



### Greetings from the HathiTrust Team

Wow! It's been a busy few months for us at HathiTrust since we put out our last update in early summer. At that time Sandra McIntyre, Lizanne Payne, and Heather Christenson had just joined our team, and we got to know each other by launching right into things. By the end of the summer the Ann Arbor staff had relocated to new offices, and to celebrate we held a one day meeting that included representation from all HathiTrust services at the University of Michigan, California Digital Library, and the Research Center at Indiana University and the University of Illinois.

But as this newsletter shows, we've been working on all fronts. In late June we proudly announced a [new partnership with the National Federation of the Blind](#) to expand access to HathiTrust to its members and users of its services. Over the summer we developed plans for launching our [Shared Print Program](#) and for re-starting copyright review of US publications dated 1923-1963, and got them underway by early fall. We have launched a new working group, focused on quality issues, and advisory committees for both the Shared Print Program and the [US Federal Documents Program](#). In August we released [the results from our survey of HathiTrust Members on our collection priorities](#), which has directed us to stay focused on our historical strength in collecting digitized books from research collections. Meanwhile the HathiTrust Research Center teams have been developing new outreach efforts and reviewing our policies for non-consumptive research using in-copyright works. We've added several new members and are working on several others. Finally, we held our annual Board of Governor's election and announced our new members. This newsletter explains all of this activity and more.

Looking ahead to the remainder of the year, you can expect to hear from us on a couple of critical matters. Our annual member meeting will be held on November 10, giving us a chance to hear directly from many of you about our work and plans. We are also finishing up the annual budget and fees process for 2017, and will send these to members for approval by electronic ballot before the end of October. We expect 2017 to be one of our most prolific years in a long time and as ever we are thrilled to be able to work with you all!

Mike Furlough  
Executive Director



*HathiTrust staff at The University of Michigan Hatcher Graduate Library in June 2016.  
Back Row, L-R: Sebastien Korner, Tim Prettyman, Jon Rothman, Bill Dueber, Heather Christenson, Josh Steverman, Tom Burton-West.  
Middle Row, L-R: Valerie Glenn, Roger Espinosa, Aaron Elkiss, Chris Powell, Lizanne Payne.  
Front Row, L-R: Kristina Eden, Angelina Zaytsev, Mike Furlough, Melissa Stewart, Sandra McIntyre, John Weise.*





### Top News

#### 2016 Board of Governors Election Results

The members of HathiTrust have elected four library directors to serve on its [Board of Governors](#) beginning in 2017. **Kevin L. Smith**, Dean of Libraries at the University of Kansas, and **Sarah E. Thomas**, Vice President for the Harvard Library and University Librarian, Harvard University, will serve three-year terms ending in December 2019. **Beth McNeil**, Dean of Libraries at Iowa State University, has been reelected to the Board of Governors to a term that will conclude in December 2018. **Joseph P. Lucia**, Dean of Libraries at Temple University, will serve a one-year term concluding in December 2017. “Our colleagues have elected an outstanding group of leaders,” said Richard Clement, past chair of the Board and chair of the 2016 Nominating Committee. “Their deep knowledge and range of expertise brings still more depth to the HathiTrust Board of Governors.”

Exiting the Board this year are Richard Clement, University of New Mexico, and Robert Wolven, Columbia University, who is retiring one year prior to the end of his elected term. Anne Kenney, Cornell University, has announced plans to retire in April 2017. Betsy Wilson, Chair of the HathiTrust Board, noted that “Rick, Bob, and Anne have all provided outstanding service and leadership on the board over the last few years and I thank them on behalf of the Board. We all are looking forward to Beth’s continued service and to working with Kevin, Sarah, and Joe.”

Read more at: <https://www.hathitrust.org/hathitrust-announces-2016-board-of-governors>.

#### 2016 HathiTrust Member Meeting

The 2016 HathiTrust Member Meeting will be held on November 10, 2016 at the Big Ten Center in Rosemont, IL. Attendees will learn more about HathiTrust’s collections plans, and our shared print, federal documents, and copyright review programs. The HathiTrust Research Center will present on its developing services and outreach programs, and a panel of members will relate HathiTrust’s activities to their own. Unfortunately, because of space constraints we can reserve only one spot for a representative of each HathiTrust member library. Presentations and notes from the day will be posted at <https://www.hathitrust.org/hathitrust-2016-member-meeting>.

