Follow up Information

Caring for Digital Materials:
Digital preservation networks and collaboration

Liz Bishoff and Tom Clareson, instructors

Sample Documents

1. *It would be great if there was a Request for Proposal model that we could look at for digital preservation projects.*

   We don't know of any current examples of RFPs for digital preservation projects. However, the Digital Preservation Coalition, a U.K. organization dedicated to digital preservation, publishes the *Digital Preservation Handbook*, which includes information on working with third party vendors in the sections "Institutional Guide" and "Organisational Activities" (http://www.dpconline.org/advice/preservationhandbook)*:* While not a specific model RFP, the document details many of the elements to consider when developing an outsourcing RFP.

   The Society of American Archivists offers a workshop on developing RFPs for Recordkeeping Systems (http://saa.archivists.org/events/developing-specifications-and-rfps-for-recordkeeping-systems-1358/377/), one of their Digital Archives Specialist certificate courses, and that might be helpful to you.

2. *What resources are there for the long-term management of digital assets (e.g. migration, emulation, etc.)?*

   Emulation and migration are both used in the long-term management of digital assets. The Digital Preservation Tutorial (now hosted at the MIT Libraries) gives a general overview of these and other strategies:
   http://www.dpworkshop.org/dpm-eng/terminology/strategies.html

   For a more detailed overview of emulation, you might look at a recent post on *The Signal*, the Library of Congress Digital Preservation blog, which interviewed Dirk von Suchodoletz from from the Department of Computer Science at the University of Freiburg, Germany (http://blogs.loc.gov/digitalpreservation/2012/12/the-foundations-of-emulation-as-a-service-an-interview-with-dirk-von-suchodoletz-part-one/ and http://blogs.loc.gov/digitalpreservation/2012/12/the-foundations-of-emulation-as-a-service-an-interview-with-dirk-von-suchodoletz-part-two/)

   Some interesting work on migration has been at The National Center for Supercomputing Applications (NCSA) at University of Illinois at Urbana-Champaign, one of the partners in the National Archives' Applied Research program. One of the tools they have developed,

Polyglot (http://isda.ncsa.illinois.edu/drupal/software/Polyglot), supports migration by identifying the software needed to convert from a given format into the desired preservation format, which may require multiple conversions. Another project, the Conversion Software Registry (CSR) (http://isda.ncsa.illinois.edu/drupal/software/CSR)provides details about the available migration options.

3. *Examples of digital policies would be very helpful for small institutions like archives/museums.*

   A list of sample plans and policies has been posted to the course website, http://www.connectingtocollections.org/courses/caring-for-digital-materials/   Most of these are created by and for larger institutions, such as university libraries, but they are still an excellent resource for all institutions.

4. *Can you recommend good sources for digital preservation plan templates that we might build off of?*

   A list of sample plans and policies has been posted to the course website, http://www.connectingtocollections.org/courses/caring-for-digital-materials/ Most of these are created by and for larger institutions, such as university libraries, but they are still an excellent resource for all institutions.

Specific Platforms and Tools

This list is not intended to be exclusive and represents questions posed during the webinar on April 15, 2013.

1. **Ex Libris Rosetta.** ExLibris is a vendor of integrate and digital library service software. Rosetta is their preservation software. See http://www.exlibrisgroup.com/category/RosettaOverview

2. **ResourceSpace.** Resource Space is an asset management system designed primarily for access. In looking at the documentation on their website (http://www.resourcespace.org/features) there is no reference to preservation or to selected preservation tools.   It looks like a system that supports access to collections not preservation.

3. **Shared Shelf**. Shared Shelf is associated with ARTstore. Although it's primarily a cataloging and media management system that institutions can subscribe to in order to upload, catalog, manage and make their images, audio, video, and documents available to users alongside the Digital Library collections for search and discovery in their teaching and research, as part of the Shared Shelf annual subscription fee, ARTstore does offer digital preservation support. See http://www.artstor.org/shared-shelf/s-html/features-preserve.html.

4. **CONTENTdm.** CONTENTdm is a commercial service provided by OCLC. The software can be run by OCLC as a service, but some institutions have installed the software locally. OCLC's CONTENTdm integrates with OCLC's Digital Archive (http://www.oclc.org/digital-archive.en.html) and DuraSpace's DuraCloud (http://www.duracloud.org/), which are both preservation systems. Members of the MetaArchive (http://www.metaarchive.org/) have uploaded data into the Private LOCKSS network that they operate.

5. **PastPerfect.** Past Perfect is a digital content management system that provides discovery and access. Institutions that locally manage their PP system will need to take responsibility for storage and preservation. If you use the hosted service (http://www.museumsoftware.com/pponline.html) you will need to find out what Museum Software does for storage and back-up. It might also be worthwhile to talk with folks like OCLC and DuraSpace to find out if the PP files could be uploaded to those systems and if so what is needed.

6. **Minisis.** Minisis out of Vancouver BC (http://www.minisisinc.com/page/18) offers a set of products that support museum collection management including archives, library and museum collections. Minisis at a recent meeting indicated that they will be introducing digital preservation services.

7. **Mozyhome.** Mozyhome is another online backup service, with some features similar to Google Drive and Dropbox. They're good for backups and access from multiple locations, but they don't provide a real preservation service.

