



Update On February Activities

March 13, 2015

Top News

Anne Kenney Appointed to HathiTrust Board of Governors

Anne R. Kenney, Carl A. Kroch University Librarian at Cornell University, has joined the HathiTrust Board of Governors. Kenney joined Cornell in 1987 and became University Librarian in 2008. She is known for pioneering work in developing standards for digitization and research in digital preservation. She currently serves on the Board of the Council on Library and Information Resources, and is also a fellow and past-president of the Society of American Archivists. Kenney's term on the Board of Governors will last for the remainder of 2015 to temporarily fill a vacancy left by the resignation of Patricia Steele of the University of Maryland. HathiTrust will hold elections later this year to fill this seat and to replace two other Board members whose terms are expiring. For a complete list of Board members and their terms please visit: http://www.hathitrust.org/board_of_governors.

New Zephir Metadata Analyst

California Digital Library welcomed Dana Jemison as the new Zephir team metadata analyst. She will be taking over these duties from Renata Ewing. Dana comes to CDL from the University of California, Berkeley where she worked in the Library Systems Office, and was formerly of the Research Libraries Group where she worked in Research and Development.

Departure of Long-time University of Michigan Staff Member

Cory Snavely, Manager of Library IT Core Services at the University of Michigan announced his departure to join the Lawrence Berkeley National Laboratory. Under his guidance the University of Michigan established and scaled the underlying systems that support HathiTrust's Trustworthy Digital Repository - from servers, storage, and system architecture, to ingest and auditing processes, content specifications, security, and so much more. He also played an instrumental role in the establishment of technical infrastructure for the Digital Preservation Network (DPN), and HathiTrust's configuration as a DPN node. We are grateful for all of the contributions Cory has made, and wish him well in his new position.

HathiTrust Research Center UnCamp

Please join us for the third annual HTRC UnCamp at the University of Michigan, March 30-31. The agenda, list of participants, and registration information are available on the UnCamp [event page](#).

Calendar for Print Holdings Data and Content Estimates

We will be issuing a call for member print holdings information and estimates of content to be deposited at the beginning of April this year, with the information due by June 30. We are moving the schedule of receiving this information forward in order to have the 2016 budget and fees prepared for member voting earlier in the fall.

March Forecast

Deploy improvements to accessibility features of PageTurner

Incorporate coordinate OCR into PDFs of HathiTrust content

Test Solr4 plug-in that reduces memory use in indexing

Enhance registration process for staff who have special access to materials in HathiTrust

HathiTrust on the Road

HathiTrust administrative staff will be attending the following upcoming meetings. Please get in touch if you would like to meet with us there:

Mike Furlough, Jeremy York, Angelina Zaytsev, DPLAfest, Indianapolis, April 17-18, 2015.

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





Update On February Activities

Board of Governors

The HathiTrust Board of Governors met by phone on February 23, 2015 and addressed the following topics.

Board Membership

To fill a vacancy left by Patricia Steele's resignation, the Board appointed Anne Kenney, Carl A. Kroch University Librarian, Cornell University, to serve through the end of 2015. An elected replacement to fill the remaining year of Steele's term will be elected during the next regular Board elections.

Government Documents Initiative

The Board reviewed the report of the Government Documents Initiative Advisory and Working Group, and the recommendations of the Program Steering Committee for further action on the Government Documents Initiatives. Executive Director Mike Furlough also reported on progress on the Government Documents Registry and discussions with the Government Publications Office. Discussion focused on the need for continued investment in the Initiative, including potentially new staffing and digitization. Furlough will draw upon the report and the PSC recommendations to draft a preliminary implementation plan, including staffing, for consideration at the next Board of Governors Meeting. The Advisory and Working Group report will be made public soon.

Bylaws Revisions

The Board reviewed proposed changes for the Bylaws and the schedule for member voting.

2015 Planning Calendar

The Board reviewed a schedule for major actions to be taken in 2015. These include:

- Appointment of the 2015 Nominations Committee (spring)
- Review of the recommendations of the Shared Print Planning Task Force (spring)
- Acting on recommendations of the Government Documents Initiative Advisory and Working Group. (spring)
- Appointment of new Program Steering Committee members (spring)
- Completion of an MOU with Michigan for the operation of the repository infrastructure and hosting HathiTrust administrative functions. (summer/fall)
- Financial and strategic planning. (summer/fall)
- Election of new Board of Governors members. (fall)
- 2nd Annual Member Meeting (fall)

You can follow HathiTrust on [Twitter](#) or [Facebook](#)

[Subscribe to email updates](#)
(via Google Groups)

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





Update On February Activities

Michigan/HathiTrust MOU

During 2015 The Board of Governors and the University of Michigan will develop a Memorandum of Understanding to document Michigan's roles in hosting the administrative, financial, and technical operations of HathiTrust. The Board reviewed the current status of this effort.

Financial and Strategic Planning

The Board will oversee the development of a long-term budget plan for HathiTrust in the coming year. Mike Furlough led a discussion of major factors to consider in developing this plan, and methods of gathering data for it.