#### Copyright Reviews Open More Titles in HathiTrust

Earlier in 2016 eleven HathiTrust members volunteered to continue review of the UK, Australian and Canadian candidate works using the Copyright Review Management System (CRMS). This has allowed us to complete reviews of these candidates, and by the end of this year we will finish with a grand total of over 290,000 volumes. The progress we’ve made in identifying public domain works is due to the collaborative and generous work of volunteers at HathiTrust institutions. Together we make a great difference in access to public domain books.

#### New Members Join HathiTrust

We are very happy to announce the newest members of HathiTrust:

- The Claremont Colleges
- DePaul University
- West Virginia University

This brings our total membership to 121.

#### Members Joining in 2016

- Amherst College
- Bucknell University
- The Claremont Colleges
- DePaul University
- Haverford College
- Tulane University
- Washington State University
- Wesleyan University
- West Virginia University
- Williams College





Starting in January 2017 the HathiTrust Copyright Review Program will recommence review of 1923-1963 U.S. publications. This is a time period when U.S. law required publications to conform with formalities such as copyright registration renewal. We have found that at least 50% of these publications never had their copyright renewed and are now public domain. As the HathiTrust collection continues to grow, there are 78,000 new works for us to review and more anticipated soon. Along with this project we also have an exciting opportunity to begin an official project to review U.S. state government documents 1923-1977. This was previously run as a small pilot with such great results (over 70% were determined to be public domain!) that we are expanding it to an official project. Improving public access to state government publications is an important area where we can make a difference.

To complete these projects we announced a call for volunteer copyright reviewers. Although the deadline for volunteering is October 15, 2016, late inquiries are welcome. More information is available on the HathiTrust webpage [Participating in Copyright Projects](#), and a recording of the webcast “Participation in HathiTrust copyright projects” is available on the HathiTrust Youtube channel. Please contact Kristina Eden at [keden@hathitrust.org](mailto:keden@hathitrust.org), 734.764.9602 for details.

### U.S. Federal Documents Program

We're pleased to announce the appointment of a new [Federal Documents Advisory Committee \(FDAC\)](#). FDAC members are: Prue Adler, Association of Research Libraries; Ivy Anderson, California Digital Library; Kristen Clark, University of Minnesota; Beth Dupuis, University of California, Berkeley; Michael Norman, University of Illinois; Judith Russell, University of Florida; Barbara Selby, University of Virginia; Heather Christenson, HathiTrust (non-voting chair). The Committee, a successor to the [Government Documents Initiative Planning and Advisory Group](#), is already in action, advising on the overall program strategy and plans for a two year horizon being developed by Program Officer Heather Christenson. HathiTrust's focused efforts regarding U.S. federal documents continue to emphasize collection development and management. HathiTrust staff are conducting an analysis to develop a “collection profile” to characterize the federal documents corpus in HathiTrust.

[The U.S. Federal Documents Registry](#) is now in beta. Mike Furlough and Valerie Glenn recently presented on the Registry work at the ALA and IFLA conferences ([read more](#)). HathiTrust Registry development staff held informal meetings with Zephir staff to share their wealth of knowledge regarding duplication detection and the nuances of handling federal documents bibliographic data.

As of Oct. 1, HathiTrust's collection included over 763,000 U.S. federal documents.

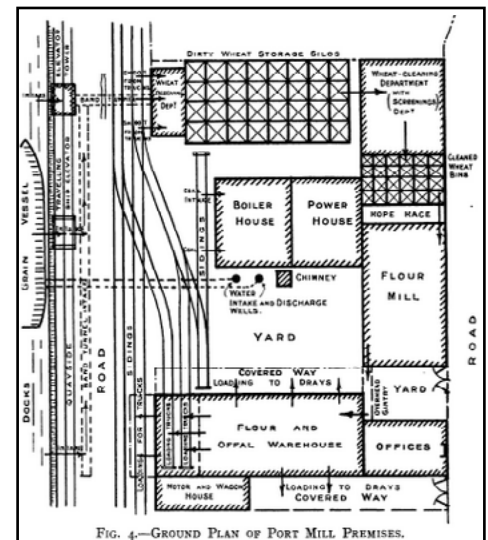


FIG. 4.—GROUND PLAN OF PORT MILL PREMISES.

