2018 Member Meeting

October 30, 2018

Welcome

John Culshaw, University of Iowa
Chair, HathiTrust Board of Governors

Agenda: Morning

9:00am Welcome | Logistics

John Culshaw, University of Iowa and Chair, HathiTrust Board of Governors

9:10am Report to Members

Mike Furlough, Executive Director

9:40am 2019 Budget Presentation

Günter Waibel, California Digital Library and Treasurer, HathiTrust Board of Governors

9:55am Membership Criteria and Fee Model Presentation w/ Q&A

Wendy Lougee, University of Minnesota and Past Chair, HathiTrust Board of Governors

10:25am BREAK (15 min)

Agenda: Late Morning

10:40am HathiTrust Research Center (HTRC) Updates

John Walsh, Indiana University and Director, HathiTrust Research Center
J. Stephen Downie, University of Illinois, Co-Director, HathiTrust Research Center

11:10am "Ask Me Anything About HathiTrust"

Jessica Rohr, Member Engagement Specialist
HathiTrust Staff

11:55am Preview of Afternoon

Günter Waibel

12:00pm LUNCH

Agenda: Afternoon

1:00pm Focus: Our Users and Discovery (Table Discussions)
Jessica Rohr

2:30pm BREAK (15 min)

2:45pm Member Lightning Talks
Karla Strieb, Ohio State University and Chair Program Steering Committee
Michelle Paolillo, Cornell University
Marie Concannon, University of Missouri
William Schlaack and Angela Waarala, University of Illinois at Urbana-Champaign
John Shanahan and Ana Lucic, DePaul University

3:15pm Recap of Day and Future Meetings
Mike Furlough

3:30pm Adjourn

Board of Governors

Through 2018
Kenning Arlitsch, Montana State University
Yolanda Cooper, Emory University
John Culshaw, University of Iowa (Chair)
James Hilton, University of Michigan
Wendy Lougee, University of Minnesota (Past Chair)
Beth McNeil, Iowa State University*
Kevin Smith, University of Kansas
Virginia Steel, University of California, Los Angeles
Winston Tabb, Johns Hopkins University*
Sarah Thomas, Harvard University*
Günter Waibel, California Digital Library (Incoming Chair/Treasurer)
Carolyn Walters, Indiana University

Newly Elected Members:
Mimi Calter, Stanford University
Robert McDonald, University of Colorado Boulder
Holly Mercer, University of Tennessee, Knoxville

*Service ends December 31, 2018

Program Steering Committee

Daniel Dollar, Yale University
Maria Estorino, University of North Carolina at Chapel Hill
Sharon Farb, University of California, Los Angeles
Joseph Hafner, McGill University
Anne Houston, Lafayette College
Rebecca Lubas, Claremont Colleges Library
Holly Mercer, University of Tennessee, Knoxville
Sandra McIntyre, HathiTrust
Sarah Shreeves, University of Arizona
Michael Stoller, New York University
Karla Strieb, Ohio State University (Chair)
Evviva Weinraub, Northwestern University

Committees and Working Groups

Collections Committee
Heather Christenson, HathiTrust
Marion Frank-Wilson, Indiana University
Denise Hibay, New York Public Library
Jeff Kosokoff, Duke University
Michael Neubert, Library of Congress
JoAnne Newyear-Ramirez, UC Berkeley
Carmelita Pickett, University of Virginia
Kathryn Stine, California Digital Library
Tom Teper, University of Illinois
Wade Wyckoff, McMaster University
Nicholas Wolf, New York University (Chair)

Metadata Use & Strategy Advisory Group
Sever Bordeianu,
University of New Mexico
Tim Cole, University of Illinois (co-chair)
Kristina Eden, HathiTrust
Stephen Hearn, University of Minnesota (co-chair)
Rebecca Lubas, Claremont Colleges (PSC liaison)
Sandra McIntyre, HathiTrust
Angelina Zaytsev, HathiTrust

Quality Assurance and Standards Working Group
Aaron Elkiss, University of Michigan
Paul Fogel, California Digital Library (Chair)
Natalie Fulkerson, HathiTrust
Janet Gertz, Columbia University
Peter Gorman, University of Wisconsin-Madison
Kat Hagedorn, University of Michigan
Sandra McIntyre, HathiTrust
Michelle Paolillo, Cornell University
Angelina Zaytsev, HathiTrust

Committees and Working Groups

Federal Documents Advisory Group

Heather Christenson (HathiTrust, non-voting Chair)
Prue Adler, Association of Research Libraries
Ivy Anderson, California Digital Library
Kirsten Clark, University of Minnesota
Beth Dupuis, University of California, Berkeley
Michael Norman, University of Illinois
Judith Russell, University of Florida
Sarah Shreeves, University of Arizona Libraries
(Program Steering Committee Liaison)

Shared Print Advisory Group

Lizanne Payne (HathiTrust, non-voting Chair)
Ivy Anderson (California Digital Library)
Daniel Dollar (Yale University)
Dan Heuer (Bucknell University)
Ken Peterson (Harvard University)
Matthew Revitt (Maine Shared Collections Strategy)
Matthew Sheehy (Brandeis University)
Tom Teper (University of Illinois at Urbana-Champaign)
Ben Walker (University of Florida)
Heather Weltin
(University of Wisconsin, Madison)

2018 Member Meeting Program Committee

Heather Christenson, HathiTrust, Chair of 2018 Member Meeting Program Committee
Eleanor Dickson, HathiTrust Research Center, University of Illinois
Geoff Harder, University of Alberta
Rebecca Lubas, Claremont Colleges
Jessica Rohr, HathiTrust
Melissa Stewart, HathiTrust
Günter Waibel, California Digital Library

Because you need to know...

Melissa Stewart

Connecting to WiFi

SSID: HathiTrust Wi-Fi

Password: HTMM18

The Basics

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Social Media

If you're tweeting...

#HTMM18

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Collaborative Notes in Google

http://bit.ly/HTMM18TEN

Join the Party Elephant!

HATHI TRUST

celebrating ten years

HathiTrust at 10

Mike Furlough, Executive Director

The HathiTrust Team: Here, There, and Everywhere

Heather Christenson, Program Officer, Federal Documents and Collections

Eleanor Dickson, Digital Humanities Specialist, HTRC

Natalie Fulkerson, Collection Services Librarian

Mike Furlough, Executive Director

Sandra McIntyre, Director of Services and Operations

Lizanne Payne, Shared Print Program Officer

Jessica Rohr, Member Engagement and Communications Specialist

Melissa Stewart, More than Ass't to the Exec Director

Angelina Zaytsev, User Services Librarian

Kristina Eden, Copyright Review Program Manager

Plus dozens of people at the **University of Michigan, Indiana University, California Digital Library, and the University of Illinois**.

And the still dozens more **contributing staff** at HathiTrust member institutions participating in **Copyright Review** and **HathiTrust User Support.**

Thank you Lizanne Payne!

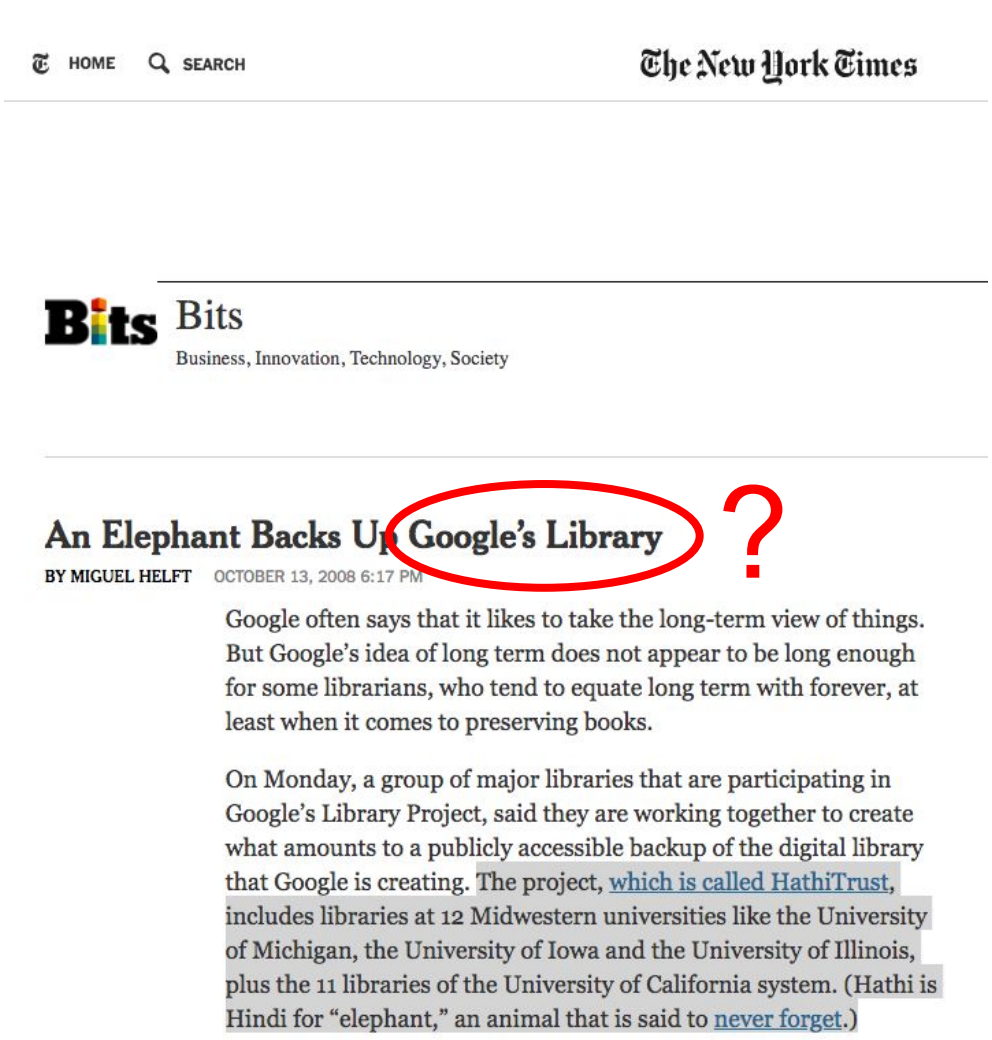
HATHI TRUST
CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •

Welcome to 17 New Members in 2018

Brigham Young University
Carleton College
Elon University
Grinnell College
McMaster University
New Mexico State University
Texas State University
University at Buffalo
University of Cincinnati
University of Colorado Boulder
University of Hawai`i
University of Nevada, Reno
University of Richmond
University of Toronto
University System of Georgia
Wayne State University
Whitman College

HATHI TRUST

celebrating ten years



HOME SEARCH

The New York Times

Bits

Business, Innovation, Technology, Society

An Elephant Backs Up Google's Library

BY MIGUEL HELFT OCTOBER 13, 2008 6:17 PM

Google often says that it likes to take the long-term view of things. But Google's idea of long term does not appear to be long enough for some librarians, who tend to equate long term with forever, at least when it comes to preserving books.

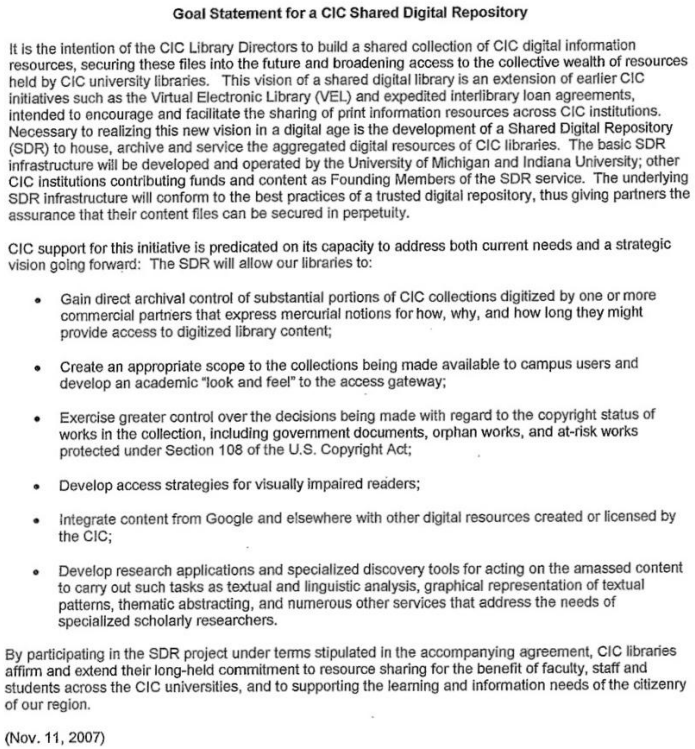
On Monday, a group of major libraries that are participating in Google's Library Project, said they are working together to create what amounts to a publicly accessible backup of the digital library that Google is creating. The project, which is called HathiTrust, includes libraries at 12 Midwestern universities like the University of Michigan, the University of Iowa and the University of Illinois, plus the 11 libraries of the University of California system. (Hathi is Hindi for "elephant," an animal that is said to never forget.)

October 13, 2008

HathitTrust launches with 12 founding members from the Committee on Institutional Cooperation and the 11 libraries of the University of California.

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •

HATHI TRUST



Goal Statement for a CIC Shared Digital Repository

It is the intention of the CIC Library Directors to build a shared collection of CIC digital information resources, securing these files into the future and broadening access to the collective wealth of resources held by CIC university libraries. This vision of a shared digital library is an extension of earlier CIC initiatives such as the Virtual Electronic Library (VEL) and expedited interlibrary loan agreements, intended to encourage and facilitate the sharing of print information resources across CIC institutions. Necessary to realizing this new vision in a digital age is the development of a Shared Digital Repository (SDR) to house, archive and service the aggregated digital resources of CIC libraries. The basic SDR infrastructure will be developed and operated by the University of Michigan and Indiana University; other CIC institutions contributing funds and content as Founding Members of the SDR service. The underlying SDR infrastructure will conform to the best practices of a trusted digital repository, thus giving partners the assurance that their content files can be secured in perpetuity.

