HathiTrust and Shared Print Agreements in North America

Lizanne Payne
HathiTrust Shared Print Program Officer
elpayne@hathitrust.org
https://www.hathitrust.org/shared_print_program
Shared Storage vs Shared Print

Shared Storage Facility
• Libraries share space and services in a shelving facility

Shared Print Collection
• Libraries agree to retain and provide access to certain holdings for a specified time period

Explicit retention commitment is key!
## Largest North American Shared Print Monograph Programs

<table>
<thead>
<tr>
<th>Group</th>
<th>Country</th>
<th>Committed Volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Collections &amp; Preservation Consortium (ReCAP)</td>
<td>USA</td>
<td>11 million</td>
</tr>
<tr>
<td>Eastern Academic Scholars’ Trust (EAST)</td>
<td>USA</td>
<td>8.6 million</td>
</tr>
<tr>
<td>Virtual Library of Virginia (VIVA)</td>
<td>USA</td>
<td>3.5 million</td>
</tr>
<tr>
<td>Washington Research Library Consortium (WRLC)</td>
<td>USA</td>
<td>2.9 million</td>
</tr>
<tr>
<td>Academic Libraries of Indiana</td>
<td>USA</td>
<td>1.7 million</td>
</tr>
<tr>
<td>Maine Shared Collections Cooperative</td>
<td>USA</td>
<td>1.4 million</td>
</tr>
<tr>
<td>Statewide California Electronic Library Cooperative (SCELC)</td>
<td>USA</td>
<td>1.2 million</td>
</tr>
<tr>
<td>Council of Prairie &amp; Pacific University Libraries (COPPUL)</td>
<td>Canada</td>
<td>1.2 million</td>
</tr>
<tr>
<td>Connect New York</td>
<td>USA</td>
<td>850,000</td>
</tr>
<tr>
<td>Tri-Universities Group (TUG)</td>
<td>Canada</td>
<td>340,000</td>
</tr>
</tbody>
</table>
Where does HathiTrust Fit?

HathiTrust is
• a digital library enabling broad access
• a trusted digital preservation service
• an organization with over 130 research library partners worldwide

Digitization Source

- 16.09 million digitized volumes
- 7.87 million book titles
- 435,000 serial titles

6.02 million volumes open for reading, includes public domain & CC-licensed items

Google 95%
Internet Archive 3%
Member Digitized 2%
HathiTrust Shared Print Program

HathiTrust is establishing a distributed shared print collection that mirrors book titles in the HathiTrust digital library

- Not looking at member library collections as a whole
- Securing print preservation to match digital preservation (not attempting to preserve the actual copy that was digitized)
HathiTrust Shared Print Phase 1
Summer 2016 – December 2017 (18 months)

- 49 libraries
- 4.8 million monograph titles
- 16 million monograph volumes
- Monographs matching HathiTrust
- 25-year retention commitment
- No validation required
- Lend with local ILL
HathiTrust Shared Print Retention Libraries

<table>
<thead>
<tr>
<th>Arizona State University</th>
<th>New York Public Library</th>
<th>University of Florida</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brandeis University</td>
<td>Northwestern University</td>
<td>University of Illinois, Urbana-Champaign</td>
</tr>
<tr>
<td>Brown University</td>
<td>Ohio State University</td>
<td>University of Iowa</td>
</tr>
<tr>
<td>Bryn Mawr College</td>
<td>Princeton University</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>Claremont Colleges</td>
<td>Swarthmore College</td>
<td>University of Minnesota</td>
</tr>
<tr>
<td>Colby College</td>
<td>Syracuse University</td>
<td>University of Missouri</td>
</tr>
<tr>
<td>Columbia University</td>
<td>Tufts University</td>
<td>University of Notre Dame</td>
</tr>
<tr>
<td>Duke University</td>
<td>University of Alberta</td>
<td>University of Pennsylvania</td>
</tr>
<tr>
<td>Emory University</td>
<td>University of California</td>
<td>University of Queensland</td>
</tr>
<tr>
<td>Georgia Tech University</td>
<td>University of California, Merced</td>
<td>University of Texas at Austin</td>
</tr>
<tr>
<td>Getty Research Institute</td>
<td>University of California, San Diego</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Harvard University</td>
<td>University of California, Santa Cruz</td>
<td>University of Washington</td>
</tr>
<tr>
<td>Indiana University</td>
<td>University of California Northern Regional Library Facility (NRLF)</td>
<td>University of Wisconsin-Madison</td>
</tr>
<tr>
<td>Iowa State University</td>
<td>University of California Southern Regional Library Facility (SRLF)</td>
<td>Washington University in St. Louis</td>
</tr>
<tr>
<td>Johns Hopkins University</td>
<td>University of Chicago</td>
<td>Yale University</td>
</tr>
<tr>
<td>Lafayette College</td>
<td>University of Delaware</td>
<td></td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McGill University</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

49 libraries out of 130+ members!

