



Emory Libraries’ Digital Collections and Preservation Policies

Last Revised: October 2021

Table of Contents

Digital Collections Development Policy	Page 2
Digital Preservation Policy	Page 4
Digital Object Retention Policy	Page 7
Third-Party Dissemination Policy	Page 10
Glossary of Terms	Page 13

These policies guide the creation and maintenance of centrally managed and/or permanent digital collections at Emory Libraries. These policies are reviewed on at least a three-year cycle, but the primary policy steward may initiate an earlier review.

Digital Collections Development Policy

Document Status: Final

Prepared by: Chris Palazzolo, Emily Porter, Kathryn Michaelis, Lisa Macklin, Katherine Fisher

Reviewed by: LTDS Leadership, Collections Steering Committee

Approved by: Library Cabinet

Date Approved: November 4, 2021

Revision Cycle: 3 years

Last reviewed: October 2021

Next Revision: September 2024

Primary Policy Steward: Head of Collections

Purpose

Emory Libraries' digital collection and preservation efforts are designed to ensure long-term access to an increasing volume of digital materials that have intellectual significance and long-term value, and to deliver **digital collections** in an environment that allows for ease of searching, browsing, retrieval, and reuse. The Digital Collections Development Policy parallels and aligns with Emory Libraries' other collection development policies and practices, which support teaching, learning, and research needs of faculty, students, and staff.

Currently, digital content is stored and managed in a number of locations, as part of a larger ecosystem of preservation services. This policy statement describes what streams of content are appropriate specifically for deposit into Emory's centrally funded and managed preservation repository, Emory Digital Collections. Individual licensed electronic materials are not under the purview of this policy. Individual libraries across campus are responsible for determining what content, physical or digital, is in scope for selecting and acquiring based on subject relevance and archival value. The Collections Steering Committee (CSC) and its subcommittees, in tandem with digital preservation program staff, inform and facilitate the collection development and collection management practices for Emory Digital Collections. The Operations and Management Team will provide guidance, as needed, on considerations of long-term storage costs that may affect scale or scope of building digital collections.

Principles for Building Digital Collections at Emory

Emory Libraries' digital collections will:

- Support the curriculum, research, and teaching missions of the university
- Enhance and widen access and use, both nationally and internationally, of Emory's rare and unique collections

Selection of Materials for Ingest into Emory Digital Collections

Selection is based on traditional collection development principles including authority, originality, curricular and research relevance, timeliness, breadth and/or depth of coverage, demand, and support of Emory's distinctive collections.

Proposed materials for inclusion in Emory Digital Collections will be evaluated for broad and enduring value. Added value components to be considered include:

- suitability for online presentation;
- intellectual control (metadata);
- improvement of resource sharing;
- advancement of collaboration;
- enhancement of access;
- intrinsic research value; and
- rarity or uniqueness of assets/content.

Emory Libraries considers the following criteria to identify materials for ingest into Emory Digital Collections for purposes of long-term access. More specific guidance on preservation characteristics is reflected in the *Digital Preservation Policy*. In order for Emory Libraries to consider ingesting digital collections into Emory Digital Collections for long-term access, digital collections should:

- Have a plan for dissemination that includes the potential for either making the material widely accessible, either now or in the future, or providing restricted access to rare, at-risk materials not otherwise available;
- Be unique or have significant research value or mitigate inherent risks for materials in analog formats. This might encompass intellectual content, carrier/format, or digital enhancements. This potentially includes preservation copies of rare, fragile, or unique collections for long-term access or data analysis, and materials that are at risk of being lost due to an obsolete media or file format;
- Meet our ethical obligations and align with existing collection policies and standards, including Emory's Open Access Policy and donor agreements;
- Abide by relevant legal statutes governing preservation and distribution of these materials, including U.S. Copyright Law;
- Meet the guidelines for descriptive metadata outlined in the Core Metadata Standards.

Digital Preservation Policy

Document Status: Final

Prepared by: Chris Palazzolo, Emily Porter, Kathryn Michaelis, Lisa Macklin, Katherine Fisher

Reviewed by: LTDS Leadership, Collections Steering Committee

Approved by: Library Cabinet

Date Approved: November 4, 2021

Revision Cycle: 3 years

Last reviewed: October 2021

Next Revision: September 2024

Primary Policy Steward: Digital Preservation Program Manager

Purpose

Emory's Digital Preservation Policy outlines the purpose and scope of Emory Libraries' digital preservation program as well as general principles to which Emory Libraries will adhere in order to ensure long-term access to its digital collections. This policy is an integral part of the Libraries' comprehensive digital preservation program. This work is done in support of Emory University's mission "to create, preserve, teach, and apply knowledge in the service of humanity."