CIC support for this initiative is predicated on its capacity to address both current needs and a strategic vision going forward: The SDR will allow our libraries to:

- Gain direct archival control of substantial portions of CIC collections digitized by one or more commercial partners that express mercurial notions for how, why, and how long they might provide access to digitized library content;
- Create an appropriate scope to the collections being made available to campus users and develop an academic "look and feel" to the access gateway;
- Exercise greater control over the decisions being made with regard to the copyright status of works in the collection, including government documents, orphan works, and at-risk works protected under Section 108 of the U.S. Copyright Act;
- Develop access strategies for visually impaired readers;
- Integrate content from Google and elsewhere with other digital resources created or licensed by the CIC;
- Develop research applications and specialized discovery tools for acting on the amassed content to carry out such tasks as textual and linguistic analysis, graphical representation of textual patterns, thematic abstracting, and numerous other services that address the needs of specialized scholarly researchers.

By participating in the SDR project under terms stipulated in the accompanying agreement, CIC libraries affirm and extend their long-held commitment to resource sharing for the benefit of faculty, staff and students across the CIC universities, and to supporting the learning and information needs of the citizenry of our region.

(Nov. 11, 2007)

Preface from the Original "Memorandum of Agreement on Collaborative Principles for a CIC Shared Digital Repository"

--November 2007

HATHI TRUST

a shared digital repository

Welcome to the Shared Digital Future HathiTrust Launched

HathiTrust is a bold idea with big plans.

As a digital repository for the nation's great research libraries, HathiTrust (pronounced hah-tee) brings together the immense collections of partner institutions.

HathiTrust was conceived as a collaboration of the thirteen universities of the Committee on Institutional Cooperation and the University of California system to establish a repository for these universities to archive and share their digitized collections. Partnership is open to all who share this grand vision.

HathiTrust is a solution.

To prospective partners, HathiTrust offers leadership and reliability.

It provides a no-worry, pain-free solution to archiving vast amounts of digital content. You can rely on the expertise of other librarians and information technologists who understand your needs and who will address the issues of servers, storage, migration, and long-term preservation.

You will shape HathiTrust.

Growing the world's largest library won't happen overnight. You've heard of other digital libraries. This one is different in concept and scale. Its greatest promise—and challenge—rests in defining how to serve researchers in the digital age. Together we will develop and refine the services needed to search and use such a large digital collection, and realize collectively our greatest potential as a library community. Together we will make HathiTrust available, to everyone, anywhere, any time. Together we will shape the future.

Bold idea. Big plan. Join HathiTrust.

» Printer-friendly version

Current Partners

California Digital Library	University of California San Diego
Indiana University	University of California San Francisco
Michigan State University	University of California Santa Barbara
Northwestern University	University of California Santa Cruz
The Ohio State University	The University of Chicago
Penn State University	University of Illinois
Purdue University	University of Illinois at Chicago
University of California Berkeley	The University of Iowa
University of California Davis	University of Michigan
University of California Irvine	University of Minnesota
University of California Los Angeles	University of Wisconsin-Madison
University of California Merced	University of Virginia
University of California Riverside	

September Activities

- HathiTrust Website Update
- Indiana Mirror Site - Full Remote Management
- Storage Increase and Reorganization
- Large-Scale Search
- Public Interface Planning
- Wisconsin Ingest
- New Growth
- October Forecast

Read the full update >>

Currently Digitized

2,145,653 volumes
750,978,550 pages
80 terabytes
25 miles
1,743 tons
337,820 volumes (~16% of total)
in the public domain

Statistics info >>

HathiTrust in October 2008

(thanks Wayback Machine!)

Our 2008 Access Strategy?

HATHI TRUST
a shared digital repository

About | Access Information | Partnership | Repository & Services | Updates
Access To HathiTrust | Bibliographic Data Distribution | Rights Management

Access to HathiTrust

Where's the search interface to all of the content in HathiTrust? There is currently no single global interface to the body of content in HathiTrust repository.

Be assured we are working with our partners to accomplish this—a way that takes into account the large number of different sources of metadata and the different strategies each of the partners has for managing that information.

Good, useful, technology takes time.... and the strength and insight born of collaborative work.

But if you want a peek...

We have made it possible for institutions to load bibliographic records for public domain materials in HathiTrust into their own catalogs, and to access limited metadata for other materials. For information about these processes, please visit Bibliographic Data Distribution.

Users interested in viewing content in HathiTrust may want to visit some of the institutions that have loaded records:

- The University of Michigan Library loads all records for HathiTrust content. You can find HathiTrust content by searching this part of their online catalog, Mirlyn.
- The University of Chicago Library adds records for public domain materials into their discovery system, Lens.

Stay tuned!

"There is currently no single global interface to the body of content in HathiTrust Repository.

Be assured we are working with our partners to accomplish this...

Good, useful, technology takes time...."

Access in 2010

HATHI TRUST Digital Library

About Help Feedback

Search | Public Collections | My Collections

Catalog Search

Search information **about** the items.

Find

All Fields

Full view only

Search Tips:

Phrase Searching: Use quotes to search an exact phrase: e.g., "occult fiction"

Wildcards: Use * or ? to search for alternate forms of a word. Use * to stand for several characters, and ? for a single character: e.g., optim* will find optimal, optimize or optimum; wom?n will find woman and women.

Boolean Searching: Use AND and OR between words to combine them with Boolean logic: e.g., (heart OR cardiac) AND surgery will find items about heart surgery or cardiac surgery.

Full-text Search

beta Search words that occur **within** the items.

Find

Full view only

Search Tips:

Phrase Searching: Use quotes to search an exact phrase: e.g., "occult fiction"

Multiple Term Searching: When your search terms are not quoted phrases, avoid common words (such as: 'a', 'and', 'of', 'the', etc.) to speed up your search.

Boolean Searching: Use AND and OR between words to combine them with Boolean logic: e.g., heart OR cardiac will find items containing the word heart or the word cardiac; heart AND cardiac will find items containing both words. Use a minus (-) to remove words from the result e.g., heart -cardiac will find items containing the word heart that do not include the word cardiac.

Note

Full-text search does not currently include functionality in use elsewhere in HathiTrust (e.g., there is no sorting or collection-related functionality). Read more about our "large-scale search".

Collections

Browse, search, or make HathiTrust collections.

View Public Collections

Collections are a way to group items for private or public use. Once grouped, you can search the full-text of all items within a collection. More information about creating collections.

Featured Collection:

UM Press

Univ. of Michigan Press's full-text titles available in HathiTrust

October 2011

53 members conduct HathiTrust's Constitutional Convention to formally organize and charter the organization

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Since the Convention, HathiTrust has....

...added 90+ new members

...formed and organized under a new structure of member-driven governance.

...added full time executive leadership and expanded staff to support the organization

...developed and launched key infrastructure at three members other than U of Michigan

...expanded direct co-investment in the HathiTrust Research Center at Indiana University and U of Illinois

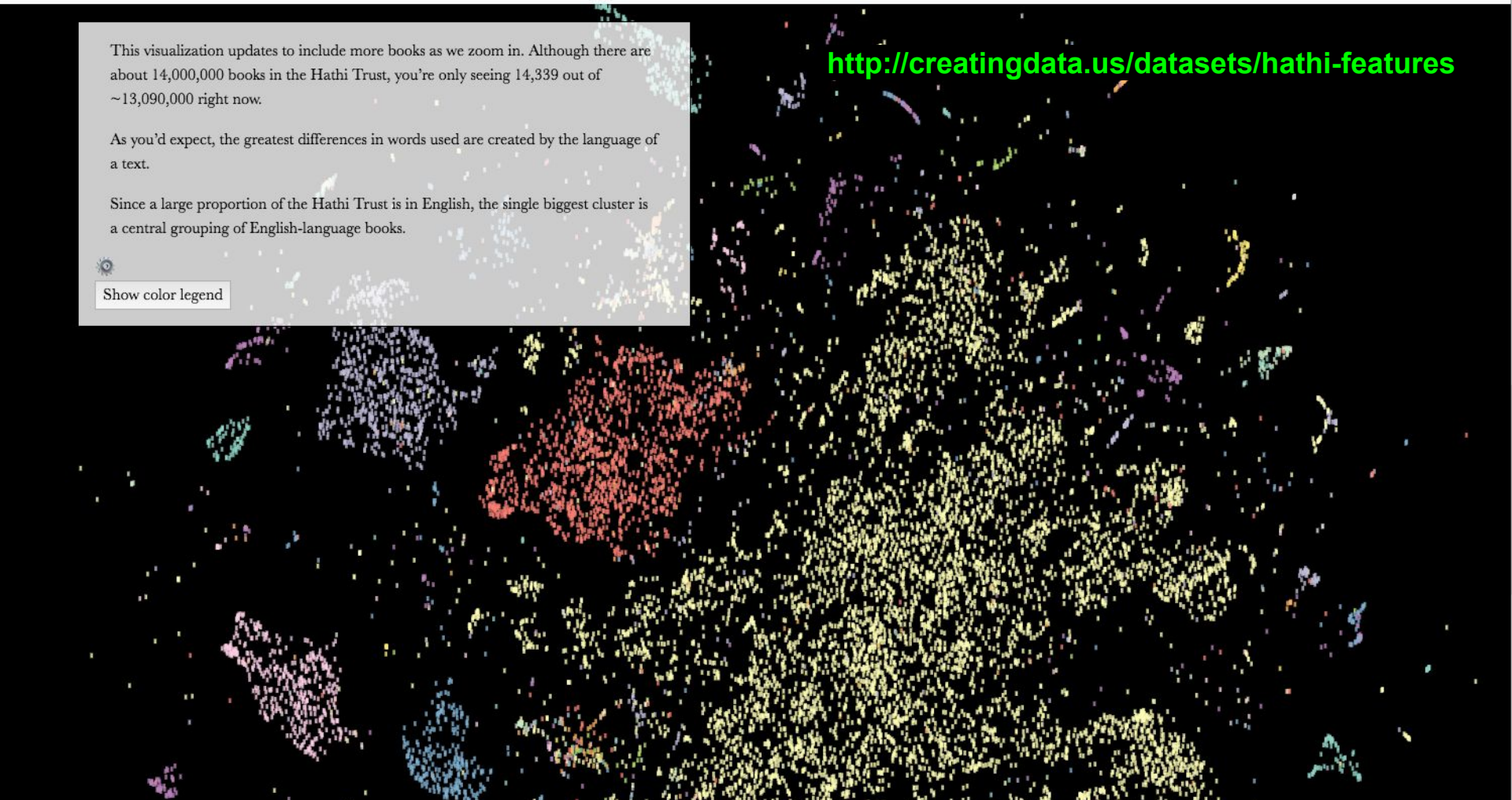
...added 7 million volumes to the collection, including 3 million open volumes.

...expanded print disabled services to the entire HathiTrust collection.

...launched the Shared Print Program, completing the 1st and beginning 2nd phases

...launched the US Federal Documents program.

...transitioned grant-funded copyright review into an ongoing service.


This visualization updates to include more books as we zoom in. Although there are about 14,000,000 books in the Hathi Trust, you're only seeing 14,339 out of ~13,090,000 right now.
As you'd expect, the greatest differences in words used are created by the language of a text.
Since a large proportion of the Hathi Trust is in English, the single biggest cluster is a central grouping of English-language books.
Show color legend
http://creatingdata.us/datasets/hathi-features

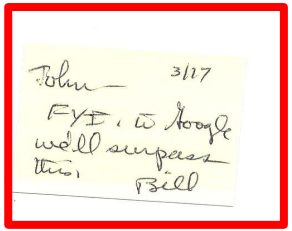
Most Useful Word of 2002

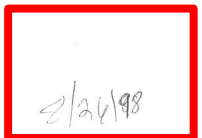
The American Dialect Society

The data below are estimates related to the digital conversion of 50,000 volumes (the average number of volumes transferred to Buhr annually). These estimates assume an average of 300 pages per volume, and reflect savings gained from converting a large number of pages (in this case, 15 million pages). These data would not be reflective of a smaller-scale conversion project.

Estimated Annual Digital Conversion Costs for Buhr Remote Shelving Facility

Category	Cost per page	Total cost	Average cost per volume
Scanning	10 cents	$1,500,000	$30.00
OCR/Markup	11 cents	$1,650,000	$33.00
Servers		$ 50,000	$ 1.00
RAID Storage		$ 60,000	$ 1.20
Development	1 cent	$ 150,000	$ 3.00
Equipment Amortization	1 cent	$ 150,000	$ 3.00
Total		**$3.56 million** + costs for connectivity and maintenance	**$71.20** + costs for connectivity and maintenance





CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Image of earliest cooperative agreement between Google and the University of Michigan, ~November 2002. Held in HathiTrust Administrative Files.

University of Michigan -- Google Project plan, Revision 9/1/2004 CONFIDENTIAL. Held in HathiTrust Administrative Files.

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Advertisement

Google Is Adding Major Libraries to Its Database

By JOHN MARKOFF
and EDWARD WYATT
Published: December 14, 2004

Google, the operator of the world's most popular Internet search service, announced today that it had entered into agreements with some of the nation's leading research libraries and Oxford University to begin converting their holdings into digital files that would be freely searchable over the Web.

It may be only a step on a long road toward the long-predicted global virtual library. But the collaboration of Google and research institutions that also include Harvard, the University of Michigan, Stanford and the New York Public Library is a major stride in an ambitious Internet effort by various parties. The goal is to expand the Web beyond its current valuable, if eclectic, body of material and create a digital card catalog and searchable library for the world's books, scholarly papers and special collections.

Google - newly wealthy from its stock offering last summer - has agreed to underwrite the projects while also adding its own technical abilities to the task of scanning and digitizing tens of thousands of pages a day at each library.

Enlarge This Image

Thor Swift

A book is scanned at Stanford University. Google's plans for digital files include the University of Michigan and the New York Public Library.

