Cornell University Library Digital Preservation Policy Framework
December 2004

Purpose
This document formalizes Cornell University Library's (CUL) continuing commitment to the long-term preservation of its diverse and extensive range of digital assets. CUL recognizes that a fully implemented digital preservation program has a reliable and sustainable digital archive at its core, compliant with prevailing standards and practice. This program contributes to the University's mission to enrich the intellectual life of the University by fostering information discovery and intellectual growth, nurturing creativity, partnering in the development and dissemination of new knowledge, and ensuring access to this corpus of information over time. CUL is committed to realizing this digital preservation program vision.

Mandate:
CUL's mandate for digital preservation is at least five-fold:
- *Scholarship*: Realizing Cornell University's mission led to the maintenance of a library system to support scholarship, teaching, and learning. As more resources and services associated with these functions become digital, CUL's responsibilities must expand to include the identification, stewardship, and preservation of designated digital content.
- *Institutional records*: The University has charged CUL with maintaining the University Archives by collecting and preserving university records, including those in electronic format.
- *Legal obligations:* The University has mandated responsibilities to preserve and maintain access to certain digital objects, as well as responsibilities as a designated land grant institution. Some legal obligations derived from Federal and State laws require us to maintain files in an archival fashion.
- *Organizational commitment:* CUL's commitment to digital preservation is explicitly cited in the Library's *Goals and Objectives* as Goal II, objective 1, which calls for establishing a common depository system capable of ensuring systematic management and long-term preservation of digital collections, and in Goal II, objective 6, which calls for working with the University to establish a program to archive the university's electronic records. CUL has identified developing an OAI S-compliant depository as a top priority.
- *Consortial and contractual obligations:* CUL has consortial obligations and contractual agreements to assume or share in the responsibility for preserving designated digital content.

Objectives
The primary intention of the digital preservation program is to preserve the intellectual and cultural heritage important to the University. The program's objectives are to:
- identify, through systematic selection, digital assets to be preserved across new generations of technologies
- include in the scope of the program materials that originated in digital form (born digital) and those that were converted to digital form through a digitization process
- protect CUL's digital investments through a fully-implemented digital preservation program
- demonstrate organizational commitment through the identification of sustainable funding for the program and its digital archive (the Common Depository System)
- identify and support the core working team to develop and operate the digital archive and provide requisite training and development as needed
- develop a cost-effective program through means such as, system-wide integration, shared responsibilities, and automating human-intensive efforts
- comply with prevailing community standards for digital preservation and access
- seek, expand, and develop digital preservation methods that are appropriate for Cornell and promote inter-institutional collaboration

Scope
The digital preservation program is responsible for identifying, securing, and providing the means to preserve and ensure ongoing access to selected digital assets. Not all of the digital content CUL creates or acquires will be preserved. Cornell commits to these classes of objects with associated preservation priorities and levels of commitment:

- *Priority 1: born digital materials* - Rigorous effort will be made to ensure preservation in perpetuity of material selected for preservation, both library resources and institutional records.
- *Priority 2: digitized materials (no available analog)* - Every reasonable step will be taken to preserve materials without a print analog, when redigitizing is not possible or no 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.
- *Priority 3: 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 redigitizing as needed will be weighed against the cost of preserving the existing digital objects
- *Priority 4: items and other materials* – No preservation steps will be taken for ephemeral materials such as, materials scanned for E-reserve and Document Delivery, odds and ends of collections, portions of text, and content that is deemed unessential to the comprehensiveness of collections.

Operating Principles
CUL avows that the digital preservation program will:

- comply with the Open Archival Information System (OAIS) reference model standard in the development of the digital archive
- adhere to prevailing community-based standards in developing and maintaining its organizational and technological context
- participate in the development of digital preservation standards and their promulgation
- commit to ensuring that the digital archive is as interoperable as possible by utilizing open source options whenever feasible and to working with other organizations on core issues to enable interoperability
- define a sustainability plan for the digital archive that ensures its cost-effective, transparent, and auditable management over time
- develop the digital archive to maximize scalability, flexibility, and reliability
- ensure consistent, documented policies, procedures, and practices for the program and the operation of the digital archive within a distributed environment (across CUL and the university and beyond)
- employ appropriate storage management for digital content, utilizing on-line, near-line, and off-line storage as appropriate
- manage hardware, software, and storage media containing archival copies of digital content in accordance with environmental, quality control, security, and other standards and requirements
- establish procedures to meet archival requirements pertaining to the provenance, chain of custody, authenticity, and integrity (bit-level and content) of institutional records and assets
- define policies and procedures for the preservation and availability of digital assets respectful of intellectual property ownership and rights

Roles and Responsibilities
The traditional role of librarians and archivists has included preserving the intellectual record of the research community in analog formats. The same is true for research materials in digital form. The University Library has primary responsibility for digital preservation at Cornell, but digital preservation is a shared responsibility across the University. All responsible officers of the University, content creators, and disseminators with a custodial role for digital content identified for long-term preservation have a responsibility to actively contribute to the intent and priorities necessary to fulfill this policy.