Mandate

Emory Libraries recognizes that its digital collections are an integral part of its overall holdings, and acknowledges the necessity of preserving those digital collections in a way that is scalable, programmatic, in alignment with established digital preservation best practices, and mindful of the financial resources and labor that have been invested in the creation and acquisition of these digital collections. Specifically, Emory Libraries commits to preserving its digital collections to support the following objectives:

- **Scholarship and teaching:** Emory Libraries commits to preserving and providing access to its digitized and born-digital assets to support instruction, research, and scholarship.
- **Legal obligations:** Emory Libraries honors commitments made to preservation and ongoing access of digital assets as part of agreements such as deeds of gift and sale, donor agreements, contracts with external partners, and other legally binding agreements.
- **Institutional records:** Emory Libraries enables the preservation of electronic Emory University records designated as permanent in [retention schedules](#) and identified by the University Archives as having continuing and enduring value (see [Policy 5.21, Records Management](#)).

Scope

Emory's digital preservation program is responsible for securing, providing the means to preserve, and ensuring long-term access to Emory's **permanent digital collections**. Digital collections suitable for long-term preservation include those in Emory Digital Collections as well as those held in other systems that make up the Emory Libraries' **preservation ecosystem**.

Permanent digital collections are those that have been designated by the Emory Libraries as appropriate for long-term preservation, and can include content from the following categories:

- Digitized versions of physical assets held by the Emory Libraries
- Born-digital material acquired as part of (or entire) archival or special collections
- Digitized or born-digital non-custodial content that the Emory Libraries have committed to stewarding through memorandums of understanding or similar agreements
- Electronic theses and dissertations
- Emory faculty research materials submitted to Emory Libraries' open access research repository

Principles

Emory Libraries commits to adhering to the following principles that guide our long-term preservation of digital collections:

- We will strive to follow guidelines and implement standards that are agreed upon by the digital preservation community.
- We will maintain a secure environment for the long-term preservation of digital content.
- We will capture and preserve these required essential characteristics of digital content necessary for long-term use.
 - All of the content-bearing files that make up a **digital object**, including supplemental materials
 - Information about the relationships between the content files that make up a digital object
 - **Representation information** that makes the object interpretable to humans and machines
 - At a minimum, the descriptive metadata specified in the adopted and Library Cabinet-approved [Core Metadata Guidelines](#)
 - The copyright, license, and access status of digital objects and/or collections
 - Information about the **chain of custody** of digital objects and/or collections
 - **Fixity** information about digital objects, which is captured regularly and monitored over time

Compliance with these principles is necessary for the Emory Libraries digital preservation program to ensure long term preservation and access. Failure to meet the principles outlined here may result in an inability to ingest a digital object into a preservation system or guarantee the digital object's functionality is preserved.

Roles and Responsibilities

The roles of key participants in Emory Libraries' digital preservation program include:

- **Administrators:** leaders at the library and division levels responsible for providing material support and advocacy for digital preservation
- **Digital content managers:** individuals tasked with making and implementing decisions about the management and preservation of digital collections content
- **Digital collections stakeholders:** archivists, curators, content stewards and others who manage the collecting areas to which digital collections belong. Responsible for selecting collections for digitization and preservation and providing the necessary information to support preservation and access for those collections.
- **Technology support providers:** Libraries and university staff who develop and support the software applications and hardware used for digital preservation purposes and monitor technical best practices in the field
- **Technology leadership:** The Libraries' technology leadership sets the technology strategy for digital preservation, ensures it is financially resourced, has adequate staffing, and works with stakeholders to identify future directions for its use.

Third-Party Preservation Services

In some cases, digital content in permanent digital collections owned or stewarded by Emory Libraries may be determined to be preserved by a third-party service to supplement our local capacity and preservation services. The proposal to join or contribute to third-party preservation services should be made by LTDS leadership in a Decision Document or Business Case and approved by the Vice Provost for Libraries and Museum.

