

HathiTrust 2017

Member Meeting

Charting our Strategic Directions

NOVEMBER 2, 2017



Welcome!

Wendy Lougee

University Librarian and Dean of Libraries, McKnight
Presidential Professor University of Minnesota

Chair, HathiTrust Board of Governors



Housekeeping

- Space orientation
- How to connect
- How to tweet: #HTMM17
- Who can help you
- What we are doing and when



NOTES

<https://goo.gl/TssEm1>



Agenda

- 9:00 Welcome and Introductions
Wendy Lougee, University of Minnesota and
Chair, HathiTrust Board of Governors
Mike Furlough, Executive Director, HathiTrust
- 9:30 Lightning Talks: Tell your HathiTrust story!
- 10:00 HathiTrust Budget Planning
John Culshaw, University of Iowa and Incoming Chair, Treasurer,
HathiTrust Board of Governors
Mike Furlough
- 10:15 BREAK (15 minutes)



Agenda continued

- 10:30 Introduction and Discussion: Vision, Organizational Values, and Themes
- 11:15 Discussion: Strategic Area 1 Impact
- 12:00 Lunch Break
- 1:00 Discussion: Strategic Area 2 Enhance
- 1:45 Break
- 2:00 Discussion: Strategic Area 3 Grow
- 2:45 Summary and wrap-up



Introduce yourselves

NEXT: Lightning Talks



Lightning Talks: Tell your HathiTrust Story

WHAT HAS HATHITRUST ENABLED YOUR LIBRARY AND/OR YOUR USERS TO DO?

WHAT CHALLENGE IS YOUR LIBRARY FACING THAT HATHITRUST COULD ADDRESS?



Lightning Talks: Tell your HathiTrust story

What can HathiTrust tell us about our profession's history?

Eric Novotny, Penn State University

Direct from private hands to the wide world: HathiTrust enables new discovery of old resources.

Sarah Pritchard, Northwestern University

Linking HathiTrust Resources to User Needs at University Archives.

Nicholas Wolf, New York University

Mission? Permission! Adventures in opening volumes to full-view at scale.

Michelle Paolillo, Cornell University

Opening Volumes at Scale.

Paul Fogel, California Digital Library

The Power of the Collective: Shared Print.

Rebecca L. Lubas, Claremont Colleges Library

HathiTrust Digital Surrogates: qualifying manifestations?

John Butler, University of Minnesota

Metadata Enhancement for HathiTrust.

Judy Russell, University of Florida



Using Hathi to Analyze Early Library Science Journals

Eric Novotny, ecn1@psu.edu



**Annual report of the Trustees of the State
Library. 57 (1875)**

by New York State Library.
Published 1875

[Catalog Record](#) [Full view](#)

In my collections:

Pre-1923 Library
Journals



Library journal. v.1 (1876-77)

by American Library Association. Library Association.
Published 1877

[Catalog Record](#) [Full view](#)

In my collections:

Pre-1923 Library
Journals



**Transactions and proceedings of the
Conference of librarians held in London
October 1877.**

by Conference of Librarians (London)
Published 1878

[Catalog Record](#) [Full view](#)

In my collections:

Pre-1923 Library
Journals



Library journal. vol. 2 (1877-78)

by Library Association. American Library Association.
Published 1878

[Catalog Record](#) [Full view](#)

In my collections:

Pre-1923 Library
Journals



Library journal. v.3 1878 Mar-Dec

In my collections:

From Private Hands to the Wide World:

HathiTrust Enables New Discovery of Old Sources

- ❖ *Le Livre nommé la Discipline d'amour divine*
- ❖ French, decorated manuscript on parchment
- ❖ One of only two known mss. copies; unpublished; the other one is not publicly accessible.
- ❖ Acquired 2015; **in private hands since ca. 1500**
- ❖ Workflow: conservation, cataloging, digitizing
- ❖ Now usable by anyone, anywhere!



Similar Items

[Le livre nommé Le Canarien : textes français de la conquête des Canaries au XVe siècle /](#)

By: Aznar Vallejo, Eduardo.
Published: (2008)

[Le quart livre /](#)

By: Rabelais, François, ca. 1490-1553?
Published: (1993)

[Le rousier des dames, siue, Le pelerin damours /](#)

By: Desmarins de Masan, Bertrand.
Published: (1852)

[Pantagruel : le quart livre /](#)

By: Rabelais, François, ca. 1490-1553?
Published: (1961)

[Le livre de paix /](#)

By: Chastellain, Georges, 1405?-1475.
Published: (2006)

[Les grandes et inestimables cronicques : du grant et enorme geant Gargantua : contenant sa genealogie, la grâdeur et force de son corps. Aussi les merveillieux faictz darmes quil fist pour le roy Artus, côme verrez cy apres.](#)

By: Rabelais, François, ca. 1490-1553?
Published: (1910)

[Cite this](#)

[Export citation file](#)

