechanism for storing basic metadata information about the BagIt bag itself such as Internal and External Identifiers, descriptions of the resource(s) encapsulated by the BagIt bag, and the person or body responsible for the BagIt bag.

The UNT Libraries' IP profile assumes bags are based on the 0.97 version of the BagIt Specification.

NaMaste File

The UNT Libraries use the NaMaste specification within the IP format to provide a quick human-readable indicator that a given bag is an Information Package, and to verify that it isn’t
another kind of BagIt nag that may not be intended as a UNTL Information Package. The UNT Libraries also use the existence of this file as a “sanity check” for the ingest process of IPs into the repository system; i.e., if the NaMaste file is not present and of the correct format then it is quickly handled as “badly formed” IP and moved into a quarantine space with error information explaining why ingest failed for the resource.

The UNT Libraries’ SIP profile requires the existence of a NaMaste file named “UNTL-SIP-1.0” in the root of the BagIt bag. Internal text inside this file is the following string “until aip 1.0”

The UNT Libraries’ AIP profile expects the existence of a NaMaste file called “UNTL-AIP-1.0” in the root of the BagIt bag. Internal text inside of this file consists of “what does it say inside.”

Coda Directives

The UNT Libraries pass processing information to both the SIP-to-AIP workflow and the AIP-to-ACP workflow using a configuration file called “coda_directive.py”. The Coda Directives include information related to how the SIP-to-AIP conversion should proceed. For example, there may be instructions about whether the processor should assume a standard file-naming convention locally referred to as “MagickNumbers,” or if the processor should assume a logical grouping of files into a file bundle based on like filenames, once extensions have been removed. Each of these local bundles are referred to as a “fileSet.” Directives can be set for each of the “manifestations” within the digital object folder. Further instructions regarding the conversion from AIP-to-ACP may be included, e.g., if a file should be tiled for a zoom interface.

An example coda-directives.py file is included in Appendix 1.

Bag-info.txt

While the bag-info.text file is a part of the BagIt specification, the UNT Libraries have specific assumptions in mind when creating these files. Information included in this file should be sufficient to identify what project or event was responsible for this object, and/or that the object is part of the general collection building process of the UNT Libraries. If it is collection- or project-based, the bag-info.txt should include basic information about the contributor and project that were responsible for creating the project. While not expected to be exhaustive, statements that can apply to all of the files held within the object (e.g., object consists of TIFF images and master OCR files processed with PrimeOCR) might be helpful in the future. The oxum field must be present with an oxum value that represents the contents in the /data/ directory.

An example bag-it.txt file is included in Appendix 2.
Organizational Concepts

The UNT Libraries’ Information Packages and the core data model used by the Aubrey access system assume that digital objects are either created or packaged in a way that can be easily understood by other systems. This section describes the organizational components used in the Information Packages throughout the system.

File

A File is the smallest unit in the Information Package. A File can be of any type, though the SIP-to-AIP conversion expects and handles certain formats as preferred and others as non-preferred formats. File names can consist of letters, numbers, spaces, periods, hyphens and underscores. While there are no specific requirements for file-naming conventions, generally the UNT Libraries encourages content creators to use locally-meaningful file-naming conventions when naming their files, and to try to use standard ASCII character sets whenever possible. It is advisable that file names include a standard extension appropriate to the file format. For example: .tiff, .tif, .pdf, .jpeg, .jpg, etc.

fileSet

A fileSet is a bundle of files grouped logically based on the basename of a filename; these are typically multiple files that represent a single unit. For example, a scan of a book page may be saved as a TIFF file, however, for printed text, digital libraries will create an OCR file for that page and possibly a TXT file of the character information from the OCR file, without coordinate information, that can be used for indexing. In this case we may have three files to represent a single page of text in a book: 0001.tif, 0001.ocr, and 0001.txt. If one takes the basename 0001 and groups the files together using this basename, we have what the UNT Libraries refers to as a fileSet.

These fileSets are useful in the UNTL model for attaching information such as pagination data, annotation data, or labels for video and audio clips. A fileSet designates a primary File which forms the basis for other processing instructions in the ingest workflow. These primary Files are taken from the list of preferred formats\(^3\). The example given above would have “0001.tif” assigned as the primary File in the fileSet by the SIP-to-AIP conversion process.

The coda_directive.py has a directive that allows the creator of the SIP to identify if a resource was created using naming conventions that make use of the concept of fileSets. For other kinds of items (e.g., photographs, or handwritten manuscript pages) the SIP-to-AIP conversion step is directed to process each File it encounters as an individual fileSet.
**Manifest**

The UNT Libraries allow for multiple “representations” of a resource to be stored logically together within the local data model. For example: a single PDF document submitted as an Electronic Theses and Dissertation (ETD) to the repository would be normalized, for preservation purposes, into its constituent parts (or pages). The result is generally a series of image files, with one image per page of the original PDF. In this example there are two manifestations (PDF and image files) that are deposited into the repository in the single IP. Manifestations are designated in the UNT Libraries’ SIP Specification by directories in the root of the /data/ directory in the BagIt bag with a notation expressing order and file types, e.g., 01_jpg, 01_tif, 01_pdf, or in the case of the example above, 01_jpg and 02_pdf. The numeric prefix helps designate the order of manifestation, with the lower number designating the priority for presentation to an end user in one of the Aubrey access systems; the text string denoting the file type is used to provide a cue as to the type of manifestation. This text string should be changed to reflect the needs of the manifestation, but users should strive for consistency over a large set of resources. Common text strings include jpg, tif, pdf, doc, data, raw, and jp2.

For each manifestation there is a section within the coda_directives.py file, which governs how it is processed during the SIP-to-AIP and AIP-to-ACP conversion steps. Additional information notes: if page numbers can be automatically extracted from the filenames because they use a common pattern, if the resource follows standard fileSet groupings based on the file-naming conventions, if the resource consists of preferred file formats, or if there are files in the resource that are not on the list of preferred formats for preservation. Finally, instructions are given for the AIP-to-ACP conversion, such as the need to create tiles for the resource, extra instructions as to the size or resolution of derivative images for the ACP, and finally if the manifestation should be included in the final ACP and exposed via one of the Aubrey access systems.

**METS Record**

A METS (Metadata Encoding and Transmission Standard) record conforming to the UNTL METS Profile is included in both the Archival Information Package (AIP) and the Dissemination Information Package (DIP). This METS record contains metadata about all of the Files in the IP, the role of each File within a fileSet, the sequence of a fileSet within a manifestation, and information such as pagination labels, annotations, audio track names, or video clip names. The sequence of manifestations within the IP is also encoded in the METS record. For AIPs and DIPs there are additional preservation metadata files encoded within the METS record using the PREMIS Data Dictionary in the PREMIS Version 2.0 XML Scheme. The METS record also stores the output of the Unix File command for each File in the IP. Each File in the IP has an associated and referenced JHOVE output file that was created during the SIP-to-AIP conversion process and which is stored as part of the AIP and DIP format.
Requirements
The key words “MUST”, “MUST NOT”, “REQUIRED”, “SHALL”, “SHALL NOT”, “SHOULD”,
“SHOULD NOT”, “RECOMMENDED”, “MAY”, and “OPTIONAL” in this document are to be
interpreted as described in [RFC2119]

An implementation is not compliant if it fails to satisfy one or more of the MUST and REQUIRED
level requirements for the protocols it implements. An implementation that satisfies all the MUST
or REQUIRED level and all the SHOULD level requirements for its protocols is said to be
“unconditionally compliant”; one that satisfies all the MUST level requirements but not all the
SHOULD level requirements for its protocols is said to be “conditionally compliant.”

UNTL-SIP Specification
A UNTL-SIP MUST be a fully valid BagIt bag that can be validated by standard BagIt software.
The UNTL-SIP specification MUST NOT use the “holey bag” features as defined in the BagIt
specification. All files that constitute the digital object MUST be stored in the /data/ folder. At
the time of writing, the UNTL-SIP makes use of MD5 cryptographic hash functions for fixity;
therefore a manifest-MD5.txt file MUST be used for the hash values for the files in the /data/
folder.

A valid NaMaste file called 0=untl_sip_1.0 MUST be in the root of the BagIt bag designating that
the bag is a UNTL-SIP.

A valid coda_directives.py file is REQUIRED with appropriate processing instructions for the
digital object.

A valid bag-info.txt file containing information about the digital object being deposited is
REQUIRED. The oxum field MUST be present with an oxum value that represents the contents
in the /data/ directory.

The contents of the /data/ directory MUST include Manifestation folders, and SHOULD include a
metadata.xml file in the UNTL Metadata format, and an OPTIONAL note.txt file.

Care should be taken to remove hidden or system files (often invisible when processing on
Windows, OSX, or Linux systems) as they will be considered part of the digital object if they are
submitted with the SIP. If included, these files can cause the SIP-to-AIP conversion process to
fail, or may be processed successfully into an AIP and later an ACP in an unintended manner.

Locally the UNT Libraries make use of a Python tool called SIPmaker.py, which helps to
automate the creation and packaging of folders of content into valid UNTL-SIPs.
An example of the UNTL-SIP structure is in Appendix 3.

**UNTL-AIP Specification**

A UNTL-AIP **MUST** be a fully-valid BagIt bag that can be validated by standard BagIt software. The UNTL-AIP specification **MUST NOT** use the “holey bag” features as defined in the BagIt specification. All files that constitute the digital object **MUST** be stored in the /data/ folder. At the time of writing the UNTL-AIP makes use of MD5 cryptographic hash functions for fixity; therefore a manifest-MD5.txt file **MUST** be used for the hash values for the files in the /data/ folder.

A valid NaMaste file called 0=untl_aip_1.0 **MUST** be in the root of the BagIt bag designating that this bag is a UNTL-AIP.

A valid coda_directives.py file is **REQUIRED** with appropriate processing instructions for the digital object.

A valid bag-info.txt file containing information about the digital object being deposited is **REQUIRED**. The oxum field **MUST** be present with an oxum value that represents the contents in the /data/ directory.

Within the /data/ directory there **MUST** be only two directories: a /data/data/ directory containing the contents of the original UNTL-SIP as ingested into the system, and a /data/metadata/ directory containing metadata generated during the ingest process that is stored as part of the UNTL-AIP.

In addition to the /data/data/ and /data/metadata/ directories there is a **REQUIRED** METS metadata record with the filename of <NAME>.mets.xml where the value of <NAME> is the ARK Name assigned to the resource upon ingest into the UNTL Libraries Digital Collections. This ARK Name, when appended to the UNTL Libraries’ Name Assigning Authority M (NAAM), creates a full ARK identifier in the format of ark:/67531/<NAME>. An example of this is ark:/67531/metapth1234

A highly suggested but **OPTIONAL** metadata.xml file in the UNTL metadata format and an optional note.txt file can be stored in the /data/ directory.

No other files or folders than those mentioned above may be stored in the /data/ directory.

Locally the UNTL Libraries make use of a Python tool called makeAIP.py, which automates the creation and packaging of a UNTL-SIP into a valid UNTL-AIP.

An example of the UNTL-AIP structure is in Appendix 4.
**UNTL-DIP (ACP) Specification**

Note: For historical reasons the UNT Libraries refers to its Dissemination Information Package (DIP) using the name Access Content Package (ACP). The specification of the UNTL-DIP is the same as that for the UNTL ACP and they are interchangeable for the purpose of mapping to the OAIS Reference Model.

A UNTL-DIP **MUST** be a fully-valid BagIt bag that can be validated by standard BagIt software. The UNTL-DIP specification **MUST NOT** use the “holey bag” features as defined in the BagIt specification. All files that constitute the digital object **MUST** be stored in the /data/ folder. At the time of writing the UNTL-DIP makes use of MD5 cryptographic hash functions for fixity; therefore a manifest-MD5.txt file **MUST** be used for the hash values for the files in the /data/ folder.

A valid NaMaste file called 0=untl_acp_1.0 **MUST** be in the root of the BagIt bag designating that this bag is a UNTL-DIP.

A valid coda_directives.py file is REQUIRED with appropriate processing instructions for the digital object.

A valid bag-info.txt file containing information about the digital object being deposited is **REQUIRED**. The oxum field **MUST** be present with an oxum value that represents the contents in the /data/ directory.

Within the /data/ directory there **MUST** be only one directory, a /data/web/ directory containing the dissemination or access versions of the originally deposited files that were generated from the UNTL-AIP.

In addition to the /data/data/ and /data/metadata/ directories there is a **REQUIRED** METS metadata record with the filename of <NAME>.mets.xml where the value of <NAME> is the ARK Name assigned to the resource upon ingest into the UNT Libraries’ Digital Collections.

A **REQUIRED** UNTL metadata record in the UNTL metadata format (validates with the UNTL Metadata Schema) with the filename of <NAME>.mets.xml where the value of <NAME> is the ARK Name assigned to the resource upon ingest into the Digital Collections.

No other files or folders than those mentioned above may be stored in the /data/ directory.

Locally the UNT Libraries make use of a Python tool called makeACP.py, which automates the creation and packaging of a UNTL-AIP into a valid UNTL-DIP (ACP).

An example of the UNTL-ACP structure is in Appendix 5.
References

Appendix 1: Example coda_directives.py files

Example 1: coda_directives.py file for a single manifestation item.

```python
""
    these are directives on a per-manifestation basis
    they'll be indexed by pathname relative to the data directory
""
# if certain filetypes need certain Jhove modules, put them here
# jhove_matcher = {
#    (r".*\.log", "ASCII-hul"),
#    (r".*\.cdx", "ASCII-hul"),
#}
manifestation_directives = {
    "01_tif": {
        # Does this manifestation use "magicknumbers" filenaming?
        "use_magick_numbers" : True, # Default False

        # Does this manifestation follow the UNTL fileSet layout?
        "untl_filessets" : False, # Default True

        # Does this manifestation need to be added to the
        # Access Content Package (ACP) or not added?
        "add_to_acp" : False, # Default True

        # What label should be added for this manifestation?
        "label" : "tif", # String used for manifestation label

        # What should the maximum size be for an image across the horizontal?
        "max_width": 1200, # Default 1500

        # Do we want to create Zoomify tiles for this manifestation?
        "make_tiles" : True, # Default False

        # Do we want to absolutely require thumbnails to be created for this manifestation?
        "require_thumbnails" : False, # Default True
    }
}
```

Example 2: coda_directives.py file for a multi-manifestation item.

```python
""
    these are directives on a per-manifestation basis
    they'll be indexed by pathname relative to the data directory
""
```
# if certain filetypes need certain Jhove modules, put them here
#
jhove_matcher = {
  "r".\".log", "ASCII-hul"),
  "r".\".cdx", "ASCII-hul"),
}

manifestation_directives = \{
  "01_tif": {  
    #Does this manifestation use "magicknumbers" filenaming?
    #"use_magick_numbers" : True, # Default False

    #Does this manifestation follow the UNTL fileSet layout?
    #"untl_filessets" : False, # Default True

    #Does this manifestation need to be added to the
    #Access Content Package (ACP) or not added?
    #"add_to_acp" : False, # Default True

    #What label should be added for this manifestation?
    #"label" : "tif", #String used for manifestation label

    #What should the maximum size be for an image across the horizontal?
    #"max_width": 1200, #Default 1500

    #Do we want to create Zoomify tiles for this manifestation?
    #"make_tiles" : True, #Default False

    #Do we want to absolutely require thumbnails to be created for this manifestation?
    #"require_thumbnails" : False, #Default True
  },
  "02_pdf": {  
    #Does this manifestation use "magicknumbers" filenaming?
    #"use_magick_numbers" : True, # Default False

    #Does this manifestation follow the UNTL fileSet layout?
    #"untl_filessets" : False, # Default True

    #Does this manifestation need to be added to the
    #Access Content Package (ACP) or not added?
    #"add_to_acp" : False, # Default True

    #What label should be added for this manifestation?
    #"label" : "pdf", #String used for manifestation label

    #What should the maximum size be for an image across the horizontal?
    #"max_width": 1200, #Default 1500

    #Do we want to create Zoomify tiles for this manifestation?
    #"make_tiles" : True, #Default False

    #Do we want to absolutely require thumbnails to be created for this manifestation?
    #"require_thumbnails" : False, #Default True
  }
}
Appendix 2: Example bag-info.txt file

Example bag-info.txt file for an item.

