Digital Collection Development and Preservation Policy

Purpose

The Digital Collection Development Policy is designed to address the full lifecycle of both digitized and born-digital content. The policy is an essential component of the Digital Library Program (DLP). It outlines principles for the selection, creation, acquisition, ingest, maintenance (of bit-streams, content, and metadata) and dissemination/access of digital collections at Emory. Digital collections, as defined under this policy, are subject to the same criteria for selection, retention and preservation as print materials, despite the difference in format. This policy addresses the additional criteria that are inherent with resources in digital format. The policy herein should be reviewed on an annual basis. The hope is that the outlined policies and parameters will serve as a means to establish an efficient ongoing digitization program at Emory.1

Principles for Building Digital Collections at Emory

The implementation of this policy will be guided by the following principles. Emory Library’s 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 and its collection strengths
- Preserve rare, fragile collections for long-term access, or materials that are at risk of being lost due to an obsolete media or file format in line with U.S. Copyright Law Section 108 requirements.

Scope of Policy

In terms of materials, this policy applies to materials in all formats in Emory Libraries collections (acquired by, purchased for, or donated to Emory Libraries). This policy covers born-digital acquisitions as well as any materials for which Emory has secured rights and permissions or for which it is the corporate author. Selected research datasets may also fall under this policy.2 The policy does not include/cover one-off digitization requests made by patrons.

Any library or LITS unit that utilizes central library functions (digitization, archiving, curation, access, and/or discovery) and wishes to have content preserved in a central repository must adhere to these policies. To emphasize, any activities that involve digitization, digital preservation and/or access and discovery of digital collections should be

---

1 This policy does not cover images or image collections that are currently available in Artstor Shared Shelf. The Shared Shelf digital images collections policy outlines the various guidelines for the inclusion of digital images in Shared Shelf. As noted in the SS policy, any digitization and/or preservation requests will continue to be governed by the Digital Collection Development and Preservation Policy.

2 Research datasets within scope of this document will be selected in consultation with LITS units responsible for providing data support and services to Emory researchers.
reviewed under this policy. Membership on the governance groups outlined here will be predicated upon adherence to these policies.³

Selection of Materials for Digitization and Oversight

Emory Libraries subscribes to the following criteria for digitization, which will be used to prioritize digitization of materials. In order to document digitization efforts and plan work and manage capacity, all digitization and preservation efforts should come through and be reviewed by the Digital Collections Steering Committee and relevant subcommittee, including all digitization and/or preservation related to ECDS digital projects, library exhibitions, and grant proposals.

- Potential for making collections widely accessible
- Materials should be limited to works (a) in the public domain; or (b) for which Emory is the corporate author of all content; or (c) for which Emory has secured permission to digitize and display; or (d) orphan works for which due diligence has been undertaken to locate copyright owners and approval from appropriate administrator has been secured.
- Preservation of rare, fragile collections for long-term access or data analysis, or materials that are at risk of being lost due to an obsolete media or file format. U.S. Copyright Law Section 108 requirements must be met.
- Digitization of collections should enhance the intellectual value and utility of the existing print or audiovisual collection
- Material to be digitized should not have been digitized elsewhere, unless digitization by Emory will add significant value.
- Prospective material should have sufficient descriptive metadata, per relevant guidelines (http://metadata.emory.edu) at least at the collection-level (but preferably at the item level) or must include plans for metadata creation or enhancement. Basic information on copyright status/permissions and conditions of access and usage for digital works in the collection should be included.
- Proposals for digitization should include a method to define and determine potential use of digitized collections.
- Proposed collections for digitization with a sense of timeliness (e.g. coincide with University event, historical milestone, important anniversary) may be prioritized over other proposals.
- An appropriate delivery system exists or a plan is in place to provide for distribution, such as Primo.
- Emory will provide access to the Dataverse for locally generated research datasets falling outside the collection policies of already established national and domain data archives and repositories.

³ Note that there is also a university-wide digital scholarship and pedagogy committee (http://it.emory.edu/about_us/it_governance/instructional_tech_members.html) whose mission is to “support IT proposals that advance the missions of digital scholarship, teaching and learning in the classroom.” The policy and governance structure outlined here focuses on the creation and maintenance of digital collections.
Implementation of Policy

The Digital Collections Steering Committee is charged with implementing the processes needed to effectuate this policy. The DCSC and its subcommittees serve as the collection development and collection management arm of the Digital Library Program. In order to document digitization and preservation efforts and plan work and manage capacity, all digitization efforts should come through and be reviewed by the Digital Collections Steering Committee (detailed under separate cover), including all digitization and/or preservation related to ECDS digital projects, library exhibitions, and grant proposals that utilize library collections.4

Ongoing Activities of the Steering Committee include:

- Maintains, executes, and periodically assesses Digital Collections Development Policy.
- Approves high-level Digital Preservation policy and levels of commitment.
- Development of clearer parameters and priorities for the development of digital collections.
- Maintains documentation (via a web presence and Box) for publicizing digitized collections, collection policies, and access to digitized content.
- Reviews, approves, prioritizes recommendations regarding the digitization and intake of new digital collections (including born or existing-digital) and related project proposals.
- Recommends external preservation partnerships (such as LOCKSS, CLOCKSS, Portico, and DPN) to Library Cabinet.
- Reviews digitization components of major grants.
- Oversees Digital Collections subcommittees and working groups.