8. **Carbonite.** Carbonite is another online backup service, with some features geared for businesses with multiple employees and computers. They're good for backups and access from multiple locations, but they don't provide a real preservation service

General questions about vendor-provided services

1. *What is a "ball park" figure of cost for one of these services such as DuraCloud or Preservation services?*

   The price varies by vendor and you need to contact them providing information on the type of service you're looking for the volume of data, size of institution, etc. Most of the hosted services are charging an annual subscription fee which is based on amount of data stored there may be additional fees for uploading and downloading the data. The software-only services such as Rosetta have a software licensing fee and then annual maintenance and volume of data fee. LOCKSS has an annual tiered membership fee. It's geared towards academic libraries and the fee is based on student FTE. There is no charge for the LOCKSS software.

2. *Are any of the companies approved for use by federal agencies?*

There does not seem to be "one" system for digital preservation recommended to all federal agencies. Different agencies have contracted with different providers for online backup and preservation services. NARA is developing its Electronic Records Archive (http://www.archives.gov/era/) for use by government agencies, but it is not yet fully available to many agencies.

3. *We joined a regional obituary indexing project and now we enter indexing info directly into their collection. We assume they are taking proper care of all the data they receive from all their partners, but what should we be asking them?*

    Don't assume anything. Check to find out what they are doing for storage and back-up. Then ask about preservation strategies for the files. Don't be surprised if they say they are backing up the files. You should work with them to explore strategies for digital preservation.

4. *I would like to join a collaborative preservation project. What is the status of TIPR, PeDALS, CINCH, etc.?*

    There are a variety of collaborative digital preservation efforts and they take on a variety of different activities. There are several state based collaborative digital preservation programs such as the Alabama Digital Preservation Network (ADPnet) and collection focused collaboration such as the MetaArchive's EDT preservation program. The MetaArchive, operated by Educopia Institute is a membership based organization open to all types of cultural heritage organizations. All of these operate using LOCKSS software.

    LOCKSS is the largest and best known collaborative digital preservation program, using distributed network approach to preserving e-journals. Portico provides for a centralized preservation e-journals and is working on preserving other digital resources. Libraries and other cultural heritage organizations concerned with the preservation of e-journals can subscribe to Portico.

    Persistent Digital Archives and Library System (PeDALS) Project (http://www.pedalspreservation.org/Default.aspx) was a project lead by the Arizona State Library and Archives and was funded in part through the National Digital Information Infrastructure Preservation Program 2008-2012. The project is now a loosely organized collaborative.

    There have been a variety of other collaborative digitization projects federally funded including the Florida Center for Library Automation's DAITSS (http://daitss.fcla.edu/), and the University of California at San Diego Chronoplis project (http://chronopolis.sdsc.edu/).

    In the last year the Digital Preservation Network (DPN) (http://www.dpn.org/) was formed to ensure that the complete scholarly record is preserved for future generations. DPN uses a federated approach to preservation. The higher education community has created many

digital repositories to provide long-term preservation and access. By replicating multiple dark copies of these collections in diverse nodes, DPN protects against the risk of catastrophic loss due to technology, organizational or natural disasters. DPN is welcoming new members.

In addition to projects that have implemented digital preservation programs, there are a variety of collaboratives that focus on tool development, education and research.

- National Digital Stewardship Alliance (http://www.digitalpreservation.gov/ndsa/) is the successor organization to National Digital Information Infrastructure and Preservation Program (NDIIPP) (http://www.digitalpreservation.gov) a program established by Congress in 2000 and administered by the Library of Congress. In 2010, the Library of Congress announced the formation of the National Digital Stewardship Alliance (NDSA), a collaborative effort to continue the work begun in NDIIPP. Any organization interested in digital preservation may join the NDSA.

  A list of partners and the projects in which they are involved is available at http://www.digitalpreservation.gov/partners/.

There have been many collaborative projects to develop digital preservation tools and software. Nationally and internationally there are many collaborative efforts including:
- Planets Project funded by the European Union. The Planets project ended on 31 May 2010. Planets results will be maintained and developed by a follow-on organisation called the Open Planets Foundation (OPF). OPF is a not-for-profit company, registered in the UK. To find out more about the OPF and how to join, please visit: www.openplanetsfoundation.org.

- CINCH (Capture INgest & Checksum Tool) (http://cinch.nclive.org/Cinch/site/page?view=project): CINCH was a collaborative effort funded by a 2011 Institute of Museum and Library Services Sparks! Ignition grant to create tools to support digital preservation.

- TIPR (Toward Interoperable Preservation Repositories) (http://wiki.fcla.edu/TIPR): A three-year project (2008-2011) funded by a National Leadership Grant from the Institute of Museum and Library Services. This was a proof of concept, and participants are now investigating similar interoperability options between services such as MetaArchive (http://www.metaarchive.org/), Chronopolis (http://chronopolis.sdsc.edu/), and DuraCloud (http://www.duracloud.org/), often using the BagIt specification developed by the Library of Congress and the California Digital Library (http://www.digitalpreservation.gov/documents/bagitspec.pdf).

Digital Preservation and the Cloud

1. *I've been told the cloud can crash with all data lost. Is that a real problem? We have opted*

for gold backed DVDs for now.

"The cloud" is really one or more servers somewhere that you don't have access to. While there may be glitches -- such as the Amazon issues that have gotten press recently -- they are still one viable approach to backups, and to some extent of preservation services. You need to look at the agreement/contract for each of the cloud based services—what they guarantee depends on the service.   The cloud should be just one of the back-up/storage options, not the only one.   Lastly CD's and DVDs are no longer considered best practice for storage. External hard drives are a little better, but you should be storing on servers. CD/DVD/flash drives are transport media.

Other questions? Email info@heritagepreservation.org and include "Caring for Digital Materials" in the subject. Or join the community at http://www.connectingtocollections.org/ and ask your questions there!