

# Digital Preservation: Roadmap To NDSA Level Two

## A Project of the Five College Digital Stewardship Committee

Abigail Baines, Hampshire College; Meghan Bergin, UMass, Amherst; Sarah Goldstein, Mount Holyoke College (chairperson); Eitan Kensky, Yiddish Book Center; Elisa Lanzi, Smith College; Aaron Rubinstein, UMass, Amherst; Laura Shea, Mount Holyoke College; Sarah Walden, Amherst College; Bryn Geffert, Amherst College (FCLC Liaison)

May 8, 2017

## Background

The Digital Stewardship Committee (DSC), reporting to the Five College Librarian's Council, committed several initial meetings to a series of digital stewardship "pain points" exercises. Through these exercises we sought to understand where we overlapped in digital stewardship-related challenges, and of those which ones might most benefit from collaboration or other shared activities, projects, or practices. We identified the following areas as important and worthy of consideration in our work as a committee:

- Preservation
- Metadata
- Born digital accessioning
- Web archiving
- Digital Asset Management, especially of large digital video files
- Global discovery and access
- Technology and programming support
- Building a culture of stewardship awareness
- Knowledge sharing and staff development

We concluded that *digital preservation* is an area of both urgency and critical importance to all of our institutions. While we have all independently made progress in preservation activities, gaps remain and further implementation has been slow. Thus, the DSC has determined that ensuring consortium-wide progress on digital preservation should be our main focus going forward. This goal is supported by our charge, and reinforces our libraries' critical role in providing leadership around digital stewardship.

## Goal

Our goal is to develop a set of implementation recommendations for aligned digital preservation practices we are referring to as: *Roadmap To NDSA Level Two*. In 2013, the [National Digital Stewardship Alliance \(NDSA\)](#) published a tiered set of recommended practices, "Levels of Digital Preservation," (see addendum below) to help organizations work towards progressively improved digital preservation activities. While not designed to assess more holistic preservation actions such as explicit policies or staffing, "Levels of Digital Preservation" is nonetheless extremely helpful in breaking out reasonable goals for establishing and communicating basic

principles of **best practices in data storage** including fixity, integrity, metadata, and formats (Level Two includes having three complete copies of data and ensuring geographic separation of at least one of those copies, for instance). The DSC feels that the practices outlined in NDSA Level Two are achievable for the institutions in our consortium and extended consortium partners, and, just as important, are also easy to understand. By focusing on getting to the NDSA Level Two as a goal, we can **establish a common baseline of preservation infrastructure** and **gain experience working collaboratively**; this will in turn enable us to **move forward effectively on progressively more comprehensive digital preservation activities**.

To achieve this goal, the DSC will:

1. Gather information on the current state of storage and preservation among stakeholders of participant institutions:
  - a. Review past assessments and data collected by 5C groups
  - b. Create appropriate and expedient methods by which to gather necessary new information
  - c. Communicate and liaise with potential participants/respondents
  - d. Reach out to other cultural organizations or consortia with experience or initiatives in this area
  - e. Organize and evaluate resulting information
2. Investigate digital preservation services vendor solutions and consult with institutions that are currently using vendors for similar services (for example those found in the [Ohio Five Digital Preservation Task Force Report](#)).
3. Coordinate with other Five College committees, task forces, and working groups or staff as appropriate.
4. Create recommendations for aligned digital preservation practice to accomplish NDSA Level Two status for participating institutions.

## **Timeline**

We anticipate the timeline for this work to be approximately late January 2017 to the end of the October 2017.

## **Participants**

The DSC will, in keeping with the diverse makeup of the committee itself, as well as the original charge of the DSC, include Five College Libraries, Archives and Special Collections, and Museums including Museums 10 partners. The DSC also plans to include appropriate members of the Five College Information Technology community, either as a member(s) of the Committee or as partners in the project.

## Benefits

- Reduce risk: Getting to NDSA Level 2 will mean that it is less likely that critical digital collections could be lost or irreparably damaged. For example, loss of “born digital” student thesis scholarship would resonate beyond the libraries.
- Build and grow: Establishing standardized, minimally-viable preservation and digital storage practices lets us then progressively increase scope and scale of collections and assets being preserved.
- Sustainability: Pooling resources and establishing shared processes and activities ensures that individual institutions are more likely to continue with preservation practices.
- Leverage resources: It will allow us to leverage the role of Five Colleges as a leader in supporting collaborative solutions for a shared set of “wicked problems”
- Level the playing field: Regardless of their staff size or resources, participation in this project will help institutions consider options for digital preservation and help stabilize and secure their most valuable digital collections.
- Build community: Working towards this goal will facilitate and encourage much-needed collaboration and communication among regional LAMS (Libraries, Archives, and Museums); as well as associated IT partners responsible for information and asset security on our campuses.

## Challenges

- The potential cost of vendor solutions may be a barrier to some institutions, even at consortial pricing.
- Institutions may lack capacity (staff or technical expertise) to participate or implement recommended activities/changes.
- Different systems, technology infrastructure, and policies at institutions may be a barrier to aligning new shared technology or practices.

