Digitization and Aggregation
Enabling a Print Network

Mike Furlough
Executive Director, HathiTrust
15 August 2016
2016 IFLA World Library and Information Congress
Columbus, OH USA
Would Shakespeare ride a bicycle? All answers in this 1896 @hathitrust pamphlet hdl.handle.net/2027/udel.3174... ht @AngelinaNZ

wow, @hathitrust provides us with this outlandishly lovely gem: a coloring book for dentists! bit.ly/1AnHlo

I found one of my key documents digitized by @hathitrust from @npl - that's saved me a lot of research time and train fare!

Love that HathiTrust is an integral part of UC’s plan. #gpodc15

All 754 vols of the NUC Pre-1956 Imprints are available online via HathiTrust, @ALCTS: on.fb.me/1NRD1Ko

Hathitrust: an amazing digital library and resource collection: hathitrust.org #hstr3912 #hstr400 #hc12181

Happy to see a fully digitised version of my dissertation online, thanks to google + Hathitrust hdl.handle.net/2027/uc1.b2714...

To see how somethings never change, read John Cotton Dana's "Library Problems" hdl.handle.net/2027/njp.32101... Thank you @hathitrust.

This historian loves tech! Thanks @internetarchive and @hathitrust who saved my paper from death by early closing hours. #yesallmedievalists
HathiTrust in August 2016

• 14.66 million total items
  – 7.3 million book titles
  – 402,000 serial titles
  – 746,00 US federal government documents

• 5.66 million items open for reading
  – Includes public domain & CC-licensed items

The collection includes (mostly) published materials in bound form, (mostly) digitized from library collections.
HathiTrust Titles by Language August 2016

- English, 49.27%
- German, 8.92%
- French, 7.04%
- Spanish, 6.78%
- Russian, 3.31%
- Chinese, 3.20%
- Italian, 2.51%
- Japanese, 2.89%
- Arabic, 1.58%
- All others, 12.87%
HathiTrust Volumes by View/Rights Status August 2016

- FULL / Creative Commons, 0.20%
- FULL-US Federal Docs, 5.09%
- FULL / PD US, 13.97%
- FULL / PD Worldwide, 19.02%
- LIMITED VIEW / IC, 61.72%
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 110 research libraries partnering to develop its programs.
• A range of transformative programs enabled by working at a very large scale.
HathiTrust Members

Allegheny College
American University of Beirut
Arizona State University
Auburn University
Baylor University
Boston College
Boston University
Brandeis University
Brown University
Bryn Mawr College
Bucknell University
Carnegie Mellon University
Case Western Reserve
Colby College
Columbia University
Cornell University
Dartmouth College
Duke University
Emory University
Getty Research Institute
Georgetown University
Georgia Tech
Harvard University Library
Haverford College
Indiana University
Iowa State University
Johns Hopkins University
Kansas State University
Lafayette College
Library of Congress
Massachusetts Institute of Technology
McGill University
Michigan State University
Montana State University
Mount Holyoke College
New Mexico State University
New York Public Library
New York University
No Carolina Central University
North Carolina State University
Northeastern University
Northwestern University
Oklahoma State University
The Ohio State University
The Pennsylvania State University
Princeton University
Purdue University
Rutgers University
Smith College
Stanford University
State University System of Florida
Swarthmore College
Syracuse University
Temple University
Texas A&M University
Texas Christian University
Texas Tech University
Tufts University
Universidad Complutense de Madrid
University of Alabama
University of Alberta
University of Arizona
University of British Columbia
University of Calgary
University of California
Berkeley
Davis
Irvine
Los Angeles
Merced
Riverside
San Diego
San Francisco
Santa Barbara
Santa Cruz
California Digital Library
The University of Chicago
University of Connecticut
University of Delaware
University of Houston
University of Illinois
University of Illinois at Chicago
The University of Iowa
University of Kansas
University of Maine
University of Maryland
University of Massachusetts, Amherst
University of Miami
University of Michigan
University of Minnesota
University of Missouri
University of Nebraska-Lincoln
University of Nevada-Las Vegas
University of New Mexico
The University of North Carolina at Chapel Hill
University of Notre Dame
University of Oklahoma
University of Pennsylvania
University of Pittsburgh
University of Queensland
University of Rochester
University of Tennessee, Knoxville
University of Texas
University of Utah
University of Vermont
University of Virginia
University of Washington
University of Wisconsin-Madison
University of Wyoming
Utah State University
Vanderbilt University
Virginia Tech
Wake Forest University
Washington University
Washington State University
Wesleyan University
Williams College
Wesleyan University
Yale University Library
Many North American Collective Print Efforts*