*Processes of flour manufacture,  
By: Amos, Percy A., d. 1916.  
Published: (1912)*







### Shared Print Program Prepares Launch

The Shared Print Program is ready to launch, with a Shared Print Advisory Committee (SPAC) and over 50 HathiTrust partner libraries participating in the final planning and initial implementation phase. The goal of the program is to secure retention commitments for print holdings that correspond to monograph titles in HathiTrust.

During Phase 1, the participating libraries will play two primary roles:

- Serve as planning partners with SPAC to finalize the policies, business model and MOU that will govern the HathiTrust Shared Print Program, and
- Work with HathiTrust to identify which of the library's matching print monograph holdings the library will agree to retain.

At the end of Phase 1 in spring/summer 2017, the participating libraries will have a chance to review the final policies and MOU and to make a decision about going forward with the proposed retention commitments.

Phase 1 participants include libraries at universities and colleges from across the United States plus two libraries in Canada and one in Australia. For a full list of participating libraries and SPAC members, see the HathiTrust program website at [https://www.hathitrust.org/shared\\_print\\_program](https://www.hathitrust.org/shared_print_program).

### Program Steering Committee Focuses on Collections and Quality

The HathiTrust Collection Committee's [Collection Priorities Survey Analysis Report](#), with recommendations endorsed by the Program Steering Committee (PSC) and the Board of Governors, has prompted the PSC to take action on a number of major collections matters related to the focus and quality of the digital corpus.

First, the PSC revised the [Collections Committee's charge](#) in July to give increased attention to directions recommended by the report. The new charge affirms a continued focus on HathiTrust's core and distinctive collection strength of digitized "print" and charges the committee to develop a collection development strategy to enhance the comprehensiveness of prioritized collections within that scope. While HathiTrust's collection priorities will remain focused on "print" (meant broadly) for the foreseeable future, the boundary between digitized print and other text formats has become increasingly fluid. In support of possible future directions, the charge also asks the Collections Committee to initiate an exploration of expanding HathiTrust's format scope from "print" to textual materials of a variety of sorts. Finally, the charge requests that the Committee help to inform collection management reporting and analytics planning and development that HathiTrust may undertake as a service to members. This is in response to a number of identified needs and interests at the individual institutional level, and in support of collective collection strategies (e.g., Shared Print).



English gold-tooled binding in calf, inlaid with vellum. Belonged to Queen Elizabeth, and was probably bound by John Day. Gospels in Anglo-Saxon. London, 1571. *British Museum.*

*Cantor lectures on decorative bookbinding. By Cyril Davenport, 1848-1941.*





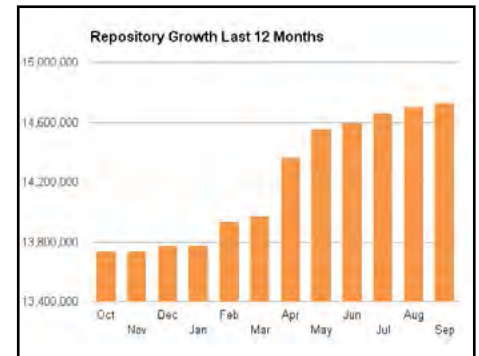
The Collection Committee’s Report also recommended that “HathiTrust should take steps to improve the quality of the existing corpus, which would include addressing scanning, processing, and metadata errors.” In response, the PSC completed its plan to form a new [Quality Assurance and Standards Working Group](#). This group, led by Paul Fogel of California Digital Library, will examine needs and recommend strategies, processes and techniques for making scalable and persistent improvements to the quality of digital facsimiles made accessible and preserved by HathiTrust. At a high level, the group will be guided by the PSC’s recent revision of the [HathiTrust Commitment to Quality statement](#).