Bag-Size: 156.09M
Bagging-Date: 2015-03-04
CODA-Ingest-Batch-Identifier: d246388c-c75a-465f-b2da-a44627ed9182
CODA-Ingest-Timestamp: 2015-03-04T21:06:16-0600
Contact-Email: mark.phillips@unt.edu
Contact-Name: Mark Phillips
Contact-Phone: 940-565-2415
External-Description: Newspaper issues from multiple titles funded through the
support of grants awarded to individual partner libraries, historical
societies, and genealogical societies. Content was digitized from microfilm
using NDNP standards by iArchives and processed for inclusion in the Portal to
Texas History. Master files are tiff images with accompanying OCR and bounding
box files.
External-Identifier: ark:/67531/metaph592269
Internal-Sender-Identifier: 1947082201
Organization-Address: P. O. Box 305190, Denton, TX 76203-5190
Payload-Oxum: 163567298.34
Source-Organization: University of North Texas Libraries

Appendix 3: Example SIP structure.

Example SIP file structure for a single manifest item.

Giddings-Box1-Folder3-1871-02-ChappellHill-SM
├── 0-untl_sip_1.0
│    └── bag-info.txt
│    └── bagit.txt
│    └── coda_directives.py
│    └── data
│       └── 01_tif
│          ├── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_01.pro
│          ├── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_01.pro.xml
│          ├── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_01.tif
│          └── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_01.txt
│          └── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_02.pro
│          └── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_02.pro.xml
│          └── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_02.tif
│          └── Giddings-Box1-Folder3-1871-02-ChappellHill-SM_02.txt
│          └── metadata.xml
│ └── manifest-md5.txt

Appendix 4: Example AIP structure.

Example AIP/DIP file structure for a single manifest item.
Appendix 5: Example DIP/ACP structure.

Example Dissemination Information Package (DIP) or Access Content Package (ACP) file structure for a single manifest item.

```
metaphth594985/
  ├── 0-untl_acp_1.0
  │    └── bag-info.txt
  │    └── bagit.txt
  │    └── coda_directives.py
  ├── data
  │    └── mets.xml
  │    └── untl.xml
  └── manifest-md5.txt
```

```
Coda Repository Screenshots

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Coda Repository Screenshots

- Coda Dashboard
- Bag List
- Bag Detail
- URLs to Reconstitute Bag
- ATOM Feed for Bag
- Storage Node List
- Storage Node Detail
- PREMIS Event Listing
- PREMIS Event Detail
- Validate Dashboards
- Validate Entry
- Validate Entry after Prioritization for On-Demand Validation
- Validation Statistics
- Coda Search Interface
- Coda OAI-PMH Interface
- Coda Stats Pages
- Replication Queue
- About and API Documentation
- Atom Feed For Coda
- AtomPub Interface
- ResourceSync Interface
Coda Dashboard
Dashboard

The Coda system acts as a digital archive for items in the UNT Libraries' Digital Collections. This dashboard presents a non-technical overview.

- 1,235,402 Bags
- 263.9 TB Disk Space Used
- 124,012,020 Files
- 4,722,656 PREMIS Events
- 0 Queue Entries
- 0 Validation Entries

Search text within a bag
Bag List
<table>
<thead>
<tr>
<th>Ark ID</th>
<th>Bagged Date</th>
<th>URLs</th>
<th>ATOM</th>
<th>Size</th>
<th># Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>aric/67531/codaq93w</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.2 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq93x</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.4 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq93y</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>41.3 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq93z</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>42.9 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq940</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>31.2 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq941</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.4 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq942</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>41.4 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq943</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>31.8 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq944</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>39.5 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq945</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>29.9 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq946</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.8 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq947</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>38.0 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq948</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>32.1 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq949</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.9 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94a</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>41.2 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94b</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.8 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94c</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>39.4 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94d</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.3 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94e</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>41.4 MB</td>
<td>6</td>
</tr>
<tr>
<td>aric/67531/codaq94f</td>
<td>June 9, 2015</td>
<td>urls</td>
<td></td>
<td>40.3 MB</td>
<td>6</td>
</tr>
</tbody>
</table>
Bag Detail
Bag Info Details:

Payload-Oxum: 48.6 MB, 6 files
Contact-Name: Mark Phillips
Contact-Email: mark.phillips@unt.edu
Bag-Size: 48.65M
Internal-Sender-Identifier: 2012.201.B0370.0626
External-Identifier: ark:/67531/codalbcd
Organization-Address: 800 Nazih Zuhdi Drive, Oklahoma City, OK 73105
Contact-Phone: 940-369-7809
Bagging-Date: 6 months, 3 weeks ago
External-Description: Photographs digitized from the Oklahoma Publishing Company Collection held by the Oklahoma Historical Society. Items are tiff images.
CODA-Ingest-Batch-Identifier: 50d33e87-7c97-40d4-a5bf-5088f01a0590
Source-Organization: Oklahoma Historical Society

There are 3 premis events associated with ark:/67531/codalbcd:

<table>
<thead>
<tr>
<th>Event ID</th>
<th>Event Date</th>
<th>Event Status</th>
<th>Linked Object(s)</th>
<th>Classified Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>e1af630c71486bf9036f5d7d2c1bb3321</td>
<td>2014-11-20 17:47:05</td>
<td>Success</td>
<td>% ark:/67531/codalbcd</td>
<td><a href="http://purl.org/net/unti/5uluaries/preservationEvents/Ingestion">http://purl.org/net/unti/5uluaries/preservationEvents/Ingestion</a></td>
</tr>
<tr>
<td>75867b03c1f044a901164854a4ed010f10e7</td>
<td>2014-11-22 13:10:40</td>
<td>Success</td>
<td>% ark:/67531/codalbcd</td>
<td><a href="http://purl.org/net/unti/5uluaries/preservationEvents/ffityCheck">http://purl.org/net/unti/5uluaries/preservationEvents/ffityCheck</a></td>
</tr>
<tr>
<td>fed3396625c440d8f5b40530e40dcb1</td>
<td>2014-11-22 13:10:41</td>
<td>Success</td>
<td>% ark:/67531/codalbcd</td>
<td><a href="http://purl.org/net/unti/5uluaries/preservationEvents/replcatiation">http://purl.org/net/unti/5uluaries/preservationEvents/replcatiation</a></td>
</tr>
</tbody>
</table>
URLS to Reconstitute Bag
ATOM Feed for Bag
Storage Node List
## Storage Node Status

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
<th>Capacity</th>
<th>Used</th>
<th>Available</th>
<th>Full</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>code-001</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-001/">http://libdigirch10.library.unt.edu/coda-001/</a></td>
<td>23.0TB</td>
<td>22.8TB</td>
<td>169.0 GB</td>
<td>99.28%</td>
<td>Nov. 1, 2014, 10:06 p.m.</td>
</tr>
<tr>
<td>code-003</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-003/">http://libdigirch10.library.unt.edu/coda-003/</a></td>
<td>23.0TB</td>
<td>22.9TB</td>
<td>134.7 GB</td>
<td>99.43%</td>
<td>Oct. 31, 2014, 10:58 a.m.</td>
</tr>
<tr>
<td>code-004</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-004/">http://libdigirch10.library.unt.edu/coda-004/</a></td>
<td>23.4TB</td>
<td>23.2TB</td>
<td>147.9 GB</td>
<td>99.38%</td>
<td>Nov. 8, 2014, 3:56 p.m.</td>
</tr>
<tr>
<td>code-005</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-005/">http://libdigirch10.library.unt.edu/coda-005/</a></td>
<td>23.4TB</td>
<td>23.3TB</td>
<td>118.3 GB</td>
<td>99.51%</td>
<td>Nov. 6, 2014, 6:41 a.m.</td>
</tr>
<tr>
<td>code-007</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-007/">http://libdigirch10.library.unt.edu/coda-007/</a></td>
<td>22.6TB</td>
<td>22.4TB</td>
<td>146.1 GB</td>
<td>99.37%</td>
<td>June 6, 2014, 3:53 p.m.</td>
</tr>
<tr>
<td>code-008</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-008/">http://libdigirch10.library.unt.edu/coda-008/</a></td>
<td>24.0TB</td>
<td>23.8TB</td>
<td>168.3 GB</td>
<td>99.32%</td>
<td>Nov. 11, 2014, 9:10 a.m.</td>
</tr>
<tr>
<td>code-010</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-010/">http://libdigirch10.library.unt.edu/coda-010/</a></td>
<td>24.0TB</td>
<td>23.9TB</td>
<td>139.1 GB</td>
<td>99.43%</td>
<td>Jan. 31, 2015, 10:10 a.m.</td>
</tr>
<tr>
<td>code-011</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-011/">http://libdigirch10.library.unt.edu/coda-011/</a></td>
<td>24.0TB</td>
<td>23.5TB</td>
<td>535.4 GB</td>
<td>97.82%</td>
<td>June 10, 2015, 6:37 a.m.</td>
</tr>
<tr>
<td>code-012</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-012/">http://libdigirch10.library.unt.edu/coda-012/</a></td>
<td>24.0TB</td>
<td>23.5TB</td>
<td>15.5 TB</td>
<td>35.54%</td>
<td>June 9, 2015, 3:24 p.m.</td>
</tr>
<tr>
<td>code-013</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-013/">http://libdigirch10.library.unt.edu/coda-013/</a></td>
<td>24.0TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>code-014</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-014/">http://libdigirch10.library.unt.edu/coda-014/</a></td>
<td>22.6TB</td>
<td>32.8 MB</td>
<td>22.6 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>code-015</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-015/">http://libdigirch10.library.unt.edu/coda-015/</a></td>
<td>24.0TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:03 p.m.</td>
</tr>
<tr>
<td>code-016</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-016/">http://libdigirch10.library.unt.edu/coda-016/</a></td>
<td>24.0TB</td>
<td>32.8 MB</td>
<td>24.0 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:04 p.m.</td>
</tr>
<tr>
<td>code-017</td>
<td><a href="http://libdigirch10.library.unt.edu/coda-017/">http://libdigirch10.library.unt.edu/coda-017/</a></td>
<td>22.6TB</td>
<td>32.8 MB</td>
<td>22.6 TB</td>
<td>0.0%</td>
<td>Jan. 21, 2015, 3:04 p.m.</td>
</tr>
</tbody>
</table>

**Totals:**
- Capacity: 386.9 TB
- Used: 294.3 TB
- Available: 134.6 TB
- Full: 66.25%
Storage Node Detail
<table>
<thead>
<tr>
<th><strong>Storage Node Status</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name</strong></td>
</tr>
<tr>
<td><strong>URL</strong></td>
</tr>
<tr>
<td><strong>Capacity</strong></td>
</tr>
<tr>
<td><strong>Used</strong></td>
</tr>
<tr>
<td><strong>Available</strong></td>
</tr>
<tr>
<td><strong>Full</strong></td>
</tr>
<tr>
<td><strong>Last Updated</strong></td>
</tr>
</tbody>
</table>
PREMIS Event Listing
There are 4,722,656 total events. Here are the 10 most recent events.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Date</th>
<th>Event Type</th>
<th>Linked Object(s)</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>687891f638414ede91f1c7395142f69e</td>
<td>June 10, 2015, 6:37 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codahb01</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>fb7050668fa296a4236af2354d6fba00e1</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codahb01</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>sfb70e297f3bd674440bd4f473e9b3b</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codahaz</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>75942b29818640ff95e38f56c76885f18</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghay</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>c2f56e12824e4dfc8b3e9f3f001e</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghax</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>e14af46a67b045ff40a01ae844b3729</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghaw</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>c32af9118774f6b6917b860d9864ae9</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghav</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>3ed500a3e454b29545614d4565b180</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghau</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>69f5d7b466cd5f98cb7df46b0929e87</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghst</td>
<td>SUCCESS</td>
</tr>
<tr>
<td>69c76f595664751b97b5f6c3559569b</td>
<td>June 10, 2015, 6:36 a.m.</td>
<td><img src="http://purl.org/net/untl/vocabularies/preservationEvents/ingestion" alt="Image" /></td>
<td>arlc67531/codaghas</td>
<td>SUCCESS</td>
</tr>
</tbody>
</table>
PREMIS Event Detail
<table>
<thead>
<tr>
<th>Event Identifier</th>
<th>48789196384f4e3e9117c7392142f6e9e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier Type</td>
<td><a href="http://purl.org/net/until/vocabularies/identifier-qualifiers/#UUID">http://purl.org/net/until/vocabularies/identifier-qualifiers/#UUID</a></td>
</tr>
<tr>
<td>Event Type</td>
<td><a href="http://purl.org/net/until/vocabularies/preservationEvents/#ingestion">http://purl.org/net/until/vocabularies/preservationEvents/#ingestion</a></td>
</tr>
<tr>
<td>Event Date</td>
<td>June 10, 2015, 6:37 a.m.</td>
</tr>
<tr>
<td>Event Details</td>
<td>Ingest of codagb1 performed by CODAProcess.py, from dropbox at /data10/coda-011_drobox/1.ToCODA</td>
</tr>
<tr>
<td>Date Added</td>
<td>June 10, 2015, 6:37 a.m.</td>
</tr>
<tr>
<td>Linking Agent Identifier Value</td>
<td><a href="http://coda.library.unr.edu/agent/codaingest">http://coda.library.unr.edu/agent/codaingest</a></td>
</tr>
<tr>
<td>Linking Agent Identifier Type</td>
<td><a href="http://purl.org/net/until/vocabularies/identifier-qualifiers/#URL">http://purl.org/net/until/vocabularies/identifier-qualifiers/#URL</a></td>
</tr>
<tr>
<td>Linking Agent Role</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td><a href="http://purl.org/net/until/vocabularies/eventOutcomes/#success">http://purl.org/net/until/vocabularies/eventOutcomes/#success</a></td>
</tr>
<tr>
<td>Outcome Detail</td>
<td></td>
</tr>
<tr>
<td>Linked Objects</td>
<td>arc:/67531/codagb1</td>
</tr>
</tbody>
</table>

ATOMPUB
Validate Dashboard
## Recent Activity

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Added</th>
<th>Last Verified</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>aric/67531/codao2t9</td>
<td>May 15, 2015, 6:45 p.m.</td>
<td>June 10, 2015, 7:08 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao6f</td>
<td>June 5, 2015, 5:02 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoB26k</td>
<td>June 5, 2015, 1:20 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoDr8</td>
<td>June 5, 2015, 4:04 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoH23</td>
<td>June 5, 2015, 8:25 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoCzo</td>
<td>June 5, 2015, 7:21 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao4Zq</td>
<td>June 5, 2015, 6:57 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoHen</td>
<td>June 5, 2015, 4:08 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao8w</td>
<td>June 5, 2015, 7:19 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoH00</td>
<td>June 5, 2015, 5:01 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoHws</td>
<td>June 5, 2015, 5:15 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoPh</td>
<td>June 5, 2015, 5:57 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao82p</td>
<td>June 5, 2015, 7:55 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao5p</td>
<td>June 5, 2015, 1:45 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoJf</td>
<td>June 5, 2015, 5:26 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoD5</td>
<td>June 5, 2015, 1:16 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao8rc</td>
<td>June 5, 2015, 6:43 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao2tw</td>
<td>May 4, 2015, 1:30 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoBfj</td>
<td>June 5, 2015, 1:05 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codooB5t</td>
<td>June 5, 2015, 4:31 p.m.</td>
<td>June 10, 2015, 7:06 a.m.</td>
<td>Passed</td>
</tr>
</tbody>
</table>
Validate Entry
<table>
<thead>
<tr>
<th>Identifier</th>
<th>i ark:67531/codao29k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Passed</td>
</tr>
<tr>
<td>Verified Date</td>
<td>June 10, 2015, 7:10 a.m.</td>
</tr>
<tr>
<td>Added Date</td>
<td>May 4, 2015, 11:36 a.m.</td>
</tr>
<tr>
<td>Priority</td>
<td>0</td>
</tr>
<tr>
<td>Priority Change Date</td>
<td>June 10, 2015, 7:10 a.m.</td>
</tr>
<tr>
<td>Server</td>
<td>llbdlarchive28.library.unt.edu</td>
</tr>
</tbody>
</table>
Validate Entry after Prioritization for On-Demand Validation
<table>
<thead>
<tr>
<th>Identifier</th>
<th>i ark:67521/codao29k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Passed</td>
</tr>
<tr>
<td>Verified Date</td>
<td>June 10, 2015, 7:09 a.m.</td>
</tr>
<tr>
<td>Added Date</td>
<td>May 4, 2015, 11:36 a.m.</td>
</tr>
<tr>
<td>Priority</td>
<td>1</td>
</tr>
<tr>
<td>Priority Change Date</td>
<td>Jan. 1, 2000, midnight</td>
</tr>
<tr>
<td>Server</td>
<td>lbcdigarch29.library.unt.edu</td>
</tr>
</tbody>
</table>
### Prioritized Items

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Added</th>
<th>Last Verified</th>
<th>Date Prioritized</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>aric/67531/codaoc86</td>
<td>June 5, 2015, 7:01 p.m.</td>
<td>June 10, 2015, 7:09 a.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
</tbody>
</table>

### Recent Activity

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Added</th>
<th>Last Verified</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>aric/67531/codao9k9</td>
<td>June 5, 2015, 4:26 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao5zm</td>
<td>June 5, 2015, 8:20 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaokp2</td>
<td>June 5, 2015, 8:17 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao60v</td>
<td>June 5, 2015, 7:31 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao0rj</td>
<td>June 5, 2015, 1:34 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaoe05</td>
<td>June 5, 2015, 11:50 a.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao29k</td>
<td>May 4, 2015, 11:36 a.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao0nf</td>
<td>June 5, 2015, 8:10 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao2yl</td>
<td>May 15, 2015, 6:52 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao0km</td>
<td>June 5, 2015, 11:43 a.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao79y</td>
<td>June 5, 2015, 1:53 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao6by</td>
<td>June 5, 2015, 2:35 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao8ln</td>
<td>June 5, 2015, 5:37 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao6f5</td>
<td>June 5, 2015, 3:17 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codao210</td>
<td>May 15, 2015, 6:45 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaw8r</td>
<td>April 14, 2015, 7:14 p.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
<tr>
<td>aric/67531/codaof61</td>
<td>June 5, 2015, 11:45 a.m.</td>
<td>June 10, 2015, 7:10 a.m.</td>
<td>Passed</td>
</tr>
</tbody>
</table>
Validation Statistics
Coda Search Interface
<table>
<thead>
<tr>
<th>Ark ID</th>
<th>Bagged Date</th>
<th>URLs</th>
<th>ATOM</th>
<th>Size</th>
<th># Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>ark:/57831/coda25f</td>
<td>Oct. 3, 2010</td>
<td>url</td>
<td>ATOM</td>
<td>14.8 MB</td>
<td>4</td>
</tr>
</tbody>
</table>
Coda OAI-PMH Interface
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```xml
http://www.openarchives.org/OAI/2.0/OAIFRME.xsd">  
<responseDate>2015-06-10T13:13:25Z</responseDate>  
<request verb="Identify">http://ods.library.unt.edu/oai</request>  
</OAId>
```

