



Update On April Activities

May 13, 2011

Top News

New PageTurner

HathiTrust released new functionality for the PageTurner application in April, improving the way volumes in the repository can be viewed and used. Enhancements to the PageTurner include:

- New views that allow users to scroll through volumes, flip pages similar to a physical book, and view thumbnail images of all pages in a volume
- Reorganized and streamlined interface including prominent display of copyright status, and re-positioning of navigation features
- Quick-copy links to volume pages in addition to permanent volume URLs
- Improved user experience for full book PDF downloads

Development of the new functionality was initiated by staff at the California Digital Library (CDL) in HathiTrust's collaborative development environment, and completed by staff at the University of Michigan. The Usability Working Group provided input and feedback on the interface design. The new views were built using Open Library's open source [BookReader](#). The thumbnail view was created specifically for HathiTrust by CDL staff, and has been incorporated as a standard feature in the core BookReader software.

We welcome comments and feedback on the new PageTurner. Please use the "Feedback" link that appears in the upper right corner of the page when viewing HathiTrust volumes, or email feedback@issues.hathitrust.org.

Support for Publishing

HTPub is an effort of the MPublishing Division of the University of Michigan Library to enable the use of HathiTrust as a platform for publishing open access electronic journals. It was first reported on in the [Update on October 2010 Activities](#), and has been in planning stages over the winter. MPublishing recently hired a summer intern who will be working with Michigan staff to define requirements for archival objects produced through HTPub. Michigan is in the process of hiring two full-time positions to support the work of the initiative. More information is available on the [HTPub project page](#).

MDL Images

John Butler of the University of Minnesota, John Weise of the University of Michigan, and project consultant Eric Celeste briefed CNI membership at the Spring 2011 Membership Meeting on the Minnesota Digital Library-HathiTrust image content prototype project. A summary of the project and slides for the presentation are available on the [HathiTrust website](#). Access to the images, now in the HathiTrust repository, will be enabled in late May or June. MDL has yet to draw conclusions regarding deposit of images in HathiTrust beyond the prototype phase. However, much has been learned throughout the project and HathiTrust intends to use the

May Forecast

- Continue work on the Data API security requirements
- Continue work on Full-text search enhancements

Papers & Presentations

Heather Christenson "[HathiTrust: A Research Library at Web Scale](#)"

J. Butler, J. Weise, E. Celeste "[Minnesota Digital Library and HathiTrust](#)"

Jon Wilkin, Jon Stroop, and Marvin Bielawski on [HathiTrust](#)

[New Partners Webinar](#)

All HathiTrust papers, presentations, and reports are available at <http://www.hathitrust.org>

There's an
elephant in
the library.





Update On April Activities

prototype and the experience gained as a base for developing general image ingest specifications that can be used for ingest of images from partner libraries.

Ingest Reports

HathiTrust has begun to post weekly reports on the ingest status of content submitted by partner institutions. The [reports](#) are available on the HathiTrust website, as well as a [description](#) of the information the reports include.

Ingest

Local Digitization Ingest

Michigan staff worked with Universidad Complutense de Madrid, Yale University, and the University of Illinois in April on ingest of locally-digitized volumes. We expect to begin ingest of volumes from Madrid in May, as well as the full set of volumes from Yale (a sample was ingested in December).

Harvard University

Ingest of an initial set of more than 50,000 volumes from Harvard University was completed in April.

Working Groups

Collections

The Collections Committee continues to work on a series of recommendations regarding duplicate volumes in HathiTrust, coordinated print management, and responding to users requests to contribute volumes to the repository. A draft discussion paper on duplicates will be shared with the Strategic Advisory Board in June for initial feedback.

Communications

The Communications Working Group finished a round of new partner webinars on April 12 and 15th. The webinars were well-attended and generated questions and rich discussion. The webinar [slides](#) and [audio recording](#) are available on the HathiTrust website. The working group also continued to craft a Facebook presence for HathiTrust, plan for a HathiTrust blog, and develop informational materials for use by partner libraries.