Improvements to PageTurner (Late Breaking)

HathiTrust released a number of changes to the PageTurner interface, reducing complexity and improving presentation of items while simplifying the underlying code to facilitate future development. The cache on some browsers may need to be cleared to view the improvements. The full list of changes includes the following:

- Toolbars are fixed at the upper right of the page and never scroll out of view
- The global search and login options have been moved to the navigation bar and are always available
- Accuracy of scrolling in the thumbnail view is improved
- Reader views now update the page "size" parameter, allowing users to retain the same size of page when returning to or refreshing a page
- "Flip" view performance is improved in the Internet Explorer browser version 9

Ingest

Locally-digitized Content

HathiTrust corresponded with Boston College, Northwestern University, University of Maryland, Cornell University, and University of Washington about ingest of locally-digitized materials. The University of Missouri deposited one volume as a precursor to future ingest.

Internet Archive-digitized Content

HathiTrust continued to ingest dissertations from the University of Massachusetts.

Bibliographic Data Management

The California Digital Library (CDL) loaded 78,092 new and 142,327 updated bibliographic records into Zephir.

Projects

There's an
elephant in
the library.™





Update On February Activities

March 13, 2015

Copyright Review

A summary of the determinations from HathiTrust copyright review activities in February is given below. See [CRMS-US](#) and [CRMS-World](#) for further information.

	February		Overall	
	Public Domain	All Determinations	Public Domain	All Determinations
CRMS-US	1,268	1,901	169,374	320,593
CRMS-World	4,477	8,121	96,734	182,633
Total	5,745	10,022	266,108	503,226

Government Documents Registry

HathiTrust staff continued to refine the process for detecting relationships between US federal government documents records (including duplicates), and to analyze the overlap between agency authority entries in VIAF and an initial set of Registry records. To date, staff have pre-processed 19 million bibliographic records from this initial set and records submitted in response to HathiTrust's [call for records](#). Staff also began planning for a public version of the Registry, including decisions about the indexing tool (Solr) and discovery interface (Blacklight) to use and specific fields to index and display. Further information about the specifications will be forthcoming.

HathiTrust Research Center Updates

Advanced Collaborative Support (ACS)

HTRC held an online kickoff meeting with each of the three inaugural ACS projects, in addition to an already ongoing ACS project with University of Toronto. All the projects have started and each will deliver a final report.

HTRC Services 3.0 Final Release on Feb 27 2015

The HTRC team made the final release for 3.0 on Feb 27 2015. This version addresses more than 50 issues reported over the period of the Beta test. Thanks to everyone who reported a problem or made a suggestion.

The main improvement for the 3.0 release is an updated user account registration process with more intuitive email recognition, and a User Agreement that more clearly spells out user responsibilities for handling HathiTrust data.

Other features introduced with the 3.0 Beta release include:

- Data Capsule - a secure environment for non-consumptive research
- More welcoming home page and portal

There's an elephant in the library.™





Update On February Activities

March 13, 2015

- Enhanced workset builder functionality
- Automatically saving jobs upon completion
- Corrected use of faceted search
- Single sign-on (except for Data Capsule and Workset Builder)

mPach

The mPach project has been put on hold indefinitely as HathiTrust and the University of Michigan reevaluate needs and opportunities in digital publishing that have emerged since Michigan began the mPach project in 2011. mPach was conceived as a suite of tools to enable direct publishing of open access journals into HathiTrust's preservation and access environment. Michigan remains strongly committed to providing robust long-term preservation and access services for digital publications, and HathiTrust remains strongly committed to supporting a variety of formats of textual publications, including born-digital and newly published materials. Michigan and HathiTrust are reconsidering how best to meet these goals, and have determined that the particular suite of tools and workflows envisioned for mPach do not align with current needs and trajectories. HathiTrust will be providing updates as planning for support of born-digital and other types of textual materials moves forward.

Development Updates

Development updates and activities by HathiTrust institutions included the following:

Access, Authorization, and Authentication:

- Fixed a bug in the Data API key expiration notification process.
- Added support for a new rights attribute to restrict access to materials that are in the public domain but must remain closed due to privacy concerns.
- Added criteria to the information that is used to restrict access to items in HathiTrust based on the location of the user (e.g., for materials that are public domain only when viewed from the United States).

Full-text Search

- Completed re-indexing of the entire repository using Solr4, including making needed adjustments to the indexing process to accommodate differences in Solr4.
- Prepared to deploy enhancements to the use of date information, which will significantly improve the ability to facet and limit searches by date of publication. The enhancements will be put into production in early March.
- Created a plug-in for Solr4 to reduce memory use. Testing of the plug-in will proceed in March.

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





Update On February Activities

- Received and installed an early release of a production-quality software fix for the high-performance storage system to address performance and stability problems. Staff are currently working closely with the storage vendor on the final steps of configuring and securing the system.
- Completed software upgrades on the 40Gb networking equipment which supplies connectivity to the storage system. Further testing and a gradual production phase-in are expected in late March or early April.

Storage Replacement Cycle

- Completed installation and replacement of storage for the 2015 cycle. Retired storage is currently undergoing security wiping before being taken off-site for disposition.