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

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

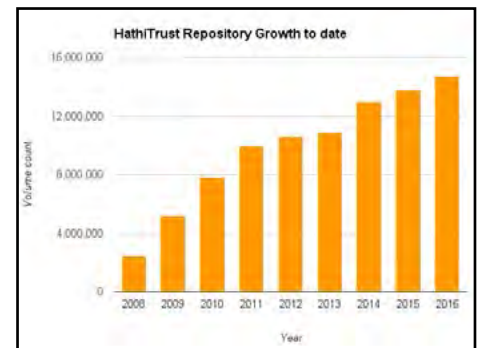
### Board of Governors Report

The Board of Governors held its Summer meeting by phone on August 31. During the meeting Board members approved the purchase of new hardware to support advanced computing at the HathiTrust Research Center. The additional capacity will advance the Research Center’s work on leading edge problems in large scale, secure text analysis, including the Data Capsule. The Board also approved the slate of nominees for the 2017 Board of Governors election, as well as the continuation of the Copyright Review Program. The Board will hold a special meeting by phone on October 19 to review and approve the proposed 2017 budget to be voted on by the membership later this year. They will also hold an in person meeting on Wednesday, November 9 at the Big Ten Center in Chicago.



### Research Center Develops Non-Consumptive Research Policy

A HathiTrust Research Center task force worked during Summer 2016 to prepare the Non-Consumptive Research Policy. The policy defines “non-consumptive research” and the levels of access permitted at various stages of this research using the HTRC Data Capsule services. Its goal is to ensure that HathiTrust facilitates the widest possible variety of non-consumptive research projects using the HathiTrust corpus, while remaining clearly within the bounds of the fair uses that courts have recognized. More generally, the policy aims to achieve the same goals as copyright itself: to promote progress in the discovery and spread of knowledge, without harming the commercial interests of authors, publishers, and other stakeholders.



The policy is nearing final publication. The task force members included:

- Eleanor Dickson, University of Illinois (Chair)
- Brandon Butler, University of Virginia
- Aaron Elkiss, University of Michigan
- Bobby Glushko, University of Western Ontario
- Robert McDonald, Indiana University
- Sandra McIntyre, HathiTrust Operations
- Leanne Mobley, Indiana University
- Naz Pantaloni, Indiana University





### Ingest and Development Updates

#### Ingest

Between July 1 and October 1, 2016, 129,914 digital items were added to the HathiTrust collection. During this period, University of Notre Dame contributed locally-digitized materials for the first time. HathiTrust also continued work with Knowledge Unlatched on a pilot project to ingest open access, born-digital PDF files (see <http://www.knowledgeunlatched.org/frequently-asked-questions/> for more information).

Work began in September on a new tool that will allow partner institutions submitting content directly with HathiTrust to more easily validate content prior to submission. Partners can follow development at [https://github.com/hathitrust/ht\\_sip\\_validator](https://github.com/hathitrust/ht_sip_validator).

#### Full Text Search

A new Solr index of HathiTrust's collection was completed and put into production in September. Application changes, to be released in October, will take advantage of the new index to provide improved limiting and faceting of search results by language. Live testing of alternative relevance-ranking algorithms using the new logging framework will also be put into production in October.

#### Storage/Infrastructure

In July, we completed a consolidation and move within the Michigan data center for all HathiTrust equipment. As part of this process, faster intra-data-center networking was deployed for the storage and web infrastructure. In September, staff replaced the servers that index full text for HathiTrust search. Both upgrades have resulted in performance improvements across services at the Michigan site. Networking improvements for the Indiana site as well are planned for the fall semester.

Work is also planned during the fall to replace the MySQL database server at each site; in addition, a third read-only database server will be deployed to support reporting and analytical needs. This work should generally improve responsiveness and reliability of HathiTrust's user-facing services.

#### HTRC Data Capsule Major System Upgrade

The HTRC Data Capsule is going through major system upgrade including co-location of additional servers from University of Illinois at Urbana-Champaign to the Indiana University Data Center. This will offer an increased total system capacity by a factor of 10 when the upgrade is completed in November 2016 and will increase the size of virtual machines available for HTRC Data Capsule use.

#### Volumes Added

Ingest numbers and Collection statistics are updated daily and can be found on our website here: [https://www.hathitrust.org/visualizations\\_deposited\\_volumes\\_current](https://www.hathitrust.org/visualizations_deposited_volumes_current)

#### Development Forecast

Continue work on a unified logging framework for HathiTrust applications.