ARTICLE TOOLS

E-Mail This Article

HATHI TRUST · CELEBRATING TEN YEARS · 2008-2018 · CELEBRATING TEN YEARS · 2008-2018 ·

(113) Places.	Lat. N.	Lon. W.
	° ′ ″	° ′ ″
St. Nicholas Mole Pt.	19 49 20	73 27 30
Port à Piment . .	19 35 0	72 57 3
St. Mark's Point .	19 2 18	72 50 0
Arcadins, South end	18 45 0	72 38 30
Port au Prince,		
— The Road . .	18 33 42	72 23 45
Gonave I., N. E. Pt.	18 48 35	72 49 47
— West Point . .	18 57 0	73 19 0
Petite Goave,		
Hummock . .	18 26 51	72 50 30
Rocheleis Reef,		
— Pirogues . .	18 37 20	73 12 0
Point Jeremie . .	18 39 57	74 6 0
Cape Donna Maria .	18 37 25	74 25 0
Point des Irois . .	18 22 23	74 30 30
Cape Tiburon . .	18 19 25	74 27 32
Point à Gravois .	18 1 3	73 55 0
Isle à Vache, E. Pt.	18 4 0	73 31 30
Les Cayes, Town .	18 11 10	73 44 0
St. Louis, Old Fort	18 14 27	73 31 30
Aquin Bay,		
— Diamond Rock	18 13 48	73 20 0
Cape Jaquemel .	18 11 20	72 33 30
Alto Vela Rock .	17 28 11	71 40 30
Point Salinas . .	18 12 40	70 49 30
St. Domingo, City .	18 28 40	69 59 37
Sta. Catalina I., W.Pt.	18 19 0	69 1 0
Sama I., S. E. Point	18 12 0	68 31 30
Point Espada . .	18 19 48	68 30 0
Cape Engannо . .	18 34 30	68 20 30
Cape Samana . .	19 15 40	69 6 15
Cape Cabron . .	19 22 0	69 12 50
Old Cape François	19 40 30	69 56 30
Point Isabella . .	19 58 43	71 12 0
The Grange . .	19 54 35	71 43 10
Cape Haytien, City	19 46 20	72 14 0
Tortuga Island, W. Pt.	20 5 20	72 54 56
— East Point . .	20 3 33	72 34 10
Port Paix . . .	19 56 0	72 43 0
Port à L'Ecu . .	19 55 8	73 6 0
Mona Island, middle .	18 6 0	67 49 0
Menica	18 9 0	67 57 0
Zacheo	18 23 48	67 27 30
C. St. Juan, or N.E.Pt.	18 24 20	65 39 0
St. Juan, Castle .	18 29 10	66 6 52
Pt. Bruquin or N.W. Pt	18 31 18	67 7 0
Aguadilla Town .	18 25 10	67 7 0
Algarroba Point .	18 14 0	67 7 17
Cape Roxo . . .	17 58 30	67 7 30
Puerto Guanica .	17 57 44	67 11 0
Caxa de Muertos .	17 50 30	66 52 45
St. Thomas' Harbour, — Fort Christian .	18 20 5	64 57 0
Var. 2° 30′ E.		
St. John's S. E. Point	18 19 0	64 44 20
Virgin Gorda, E. Pt.	18 31 7	64 21 30
Anegada, S. E. Point	18 43 48	64 16 30

(114) Places.	Lat. N.	Lon. W.
	° ′ ″	° ′ ″
St. Croix or Sta. Cruz, — Christianstad .	17 44 8	64 43 30
Sombrero . . .	18 38 0	63 29 30
Anguilla, S. W. Point	18 11 0	63 16 0
— North East Point	18 16 0	63 2 30
St. Martin, N.W. Pt.	18 4 0	63 14 0
St. Bartholomew, E.Pt.	17 53 35	62 52 30
Saba, middle . .	17 39 20	63 19 0
St. Eustatius, Road	17 29 30	63 4 30
St. Christopher,		
— Basse Terre .	17 19 30	62 49 15
Nevis, Town . .	17 10 30	62 43 0
Barbuda, N. Point	17 43 0	61 52 0
Antigua,		
— St. John's Road	17 9 40	61 57 0
— English Harbour	17 3 0	61 50 0
Redonda. . . .	16 55 30	62 26 0
Monserrat, N.E. Pt.	16 48 0	62 16 40
Guadaloupe,		
— Basse Terre .	15 59 30	61 47 30
— Old Fort. . .	15 56 0	61 45 30
— Chateau Point .	16 12 30	61 12 0
— North Point .	16 30 0	61 37 0
Aves or Bird's I.	15 40 56	63 35 0
Deseada, N. E. Point	16 20 0	61 6 30
Mariegalante,		
— Basse Terre .	15 52 15	61 22 0
Saintes, S.W. end	15 51 20	61 44 0
Dominica,		
— Pr. Rupert's Bay	15 35 30	61 33 0
— Roseau . . .	15 18 23	61 28 0
— Scot's Head .	15 12 0	61 26 0
Martinique,		
— Fort Royal . .	14 35 55	61 7 30
— Port St. Pierre	14 44 30	61 11 30
— Macoubs Point	14 52 40	61 11 30
— Saline Point .	14 23 50	60 56 10
Var. 2° 47′ E.		
— Diamond Rock	14 26 30	61 6 0
St. Lucea,		
— Careenage, entr.	13 57 0	61 7 0
— Point Piton .	13 42 0	61 13 0
— Moulacuise Point	13 35 0	61 2 0
St. Vincent,		
— Kingston . .	13 9 0	61 15 0
Barbadoes,		
— North Point .	13 19 0	59 44 0
— Bridgetown .	13 5 30	59 43 15
— South Point .	13 3 0	59 39 0
Granadillos,		
— Becouya, N. Point	13 3 0	61 13 0
— Cariaco. N. Point	13 32 0	61 30 0
Grenada,		
— North East Point	12 14 0	61 42 0
— Fort St. George	12 2 54	61 49 40
— Point Salines .	11 58 0	61 52 0
Tobago, N. E. Point	11 22 0	60 27 0
— Man of War Bay	11 22 0	60 32 0
— South West Point	11 9 0	60 52 0
Trinidad,		
— Boca de Navois	10 41 45	61 48 0

Hayti or St. Domingo. Caribbee Islands. Porto Rico. Virgin Islands.

Digitized by Google

CELEBRATING TEN YEARS • 2008-2018 • HATHI TRUST

Image of correspondence: Peter Olson, Random House to Mary Sue Coleman, University of Michigan, June 10, 2005. Document held in HathiTrust administrative files.

Image of correspondence Patricia Schroeder, American Association of Publishers, June 9, 2005. Document held in HathiTrust administrative files.

Arts, Briefly; Major Publishers Sue Google

By EDWARD WYATT OCT. 20, 2005

Five large American publishing companies filed a lawsuit against Google yesterday, charging that the company's Google Print for Libraries program -- which plans to make searchable, digital copies of nearly all the books in five major research libraries -- violates the rights of the publishers under the federal Copyright Act. McGraw-Hill, Pearson Education, Penguin Group, Simon & Schuster and John Wiley & Sons filed a joint lawsuit in the Federal District Court in Manhattan, seeking a declaration that the Google Print for Libraries program violates copyright and requesting an injunction against the scanning of printed works without the permission of the copyright holder. The lawsuit is similar to one filed last month by three authors whose works are included in the University of Michigan library, one of the libraries whose collections Google plans to digitize. By filing a similar suit, the publishers put considerably more muscle behind the action against Google. Google has maintained that its program is covered by the fair use provision of copyright law and that it therefore does not need permission of [illegible]brary project. The program makes only small [illegible]available to users online; Google likens the [illegible]alog. EDWARD WYATT

Writers Sue Google, Accusing It of Copyright Violation

By EDWARD WYATT SEPT. 21, 2005

Three authors filed suit against Google yesterday contending that the company's program to create searchable digital copies of the contents of several university libraries constituted "massive copyright infringement."

The lawsuit, filed in United States District Court in Manhattan, is the first to arise from the Google Print Library program, the fledgling effort aimed at a searchable library of all the world's printed books.

Google intends to make money from the project by selling advertising on its search pages, much as it does on its popular online search-engine site.

The plaintiffs, who are seeking class-action status, also include the Authors Guild, a trade group that says it represents more than 8,000 published authors. Listed as plaintiffs in the suit are Daniel Hoffman, a former consultant in residence at the Library of Congress and the author of many volumes of poetry, translation and literary criticism; Betty Miles, an author of children's and young adult fiction; and Herbert Mitgang, the author of a biography of Abraham Lincoln as well as novels and plays. Mr. Mitgang is a former cultural correspondent and editorial writer for The New York Times.

Lawsuit Seeks the Removal of a Digital Book Collection

By JULIE BOSMAN SEPT. 12, 2011

Three major authors' groups and eight individual authors filed suit against a partnership of research libraries and five universities on Monday, arguing that their initiative to digitize millions of books constituted copyright infringement.

The lawsuit, filed in United States District Court for the Southern District of New York, contends that "by digitizing, archiving, copying and now publishing the copyrighted works without the authorization of those works' rights holders, the universities are engaging in one of the largest copyright infringements in history."

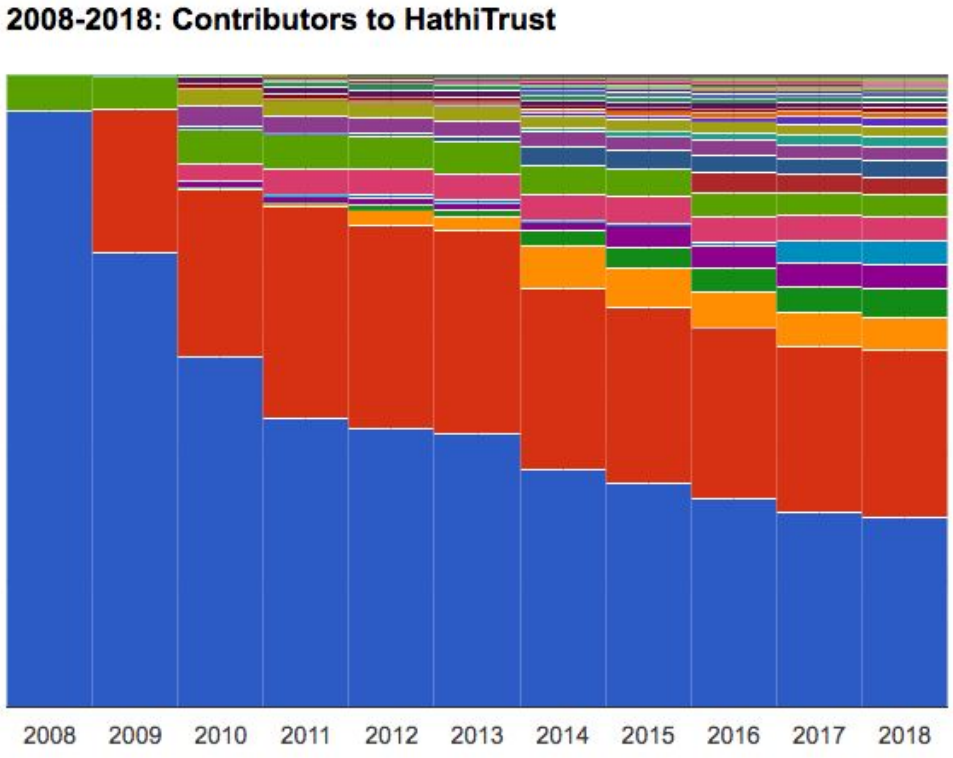
The plaintiffs in the lawsuit are the Authors Guild, the Australian Society of Authors and the Québec Union of Writers. Individual authors include Pat Cummings, Roxana Robinson and T. J. Stiles.

"We've been greatly concerned about the seven million copyright-protected books that HathiTrust has on its servers for a while," said Paul Aiken, executive director of Authors Guild, an industry group that says it represents more than 8,500 authors. "Those scans are unauthorized by the authors." HathiTrust is the name of the partnership of libraries.

“I cannot imagine a definition of fair use that would not encompass the transformative uses made by Defendants’ M[ass] D[igitization] P[rogram] and would require that I terminate this invaluable contribution to the progress of science and cultivation of the arts that at the same time effectuates the ideals espoused by the ADA.”

Judge Harold Baer, October 12 2012, in re A*uthors Guild v. HathiTrust*

HathiTrust changes

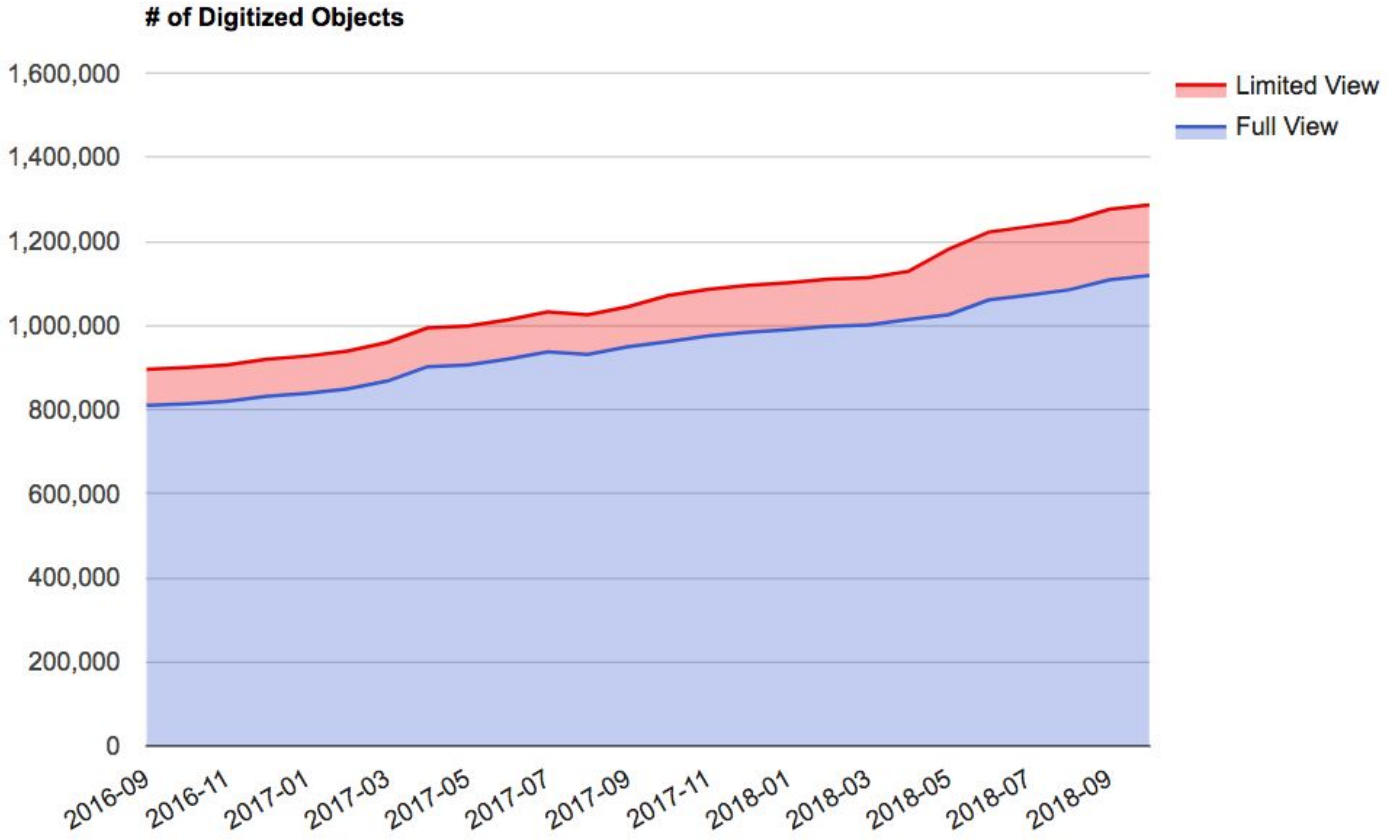


From a snapshot analysis of the collection at

https://is.gd/ht10collection

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Growth of US Federal Documents Collection



CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Today: HathiTrust at Human Scale

16.77 million volumes physically moved at least four times each: **70+ million touches of these volumes.**

8.58 million catalog **records created, converted, manipulated by catalogers**

5.87 billion pages….