Processes

- Develops a selection and approval mechanism for collections and assets destined for the repository and/or preservation.
- Develops a selection and approval mechanism for collections and assets in need of digitization or reformatting efforts.
- Create a communication plan to inform relevant stakeholders about which collections and assets are destined for the repository.
- Record decisions and outcomes of digital collection assessments.
- Develop minimal requirements for a collection to be included in the repository (metadata requirements, rights review, curator/steward, etc).
- Assigns individual to maintain an inventory of existing digital content – both in and out of Fedora – and provide a recommendation based on the above minimal requirements, for inclusion in the repository.

4 For example, if Events/Exhibitions wishes to digitize, preserve, or disseminate content from an exhibit, such plans will need to be reviewed by the DCSC along with relevant subcommittees.
Policies

- Review and update the Digital Collections policy; ensure that the policy provides guidance regarding standards for collections.
  - Establishes a Task Force for Digital Collections Policy Revision
- Reviews and Approves policies from the Digital Assets Working Groups (preservation commitment, external partnerships, retention, dissemination, etc)

There will be several DCSC Subcommittees, some permanent and ongoing, others more ad-hoc.

Selection and Review Subcommittee

Activities

- Assessment (technical, rights/permissions, metadata) and selection of collections for digitization and born digital materials in light of collection priorities and criteria established by the Steering Committee, for ingest into the repository
- Preservation recommendation for projects based on the Preservation Policy (Levels of Commitment)
- Recommendations for Dissemination and Distribution in conjunction with digital steward, of digital assets
- Recommendations to Steering on workflow revisions, additional funding or storage needs, preservation and dissemination/distribution channels for digital assets

Digital Assets Policy Working Groups (Ad Hoc)

These ad hoc groups will work on the development of policies related to the management and distribution of digital assets. Policies are to be approved by DCSC. Membership.

Groups may work with other existing groups, such as Metadata and Research Data Management to document policies.

- Some potential policies for drafting:
  - Levels of Commitment (digital preservation policy)
  - Submission agreements (review existing agreements and extend if needed to account for preservation actions)
  - Policies towards external partnerships as they relate to preservation
  - Documentation of current dissemination and distribution policies
  - Retention Policy

Note: More policy, approval guidelines, and processes may be needed. This is not intended to be an exhaustive list.
Glossary

The following terms are employed throughout this policy document:

**Digital Collection:** A digital collection consists of digital objects that are selected and organized to facilitate their discovery, access, and use. Objects, metadata, and the user interface together create the user experience of a collection. Note that digital collections in this context refer to unique, Emory-created collections of digitized or born digital content for delivery to an Emory Libraries repository or discovery tool.


**Archival Object:** An archival object is not the same thing as an archived file. Most often, an archival object will consist of a set or bundle of related files with associated metadata, for example, a recording in several formats.

**Preservation and Stewardship:** Digital preservation combines policies, strategies and actions to ensure the accurate rendering of authenticated content over time, regardless of the challenges of media failure and technological change. Digital preservation applies to both born digital and reformatted content. Digital preservation policies document an organization’s commitment to preserve digital content for future use; specify file formats to be preserved and the level of preservation to be provided; and ensure compliance with standards and best practices for responsible stewardship of digital information.

*(lifted from the ALCTS Preservation and Reformatting Section, Working Group on Defining Digital Preservation, [http://www.ala.org/alcts/resources/preserv/defdigpres0408](http://www.ala.org/alcts/resources/preserv/defdigpres0408)).*

**Research Data:** Research data is commonly defined as “the material collected or generated in the course of conducting research”⁵, including that information which is “necessary to validate research findings”⁶. It may take the form of raw, processed, analyzed, or published data, and the associated code, software, algorithms and metadata required to use the data⁷.

---

Chris Palazzolo (Head of Collection Management), Jennifer Doty (Data Management Specialist, ECDS), Lisa Macklin (Director, Scholarly Communications Office), Kyle Fenton (Head, Digitization and Digital Curation) and Erika Farr (Head of Digital Archives, MARBL), Emily Porter (Manager, Digital Library Program), Rosalyn Metz (Head of Digital Technology and Systems)