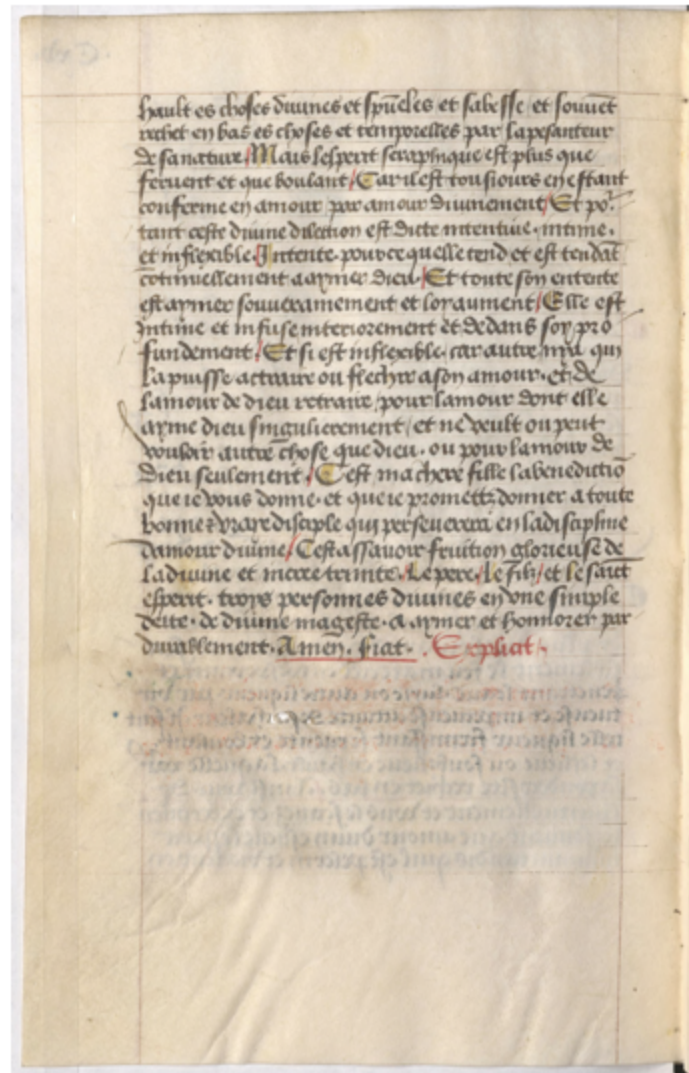
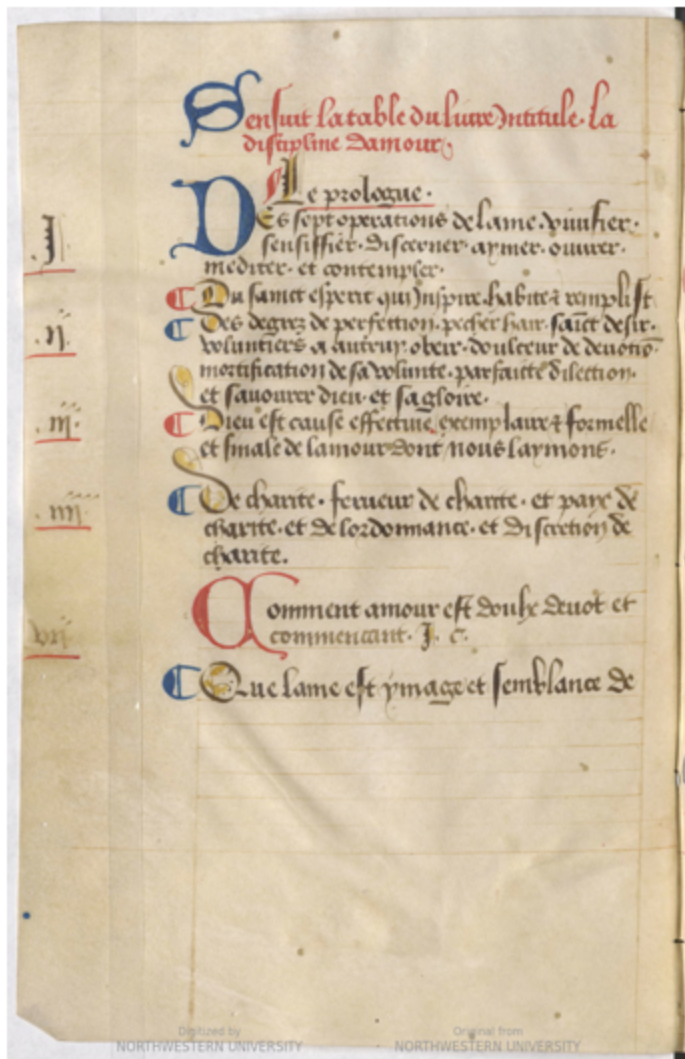
Cy commence le livre nom[m]e la discipline damour divine.