3+ billion page turns by thousands of scanning staff

TABLE 22. PROPORTIONAL DEPTHS OF NEUTRAL AXIS

The table below gives the proportional depths of the neutral axis calculated from formula (6) on page 420 for various percentages of steel and moduli of elasticity. Its use is *not* advised for ordinary calculations of moments of resistance and dimensions of beams or slabs, because it presents no mean s of determining, without further calculation, the stress in the steel or the concrete, and therefore is liable to lead to uneconomical design. Its principal use is for determining the moment of resistance, and consequently the safe loading for beams already built.

Proportional Depth of Neutral Axis for Beam for different per cents of Steel and various assumed ratios of elasticity.. (See p. 310.)

p = Ratio of area of steel section to area of cross-section of beam above steel	Ratio of depth of neutral axis to depth of beam below most compressed surface of beam. Ratios of Modulus of Elasticity of Steel to Modulus of Elasticity of Concrete in Compression, $\frac{E_s}{E_c} = n$						
	6	7.5	10	12	30	35	40
0.001	0.10	0.115	0.132	0.143	0.217	0.232	0.246
0.002	0.184	0.159	0.181	0.196	0.292	0.311	0.328
0.003	0.173	0.191	0.217	0.235	0.344	0.365	0.384
0.004	0.196	0.217	0.246	0.266			0.428
0.005	0.217	0.239	0.270	0.292			0.464
0.006	0.235	0.258	0.292	0.314			
0.007	0.251	0.276	0.311	0.334			
0.008	0.266	0.292	0.328	0.353			
0.009	0.279	0.306	0.344	0.369			
0.010	0.292	0.320	0.358	0.384			
0.012	0.315	0.344	0.384	0.402			
0.014	0.334	0.364	0.407	0.436			
0.016	0.353	0.384	0.428	0.457			
0.018	0.369	0.402	0.446	0.476			
0.020	0.384	0.418	0.463	0.493			
0.030	0.446	0.483	0.531	0.562			
0.040	0.493	0.531	0.580	0.611			
0.050	0.531	0.569	0.618	0.649			

Digitized by Google

THE

FAIRY MYTHOLOGY.

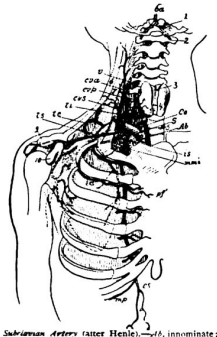
NORTHERN ISLANDS.

...asborgen slaar
...t grönskande haar
...ie drägten.

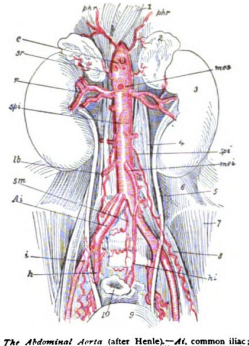
STAGNELII.

...arp on the glass castle plays,
...omb out their green hair always,
...e their shining white clothes.

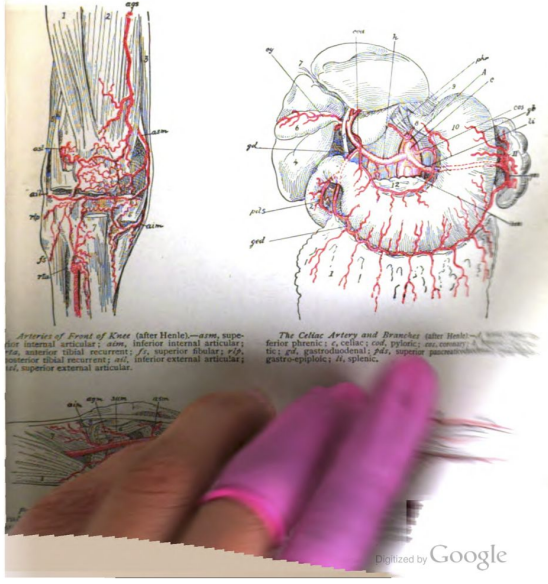
Digitized by Google



Subclavian Artery (after Henle).—*Ab*, innominate; *Cc*, common carotid; *S*, subclavian; *v*, vertebral; *Ba*, basilar; *mam*, internal mammary; *ts*, superior epigastric; *mf*, musculo-phrenic; *is*, superior intercostal; *cvp*, deep cervical; *ts*, inferior thyroid; *cvs*, ascending cervical; *cvs*, superficial cervical; *ts*, suprascapular; *ts*, posterior scapular.



The Abdominal Aorta (after Henle).—*Ai*, common iliac; *i*, external iliac; *h*, internal iliac; *sm*, middle sacral; *phr*, inferior phrenic; *lb*, lumbar; *c*, celiac; *mes*, superior mesenteric; *mes*, inferior mesenteric; *sr*, capsular; *r*, renal; *spi*, internal spermatic; *hi*, internal hemorrhoidal.



Arteries of Front of Knee (after Henle).—*asm*, superior internal articular; *aim*, inferior internal articular; *ata*, anterior tibial recurrent; *fi*, superior fibular; *ptp*, posterior tibial recurrent; *aei*, inferior external articular; *ase*, superior external articular.

The Celiac Artery and Branches (after Henle).—*d*, inferior phrenic; *c*, celiac; *gsd*, pyloric; *gsc*, coronary; *hs*, splenic; *gd*, gastroduodenal; *pds*, superior pancreaticoduodenal; *gsi*, gastro-epiploic; *hs*, splenic.

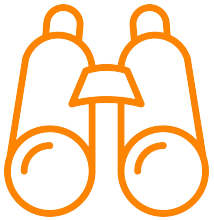
Digitized by Google

Our job....

2008-2018 • CELEBRATING TEN YEARS •
HATHI
TRUST
2008-2018 • CELEBRATING TEN YEARS •

Mission *and Vision*

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



Over the next five years ***HathiTrust will be a vital catalyst for emerging forms of research, teaching, and learning that engage the scholarly and cultural record***.

empower ☆

HathiTrust empowers its member libraries and their patrons through a durable, sustainable, and transparent organization and open infrastructure.

In 2019

- Begin development of new roadmap for infrastructure and users' interaction with it.
  - Add expertise in key areas, including technology leadership and discovery.

- *Adjust our fee model and member criteria to support diversity in HathiTrust membership institutions*

- Develop forward-looking collections Strategy

- *Finalize & implement of a Metadata Sharing Policy*

Metadata Sharing Policy: DRAFT

"To encourage discovery of materials, research, and for a variety of current and future uses, HathiTrust provides open access to HathiTrust-collected and -managed metadata … where it has been determined that legal, contractual, policy, ethical, practical, and strategic considerations allow."

AKA--as open as possible, with as few limits or restrictions as possible.

Rolling assessment needed to implement policy for each type of metadata, e.g., rights, quality, print retention, print holdings, Extracted Features, preservation, structural (METS), bibliographic.

COMMENT PERIOD OPENS SOON: At least one webcast presenting and discussing the policy in December timeframe with opportunities for member input.

enhance

In 2019

- Public Domain Day: 40,000+ new public domain works
  - Plus: continuing copyright reviews (1924-1963)
- Completion of Phase 2 of the Shared Print Program
  - Plus: A new Shared Print Program Officer
- Bringing Research Center innovations to the corpus
- Expanding outreach activities for HathiTrust and the Research Center

HathiTrust bolsters content discovery and diversifies uses of the scholarly record through services that improve the quality of the collection and the user experience.

transform

In 2019:

- Exploring new paths for print disabled services
  - Impact of ratification of the Marrakesh Treaty
- Shared Print
  - Engage national/international coordination
- Public policy watch
  - USMCA (post-NAFTA) and Canadian Copyright Term
  - Legislative reform of US Federal Depository Library Program
  - Privacy issues and the “Right to Be Forgotten”

HathiTrust addresses significant challenges libraries cannot independently confront to advance innovative forms of research, pedagogy, and public engagement.

Photo of George Harrison

Photo of Ringo Starr

Photo of York Ono

Photo of Pete Best

Photo of Stu Sutcliffe

Photo of Astrid Kercherr

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

Image of Beatles performing Hey Jude, on David Frost Show, September 4, 1968, surrounded by audience of fans.

CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST
CELEBRATING TEN YEARS • 2008-2018 •

Questions....

Thank you for being part of HathiTrust!

HathiTrust's 2019 Budget

Günter Waibel
Mike Furlough

2019 Budget Approval Process

HathiTrust members review and vote to approve “the total of the budget of HathiTrust and the total amount of dues, fees or other charges.”

Approval requires a simple majority of weighted votes.

Electronic ballot has been sent to all member representatives.
Voting opened on October 22; closes at 11:59 ET (US) Nov. 12, 2018.

All member representatives and attendees of this meeting have received a detailed memo describing our budget plan for 2019.

HathiTrust Budget: Sources of Income

Member Fees

- Annual fees for returning members
- Prorated fees for members who join mid year

Investment income

- Interest paid on all HathiTrust funds held in University of Michigan financial systems
- Tied to quarterly T-Bill rate

HathiTrust Budget: Expenses

Operations

- Administration of organization and services
  - Staff, office, travel, meetings, communications, copyright review
- Infrastructure
  - Equipment, data centers, backup, networking, staff
- Contracted Services
  - Zephir (CDL)
  - Administration, hosting (Michigan)

Programs

- US Federal Documents Program
- Shared Print Program
- HathiTrust Research Center

Expenses include recurring/annual costs and "one-time" project-based costs.

Planning for Growth, Renewal and Contingencies

Strategic Investment Funds: Supports innovative projects and opportunities that advance our Strategic Directions and services. In the past these have been referred to as Programmatic funds.

Infrastructure Renewal Fund: annual savings to support the lifecycling of major equipment and systems that support preservation of and access to the collection.

General (Board) Reserves: held aside for uses approved by Board of Governors in event of organizational emergencies. Target amounts should cover expenses for between 6 and 8 months, based on a rolling three year average.

HathiTrust 2019 Budget Summary

AREA/ACTIVITY	2018 PLANNED	2019
INCOME		
Current Member Fees	**3,335,396**	**3,587,445**
Investment Income	40,000	100,000
Prorated new member fees	105,000	90,000
TOTAL INCOME	3,480,396	3,777,445
INFRASTRUCTURE RENEWAL		
Starting Balance	1,309,059	2,052,587
Allocations/Transfers IN	743,528	569,710
Expenses/Transfers OUT	-	(150,000)
END BALANCE	2,052,587	2,472,297
GENERAL (BOARD) RESERVE		
Starting Balance	1,281,628	1,281,628
Allocations/Transfers IN	-	31,234
Expenses//Transfers OUT	-	-
END BALANCE	1,281,628	1,312,862
OPERATIONS		
Starting Balance	1,759,006	1,694,006
Allocations/Transfers IN	2,266,616	2,434,901
Expenses/Transfers OUT	(2,331,616)	(2,799,534)
END BALANCE	1,694,006	1,329,373
PROGRAMS		
Starting Balance (Strategic Investment Funds)	2,578,530	2,259,122
Allocations/Transfers IN	550,252	772,835
Expenses/Transfers OUT	(869,660)	(1,240,209)
END BALANCE	2,259,122	1,791,748
EOY NET CASH BALANCE (sum all END BALANCES)	**7,287,343**	**6,906,280**

2019 Fee Calculations & Factors

Cost per volume = total fees/number of volumes

All partners share costs for public domain volumes:

(Public Domain volumes x Cost per volume)/Total Members

For 2019: $9275

Total number of members: 143 (+13 new members in 2019)

Members share costs for in-copyright volumes based on holdings

(InCopyright Volume$_N$) = Cost per volume/number of holding members

Cost varies per member

If cost per volume is $0.20

If 2 members hold a HathiTrust title: $0.10 for both members

If 10 members hold a HathiTrust title: $0.02 for all 10 members

2019 Fee Analysis

Total fees collected increased by +7.55%

Most members saw a fee decrease

- Increased recurring costs for staff and associated costs, member meeting, hosted services, inflationary costs.
- A steady growth in the number of volumes kept cost per volume lower
- Significant growth in membership allocated costs more widely..

Average fee change: -0.99%
Median fee change: -1.43%

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

Proposals: Membership Criteria & Fee Model

Wendy Lougee
Mike Furlough

Membership Criteria and Fee Models

The Bylaws state: The Board proposes, the Members dispose

- The total of the annual budget and fees
- The formula or the criteria for eligibility for membership
- The principle proposed by the Board of Governors by which the dues, fees, or other charges are to be levied against the Members

The current financial model was implemented with new governance in 2013.

Member criteria have not been formally proposed and approved.

