

Catching Up: Creating a Digital Preservation Policy After the Fact

Jennie Levine Knies, Manager, Digital Programs and Initiatives, University of Maryland Libraries

Robin C. Pike, Manager, Digital Conversion and Media Reformatting, University of Maryland Libraries

Abstract

The University of Maryland Libraries have been managing digital content for about a decade without an official digital preservation policy. The Libraries' 2013 Strategic Plan called for the development of preservation plans, including "bit-preservation, digital object preservation, and metadata management pertaining to preservation of digital objects and their aggregations." A task force examined policies at other institutions and evaluated needs and strategic directions. The resulting high-level policy and appendix based on the Trusted Repository Audit Checklist (TRAC) will direct all future planning, policies, and implementation. This case study will describe the process and planned steps towards implementation.

History and Mandate

The University of Maryland (UMD) Libraries is the largest library system in the Washington, D.C.-Baltimore area, and serves 37,000 students and over 4,000 faculty members of the College Park campus. The UMD Libraries rank 39th among the 115 members of the Association of Research Libraries, and 10th in electronic resources as a percentage of total library materials.¹

Digital Collections and the Digital Repository at the University of Maryland (DRUM) are the two primary repositories for UMD-owned and created digital content. Digital Collections uses the Fedora architecture and provides access to digital images, audio, and video. Among its premier collections is the Gordon W. Prange Digital Children's Book Collection, a growing repository of unique Japanese children's books published during the post-WWII occupation (1945–1949). At the time of writing, approximately 5,000 of the books from this massive preservation-digitization project are available online. By September 2014, that number will be 8,000. The entire project aims to digitize 78,000 volumes. Other research collections include *University AlbuM*, a continually growing resource that currently contains over 3,000 historic university images, and close to 800 films of historic UMD football games. Other collections of note include 5,000 postcards featuring historic buildings and sites across the country, and numerous manuscripts documenting the Civil War and slavery in Maryland. Digitization is currently underway on the correspondence of the American author Katherine Anne Porter. By the end of 2013, we will add several hundred of her letters to Digital Collections. DRUM is built on the DSpace platform and serves as the repository for electronic theses and dissertations, technical reports, and faculty research. It currently contains over 14,000 publications, including close to 9,000 theses and dissertations produced at UMD since 2003. While Digital Collections and DRUM began in the mid-2000s, they were not part of the same department until 2011. Much of the digital preservation work performed is ad hoc, manual, and focused more on backup and restoration of files than on preservation of holistic systems.

The UMD Libraries manages other digital content including books in the Internet Archive, web-archives in the Internet Archives's Archive-It system, and will soon be submitting content to the HathiTrust. We recently launched an instance of the Open Journal Systems software and are starting to examine management of electronic records and born-digital content within manuscript collections, university records, and the preservation and accessibility of research data sets. Although a platform or suite of tools for these activities has not been chosen, we would like to plan for how to best manage these acquisitions moving forward.

A number of administrative and organizational changes at the UMD Libraries in recent years provided the support for the development of a digital preservation policy. The UMD Libraries hired a new dean in 2009, who in turn, hired an Associate Dean for Information Technology (now Digital Systems and Stewardship) in 2010. Management for Digital Collections, DSpace, and all other digital initiatives moved to the Digital Systems and Stewardship (DSS) Division in 2011 as a new department, Digital Programs and Initiatives (DPI). Reformatting services formerly managed by Digital Collections also moved to a new department, Digital Conversion and Media Reformatting (DCMR), which focused specifically on digitization. The establishment of DCMR provided the UMD Libraries with more staff, time, and expertise to devote to the act of digitization, which meant the creation of more digital assets, including audio and video. It also allowed the staff in DPI to focus more on devising policy surrounding digital preservation and management of digital assets.

During this time, the UMD Libraries also developed a strategic plan, which included the mandate to: "Develop preservation plans for collections, including bit-preservation, digital object preservation, and metadata management pertaining to preservation of digital objects and their aggregations." The 2013 revision to the strategic plan tasked DSS to "develop digital preservation policies for the Libraries via a working group to inform and review the project" with the objective to "provide authenticity, discovery, and access to digital assets for current and future generations." The task force consisted of the DPI manager, the DCMR manager, the special collections digital projects librarian, the curator for Special Collections in Performing Arts, and the UMD Libraries' Head of Preservation. Their first meeting was April 8, 2013.