Selection and Acquisition
Selection principles and guidelines for content that enhances the scope of CUL's collections is well-established for print materials and is being extended to digital content using CUL-approved criteria that are regularly reviewed and updated (see Appendix C). The criteria include factors for determining when the "look and feel" of digital content as created should be retained. Materials selected for inclusion in the digital preservation program carry with it CUL's commitment to preserve the materials for as long as needed or desired, with an explicit statement of the relevant retention requirements.

Preservation
Digital preservation activities at CUL include:
- selecting material for long-term access based on scholarly value and technical capabilities
- securing digital assets through physical custody or other arrangements
- actively monitoring technology
- reviewing licenses for preservation implications
- promulgating good practice
- providing support services
- ensuring quality creation of digital materials
- negotiating for deposits
- responding to technology obsolescence through migration or other strategies
- establishing reliable, comprehensive, archival storage
- providing security for content and systems
- collaborating with other institutions to provide digital preservation services

In acquiring materials for the common depository, CUL must continually define and communicate levels of preservation appropriate to each type of format. Rather than limit inclusion in the digital preservation program to those formats with known and feasible preservation approaches, a level of preservation is associated with each format. The preservation level and the priority, as described above under Scope, determine the preservation strategy for materials. As new formats emerge, a level of difficulty required to preserve each type is assigned. The resources required to preserve are balanced against the overall need. This selection process applies to both new and existing content. CUL will review existing digital content for preservation as priorities dictate and resources allow.

Access/Use
Access to preserved digital content is provided using the most up to date technology available at the time of use. When retaining the look and feel is deemed necessary, CUL will seek to enable the original versions of the digital objects to be rendered over time. CUL complies with access restrictions as defined in relevant laws, regulations, licenses, and deposit agreements. Appropriate preservation plans to make rendering the original version possible are devised on a case-by-case basis and revised as needed. Without the preservation of digital materials, access would not be possible and essential cultural heritage materials would be at risk.

Challenges
There are recognized challenges in implementing an effective digital preservation program, such as:
- *Rapid growth and evolution:* Technology that enables the variety of formats and dissemination mechanisms changes rapidly. Establishing a program that is responsive to change is a huge challenge.
- *Sustainability*: The need for good cost models and affordable programs is widely acknowledged, yet still unaddressed. The scale is based on the level of commitment. CUL requires sufficient funding for startup and major improvements, as well as designated library funding to sustain ongoing preservation efforts. The program should reflect reasonable expectations of requisite resources, i.e., CUL should not promise more than can be delivered.

- *Content provider partnerships*: Working with creators and providers of crucial content to employ appropriate maintenance prior to deposit that will facilitate future preservation.
- *Enabling full preservation*: Moving from well-managed digital collections to preserved collections in the true sense of the term requires institutional effort, partnership development, and a financial commitment.
- *Flexibility*: To respond to evolving technological capabilities and changing user expectations, the digital archive must revise continually the definition of the dissemination information package (DIP) for delivery to an expanding array of content delivery platforms. These requirements must be met without jeopardizing the ongoing care of the digital content.

Cooperation/Collaboration
CUL is committed to collaborating within the University and with other institutions to: advance the development of the digital preservation program, share lessons learned with other digital preservation programs, extend the breadth of our available expertise, and extend the digital content that is available within a broad information community to CUL users through cooperative efforts.

Appendix A. Terms and Resources

Definitions

Proposed terms to define:

- Access
- Deposit
- Deposit Agreement
- Deposit Requirements
- Depositors
- Digital Archive
- Digital Repository
- Institutional Repository
- Preservation
- Selection
- Stakeholders
- Stewards

References

Primary
- *Cornell University Library Goals and Objectives, 2002-2007*
  http://www.library.cornell.edu/Admin/goals/objectives.html
- *Report of the Digital Preservation Policy Working Group on Establishing a Central Depository for Preserving Digital Image Collections[1]*
  http://www.library.cornell.edu/preservation/IMLS/image_deposit_guidelines.pdf
- *Trusted Digital Repositories: Attributes and Responsibilities*, RLG-OCLC, May 2002
  http://www.rlg.org/longterm/repositories.pdf
- *Reference Model for an Open Archival Information System (OAIS)*, CCSDS, July 2002
  http://windsor.gsfc.nasa.gov/standards/ccsds/pdf/CCSDS-650.0-B-1.pdf

Secondary
- *Preservation Metadata and the OAIS Information Model*, OCLC-RLG, June 2002.
  http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
- Metadata Encoded Transmission Standard (METS)
  http://www.loc.gov/standards/mets/
- Open Archives Initiative (OAI)
  http://www.openarchives.org/

---

[1] The Central Depository for Digital Images was renamed the Common Depository to include digital formats that might be designated for preservation and to address the possibility that the depository might be distributed, but would be implemented using common policies and practices.