Language(s):	French, Middle (ca. 1300-1600)
Published:	[between 1480 and 1500]
Subjects:	Manuscripts, French > Illinois > Evanston . Illumination of books and manuscripts > Specimens . Devotional literature, French > 15th century .
Note:	Initials in blue, red and yellow; table of contents on vellum; ruled. Published as: Le livre de la discipline d'amour divine. Paris: Regnault/Chaudière, 1519. Leaves 8, 124-128 blank except for ruling. Mss. (transcripts, handwritten) Owned by: Count Justin MacCarthy-Reagh (1744-1811); bound for him by Richard Wier, 1792; Sir Thomas Phillipps (1792-1872); Joost R. Ritman (b. 1941-).
Physical Description:	8 unnumbered leaves, vjxx leaves (that is, 120), (29 lines), bound : illuminations ; 22 cm
Locate a Print Version:	Find in a library

Viewability:

[Full view](#)

(original from Northwestern University)



A dialogue between Love, Reason, and the Soul

Linking HathiTrust Resources to User Needs at University Archives

Janet Bunde, University Archives
Nicholas Wolf, Data Services

nicholas.wolf@nyu.edu; orcid.org/0000-0001-5512-6151



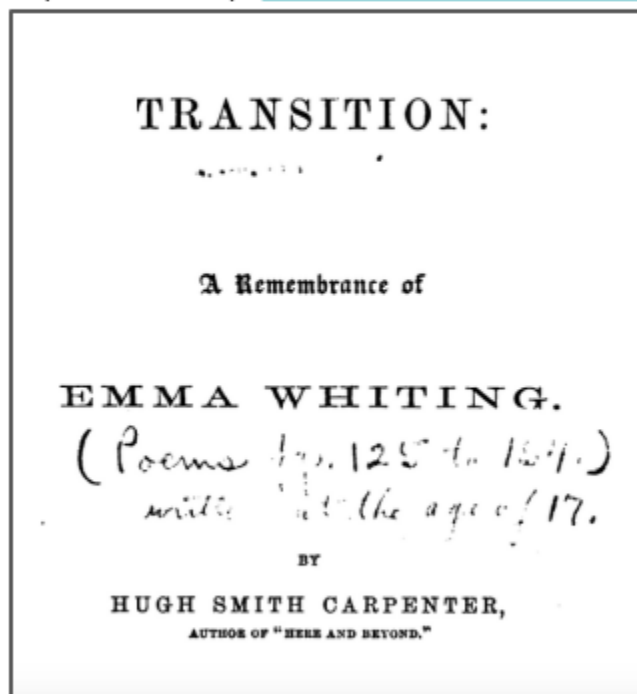
Goals and Intentions

- Track NYU alumni who appear in the textual record
- Facilitate compiling of information to fulfill patron requests for information about past students
- Explore integration of Hathi content with University Archives' digital presence and collection highlights
- Track intersection of NYU as a university and the textual record



Goals and Intentions

Example: Hugh Smith Carpenter (1824-1899), [*Transition: A Remembrance of Emma Whiting*](#) (1863)



Operationalization

- Robert Marchman (rlmv) [Python wrapper](#) for Hathi Bibliographic, SOLR, and Data APIs.
 - Search author and title fields in Bibliographic API
 - Search full-text field of SOLR API wrapper (need to regroup, discontinued SOLR API)
- Search on person, “University of the City of New-York” (and variations), place
- Limit results examined to first 500; average number of results/term ~ 5 (person) and ~ 11,000 (NYU)
- Workflow:
 - Retrieve results matching all terms in author, title, and full-text fields, compile doc IDs
 - Retrieve all OCRd plain text from Data API using doc IDs.
 - Search for instances of terms in full-text context; capture ~ +/- 15 word context



Next Steps

- Broaden terms to include keywords “catalogues” and “directories” so as to track down alumni listings
- Capture author ids to facilitate complete author bibliography list
- Expand search fields to include “publisher”
- Limit full-text search results to close-proximity keyword terms (adjacent terms)
- Track NYU-related publications outside of NYU by holding library; map locations of holdings
- Successor to SOLR API?

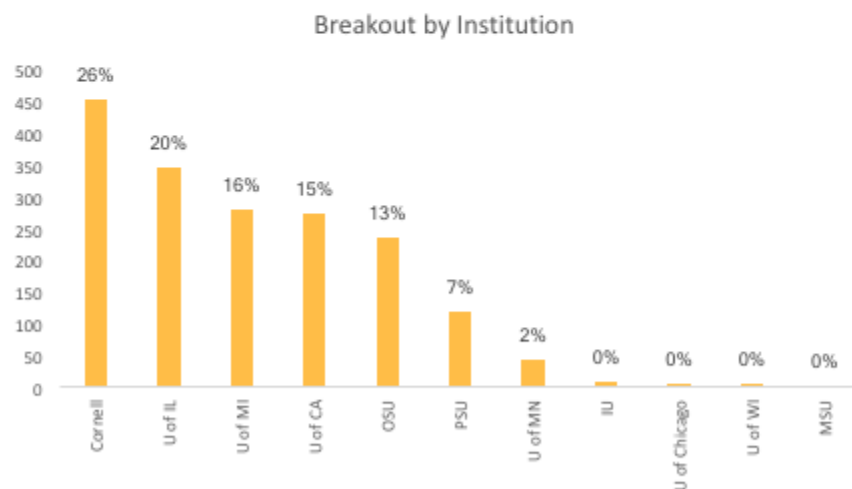




Mission? Permission!