## Appendices

### Appendix 1. NDSA Levels of Digital Preservation

See also: <http://www.digitalpreservation.gov:8081/ndsa/activities/levels.html>

**Table 1: Version 1 of the Levels of Digital Preservation**

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	<ul style="list-style-type: none"> <li>- Two complete copies that are not collocated</li> <li>- For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system</li> </ul>	<ul style="list-style-type: none"> <li>- At least three complete copies</li> <li>- At least one copy in a different geographic location</li> <li>- Document your storage system(s) and storage media and what you need to use them</li> </ul>	<ul style="list-style-type: none"> <li>- At least one copy in a geographic location with a different disaster threat</li> <li>- Obsolescence monitoring process for your storage system(s) and media</li> </ul>	<ul style="list-style-type: none"> <li>- At least three copies in geographic locations with different disaster threats</li> <li>- Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems</li> </ul>
File Fixity and Data Integrity	<ul style="list-style-type: none"> <li>- Check file fixity on ingest if it has been provided with the content</li> <li>- Create fixity info if it wasn't provided with the content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity on all ingests</li> <li>- Use write-blockers when working with original media</li> <li>- Virus-check high risk content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity of content at fixed intervals</li> <li>- Maintain logs of fixity info; supply audit on demand</li> <li>- Ability to detect corrupt data</li> <li>- Virus-check all content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity of all content in response to specific events or activities</li> <li>- Ability to replace/repair corrupted data</li> <li>- Ensure no one person has write access to all copies</li> </ul>
Information Security	<ul style="list-style-type: none"> <li>- Identify who has read, write, move and delete authorization to individual files</li> <li>- Restrict who has those authorizations to individual files</li> </ul>	<ul style="list-style-type: none"> <li>- Document access restrictions for content</li> </ul>	<ul style="list-style-type: none"> <li>- Maintain logs of who performed what actions on files, including deletions and preservation actions</li> </ul>	<ul style="list-style-type: none"> <li>- Perform audit of logs</li> </ul>
Metadata	<ul style="list-style-type: none"> <li>- Inventory of content and its storage location</li> <li>- Ensure backup and non-collocation of inventory</li> </ul>	<ul style="list-style-type: none"> <li>- Store administrative metadata</li> <li>- Store transformative metadata and log events</li> </ul>	<ul style="list-style-type: none"> <li>- Store standard technical and descriptive metadata</li> </ul>	<ul style="list-style-type: none"> <li>- Store standard preservation metadata</li> </ul>
File Formats	<ul style="list-style-type: none"> <li>- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs</li> </ul>	<ul style="list-style-type: none"> <li>- Inventory of file formats in use</li> </ul>	<ul style="list-style-type: none"> <li>- Monitor file format obsolescence issues</li> </ul>	<ul style="list-style-type: none"> <li>- Perform format migrations, emulation and similar activities as needed</li> </ul>

## Appendix 2. Digital Stewardship Committee charge

The Five College Digital Stewardship Committee (DSC) advocates for cooperative, formal and informal digital stewardship projects, and, where feasible, advances solutions that lead to enhanced, integrated discovery and preservation of our local digital resources. The work of the DSC covers digital curation, digital preservation, and digital asset management, i.e., activities related to the digital object lifecycle and managing collections over time. The DSC ensures transparency, communication, and continuity between Five College digital projects and encourages and supports digital stewardship-related collaborations of any size between Five College institutions.

Established to be forward-thinking and action-oriented, the Digital Stewardship Committee is charged with scanning the horizon to ensure that Five College stewardship of digital assets is both informed by, and also contributes to, regional and national practices. Anticipating and studying trends supports the Five College Consortium's ability to, "Seek to ensure that major investments made by each campus are undertaken from the earliest point of exploration with advice and consultation with colleagues from the other campuses, enhancing the possibility of other campuses joining in any major endeavor either immediately or at a later date." (Optimizing the Consortial Advantage by 2020: A Strategic Plan for Five Colleges, Incorporated 2010-2020, 3.c)

The Digital Stewardship Committee has the following responsibilities:

- Play a key leadership role by providing recommendations and developing policy in areas pertaining to the committee's charge and expertise.
- Collaborate within the five colleges, with FCLC committees and others including the Museums, Information Technology units, grants, and other initiatives that further the consortium's goals;
- In consultation with FCLC create task forces and working groups as needed;
- Serve as a resource to other Five College committees and task forces;
- Sponsor conversations with other committees and across peer groups in the Five Colleges;
- Identify professional development areas that foster skill development in areas supporting the FCLC strategic priorities;
- Partner with the Professional Development Committee to develop opportunities for Five College library staff;
- Identify, develop and seek funding for collaborative initiatives that directly support the FCLC strategic priorities;

The Five College Libraries Digital Stewardship Committee. We would like to invite you and your institutions to participate in a project we're undertaking that would help us create a set of recommendations to guide basic preservation-consistent file storage practices amongst the Five College libraries, archives, and museums, including the Museums10 partners. By establishing a common baseline of preservation infrastructure together we can gain experience working collaboratively; this will in turn enable us to move forward effectively on progressively more comprehensive digital preservation activities.

For many of our institutions, the digitization of collections (original works of art, physical archives materials, born-digital files, etc.) has become a fundamental part of our everyday practice. Our organizations invest a great deal of resources toward the creation of these individual digital repositories. In many cases, we face similar challenges in maintaining the long-term access and preservation of these valuable assets, leaving them vulnerable to loss and data corruption and puts our investments in them at risk. The Digital Stewardship Committee is working to try to address this challenge by finding practical, collaborative solutions (including possible vendor-provided services) that will greatly benefit both Five Colleges and Museums10.