194
There are 4 total agents.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://coda.library.unt.edu/agent/codaIngest">http://coda.library.unt.edu/agent/codaIngest</a></td>
<td>Coda Ingest Agent</td>
<td></td>
</tr>
<tr>
<td><a href="http://coda.library.unt.edu/agent/codaMigration">http://coda.library.unt.edu/agent/codaMigration</a></td>
<td>Coda Migration and Verification</td>
<td></td>
</tr>
<tr>
<td><a href="http://coda.library.unt.edu/agent/codaReplication">http://coda.library.unt.edu/agent/codaReplication</a></td>
<td>Coda Replication and Verification</td>
<td></td>
</tr>
<tr>
<td><a href="http://coda.library.unt.edu/agent/codaValidation">http://coda.library.unt.edu/agent/codaValidation</a></td>
<td>Coda Validation</td>
<td></td>
</tr>
<tr>
<td><strong>Agent ID:</strong></td>
<td><a href="http://coda.library.unt.edu/agent/codal">http://coda.library.unt.edu/agent/codal</a> ingest</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Agent Type:</strong></td>
<td>Software</td>
<td></td>
</tr>
<tr>
<td><strong>Agent Name:</strong></td>
<td>Coda Ingest Agent</td>
<td></td>
</tr>
<tr>
<td><strong>Agent Note:</strong></td>
<td>This agent is responsible for processing archival information packages, checking that they are valid and then ingesting them into Coda.</td>
<td></td>
</tr>
</tbody>
</table>
Coda Stats Pages
Replication Queue
<table>
<thead>
<tr>
<th>Identifier</th>
<th>Status</th>
<th>Harvest Start</th>
<th>Harvest End</th>
<th>Oxum</th>
</tr>
</thead>
<tbody>
<tr>
<td>aric/67531/codapc7u</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>70342620.6</td>
</tr>
<tr>
<td>aric/67531/codapd97</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>70309877.6</td>
</tr>
<tr>
<td>aric/67531/codapd5k</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>70203621.6</td>
</tr>
<tr>
<td>aric/67531/codaphi2</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>70191586.6</td>
</tr>
<tr>
<td>aric/67531/codapf6</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>70157869.6</td>
</tr>
<tr>
<td>aric/67531/codapdc7</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>69972991.6</td>
</tr>
<tr>
<td>aric/67531/codaph6t</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>69772119.6</td>
</tr>
<tr>
<td>aric/67531/codapbos</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>69823234.6</td>
</tr>
<tr>
<td>aric/67531/codapby6</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>69869484.6</td>
</tr>
<tr>
<td>aric/67531/codapgsf</td>
<td>Completed</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td>69822497.6</td>
</tr>
<tr>
<td>Ark</td>
<td>ark:/67531/codapo7u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>Completed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvest Start Date</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvest End Date</td>
<td>June 10, 2015, 6:05 a.m.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxum</td>
<td>70342620.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Queue Position</td>
<td>29445</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
About and API Documentation
Introduction

The UNT Coda App is a tool for validating and retrieving data related to coda and premis events. If you have questions or comments about the UNT Coda App, please contact Joey Liechty at joey.liechty@unt.edu.

Three Sections

Notice there are three sections to The Coda App. The first section is CODA. This section will detail bags information and statistics, storage node status, and also serves up the sitemap xml files. There is also a system in place to validate the contents of BagIT structures. The second section is PREMIS EVENTS. Premis Events serve as a status indicator for different events as they relate to the coda system, such as data migration events. This section gives detailed on the agents, which give information related about the transfer from underlying medium storage to another, a list of the most recent events, and a search utility for the events. The third and final section is QUEUE. The queue shows a listing of to-be-ran and completed process, such as a harvest event. There is also a status page and search function.