Adventures at Cornell in opening volumes to full view at scale

1.7K volumes of Cornell publications released to full view



“Partnering at scale”

- Share institutional knowledge
 - only 26% affected volumes are Cornell deposits
- Identify CUL vols (6K)
- Flexibility of process
- Documented process available on request



Michelle Paolillo, Cornell University– map6@cornell.edu

HathiTrust Annual Meeting 11/2/2017



Opening Volumes at Scale

PAUL FOGEL, CALIFORNIA DIGITAL LIBRARY

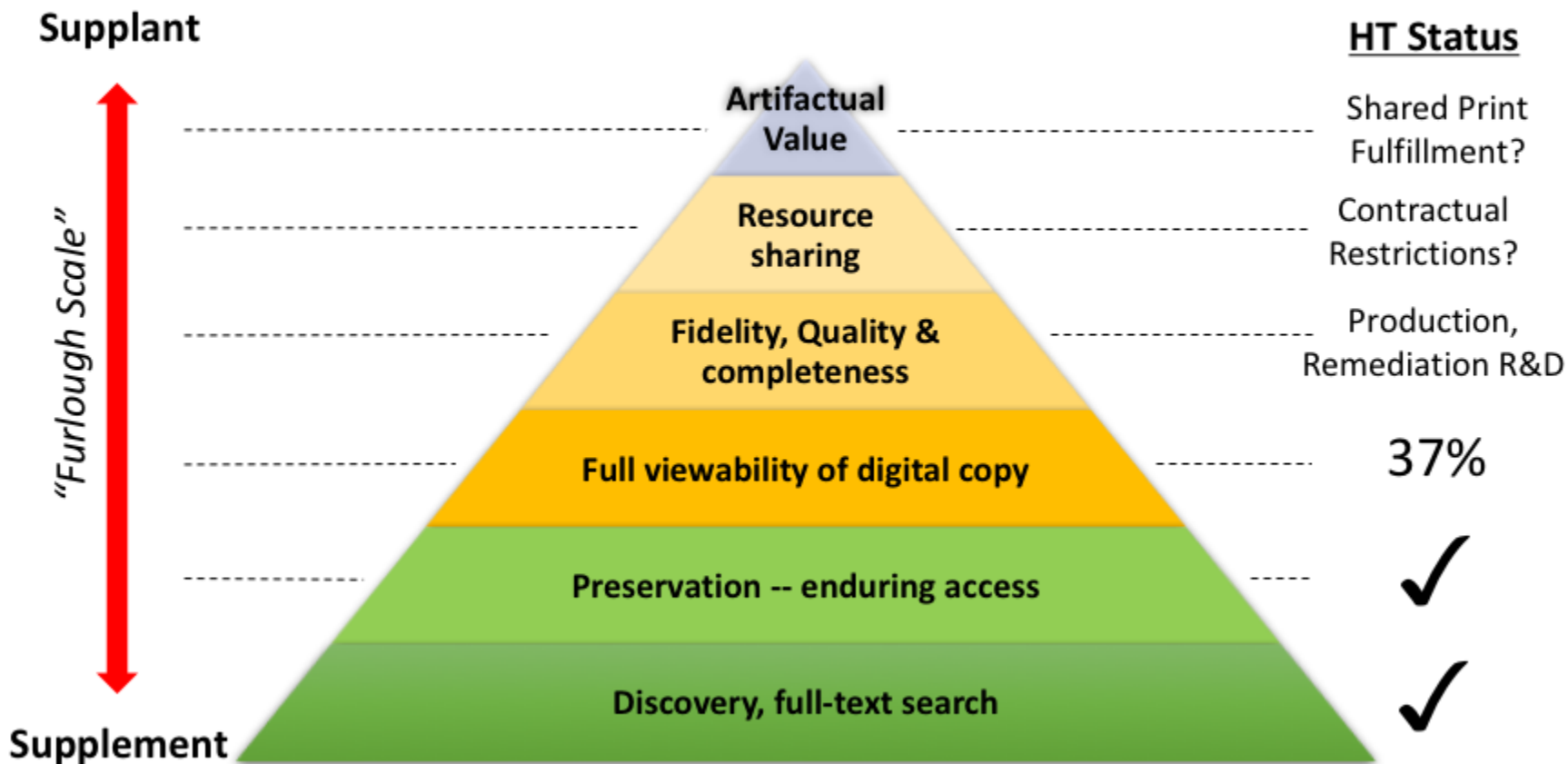


Claremont Colleges Library: Shared Print is Active Print

- Vision: transform the library from a warehouse of books to collections that inspire interaction
- Joining Shared Print is an important part of the strategy
- CCL joined two Shared Print at almost the same time: HathiTrust and SCELC (mostly CA private institutions)
- CCL contributed more to SCELC because we double-committed
- Power of the Collective: developments with OCLC Shared Print services

HathiTrust Digital Surrogates: qualifying manifestations?

