Stewarding the Monographic Collective Collection: Charting the Present

Devon Smith – OCLC
Alison Wohlers – Partnership for Shared Book Collections
origin story

• The Partnership for Shared Book Collections serves 17 shared print monograph programs including 388 libraries.

• That service includes scaled research to support:
  o evidence-based decision-making and best practices for shared print and collective collection management;
  o cost-effective retention and access to shared print monograph collections;
  o preservation and long-term accessibility of print resources; and
  o expanding the collective of shared print programs and the value they provide.
Tier 1 priority research for the Partnership

• Research questions...
  o What is the corpus of content we are looking to protect/preserve as part of Partnership programs (monographs)?
  o What is not being adequately retained (or potentially over retained)?

To answer these questions, we needed to be able to look at the broad landscape of *both retained and unretained* print book collections.
OCLC Analysis Approach

• Analysis of current registered monograph print retentions in the US and Canada
  • Refine by state/province and shared print program
  • Further refine by subject, language, decade of publication, place of publication, retention expiration date, scarcity
• Compare the landscape of print retentions against all monographic bib holdings
What we'll share today

• Share initial findings on the landscape of current registered retention commitments of monographs
  o Gaps in data and participation
  o Data anomalies

• Frame the next steps for the analysis
  o Comparing registered retention commitments against unretained holdings
  o Anticipated timeline and deliverables
Data Overview

• Who is represented in the analysis?
  – Institutions that have registered their commitments in OCLC’s Shared Print Registry

• What is the state of the data?
  – 34 million registered monographic retention commitments
  – 99% updated in last 5 years
  – Normal errors that need normal cleaning

• What challenges does the state of data present?
  – Multiple overlapping commitments complicate the analysis
    • One item committed to multiple programs
    • Multiple commitments for the same manifestation
  – Unregistered data limits the extent of the analysis
Institutions with monographic retention commitments

OCLC 482

Participating institutions

Partnership 335

OCLC 265

Partnership 118

217

Source: OCLC Shared Print Registry, January 2024
Registered monographic retention commitments
US and Canada

US: 30,536,042
CA: 3,471,572
Total: 34,007,614
Source: OCLC Shared Print Registry, January 2024
Total registered monograph commitments per state

Source: OCLC Shared Print Registry, January 2024
Total registered monograph commitments per state

Source: OCLC Shared Print Registry, January 2024
Total registered monograph commitments per state

Source: OCLC Shared Print Registry, January 2024
Total registered monograph commitments per state

Source: OCLC Shared Print Registry, January 2024
Institutions per state with registered monograph commitments

Source: OCLC Shared Print Registry, January 2024
Registered monographic retentions by program (>10000)

<table>
<thead>
<tr>
<th>Program</th>
<th>Retention Libraries</th>
<th>Total Retentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>HathiTrust</td>
<td>97</td>
<td>11,159,354</td>
</tr>
<tr>
<td>EAST</td>
<td>147</td>
<td>10,729,007</td>
</tr>
<tr>
<td>ReCAP</td>
<td>1</td>
<td>2,998,048</td>
</tr>
<tr>
<td>WRLC</td>
<td>8</td>
<td>2,643,788</td>
</tr>
<tr>
<td>SCELC</td>
<td>35</td>
<td>2,128,218</td>
</tr>
<tr>
<td>PBUQ</td>
<td>18</td>
<td>1,971,437</td>
</tr>
<tr>
<td>Florida Academic Repository</td>
<td>1</td>
<td>1,604,193</td>
</tr>
<tr>
<td>Maine SC</td>
<td>13</td>
<td>1,377,897</td>
</tr>
<tr>
<td>MI-SPI</td>
<td>4</td>
<td>596,776</td>
</tr>
<tr>
<td>FDLP</td>
<td>1</td>
<td>493,385</td>
</tr>
<tr>
<td>COPPUL SPAN</td>
<td>13</td>
<td>336,105</td>
</tr>
<tr>
<td>CNY</td>
<td>4</td>
<td>218,970</td>
</tr>
<tr>
<td>CI-CCI</td>
<td>5</td>
<td>112,875</td>
</tr>
<tr>
<td>Scholars Trust</td>
<td>4</td>
<td>72,902</td>
</tr>
<tr>
<td>CUSP</td>
<td>6</td>
<td>63,940</td>
</tr>
<tr>
<td>Macalester Scarcely Held Titles</td>
<td>1</td>
<td>16,953</td>
</tr>
</tbody>
</table>

Source: OCLC Shared Print Registry, January 2024
Timeline

• Started data engineering July 2023
• Started analyzing September 2023
• Targeting analysis completion February 2024
• Will produce a public report, mid 2024
  – Including a website with datasets
• Regular updates of analysis as registration grows
Q&A?

• Join us at the PAN Watercooler!
Thank you

Alison Wohlers – Partnership for Shared Book Collections
Devon Smith – OCLC

Because what is known must be shared.