Academic Libraries of Indiana Shared Print Project

David Lewis
April 2013

What is the Shared Print Project? The project is a study of the print book holdings of Indiana academic libraries. The study will look at circulation and holdings in both Indiana and nationally. The intent of the study is to help individual libraries identify books in their collections that have had little or no use and are wildly held, so that these items can be considered for withdrawal with reasonable assurance that access will persist. The study will also identify books that are unique or rarely held so that libraries can commit to keeping them and assure that they are held in a manner than secures and preserves them.

Expected Results.

1. The study should provide participating libraries an opportunity to weed their collections in a way that will have a minimum impact on their users. This will create opportunities to recapture space that can be repurposed for other library or campus needs.

2. By conducting the project together we can develop a statewide strategy for print books collections. We anticipate that IU Bloomington, Notre Dame, and Purdue will be resource collections. However, it may make sense for groups like PALNI or libraries in the northwest part of the state to coordinate their collections.

3. We also anticipate that the study will provide the basis for informing and influencing ongoing collection development coordination across Indiana’s academic libraries. The study will provide a rich and detailed look at the print book collections across the state that might be used as the basis for collection development going forward.

Who will participate? We know that IU Bloomington, Notre Dame, Purdue, and the IVYTech libraries will not participate. Other academic libraries can choose to participate if they wish.

How will it work? The most likely approach will be for ALI to contract with Sustainable Collection Services (SCS, see: http://sustainablecollections.com), a consulting company with considerable experience and success doing this sort of study. SCS works with bibliographic and circulation data that is extracted from library catalogs. This data is then checked against OCLC and several other
sources to establish how wildly held the item is. This data allows a library to select a set of rules to create a list of candidates for withdrawal. See the examples in the appendix for how this might work.

**What will it cost?** The costs will be between 2.5 cents and 4 cents per bibliographic record. ALI hopes to raise some external support to reduce the price to the lower of these two figures.

**What are the next steps?** Rick Lugg from SCS will give a presentation at the ALI annual meeting on May 9. Sue Ward from Purdue, which is currently conducting a study with SCS, will also speak. After the annual meeting ALI will survey members to see who is interested in participating in the project. We anticipate that by the late summer we will be looking for firm commitments from libraries. If we can keep to this schedule, it should be possible to complete the study in the 2013/14 academic year.

**How will the project be managed?** We expect that a small coordinating group made up of staff from participating libraries will work with SCS and guide the project. Individual participating libraries will need to appoint a project manager to work with the ALI coordinating group.
Appendix A: How a Collection Study Works

The following example comes from Drake University.

The bibliographic and circulation data is extracted from the OPAC and filtered to get a clean list of paper books. Various classes, for example government documents, can be included or excluded as desired.

### Collection Summary: Filtered Record Counts

<table>
<thead>
<tr>
<th></th>
<th>Title Count</th>
<th>Item Count</th>
<th>Percent of Filtered Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Bib Records</td>
<td>270,286</td>
<td>286,073</td>
<td>N/A</td>
</tr>
<tr>
<td>All Bib Records - Filtered</td>
<td>196,137</td>
<td>207,218</td>
<td>100%</td>
</tr>
</tbody>
</table>

The data can then be sorted by the amount of use various items received.

### Collection Summary: Total Uses Since 1992 (20 years)

<table>
<thead>
<tr>
<th>Uses</th>
<th>Circulation: 0 Charges</th>
<th>In-House Uses: &gt; 0</th>
<th>Total Uses: 0 (charges and in-house uses combined)</th>
<th>Total Uses: 1 (charges and in-house uses combined)</th>
<th>Total Uses: 2 (charges and in-house uses combined)</th>
<th>Total Uses: 3 (charges and in-house uses combined)</th>
<th>Total Uses: &gt; 3 (charges and in-house uses combined)</th>
<th>Total Uses: &gt; 5 (charges and in-house uses combined)</th>
<th>Circulation: last charge date &gt; 2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>138,415</td>
<td>47,394</td>
<td>129,933</td>
<td>58,916</td>
<td>35,459</td>
<td>23,472</td>
<td>78,786</td>
<td>50,610</td>
<td>92,521</td>
</tr>
<tr>
<td>4</td>
<td>145,559</td>
<td>48,542</td>
<td>139,573</td>
<td>60,463</td>
<td>35,976</td>
<td>23,730</td>
<td>80,482</td>
<td>51,592</td>
<td>97,824</td>
</tr>
<tr>
<td>5</td>
<td>43%</td>
<td>14%</td>
<td>41%</td>
<td>18%</td>
<td>11%</td>
<td>7%</td>
<td>24%</td>
<td>15%</td>
<td>29%</td>
</tr>
</tbody>
</table>
The data can be matched with OCLC and other sources to establish how rare or widely held an item is.

**Collection Summary: WorldCat Holdings Distribution**

<table>
<thead>
<tr>
<th></th>
<th>Titles</th>
<th>Items</th>
<th>Percent of Filtered Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>&gt; 100 holdings in USA - WorldCat (includes Drake's Holding)</td>
<td>266,495</td>
<td>279,752</td>
</tr>
<tr>
<td>13</td>
<td>&gt; 50 holdings in USA - WorldCat (includes Drake's Holding)</td>
<td>289,583</td>
<td>304,612</td>
</tr>
<tr>
<td>14</td>
<td>&lt; 5 holdings in USA - WorldCat (includes Drake's Holding)</td>
<td>10,178</td>
<td>10,834</td>
</tr>
<tr>
<td>15</td>
<td>&gt; 5 holdings in Iowa - WorldCat (includes Drake's Holding)</td>
<td>152,556</td>
<td>160,149</td>
</tr>
<tr>
<td>16</td>
<td>Unique in Iowa - WorldCat</td>
<td>28,062</td>
<td>30,560</td>
</tr>
</tbody>
</table>

By combining this information, various strategies can be developed and tested to create lists for withdrawal. It should be noted that these rules could include breakdowns by classification if that is desired. Multiple lists can be created for withdrawal with or without review. It might also make sense to create comparison groups like PALNI or be regions within the state.
### Examples of Withdrawal and Preservation Scenarios

<table>
<thead>
<tr>
<th>Withdrawal Candidates 1 - Published before 1990; 0 total uses; not a Drake faculty publication; and held by more than 100 US libraries.</th>
<th>Titles</th>
<th>Items</th>
<th>% of Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>65,687</td>
<td>68,902</td>
<td>20%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Withdrawal Candidates 2 - Published before 1990; fewer than 3 total uses; not a Drake faculty publication; and held by more than 100 US libraries.</th>
<th>Titles</th>
<th>Items</th>
<th>% of Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>116,939</td>
<td>122,124</td>
<td>36%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Withdrawal Candidates 3 - Published before 1990; fewer than 3 total uses; 0 charges since 2002; not a Drake faculty publication; held by more than 100 US libraries; and never reviewed in CHOICE.</th>
<th>Titles</th>
<th>Items</th>
<th>% of Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>105,672</td>
<td>110,439</td>
<td>32%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Withdrawal Candidates 4 - Published before 2002; 0 total uses; not a Drake faculty publication; held by more than 50 US libraries; held by more than 5 libraries in Iowa; and never reviewed in CHOICE.</th>
<th>Titles</th>
<th>Items</th>
<th>% of Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>34,469</td>
<td>35,914</td>
<td>11%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preservation Candidates 1 - Fewer than 5 US holdings and not in Hathi Trust</th>
<th>Titles</th>
<th>Items</th>
<th>% of Item Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>10,155</td>
<td>10,810</td>
<td>3%</td>
</tr>
</tbody>
</table>
Appendix B: Overview from 4/15/2013 Presentation to ALI Task Force

The SCS Approach: Data-Driven Decisions

- Circulation & other use data (ILL, in-house)
- Holdings in other libraries (peer, regional, national)
- Secure digital copy (Hathi)
- Authoritative title lists
- [Commercial availability]
- Library-defined rules
- Interactive rule sets

Step 1: Collect and prepare the library's data

- Bibliographic, item, circulation, and holdings data extracted, transformed, and loaded to an SCS postgres database
  - Filter out-of-scope bib records
    (eBooks, maps, scores, DVDs, Gov Docs)
  - Eliminate duplicate bib records
  - Normalize call numbers
  - Eliminate trailing spaces in control numbers
  - Validate OCLC numbers
  - Match bib records on OCLC number (with title-string check)
  - LCCN/title-string lookups for records lacking OCLC#s
  - Identify and accommodate unusual implementations of MARC
  - Map item-level data and interpret codes
Step 2: Match titles to external data sources

- OCLC WorldCat
  - Holdings: Global
  - Holdings: US
  - Holdings: State
  - Holdings: Peers/Consortial Partners
- HathiTrust
  - Public domain
  - In-Copyright
- CHOICE
Step 4: Scenario Development; Combine Factors

- Usage levels, circulation thresholds
- Print Holdings
  - Service copies
  - Archival copies
- Digital Copies
  - Hathn PD
  - Hathn IC
- Authoritative Title Lists
  - CHOICE, CHOICE DAT
- Date restrictions
  - Date acquired
  - Publication date

Step 6: Review and approve lists

- Share with selectors and stakeholders as needed
- Review options include
  - Reviewing the excel lists themselves
  - Flagging titles for in-stack review
  - Moving titles to a staging area for review
Step 7: Perform Deselection Workflows

- Post-decision tasks
  - Batch Bib record maintenance
  - Batch item and holding record maintenance
  - Physical re-processing (for transfer)
  - Disposition logistics

- Integration into ongoing operations
  - Workflow definition
  - Scheduling
  - Sustainability/Carrying Capacity definition