Availability

Repository

Cumulative 12-month availability of repository access (user-facing applications): 99.972% (+0.008%)

No outages were reported in February.

Papers and Presentations

Mike Furlough, "[HathiTrust: An Update for the University of California](#)," California Digital Library, Oakland, CA, February 23, 2015.

Jeremy York, "[Preservation With A Purpose: End User Services in HathiTrust Digital Library](#)," Rutgers University, February 24, 2015.

Mike Furlough, "[HathiTrust, Collective Action, and Local Services](#)," University of California, Davis, Davis, CA, February 26, 2015.

Mike Furlough, "[The HathiTrust Research Corpus](#)," University of California, Davis, Innovating Communication in Scholarship (ICIS) Seminar, Davis, CA February 26, 2015.

Mike Furlough, "[HathiTrust, Collective Action, and Local Services](#)," University of California, Berkeley, Berkeley, CA, February 26, 2015.

HathiTrust Research Center

Beth Namachchivaya, "[HTRC Update](#)," ALA Midwinter Chief Collection Development Officers Meeting, January 31, 2015.

Beth Namachchivaya, "[HathiTrust Research Center & Support for Text Data Mining Research: Opportunities for Campus Collaboration](#)," University of Illinois, February 20, 2015.

There's an
elephant in
the library.™





Update On February Activities

Harriett Green, Sayan Bhattacharyya, “The Savvy Researcher - [Slides, Workshop-handout, Plan/Outline](#),” Workshop series on HTRC 3.0 at the University of Illinois Scholarly Commons, February 16, 2015.

Beth Plale, Dirk Herr-Hoyman, Jiaan Zeng, Zong Peng and Miao Chen, Workshop on HTRC 3.0, Indiana University February 25, 2015.

J. Stephen Downie, “HathiTrust: Large-Scale Repository in the Humanities -- Unlocking the Secrets of 4.6 Billion Pages,” Hong Kong University of Science and Technology, February 27, 2015.

Megan Senseney “[Digital Humanities and the HathiTrust Research Center](#),” [Illinois Digital Humanities Symposium](#), February 28, 2015.

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





Update On February Activities

Total Volumes Added	February	Overall
Boston College	0	3,263
Columbia University	0	73,396
Cornell University	5,458	515,744
Duke University	0	8,206
Emory University	0	52
Getty Research Institute	568	20,130
Harvard University	7	838,122
Indiana University	165	529,766
Keio University	8	90,120
Knowledge Unlatched	0	28
Library of Congress	0	108,892
McGill University	0	893
New York Public Library	9,721	304,604
North Carolina State University	0	3,196
Northwestern University	37	56,992
Ohio State University	677	69,094
Penn State University	507	389,220
Princeton University	4	252,841
Purdue University	0	47,488
Sterling & Francine Clark Art Institute	0	358
Texas A&M	0	2,446
Universidad Complutense	19	117,310
University of Alberta	0	76,106
University of California	10,143	3,625,049
University of Chicago	4,264	56,402
University of Connecticut	0	4,637
University of Delaware	0	48
University of Florida	0	9,866
University of Illinois	9,992	339,128
University of Massachusetts, Amherst	3	12,007
University of Michigan	4,934	4,721,293
University of Minnesota	140,539	333,663
University of Missouri	1	1
UNC - Chapel Hill	0	17,025
University of Virginia	0	51,207
University of Wisconsin	32	561,126
Utah State University	0	117
Yale University	0	23,832
Total	187,079	13,263,668

Public Domain (~37% of total)

Total*	69,308	4,967,590
--------	--------	-----------

*Includes works opened via copyright review and rights holder permissions.

Most-accessed volumes

Quicksand, by Nella Larsen.
Modern California Houses: Case Study Houses, 1945-1962, by Esther McCoy.
The lesson of Japanese architecture, by Jiro Harada.
The Human Figure, by John H. Vanderpoel
Roster of the Confederate soldiers of Georgia, 1861-1865, v.2.
Godey's Magazine, v.40-41, 1850.
Solid mensuration, by Willis F. Kern and James R. Bland.
Roster of the Confederate soldiers of Georgia, 1861-1865, v.1.
The Five Laws of Library Science, by S. R. Ranganathan.
History of wages in the United States from Colonial times to 1928, United States Department of Labor.

There's an elephant in the library.™





Update On February Activities

User Support Issues	February	January
Content	227	158
Quality	211	143
Collections	12	15
Cataloging	164	142
Access and Use	157	121
Copyright	105	76
Permissions	18	8
Takedown	0	0
Print on Demand	0	0
Inter-library loan	2	0
Full-PDF or e-copy requests	12	11
Datasets	5	2
Data Availability and APIs	3	1
Reuse of content	6	1
Web applications	41	28
Functionality problems	25	12
Problems with login specifically	1	0
General questions about login	3	0
Partners setting up login	0	1
Usability issues	0	0
Feature requests	1	3
Partner Ingest	7	6
General	134	103
Partnership	8	9
Miscellaneous	126	94
Total	730	558

*See [User Support Working Group Issue Types](#) for a description of the types of issues included in each category.

There's an elephant in the library.™