-- John Butler



Florida Thesis Project

Sample records from the pilot project clearly demonstrate the enhanced metadata:

Subjects	
Subjects / Keywords:	alligator, detectability, estimates, habitat, population, sightability, survey Interdisciplinary Ecology -- Dissertations, Academic -- UF
Genre:	Electronic Thesis or Dissertation bibliography (marcgt) theses (marcgt) Interdisciplinary Ecology thesis, M.S.

Original Record



Enhanced Record



Subjects	
Subjects / Keywords:	alligator, detectability, estimates, habitat, population, sightability, survey Interdisciplinary Ecology -- Dissertations, Academic -- UF Alligators (JSTOR) Population estimates (JSTOR) Aquatic habitats (JSTOR) Wet prairies (JSTOR) Vegetation (JSTOR) Night lights (JSTOR) Water depth (JSTOR) Area Surveys (JSTOR) Florida -- Everglades Florida -- Kissimmee Florida -- Franklin County Florida -- Taylor County
Genre:	Electronic Thesis or Dissertation bibliography (marcgt) theses (marcgt) Interdisciplinary Ecology thesis, M.S.



Lightning Questions



Financial Planning and 2018 Budget Report

**JOHN CULSHAW, UNIVERSITY OF IOWA AND TREASURER/
INCOMING CHAIR OF BOARD OF GOVERNORS**

MIKE FURLOUGH, HATHITRUST



2018 Budget Approval

- Annual budget must be approved by membership before we can invoice for 2018.
- Approval requires a simple majority of weighted votes.
- We will not vote today.
- Electronic ballot has been sent to all member representatives.
- Voting closes at 11:59 ET (US) Nov 15, 2017.



Budget Structure

Revenue

Partner fees
Interest/investment income

Expenses

Operations

- Administration
 - Staff, office, travel, meetings, communications, copyright review
- Infrastructure
 - Equipment, data centers, backup, networking, staff
- Contract Services
 - Zephyr (CDL)
 - Administration, hosting (Michigan)

Programmatic Activities

- Federal Documents Initiative
- Print Monographs Archive
- HathiTrust Research Center
 - (Indiana and Illinois)
- Print disability services assessment/expansion

Reserve

Infrastructure life cycling and contingencies



Fee Model & 2018 Factors

- Determine total operations and infrastructure costs.
- All partners share costs for public domain volumes:
 $(Public\ Domain * Cost\ per\ volume * X) / Total\ Members$
- Members share costs for in-copyright volumes based on holdings
 - For a given in-copyright volume:
 $(Cost\ per\ volume * X) / Holding\ Members$
- For 2018:
 - C = \$0.1758 per volume per year (base only)
 - X = 1.2 (multiplier for programmatic funds)
 - 130 members (including 10 new)



2018 Fee Analysis

- Total fees collected increased by +3.84%
- Individual member fee increases are not uniform, and generally decreased.
 - Recurring costs for 2018 did not increase significantly.
 - A steady growth in the number of volumes has resulted in a lower operating cost per volume
 - Our membership has increased by a total of ten institutions.
- Average fee change: -3.40%
- Median fee change: -4.74%
- High: 58.15%*
- Low: -12.33%

*NOTE: A few members saw greater change due to corrected holdings data and/or deposits.



Proposed 2018 HathiTrust Budget (for approval)

	BUDGET	ACTIVITY	2017 PLANNED	2018 PLANNED
A	FUNDING SOURCES			
B		Current Member Fees	\$3,212,125	\$3,335,596
C		Investment Income	\$40,000	\$40,000
D		New member fees (pro-rated)	\$125,000	\$75,000
E		TOTAL REVENUE	\$3,377,125	\$3,450,596
F	RESERVE FUNDS			
G		Total Allocations/Transfers	\$943,997	\$633,528
H		Total Expenses	\$-	\$(180,000)
I		NET CHANGE	\$943,997	\$288,528
J	OPERATIONS			
K		Total Allocations	\$2,158,243	\$2,261,135
L		Total Expenses/Transfers	\$(2,418,712)	\$(2,361,135)
M		NET CHANGE	\$(260,469)	\$(100,000)
N	PROGRAMS			
O		Total Allocations	\$535,354	\$555,933
P		Total Expenses/Transfers	\$(1,443,905)	\$(1,114,904)
Q		NET CHANGE	\$(908,551)	\$(558,971)
R				
S	TOTALS			
T		Revenue (all sources)	\$3,377,125	\$3,450,596
U		Allocations/Transfers	\$3,637,594	\$3,450,596
V		All expenses	\$(3,862,617)	\$(3,656,039)
W		NET CHANGE	\$(225,023)	\$(205,443)



Board Review of the Budget and Fee Model

Current fee model proposed 2011, implemented 2013

- Value-based and values-based
- Complex and costly
- Intricately linked to concept of shared/collective collections

2017 Membership and Finance Committee recommended:

- Ensuring a diverse, robust membership
- Look to simplify the process and the model
- Address year-to-year volatility in member fees
- Explore tiered fees for public domain costs
- Develop multi-year budgeting models



Status of the review

Consultant has worked with us since spring

Board has reviewed recommendations

Elements considered (in no order)

- Introduction of tiers in the model
- Extensive reliance upon holdings in the fee model
 - Frequency of the holdings analysis
 - Can other data be used?
- Length of phase-in, if adjustments are made

Active discussion continues



Assumptions and Timeline

Vision and strategy should drive the budget.