- WEST
- EAST
- MI-SPI
- CI-CCI
- VIVA
- OCUL
- Five Colleges
- ReCAP
- CIC
- FLARE
- COPPUL
- Connect New York
- ASERL
- WRLC

Not complete, and not including substantial consolidated holdings of single universities, e.g., Harvard, University of California
<table>
<thead>
<tr>
<th>Contributor</th>
<th>Volumes</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Michigan</td>
<td>4,713,370</td>
<td>32.14%</td>
</tr>
<tr>
<td>University of California</td>
<td>3,809,278</td>
<td>25.98%</td>
</tr>
<tr>
<td>Harvard University</td>
<td>841,968</td>
<td>5.74%</td>
</tr>
<tr>
<td>Cornell University</td>
<td>585,027</td>
<td>3.99%</td>
</tr>
<tr>
<td>University of Wisconsin - Madison</td>
<td>561,945</td>
<td>3.83%</td>
</tr>
<tr>
<td>Indiana University</td>
<td>530,756</td>
<td>3.62%</td>
</tr>
<tr>
<td>University of Illinois at Urbana-Champaign</td>
<td>507,428</td>
<td>3.46%</td>
</tr>
<tr>
<td>University of Minnesota</td>
<td>487,376</td>
<td>3.32%</td>
</tr>
<tr>
<td>The University of Texas</td>
<td>460,137</td>
<td>3.14%</td>
</tr>
<tr>
<td>Pennsylvania State University</td>
<td>390,338</td>
<td>2.66%</td>
</tr>
<tr>
<td>New York Public Library</td>
<td>319,481</td>
<td>2.18%</td>
</tr>
<tr>
<td>Princeton University</td>
<td>253,427</td>
<td>1.73%</td>
</tr>
<tr>
<td>The Ohio State University</td>
<td>153,565</td>
<td>1.05%</td>
</tr>
<tr>
<td>University of Chicago</td>
<td>128,404</td>
<td>0.88%</td>
</tr>
<tr>
<td>Universidad Complutense de Madrid</td>
<td>120,543</td>
<td>0.82%</td>
</tr>
<tr>
<td>Library of Congress</td>
<td>108,892</td>
<td>0.74%</td>
</tr>
<tr>
<td>Keio University</td>
<td>92,500</td>
<td>0.63%</td>
</tr>
<tr>
<td>Columbia University</td>
<td>77,480</td>
<td>0.53%</td>
</tr>
<tr>
<td>University of Alberta</td>
<td>76,114</td>
<td>0.52%</td>
</tr>
</tbody>
</table>
UIUC’s titles overlap with HathiTrust is 32.13%
Shared Print Monographs Program

• Focus
  – Ensure preservation of print and digital collections by linking the two
  – Catalyze national/continental collective management of collections

• Original proposal, task force charge, & preliminary recommendations:
  – https://www.hathitrust.org/print_monograph_archiving
Background of the HathiTrust Shared Print Program

2011: Constitutional Convention “Distributed Print Monographs Archive Planning Task Force” approved

2014: HathiTrust Print Monographs Archive Planning Task Force begins work

2015: Task Force planning report issued (March) and approved by PSC and Board (June)

2016: Lizanne Payne hired as Shared Print Program Officer
Goals of the HathiTrust Shared Print Program

*From the final report of the HathiTrust Print Monographs Archive Planning Task Force (March 2015)*

- Commit to retain print holdings that mirror book titles in the HathiTrust digital collection
- Maintain a lendable print collection distributed among HathiTrust member collections
- Reflect support by and provide benefits to all HathiTrust members (not a subset)
- Build on existing shared print and resource-sharing arrangements, avoid disturbing members’ other affiliations
Proposed approach: Build up to full coverage with early voluntary commitments

*Adapted from the final report of the HathiTrust Print Monographs Archive Planning Task Force (March 2015)*

- All HathiTrust institutions (~140)
  - Retention libraries (10 – 30)
- All HathiTrust book titles ~7,300,000
  - Retention titles ~3,650,000

Ask each participant to retain at least 25% of their matching holdings

Seek to match 50% of the HathiTrust digital collection in Phase 1
ARL Book Collection

HathiTrust Monograph Titles
Thank You!