Search API

Let's dive right into the following are various search examples using curl:

```bash
# to get only the events which have failed, use the 'outcome' url argument
$ curl -L coda.library.unt.edu/event/search.json?outcome=http://purl.org/net/unt/vocabularies/eventOutcomes#failure
[
  {
    'date': '2013-05-13 14:27:29',
    'event_type': 'http://purl.org/net/unt/vocabularies/preservationEvents#fixityCheck',
    'identifier': 'Z26Bbcs6F0134c3e890d56664c98e852',
    'linked_objects': 'or.k/67531/coda1.295',
    'outcome': 'http://purl.org/net/unt/vocabularies/eventOutcomes#failure'
  },
  {
    'date': '2013-05-13 14:27:30',
    'event_type': 'http://purl.org/net/unt/vocabularies/preservationEvents#Migration',
    'identifier': '41e8bfc6e944b882db1b9d3ebf33f',
    'linked_objects': 'or.k/67531/coda1.295',
    'outcome': 'http://purl.org/net/unt/vocabularies/eventOutcomes#failure'
  }
  ...
]

# to get the events which failed after a certain date, append a 'start_date' argument
$ curl -L coda.library.unt.edu/event/search.json?outcome=http://purl.org/net/unt/vocabularies/eventOutcomes#failure&start_date=05/13/2013
[
  {
    'date': '2013-05-13 14:27:30',
    'event_type': 'http://purl.org/net/unt/vocabularies/preservationEvents#Migration',
    'identifier': '41e8bfc6e944b882db1b9d3ebf33f',
    'linked_objects': 'or.k/67531/coda1.295',
    'outcome': 'http://purl.org/net/unt/vocabularies/eventOutcomes#failure'
  },
  ...
]

# filter events to a particular bag object by giving a 'link_object_id'

208
Atom Feed For Coda
AtomPub Interface
<?xml version="1.0"?>
<service>
  <workspace>
    <title xmlns="http://www.w3.org/2005/Atom">UNT Coda Atom Publishing Protocol (APP) Interface</title>
    <collection href="http://coda.library.unt.edu/APP/node/">
      <title xmlns="http://www.w3.org/2005/Atom">nodes</title>
      <accept>application/atom+xml;type=entry</accept>
    </collection>
    <collection href="http://coda.library.unt.edu/APP/bag/">
      <title xmlns="http://www.w3.org/2005/Atom">bag</title>
      <accept>application/atom+xml;type=entry</accept>
    </collection>
    <collection href="http://coda.library.unt.edu/APP/event/">
      <title xmlns="http://www.w3.org/2005/Atom">event</title>
      <accept>application/atom+xml;type=entry</accept>
    </collection>
    <collection href="http://coda.library.unt.edu/APP/agent/">
      <title xmlns="http://www.w3.org/2005/Atom">agent</title>
      <accept>application/atom+xml;type=entry</accept>
    </collection>
    <collection href="http://coda.library.unt.edu/APP/queue/">
      <title xmlns="http://www.w3.org/2005/Atom">queue</title>
      <accept>application/atom+xml;type=entry</accept>
    </collection>
  </workspace>
</service>
Photographs digitized from the Oklahoma Publishing Company Collection held by the Oklahoma Historical Society. Items are tiff images.
ResourceSync Interface
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```xml
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" xmlns:rs="http://www.openarchives.org/rs/terms/">
  <url>
    <loc>http://coda.library.unt.edu/resourceindex.xml</loc>
    <rs:md capability="resourcemd"/>
  </url>
  <url>
    <loc>http://coda.library.unt.edu/changelist.xml</loc>
    <rs:md capability="changelist"/>
  </url>
</urlset>
```
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```xml
<sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" xmlns:rs="http://www.openarchives.org/rs/terms/">
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-001.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-002.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-003.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-004.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-005.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-006.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-007.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-008.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-009.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-010.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-011.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-012.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-013.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-014.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-015.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-016.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-017.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-018.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-019.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-020.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-021.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-022.xml</loc>
  </sitemap>
  <sitemap>
    <loc>http://coda.library.unt.edu/resourcelist-023.xml</loc>
  </sitemap>
</sitemapindex>
```
UNT Libraries’ Archival Storage Replacement Fund

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
UNT Libraries’ Archival Storage Replacement Fund

Overview

The Archival Storage Replacement Fund (ASRF) was created in 2013 by the University of North Texas (UNT) Libraries as a way of lessening the financial effect of the replacement of archival storage for the UNT Libraries’ Digital Collections.

Background

The UNT Libraries operate a large digital library infrastructure locally referred to as the UNT Libraries’ Digital Collections. This digital library platform stores archival copies of both digitized analog and collected born-digital resources used by users around the world. These collections are defined by the UNT Libraries’ Collection Development Policy for Digital Collections.

Two copies of the archival master content are stored: one copy on the UNT main campus in the Willis Library server room; a second copy is stored at the UNT Discovery Park research campus in the UNT ITSS server room, located 3.5 miles north of the main UNT campus. Each of these copies of data are stored on spinning disk technology and are part of an active replication and monitoring system developed by the UNT Libraries called Coda.

The underlying storage used by the UNT Libraries’ Digital Collections is on a refreshment schedule for every five years. The last storage refresh took place in the summer of 2012 with the next full refresh of the underlying storage technology expected for 2017.

The ASRF was started to help minimize the impact of this planned storage replacement on the UNT Libraries yearly budget by placing a portion of the replacement cost into a local account each year that will be used to replace the existing archival storage infrastructure in the future.

Assumptions

The ASRF was created with the following assumptions in place:

- UNT Libraries is committed to the long term preservation and archiving of all items accessioned into the UNT Libraries’ Digital Collections.
- UNT Libraries will operate two complete copies of the digital libraries’ content locally on spinning disk.
- Storage costs will decline over a five year period.
• Future storage technologies will offer denser and more cost effective storage platforms over time.
• Replacement solutions for archival storage will exist with similar or better performance and cost metrics than the previous generation of storage infrastructure used.
• The UNT Libraries will deposit the calculated allotment into the storage fund at the beginning of each fiscal year.
• While the ASRF is expected to cover the replacement costs of storage used by existing digital objects in the UNT Libraries’ Digital Collections, additional storage will be purchased at each refresh cycle that may add costs to the overall storage purchase and exceed the funds in the ASRF.

Calculating yearly allotments
In order to maintain adequate funding for future storage replacement costs in the ASRF, it is necessary for the UNT Libraries to contribute the calculated allotment for the given year.

To calculate the yearly allotment, the total number of terabytes (TB) stored in the UNT Libraries’ Digital Collections is used as the base storage amount with a storage Cost-Per-Year Multiplier multiplied to the base storage amount.

Base Storage Amount * Cost-Per-Year Multiplier = Yearly Allotment

Calculating the Cost-Per-Year Multiplier
With each equipment refresh, the multiplier for the following five years is set based on twice the cost per formatted and usable TB of storage purchased, once the new storage is in place.

This cost per formatted and usable TB of storage is calculated on one instance of the storage infrastructure. Costs to include in the calculations include all compute, storage, warranty and directly associated rack-networking needs for the purchase. Below is an example of how these costs are determined.

168 TB of formatted and usable TB of storage purchased for $42,000
42,000 / 168 TB= $250 per TB
$250 per TB / 5 year lifespan = $50 per TB-Cost-Per-Year
50 per TB-Cost-Per-Year * 2 copies = $100 Cost-Per-Year Multiplier

Adding to the fund.
At the beginning of each fiscal year, the Assistant Dean for Digital Libraries will alert the Administrative Office and Technology and Computer Operations (TACO) as to the current
amount of content being consumed by the UNT Libraries' Digital Collections. This number is calculated as the amount of data registered within the Coda archival system at http://coda.library.unt.edu/. This amount is used for the Base Storage Amount used in the funding calculation described above. The amount of funding needed to cover this amount of storage is then added to the ASRF for the next fiscal year.

Example Archival Storage Replacement Fund 2013-2017:

September 1, 2013 = 120 TB = $12,000 placed in Archival Storage Replacement Fund
September 1, 2014 = 160 TB = $16,000 placed in Archival Storage Replacement Fund
September 1, 2015 = 192 TB = $19,200 placed in Archival Storage Replacement Fund
September 1, 2016 = 240 TB = $24,000 placed in Archival Storage Replacement Fund
September 1, 2017 = 300 TB = $30,000 placed in Archival Storage Replacement Fund

Total Funds Available in 2017 to supplement storage replacement: $101,200

Expendung funds

It is expected that the ASRF will be used once every five years for the purchase of replacement hardware for the UNT Libraries' Digital Collections. It should be used in its entirety to cover the costs of storage and, if available, supplement the cost of new storage also purchased at that time. It is expected that this funding will be used in conjunction with funds from other funding sources such as grants, local accounts, or other funding pools as appropriate. It is expected that in order to receive the most aggressive storage costs, this funding may be pooled with other infrastructure expenditures that the UNT Libraries incurs at the same time to make the best deal for the library.

Calculated Storage Multiplier

For the period of 2012-2017, the storage multiplier is $100 per TB per year. This will be recalculated after the next storage refresh.
UNT Libraries’ Policy on Creation and Licensing of Metadata

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Sian Brannon Assistant Dean for Collection Management
Jesse Hamner Director of Research & Assessment

This work is licensed under a Creative Commons Attribution 4.0 International License.
Locally Created Cataloging and Metadata Records
Rights Policy

APPROVED: April 2015

As a way of supporting Open Access (http://openaccess.unt.edu) and the free dissemination and reuse of knowledge, the UNT Libraries makes its contribution of cataloging and metadata records available as Public Domain resources.

All locally created cataloging and metadata records by the UNT Libraries are Public Domain and are available under the Creative Commons Public Domain Dedication (CC0) (http://creativecommons.org/publicdomain/zero/1.0/).

The University of North Texas Libraries, as creator of bibliographic and descriptive metadata records, has waived all rights to them worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

A bibliographic or descriptive metadata record for some resources, such as theses and dissertations, may contain an abstract created by the author of the work. Copyright law may still apply to the abstract in the bibliographic or descriptive metadata record.

This work is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).
UNT Libraries’ Open Source Software Policy

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Will Hicks Department Head, User Interfaces
Jason Thomale Resource Discovery Systems Librarian

This work is licensed under a Creative Commons Attribution 4.0 International License.
Open Source Software Policy

The UNT Libraries issues the following policy related to the use of, contribution to, and creation and release of Open Source Software in order to support open scholarship at the University of North Texas by building upon the efforts of the UNT Open Access Policy (http://policy.unt.edu/sites/default/files/untpolicy/17.5_Open%20Access_Self-Archiving_and%20Long-Term%20Digital%20Stewardship%20for%20UNT%20Scholarly%20Works.pdf) for scholarly publishing and the Denton Declaration (http://openaccess.unt.edu/denton-declaration) for releasing and managing data.

Definitions

**Contribution** is a tangible product created by individuals that help advance an Open Source Software project. Examples of contributions include: programming, bug fixes, issue reports, feature suggestions, and documentation.

**Open Source Software** is “software that can be freely used, changed, and shared (in modified or unmodified form) by anyone” (http://opensource.org/) and is distributed under a license approved by the Open Source Initiative, e.g. one of the licenses found here (http://opensource.org/licenses).

A **Project**—specifically, an Open Source Software project—includes the entirety of collaborative activity, such as planning, discussion, design, documentation, and programming, that furthers the development of a piece of Open Source Software.

**Software** refers to computer programs, or any set of instructions intended to be executed by a machine, including (but not limited to) executable files, libraries, and scripts.

Use Open Source Software when applicable.

- The UNT Libraries supports using Free and Open Source Software for its business operations whenever open source solutions meet business needs.
Contribute to existing Open Source Software development projects.

- The UNT Libraries encourages librarians and staff members to contribute to projects for the Open Source Software that they use. Suitable projects should have a clearly documented license approved by the Open Source Initiative (http://opensource.org/licenses).

- UNT Libraries employees should ensure that the level of contribution made to an external project is adequately credited within the project’s documentation.

- If a project requires it, copyright for contributed code may be transferred to another party, but UNT Libraries units and departments are responsible for evaluating such projects on a case-by-case basis to ensure they align with relevant UNT policies (such as the Copyright Compliance Policy, No. 16.13.3; Contracts and Agreements, No. 10.4; and the Intellectual Property Policy, No. 16.13.1) before their employees participate.

Create and sustain Open Source Software projects.

- Software created in-house may be released under an Open Source License approved by the Open Source Initiative when possible.

- All software released by the UNT Libraries will include current contact information for the maintaining person, unit, or other entity, and a license statement. The BSD Three Clause License (see below) is recommended.

- Open Source Software projects should use a public version control repository to manage the software when possible. A hosting service such as GitHub (https://github.com) or BitBucket (https://bitbucket.org) is recommended.

Abide by best practices for Open Source Software development.

- The author of, or contributor to, an Open Source Software project will adhere to common codes of conduct within shared version control repositories and will address issues and other communications with the public in an timely, professional manner.
If there are situations that are not covered by this policy or if there are questions about the policy in general, please contact the Assistant Dean for Digital Libraries.

Suggested BSD 3 Clause License Template

Copyright © 2014, Regents of the University of North Texas
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

Neither the name of the University of North Texas Libraries nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

This work is licensed under a Creative Commons Attribution 4.0 International License.
Digital Libraries Division Data Loss Escalation Procedures

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Digital Libraries Division Data Loss Escalation Procedures

Introduction

While data loss is a difficult subject in any repository environment, the UNT Libraries strives to make transparent information about data loss in either the Coda repository or the Aubrey access systems that act as underlying infrastructure for the UNT Libraries’ Digital Collections.

This document is being created as a way of notating the procedures related to the escalation of events for a significant data loss event that adversely affects the UNT Libraries’ Digital Collections.

For the purposes of this document we define data loss event as any occurrence that results in the loss of preservation or access digital objects, or the metadata related to these items.

Definitions

We also define data loss into two categories, temporary data loss and permanent data loss. Data loss can occur in the format of either loss of data or a corruption of data that renders a digital object or set of digital objects unusable.

Temporary Data Loss: Data loss that can be recovered either from a replica copy of the data, or from a traditional tape backup. Often times temporary data loss may not have adverse affects on the repository’s designated community of users because of various redundancies in place.

Permanent Data Loss: Data loss that can not be recovered from in any traditional way. In this instance data is lost forever and will require regeneration. Usually this type of data loss will adversely affect the repository’s designated community of users.

Procedures

When a data loss of any kind has been identified the first step is to verify that the data loss is in fact a loss. This includes the verification from another device, and in most cases having another person verify that the data loss has occurred.
Once a data loss has been positively identified, the Assistant Dean for Digital Libraries is notified if there has been any issue related data corruption or loss in the system.

Options for how to handle the data loss is discussed with members of the Digital Libraries Division and a recovery plan is generated to attempt to correct the issue.

**Recovery Plan**

A data recovery plan describes the data loss that occurred, when and how it was identified, how it was corroborated by a second person, and information about the plan to attempt to recover the data that was identified as missing or corrupt. This recovery plan includes the individuals who will work on the process, how we will proceed with the process of recovery. This recovery plan is distributed via email and includes as recipients the Assistant Dean for Digital Libraries, Head of the Digital Projects Unit, and the members of the DPU Software Development Team. It may be necessary to include the Libraries Network Manager or Director of Facilities and Systems Division in the email to keep them abreast of the situation.

**Coda**

If data loss occurs in one of the Coda repository instances, the digital objects that are affected are identified and restored from the other Coda repository instance. The steps for recovering these items include the following:

- Validate digital object on Coda repository that is unaffected.
- Replicate digital object to Coda repository instance where it is missing.
- Perform a full validation of the item once replicated.
- Create a PREMIS Event for the validation and the successful replication in the PREMIS Event Service.
- Confirm that affected digital objects are complete and present.
- Alert those identified in the recovery plan of its completion.

**Aubrey**

If the data loss occurs in the Aubrey access system then an inventory of the affected items is conducted and a process to re-populate the affected digital objects from the Archival Information Packages (AIPs) in the Coda repository begun. This workflow includes:

- Create a list of affected items in Aubrey
- Generate list of Coda identifiers that match the public access ARKs
- Retrieve digital objects from Coda
- Regenerate Dissemination Information Packages (DIPs) to be reloaded into the Aubrey access system.
- Verify that the items affected have been fully regenerated via the Aubrey user interface.
Alert those identified in the recovery plan of its completion.

**Metadata**

If the data loss occurs in the Metadata Storage Server, then the affected files are recovered either from the nightly rsync data or if that fails from the nightly tape backups made for the system. In this event it is likely that metadata edits that were made within the previous 24 hours would not be present. A list of the records that were edited during this period is available from the Metadata Edit Event Service. Once this list is generated, persons who edited metadata records can be notified as to which records will need to be re-edited.

**Search Index**

If the data loss occurs in the Solr search index system that is not recoverable with the built-in replication and recovery procedures of Solr, the index will need to be rebuilt. This rebuilding will generally require a long indexing time so a notice to the end users of the Aubrey access systems will be put in place. During the index period it is possible to edit metadata records with the Edit system but all records may not show up as available even if they are. Once a full index has been completed and confirmed the notice to end users can be removed.

**Extended Recovery Issues**

If the recovery plan is the recovery time will significantly impact services for more than 24 hours offered by the UNT Libraries, a message is added to the affected systems that communicates the service unavailability or degraded performance until the recovery can be completed.

If the recovery plan is either unsuccessful or if the recovery time period will significantly impact services for more than 72 hours, or if the data recovery is unsuccessful or considered a *permanent data loss* the following individuals are alerted by the Assistant Dean for Digital Libraries.

- Dean of Libraries
- Associate Dean of Libraries
- Director of Facilities and Systems
- Network Manager

If the data loss will adversely affect the public access to resources in the UNT Libraries Digital Collections a notice will be applied to appropriate systems within an expected period of time to remedy the problem and an email message will be sent to appropriate collection owners and maintainers in the UNT Libraries.
If the issue cannot be resolved within a seven day period a public message will be sent to External Partners with content managed by the UNT Libraries about the situation, steps being taken to remedy the problem, and expected timelines for issue to be resolved.

**Notice of Recovery Plan Completion**

Once the data loss recovery plan has been fully executed and results in either a successful restoring of digital objects and services, or in an outcome of a permanent data loss, the Assistant Dean for Digital Libraries will communicate to those parties notified in previous steps an overview of the issue and steps that will be taken so that the same situation does not occur again.

This information will be documented on the Digital Projects Shared Storage Drive in a folder called “System Issues”.

If the data loss occurred with the Coda repository system, a Major Event is logged with the Digital Libraries Division's Major Event Log application so that a record of the event is saved and associated with affected items in the Coda repository.
UNT Digital Libraries: Risk Analysis and Management Strategy Plan

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
Risk Analysis and Management Strategies: UNT Libraries’ Digital Collections

Introduction

This document describes threats that could disrupt normal operations of the UNT Libraries’ Digital Collections and prevent the repository from fulfilling its mission to support “...the long-term collection, production, maintenance, delivery, and preservation of a wide range of high-quality digital resources and services for the UNT Community and users throughout the world.” for short or long periods of time. This document also details risk-minimization strategies, system infrastructures, operating practices, and business relationships that the UNT Libraries employ to reduce the potential impact of these threats.

