

Columbia University Libraries Policy for Preservation of Digital Resources

July 2000

1. Statement of CUL policy for preservation of digital resources

According to the Columbia University Libraries' Mission Statement (7/22/93):

The Columbia University Libraries provides Columbia faculty, students, and staff with access to information in all subject areas related to the University's academic mission and its goals. The Libraries embraces its time-honored obligations of collecting, preserving, and providing access to collections, not only for the Columbia community but also for scholars and students from throughout the world who require access to the Libraries' unique materials.

Policy:

Digital resources are part of the CUL collections and subject to the same criteria for selection and retention decisions as other media. As such, they are included under the central CUL preservation policy: ensuring that the collections remain available over the long term, through prevention of damage and deterioration; reversing damage where possible; and, when necessary, changing the format of materials to preserve their intellectual content.

As with other parts of the collections, decisions about preservation are made by selectors, curators, and bibliographers as experts on the value of the content, in consultation with the relevant technical experts, including Academic Information Systems, the Library Systems Office, and Preservation. Priorities for preservation action are based on this Policy, the CUL Strategic Plan, and available resources. When possible, decisions about the need for long-term retention are made at the time of creation, acquisition, or licensing of digital resources.

For digital resources that are deemed to be of long-term value, preservation can be defined as the actions needed to assure enduring access to the full content of those resources over time. Content has wider implications than simply assuring that a given image can be accessed. Thus, hierarchical and structural relationships among the files (e.g. the pages of a book) and metadata that make the files usable must be preserved as well as the files themselves,

Digital resources may exist in multiple versions. CUL is committed to preserving the archival version: the fullest, highest-quality available version of the resource, whenever possible; and the descriptive, structural, and administrative metadata associated with it.

Scope of preservation responsibility:

1. Responsibility for internal long-term retention and management by CUL/AcIS of:
 - a. Digital resources created by CUL for which no other versions exist, and deemed to be of long-term value.
 - b. Digital versions of resources reformatted by CUL, and deemed to be of long-term value in digital form.
 - c. Unique digital resources which are acquired by CUL (through donation or purchase) as parts of archival/manuscript collections and which are unlikely to be preserved anywhere else.
 - d. Digital records (e.g. bibliographic records, personnel records) deemed of long-term value and/or essential to CUL's functioning, and not preserved through any other arm of the university.
2. Responsibility for working externally through consortial action, licensing agreements, etc. to assure that someone (possibly but not necessarily CUL/AcIS) carries out preservation of appropriate commercially available digital resources to assure that CU faculty, staff, and students will have adequate ongoing access to these resources. Particular emphasis should be given to resources which exist in digital form only.
3. Responsibility for informing, consulting, and as appropriate coordinating with other units of Columbia University in the preservation of administrative and other digital resources to assure that CU faculty, staff, and students will have adequate ongoing access to these resources. Included here are digital resources created at CU outside of the Libraries but considered to be part of the digital library (EPIC ventures, others) and of long-term value.

Frequency with which preservation/retention policy for digital materials will be updated:

This policy will be reviewed by at the beginning of each academic year to assure timely updates as technology and experience mature, or more often if need arises.

2. Statement of CUL's commitment to lifecycle management

CUL is committed to lifecycle management of its digital resources. Guidelines and procedures for each stage have been or are being developed, and are reviewed as technology changes or other need arises. CUL will participate actively where appropriate in research, development, and implementation of new practices for preservation of digital resources.

- Development of preservation strategies, including consideration of:
- Degree of integration with storage, backup, and preservation for non-digital library resources.
 - Development and use of decision-making tools (e.g., risk analysis, usage monitoring, probability of loss calculations, cost models, etc.).

- Maintenance strategies (backups - online and/or offline, monitoring, refreshing, redundancy through mirror sites or caching, etc.).
 - Survival strategies (migration, emulation, archeology, etc.).
 - Reliance on outside consulting and archiving services, if any, contract negotiation, etc.
- Selection for long-term retention at time of digitization, acquisition, or licensing; and later reselection for retention if this decision was not made initially:
- Primary criteria: based on institutional mission, needs, priorities, and reasons for creating or acquiring the resources (related to long-term institutional mission and linked to conversion guidelines): see Selection Criteria for Digital Imaging Projects, <http://www.columbia.edu/cu/libraries/digital/criteria.htm>.
 - Secondary criteria: based on regional, national, consortial, and international responsibilities.
- Conversion guidelines:
- CUL follows relevant standards where they have been established and best practice for digital conversion, as documented by Library of Congress, the Research Libraries Group (RLG), the Digital Library Federation (DLF), and other relevant bodies.
 - See: Technical Recommendations for Digital Imaging Projects, <http://www.columbia.edu/acis/dl/imagespec.html>.
- Metadata creation and management (unique IDs and other descriptive, structural and administrative metadata, including ownership/rights management):
- CUL follows relevant standards where they have been established and best practice for metadata creation and management, as documented by Library of Congress, RLG, DLF, and other relevant bodies.
 - Digital resources deemed to be of long-term value are tracked through the Master Metadata File (MMF). See: CU Master Metadata File, <http://www.columbia.edu/cu/libraries/inside/projects/metadata/>.
- Storage (online, offline, redundancy recommendations, etc.):
- CUL follows relevant standards where they have been established and best practice for storage, as documented by Library of Congress, RLG, DLF, and other relevant bodies.
 - The archival version of the digital resource is preserved in a lossless, non-proprietary format, whenever possible.
 - Storage plan for primary responsibility materials:
 1. Resources currently in use: kept online with regular backup, refreshment, and migration.
 2. Whether online or not: all archival versions (highest resolution, fullest capture, lossless compression) are written to approved storage media and stored off-line in the Library Systems Office (LSO), with a schedule for regular refreshment, and migration.
 3. For archival versions which are not currently online: a duplicate off-line copy is created for storage at a different site.

4. All versions, online and offline, are tracked through the MMF.
- Access arrangements (database management, Web interface, access reliability, etc.):
 - CUL follows relevant standards where they have been established and best practice for access arrangements, as documented by Library of Congress, RLG, DLF, and other relevant bodies.

3. Statement of CUL's resource management policies and plans

As stated in the Six-Year Plan for the Libraries and AcIS (May 1999), the digital library is the primary on-line information interface for the University, delivering scholarly, instructional, research, administrative and personal information. The University is committed to supporting all of the operations of the digital library, including preservation. Resource management for preservation of digital resources includes:

- Technical infrastructure (equipment purchases, maintenance and upgrades, software/hardware obsolescence monitoring, network connectivity, etc.).
- Financial plan (strategy and methods for financing the digital preservation program, commitment to long-term funding).
- Staffing infrastructure (including hiring and ongoing training).
- Rights management.

4. Statement related to regional, national, consortial, and international responsibilities

Whenever it is proposed that digital resources created or acquired by CUL as part of a cooperative or consortial effort, and deemed to be of long-term value, should be preserved by another institution or organization, CUL will evaluate their capacity to provide long-term access to digital files, metadata, and functionality in order to assure that the resources will be preserved at least as well as they would be by CUL. Issues which must be taken into account include:

- Shared obligations and cooperative preservation.
- Information sharing.
- Technology support (e.g., reuse of same migration paths and strategies by others; following emerging common/recommended practice).
- Development of minimal standards for capture, management, and maintenance (to enable efficient strategies).