Applications of the QA Catalogue for Analysing Library Data

David Floyd, Chief Cataloging Librarian
Sasha Frizzell, Catalog/Metadata Management Librarian
Binghamton University
CITSIG, Core Interest Group Week, March 8, 2024
Agenda

➔ Introduction to the QA catalogue
➔ Projects we’ve completed based on QA findings
➔ Future project ideas
Web application, developed by Péter Király, that provides a visual interface for the analysis of large sets of MARC metadata.

Outputs include:

- Record completeness
- MARC coding issues
- Quantitative access point review
- Several unique analyses designed for serials metadata

### Shelf-Ready completeness

These scores are the implementation of the following paper:


The main purpose of the report is to highlight which fields of the printed and electronic book records are important when records are coming from different suppliers. 50 libraries participated in the survey, each selected which fields are important for their collection. The report listed those fields which gets the highest scores.

The current calculation based on this list of essential fields. If all data elements specified are available in the record it gets a score of 50, if only some of them, it gets a proportional score. E.g. under 250 (edition statement) there are two subfields. If both are available, it gets score 44. If only one of them, it gets the half of it, 22, and if none, it gets 0. For 1XX, 6XX, 7XX and 8XX fields, if one gets the full score of 50, the maximum score would be 28.44, which could be accessed if all the data elements are available in the record.
Can’t you get this information from your ILS?

QA tool is better at:

- Comprehensive overviews
- Addressing records and fields not indexed by the ILS

QA tool is also:

- Open-source
- Not ILS dependent
## How it works

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Collect MARC files in .mrc format</td>
<td>2</td>
</tr>
</tbody>
</table>
Projects stemming from QA Analysis

1. Vendor metadata quality assessment

Comparison of vendor metadata in the Alma Community zone, Binghamton’s Institution Zone, and Cornell University’s catalog.

Focus on if the records included classification. Additional review of subject heading quantity, quality, and source vocabularies.
Projects stemming from QA Analysis

2. Metadata enhancement project for Pleasure Reading

<table>
<thead>
<tr>
<th>path</th>
<th>label</th>
<th>records</th>
<th>occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Access</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>655 — Index Term - Genre/Form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ind1</td>
<td>Type of heading</td>
<td>12,533 (19.60%)</td>
<td>20,609 (13.64%)</td>
</tr>
<tr>
<td>$ind2</td>
<td>Thesaurus</td>
<td>12,533 (19.60%)</td>
<td>20,609 (13.64%)</td>
</tr>
<tr>
<td>$a</td>
<td>Genre/form data or focus term</td>
<td>12,533 (19.60%)</td>
<td>20,610 (13.64%)</td>
</tr>
<tr>
<td>$b</td>
<td>Non-focus term</td>
<td>2 (0.00%)</td>
<td>2 (1.00%)</td>
</tr>
<tr>
<td>$c</td>
<td>Facet/hierarchy designation</td>
<td>21 (0.03%)</td>
<td>27 (3.29%)</td>
</tr>
<tr>
<td>$v</td>
<td>Form subdivision</td>
<td>6 (0.01%)</td>
<td>7 (1.17%)</td>
</tr>
<tr>
<td>$x</td>
<td>General subdivision</td>
<td>23 (0.04%)</td>
<td>33 (1.43%)</td>
</tr>
<tr>
<td>$y</td>
<td>Chronological subdivision</td>
<td>216 (0.34%)</td>
<td>291 (1.35%)</td>
</tr>
<tr>
<td>$z</td>
<td>Geographic subdivision</td>
<td>198 (0.31%)</td>
<td>403 (1.72%)</td>
</tr>
<tr>
<td>$0</td>
<td>Authority record control number</td>
<td>281 (0.44%)</td>
<td>553 (1.97%)</td>
</tr>
<tr>
<td>$2</td>
<td>Source of term</td>
<td>11,097 (17.36%)</td>
<td>15,451 (13.39%)</td>
</tr>
<tr>
<td>$3</td>
<td>Materials specified</td>
<td>1 (0.00%)</td>
<td>4 (4.00%)</td>
</tr>
<tr>
<td>$5</td>
<td>Institution to which field applies</td>
<td>50 (0.08%)</td>
<td>77 (1.54%)</td>
</tr>
</tbody>
</table>
Projects stemming from QA Analysis

3. Enhancement of records created by Binghamton,
   First subset: language of cataloging empty
Future projects

- Enhancement of local origin records
- Evaluation of Special Collections cataloging
- Targeted authority control for faculty
- Implement into decision making workflows for large sets of vendor metadata
We have accomplished a lot already with this tool, and full disclosure, we are still at about 40% functionality. The British Library (http://141.5.102.114/bl/?tab=completeness) has a successful full implementation that shows signs of where things can go from here.
Thank you

Questions? Contact
David - DFloyd@Binghamton.edu
Sasha - SFrizzell@Binghamton.edu

QA Catalogue on GitHub
https://github.com/pkiraly/qa-catalogue