- 45 in United States
- 3 in Canada
- 1 in Australia

Average commitment = 300,000 volumes, largest commitment = 2.5 million volumes
Key Points about
HathiTrust Retention Commitments (Phase 1)

• Voluntary participation and light-weight analysis

• Required: Lendable monographs that match HathiTrust

• Encouraged:
  • Shelved in a storage facility
  • Unique or rare holdings in HathiTrust
  • Already committed to another shared print program

• No efforts to predefine desired number of copies nor to prevent duplication

• No minimum commitment for an individual library
Lightweight Process to Identify Retention Commitments

No efforts to predefine desired number of copies nor to prevent duplication

Print Holdings Match File

In-house analysis of HathiTrust member library print holdings data

Libraries determine criteria

Libraries export proposed commitments

HathiTrust created Shared Print Registry

Shared Print Commitments File
Duplication within HathiTrust Retention Commitments

No efforts to predefine desired number of copies nor to prevent duplication

- Committed by 11-18; 40'000 ; 1%
- Committed by 6-10; 860'000 ; 18%
- Committed by 2-5; 2'400'000 ; 50%
- Committed by 1; 1'500'000 ; 31%

4.8 million committed monograph titles (OCNs)
Overlap with Other Shared Print Monograph Programs

- HathiTrust only; 12'000'000
- ReCAP; 2'300'000
- UC Shared Print; 1'600'000
- EAST; 300'000
- FDLP; 111'000
- FLARE; 100'000
- COPPUL; 30'000

16 million committed volumes
HathiTrust Shared Print Phase 2
January – December 2018

Build on shared print metadata

Secure next round of retention commitments
Build on Shared Print Metadata

- Disclose retentions in OCLC WorldCat
- Build on Shared Print metadata
- Seek replacements for lost/damaged retention volumes
- Produce Retention Comparison Reports
Identify Next-Round Retention Commitments

Consider overlap, desired number of copies, possibly other commitments

- Identify not-committed HT digital monographs
- Select provider of collection analysis services
- Analyze matching library holdings
- Secure retention commitments from HT libraries
How Many HathiTrust Digital Monographs Remain to be Committed?

- **Titles with shared print commitment;**
  - 3'200'000; 40%
- **Titles without shared print commitment;**
  - 4'800'000; 60%

3.2 million *not-committed* monograph OCNs held by 135 HathiTrust libraries
Planning Issues for HathiTrust Phase 2

- Consider a “preservation commitment” that would not require lending physical volumes
- Consider the desired number of copies to retain across HathiTrust and other programs
- Consider reducing existing HathiTrust commitments where many copies are retained
Shared Print Agreements: Where are we now?

- Overlapping regional groups
- Inconsistent retention periods
- Limited visibility of retention commitments
- Accidental redundancy and gaps
- Unknown status and condition
Shared Print Agreements: Where do we want to be?

- Discoverable holdings
- Coordinated preservation priorities
- Feasible validation
- Specialized lending
Common Registry of Discoverable Holdings

• Make shared print commitments visible to other libraries

• OCLC is developing a “Shared Print Registration Service” for monographs:
  • Library uploads minimal identifying data
  • OCLC generates shared print LHRs in WorldCat and exports records for local ILS update
  • Expected spring 2018
# Early Steps Toward Federation in North America

<table>
<thead>
<tr>
<th>Center for Research Libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Agenda for Shared Print 2017 – 2026: “shift the burden of preserving the critical corpus of published serial literature from individual libraries to the collective”</td>
</tr>
<tr>
<td>• <a href="https://www.crl.edu/blogs/crls-agenda-shared-print-2017-2026">https://www.crl.edu/blogs/crls-agenda-shared-print-2017-2026</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rosemont Shared Print Alliance</th>
</tr>
</thead>
<tbody>
<tr>
<td>• “Collaboration of several regional print journal archiving programs to coordinate efforts on a larger scale”</td>
</tr>
<tr>
<td>• <a href="https://www.rosemontsharedprintalliance.org">https://www.rosemontsharedprintalliance.org</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EAST Shared Print Monograph Summit</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Meeting of key stakeholders to “consider how to ensure a robust future for shared print monographs” (April 2018)</td>
</tr>
<tr>
<td>• <a href="https://eastlibraries.org/special-highlights/shared-print-monograph-summit">https://eastlibraries.org/special-highlights/shared-print-monograph-summit</a></td>
</tr>
</tbody>
</table>
Thank you!