Criteria for choosing preservation services

On behalf of Emory Libraries, the Digital Preservation Program Manager, in consultation with other stakeholders as appropriate, will evaluate a third party's capacity to provide preservation for digital objects. General considerations for choosing a third-party service or vendor include:

- **Transparency:** The third-party service provider should offer transparency and disclosure of the policies and strategies it uses for preservation.
- **Business model:** The third-party service provider should exhibit economic sustainability and viability, transparency in business operations and costs, and provide a clearly defined service level agreement and/or terms of service.
- **Standards:** The third-party service provider should align with industry standards.
- **Ownership/Stewardship of Content:** The third-party service should allow Emory Libraries to retain and retrieve preservation copies of submitted digital content, and enable Emory to version, delete, or decommission submitted digital content per the *Digital Object Retention Policy*.

Digital Object Retention Policy

Document Status: Final

Prepared by: Chris Palazzolo, Emily Porter, Kathryn Michaelis, Lisa Macklin, Katherine Fisher

Reviewed by: LTDS Leadership, Collections Steering Committee

Approved by: Library Cabinet

Date Approved: November 4, 2021

Revision Cycle: 3 years

Last reviewed: October 2021

Next Revision: September 2024

Primary Policy Stewards: Digital Repository and Digital Preservation Program Managers

Purpose

Emory's Digital Object Retention Policy outlines situations in which Emory Libraries must remove or update digital objects in Emory Libraries centrally managed preservation services as part of normal business operations and as Emory Libraries' missions and collecting areas change with time. The purpose of the Digital Object Retention Policy is to identify under which circumstances a **fully ingested** digital object from a permanent digital collection may be versioned, deleted, or decommissioned from the Emory Digital Collections repository or Emory Libraries' centrally managed preservation services. Additional policy and procedure regarding take-downs of digital objects in Emory preservation repositories is outlined in the Emory Libraries' [Digital Collections Copyright and Content Policy](#). Items disseminated to third-party services for the purpose of discovery and/or preservation should also comply with the requirements outlined in this policy.

Defining Retention

All digital objects preserved in the Emory Digital Collections repository and additional Emory Libraries centrally managed preservation services are intended to be retained permanently. Emory's preservation services will not be used as a temporary storage facility for digital objects nor as a temporary access platform.

Digital objects are not always considered fully ingested at the time of initial deposit, but only after any required post-deposit processing of the digital object has been completed, at which point it is considered ingested. In some cases, human or system errors may result in the creation of digital objects that are accidental duplicates or are not fully complete, or that require further approval workflows to complete ingest: these in-process objects are excluded from the scope of this policy.

Versioning

Defining Versioning

Versioning is the act of updating or editing a digital object that has been fully ingested into the Emory Digital Collections repository or an additional Emory Libraries' centrally-managed preservation service. Where feasible, previous versions of the digital object should be retained for administrative and preservation purposes, but only the most recent version of the digital object

will be accessible to users. When this capability is not available, administrative metadata should be recorded to indicate that modifications to an object have occurred.

Principles of Versioning

A new version should be created when digital objects' essential preservation characteristics are modified. These characteristics are defined in the *Digital Preservation Policy* and are listed below for convenience:

- The intellectual content of the object.
- The descriptive metadata associated with the object.
- The copyright and/or license status of the object.
- The chain of custody of the object.
- The object's representation.
- Fixity for the object.

Versioning of a digital object should only be performed by a designated digital content steward, an authorized depositor, or a repository administrator. Fully ingested objects, because they are considered final, should be only formally versioned in response to significant preservation or access events. When feasible, the reasons for versioning the object should be recorded within the metadata of the digital object itself.

Decommissioning

Defining Decommissioning

Decommissioning is the act of removing a fully ingested, discoverable digital object from the Emory Digital Collections repository or an additional Emory Libraries centrally managed preservation service for which access has been granted to end users (e.g. the public or the Emory community). Decommissioning assumes the digital object's essential preservation characteristics will remain within the designated repository, but the content and metadata will no longer be accessible to end users.

Principles of Decommissioning

Decommissioning can only be considered for digital objects that meet one or more of the following criteria:

- The content in the digital object no longer aligns with its overarching collection strategy; or
- The digital object is unnecessarily duplicated; or
- The digital object no longer retains its integrity, identity, or authenticity; or
- The holding library has withdrawn the physical analog of the digital surrogate; or
- The right to distribute the digital object becomes unclear or unknown.

Before any digital object is recommended for decommissioning, reasonable efforts shall be made to ascertain that Emory Libraries has the right and responsibility to decommission the object.

