



Update On May Activities

June 14, 2013

Note from John

John Wilkin, Executive Director, HathiTrust

Some five years ago, I was appointed as the founding Executive Director for HathiTrust, and the intervening years have been tremendous. We have a long list of utterly extraordinary accomplishments, made possible by the partnership, by our governance, and by a very talented group of staff. I don't think it's a stretch to say that we've changed the world. We've helped libraries see what's possible through deep collaboration, and ours is sure to be a model for other transformative efforts. I feel deeply honored to have played a role in all of this.

Consequently, it is with mixed feelings that I share that I will be stepping down as Executive Director on August 16th as I begin my very exciting new responsibilities as the next Dean of Libraries at the University of Illinois. I follow in the footsteps of one of the most outstanding library deans in the country (and one of HathiTrust's first executive committee members), Paula Kaufman.

My friends and collaborators on the HathiTrust Board of Governors are poised to facilitate a smooth transition. A search will take place very soon, and I will continue to work with HathiTrust both to help bridge the gap and in my new capacity at Illinois. The strength of our collective individual commitments and the robustness of the organization we have put in place will ensure a bright future for HathiTrust. I want to thank everyone in the organization and at the partner institutions for all you've done to make HathiTrust great, and for the support you've given me over these years!

Top News

HathiTrust Program Steering Committee Membership Finalized

HathiTrust is pleased to announce the membership of the inaugural Program Steering Committee (PSC). The members, whose terms begin immediately, include:

- Ivy Anderson (California Digital Library)
- John Butler (University of Minnesota)
- Todd Grappone (University of California, Los Angeles)
- Mike Furlough (Penn State)
- Martha Hruska (University of California, San Diego)
- Martin Kurth (New York University)
- Erika Linke (Carnegie Mellon University)
- Matthew Sheehy (Harvard University)
- Elaine Westbrooks (University of Michigan)
- Bob Wolven, Chair (Columbia University)

The Program Steering Committee serves at the direction of the Board of Governors in the following manner:

Late-breaking news

HathiTrust is pleased to welcome Brown University to the partnership! View the [full announcement](#).

Papers & Presentations

John Wilkin and Sigrid Cordell, "[HathiTrust: The Collection and Its Uses](#)", ELI Webinar, May 6, 2013.

Beth Plale and Yiming Sun, "[Digital Humanities At Scale: HathiTrust Research Center](#)", Notre Dame, May 7, 2013.

Downie, J. Stephen, "Unlocking the Secrets of 3 Billion Pages: Introducing the HathiTrust Research Center", University of Waikato, New Zealand, May 15, 2013 and University of Canterbury, New Zealand, May 20, 2013.

Jeremy York, "[Institutional Uses of HathiTrust](#)", Maine Shared Collections Strategy Advisory Meeting, September 23, 2013. [Similar presentation](#) given at University of Maine, September 24, 2013.

There's an
elephant in
the library.™





Update On May Activities

- Reviews HathiTrust's development agenda, shaping initiatives and strategies for Board discussion and decision-making, and considering the implications of those initiatives for the future. Working Groups can be appointed and charged by the Chair of the Program Steering Committee to assist with its work.
- Reports to the Board of Governors recommended alterations in the development agenda based on such reviews.
- Based on its reviews, develops position papers for the member community to encourage debate or mobilize discussion with regard to particular issues.
- Works with the Board of Governors to develop policies for HathiTrust and its members.

More information about the Program Steering Committee is available in the [HathiTrust Bylaws](#).

Update on Constitutional Convention Ballot Initiatives

With our final governance-related steps in place, the HathiTrust Board of Governors felt it would be useful to provide a brief update on the October 2011 Constitutional Convention ballot initiatives. The [ballot initiatives are available online](#).

Distributed Monographs Archive

HathiTrust has not yet undertaken work toward the establishment of a distributed network of print monograph archives. With the appointment of the PSC, we will develop a plan and will be in communication with member institutions in coming months.

Approval Process for development initiatives

HathiTrust has not yet undertaken work on an approval process for development initiatives. In light of the governance changes that have occurred since the proposal was made, the Board has determined that it will approve future development initiatives. The Board will develop appropriate procedures for review, communication, and prioritization of proposals related to HathiTrust initiatives and development.

Governance

With the establishment of the Program Steering Committee, this work is complete. The Board of Governors was established in April 2012. Bylaws were accepted by the member institutions in February 2013, establishing 3-year terms for members of the Board. (The current elected Board members have 3, 4 or 5 year terms to stagger appointments.) The Board formed an Executive Committee and elected officers (announced Oct. 2012). The terms of the officers are enumerated at http://www.hathitrust.org/board_of_governors. Nominations were received for the PSC, and the PSC members have been appointed (see the announcement above and <http://www.hathitrust.org/psc>).