Membership and Finance Committee

Formed by the Board of Governors in 2016

- Define member criteria for approval
- Review and propose adjustments in our membership strategy
- Review our fee model to determine if adjustments were necessary

Timeline of Work

2016

- Initial investigations by Membership and Finance Committee
- Recommendations to Board of Governors (accepted)

2017

- Further investigations focused on fee model
- Consultant engaged, models considered

2018

- Final development of member criteria
- Final development of adjustments to fee model
- Presentation to membership

2019

- Membership votes on fee model and member criteria

Issues and Challenges Considered

Membership:

- Increasing diversity of members
- Affirming contribution to public good and common good of members

Membership strategy:

- Taking a proactive approach to membership recruitment
- Identifying and reducing barriers to membership

Fee model:

- Ensuring alignment with membership strategy
- Address year-to-year volatility in fees
- Incentivizing membership and participation

Bounding the Questions: Principles

- HathiTrust will remain focused on text-based content for near-term

- Program development will build on core functions, including
  - Corpus content, preservation infrastructure
  - Discovery and access
  - Support of local collection management decisions
  - Data aggregation to enable text/data mining

- Member/Financial models should continue to support public good

All of these were considered as core assumptions for the development of HathiTrust's Strategic Directions for 2019-2023.

Recommendations

Membership strategy

- Recruit a diverse, robust membership
- Recruit 5-10 new members each year
- Identify libraries with collections to deposit
- Recruit international members in a targeted manner
- Develop organizational capacity for marketing, communications, engagement
- **Define formal criteria for membership approval**

Financial

- **Explore tiered membership to encourage diversity of members**
- Develop multi-year planning framework
- Forecast program development and migration
- Reduce frequency of submission of holdings data to simplify fee volatility

Current Requirements of Members Will Not Change

Legal standing to enter into a Supporting Institution Agreement (Membership Agreement) with the University of Michigan on behalf of HathiTrust.

Operate a SAML 2.0-compliant identity provider

- such as Shibboleth or OpenAthens
- release required attributes
- maintain membership in an accepted identity federation where available and adhere to its standards

Deliver library holdings data on a regular basis, employing HathiTrust specifications.

Proposed: Criteria for Membership

- Academic and research institutions (international)
  - Not-for-profit organizations, such as universities, colleges, museums, research centers, agencies that operate one or more libraries
  - University systems (US only)
    - A multi-institution, not-for-profit higher education system, operating one or more libraries. *Systems are assessed a public domain fee for each R1 institution.* **(new)**
  - Consortia
    - Multi-institution, non-profit consortium operating services (beyond licensing) on behalf of member libraries
    - Need to be able legally to sign a HathiTrust member agreement on behalf of the consortium's participant members, provide HathiTrust with consolidated data, and verify the full participation of all members, in a manner similar to that of university systems.

Strategy: International members

We welcome inquiries from non-profit, non-U.S. research libraries, which will be reviewed on a case-by-case basis.

HathiTrust currently prioritizes membership for Canadian, Australian, New Zealand, and UK institutions.

U.S. and non-U.S. members are assessed fees in the same way, and those fees must be paid in U.S. dollars.

Some services and content may not be available in all jurisdictions if local copyright laws do not permit them or if we cannot cost-effectively implement them in those jurisdictions.

Excluded: Non-library memberships

Not eligible at this time

We will focus on recruiting members that align best with our mission, and for whom we are best positioned to develop and offer services and programs.

Deposit of materials by non-members still needs to be addressed.

Background: Fee Model Proposal

Our fee model will continue to focus on common support for a public good AND the value members derive from their participation in HathiTrust.

A greater diversity of members will strengthen HathiTrust and the library community.

Elements of our fees that support a “common good” should be equitably distributed according to the member’s ability to contribute.

We should take steps that can minimize year-to-year volatility in the assessment of fees.

Proposal Affects Public Domain Fee Only

Cost per volume = total fees/number of volumes
For 2019: $0.2147

All partners share costs for public domain volumes:

(Public Domain volumes x Cost per volume)/Total Members

For 2019: $9275

Members share costs for in-copyright volumes based on holdings

($InCopyright\ Volume_N$) = Cost per volume/number of holding members
For 2019: Varies per member

Total number of members: 143 (+13 new members in 2019)

Proposal: Tiered Public Domain Fees

Beginning in 2020:

Members are sorted into three tiers, based on their total library expenditures

- Top tier (top 20%): public domain fee would be +33% higher
- Bottom tier (bottom 20%): public domain fee -33% lower
- Middle tier (middle 60%): no change to public domain fee.

Example: base public domain fee = $9,275

Tier	Weight	Public Domain Fee
Tier 3	1.33	$12,336 = 1.33 x $9,275
Tier 2	1.00	$9,275 = 1.00 x $9,275
Tier 1	0.67	$6,279 = 0.67 x $9275

Comparing the Tiers

The majority of members (tier 2) will see no change in their fees.

Tier 3 Characteristics

- Largest total expenditures
- Larger, more unique collection holdings
- Pay the largest fees (generally above $35,000)
- Public Domain fee is 19% of average total fee in this tier
- Libraries in Tier 3 will see a ~10% **increase** in total fees

Tier 1 Characteristics

- Smallest total expenditures
- Smaller, more duplicative collection holdings
- Pay the lowest fees (but generally above $10,000 US)
- Public Domain fee is 84% of average total fee in this tier
- Libraries in Tier 1 will see a ~25% **reduction** in fees

Implementation

Proposal would not go into effect until 2020 budget year

Tiers to be finalized in early 2019, before vote.

Based upon Total Annual Expenditures (in US Dollars)

Sourced from public databases or reports

Among the HathiTrust members

- Top 20% of total expenditures = Tier 3 with a 1.33 weight
- Bottom 20% of total expenditures = Tier 1 with a 0.67 weight

Timeline: Consultation and Voting

November 2018-January 2019 Consultation Period

- More complete details with real world examples in early November.
- At least 3 webcasts led by Mike Furlough will be scheduled.
- 1-1 calls can be requested and/or we will arrange them.
- Feedback in written form always welcome.

February 2019

- Membership votes on proposals (dates TBD).

Implementation

- Criteria: Immediately upon approval
- Fees: Beginning with 2020 budget (January 2020).

Morning Break
15 minutes

NEXT: HathiTrust Research Center

HathiTrust Research Center

John Walsh, Indiana University
Stephen Downie, University of Illinois

HTRC Big News: Access to Full HathiTrust Corpus

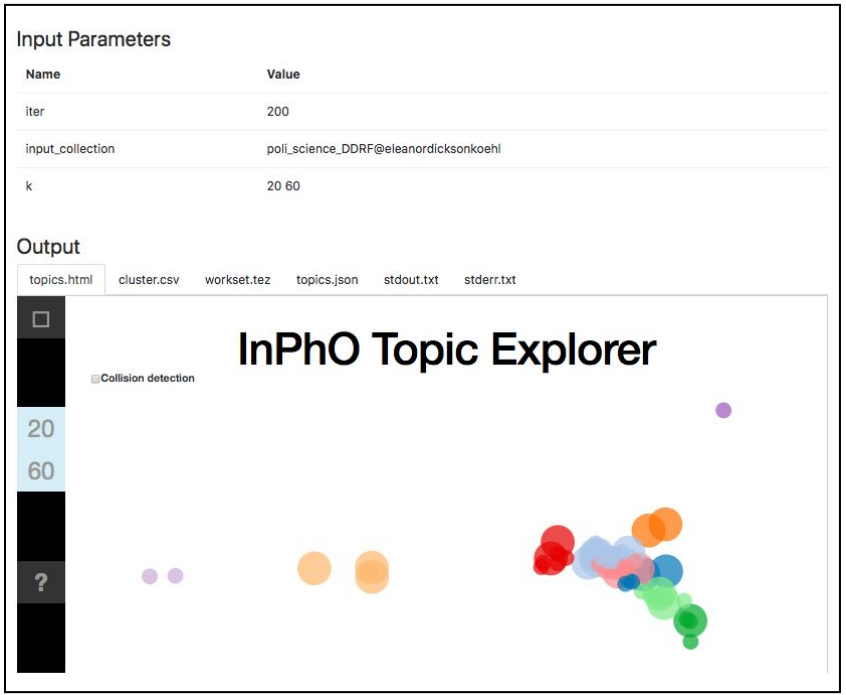
- As of last month, September 2018, HTRC tools and environments provide access to the full 16,000,000+ volumes of the HathiTrust Digital Library, *including the in-copyright content.*
- Web algorithms and downloadable datasets available to all.
- Data capsules with access to public domain content available to all.
- Data capsules with access to in-copyright content available to individuals affiliated with HathiTrust member institutions.

HTRC Services: Tools and data

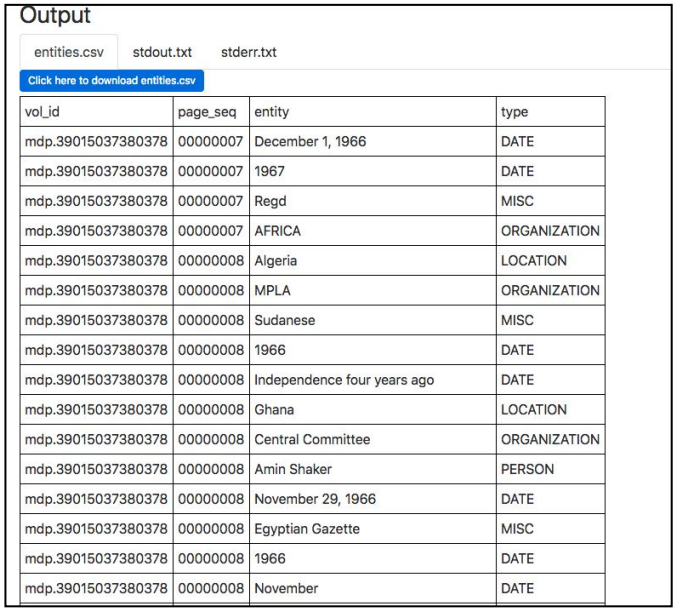
1. HTRC Algorithms
   - Pre-determined web-based tools, in addition to HT+Bookworm
2. Extracted Features dataset
   - Downloadable data
3. Data Capsule
   - Secure, web-based desktop environment

analytics.hathitrust.org

HTRC Algorithms



Topic modeling



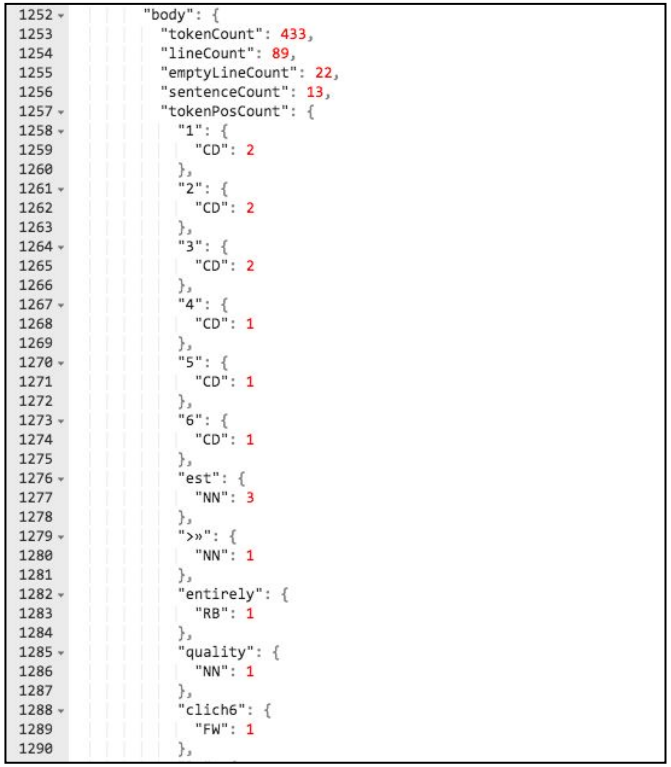
vol_id	page_seq	entity	type
mdp.39015037380378	00000007	December 1, 1966	DATE
mdp.39015037380378	00000007	1967	DATE
mdp.39015037380378	00000007	Regd	MISC
mdp.39015037380378	00000007	AFRICA	ORGANIZATION
mdp.39015037380378	00000008	Algeria	LOCATION
mdp.39015037380378	00000008	MPLA	ORGANIZATION
mdp.39015037380378	00000008	Sudanese	MISC
mdp.39015037380378	00000008	1966	DATE
mdp.39015037380378	00000008	Independence four years ago	DATE
mdp.39015037380378	00000008	Ghana	LOCATION
mdp.39015037380378	00000008	Central Committee	ORGANIZATION
mdp.39015037380378	00000008	Amin Shaker	PERSON
mdp.39015037380378	00000008	November 29, 1966	DATE
mdp.39015037380378	00000008	Egyptian Gazette	MISC
mdp.39015037380378	00000008	1966	DATE
mdp.39015037380378	00000008	November	DATE