Decommissioning Procedures

A designated digital content steward or repository administrator may recommend decommissioning a digital object if, in their best judgment, one or more criteria for

decommissioning has been met. The recommendation must be made in writing to Emory's Collections Steering Committee. Such recommendations will specify the source and/or provenance of the digital object and the reason for decommissioning. Collections Steering Committee approval is required for actual decommissioning, and all decisions will be recorded in a decision document. For recommendations to delete an object or other special circumstances, the Collections Steering Committee will refer a decision document to the Vice Provost for Libraries and Museum for final approval.

Once the object has been decommissioned from Emory Digital Collections, a **tombstone page** indicating the object is not available will be displayed in its place. To avoid administrative or system conflicts, the object's persistent identifier is retained and will not be re-used. The circumstances for decommissioning should be recorded as part of the preservation metadata for the now-dark digital object. Ideally, any Emory centrally managed preservation service with an end user interface should also follow these processes.

Deletion

Defining Deletion

Deletion is the act of removing the essential preservation characteristics of a digital object from Emory Digital Collections or an Emory Libraries' centrally-managed preservation service, including its intellectual content files. Deleting a digital object means that the digital object will no longer be maintained, preserved, or stored in Emory Digital Collections or an Emory Libraries' centrally-managed preservation service.

Principles of Deletion

Deletion can only occur for digital objects that meet one or more of the following criteria:

- Emory Libraries agrees to delete a digital object because its rightful copyright owner has issued a demand to do so or Emory received an order from a court of competent jurisdiction; or
- Emory Libraries determines that the deletion of the digital object is in the best interest of Emory Libraries; or
- Emory Libraries determines that another entity is the rightful owner of the digital object.

Before any digital object is recommended for deletion, reasonable efforts shall be made to ascertain that Emory Libraries has the right and responsibility to delete the object.

Deletion Procedures

Deletion may only be performed by designated repository administrators. Once the object has been deleted from Emory Digital Collections, a tombstone page indicating the object is no longer available may, if appropriate, be displayed in its place. While the intellectual content files will be deleted, descriptive and administrative metadata, including a record of the circumstances for deletion, should be retained for repository administrators. The object's persistent identifier will also be retained and not be re-used. Ideally, these procedures should also apply to digital content within other Emory centrally managed preservation services.

Third-Party Dissemination Policy

Document Status: Final

Prepared by: Chris Palazzolo, Emily Porter, Kathryn Michaelis, Lisa Macklin, Katherine Fisher

Reviewed by: LTDS Leadership, Collections Steering Committee

Approved by: Library Cabinet

Date Approved: November 4, 2021

Revision Cycle: 3 years

Last reviewed: October 2021

Next Revision: September 2024

Primary Policy Stewards: Head of Collections and Digital Repository Program Manager

Purpose and Scope

In order to expand access to and discoverability of Emory University's digital collections, digital content may be shared with third-party access services. This policy outlines considerations for making digital objects and/or components of a digital object available to third-party services. Although Emory Libraries may disseminate stewarded content to third parties for access, Emory University will continue to own and/or steward its digital collections.

Only digital content locally owned or stewarded by Emory Libraries is governed by this policy. This policy does not address which components of a digital object are disseminated to third-party services; this may be determined by the service itself.

The proposal to join or contribute to third-party services should be made in consultation with the LTDS leadership and Scholarly Communications (if content protected by copyright is included) and reviewed by the Collections Steering Committee (CSC). A Decision Document, and if necessary, a Business Case, should be submitted to the Vice Provost for Libraries and Museum and Libraries Cabinet for approval. Failure to seek approval may result in the Emory Libraries' being unable to comply with the requirements of the third-party service, or a granting agency, if applicable.

Dissemination for Access

Recognizing that users may desire multiple channels for accessing Emory University's digital collections, it is recommended that digital collections be as widely shared and discoverable as possible. Digital content in the public domain, or to which Emory University holds copyright, may be shared with appropriate third-party services. Digital content still protected by copyright, or with an unknown copyright status, will be reviewed for risk assessment by Scholarly Communications as part of the approval process for sharing with appropriate third-party services. Preservation copies of all disseminated digital collections will be maintained in Emory Digital Collections or an Emory Libraries' centrally-managed preservation service.

Emory Libraries will share metadata and/or digital content with third-party repositories and portals to further expose relevant Emory University digital content. Any requirements from grant or collaborative efforts for distribution of Emory University's collections should be vetted using

the criteria below by the Collections Steering Committee in consultation with LTDS leadership prior to submitting for approval.

Criteria for Choosing Dissemination Services

The following general criteria are used to determine whether Emory University's digital content should be provided to a third-party dissemination service.