June Forecast

Continue work on full-text relevance ranking.

Continue exploration of the Solr experimental block-join functionality.

Continue to design a process for indexing the full text of JATS XML articles.

Continue work to enable the delivery of JATS XML articles as PDFs.

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





Update On May Activities

Government Documents

HathiTrust has done background work on the partner [ballot initiative](#) related to digitization of US federal government documents. This work includes gathering information about the status of a variety of existing and planned government documents digitization initiatives (e.g., in the CIC), and assembling lists of reference and general information sources about digitized and non-digitized government documents. In April 2013, HathiTrust hired Valerie Glenn for a 2-year project to build a registry of US federal government documents. The purpose of the registry is to form a database so that we can identify the corpus of US federal documents--both currently digitized and those that could be digitized.

Mission and Goals

HathiTrust has not yet undertaken work to revise HathiTrust's mission and goals.

Fee for service

HathiTrust has not yet undertaken work toward the establishment of a fee-for-service model of content deposit for non-HathiTrust partners.

Nominations for User Support Working Group

The User Support Working Group is seeking nominations for up to 3 new members. This is a great opportunity to get involved in the day-to-day work of the partnership. Nominations should be submitted by July 1, 2013 via the form at <http://tinyurl.com/m9qlyyg> and include the name, title, email address, and a short description of current job duties for the nominated staff member. Additional information that might be relevant to participation in the group may be included as well. User Support members are on call at least one day per week and follow up on inquiries throughout the week, requiring between 2-4 hours of work. Staff that participate on the group will:

- Gain knowledge about HathiTrust's user base, typical problems and questions that are raised, and how these problems and questions are resolved.
- Become aware of new ways HathiTrust is being used, and features and functionality that users desire.
- Gain knowledge of HathiTrust organizational and technical infrastructure, and policies and procedures relating to copyright, access, collection development, deposit of materials, and preservation.
- Contribute to a vital component of partnership operations.

The charge for the working group is available at http://www.hathitrust.org/wg_user-support_charge.

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

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

There's an
elephant in
the library.™





Update On May Activities

Ingest

Local Digitization

HathiTrust corresponded with contacts from the University of Maryland, Texas A&M University, The University of Chicago, the University of Iowa and the University of Illinois about requirements for locally-digitized content.

Working Groups and Committees

User Experience Advisory Group

The User Experience Advisory Group concluded discussions of user interface priorities and will be sharing results soon.

User Support Working Group

A summary of User Support inquiries received in May is included at the end of the update.

Projects

Bibliographic Data Management

Coding work continues on Zephir (the new metadata management system) to reflect recently revised requirements and workflow changes. The California Digital Library (CDL) team is on target to complete loading of a subset of HathiTrust bibliographic records to its development environment in early June. This load will incorporate the removal of rights determination code as well as improved metadata extraction to support core management functions and enhanced analysis. The CDL team is working with staff at the University of Michigan on refining strategies to ensure that Zephir includes the same records as those in the current system. This month has also seen a successful test of a new backup system for Zephir. The timeline for the project is available at <http://www.hathitrust.org/htmms>.

Copyright Review

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

	May		Overall	
	Public Domain	All Determinations	Public Domain	All Determinations
CRMS-US	3,720	7,935	134,947	251,935
CRMS-World	2,794	5,439	26,606	49,492
Total	6,514	13,374	161,553	301,427

You can follow HathiTrust on [Twitter](#) or [Facebook](#)
 Subscribe to email updates (via [Google Groups](#))

There's an elephant in the library.™

www.hathitrust.org 



Update On May Activities

HathiTrust Research Center (HTRC)

The HTRC has kicked off its second phase of development, which includes a substantial focus on community-building. To that end, the HTRC devoted the month of May to user engagement, visiting with user groups at Notre Dame University and gearing up for a series of focus groups and interviews at upcoming conferences, including Digital Humanities 2013 and JCDL 2013, to gather information about researcher projects and requirements. If you are already using HTRC resources, or are interested in starting, please [join our user group mailing list](#) to contribute recommendations and participate in technical and development discussions with the HTRC team.

The HTRC is planning its second annual UnCamp for September 8-9, to be held at the iHotel and Conference Center in Champaign, IL. Save the date!

mPach

University of Michigan staff posted mockups on the mPach [website](#) of enhancements to be made to the Collections and PageTurner applications to support journal articles published using mPach. Staff updated the [project timeline](#) and prepared to travel to Texas A&M University to present on mPach to staff from the library, university press, and the Integrated Ocean Drilling Program, and discuss TAMU's envisioned use of mPach.

Development Updates

HathiTrust institutions performed the following work related to applications and Web interfaces:

Access for Users who have Print Disabilities

Staff expanded the number of simultaneous accesses available to users at HathiTrust partner institutions who have print disabilities per print copy of a volume owned by the user's institution. The number of simultaneous accesses was previously limited to the number of copies held by the user's institution. Now there is no limit on the number of simultaneous accesses as long as the institution holds or previously held a print copy.