Named entity recognition

analytics.hathitrust.org

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

HTRC Extracted Features Dataset



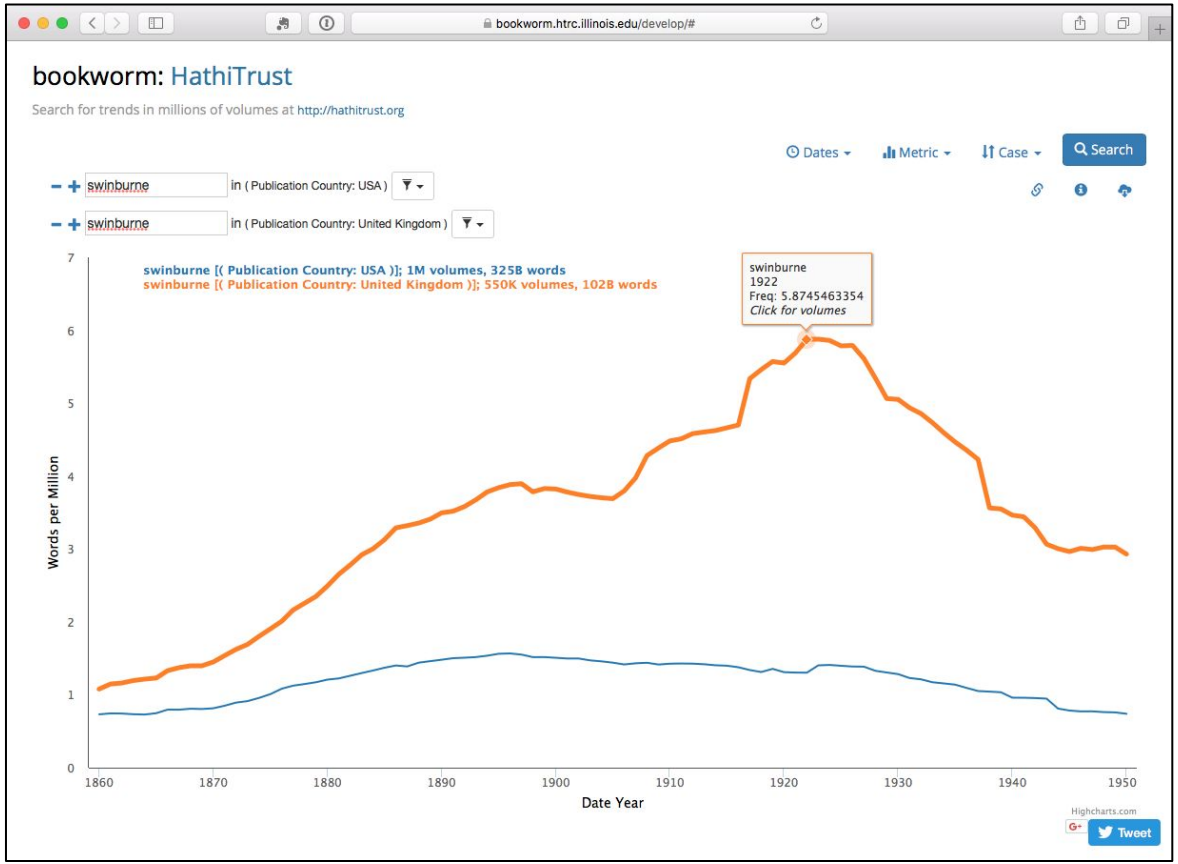
Feature data, including word & word counts

Volume and page metadata

analytics.hathitrust.org

HathiTrust+Bookworm

bookworm.htrc.illinois.edu



CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

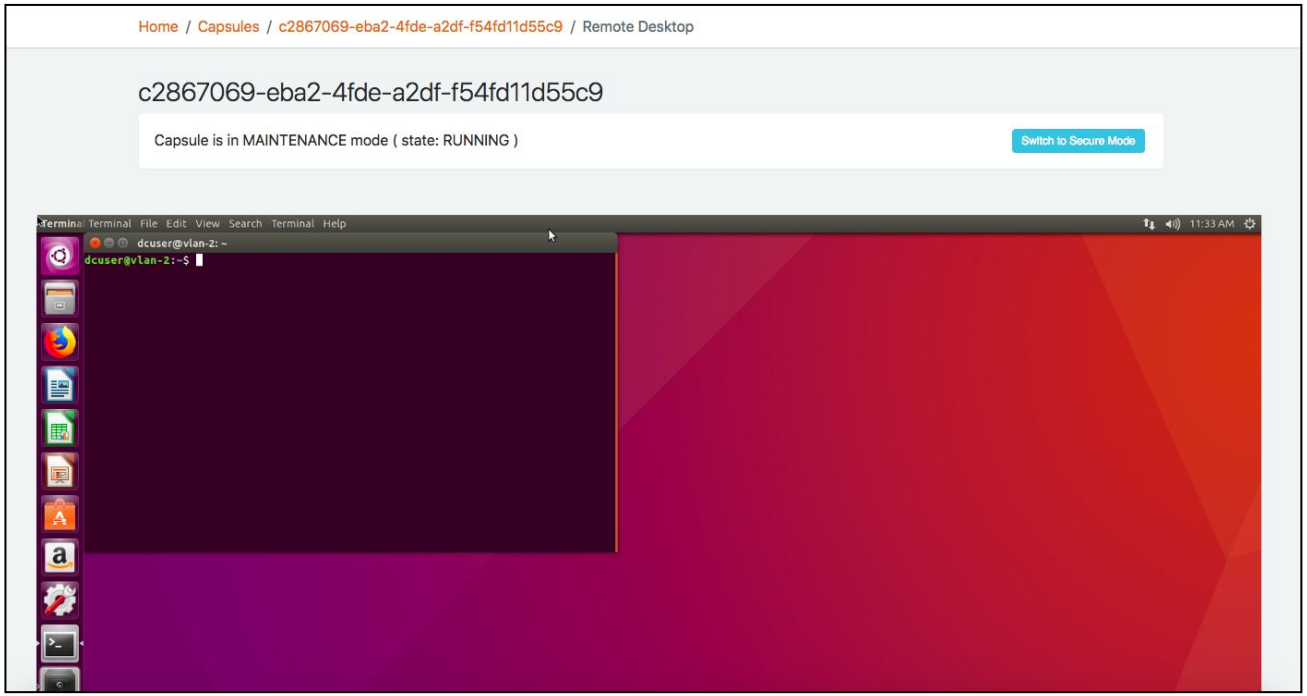
HTRC Analytics for HT Members

Tool	Function	Data access
HTRC Data Capsule	Provides a secure computing environment for text analysis on the HathiTrust corpus, using the researcher's tools of choice.	Access to copyrighted items using an HTRC Data Capsule is available ONLY to researchers affiliated with HathiTrust member institutions.

Our ability to provide data capsules is limited by finite computing resources, which resulted in the decision to offer this service as an exclusive benefit of HathiTrust membership.

See list of HathiTrust Member Institutions.

HTRC Data Capsule



analytics.hathitrust.org

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

How can HTRC support your researchers?

Research Support

- Dataset and workset creation
  - For those who need specialized support
- Monthly office hours
  - Second Wednesday of every month at 3 p.m. ET
  - go.illinois.edu/htrchelp-live
- Email reference
  - htrc-help@hathitrust.org

How can HTRC support your researchers?
Training

- Workshops
  - Digging Deeper, Reaching Further train-the-trainer curriculum for librarians
    - Final OER coming this fall! See website for current content: teach.htrc.illinois.edu
    - Plan to continue training and workshops
- Documentation and tutorials
  - HTRC Wiki: wiki.htrc.lllinois.edu
  - Programming Historian Extracted Features lesson: https://programminghistorian.org/en/lessons/text-mining-with-extracted-features

How can HTRC support your researchers?
Advanced Collaborative Support program

- Competitively awarded time and resource support
- Currently wrapping-up 4th round
- Project reports: https://wiki.htrc.illinois.edu/x/CADiAQ

Research example: Wikification as the First Step in Text Mining in Architectural History

Dan Baciu (now at University of California, Santa Barbara)

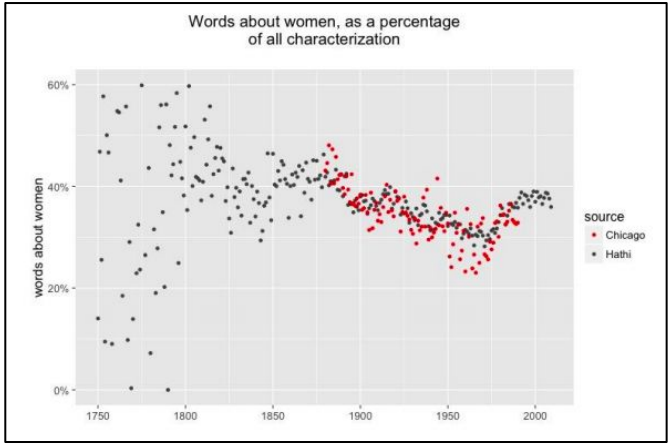
- Explores the history of the term “Chicago School” through the HT corpus.
- Used “wikifier” tool to link named entities to corresponding Wikipedia entries.
- Found many different types of “Chicago Schools” (e.g., the Chicago School of bone breakers)
- Found the use of the term as applied to architecture dating to 1889.

Project report:
https://wiki.htrc.illinois.edu/download/attachments/31588360/HTRC_ACS_report171001.pdf

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

The Transformation of Gender in English-Language Fiction

Ted Underwood (University of Illinois), David Bamman (University of California, Berkeley), and Sabrina Lee (University of Illinois)



- Used natural language processing (NLP);
- Found that depictions of men and women in fiction have become less distinct over time;
- Found the overall number of female characters and female authors declined by half in literature from 1850 to 1950.

http://culturalanalytics.org/2018/02/the-transformation-of-gender-in-english-language-fiction/

CELEBRATING TEN YEARS • 2008-2018 • CELEBRATING TEN YEARS • 2008-2018 •
HATHI TRUST

HTRC Phase 2: Organization and Service Units

Director (IU), Co-Director (Illinois) and three Associate Directors leading three service units:

- Cyberinfrastructure Operations Services (IU)
- Education and Outreach Services (HT/Michigan)
- Research Support Services (Illinois)

HTRC Phase 2: Goals

1. Provide HTRC users consistent and reliable access to existing HTRC research environments;
2. Provide HTRC users with new tools and enhancements for existing HTRC research environments;
3. Expand the HTRC user base;
4. Provide services to both HTRC research users and general HathiTrust users by developing and implementing a framework for migrating HTRC work to core HathiTrust environments;
5. Operate the HathiTrust Research Center so as to provide incentives for institutions to join as members of HathiTrust to enhance sustainability of HathiTrust and HTRC.

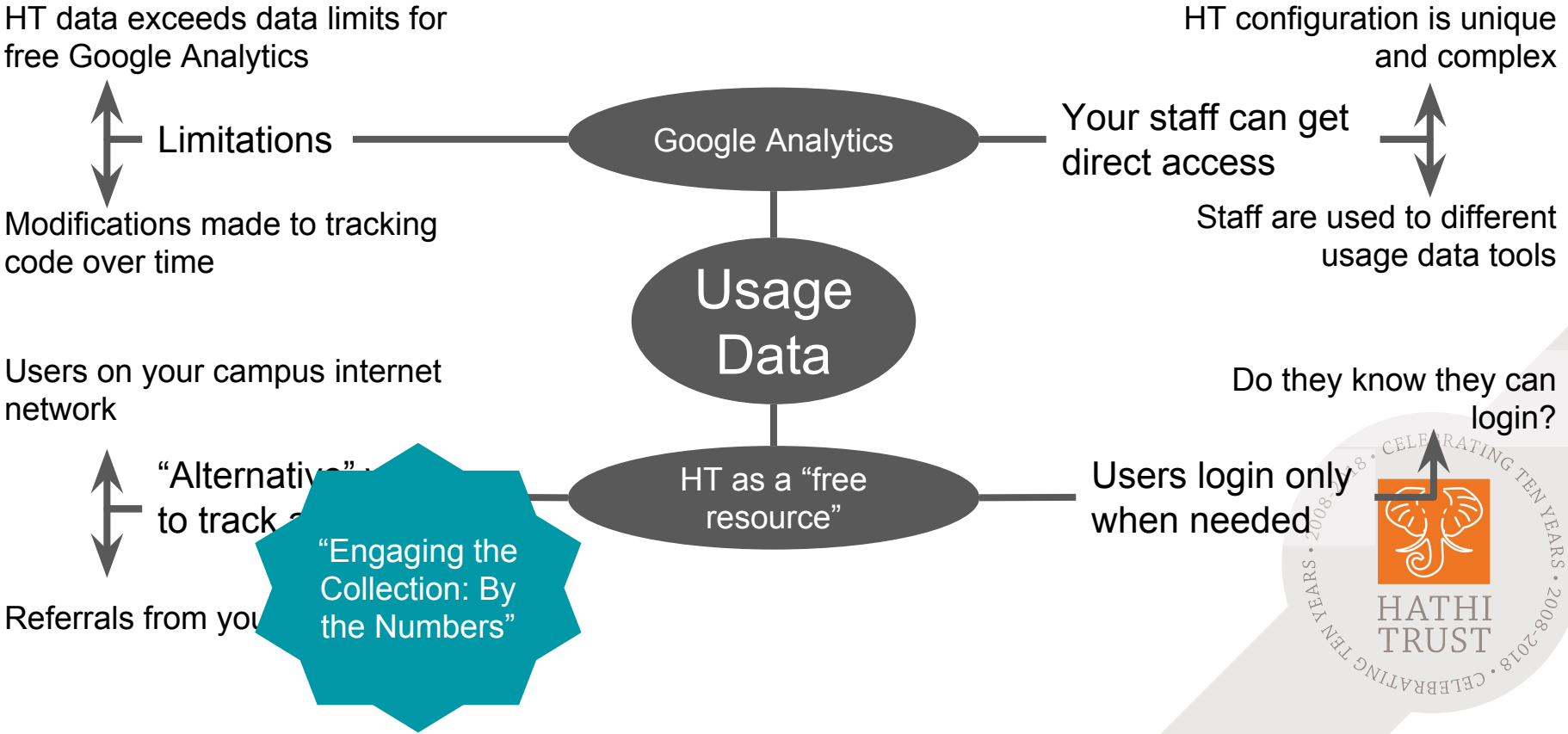
Getting Started, Support & Questions

Access tools, create an account, sign-in	analytics.hathitrust.org
Obtain support from HTRC staff	htrc-help@hathitrust.org
HTRC Co-director, J. Stephen Downie, University of Illinois	jdownie@illinois.edu
HTRC Director, John Walsh, Indiana University	jawalsh@indiana.edu

Ask Me Anything (About HathiTrust)

HathiTrust Staff

What can HathiTrust tell me about how my faculty and students use HathiTrust (HT)?



What is HathiTrust doing to improve the ingest process for contributing institutions?

Documentation is a major focus this year

- Digitization specification
- Content package requirements
- Website content

Expanded support for locally-digitized content will be a major focus in 2019

- Process improvements
- New tracking/reporting capabilities
- New validation tools

What are we doing to expand what can be ingested?

EPUB ingestion pilots:

1. Humanities Open Books program *(underway)*:
   University presses funded by NEH and Mellon Foundation to convert backlist titles and open them up as OA, with long-term preservation
2. University of Michigan Press, ebook titles *(January 2019)*
3. Knowledge Unlatched, newly open titles *(Summer 2019)*

Desired outcomes of pilots:

- Completely re-worked and more flexible processes for ingestion, display, download, and long-term preservation
- Familiarity with directions for ebook publishing and better understanding of publishers' born-digital needs
- Solid scoping for pilots with additional born-digital formats, e.g., **PDF/a** and **DAISY**

Will the Shared Print Program expand to other categories ?

