

0.1. Executive Summary

From September through December 2010 the Minnesota Digital Library (MDL) worked with HathiTrust (HT) to add content from Minnesota Reflections and the Minnesota Historical Society (MHS) to HathiTrust as a preservation archive. Nearly 50,000 images from Minnesota Reflections were shipped to HT before Christmas 2010 with another 8,000 MHS images due to be shipped soon after the holiday. HT plans to ingest these images and their metadata in early January.

While building this preservation archive MDL and HT learned a number of lessons:

- Master images at local institutions are often not formatted as required by HT and require transformation and the addition of embedded metadata. Many of these master images also lack fixity checks. The format requirements of HT may present a high-bar for potential participants.
- Items in the preservation archive require unique identifiers and that identifier namespace needs continual careful attention as new collections are added.
- Mapping metadata from local systems is difficult to routinize and would require ongoing attention in a long-term preservation effort. Properly packaging items for HT can also be time consuming. The different perspectives of MDL and HT concerning metadata has resulted in differences of directional intent over the project period, some of which have been resolved during this pilot project, and others of which must be revisited before a long-term program is undertaken.
- A programmer would be required in any long-term effort to integrate new collections, building scripts to do metadata mapping and packaging of objects.
- A trusting relationship with well defined responsibilities is required to allow for pragmatic solutions to data transfer since the MDL will likely end up with access to more information than it needs to complete the archiving task.
- Image data of the sort in Minnesota Reflections is quite a bit larger than the page images of books currently found in HT collections. This creates challenges for data transfer and package ingest.
- Local institutions may be more sensitive about image dissemination than HT expects. Rights issues have posed key challenges for MDL and HT during this pilot project. The project partners currently hold different expectations and requirements regarding rights and display.
- Once descriptions are ingested into HT, only the catalog information can usually be changed.
- This model cost about \$1 per image up front and \$0.10 per image in ongoing maintenance.

The delays imposed by the precise requirements of packaging pushed ingest into late-December, early January, well past the time this report was prepared. The lessons of ingest will have to be evaluated by the team once ingest has been completed. The lesson for the team is that whether the goals are accomplished or not, projects do have to come to an end. Both MDL and HT staff plan to continue work on ingest and retrieval into January without the project manager or preservation consultant directly involved.

Preservation services operate on a continuum from bit-level services to “full” preservation services that include the maintenance of a fully operational access-oriented content catalog. HT is in the highest end of this continuum. In this pilot project, MDL has explored the processes and workflows that would need to be actualized by a wide range of MN institutions in order to participate in HT preservation services. As MDL and HT continue to explore a potential longer-term relationship, it might be helpful to share pilot findings with representatives from across this range of institutions to see how their needs and abilities match up with this preservation service model.

To date, MDL has coordinated and hosted Minnesota Reflections and in this context, has worked with Minnesota-based institutions in a relatively informal manner. MDL now seeks to offer a new program consisting of preservation services to a Minnesota-based constituency. This project's Sponsors Group has agreed that to offer preservation services, MDL will need to create an entity with (or perhaps build into MDL) a higher level of formality in its governance, policies, and documentation than has previously been engaged in the Minnesota Reflections project.