This document was prepared by the UNT Libraries’ Digital Libraries Division in order to identify risks and document the management strategies of those risks to the UNT Libraries’ Digital Collections. While the Digital Libraries Division is responsible for the technology, workflows, and infrastructure of the Digital Collections, success of this work relies on close coordination between curators and stakeholders from other departments and divisions throughout the UNT Libraries. The strategy discussed in this document heavily references and relies upon services provided and managed by the UNT Libraries’ Facilities and Systems Division and its two units, Technology and Computer Operations (Lib-TACO) and Facilities. This document will refer to the UNT Libraries’ Digital Collections or “repository,” and the Digital Libraries Division as the primary stakeholders or “staff.” This document also refers to the designated community of the Digital Collections comprising a broad audience:

- The UNT Extended Community (UNT students, faculty, staff, alumni, and administrators)
- Contributing partners and their constituents
- The larger academic community and researchers
- The general public throughout the world

The UNT Libraries have adapted a typology of threats developed by the Scholars Portal as part of their TRAC Audit Documentation that focuses on probable and manageable events in relation to the repository’s physical environment, technology infrastructure, repository content, personnel requirements, business relationships, and legal obligations. In the original development of this typology, Scholars Portal examined risk management strategies practiced by a number of authorities in the digital curation community, exemplars include CCSDS typology, PLATTER typology, HathiTrust disaster planning typology. These typologies are included in the document below for reference.
In order to identify and document possible threats to the Digital Collections, the UNT Libraries team borrowed heavily (with permission) from the original document created by the Scholars’ Portal.

## Typologies of threats

This section introduces a number of risk analysis frameworks used by the digital curation community. The structure employed by Scholars Portal for their own Risk Analysis and Management Strategies documentation was reviewed in creating the current documentation. The UNT Libraries, having adopted the Scholars Portal framework, is including portions of that original analysis as a point of reference to prior work conducted in this area. This section outlines some typologies of threats found in the Consultative Committee on Space Data Systems’ (CCSDS) Audit and Certification of Trustworthy Digital Repositories criteria, the PLATTER framework developed by DigitalPreservationEurope (DPE), and the disaster recovery planning process used by HathiTrust.

### CCSDS typology

The *Audit and Certification of Trustworthy Digital Repositories* criteria expects that the repository “maintain[s] a systematic analysis of security risk factors associated with data, systems, personnel, and physical plant” (sec. 5.2.1). In their detailed explanation, the checklist recommends that the repository identify and manage risks related to:

- Hardware, software, communications equipment, facilities, and firewalls
- Physical environment
- Personnel, management, and administration
- Operations and service delivery
- Income, budget, reputation, and mandate
- Contractual and regulatory compliance
- Personnel knowledge and skills
- External threats and denial of service attacks
- Relationships with third parties

### PLATTER typology

The PLATTER (Planning Tool for Trusted Electronic Repositories) checklist developed by DigitalPreservationEurope (DPE) encourages repositories to focus on “foreseeable disasters” in eight general categories (p. 37-39):

- Economic upheaval
- Political upheaval
- Loss of purpose/mandate
- Technological upheaval
- Environmental upheaval
- Loss of users and/or the arrival of competition
- Loss of educated key staff
- Breach of security

**HathiTrust disaster planning typology**

As a part of their disaster recovery planning process, HathiTrust identified ten types of threats:

- Hardware failure or obsolescence and data loss
- Network configuration errors
- Network security and external attacks
- Format obsolescence
- Core utility and/or building failure
- Software failure or obsolescence
- Operator error
- Physical security breach
- Natural or manmade disaster
- Media failure or obsolescence

**Scholars Portal typology**

The typologies listed in this section represent different perspectives on disaster, ranging from abstract, unpredictable threats (“political upheaval”) to tangible, expected threats (“software obsolescence”). From these, several common themes emerge, and threats are grouped into five general categories.

**Economic, political, social, or legal threats.** Includes:

- Loss of funding or institutional host
- Loss of staff
- Contractual liability
- Regulatory liability

**Technology-related failures.** Includes:

- Hardware failure and obsolescence
- Storage media failure (includes bit rot)
- Software failure and obsolescence
- File format-related obsolescence
• Loss of critical hardware or software support

**Manmade threats.** Includes:
• Operator error
• Sabotage by insider
• Cyber attack
• Physical security incident

**Natural threats.** Includes:
• Fire
• Flood
• Severe weather (blizzard, tornado, thunderstorm)
• Earthquake

**Utility or environmental/building systems failure.** Includes:
• Power failure
• Plumbing failure
• Server room cooling failure
• Heating, air conditioning, or air quality failure

**Threat Documentation Model**

The threats presented in this document have the following categories of information present for each documented threat.

**Dependencies**

Why this threat is important to the operation of the UNT Libraries’ Digital Collections and associated services.

**Likelihood and impact assessment**

To assess the risks associated with individual threats and prioritize management strategies, in their original model, Scholars Portal estimated the likelihood and impact of each threat and plotted the combined result on a graph. The UNT Libraries have adopted this model to assess risks. Assessing threats in terms of likelihood and impact is a common strategy in disaster planning and is recommended by the Federal Emergency Management Agency (FEMA) and the National Institute of Standards and Technology (NIST). Here is an example of a graph:
Likelihood and impact are represented on the graph as a green marker. The position of each dot is speculative and serves as a guide for clarifying risks and planning management strategies.

**Likelihood** is an estimate of the frequency of a type of event, divided into three periods for the purposes of this document:

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>occurs every 0-2 years</td>
</tr>
<tr>
<td>Medium</td>
<td>occurs every 3-7 years</td>
</tr>
<tr>
<td>Low</td>
<td>occurs every 8+ years</td>
</tr>
</tbody>
</table>

These are rough estimates, and exceptional events are always possible.

**Impact** is an estimate of the effect that a threat may have on the repository’s content, services, and/or administration. The Digital Libraries Division divides impact into three categories of severity:

<table>
<thead>
<tr>
<th>Impact</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Large-scale data loss in the storage system and/or Prolonged service outage</td>
</tr>
<tr>
<td>Medium</td>
<td>Small-scale, isolated data loss in the storage system and/or</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Brief service outage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
</tr>
<tr>
<td>No data loss in the storage system and Degraded system performance and/or Disruption of repository administration, maintenance, or workflow</td>
</tr>
</tbody>
</table>

In general, high-impact events may lead to widespread data loss and/or prevent the designated community from successfully accessing information or using one or more of the repository’s services for a prolonged duration. Medium-impact events may lead to some data loss and/or brief service outages. Some medium-impact events, whether small-scale data loss or brief service outages (or both), may be corrected before affecting the designated community. Low-impact events do not lead to data loss, but may degrade system performance or disrupt administration of the repository in one way or another. For example, loss of staff does not lead to data loss, but can easily disrupt ongoing maintenance and/or interrupt projects. Please note that data loss and service outages have many causes, and this document cannot describe all of them.

**Management strategy**

Strategies that the Digital Libraries Division employ to mitigate the impact(s) of this risk.

**Identified Threats**

**Economic, political, social, or legal threats**

**Loss of funding**

**Dependencies**

The UNT Libraries’ Digital Collections operate using funding designated by the UNT Libraries in the form of wages for librarians, staff, and student assistants; technology infrastructure support; funding for digitization; and travel and training. The UNT Libraries are an academic unit within the University of North Texas, the flagship campus in the UNT System, which is a public, state university. In addition to annual funding from the UNT Libraries, the Digital Libraries Division
receives external funding in the form of digitization fees from partner institutions who contribute content to The Portal to Texas History. For that reason, the UNT Libraries’ Digital Collections are vulnerable to reductions in funding for higher education by the university, state, and federal governments. UNT Libraries’ Digital Collections are also vulnerable to governmental and university funding freezes. Cost increases that are not matched by funding increases are effectively reductions in funding.

**Likelihood and impact assessment**

Loss of funding could span a range of probabilities and impacts. Full loss of funding (illustrated) is a low-likelihood, low- to high-impact event. Full loss of funding could force the UNT Libraries’ Digital Collections into one of two scenarios: either to suspend operations until the repository can establish a new source of funding, or to force it to engage in a succession plan for the repository. A swift transition to a new source of funding minimizes the chance of a service outage. Partial loss of funding or freezes (not illustrated) are moderate-likelihood, low-impact events. Partial loss could slow or halt efforts to maintain existing collections, infrastructure, and services and could delay the development of new collections and new services. The Digital Collections expect that funding will fluctuate in relation to the financial health of the State of Texas and the university. In addition, funding could shrink if prospective partner libraries choose to work with a competitor rather than with the UNT Libraries for content services.

**Management strategy**

The Digital Libraries Division strives to provide a cost-effective model for the digitization, archiving, and access to resources held in the UNT Libraries’ Digital Collections.

The UNT Libraries have a commitment of funding for regular hardware replacement and/or upgrade. The UNT Libraries have established an Archival Storage Replacement Fund to help
reduce the impact of replacing server and storage hardware for the UNT Libraries' Digital Collections infrastructure during planned for replacement cycles.

The UNT Libraries replace hardware and storage on a five-year cycle (i.e., every five years, or less); however, media exhibiting performance problems or on a more frequent schedule may be replaced before the end of five years. Whenever possible, the UNT Libraries also purchase five-year warranties for hardware components.

In fall 2012, the UNT Libraries established The Portal to Texas History Endowment. The endowment will enable UNT to extend the impact of the Portal by creating a permanent, sustainable source of income. Funds derived from this endowment will serve as a catalyst for enhancing future technology development, acquisition, and support; for adding collections and content to the Portal; and for creating lesson plans and supporting new educational initiatives.

The UNT Libraries work with the Office of the Provost and Vice President for Academic Affairs to establish and secure funding required to support the operations of the UNT Libraries and subsequently the UNT Libraries’ Digital Collections and associated infrastructure and staffing.

**Loss of staff**

**Dependencies**

The UNT Libraries’ Digital Collections rely on knowledgeable and skilled personnel to maintain the repository and satisfy the evolving needs and expectations of the designated community. The Digital Collections also depend on experienced leaders who can manage projects, mentor personnel, and advocate on behalf of the repository. The Digital Libraries Division and the Digital Collections require stable, typical, and ideally low rates of staff turnover in order to manage projects effectively and efficiently.

**Likelihood and impact assessment**
Loss of staff is a high-likelihood, low-impact event. Loss of knowledgeable and skilled personnel is inevitable in the life of an organization, and therefore it is a high-likelihood event, but losses do not typically lead to service outages or data loss. The severity of the impact will depend on the individuals lost. Loss of staff at any level could affect the ability of the Digital Libraries Division to operate in an efficient and innovative manner. Loss of leadership could disrupt business relationships, disrupt administrative activities, and affect the reputation of the repository. The loss of a large number of staff could slow or stop efforts to maintain existing content, infrastructure, and services and delay the development of new content and new services.

Temporary loss of staff due to widespread illness (not illustrated) is a low-likelihood, low-impact event. Temporary loss of staff could slow or stop administrative activities, project development, and ongoing maintenance.

**Management strategy**

The Digital Libraries Division delegates responsibility for projects and repository administration to several people in order to reduce single points of failure.

The Digital Libraries Division uses a number of technologies including shared network drives for documentation, and an internal Wiki as a collaborative documentation portal, for knowledge management, project planning, and document sharing.

The Digital Libraries Division includes documentation of major changes to its organizational and technical infrastructure in its annual Operational Plans and other reports that are submitted to the UNT Libraries’ Administrative Offices.

The Digital Libraries Division maintains an organizational chart that identifies appropriate staff who can explain repository practices and workflows.
The Digital Libraries Division reduces the rate of staff turnover by offering its employees competitive compensation and benefits, clear and reasonable performance expectations, constructive and practical feedback, and opportunities for advancement and professional development.

The Digital Libraries Division shares copies of its yearly Operational Plan via the UNT Libraries’ UNTranet portal.

Contractual liability

Dependencies

Contractual agreements between parties involved with the UNT Libraries' Digital Collection exist between the University of North Texas, content partners, and in some situations external vendor services such as software companies and service providers. It is critical that the Digital Collections (and the greater University of North Texas) fulfill these contracted obligations.

Likelihood and impact assessment

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Contractual liability is a low-likelihood event. It is difficult to estimate the impact of contractual liability because the context and circumstances of any problem may vary widely. In some cases, contractual liability may lead to a temporary service outage while staff from the Digital Libraries Division work to resolve the problem. In other cases, the impact may be strictly administrative or financial (not illustrated). Removal of content by a Provider would be a high-impact event (not illustrated).

Management strategy

All legal documents, contracts, and memoranda of understanding are vetted and signed by the UNT Office of General Counsel (OGC) or body designated with signing authority by the OGC.
The Digital Libraries Division staff work closely with the rest of the UNT Libraries to ensure that the operations are consistent with relevant state and federal laws.

When necessary, the Digital Libraries Division seeks contextual advice from the UNT Libraries’ Director of Copyright and legal advice from the UNT Office of General Council.

Licensing agreements between UNT and content partners are governed by and construed in accordance with the laws of the State of Texas.

In all matters of ingest, data management, archival storage, dissemination, and repository administration, the Digital Libraries Division operates according to the terms of license agreements between UNT and content partners. These agreements describe usage rights in detail and provide UNT with clear directions for managing ingest processes, recording administrative metadata, and implementing security and access controls.

**Regulatory liability**

**Dependencies**

The Digital Libraries Division must design, implement, and administer the UNT Libraries’ Digital Collections in a manner that is consistent with relevant state and federal laws.

**Likelihood and impact assessment**

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regulatory liability is a low-likelihood event. It is difficult to estimate the impact of regulatory liability because the context and circumstances of any problem may vary widely. In some cases, regulatory liability may lead to a temporary service outage while staff from the Digital Libraries Division work to resolve the problem. In other cases, the impact may be strictly administrative or financial (not illustrated).
Management strategy

The Digital Libraries Division staff work closely with the rest of the UNT Libraries to ensure that the operations are consistent with relevant state and federal laws.

When necessary, the Digital Libraries Division seeks contextual advice from the UNT Libraries’ Director of Copyright and legal advice from the UNT Office of General Council.

Licensing agreements between UNT and content partners are governed by and construed in accordance with the laws of the State of Texas.

Technology-related threats

Hardware failure

Dependencies

The mission and related activities of the UNT Libraries’ Digital Collections rely on the continuous, error-free operation of servers, a storage array, and workstations. The UNT Libraries uses Dell servers and Dell RAID storage array housed redundantly in the UNT Libraries and in the UNT ITSS secure computing facility at the Discovery Park research campus. Hardware may fail due to spontaneous malfunctions, manufacturing defects, or improper operation. Points of failure include power supplies, fans, connectors, and (highly unlikely) chips and motherboard components.

Storage media failure due to mechanical failure, physical degradation, or magnetic failure is analyzed as a stand-alone topic in the section titled “Storage media failure that leads to data loss.”

likelihood and impact assessment
Hardware failure is a low-likelihood event. However, hardware failure can span a range of impacts depending on the components involved and the nature of the failure. The number of components and/or sites involved may affect the duration of any outage.

There is a risk that large-scale hardware migrations could cause temporary service outages and lead to data loss.

Management strategy

Digital Libraries Division staff work closely with personnel from the UNT Libraries’ Technology and Computer Operations (Lib-TACO) group to monitor hardware health, perform routine maintenance, and install replacements or upgrades. Lib-TACO also has a stock of surplus component for some hardware.

The Digital Libraries Division uses a variety of widely-accepted, industry-standard techniques and tools to monitor the repository's hardware platform. Systems administrators in the Digital Libraries Division and Lib-TACO receive information about system behavior and usage from a number of custom-built scripts, a Nagios monitoring program, and monitoring functionality built into the hardware. These tools warn administrators about abnormal activity such as excessive processor loads and slow response times. In addition, staff monitor critical processes, such as ingest and data management, for malfunctions and suboptimal performance.

Feedback from the designated community is an important source of information about system behavior and hardware performance. Feedback from the designated community provides valuable information about response times, page loading, and overall system performance.

The UNT Libraries have a commitment of funding for regular hardware replacement and/or upgrade. The UNT Libraries have established an Archival Storage Replacement Fund to help reduce the impact of replacing server and storage hardware for the UNT Libraries Digital Collections infrastructure during planned for replacement cycles.

The UNT Libraries replaces hardware on a five-year cycle (i.e., at least every five years) even if the hardware is functioning normally.

The UNT Libraries purchases a five-year warranty for servers and hardware components when possible.

When server or hardware changes are indicated, Digital Libraries Division staff collaborate with personnel from Lib-TACO to evaluate hardware alternatives and the timing of technology changes on a cost-benefit basis. Evaluations vary in formality according to the circumstances. The most extensive assessments take place whenever large components, such as the repository’s servers or storage array network, require replacement. When necessary, systems
administrators consult vendors for additional information and advice, taking into account the cost of hardware and future maintenance.

Large-scale hardware migration is carried out by systems administrators and programmers from the Digital Libraries Division and Lib-TACO. Staff test and evaluate new hardware in isolation before moving the repository to the new component(s). In addition to tests and checks carried out by automated monitoring programs, staff manually assess changes by examining samples of relevant content. Staff evaluate changes for their effect on the integrity and understandability of information, the speed and interoperability of the system, and the accessibility and usability of disseminated content. Whenever hardware changes involve migrating data, the repository performs checksum and file size tests to validate the integrity of information.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to local standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Storage media failure that leads to data loss

Dependencies

Storage media failure refers to mechanical failure, physical degradation, or magnetic failure in disk drives that may lead to data corruption or loss. The failure of multiple drives in the RAID array may lead to data loss and cause a service outage until Lib-TACO and the Digital Libraries Division can restore the affected content from backup or other replicated copy of the data. Media failure can be considered a subset of hardware failure, but it is analyzed here as a stand-alone topic because of the critical impact of media refreshment on the long-term preservation of digital information.

Likelihood and impact assessment
Storage media failure that leads to data loss is a low-likelihood, medium- to high-impact. The impact of storage media failure varies according to the extent of the damage, the data affected, and the duration of any service outage.

Storage media may fail in a RAID array without causing data loss. For example, the failure of a single drive in the RAID array should not lead to data loss. However, a drive failure will compromise the array’s overall redundancy.

There is a risk that disk replacement and data copying operations could cause temporary service outages.

Management strategy

The repository uses replicated Dell MD3260i and MD3060e storage arrays for redundancy of digital master files.

The repository uses replicated Dell MD3200i and MD1200 storage arrays for redundancy of access files.

Both sets of Dell storage arrays employ RAID technology for purposes of data redundancy.