Our work today is laying ground for 2019+ budgets.

2018:

- Conclude planning process
- Finalize any recommended changes to budget/fee model
- Present to membership for feedback and consent
- Any changes would not go into effect until 2019 budget year



Morning Break: 15 Minutes

NEXT: DISCUSSION OF PROPOSED VISION AND THEMES



Charting HathiTrust's Strategic Directions

HATHITRUST MEMBER MEETING
2 NOVEMBER 2017



Mission and Purpose

To contribute to research, scholarship, and the common good by collaboratively collecting, organizing, preserving, communicating, and sharing the record of human knowledge.

A trusted digital preservation service enabling the broadest possible access worldwide.

An organization with over 130 research libraries partnering to develop its programs.

A range of transformative programs enabled by working at a very large scale.



Six stages of HathiTrust

2002-2006: Mass digitization programs begin

2006-2008: Moving towards launch

2008-2011: Rapid start-up and organizing

2012-14: New governance and leadership

2014-17: Settling in and starting up

Now: Taking stock and looking out



What is different now?

The World – because we changed it.

HathiTrust

- Membership diversification
- Organizational maturity (but still adolescent?)
- Governance is addressing a wider range of challenges

Ecosystem

- Mass digitization is assumed and non-controversial
- Legal challenges have ended, but questions remain
- Cooperation and collaboration at scale is proven but still hard
- Data management is a generalized problem



The Planning Process and Timeline

Collections Priorities Survey (2015)

Planning to plan (2016)

Board/Program Steering Committee Retreat (June 2017)

Member outreach and engagement (August-October)

Member Meeting (November 2 2017)

Wrap up and finalization (November - March)



What do we hope to create?

A strategic vision that tells a compelling story

- Who are we?
- What do we value?
- How do we put that into action?

Practically

- 3 key strategic “themes” where we will make a difference
- To help us align and guide our work
- For 3-5 years
- Developed with membership engagement and endorsement.



2015 Collections Survey: Findings

Don't mess up what you already do well!

Keep building the collections and do so faster.

No strong impetus to expand collecting focus.

Quality is important!

DECISION: Keep collection focus on "books" and "book-like" materials for several years; subject to revision.



Board of Governors & Program Steering Committee Retreat June 2017

Ivy Anderson, California Digital Library

John Butler, University of Minnesota

John Culshaw, University of Iowa

Mike Furlough, HathiTrust

Todd Grappone, University of California, Los Angeles

James Hilton, University of Michigan

Martha Hruska, University of California, San Diego

Anne Kenney, Cornell University

Wendy Lougee, University of Minnesota

Joe Lucia, Temple University

Robert McDonald, Indiana University

Beth McNeil, Iowa State University

Erik Mitchell, University of California, Berkeley

Beth Sandore Namachchivaya, University of Illinois at Urbana-Champaign

Oya Rieger, Cornell University

Brian Schottlaender, University of California, San Diego

Kevin Smith, University of Kansas

Virginia Steel, University of California Los Angeles

Karla Strieb, The Ohio State University

Winston Tabb, Johns Hopkins University

Sarah Thomas, Harvard University

Günter Waibel, California Digital Library

Carolyn Walters, Indiana University

Betsy Wilson, University of Washington



Feedback: Small Group Discussions

Recognized alignment between vision and their library.

High value from membership in HathiTrust – scalability.

Shared Print most cited as potentially transformative.

Research center potentially valuable but not fully realized.

Desire to see collections expand, especially international materials.

Need to do more to get the HathiTrust story out to our users, our librarians, key administrators.

Some current services don't scale well and could hold us back.

Uncertainty about the membership strategy.



DRAFT Vision, Themes, Objectives

THIS IS NOT THE FINAL PLAN

Subject to revisions based on member feedback.

Reflects core discussions among Board and Program Steering Committee.

Consider objectives and action items as illustrative, not formal decisions.



Organizational Values

Collaboration to support

- Public good
- Transformation at scale (enhancing the ability of individual institutions to respond to change)

Participation in governance and operations

Transparency in governance

Effectiveness (getting things done)

Strategic risk-taking and innovation

Trustworthiness

Flexibility and responsiveness

Quality



Vision Statement

HathiTrust: Collaborating to advance research, teaching, and learning using the digital record of human knowledge.

