Connecting Print to Digital
Using the Collection Comparison Tool

CDL, CRL, HathiTrust Collaboration for Shared Print Infrastructure

December 10, 2020
CDL, CRL & HT Collaboration for Shared Print Infrastructure

Vision

The future of print collections is shared. **Embedding shared print within the lifecycle of library collections promotes equity of access, enriches the scholarly record and increases opportunities for research and teaching.** A well-developed collaborative and interoperable infrastructure ensures we will realize the full potential of our networks and their collective collections.

Mission

The California Digital Library, the Center for Research Libraries, and HathiTrust are committed to shared print’s integration into the scholarly ecosystem by **developing shared, interoperable infrastructure**. We see this collaboration as an opportunity to **shift the shared print paradigm so that the work is not tangential to the traditional collection, but fully integrated into the life-cycle of collections** (from acquisitions to discovery and resource-sharing). The work we are facilitating will be guided by our principles and centered on our assumptions.
Comprehensive actionable data + Equitable access to effective tools + Collective action at scale

Essential pieces for building a comprehensive collective collection
Serials
12.1 Titles

Journals
5.5 Million

PAPR2, Data in Hand, Malpas-Wood
Serials
12.1 Titles

Journals
5.5 Million

Archived
167,000
What we know

- Publication/duplication ratio uneven across disciplines
- Indeterminate disciplinary distribution
- Full corpus not reflected in the archived
Comprehensive actionable data

- Freely available
- Responsibly stewarded
- Actively growing

Multiple uses
HathiTrust Serials Metadata in PAPR


- Publisher: Western Pub. Co. [etc.]
- Print ISSN: 0072-8829
- Electronic ISSN:
- OCLC Number: 2115595
- Subject: Coins, American Catalogs.
- Subject code: C
- Publication history: [1st]- ed.; 1947-

### Print Holdings

<table>
<thead>
<tr>
<th>Archiving Institution</th>
<th>Program</th>
<th>Holdings Description</th>
<th>Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emory University (EMU)</td>
<td>Scholars Trust</td>
<td>1976-1977, 1979-1980</td>
<td>20351231</td>
</tr>
</tbody>
</table>

### Digital Holdings

<table>
<thead>
<tr>
<th>Program</th>
<th>Holdings Description</th>
</tr>
</thead>
</table>
Equitable access to tools

What goes in: Serials data

Compared against: Shared print retentions and trusted digital archives registered in PAPR

What comes back out: Reports on retained and unretained titles
<table>
<thead>
<tr>
<th>OCLC</th>
<th>ISSN</th>
<th>Bib ID</th>
<th>Title</th>
<th>Holdings</th>
<th>HLC</th>
<th>Archiver</th>
<th>Archived Title</th>
<th>Archived ISSN</th>
<th>Archived Holdings</th>
</tr>
</thead>
<tbody>
<tr>
<td>48809929</td>
<td></td>
<td>b1104799x</td>
<td>The People's cause; a journal of tariff reform,</td>
<td>1990-2000</td>
<td>SerGen</td>
<td>HTDL</td>
<td>&quot;The People's cause&quot;</td>
<td>v.1 (1889)</td>
<td></td>
</tr>
<tr>
<td>1715121</td>
<td></td>
<td>b123456x</td>
<td>A.M.A. archives of dermatology and syphilology</td>
<td>1953-1965</td>
<td>Med</td>
<td>HTDL</td>
<td>A.M.A. archives of dermatology and syphilology</td>
<td>v.63 1951 Jan-Jun; v.64 1951 Jul-Dec; v.63 yr.1951; v.64 yr.1951; v.65 1952 Jan-Jun; v.66 1952 Jul-Dec; v.65 yr.1952; v.66 1953 Jan-Jun; v.68 1953 Jul-Dec; v.67 yr.1953; v.68 yr.1953; v.69 1954 Jan-Jun; v.70 1954 Jul-Dec; v.69 yr.1954; v.70 yr.1954; INDEX: v.61-89 mi.+index yr.1950-59</td>
<td></td>
</tr>
</tbody>
</table>
Collective action at scale

1. What is the role of digital surrogates coupled with the work of print preservation in making collection management decisions?
2. What’s the role of metadata in collection management decisions?
3. What are other data sources (DEI)?
4. Other ideas?
HathiTrust Serials Metadata in PAPR

Distribution of Titles Across LC Class

- PAPR %
- Hathi %
HathiTrust Serials Metadata in PAPR

More than 200 countries of publication are represented

Fewer than 30 countries make up 86% of the corpus

50% of the titles published in the United States & England

HathiTrust serials have strengths in Latin and South American countries

Shared print holdings disclosed in PAPR have strengths in South Asian countries
HathiTrust Serials Metadata in PAPR

Distribution of First Date of Publication

- Blue: Papr count
- Orange: Hathi count
Collective action at scale

1. What is the role of digital surrogates coupled with the work of print preservation in making collection management decisions?
2. What’s the role of metadata in collection management decisions?
3. What are other data sources (DEI)?
4. Other ideas?
Please join us at the next webinar.

Demonstration of Collection Comparison Tool for Serials and Journals

January 13, 4 p.m. / EST, 3 p.m. CST / 1 p.m. PST

Register in advance for this meeting:
https://UCOP.zoom.us/meeting/register/tJ0tc-CgrTsjE9Mqh9gP4QVaTFOxM9T21-kT
Questions? Let us know!

Amy Wood, Center for Research Libraries - wood@crl.edu

Heather Weltin, HathiTrust - weltin@hathitrust.org

Alison Wohlers, California Digital Library - alison.wohlers@ucop.edu

Judy Dobry, California Digital Library - judy.dobry@ucop.edu