Appendix B. Roles and Responsibilities

This appendix builds on the Roles and Responsibilities section of the Digital Preservation Policy.

Using the OAIS Standard, CUL has identified the following stakeholder categories for the digital preservation program.

- *Producers*: Contributors 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.
- *Management*: CUL managers with responsibility for key organizational, administrative, financial, information technology, and programmatic areas including specialist positions such as: the University Archivist, the Digital Preservation Officer, the Intellectual Property Officer, and others. Management will be responsible for setting digital preservation policies and integrating them into broader organizational contexts.
- *Administrators*: Content stewards (designated staff responsible for selection and for ongoing curation of specific collections), digital preservation specialists and working teams (University Archivist, Digital Preservation Officer, Intellectual Property Officer, Metadata Services staff, Library Systems staff, supporting and technical staff at CUL, elsewhere at Cornell, and at collaborating institutions). Administrators will be responsible for the establishment of the full digital preservation program and for day-to-day management of the digital archive.
- *Consumers*: All users of digital content identified for digital preservation and stored in the digital archive, including faculty, students, alumni, researchers, staff, and, increasingly, members of consortia.
- *Third party:* an additional stakeholder will be providers of value-added services like aggregators
- *Institutional collaborators:* Corporate partners, non-profit organizations, academic researchers, cultural heritage organizations, government agencies at all levels, and all other relevant partners.

Appendix C. Selection Criteria

Two recent initiatives form the basis for the development of system-wide digital preservation selection criteria:

Central Depository grant selection criteria
http://www.library.cornell.edu/preservation/IMLS/

Cornell's IMLS-funded project, Preserving Cornell's Digital Image Collections: Implementing an Archival Strategy defined selection criteria in its final report: *Report of the Digital Preservation Policy Working Group on Establishing a Central Depository for Preserving Digital Image Collections – Part 1: Responsibilities of Transferee* defined this set of selection criteria:[2]

> *Scope*
> Digitized materials, like their analog counterparts, should primarily support teaching and scholarship at Cornell University and secondarily the needs of other communities. Selection of digital materials is based, in significant part, on their content. It is also essential that the functionality of digitized items be considered as part of the selection decision, i.e., the extent to which the items' particular digital forms affect or extend their utility for specific scholarly or educational purposes. Additional evaluation elements, enumerated in subsequent sections of these guidelines, imply that a resource meets the standards of authority and relevance that characterizes other segments of the library's holdings.
>
> *Content/Functional Criteria*
> Digital image materials selected for central deposit should reflect the same content criteria used for evaluating print collections. They should:
> - Support the curriculum, including emerging distributed learning initiatives
> - Ensure standard source availability – the identification of "core" publications
> - Facilitate faculty and student research
> - Maintain national collection strengths
> - Honor inter-institutional and other commitments
>
> These considerations are codified in the collection development policy of the Library:
> > *Cornell University Library Collection Development Policy Statements*
> > http://www.library.cornell.edu/colldev/cdhome1.html
> >
> > *Cornell Primary Collecting Responsibilities*
> > http://www.library.cornell.edu/colldev/cpr.pdf
>
> *Priorities for Deposit*
> At some point, CUL may have to make difficult choices about which materials to accept for deposit based on economic or managerial considerations. Technical and legal issues will also most likely be the determining factors in selection but content decisions may also come into play. While not restricting selection to materials that meet one or more of the following criteria, preference for central deposit could be given to digital image materials that:
> - Represent complete and credible versions of digitized resources
> - Represent thematic or format-based aggregates, rather than idiosyncratic works
> - Help create a comprehensive collection
> - Enhance access to collections by making them easier to browse, search, and use

---

[2] See Project website at: http://www.library.cornell.edu/preservation/IMLS/image_deposit_guidelines.pdf

- Increase use of collections, by bringing little-known materials to light or be widening potential readership
- Are accessible campus-wide
- Help preserve, protect, store remotely, or replace materials by providing reliable surrogates for consultation
- Enable new kinds of research, not possible in the analog form
- Do not replicate resources available to the Conrell community via other arrangements, e.g., digital resources managed at individual libraries within the system or other divisions within the University, or made accessible via other institutions, consortia, inter-institutional agreements, or existing licenses

CUL's Faculty Grants Programs
http://dcaps.library.cornell.edu/facultygrants/

CUL's Faculty Grants Program put forth the following selection criteria:
- Increase the availability, and consequently the use of a collection of demonstrated scholarly significance
- Demonstrate strong interest within the academic community for access to the collection
- Contribute significantly to the existing digital collections such as the ones included in the Cornell University Library's digital library.
- Support the University's subject strengths
- Involve interdisciplinary, collaborative, and inter-institutional resources