HathiTrust enables contemporary and emerging forms of research, teaching, and learning through collaboration that *transforms* local library capacity, adding value and *at scale* opportunities for innovative action.



Questions to consider....

Do you recognize how your library figures into this vision for the future?

Does this new direction represent where HathiTrust should be headed in the next 3-5 years?

Is anything missing?

Is this ambitious enough? Too ambitious?



Room Discussion



Themes

Impact

Grow

Enhance



Discussion Questions

(35 minutes at your table)

Does the summary accurately describe the scope of the strategic area?

Is anything missing? What is here that does not belong or is not strategically important?

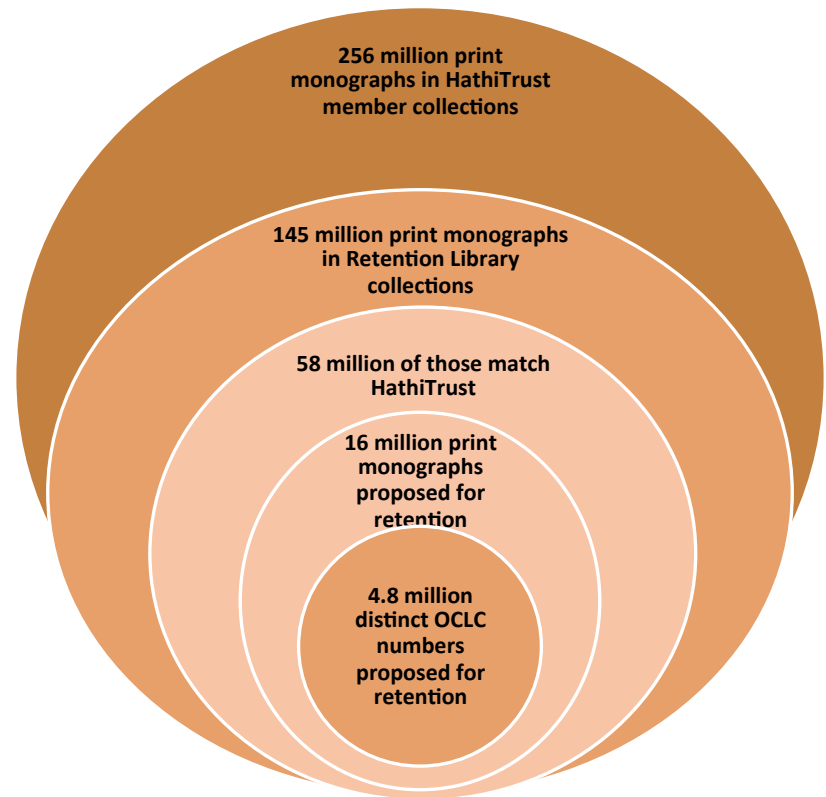
What are the priorities among the objectives?

NOTES: <https://goo.gl/TssEm1>



Impact

Generate impact by transforming infrastructure and services for the preservation and use of digitized collections, and creating a framework for sustaining and enhancing the record of scholarship at scale.



Impact: Possible Objectives

Engage potential user communities in **using, enhancing, and curating HathiTrust resources.**

Build on HathiTrust Research Center's current efforts and develop sustainable programs to support **new modes of inquiry and national research infrastructure needs.**

Cultivate **strategic partnerships with other organizations** (e.g., CRL, DPLA, DLF) to leverage HathiTrust's position to strengthen the information ecosystem.

Provide robust centralized and coordinated (e.g., print retention) services for preservation and access to HathiTrust collections to enable **transformation of local library infrastructure.**



Lunch until 1pm

FEEL FREE TO MINGLE

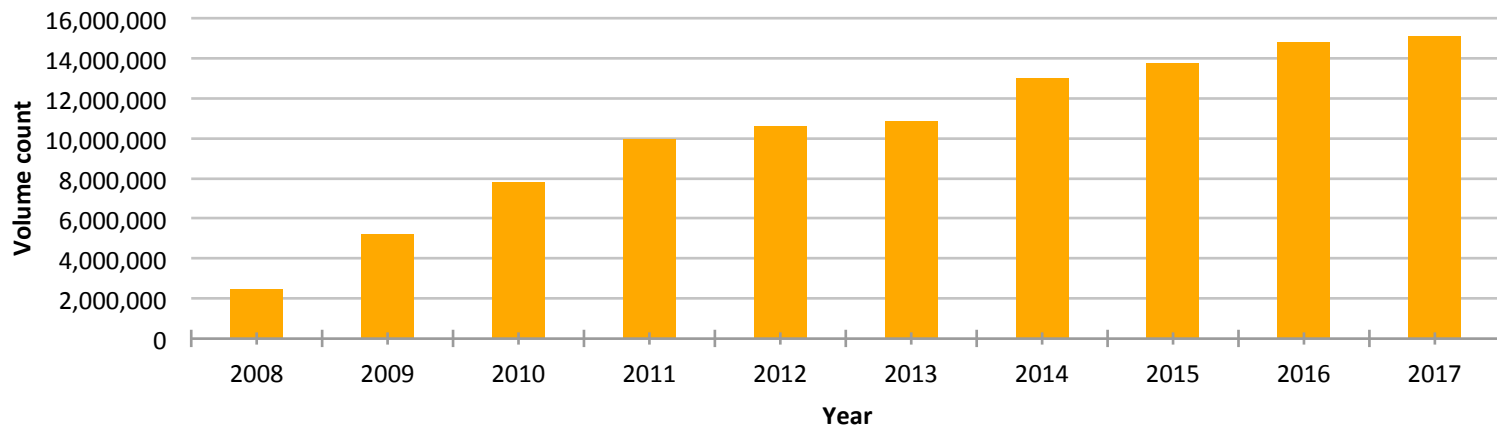
BUT PLEASE RETURN TO YOUR ORIGINAL TABLE