The Shared Print Program currently focuses on **lendable monographs**. Other categories and material types can be considered:

- Non-circulating volumes: Current policy requires lending physical volumes so other libraries can borrow them if they deselect their own copies. Some operational issues need to be resolved if non-circulating volumes are retained, such as how to identify them to other libraries during deselection

- Federal documents: There is strong interest in expanding the HathiTrust program to include matching Federal documents, perhaps building on the FDLP program.

- Serials: HathiTrust may want to coordinate with existing shared print journal efforts and ensure that HathiTrust journals are retained

What are priorities for the Federal Documents Program?

Mission: *Through coordinated and collective action, expand and enhance digital access to U.S. federal publications including those issued by GPO and other federal agencies.*

- Building a **comprehensive collection**
  - U.S. Federal Documents Registry
  - Identification of our federal documents collection
  - Intentional collection development, including Collection Framework and "gap filling"
  - More digitization!
- Improving discovery of and access to federal documents
  - Metadata improvement and enrichment
  - Development of collections and worksets
  - User needs assessment
- Community Engagement
  - Member libraries
  - TRAIL, PEGI, FDLP

Questions?

CELEBRATING TEN YEARS • 2008-2018 •
HATHI
TRUST
CELEBRATING TEN YEARS • 2008-2018 •

Coming Attractions: Afternoon Discussion

Günter Waibel

LUNCH - 60 minutes

Focus: Our Users and Discovery

Table Activity: Delve In, Take Note, Share
Jessica Rohr, Facilitator

Table Notes

https://bit.ly/2AuejWU

Understanding Users & Discovery: Activity Goals

To discuss member-driven objectives with the goal to **identify a list of member-recommended actions, investments, or activities**.

Member answers to discussion questions will **guide the development of more refined initiatives and help guide conversations about prioritization**.

Activity Outline

Each table

- Has table notes in Google Docs, a facilitator, and a notetaker.
- Will discuss 2 strategic objectives, 1 related to *users* and 1 related to *discovery*

Part I: Delving In — 20 minutes

- Discuss questions and answers by all table members
- Notetaker will capture notes in that section of notes document

Part II: Taking Note — 10 minutes

- As a table select 2 of the most interesting ideas, initiatives, activities etc.

Part III: Sharing — 10 minutes

- Report out (alternating tables odd/even)

REPEAT for second objective….

Strategic Directions 2019-2023: Enhance

HathiTrust bolsters content discovery and diversifies uses of the scholarly record through services that improve the quality of the collection and the user experience.

Objective: Understanding Users DISCUSSION

Invest in understanding the needs of users of HathiTrust and of our members.

20 minutes

- Introductions--who are you and what do you do?
- Who are the primary HathiTrust users on your campus?
- Who are the users we aren't serving that we should be serving?
- What actions can you take to help HathiTrust understand your users' needs?
- What actions could you take to connect your users to HathiTrust?

Objective: Understanding Users PRIORITIZING

Invest in understanding the needs of users of HathiTrust and of our members.

10 minutes

- Based on the discussion of the previous 20 minutes, what are the two most important actions that can be taken to connect users to HathiTrust?

Report Out
10 minutes

Objective: Discovery DISCUSSION

Redefine our discovery and access approaches to support greater interoperability and frictionless use of our collection.

20 minutes

- What are the barriers to discovery and access to HathiTrust collections on your campus?
- What trends, developments, and/or emerging services relating to discovery are you tracking and we should be attentive to?
- What could HathiTrust do to improve your users ability to discover, find, and use HathiTrust?

Objective: Discovery PRIORITIZING

Redefine our discovery and access approaches to support greater interoperability and frictionless use of our collection.

10 minutes

- Based on the discussion of the previous 20 minutes, what are the two most important ideas or suggestions you want HathiTrust to consider to support this objective?

Report Out
10 minutes

AFTERNOON BREAK
15 minutes

NEXT: Lightning Talks from Members

Lightning Talks: Looking Back, Looking Ahead

Lightning Talk Presenters

All collections are local, Part 1: Books
Michelle Paolillo, Digital Curation Services Lead, Cornell University

All collections are local, Part 2: Pages
Marie Concannon, Head, Government Information & Data Archives, University of Missouri

Can't Fight This Ingest: UIUC "Speedwagon" Tools for HathiTrust Packaging
William Schlaack, Digital Reformatting Coordinator, U of IL at Urbana-Champaign

The 'Reading Chicago Reading' Project: Successes and Challenges Using HathiTrust for In-Copyright Corpora
Ana Lucic (Library); John Shanahan (English), DePaul University

All collections are local pt. 1: books

a story from Cornell University

History fosters certain expectations…

search

advanced search

browse

MOA about

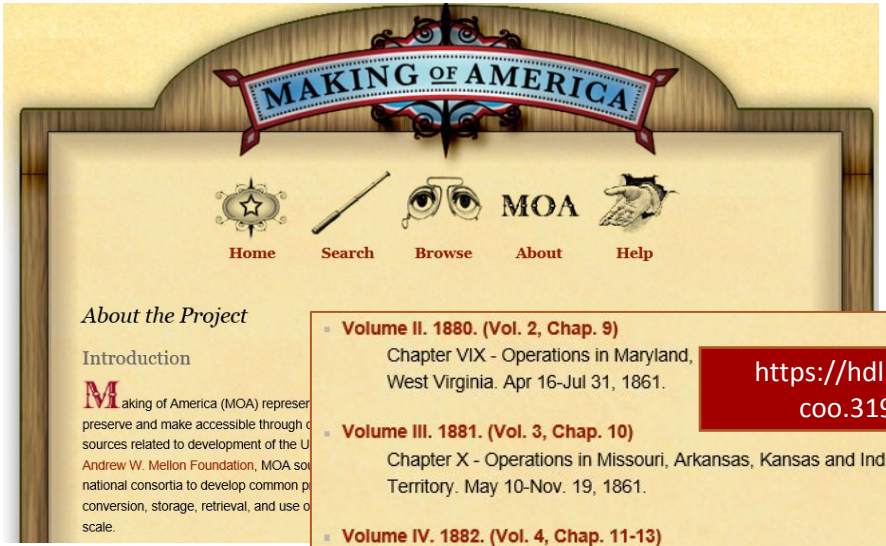
help

contact

Materials accessible here are Cornell University Library's contributions to Making of America (MOA), a digital library of primary sources in American social history from the antebellum period through reconstruction. The collection is particularly strong in the subject areas of education, psychology, American history, sociology, religion, and science and technology. This site provides access to 267 monograph volumes and over 100,000 journal articles with 19th century imprints. The project represents a major collaborative endeavor in preservation and electronic access to historical texts.

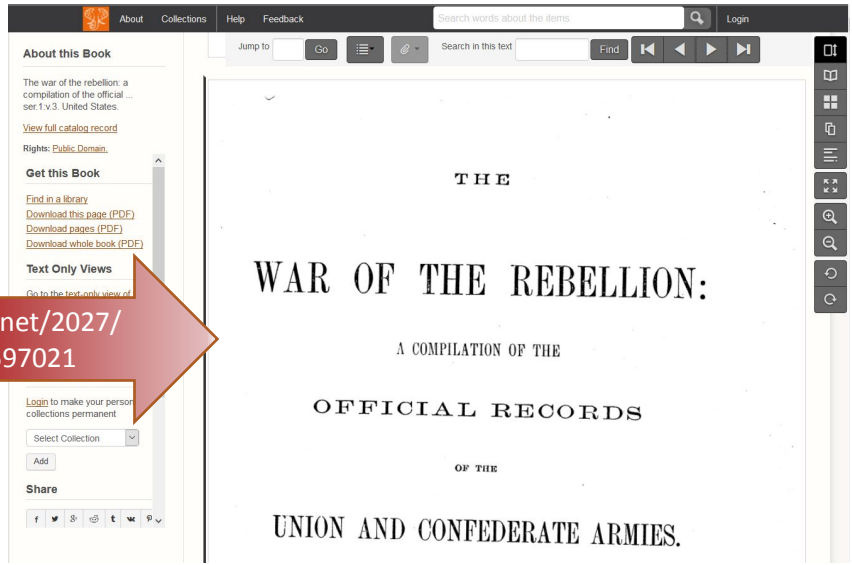
1995 2008 Migrated 2017

…a collaborative solution…



- **Volume II. 1880. (Vol. 2, Chap. 9)**
  Chapter VIX - Operations in Maryland, West Virginia. Apr 16-Jul 31, 1861.
- **Volume III. 1881. (Vol. 3, Chap. 10)**
  Chapter X - Operations in Missouri, Arkansas, Kansas and Indian Territory. May 10-Nov. 19, 1861.
- **Volume IV. 1882. (Vol. 4, Chap. 11-13)**
  Chapter XI - Operations in Texas, New Mexico, and Arizona. Jun 11, 1861-Feb. 1, 1862.
  Chapter XII - Operations in Kentucky and Tennessee. Jul 1-Nov. 19, 1861.
  Chapter XIII - Operations in North Carolina and Southeastern Virginia. Aug. 1, 1861-Jan. 11, 1862.

https://hdl.handle.net/2027/coo.31924079597021



…measuring success!

Source	Users	New Users	Sessions	Bounce Rate	Pages / Session	Avg. Session Du
Cornell MOA Collection	11,509 % of Total: 0.30% (3,833,158)	7,233 % of Total: 0.20% (3,617,837)	24,577 % of Total: 0.31% (8,020,907)	0.95% Avg for View: 3.22% (-70.53%)	51.80 Avg for View: 24.86 (108.39%)	00:23:00 Avg for View: 00:10:01 (129.77%)
1. collections.library.cornell.edu	**10,109** (86.26%)	6,558 (90.67%)	21,903 (89.12%)	1.06%	50.50	00:22:45
2. onlinebooks.library.upenn.edu	**492** (4.20%)	181 (2.50%)	778 (3.17%)	0.00%	55.04	00:19:44
3. babel.hathitrust.org	**156** (1.33%)	0 (0.00%)	389 (1.58%)	0.00%	152.72	00:43:45

~86% of users

session avg 23 min

ALL COLLECTIONS ARE LOCAL
Part 2: Pages

Generate sky-high statistics by capitalizing on page-level links in HathiTrust

Libraries
University of Missouri

Marie Concannon
Head, Government Information & Data Archives
concannonm@Missouri.edu

Libraries
University of Missouri

Prices and Wages by Decade: 1860-1869

Links to U. S. government documents listing retail prices for typical consumer purchases, and wages for common occupations. All data is for the United States unless specifically indicated.

Introduction | 1700s | 1800s | 1900s | 2000s

Wages, 1860-1869

- Wages for four common occupations in 1860
  The 1860 Census showed average wages for a farm hand, a day laborer and a carpenter (all with and without board), and to a female domestic worker with board.
  more...
- Confederate Army wages
  Chart published in the *Confederate States Almanac* (1862) shows the monthly pay of officers and privates in the regular service of Confederate states. Breaks out pay for generals, colonels, lieutenants, majors, sergeants, captains, corporals, artificers, musicians, chaplains and privates.
- Union Army officer pay, 1861-1862
  Army pay tables begin with the highest officers listed first. Page forward to find colonels, majors, captains, lieutenants, etc.
- Military pay - Civil War (Union)
  Military pay tables are cited in an 1866 book titled *War Claimant's Guide* by George W. Raff (not a government document). The book's purpose was to help people apply to the U.S. government for war-related financial reimbursement to which they were legally entitled. See Navy tables beginning on page 197, and Marine Corps officer pay on page 203. Pay for other branches of the military are elsewhere in the same volume.
- History of wages in the United States from Colonial times to 1928
  Extensive tables show average wages paid each year for a variety of occupations, reported by state and sometimes by city, often broken out by sex, in the following industries:
  - Building trades, 1840-1900
    Listed wages for bricklayers, carpenters, electricians, stationary engineers, stationary firemen, hod carriers, building laborers, marble cutters, stone masons, painters, plasterers, plumbers, stone cutters, granite cutters and tile layers
  - Clothing manufacture, 1851-1898
    Listed wages for dressmak[illegible] and sewing m[illegible]
  - Farm labor, 1841-189[illegible]
    [illegible]further time se[illegible]

Prices, 1860-1869

PRICES for FOOD and SUPPLIES

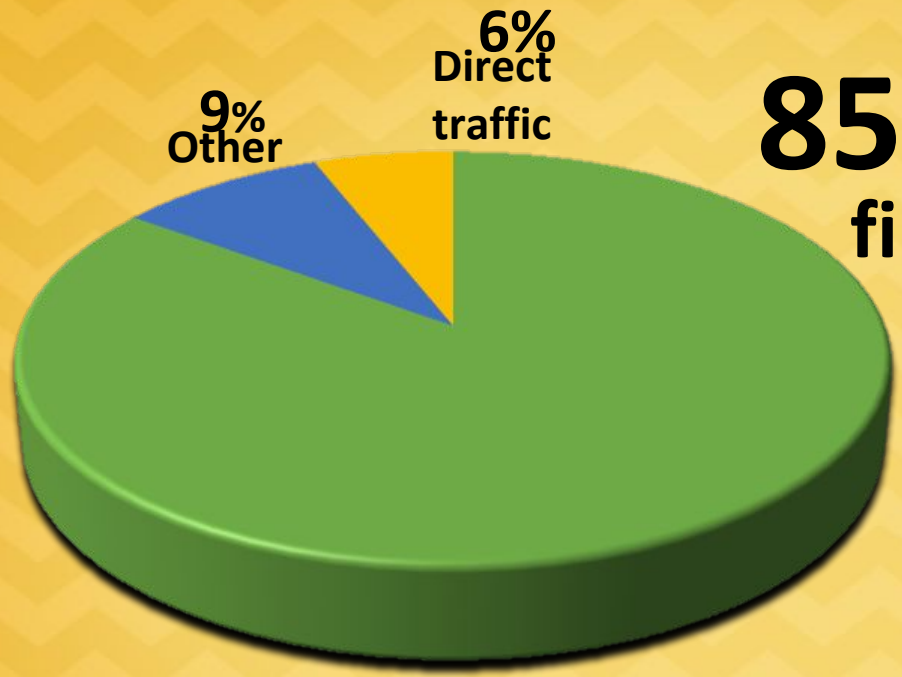
- Prices of food and dry goods by region, 1869-1870
  Average retail prices of food, cloth and other leading articles of consumption. Source: *Special Report on Immigration* published by the U.S. Treasury's Statistics Bureau.
- Retail prices by town, 1851-1880
  Shows changing retail prices over time in about 65 towns for items such as fabrics, men's heavy boots, coffee, tea, sugar, molasses, soap and starch. Flours, meats, fish, cheese, butter, beans, eggs, rice, milk and other provisions are in a separate section. Source: 1890 Census.
  more...
- Retail prices before and after the Civil War
  Comparative cost of provisions, groceries, domestic dry goods, house rent, etc. States incl[illegible] ME, NH, VT, MA, RI, CT, NY, PA, NJ, DE. Dates compared: 1860-61 and 1867-68. Source: from the Special Commissioner of the Revenue.
- Restaurant food and drink prices, 1860s
  Searchable, scanned images of about 20 menus from the 1860s. It can be useful for spot-checking individual restaurant prices.
- Board rates by city for men and women, 1851-1880
  Cost to regularly take meals at boarding establishments in selected [illegible] IN, IA, KS, KY, MO, NY, OH, PA and WV.