Staff also added a checkbox, which appears only to authenticated members of partner institutions, to the [advanced full-text search page](#), allowing users to limit a search to items held in print by their institution. This was implemented specifically to aid users who have print disabilities in identifying works from their institution that are available to them. A description of how to use the checkbox, and other information about access for users who have print disabilities, can be found at <http://www.hathitrust.org/accessibility>.

There's an
elephant in
the library.™





Update On May Activities

Full-text Search

Michigan staff discussed issues surrounding the implementation of better relevance-ranking systems for books with attendees at the Lucene Revolution Conference, including several Lucene and Solr committers. Following these discussions, staff began to investigate Lucene and Solr's [grouping and joining capabilities](#) and the potential to do relevance ranking based on the relevance of book chapters or other book parts.

Staff performed work to optimize indexing processes at times when maintenance or large updates that affect full-text indexing are underway, and began to design a process for indexing the full text of JATS XML articles.

Staff at California Digital Library improved the performance of a spelling suggestion feature that is in development, doubling the previous speed. Staff began to refactor code for the feature and implement a web service interface.

PageTurner

Staff tuned heuristics for determining whether to display volumes in the interface from left to right or right to left (depending on the language). Staff also

- corrected a bug that prevented PDFs for volumes that are read from right to left from being searchable;
- added an additional notice to PDFs generated by proxies for users who have print disabilities; and
- continued work to enable the delivery of JATS XML articles as PDFs.

Storage Hardware Replacement Cycle

HathiTrust suspended storage replacement activity in May due to complications with a storage system software upgrade. The software upgrade was completed in May, and storage due for retirement will be taken offline starting in June.

Total Volumes Added

	May	Overall
Boston College	0	2,179
Columbia University	0	65,033
Cornell University	32	419,836
Duke University	0	4,523
Harvard University	0	236,068
Indiana University	10	195,237
Library of Congress	0	89,724
North Carolina State University	0	3,196
Northwestern University	408	33,969
New York Public Library	1	288,343
Penn State	591	59,187
Princeton University	2	251,704
Purdue University	0	44,692
Universidad Complutense	0	111,982
University of California	770	3,388,795
University of Chicago	106	30,391
University of Florida	0	2,068
University of Illinois	4	109,318
University of Michigan	6,380	4,643,131
University of Minnesota	544	106,819
University of North Carolina - Chapel Hill	0	16,588
University of Wisconsin	8	555,733
University of Virginia	0	50,815
Utah State University	0	117
Yale University	0	23,678
Total	8,856	10,733,126

Public Domain (~31% of total)

Total*	12,504	3,386,026
--------	--------	-----------

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





Update On May Activities

User Support Issues

	May	April
Content	390	379
Quality	386	368
Non-partner Digital Deposit	1	1
Collections	13	9
Cataloging	135	140
Access and Use	144	145
Copyright	79	71
Permissions	10	13
Takedown	1	0
Print on Demand	0	0
Inter-library loan	0	4
Full-PDF or e-copy requests	20	21
Datasets	5	4
Data Availability and APIs	1	1
Reuse of content	7	5
Web applications	33	31
Functionality problems	13	7
Problems with login specifically	3	1
General questions about login	2	1
Partners setting up login	1	0
Usability issues	1	7
Feature requests	1	2
Partner Ingest	10	3
General	57	76
Partnership	5	12
Infrastructure	0	2
Miscellaneous	52	62
Total	769	774

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

Most-accessed volumes

Title
Mechanick exercises: or The doctrine of handy-works, by Joseph Moxon.
Quicksand, by Nella Larsen.
Mary Meigs Atwater Recipe book : patterns for handweavers / Mary M. Atwater.
A treatise on money, v.1 1930, by John Maynard Keynes.
Theoretical nuclear physics, by John M. Blatt [and] Victor F. Weisskopf.
Investigation of Korean-American relations: Report of the Subcommittee on International Organizations of the Committee on International Relations, U.S. House of Representatives, October 31, 1978.
Pathology of tropical and extraordinary diseases, edited by Chapman H. Binford and Daniel H. Connor.
Perfume and flavor materials of natural origin.
Forerunner. v.1 1909/1910
Report of the commanding general of the Army air, 1945 Dec.

Outages

HathiTrust users may have been unable to perform searches within a book on Wednesday, May 1 from 11:30am - 4:40pm due to a change in system configuration that had an unintentional side effect at one site.

HathiTrust users may have experienced slow page loading or errors in page loading on Wednesday, May 8 from 2:30-4:15pm due to a database performance problem related to a large update to partner print holdings information.

There's an elephant in the library.™