Continue work to take fuller advantage of Shibboleth and remove isolated institution specific dependencies on Cosign.

Improved full text access for National Federation of the Blind (NFB) users

Access to initial set of EPUB files

Improvements to Copyright Review Management System in response to feedback





### Organizational Updates

#### Thank You to Departing User Support Working Group Members

HathiTrust would like to recognize and thank all of the volunteers who have served on the User Support Working Group in the past 12 months. Members of the USWG are responsible for dealing with the several thousand messages we receive annually from users and partners. These volunteers come from our partner libraries and typically commit to at least one year of service, though some have served on this group for over four years. They tend to be on call 1-2 days every two weeks and respond to all sorts of inquiries from around the world about copyright, website problems, general reference questions, membership, the collection, what HathiTrust is, and so much more. Without this group of valued members, we would not be able to meet our users' and partners' needs as well as we do.

In the past year, the following members have completed their service to this group, and we thank them for the many hours they have contributed:

- Judy Ahronheim (University of Michigan)
- Charlie Heinz (University of Minnesota)
- Michelle Henley (Ohio State University)
- Kent LaCombe (University of California, Riverside)
- Dale Larsen (University of Utah)
- Daniel Mack (University of Maryland)
- Jill Wilson (Cornell University)
- Naomi Young (University of Florida)

In October, a new group of volunteers joined the USWG, bringing the number of total members up to 22. To view the current list of members, please see the charge at [https://www.hathitrust.org/wg\\_user-support\\_charge](https://www.hathitrust.org/wg_user-support_charge).

#### All-Sites Tech Teams Meetings

The first all-sites meeting of HathiTrust technical teams involved 26 participants in Ann Arbor on September 15. Attendees included technical staff from HathiTrust central, HathiTrust Research Center at both the University of Illinois and Indiana University, University of Michigan Library, and the Zephyr team from California Digital Library.

The goals of the day were to increase communication across the institutions who contribute staff time to the development of HathiTrust systems and services, as well as to focus on technical topics of interest across multiple aspects of the joint portfolio of HathiTrust services. Longer-term goals for all-sites collaboration include translation of research outcomes into production services and articulation of a broader operations vision for the HathiTrust enterprise.

#### HathiTrust on the Road

HathiTrust staff will be attending the following events in Fall 2016. Please contact us if you wish to meet us at any of these events:

The Transformation of Academic Library Collecting: A Symposium Inspired by Dan C. Hazen, Harvard University, October 20-21 - Kristina Eden, Lizanne Payne

Internet Archive Library Leaders Forum 2016, October 26-28, San Francisco - Heather Christenson

Preservation and Archiving Special Interest Group Fall 2016 Meeting, October 26-28, New York City - Sandra McIntyre

ReCAP Partners Collection Development Forum, Princeton University, October 28 - Lizanne Payne.

DLF 2016 Forum, November 7-9, Milwaukee - Heather Christenson, Sandra McIntyre, Angelina Zaytsev

XIV International Conference on University Libraries, November 16-18, Ciudad Universitaria, México - Mike Furlough

CNI Fall 2016 Member Meeting, December 12-13, Washington, D.C. - Mike Furlough







## Update On Activities: July-September 2016

October 14, 2016

Individuals from the different tech teams became better acquainted with their respective roles and responsibilities on HathiTrust-related projects, in the context of the overall goals of the entire enterprise.

The group addressed a handful of issues for near-term resolution and raised ideas for future collaboration:

- An Operations breakout group adopted a plan for improved data synchronization and explored various options for sharing a single index for certain purposes. The group also discussed additional opportunities for leveraging work across different institutions, including Collection Builder / Workset Builder coordination. Additional topics surfaced for consideration in the future, including a need for fresh attention to user experience assessment, sharing data API work, and sharing user authentication tools.
- A Metadata breakout group explored the advantages of developing a plan for storing and querying page-level metadata and examined ways to do better matching and syncing on bibliographic records, where item-level data could be advantageous, more ways to create a better record from various sources and processes, other use cases for enrichment by Zephyr, the possibility of a data warehouse, and developing guidelines for data contributors about how to use linked-data URIs in MARC records.