The Dell arrays have health-monitoring, diagnostic, and error-correction tools. The storage controllers will automatically report errors to Lib-TACO staff.

The Digital Libraries Division uses a number of tools and procedures to detect bit corruption or loss. The repository uses widely-accepted hashing techniques to generate digest values for new content and carries out regular, automated fixity checks on archived content. For each digital object, the repository generates and records MD5 values for each file associated with the object. Digest values are stored in the preservation metadata, which is separate from the objects files.
The UNT Libraries replaces storage media according to a predetermined schedule or whenever media exhibit performance problems. As a rule, the UNT Libraries refreshes its storage array media within a 5-year period (i.e. every 5 years or less) even if the drives are functioning normally and appear healthy.

The UNT Libraries purchases a 5-year warranty for hardware components when possible.

The UNT Libraries have a commitment of funding for regular hardware replacement and/or upgrade. The UNT Libraries have put in place an Archival Storage Replacement Fund to help reduce the impact of replacing server and storage hardware for the UNT Libraries Digital Collections infrastructure during planned for replacement cycles.

The procedures and tests involved in storage media change are very similar to those involved in hardware change. Please see the section title “Hardware failure,” above, for details.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Hardware obsolescence

Dependencies

The UNT Libraries’ Digital Collections rely on hardware components that work well with the repository’s applications, databases, and data and that also interoperate well with associated components inside and outside the system.

Likelihood and impact assessment
It is very difficult to assess the likelihood and impact of hardware obsolescence because it depends on the repository growth rate as well as developments in hardware, software, and user behavior. It may affect diverse users in different ways. Degraded performance is the first sign of hardware obsolescence, but service outages should be expected. The impact could vary from low to medium (and possibly high, not illustrated) in relation to the components involved, the services affected, and the duration of any service outage. Hardware obsolescence is medium-likelihood in the context of repository growth and wider technological change.

There is a risk that large-scale hardware migrations could cause temporary service outages and lead to data corruption or loss.

Management strategy

The UNT Libraries strategy for minimizing the risk of hardware obsolescence is largely the same as its strategy for hardware failure. Please see above, for details.

Software failure

Dependencies

The UNT Libraries’ Digital Collections and the Digital Libraries Division depend on a variety of software components to carry out operations and fulfill their mission. The repository uses a combination of open-source and custom-built software to ingest and transform content, manage backup, disseminate content, and deliver various services.

It is necessary for the repository to install new software or modify existing software as a part of normal operations. Software failure is typically confined to bugs or incompatibilities that appear during or after software changes. The CCSDS’ Audit and Certification of Trustworthy Digital Repositories checklist singles out security patches and firmware updates for special concern and states that they “are frequently responsible for upsetting alternative aspects of system functionality or performance” (sec. 5.1.1.4). Software failure could force the Digital Libraries
Division to suspend operations of the Digital Collections until the affected systems can be thoroughly analyzed, repaired, and tested.

**Likelihood and impact assessment**

Software failure is a high-likelihood, low- to high-impact event. It is a high-likelihood event because bugs, glitches, and incompatibilities are difficult to predict and eliminate. The Digital Libraries Division uses widely-accepted, industry-standard procedures for testing and evaluating software changes, but small errors or conflicts sometimes escape testing. The impact may vary from low to high in relation to the systems affected, the duration of any service outage, and the extent of any data loss.

There is a risk that large-scale software migrations could cause temporary service outages and lead to data corruption or loss.

**Management strategy**

Software development, testing, and improvement is an ongoing process of the Digital Libraries Division. In general, there is no time when software is not subject to monitoring and evaluation.

The Digital Libraries Division uses a variety of current, widely-accepted, industry-standard techniques and tools to monitor the repository's applications. Systems administrators and programmers receive information about system behavior and usage from a number of custom-built scripts.

Feedback from the designated community provides valuable information about accessibility, usability, understandability, and holdings. The Digital Libraries Division receives ongoing feedback about application behavior and interface design from its designated community.
Lib-TACO staff evaluate all mandatory and optional security patches and software/firmware updates on a risk-benefit basis. Staff apply all mandatory security patches and software/firmware updates. Staff may or may not apply optional patches or updates.

The UNT Libraries have a commitment of funding for regular software replacement and upgrade. The UNT Libraries have no fixed schedule for software replacement, but relies on internal performance reports and software support availability to indicate when software change is needed.

When software changes are indicated, staff in the Digital Libraries Division collaborate with personnel from Lib-TACO to evaluate software alternatives and the timing of software changes on a cost-benefit basis.

The Digital Libraries Division in collaboration with Lib-TACO tests and evaluates changes to software in isolated development environments to minimize the risk that changes will disrupt normal operations or cause data corruption or loss. Developers cannot write changes directly to the production server. In addition to automated tests and checks performed by programs, Digital Libraries Division staff manually evaluate changes by examining samples of the relevant content. Staff evaluate changes for their effect on the integrity and understandability of information, the speed and efficiency of the system, and the accessibility and usability of disseminated information.

The Digital Libraries Division retains historical versions of software so that changes to critical processes can always be reversed. The repository’s backup process makes regular copies of the code versioning system.

To support long-term infrastructure planning, Lib-TACO has an inventory of hardware and software.

**Software obsolescence**

**Dependencies**

The UNT Libraries' Digital Collections depend on software that interoperates well with hardware, data, and applications inside and outside the repository. In addition, dissemination software and web services must meet the evolving needs and emerging expectations of the repository’s designated community. The Digital Libraries Division uses a combination of commercial, open-source, and custom-built software to ingest and transform content, manage backup, disseminate content, and deliver various services.
File format obsolescence is a closely-related problem. In this document, it is analyzed as a stand-alone topic below.

**Likelihood and impact assessment**

<table>
<thead>
<tr>
<th>Impact</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is very difficult to assess the likelihood and impact of software obsolescence because it depends on the repository growth rate as well as developments in hardware, software, and user behavior. It may affect diverse users in different ways. Degraded performance is the first sign of software obsolescence, but service outages should be expected. The impact could vary from low to medium (and possibly high, not illustrated) in relation to the services affected and the duration of any service outage. Software obsolescence is high-likelihood in the context of repository growth and wider technological change.

There is a risk that large-scale software migration could cause temporary service outages and lead to data corruption or loss.

While the Digital Libraries Division makes every effort to ensure that it offers web services that meet the needs and expectations of its designated community, the repository cannot guarantee that users have up-to-date and compatible software environments on their computers. For this reason, there is a risk that some users could experience software-related problems that the repository cannot control.

**Management strategy**

The Digital Libraries Division’s strategy for minimizing the risk of software obsolescence is largely the same as its strategy for software failure. Please see above, for details.

To help users anticipate and manage certain software-related issues, the Digital Libraries Division added the following warning to The Portal to Texas History, the UNT Digital Library and The Gateway to Oklahoma History websites: "We recommend viewing this system with Firefox
3, Google Chrome 3, Internet Explorer 7, Opera 10, Safari 3, or their newer versions.” A link to the repository’s help services and feedback forum is available on the user interface.

File format-related obsolescence

Dependencies

The UNT Libraries’ Digital Collections ingest, preserve, and disseminate a variety of file formats. While the repository is not dependent on or restricted to any particular format or group of formats, it aims to use well-known, widely-accepted formats that support long-term preservation. Please see the “UNT Libraries Digital Preservation Policy Framework” and “Preferred File Formats” documentation for more information about the repository’s format policies and practices.

Likelihood and impact assessment

<table>
<thead>
<tr>
<th>Impact</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is difficult to assess the likelihood and impact of format-related obsolescence because it depends on developments in formats, software, and user behavior. With respect to the formats that the Digital Collections have preserved to date, the likelihood of obsolescence is generally low. If obsolescence occurs, the impact could vary from medium to high in relation to the duration of the problem. It can be a low-impact event (not illustrated) when the Digital Libraries Division moves proactively to migrate files from obsolete formats to accessible formats and users move swiftly to update any relevant software. The probability of a low-impact event increases when the formats in use, even though they have been superseded, remain widely known and well documented.

There is a risk that large-scale transformations could cause temporary service outages and lead to data corruption or loss.
While the Digital Collections make every effort to preserve and disseminate formats that meet the needs and expectations of the designated community, the repository cannot guarantee that users have up-to-date and compatible software on their computers. For this reason, there is a risk that some users could experience format-related problems that the repository cannot control.

If format transformation is necessary, the Digital Libraries Division will give priority to maintaining the intellectual content contained in an individual object over preserving its appearance or a specific presentation. Consequently, some users may perceive that information has been altered or lost.

**Management strategy**

The repository’s first line of defense against format obsolescence is its ongoing efforts to preserve and disseminate formats that meet the needs and expectations of its designated community. To this end, the Digital Libraries Division monitors the digital curation field for broad trends and emerging standards, carries out extensive usability testing, and solicits feedback from its designated community.

The repository does not have a prescribed threshold or metric that would initiate format migration. The necessity and urgency of format migration will be evaluated on a case-by-case, cost-benefit basis.

To ensure the integrity of information during a large-scale transformation, the Digital Libraries Division would perform extensive testing, validation, and logging in an isolated development environment before initiating the transformation. After transformation, random samples of migrated content would be examined manually to assess their fidelity to the original. The migration and any related information would be recorded as an event in the preservation metadata for each object and the Digital Collections would retain the original objects in the repository.

While the Digital Libraries Division does not have an explicit list of accepted file types, a documented list of commonly accepted formats and standardized normalization procedures provides a framework. This documentation is included in the Digital File Formats documentation on the Digital Projects Unit’s website.

**Loss of critical hardware or software support**

**Dependencies**

The UNT Libraries’ Digital Collections rely on hardware and software supplied by commercial vendors to fulfill its mission. In particular, the Dell servers, Dell storage arrays, and VMWare
virtualization environments are central to the ongoing operation of the repository. Vendors could withdraw support if they discontinue a product, if their business strategies change, or if they encounter financial difficulties.

**Likelihood and impact assessment**

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td><img src="image" alt="Green Circle" /></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Loss of support may not lead to service outages in the short-term, but it could force the repository to carry out substantial migrations. While disruptions may be limited to the repository’s administration and maintenance, the large-scale migration of data, databases, or applications could lead to service outages and data corruption or loss.

**Management strategy**

For servers, storage arrays, and virtualization environments, Lib-TACO works with well-known, commercially-successful vendors who have large user bases.

The UNT Libraries employ a number of experienced systems administrators and programmers to oversee the repository’s technical operations at various levels. When hardware and software changes are indicated, they work closely with systems administrators in Lib-TACO to evaluate alternatives, timing, and expected outcomes. For more information about the Digital Libraries Division’s procedures for hardware and software change please see sections above.

**Manmade threats**

**Operator error**

**Dependencies**
The successful operation of the UNT Libraries’ Digital Collections relies on cautious, skilled personnel to design systems, write software, perform routine maintenance, and correct errors. Minimizing or preventing operator error is crucial to protecting the long-term integrity of information in the repository and ensuring the reliability of the organization’s services.

**Likelihood and Impact Assessment**

![Likelihood and Impact Assessment Diagram](image)

Operator error is a high-likelihood, low- to medium-impact event. People will inevitably make mistakes in the course of their work, and therefore operator error is a high-likelihood event. The impact will vary according to the systems and information involved. In general, it is unlikely that a single error will lead to widespread data loss or a prolonged service outage. However, efforts to fix operator error may lead to temporary service outages.

**Management Strategy**

The Digital Libraries Division has automated its critical processes in order to minimize the risk of operator error. Ingest, data management, archival storage, and dissemination are carried out by applications and use industry-standard error detection and quality-control measures.

Lib-TACO grants authorizations and administers access controls with the intention of maintaining a high level of security and stability. The UNT Libraries’ Network Manager in collaboration with the Digital Libraries Division authorizes each staff member with limited access to system functionality based on his or her assigned duties. The UNT Libraries: TRAC Conformance Document section C3.3 provides a general outline of the relationship between staff roles and specific duties. Only systems administrators can make changes to access controls.

Only systems administrators and core developers can write changes to the production servers or file system.
Regular fixity checks help to detect unauthorized changes to AIPs, after which Digital Libraries Division staff can initiate recovery processes and revert the AIP to a known good state.

Only Lib-TACO systems administrators have access to the UNT Libraries server room. Only Lib-TACO can grant authorization to enter the facility.

Users retrieve disseminated information from read-only HTTP-based web services and not from the core storage directories as a mounted file system.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Sabotage by insider

Dependencies

In order to carry out mission of the UNT Libraries’ Digital Collections, the Digital Libraries Division provides its employees with appropriate and limited access to information and technologies that are subject to licenses, agreements, terms of service, access policies, and security controls that are not generally accessible by the designated community. The operation of the UNT Libraries' Digital Collections depends on the discretion, confidentiality, and lawful behavior of its employees.

Likelihood and impact assessment

<table>
<thead>
<tr>
<th>Impact</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>
Sabotage is a low-likelihood event. The impact may be medium or high because sabotage is a deliberately malicious act by an individual (or group) who knows, to a certain extent, how the system has been configured and protected. It may be difficult or time-consuming to identify the extent of the sabotage and repair the damage. Efforts to repair any damage may require temporary service outages.

**Management strategy**

The Digital Libraries Division reduces the risk of sabotage by offering its employees competitive compensation and benefits, clear and reasonable performance expectations, constructive and practical feedback, opportunities for advancement and professional development, and confidential mechanisms for dispute resolution.

UNT Libraries’ Digital Libraries Division’s strategy for minimizing unauthorized access to the repository is largely the same as its strategy for operator error. Please see above, for details.

**Cyber attack**

**Dependencies**

The operation of the UNT Libraries' Digital Collections relies on the continuous, error-free operation of its technology infrastructure and on the authenticity and integrity of data to carry out its mission and fulfill its obligations. The Digital Collections are committed to ensuring that the designated community can access authentic information in a safe and efficient manner. UNT must have security controls that minimize violations of its access policies and licenses.

**Likelihood and impact assessment**

The UNT Libraries expect that cyber attacks will occur often, though the impact will vary considerably in relation to the nature of the attack. On the one hand, some cyber attacks will
have little (or no) effect on data or services. These attacks include programs designed to use
Digital Collections as a platform or proxy for attacks on other targets. On the other hand, some
attacks will affect service in obvious ways, either by impairing performance, causing service
outages, or by damaging data. These attacks include automated content harvesting,
denial-of-service attacks, and purely destructive attacks. Efforts to close the vulnerability, repair
the affected components, and restore data could lead to temporary service outages.

Management strategy

The Digital Libraries Division staff work closely with Lib-TACO personnel to minimize the risks
associated with cyber attack. The UNT Libraries use current, widely-accepted, industry-standard
procedures to reduce vulnerability and respond to attacks.

Lib-TACO deploys a variety of techniques and tools to monitor the repository’s hardware and
software. Systems administrators receive information about system behavior and usage from a
number of custom-built scripts, a Nagios monitoring program, malware and virus detection
programs, and monitoring functionality built into hardware. These tools warn administrators
about abnormal activity such as excessive processor loads and slow response times. In
addition, staff monitor critical processes, such as ingest and data management, for malfunctions
and suboptimal performance.

Lib-TACO and Digital Libraries Division staff evaluate all mandatory and optional security
patches and software/firmware updates on a risk-benefit basis. Staff apply all mandatory
security patches and software/firmware updates. Staff may or may not apply optional patches or
updates.

The Digital Libraries Division and Lib-TACO have network protection mechanisms in place,
including firewalls, proxies, packet filtering routers, intrusion detection systems, and malware
and virus detection programs.

Regular fixity checks help to detect unauthorized changes to AIPs, after which Digital Libraries
Division staff can initiate recovery processes and revert the AIP to a known good state.

The Digital Libraries Division does not limit bandwidth or data transfer volume for users unless
there is a clear breach of security or terms of service. Pre-defined limits (caps) can sometimes
degrade service for legitimate users. The Digital Libraries Division and Lib-TACO monitor
system performance and usage in order to identify and possibly block IP addresses that appear
to be harvesting content that causes a degradation in service for other users and may block
these IP addresses.

Access policies for the UNT Libraries' Digital Collections are exercised by a combination of
IP-based and Local Directory Authorization Protocol (LDAP) authentication that uses the UNT
Active Directory system and the UNT EUID authentication mechanism.
Users retrieve disseminated information from read-only HTTP based Web services and not from the core storage directories as a mounted file system.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Physical security incident

Dependencies

The Digital Libraries Division depends on physical security in order to protect its staff and safeguard the long-term authenticity and integrity of its data, databases, applications, and administrative documents. The Digital Libraries Division relies on security controls provided by the UNT Libraries and the University of North Texas.

Likelihood and impact assessment

Physical security incidents in connection with the Digital Libraries Division, its offices and the UNT Libraries server room are low-likelihood events, though the impact may vary in relation to the staff involved and the systems affected. Incidents in connection with other areas of the Willis Library are medium-likelihood, low-impact events (not illustrated).

Management Strategy
Physical access to the Digital Libraries Division office space is managed by the UNT Libraries Facilities and Systems.

The UNT Police are responsible for the physical security of all campus properties including Willis Library. The UNT Police work in collaboration with the UNT Libraries Facilities and Systems Division.

The UNT Libraries Facilities and Systems provide an on-call service for situations when staff or librarians within the library have security issues. This service is implemented as a “Help Me” button available on all staff machines.

UNT Libraries Facilities and Systems coordinates physical security and access control for the Digital Libraries Division administrative offices. Entry to the Digital Libraries Divisions’ administrative office is controlled by electronic locks at all hours.

UNT Police actively patrol the UNT Denton campus and visit the Willis Library building as a part of their regular route. The UNT Police are available 24/7 in an on-call capacity if needed.

Access to the UNT Libraries’ off-site storage facility for backup copies at the UNT System Data Center at the UNT Discovery Park research campus is restricted to authorized personnel at all hours. Access to this computing facility is controlled by UNT IT Shared Services.