HEALTH COSTS

- Price of ar[illegible]al limbs - 18[illegible]
  Congre[illegible]priated [illegible]p U[illegible]
  price [illegible]

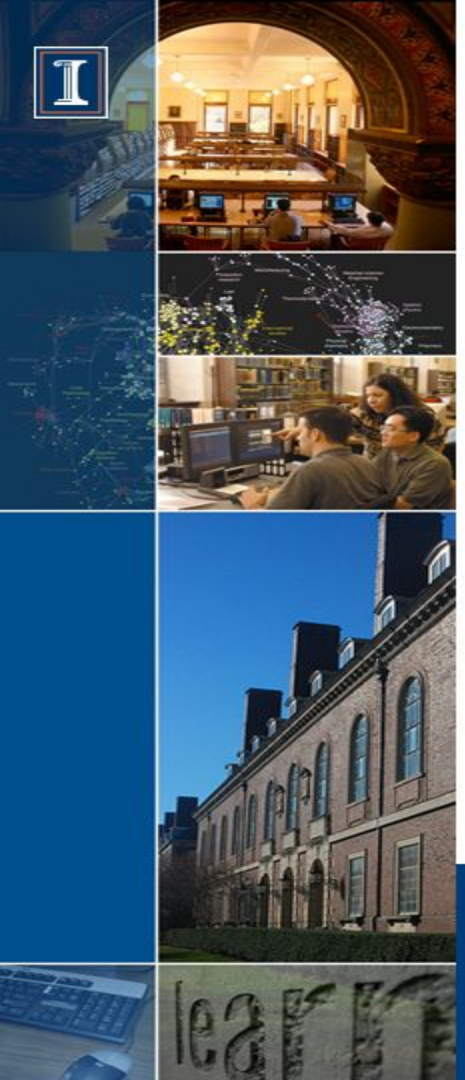
One million hits from 167 countries in 15 months!

Libraries
University of Missouri
Title expresses intent clearly
Prices and Wages by Decade: 1860-1869
Links to U. S. government documents listing retail prices for typical consumer purchases, and wages for common occupations. All data is for the United States unless specifically indicated.
Introduction
1700s
1800s
1900s
2000s
Wages, 1860-1869
Wages for four common occupations in 1860
The 1860 Census showed average wages for a farm hand
and without board), and to a female domestic worker with
more...
Confederate Army wages
Chart published in the Confederate States Almanac (1862) shows the monthly pay of officers and privates in the regular service of Confederate states. Breaks out pay for generals, colonels, lieutenants, majors, sergeants, captains, corporals, artificers, musicians, chaplains and privates.
Union Army officer pay, 1861-1862
Army pay tables begin with the highest officers listed first. Page forward to find colonels, majors, captains, lieutenants, etc.
Military pay - Civil War (Union)
Military pay tables are cited in an 1866 book titled War Claimant's Guide by George W. Raff (not a government document). The book's purpose was to help people apply to the U.S. government for war-related financial reimbursement to which they were legally entitled. See Navy tables beginning on page 197, and Marine Corps officer pay on page 203. Pay for other branches of the military are elsewhere in the same volume.
History of wages in the United States from Col
Extensive tables show average wages paid each year for
cupations, reported by
state and sometimes by city, often broken out by sex, in
lowing industries:
Building trades, 1840-1900
Listed wages for bricklayers, carpenters, electricians, stationary engineers, stationary firemen, hod carriers, building laborers, marble cutters, stone masons, painters, plasterers, plumbers, stone cutters, granite cutters and tile layers
Clothing manufacture, 1851-1898
Listed wages for dressmal
and sewing m
Farm labor, 1841-189
further time se
Easy navigation by decade
Annotations optimize discovery
Page level links
Prices, 1860-1869
PRICES for FOOD and SUPPLIES
1869-1870
other leading articles of consumption. Source: Special Report on Immigration published by the U.S. Treasury's Statistics Bureau.
Retail prices by town, 1851-1880
Shows changing retail prices over time in about 65 towns for items such as fabrics, men's heavy boots, coffee, tea, sugar, molasses, soap and starch. Flours, meats, fish, cheese, butter, beans, eggs, rice, milk and other provisions are in a separate section. Source: 1890 Census.
mor
ME, NH, VT, MA, RI, CT, NY, PA, NJ, DE. Dates compared: 1860-61 and 1867-68. Source: from the Special Commissioner of the Revenue.
Restaurant food and drink prices, 1860s
Searchable, scanned images of about 20 menus from the 1860s. It can be useful for spot-checking individual restaurant prices.
city for men and women, 1851-1880
Cost to regularly take meals at boarding establishments in selected t
IN, IA, KS, KY, MO, NY, OH, PA and WV.
HEALTH COSTS
Price of ar
al limbs - 18
Congre
priated
lp U
price
One million hits from 167 countries in 15 months!



85% of visitors find us by Googling a question in natural language

Prices and Wages by Decade,
https://libraryguides.missouri.edu/pricesandwages
University of Missouri Library



Can’t Fight This Ingest: UIUC “Speedwagon” Tools for HathiTrust Packaging

Henry Borchers – Digital Library Technical Coordinator
William Schlaack – Digital Reformatting Coordinator
Angela Waarala – Digital Projects Coordinator

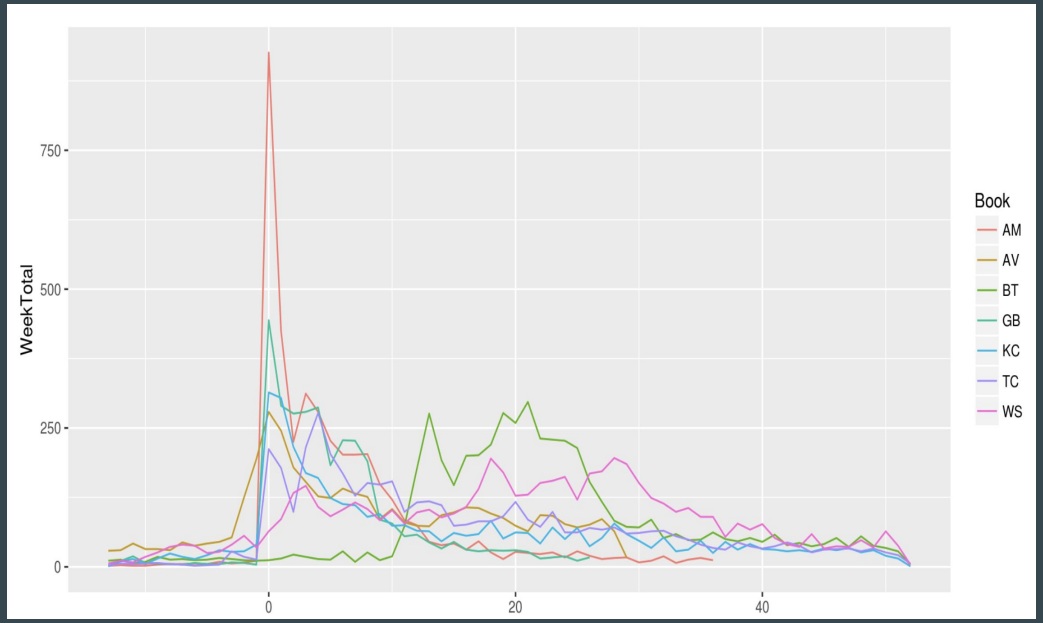
UNIVERSITY LIBRARY
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

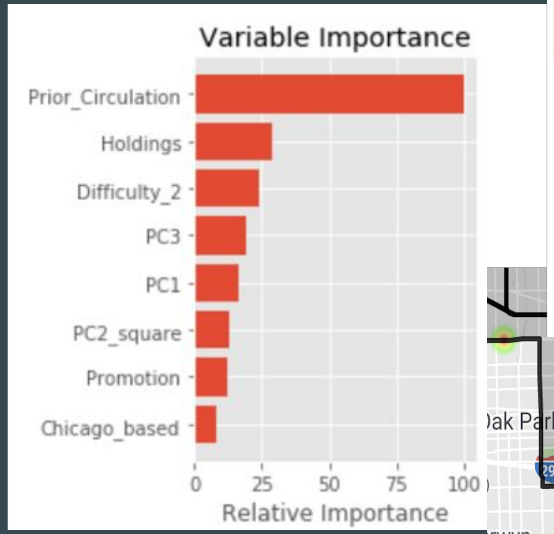
The ‘Reading Chicago Reading’ Project: Successes and Challenges Using HathiTrust for In-Copyright Corpora

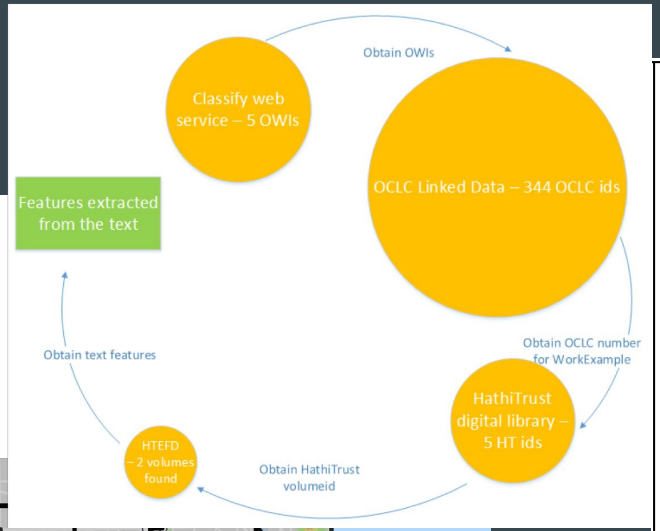
Co-PIs:
Robin Burke (Computing)
Ana Lucic (Library)
John Shanahan (English)

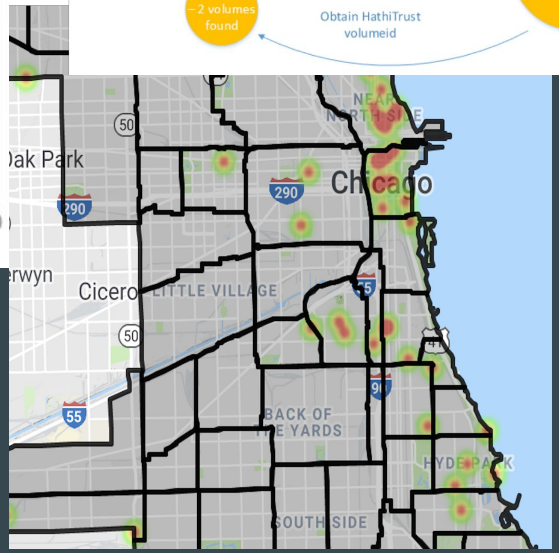
DePaul University

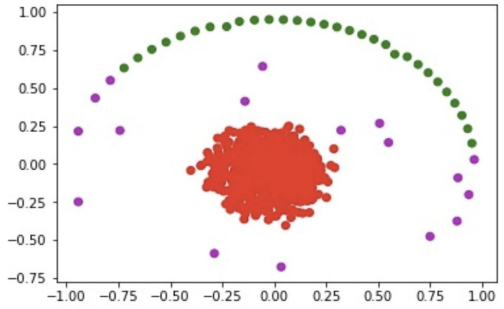
Reading Chicago Reading
DePaul University



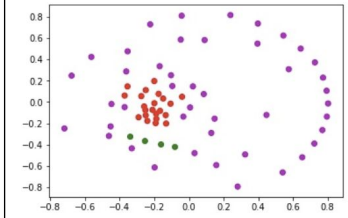








The House of Mirth by Edith Wharton
https://babel.hathitrust.org/cgi/pt?id=aeu.ark:/13960/t77s91d3g



The Rhyme of the Ancient Mariner by Samuel Coleridge
https://babel.hathitrust.org/cgi/pt?id=coo1.ark:/13960/t7cr6ch7n

John Shanahan
jshanah1@depaul.edu

Code on our blog:
https://dh.depaul.press/reading-chicago/

Questions?

CELEBRATING TEN YEARS • 2008-2018 •
HATHI
TRUST
CELEBRATING TEN YEARS • 2008-2018 •

Closing Remarks

Mike Furlough

Thoughts…

CELEBRATING TEN YEARS • 2008-2018 •
HATHI
TRUST
CELEBRATING TEN YEARS • 2008-2018 •

Actions you can take…

Tell us about today

Vote on November 6!

Vote (on the HathiTrust budget) by November 12!

Ask us anything

In the coming weeks, look for….

A request for your evaluation of this meeting

A brief report from this meeting

Details on proposals regd the fee model and member criteria

- Registration information and dates for webcasts (Nov-Jan)
- More details on how it will work
- Opportunities for 1-1 conversations
- Information about a scheduled vote

A webcast on metadata sharing policy and plans (early December)

Planning for next time….

Location

Theme and focus

Audience

Thank yous....

CELEBRATING TEN YEARS • 2008-2018 •
HATHI
TRUST
CELEBRATING TEN YEARS • 2008-2018 •

Travel safely!

HATHI TRUST

celebrating ten years