Grow

Expand the collection, increase membership, and enlarge the user base to enable expanded services and increase HathiTrust's impact on research, teaching, and learning

HathiTrust Repository Growth to Date



Grow: Possible Objectives

Grow the corpus to make more resources available and drive development of **more intentional digitization strategies** in LAM communities.

Increase membership to enable collection growth, make member benefits more widely available, and increase capacity through shared resources or expertise.

Develop services, programs, and tools to **market HathiTrust to end users at member campuses**.

Cultivate a community of researchers who are eager to expand conventional uses of the collection.



Discussion Questions (35 minutes at your table)

Does the summary accurately describe the scope of the strategic area?

Is anything missing? What is here that does not belong or is not strategically important?

What are the priorities among the objectives?

NOTES: <https://goo.gl/TssEm1>



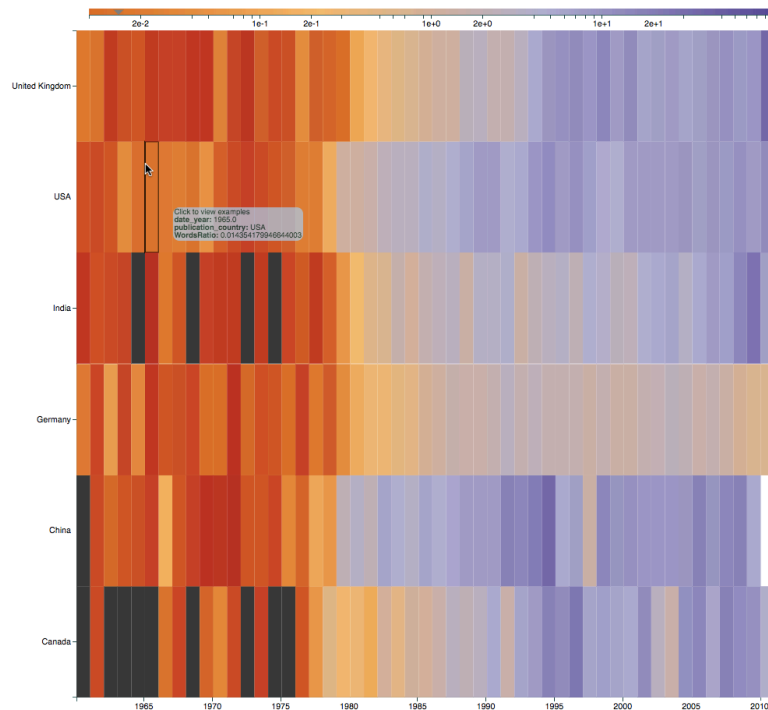
Afternoon Break: 15 Minutes

NEXT: DISCUSSION OF FINAL THEME AND WRAP UP



Enhance

Add value for members and the public by improving access to the collection, improving the quality of the images, text, and metadata, and enhancing the overall user experience. These enhancements are needed to attract and retain the level of use required to create the type of impact HathiTrust envisions.



Enhance: Possible Objectives

Investigate and implement methods to **improve content and metadata quality** to improve user experience and discoverability of open content.

Develop interfaces that provide for **different modes of discovery**, e.g. allow users to visualize or analyze the content of the collection and specific items within it.

Enhance **services for print disabled users**.

Engage researcher communities to **enhance overall usability of and access to HathiTrust collections**.

Continue to emphasize copyright review as an important strategic HathiTrust activity.



Discussion Questions (35 minutes at your table)

Does the summary accurately describe the scope of the strategic area?

Is anything missing? What is here that does not belong or is not strategically important?

What are the priorities among the objectives?

NOTES: <https://goo.gl/TssEm1>



Drawing to a close



Reflections



What is next?

Within the next day or so:

- Evaluation form
- Links to notes and slides

Over the next few weeks:

- Review and processing of today's discussions
- Start to adjust vision and themes as needed

Over the next couple of months:

- Complete final drafts of vision, themes
- Begin prioritizing objectives and plot out implementation
- Confirm plans for budget/fee model

By late winter:

- Report back to you all



Thank you!

GOODBYE!