Both breakout groups identified topics for follow-up leading up to an all-sites meetings in Chicago at the Big Ten Center planned for February 2017. The group as a whole discussed ideas for continuing the collaborative discussions, including additional video-conferencing sessions, webinars, sharing Slack channel(s), and developing use cases. For more details on the meeting, please see the group notes, collaboratively created, at <https://goo.gl/cX1FTU>.

### HTRC Scholarly Commons Team Adds Staff

The HathiTrust Research Center Scholarly Commons team is proud to announce the addition of new personnel, Ewa Zegler-Poleska, who has joined the team as part of the IU SOIC IDEASc Program funded by IMLS (<http://info.ils.indiana.edu/IDEASc/>). As part of the program Zegler-Poleska will work with the Scholarly Commons team on its extended outreach curriculum as part of the IMLS funded grant, “Digging Deeper, Reaching Further: Libraries Empowering Users to Mine the HathiTrust Digital Library Resources”. Pilot workshops for the grant will take place at Indiana University, the University of Illinois, Northwestern University, Lafayette College, and the University of North Carolina at Chapel Hill during October and November, 2016.



*Breakout session during All-Sites Tech Teams Meeting, September 2016.  
Clockwise from bottom left: John Weise, Rachel Vacek, Nabeela Jaffer, Sebastien Korner, Ben Howell, Boris Capitanu, Marie Ma, Tom Burton-West.*



*Breakout session during All-Sites Tech Teams Meeting, September 2016.  
Clockwise from bottom left: Tim Prettyman, Chris Powell, Josh Steverman, Bill Dueber, Jon Rothman, Tim Cole, Aaron Elkiss, Kat Hagedorn, Roger Espinosa, Martin Warin.*







## Update On Activities: July-September 2016

October 14, 2016

### Most-accessed volumes July-Sept 2016

Circular to bankers. 1833-43 INC.
Circular to bankers. no. 209-284 (1832/33)
The ordeal of Mansart, by W.E.B. Dubois
Quicksand, by Nella Larsen.
Fatigue of metals and structures, by H.J. Grover, S.A. Gordon, and L.R. Jackson
Locomotive Cyclopedia of American Practice, 1950-52: Definitions, Drawings and Illustrations of Diesel, Steam, Electric and Turbine Locomotives for Railroad, Industrial and Foreign Service; Their Parts and Equipment; Descriptions and Illustrations of Locomotive Shops and Servicing Facilities.
The Five Laws of Library Science, by S. R. Ranganathan
Annual Report of the United States Geological Survey to the Secretary of the Interior, v.20:5(1898-1899).
The London Stage, 1660-1800, pt.3 v.2.
Return to Life Through Contrology, by Joseph H. Pilates.

A summary of the determinations from HathiTrust copyright review activities in Summer 2016 is given below. See [HathiTrust Projects: Copyright Review](#) for further information.

	June-Sept.		Overall	
	Public Domain Determinations	All Determinations	Public Domain Determinations	All Determinations
CRMS-US	4,927	6,726	183,965	340,848
CRMS-World	3,522	11,458	152,526	291,833
Total	9,449	18,180	336,491	632,721

### User Support Issues

	Jun-Sept	Mar-May
<b>Content</b>	<b>387</b>	<b>127</b>
Quality	245	108
Collections	125	17
<b>Cataloging</b>	<b>380</b>	<b>155</b>
<b>Access and Use</b>	<b>346</b>	<b>346</b>
Copyright	235	156
Permissions	37	28
Takedown	0	2
Print on Demand	1	0
Inter-library loan	5	2
Full-PDF or e-copy requests	80	69
Datasets	51	8
Data Availability and APIs	2	2
Reuse of content	24	19
<b>Web applications</b>	<b>122</b>	<b>74</b>
Functionality problems	43	46
Problems with login specifically	13	6
General questions about login	1	0
Partners setting up login	0	1
Usability issues	0	0
Feature requests	5	1
<b>Partner Ingest</b>	<b>66</b>	<b>65</b>
<b>General</b>	<b>338</b>	<b>225</b>
Partnership	33	19
Miscellaneous	304	206
<b>Total</b>	<b>1777</b>	<b>1150</b>

\*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.™