UNT Libraries Facilities and Systems coordinates physical security and access control for the UNT Libraries Willis server room. Only UNT Libraries Facilities and Systems can grant authorization to enter the facility, and access is restricted to authorized personnel.

**Natural threats**

**Fire**

**Dependencies**

The Digital Libraries Division offices, work areas, and primary systems are located in a large building and are vulnerable to fire. The Digital Libraries Division depends on fire detection and suppression systems provided by the UNT Libraries and the University of North Texas Facilities Fire Safety department.

**Likelihood and impact assessment**
Fire is a low-likelihood event. The impact of a fire will vary according to the personnel and systems affected. Prolonged service outages are possible. Even if there is no evident damage to the Digital Collections systems, a fire could force the UNT Libraries to shut down service in order to permit assessment and repair of the facility. A fire elsewhere in the building or on the campus could affect network cables that deliver the repository’s information to the internet, effectively causing a service outage for a period of time. A fire in the surrounding region could affect electrical power for short or long periods of time. A fire that prevents staff from carrying out their duties (e.g., if staff cannot access the office) will likely be a temporary, low-impact event.

**Management strategy**

The Digital Libraries Division complies with emergency preparedness and response procedures designed by the UNT Libraries and the University of North Texas Facilities Fire Safety department.

In the event of evacuation, staff will follow procedures outlined in the UNT Libraries Emergency Manual

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Fire suppression in the UNT Libraries server room is based on heat and smoke detectors and automatic Halon fire suppression system.

The UNT Libraries have fire detection and suppression systems throughout the Willis Library.
Upgrades and expansion to existing wet fire suppression is in the planning/implementation process as part of a Willis Library Mechanical Electric and Plumbing (MEP) refresh scheduled for 2015-2016.

Flood
For plumbing failure in Willis Library, see “Plumbing” below.

Dependencies
In order to ensure the long-term preservation of content and the reliable and efficient dissemination of information to its designated community through the UNT Libraries’ Digital Collections, the Digital Libraries Division relies on stable environmental conditions in its offices, the UNT Libraries server room, and its off-site storage facility. In addition, the Digital Libraries Division depends on stable, continuous electrical power.

Likelihood and impact assessment

<table>
<thead>
<tr>
<th>Impact</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>

Flooding due to rain, snowmelt, or utility failure is a low-likelihood event in Denton and/or the surrounding area. The impact will vary according to the location and severity of the flood. There is low- to medium- risk of physical damage to the Digital Collections because the repository and its technology infrastructure are located on the lower level of a large building. Additionally, flooding in the basement of the building could affect network cables that carry information from the repository to the internet. Flooding that causes extended electrical failures could lead to service outages (see below, for analysis of electrical failure). For these reasons, the impact of flooding may be high (not illustrated). The likelihood of high-impact events is low. Flooding that prevents staff from carrying out their duties (e.g., if staff cannot travel to the office) is likely a temporary, low-impact event.
The Willis Library is located in a Zone X area — i.e., an “Area determined to be outside of the 0.2% annual chance of floodplain” — according to the Federal Emergency Management Agency (FEMA) Flood Insurance Rate Map Number 48121C0360G revised April 18, 2011.

Management strategy

The Digital Libraries Division complies with emergency preparedness and response procedures designed by the UNT Libraries and the University of North Texas. In the event of evacuation, staff will follow procedures outlined in UNT Libraries’ Emergency Manual.

For local water problems, the Digital Libraries Division depends on facilities management provided by the UNT Libraries and UNT Facilities.

The Digital Libraries Division operations are split between the lower level and the third floor of the Willis Library. If flooding occurred in the lower level, temporary offices could be arranged either on the third floor or somewhere else in the building.

While the UNT Libraries Server Room is located in the lower level, several strategies to minimize risk of damage due to flooding are in place:

- All servers are mounted on racks and have uninterruptible power supply (UPS) systems
- Water detection systems alert network administrators if water is present on the floor in the server room
- Power in the server room is positioned overhead and never run along the floor

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Severe weather (thunderstorm, blizzards, tornado)

Dependencies

The Digital Libraries Division depends on safe and typical weather conditions in order to carry out day-to-day operations, safeguard archival storage, and disseminate information via the UNT Libraries’ Digital Collections to its designated community.

Likelihood and impact assessment
Severe weather is a high-likelihood, low- to medium-impact event. Tropical storms, tornadoes, severe thunderstorms, heavy rainfall, heavy snowfall, and other extreme weather conditions have occurred in Denton and/or the surrounding region. The chief risk to the Digital Collections are from weather events that cause extended power failures. While the repository and associated server infrastructure have an uninterruptible power supply (UPS) for short-term operation of the storage array and servers, it does not have a live mirror site. Prolonged power failures could lead to service outages. For that reason, severe weather can be a high-impact event (not illustrated). The likelihood of high-impact events is low. Severe weather that prevents staff from carrying out their duties (e.g., if staff cannot travel to the office) is likely a temporary, low-impact event.

**Management strategy**

The Digital Libraries Division complies with emergency preparedness and response procedures designed by UNT Libraries and the University of North Texas. In the event of evacuation, staff will follow procedures outlined in UNT Libraries’ Emergency Manual.

The Willis Library building is a large, thick-walled, reinforced concrete structure. The UNT Libraries’ server room is located in the lower level of the building, away from the above-ground exterior walls and windows.

The Digital Libraries Division depends on power infrastructure and facilities management provided by the UNT Libraries and UNT Facilities. See below for details about the Digital Libraries Division management strategy for electricity-related problems.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.
Earthquake

Dependencies

In order to ensure the long-term preservation of content and the reliable and efficient dissemination of information to its designated community through the UNT Libraries’ Digital Collections, the Digital Libraries Division relies on stable environmental conditions in its offices, the UNT Libraries server room, and its off-site storage facility. In addition, the Digital Libraries Division depends on stable, continuous electrical power.

Likelihood and impact assessment

Earthquake is a low-likelihood, low-impact event for the City of Denton. The University of North Texas and the UNT Libraries are located in a location with a seismic design category of Zone A (lowest risk) by the Federal Emergency Management Agency (FEMA).

Management strategy

The Digital Libraries Division complies with emergency preparedness and response procedures designed by UNT Libraries and the University of North Texas. In the event of evacuation, staff will follow procedures outlined in the UNT Libraries’ Emergency Manual.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.
Utility or environmental/building system failure

Electrical failure

Dependencies

The UNT Libraries’ Digital Collections require continuous, stable electrical power at Willis Library to maintain the collections, run various services, and disseminate information to the designated community. The repository depends on electrical infrastructure within the building and on the electrical utilities that serve Denton and northern Texas. In addition, the continuity and stability of electrical power at UNT’s Discovery Park research campus that houses off-site replicated storage is important for maintaining environmental conditions and security systems.

Likelihood and Impact Assessment

Electrical failure is a high-likelihood, low- to medium-impact event. The impact will vary according to the duration of the power failure. The uninterruptible power supply (UPS) can keep the repository’s storage array and servers running for a short period of time. Since the repository does not have a live mirror site, extended electrical problems could lead to service outages. Prolonged outages are low-likelihood (not illustrated).

Power spikes due to lightening strikes or utility irregularities (not illustrated) are a species of electrical failure and could damage the repository’s hardware, software, and data. These are low-likelihood events.

Management Strategy

The Digital Libraries Division complies with emergency preparedness and response procedures designed by UNT Libraries and the University of North Texas. In the event of evacuation, staff
will follow procedures outlined in UNT Libraries’ Emergency Manual. The UNT Libraries’ have an auxiliary power source that will provide adequate lighting for safe evacuation.

Electricity and power infrastructure for the Digital Libraries Division is managed by Lib-TACO and the UNT Libraries in conjunction with UNT Facilities.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Lib-TACO maintains per-rack uninterruptible power supplies (UPS) for the repository’s storage arrays and servers. The UPS provides continuous power in the event of short electrical disruptions. If an electrical failure lasts longer than the battery life of the UPS, the power supply will initiate a controlled shutdown of the systems.

Electricity for the repository’s server and storage passes through a power conditioner and protection circuit that ensures stable, clean power.

**Plumbing failure**

**Dependencies**

In order to ensure the long-term preservation of content and the reliable and efficient dissemination of information to its designated community through the UNT Libraries’ Digital Collections, the Digital Libraries Division relies on stable environmental conditions in its offices, the UNT Libraries server room, and its off-site storage facility.

**Likelihood and Impact Assessment**
Plumbing failures that affect the Digital Collections are low-likelihood events, but the impact will vary from low to medium depending on the systems involved and personnel affected. Water leaks that contact electrical devices are a serious threat to human safety, and therefore flooding could force the Digital Libraries Division and Lib-TACO to shut down systems until the situation can be evaluated and resolved. The worst-case scenario is water flooding into the UNT Libraries’ server room from the floors above. Plumbing failures in other parts of the building could affect electrical power and/or network cables that carry the repository’s content to the Internet.

**Management Strategy**

The Digital Libraries Division depends on facilities management provided by the UNT Libraries and UNT Facilities. Lib-TACO monitors and manages conditions in UNT Libraries’ server room.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

**Server room cooling failure**

**Dependencies**

Cooling of servers and storage arrays is important for the integrity of data, databases, and applications. Server room cooling depends on electrical power for operation. See above for the Digital Libraries Division’s risk analysis and management strategies for electrical failure.

**Likelihood and Impact Assessment**
Heating, air conditioning, and air quality failures are medium-likelihood, medium to high-impact events because they are dependent on electrical power. Any power failure in the Willis Library building or the surrounding region could shut down environmental systems in the Digital Libraries Division offices. Local or regional power failures may shut down systems at the UNT System Data Center at the UNT Discovery Park research campus (the off-site server facility). Temporary environmental problems that affect human comfort may disrupt the administration of the repository but should not lead to service outages or data loss. In the case of a long-term disrupt in power to the server room cooling system, Lib-TACO will shutdown systems so that they do not overheat and suffer permanent damage. Extended cooling problems could lead to service outages. Prolonged outages are low-likelihood (not illustrated).

Management Strategy

The Digital Libraries Division depends on facilities management provided by the UNT Libraries and UNT Facilities. Lib-TACO monitors and manages environmental conditions in the UNT Libraries’ server room.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

Lib-TACO operates a redundant cooling infrastructure in the UNT Libraries’ server room. There is adequate capacity to operate the cooling systems in the event that one should fail because of mechanical or electrical issues.

Upgrades and expansion to existing server room cooling infrastructure is in the planning/implementation process as part of a Willis Library Mechanical Electric and Plumbing (MEP) refresh scheduled for 2015-2016.
Heating, air conditioning, or air quality failure

Dependencies

In order to ensure the long-term preservation of content and the reliable and efficient dissemination of information to its designated community through the UNT Libraries’ Digital Collections, the Digital Libraries Division relies on stable environmental conditions in its offices, the UNT Libraries server room, and its off-site storage facility.

Likelihood and Impact Assessment

![Likelihood and Impact Assessment Chart]

Heating, air conditioning, and air quality failures are low-likelihood, low-impact events. Any extended power failure in the Willis Library building or the surrounding region could shut down environmental systems in the Digital Libraries Division offices. Local or regional power failures may shut down systems at the UNT System Data Center at the UNT Discovery Park research campus. Temporary environmental problems that affect human comfort may disrupt the administration of the repository but should not lead to service outages or data loss.

Management Strategy

The Digital Libraries Division depends on facilities management provided by the UNT Libraries and UNT Facilities. Lib-TACO monitors and manages environmental conditions in the UNT Libraries’ server room.

The Digital Libraries Division and Lib-TACO carry out regular backup of data, databases, and applications according to standard backup practices. These backups are intended to serve as the basis for restoration of server infrastructure materials in the event of data corruption or loss. Data stored as part of the UNT Libraries’ Coda repository is migrated between two storage
facilities, one in the Willis Library, and one in the UNT System Data Center at the UNT Discovery Park research campus.

If conditions occur that would require the temporary displacement of Digital Libraries Division staff, arrangements can be made for these staff to work remotely.

Upgrades and expansion to existing heating, cooling, and air quality infrastructure is in the planning/implementation process as part of a Willis Library Mechanical Electric and Plumbing (MEP) refresh scheduled for 2015-2016.

References


http://www.library.unt.edu/policies/ accessed September 14, 2015
UNT Libraries Digital Workflow Diagram

Date: October 2015

Version: 1.0

Contributors:

- Mark Phillips  Assistant Dean for Digital Libraries
- Ana Krahmer  Supervisor, Digital Newspaper Unit
- Hannah Tarver  Department Head, Digital Projects Unit
- Daniel Alemneh  Supervisor, Digital Curation Unit
- Laura Waugh  Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
UNT Libraries Digital Workflow Diagram

- Content Acquisition
- Digitized Content
- Born Digital Content
- Harvested Web Content
- Content Ingest
- Coda (Preservation)
- Aubrey (Access)
- Metadata Creation / Curation
UNT Libraries Ingest Workflow Diagram

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Ana Krahmer Supervisor, Digital Newspaper Unit
Hannah Tarver Department Head, Digital Projects Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
UNT Libraries Ingest Workflow Diagram

Content Acquisition

Content Staging

SIP Creation

Package Verification

AIP Creation

Coda

DIP Creation

Aubrey
UNT Libraries SIP-to-AIP Conversion Workflow

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Hannah Tarver Department Head, Digital Projects Unit
Ana Krahmer Supervisor, Digital Newspaper Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
UNT Libraries SIP-to-AIP Conversion Workflow

Introduction

The UNT Libraries has a formal definition of the required features and elements for a Submission Information Package (SIP), documented in the UNT Libraries OAIS Information Package Specification. That document, while describing what should be present in a valid Submission Information Package (SIP) and Archival Information Package (AIP), does not cover the process used for converting a SIP to an AIP. The goal of this document is to establish the process and document the workflow for the conversion of a SIP to an AIP for the UNT Libraries’ Digital Collections.

Ingest Dropbox

The conversion process starts when a SIP is submitted to one of the different ingest dropboxes configured for the ingest process. The difference in the dropboxes is in how the names (ARK identifiers) for objects are assigned. One dropbox per namespace is assigned. We currently create identifiers under the metapth, metadc and metarkv namespace.

Dropboxes define the steps involved in the ingest process of our digital content. They are responsible for creating the Archival Information Packages (AIPs) and Dissemination Information Packages (DIPs). Locally, DIPs are called Access Content Packages or ACPs for our Coda repository and Aubrey access systems. A dropbox encapsulates a series of steps that are executed in sequential order on submitted digital objects. These steps are represented by folders in the dropbox that hold the input, output, and errors for each of the steps. The management python scripts such as makeAIP.py and makeACP.py are used to convert a SIP to an AIP and an AIP to a DIP/ACP.
A dropbox is organized like this:

dropbox/
| 0.Staging/
| 1.ToAIP/
| 2.ToAIP-Error/
| 3.ToACP/
| 4.ToACP-Error/
| 5.ToArchive/
| 6.ToAubrey/
| 7.ToAubreySorted/
| 8.ToAubreySorted-Error/
| dropbox_config.py
| makeACP.py
| makeACPSort.py
| makeAIP.py
| moveToCODA.sh

Verification steps

The verification steps that makeAIP.py execute before processing a submitted SIP include the following:

1. Check to see that the proper utilities and versions are installed on the ingest server.
2. Check that the appropriate number server is online (number server provides object names)
3. Check the Namaste tag in the bag to make sure it is a SIP bag
4. Fully validate the bag based on the BagIt specification
5. Verify the conformance to UNTL-FileSets for each manifestation in the SIP.

Objects to ingest are loaded by an ingest operator into the 1.ToAIP folder and then the makeAIP.py script is executed, which walks through the five steps listed above to determine if a supplied SIP can be converted into a UNTL-AIP.

If makeAIP.py fails on check 1 or 2 in the above list, the whole process stops. If these two steps pass, makeAIP.py will begin multiple sub-processes that allow for parallel processing of digital objects and therefore take advantage of available processors on the ingest server. If there is an error in one of the 3, 4 or 5 checks in the list above, the process moves the object to the 2.ToAIP-Error folder with log file describing the error encountered and will continue to process the next object in the queue. When makeAIP.py is finished running, it will let the operator know how many SIPS were processed, the number that were successfully converted into AIPs, and the number that failed.
SIP to AIP Conversion

Once the makeAIP.py script verifies that the submitted SIP is valid and well formed, it executes a series of steps to complete the AIP creation. The steps include the following:

1. Retrieve instructions from coda_process.py
2. If magicknumbers is set to true, verify that files are valid magicknumbers and use these to create sequential ordering for the resulting METS document.
3. If pageNumbers.txt is present use these values as order labels in the resulting METS document.
4. Identify primary bitstream of fileSet
5. Process all files and create a PREMIS Object XML file, JHOVE output stream and UNIX File output.
6. Associate optional order labels, or annotations with fileSets.
7. Generate a UNTL METS Archival Information Package Profile XML document conforming to the METS Profile -
   http://www.loc.gov/standards/mets/profiles/00000045.xml

If all of these steps are completed successfully then an AIP is created in the 3.ToACP folder and its resulting BagIt bag is updated with the newly created JHOVE and METS files. If there were any errors encountered (usually caused by malformed filesets or incorrectly formatted magicknumbers) it is moved to the 2.ToAIP-Error folder and the issue is logged with the object so it can be fixed and reprocessed in the future.
Example SIP and AIP

Input SIP

1999.001.001/
|-- 0=UNTL_SIP_1.0
|-- bag-info.txt
|-- bagit.txt
|-- coda_directives.py
|-- data/
| |-- 01_tif/
| | |-- 1999.001.001_01.tif
| | |-- 1999.001.001_02.tif
| `-- metadata.xml
`-- manifest-md5.txt

Resulting AIP

metaph1234/
|-- 0=UNTL_AIP_1.0
|-- bag-info.txt
|-- bagit.txt
|-- coda_directives.py
|-- data/
| |-- data/
| | |-- 01_tif/
| | | |-- 1999.001.001_01.tif
| | | |-- 1999.001.001_02.tif
| | `-- metadata/
| | | |-- 17711fdb-2e25-4566-bf3f-da172a12190.jhove.xml
| | | |-- 9639aaca-397a-4c90-8851-52b6f04c4d8d.jhove.xml
| | `-- metadata.xml
| `-- metaph1234.aip.mets.xml
`-- manifest-md5.txt
UNT Libraries AIP-to-DIP/ACP Conversion Workflow

Date: October 2015

Version: 1.0

Contributors:

Mark Phillips Assistant Dean for Digital Libraries
Hannah Tarver Department Head, Digital Projects Unit
Ana Krahmer Supervisor, Digital Newspaper Unit
Daniel Alemneh Supervisor, Digital Curation Unit
Laura Waugh Repository Librarian for Scholarly Works

This work is licensed under a Creative Commons Attribution 4.0 International License.
UNT Libraries AIP-to-DIP/ACP Conversion Workflow

Introduction

The UNT Libraries has a formal definition of the required features and elements for an Archival Information Package (AIP) documented in the UNT Libraries OAIS Information Package Specification. That document, while describing what should be present in a valid Archival Information Package (AIP) and Dissemination Information Package (DIP), does not cover the process used for converting an AIP to a DIP. The goal of this document is to establish the process and workflow for the conversion of an AIP to a DIP for the UNT Libraries’ Digital Collections. Historically the UNT Libraries has referred to its Dissemination Information Package (DIP) locally as the Access Content Package (ACP) and for the purpose of this document they are equivalent and can be used interchangeably.

Ingest Dropbox

Dropboxes define the steps involved in the ingest process of our digital content. They are responsible for creating the AIPs and DIPs for our repository, locally referred to as the Coda repository and Aubrey access systems. A dropbox encapsulates a series of steps that are executed, in sequential order, on the submitted digital objects. These steps are represented by folders in the dropbox that hold the input, output, and errors for each step. The management python scripts, such as makeAIP.py and makeACP.py, are used to convert a SIP to an AIP and an AIP to a DIP/ACP, respectively.

The conversion process proceeds after the successful conversion from a submitted Submission Information Package (SIP) to an AIP that is covered in the document UNT Libraries SIP-to-AIP Conversion Workflow. As can be seen from the dropbox organization list below, this newly created AIP resides in the 3.ToACP folder, which is designated by Dropbox during ingest.
A dropbox is organized like this:

```
dropbox/
|-- 0.Staging/
|-- 1.ToAIP/
|-- 2.ToAIP-Error/
|-- 3.ToACP/
|-- 4.ToACP-Error/
|-- 5.ToArchive/
|-- 6.ToAubrey/
|-- 7.ToAubreySorted/
|-- 8.ToAubreySorted-Error/
|-- dropbox_config.py
|-- makeACP.py
|-- makeACPSort.py
|-- makeAIP.py
`-- moveToCODA.sh
```

### Verification steps

The verification steps that `makeACP.py` executes before processing an AIP include the following:

1. Checks to see that the proper utilities and versions are installed on the ingest server.
2. Check the Namaste tag in the bag to make sure it is an AIP bag.
3. Quickly validate the BagIt bag using the oxnum value in the bag-info.txt file.

Objects to convert to DIP are present from the previous SIP to AIP process and reside in the 3.ToACP folder, the `makeACP.py` script is executed, which walks through the three steps listed above to determine if a supplied AIP can be converted into a UNTL-DIP.

If `makeACP.py` fails on check 1 in the above list, the whole process stops, if this step passes then `makeACP.py` will start multiple sub-processes that allow for parallel processing of digital objects and therefore takes advantage of available processors on the ingest server. If there is an error in either step 2 or 3 in the list above, the process moves the object to the 4.ToACP-Error folder with a log file describing the error encountered and continues to process the next object in the queue. When `makeACP.py` is finished running it will let the operator know how many AIPs were processed, the number that were successfully converted into a DIP/ACP, and the number that failed.
AIP to DIP/ACP Conversion

Once the makeACP.py script verifies that the submitted AIP is valid and well formed, it executes a series of steps to complete the DIP creation. The steps include the following:

1. Retrieve instructions from coda_process.py
2. Based on the primary bitstream in each fileSet, create a web derivative file format for each fileSet.
3. If a metadata.xml record is not present generate a generic metadata.xml record.

If all of these steps are completed successfully then a DIP/ACP is created in the 6.ToAubrey folder and the AIP that was processed is moved to the 5.ToArchive folder. If there were any errors encountered, the AIP is moved to the 4.ToACP-Error folder and the issue is logged with the object so it can be fixed and reprocessed in the future.
Example AIP and DIP/ACP

Input AIP

metaphth1234/
|-- 0=UNTL_AIP_1.0
|-- bag-info.txt
|-- bagit.txt
|-- coda_directives.py
|-- data/
  |-- data/
  | |-- `-- 01_tif/
  | | |-- 1999.001.001_01.tif
  | | |-- `-- 1999.001.001_02.tif
  | |-- metadata/
  | | |-- 1771lfdb-2e25-4566-bf3f-daal172a12190.jhove.xml
  | | |-- `-- 9639acaoe-397a-4c90-8851-52b6f04c4d8d.jhove.xml
  | |-- metadata.xml
  | |-- `-- metaphth1234.aip.mets.xml
  |-- `-- manifest-md5.txt

Resulting DIP/ACP

metaphth1234/
|-- 0=UNTL_ACP_1.0
|-- bag-info.txt
|-- bagit.txt
|-- coda_directives.py
|-- data/
  |-- `-- metaphth1234.mets.xml
  | |-- `-- metaphth1234.until.xml
  |-- `-- web/
  | |-- `-- 01_tif/
  | | |-- 1999.001.001_01.jpg
  | | |-- 1999.001.001_01.medium.jpg
  | | |-- 1999.001.001_01.square.jpg
  | | |-- 1999.001.001_01.thumbnail.jpg
  | | |-- `-- 1999.001.001_02.jpg
  | | | |-- 1999.001.001_02.medium.jpg
  | | | |-- 1999.001.001_02.square.jpg
  | | | |-- `-- 1999.001.001_02.thumbnail.jpg
  | |-- `-- manifest-md5.txt