Total Volumes Added

	April	Overall
Columbia University	3	58,483
Cornell University	40,729	311,110
Harvard University	52,709	52,709
Indiana University	893	183,881
Library of Congress	0	71,418
NYPL	0	258,691
Penn State University	18	39,016
Princeton University	8,810	237,034
University of California	41,512	2,408,727
University of Chicago	0	5,172
University of Illinois	0	14,501
University of Madrid	15,486	103,797
University of Michigan	19,974	4,338,368
University of Minnesota	1,419	84,985
University of Wisconsin	10,602	454,332
Yale University	0	161
Total	192,155	8,662,385

Public Domain (~27% of total)

Total*	181,909	2,386,430
--------	---------	-----------

*This count includes volumes already in the repository to which rightholders have opened access

There's an elephant in the library.





Update On April Activities

Usability

The Usability Working Group made significant progress in April in developing a set of personas for HathiTrust users and scenarios of use. To help inform this draft set, the group has been gathering real life use cases from user feedback, reference interactions with users, and uses of HathiTrust that have been posted in blogs and tweets. It has also been analyzing HathiTrust usage statistics for trends. The personas and scenarios are intended to inform development and policy-making surrounding HathiTrust applications and interfaces. The group anticipates having the draft set of personas and scenarios ready to share with partner institutions and other HathiTrust working groups in May. The personas will be refined over time as additional use cases are assembled and user research conducted.

The Usability Group is still accepting volunteers to join the new User Experience Special Interest Group (UX-SIG), reported in [February's update](#). Please contact Suzanne Chapman (suzchap@umich.edu) if you are interested in joining this group or have any questions about participation.

User Support Working Group

During March and April, the chair of the User Support Working Group chair coordinated with staff members at the University of Michigan who have been handling user feedback for HathiTrust, to configure a partner-wide issue tracking system using JIRA. User Support members began accessing the system in April and observing the preliminary processes that had been put in place. The working group will assume responsibility for responding to issues and directing feedback as appropriate to partner institutions and working groups in May. Michigan staff will continue to play an integral role in addressing issues related to content quality and bibliographic metadata.

Projects

IMLS Quality Grant

The grant project team continued to refine definitions for the preliminary set of quality errors they have identified within volumes, and make improvements to the quality review application interface. The team continued to focus on dual review of volumes (two reviewers coding the same set of volumes) to identify problematic error definitions and refine descriptive wording to better illustrate each error type. The team also revised definitions for the scale of severity that is applied to errors, in order to improve inter-coder consistency. A second sample of 10 public domain volumes was reviewed by project staff to provide sufficient data for the project statistician to develop appropriate sampling techniques for Phase Two of the project: production level coding. The University of Minnesota will be joining in data collection efforts and will begin remote reviewing in the next two months after a series of training sessions with members of the project team. Background information

You can follow HathiTrust on Twitter at <http://www.twitter.com/hathitrust>

There's an
elephant in
the library.





Update On April Activities

on the project can be found on the [grant projects page](#) of the HathiTrust website.

Development Updates

Bibliographic Data Management

The HathiTrust Metadata Management System team completed development of the core database system in April, as well as an API to export bibliographic data in XML format. Approximately 200,000 records have been loaded into the system for initial testing. The team is analyzing MARC records from current content-contributing partner institutions, received from the University of Michigan, looking for irregularities and performing a general survey of the record set. CDL staff continue to interview for a Principal Metadata Analyst. Details on the project are available at <http://www.hathitrust.org/htmms>.

Data API

Staff at Michigan have completed a rough draft of requirements for improved security in the Data API based on symmetric key cryptography. The draft will be made available for comment in the near future.

Development Environment

New MySQL servers installed in the development environment by staff at the University of Michigan have boosted performance of print holdings database operations by an order of magnitude. Similarly-configured servers will be installed in the production environment in May.

Full-text Search

Michigan staff began development work on priority features for full-text search as identified in the [Full-Text Search Working Group's report](#). The implementation team is focusing initially on relevance ranking of search results based on a combination of full-text OCR and bibliographic metadata, and on faceting of results using bibliographic metadata. The goal is to release significant new features that use the bibliographic data to enhance full-text search results by July 1, 2011.

Storage Replacement Cycle

All replacement storage equipment at the Michigan and Indiana storage sites is online and in use. The storage equipment that was replaced is being wiped for security purposes by staff at the University of Michigan and will be traded in for a credit on new storage that will be purchased in June 2011.

Outages

There were no outages in April.

**There's an
elephant in
the library.**