About the Policy

This task force began by examining policies from other institutions and evaluating the needs and strategic direction of the UMD Libraries in terms of collection development, curation, and services to the campus community.² In 2006, a similar group at the UMD Libraries had developed a "Digital Curation Policy" that addressed many issues still relevant seven years later. For various reasons, this policy document never made it past draft form, and was not officially implemented; however, the task force did review its contents to determine which portions could be incorporated into the new policy.

Early in the process, the task force made a decision to make this digital preservation policy both brief and useful. To achieve this, we focused on purely high-level statements, and spent a large portion of the four months authoring the policy determining whether or not a statement was policy or procedure. If the latter, we dropped the statement from the policy. Ultimately, the policy ended up consisting of a little more than four pages, with an extended, ten-page appendix. The primary sections were: Mission, Mandate, Digital Preservation at the UMD Libraries, Scope of Digital Preservation at the UMD Libraries, Challenges of Digital Preservation at the UMD Libraries, Roles and Responsibilities, Training and Education, Review and Evaluation, and Implementation Strategy.

The task force had to decide upon language to use that would be clear and consistent throughout the policy. To that end, the most-heavily discussed section was "Roles and Responsibilities." We wanted to make it clear to readers that "digital preservation" is not an activity confined to one person or department, or a specific genre of digital content; rather, it would require cooperation and understanding from stakeholders throughout the organization. We identified four primary user groups, and drew language from the Reference Model for an Open Archival Information System (OAIS), which contained active submission and ingest roles: producers, collection managers, administrative managers, and consumers. We briefly defined each group, and identified specific responsibilities related to digital preservation in job descriptions, collection policies, and through other official planning documents.

The "Mission" section consists of one paragraph explaining the UMD Libraries' commitment to preserving content selected for retention by collection managers. It clarifies that the policy refers to digital preservation of digital content both born-analog and born-digital. It also defines the audience for the policy: UMD Libraries employees, digital content contributors, donors, and users. The "Mandate" section outlines the impetus for the existence of the policy and references the 2013 UMD Libraries' Strategic Plan.

Where is the actual "policy"? The "Digital Preservation at the UMD Libraries" section contains most of what people might consider the detailed language about digital preservation and explains what we will actually do. On a very broad level, we promise to provide authenticity, discovery, and access to digital assets for current and future generations. In order to do this, several high-level events are listed. For example, the first item in our list states that we will perform "regular checks on the integrity of stored content." We do not address *how* we will do this, nor do we elaborate on what or where the stored content might be. That information will be defined by future policies and working documents, as outlined in the appendix to the policy. Likewise, the policy explains that we will enable "uninterrupted (not necessarily instant) access to digital content over time as technology for digital content evolves." This wording implies we will monitor and take into account how systems change over time and strive to make content available either through migration or emulation, but we do not say specifically how or when we might perform such actions. Related to that, we also mention "periodic review of preferred digital formats and digital metadata standards," and we promise to comply with standards and best practices.

In this section, we also addressed the cost of digital preservation; tied into this discussion is the concept of appraisal. Our digital preservation policy explicitly does not claim to preserve all digital content, forever. Instead, it calls upon collection managers to appraise digital assets to "determine duration of digital preservation and access." We promise to provide a range of digital preservation services (from bit-level to full migration), based on priorities, content, format, and other factors, including adequate financial and organizational resources. The intent of the language is not to provide future generations with an excuse for inaction, but rather, to impress upon everyone that this is a vital activity for our organization, and one that must be supported at a programmatic level.

To define the scope of the policy, we turned to the UMD Libraries' *Collection Development Policy*, which states that "The University of Maryland Libraries is committed to ensuring the preservation and long-lasting availability of its research collections and resources in all formats."³ Because UMD Libraries acquires a wide variety of digital content and we cannot predict what formats and content we will be asked to preserve in the future, we broadly defined the scope, promising to preserve digital assets "created using any type of application or on any computing platform" and "delivered on any digital media." We also discussed our concern about duplication of effort and requests to preserve content that might be preserved elsewhere. Therefore, we added two additional points to our scope—content must be "unique to the University of Maryland Libraries' collections" and "in danger of obsolescence or loss." For example, the Libraries should provide access to licensed content on purchased CDs available for the life of the media and the span of interoperability with common operating systems, but should not commit to preserving or making the content accessible past the lifespan of the media. Rather than impose limits on what we will preserve, the intention of this wording was to help collection managers and others in our organization think deeply about prioritization and the impact of their digital preservation decisions. We avoided creating a matrix or decision tree in this document, although the earlier 2006 document did contain a helpful tool for determining preservation activity. We know now, from experience, that each decision may contain so many variables and factors that we did not want to try to anticipate them in a prescribed way.