- **Open Access:** The third-party service should not put digital content behind a pay wall or be commercialized.
- **Facility of Ingest:** The third-party service's method of ingest for digital content should be documented and content should be easily transferred from Emory Libraries.
- **Recognition of Emory:** The third-party service should allow for identification of Emory University as the contributor of the digital content to highlight its contribution to the greater corpus, potentially leading researchers to explore Emory University's unique digital collections more deeply; digital content should be able to be grouped as Emory collections.
- **Accessibility:** The third-party service should demonstrate a commitment to and compliance with current accessibility standards, such as WCAG.
- **Ownership and Maintenance of Content:** The third-party service should allow Emory Libraries to retain preservation copies of submitted digital content and also enable Emory to version, delete, or decommission submitted digital content per the *Digital Object Retention Policy*.
- **Protection of Patron Privacy:** The third-party service should not require personally identifiable information to use the digital content and should only report user activity at an aggregate level.
- **Search Capability and Functionality:** The third-party service should provide search and discovery tools that allow users to locate and manage their searches and the digital content.
- **Variety of file formats:** The third-party services should support the ability to export a variety of file formats for users to access and download digital content.
- **Availability of content:** The third-party service should make digital content available to a wider audience rather than restricted to institutional members.

Glossary of Terms Used in this Policy

- **Chain of custody:** the sequence of individuals or organizations that have created, modified, or transferred materials throughout their lifecycle. Chain of custody can be recorded at the individual digital object level as well as at the collection level.
- **Dark object/dark archive:** an object or set of objects not intended to be accessed by end-users and only available to a select few staff. By contrast, the Emory Digital Collections repository is a "grey" archive in which some material may be restricted at the time of ingest, with the intent of making it accessible in the future.
- **Decommissioning:** the act of removing a fully ingested digital object from the Emory Digital Collections repository or an additional Emory Libraries-approved preservation system for which access has been granted to end users (e.g. the public or the Emory community)
- **Deposit:** initial submission of digital content prior to any processing for ingest into Emory Digital Collections.
- **Descriptive Metadata:** describes content for search and discovery contexts -- helps connect users to resources and provides important context about a resource once it is discovered. This type of metadata drives the ability to search, browse, sort, and filter information.
- **Digital Collection:** an aggregation of digital materials managed collectively and based on provenance, function, format, or subject. These include digital materials created by Emory University, unique born-digital materials collected and stewarded by Emory Libraries and published digital materials that require local hosting and rights management.
- **Digital Content Steward:** a designated individual who oversees different preservation activities on the part of an object or collection.
- **Digital Object:** an individual intellectual entity, comprised of metadata, content files, and may also contain supplemental files.
- **Emory Digital Collections:** the Emory Libraries' centrally funded and managed preservation repository, as implemented through the Digital Library Program. Also known as "Cor" and "DLP".
- **Emory Libraries:** the five Emory libraries (Woodruff, Pitts Theology Library, Woodruff Health Sciences Library, Oxford College Library, and the Macmillan Law Library).
- **Fixity:** the property of a digital file remaining unchanged over time. Fixity is commonly monitored through the periodic generation and verification of values called checksums.
- **Fully Ingested:** objects that have completed their initial deposit into Emory Digital Collections or another Emory Libraries centrally-managed preservation service, and any required post-deposit processing or quality assurance has also been performed on the digital object.

- **Ingest:** the process by which a preservation repository gathers and/or generates the components of a digital object that will be preserved.
- **Permanent Digital Collections:** Those collections designated by the Emory Libraries as appropriate for long-term preservation.
- **Preservation Ecosystem:** the set of centrally-managed services that have been implemented through a formal approval process, contain permanent digital collections, and meet minimum requirements for long-term digital preservation.
- **Preservation Metadata:** information detailing preservation activities applied to a digital object in the repository such as system events, human-initiated workflows, or an audit trail of modifications. Additional information regarding the composition and encoding of digital surrogates specifically is recorded through technical/characterization metadata.
- **Representation Information:** information about the structure and meaning of the bits that make up a digital object. Representation information is necessary to convert bit sequences into meaningful information.
- **Tombstone page (placeholder page):** a user-facing page that is generated when content has been decommissioned or deleted. The page should not explain why the content has been removed but should provide contact information should the user believe they have reached the tombstone page in error.
- **Versioning:** the act of updating or editing a digital object that has been fully ingested into the Repository