HathiTrust Digital Library

Update On August Activities

Top News

New Institutions Offer Access To Orphan Works

Duke, Cornell, Emory, and Johns Hopkins universities, and the University of California announced in August that they will begin offering users at their institutions access to orphan works in HathiTrust where print copies of the works are held in their library collections. Approximately 160 orphan works candidates have been identified in HathiTrust to date by the University of Michigan in a pilot effort funded by HathiTrust. The total number of orphan works in HathiTrust is unknown, but John Wilkin, HathiTrust's Executive Director, has estimated that the total proportion of orphan works could be as high as 50% of the entire collection. The currently identified orphan works candidates are listed in a public online catalog and will be considered to be orphan works 90 days from the time of their posting if they are not claimed by copyright holders. As orphan works identification in HathiTrust moves from a pilot to production phase, it is expected that the review of volumes will be expanded to multiple institutions similar to the existing Copyright Review Management System. More information about the Orphan Works Project is available at http://www.lib.umich.edu/orphan-works. Institutions that have previously announced their intention to offer access to orphan works under the same terms as above include the University of Florida, the University of Michigan, and the University of Wisconsin-Madison.

Access For Users Who Have Print Disabilities

The Update on July 2011 Activities outlined the general framework under which access to orphan works will be provided, as well as access to orphan works and other in-copyright works in HathiTrust for users who have print disabilities. Access in both cases is contingent upon print copies of volumes being held currently or at one time by partnering libraries, and is provided to users who are authenticated via Shibboleth. The Shibboleth attribute and particular attribute values that partner institutions must use to enable access for users who have print disabilities or their proxies are available at http://www.hathitrust.org/shibboleth. These were determined by a small working group comprised of members from the University of Iowa, University of Illinois, and University of Michigan. Institutions may populate these values effective immediately to gain access for their users.

Staff at the University of Michigan are in the process of making enhancements to the HathiTrust catalog and full-text search applications that will allow users to search volumes based on the volumes' availability to themselves specifically, or to their particular institution. This work is targeted to be complete in early- to mid-October, in conjunction with the availability to partner institutions of the first orphan works.

Constitutional Convention News

Partner institutions are in the process of submitting ballot proposals on a variety of topics for consideration at the HathiTrust Constitutional Convention. Some of the

Late Breaking News

The University of Connecticut announced membership in HathiTrust in early September, and OCLC and EBSCO announced plans to integrate the HathiTrust full-text index in their discovery offerings. We are very pleased to welcome the University of Connecticut, and to be expanding users' ability to find and use materials in HathiTrust's collections.

September Forecast

- Continue work on interface enhancements to indicate availability of content on an institutional or individual basis as appropriate
- Continue work to improve mechanisms for throttling

Update On August Activities

topics include governance structure, content deposit, approval processes for partner initiatives, and a distributed strategy for archiving print monographs. The deadline for submitting proposals for the Convention is September 9, 2011. After a brief period of collation, the proposals will be posted publicly to HathiTrust's Constitutional Convention web page, where further information about the Convention is available.

HathiTrust Goes Mobile

University of Michigan staff released a beta mobile interface for searching and viewing volumes in HathiTrust: http://m.hathitrust.org/. It is currently considered to be a "soft release" for testing purposes, but is available without restrictions. Once any issues that arise are worked out, the mobile site will be publicized more broadly and mobile users who visit the regular site will be automatically redirected to the mobile version. Although it is designed for small screens, the mobile interface also works in a regular web browser. If you have any questions or would like to submit comments on the new interface, please send them using the "feedback" link on HathiTrust pages, or email Suzanne Chapman (suzchap@umich.edu).

Ingest

Local Digitization Ingest

58 of approximately 217 rare manuscripts and incunabula were ingested from Universidad Complutense de Madrid in August. Staff at Michigan and Madrid continued to work on transfer of the remaining volumes to Michigan, and their subsequent transformation for deposit. Michigan staff coordinated with staff at Northwestern University and the universities of Iowa, Pittsburgh, and Utah on ingest of locally-digitized volumes.

Utah State University Press

HathiTrust ingested 46 of approximately 300 volumes to be contributed by the Utah State University Press. The USU Press is the second university press to deposit back file publications in HathiTrust on an open access basis. In return for open access, HathiTrust is archiving these volumes free of cost.

Total Volumes Added

	August	Overall
Columbia University	41	64,042
Cornell University	12,237	357,331
Harvard University	87	52,814
Indiana University	1,252	186,139
Library of Congress	0	71,418
North Carolina State University	2,954	2,954
Northwestern University	5,184	5,184
NYPL	215	259,043
Penn State University	166	39,369
Princeton University	4,187	245,782
University of California	55,403	3,039,063
University of Chicago	1,569	8,036
University of Illinois	0	14,501
Universidad Complutense	201	108,155
University of Michigan	27,448	4,432,297
University of Minnesota	606	88,251
University of Wisconsin	9,508	497,539
University of Virginia	4	47,308
Utah State University	46	46
Yale University	0	18,385
Total	121,108	9,537,657

Public Domain (~27% of total)

	August	Overall
Total*	84,644	2,593,035

*Includes volumes opened through copyright review or rights holder permissions.

Update On August Activities

Google and Internet Archive

HathiTrust began ingest of the first Google-digitized volumes from Northwestern University in August and prepared for ingest of Google-digitized volumes from Purdue University. HathiTrust also ingested a set of nearly 3,000 volumes from North Carolina State University, digitized by the Internet Archive.

Working Groups and Committees

Collections

The Collections Committee submitted its ballot initiative for a Distributed Print Monographs Archive for the Constitutional Convention in August. A draft recommendation for the treatment of duplicates was slightly delayed by August vacations but will be shared with the Strategic Advisory Board soon for feedback and direction about next steps. The committee will turn its attention next to a process for responding to individual requests and offers to include additional materials in HathiTrust, among other pending items on its work agenda.

Communications

The Communications Working Group welcomed Stacy Kowalczyk of Indiana University as a new representative from the HathiTrust Research Center. A small team composed of members of the working group and operational staff from the Research Center met to identify specific avenues for developing the Research Center's presence on the HathiTrust website and other communication activities.

Usability

The Usability Group, now one year old, undertook a review of its progress and discussed opportunities for fine-tuning the group's mission. The result of the review was a decision, approved by the Executive Committee, to change from a "working" group to an "advisory" group. To reflect this development, the name of the group has changed to the "HathiTrust User Experience (UX) Advisory Group." As an advisory group for operational needs, the group will continue to report to the HathiTrust Director. It will also continue to manage the HathiTrust UX-SIG email group, participate in other HathiTrust committees as a liaison for UX-related issues, and advise development staff on user interface designs, development priorities, and usability priorities. The group will not play a role specifically in the implementation of usability studies or interfaces for HathiTrust applications and services.

The UX advisory group continued to review and track feedback received via the User Support Group to help discover issues related to usability. Both the UX Advisory Group and the UX-SIG group were given an opportunity to provide early feedback on the HathiTrust mobile beta interface.

Update On August Activities

User Support

The adjacent table is a summary of the issues received by the User Support Working Group in August.

Projects

IMLS Quality Grant

Review of the first sample of 1,000 randomly selected volumes from HathiTrust continued in August. As of late August, over half the sample had been coded by reviewers at the University of Michigan and University of Minnesota, amounting to 54,635 out of 100,000 total pages (systematic sampling is being used to select 100 pages within each volume).

In August, the project team also hired a team leader to coordinate logistics for conducting physical review of all of the volumes in the 1,000-volume sample. The team developed and finalized a short survey that will capture certain physical characteristics of the volumes and confirm bibliographic data. An interface for entering data from the survey is under development, and physical review of the original volumes is expected to commence in mid-September. The grant team also worked to finalize sampling parameters for the second production sample of volumes, which is set to begin in late September. The grant Advisory Board met in late August to provide feedback on the team's progress to date and guidance on future directions. Additional information on the project can be found at http://www.hathitrust.org/grants.

HTPub

The University of Michigan has hired a full-time software developer for the HTPub project. The developer will be focusing on the content ingest and publication management components of the HTPub system, and will begin work on September 19.

HathiTrust Research Center

The HathiTrust Research Center (HTRC) has heard positive responses on about 70% of invitations it sent out to form an HTRC advisory board. The advisory board is anticipated to guide the HTRC in setting resource allocation policies, being good stewards of the research data and outputs, and securing additional resources to make HTRC a stable entity for years to come. HTRC has been in discussions with HathiTrust to develop an integrated web presence, and is working on a draft of

User Support Issues

	August	July
Content	**110**	**90**
Quality	96	89
Non-partner Digital Deposit	3	1
Collections	8	2
Cataloging	**26**	**20**
Access and Use	**111**	**81**
Copyright	58	52
Permissions	23	2
Takedown	2	0
Print on Demand	6	36
Inter-library loan	0	9
Full-PDF or e-copy requests	14	13
Datasets	1	0
Data Availability and APIs	1	1
Reuse of content	7	2
Web applications	**27**	**23**
Functionality problems	5	7
Problems with login specifically	1	6
General Questions about login	3	4
Partners setting up login	4	6
Usability issues	11	3
Feature requests	7	8
Partner Ingest	**2**	**2**
General	**59**	**23**
Partnership	13	6
Infrastructure	1	0
Miscellaneous	45	17

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

HathiTrust Digital Library

Update On August Activities

policies for access to and use of HTRC resources.

Development Updates

Bibliographic Data Management

The California Digital Library development team continues to make improvements to the core metadata management system and work on issues related to integration with HathiTrust systems. The team is also preparing a production virtual environment for performance and load testing, and has started work on a migration verification strategy that will use Z39.50 to compare bibliographic data that has been loaded from the University of Michigan with the same data in its pre-load state. The newly-hired metadata analyst for the project will start on September 12, 2011. Finally, the team has given a name to the HathiTrust metadata management system software - Zephir.

Data API

Preparations to allow access for users who have print disabilities and access to orphan works took precedence over the ongoing HathiTrust Data API security work in August, though this work remains a high priority. The API enhancements are described at http://bit.ly/jozHQK. Interested parties are invited to submit comments and feedback to feedback@issues.hathitrust.org.

Full-text Search

Staff at Michigan and the California Digital Library (CDL) continued to make progress on the full-text search tasks identified as high priority by the HathiTrust Full-Text Working Group. Michigan successfully replaced the XPat search engine, which had been used since the launch of HathiTrust for searching within a book, with Solr/Lucene. This move has improved the order in which results are displayed for multi-word searches in "within book search," the group's third highest priority. CDL's work on a spelling suggestion feature continued as well.

PageTurner

Michigan staff made progress on enabling application-level throttling in HathiTrust applications and will continue this work in September. A proof of concept was implemented and shows promise. Application-level throttling will give HathiTrust finer-grained control over when and when not to throttle user access (block access for short periods of time). This will allow HathiTrust to maintain compliance with third party agreements on content and provide a consistent experience for all users, while offering fewer interruptions to routine activities such as browsing thumbnails of content or scrolling quickly through a volume.

Outages

HathiTrust may have been inaccessible for some users from approximately

HathiTrust Digital Library

Update On August Activities

5:00pm - 5:20pm EDT on Tuesday, August 16 due to network connectivity problems at the Indianapolis site. The problems were intermittent, preventing normal failover mechanisms from triggering.

HathiTrust sends notice upon discovery and resolution of unscheduled outages and in advance of scheduled outages and maintenance work that may result in an outage. We welcome and encourage additional recipients for these notices. If your institution is not receiving outage notifications and would like to, please contact feedback@issues.hathitrust.org.

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

There's an elephant in the library.™

www.hathitrust.org