The section outlining challenges of digital preservation almost did not end up in the policy. Although many of the other sample policies we reviewed contained this type of section, the task force was afraid that it would sound non-committal. While working on this section, the Associate Dean for Digital Systems and Stewardship suggested we think of real-life scenarios that pose challenges, and how our policy might try to address them. Since donors are also potential readers of this policy, we sought to make this section as educational as possible. The challenges do not represent actions that we *cannot* do but rather identify the intricacies and

complications related to working with digital assets, such as unstable carriers, hardware and software obsolescence, access barriers such as password protection or encryption, poor description, and loss of contextual information. Despite these challenges, however, we note that we will "continually work to mitigate these risks through policy and technological development."

Because we emphasize the need for organization-wide collaboration in implementing digital preservation policies, we included a section that addresses the importance of training and education, including the necessity of updating job descriptions. We identified five areas for focusing education: general awareness; information lifecycle management; information storage management and systems; maintenance, best practices and standards; and legal issues and university policies. Our hope is that, when necessary, the UMD Libraries will support training and educational opportunities to better practice digital preservation. The policy concludes with instructions for annual review and evaluation of the digital preservation policy, and an implementation strategy.

Plans for Implementation and Conclusion

The UMD Libraries' *Digital Preservation Policy* defines the overarching philosophy regarding digital preservation, including local and remote systems used to manage digital assets. Implemented after evaluation and approval from the Libraries' Dean, Associate Deans, and Directors, the policy requires a detailed digital preservation plan that incorporates a number of supplemental policy and procedure documents. These separate documents will alleviate the need to extensively revise the high-level policy, focusing editorial work on the supporting documents, as well as providing functional continuity.

The final effort of the task force involved reviewing the *Trustworthy Repositories Audit and Certification (TRAC): Criteria and Checklist* as a framework for the appendix outline, which list the types of policies and procedures that will together comprise a comprehensive, sustainable digital preservation program at the UMD Libraries.⁴ Included in this outline are guidelines for establishing governance and organizational support, digital object management, and technological infrastructure and security. We also assigned broad responsibility among the Libraries' divisions to each item in the appendix. Many of the relevant policies and procedures already exist, and the initial work of implementation will involve identifying, collating, and revising existing policies and procedures, and assigning staff to draft new policies and procedures where gaps exist. We plan to use TRAC and the NDSA *Levels of Digital Preservation* as targets. Revising and drafting new policies and procedures will begin in 2014. The completed policy is available at: <http://hdl.handle.net/1903/14745>.

Notes

1. "About Us," University of Maryland Libraries, accessed November 12, 2013, <http://www.lib.umd.edu/about>.
 2. Policies reviewed and frequently referred to included: "Digital Preservation Policy," HathiTrust, accessed November 22, 2013, <http://www.hathitrust.org/preservation>.
- "Digital Preservation Policy, 4th Edition," National Archives of Australia, accessed November 22, 2013, <http://www.nla.gov.au/policy-and-planning/digital-preservation-policy>.
- "Digital Preservation Policy," University of Utah Library, accessed November 22, 2013, <http://www.lib.utah.edu/collections/digital/digital-preservation.php>.
- "Digital Preservation Strategy," British Library, accessed November 22, 2013, http://www.bl.uk/aboutus/stratpolprog/collectioncare/discovermore/digitalpreservation/strategy/BL_DigitalPreservationStrategy_2013-16-external.pdf.
3. "Levels of Preservation," National Digital Stewardship Alliance (NDSA), accessed November 22, 2013, <http://www.digitalpreservation.gov/ndsa/activities/levels.html>.
 3. "Collection Development Policy for the General Collections," University of Maryland Libraries, accessed: November 12, 2013, <http://www.lib.umd.edu/collections/collection-development-policy>.
 4. "Trustworthy Repositories Audit and Certification (TRAC): Criteria and Checklist," OCLC and CRL, accessed: November 12, 2013, http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf.