5-20-2021

‘Theses’ Going to be Good!: A How to Guide on Dealing with Large Complex Cataloging Projects

Paul Daybell  
*Utah State University*

Becky Skeen  
*Utah State University*

Liz Woolcott  
*Utah State University*

Melanie Shaw  
*Utah State University*

Seth Westernburg  
*Utah State University*

Follow this and additional works at: [https://digitalcommons.usu.edu/lib_present](https://digitalcommons.usu.edu/lib_present)

Part of the [Library and Information Science Commons](https://digitalcommons.usu.edu/lib_present)

---

**Recommended Citation**  

This Presentation is brought to you for free and open access by the Libraries at DigitalCommons@USU. It has been accepted for inclusion in Library Faculty & Staff Presentations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
‘Theses’ Going to be Good!: A How to Guide on Dealing with Large Complex Cataloging Projects
Hello!

Paul Daybell
Archival Cataloging Librarian
Utah State University, Merrill-Cazier Library
paul.daybell@usu.edu

Becky Skeen
Special Collections Cataloging Librarian
Utah State University, Merrill-Cazier Library
becky.skeen@usu.edu

Liz Woolcott
Head, Cataloging and Metadata Services
Utah State University, Merrill-Cazier Library
liz.woolcott@usu.edu

Melanie Shaw
Music, Serials, and Batch Cataloger
Utah State University, Merrill-Cazier Library
melanie.shaw@usu.edu

Seth Westenburg
Student Supervisor, Cataloging Assistant
Utah State University, Merrill-Cazier Library
seth.westenburg@usu.edu
Outline

1. Background & Purpose
2. Collection Preparation
3. Cataloging Process
4. Common Problems
5. Pros and Cons of Chosen Model
6. Lessons Learned & Next Steps
Background & Purpose
Special Collections & Archives Barcoding Project

- Multi-year project
- Cataloging/Barcoding
- 250,000 items in SCA
- Metadata added into Sierra and ArchivesSpace
SCA Theses Collection

- Date range – 1920 to present
- Approximately 22,000 print theses/dissertations
- Also available in different formats such as microform and digital
- 2 shelving systems
- Legacy copies considered archival preservation copies and not consistently cataloged (e.g. mixed format records)
Collection Preparation
Getting Started

- Pulled and reshelved entire collection
  - Merged two call number systems into one
- Organize collection according to new classification scheme
  - Year
  - Author Last Name (alphabetical)
  - Author First Name (alphabetical)
  - Title (alphabetical)
3 Cataloging Process
Our process...

- Collect and clean dataset
- Compare Shelflist
- Batch assignment
- Batch processing
- Physical processing
Collect and Clean Data

Data Collection
- Extract all existing theses records from catalog
  - All formats
  - USU theses only
- Export fields into spreadsheet
  - OCLC #
  - Material Type
  - Record Bib Number
  - 100
  - 245
  - 260 |c & 264 |c
  - 300
  - 500
  - 502
  - 533
  - 590
  - 655
  - 690

Data Clean-up/Parsing
- De-dupe Titles
  - Follow selection criteria
- Isolate OCLC #
- Split 100 field into first, last, suffix, years
- Split 245 |c
- Split 300 into pagination and illustrations.
  - Standardize both
- Split and standardize 502 into:
  - Theses type
  - Degree
  - University
  - Department
- Copy 500, 590s, and 690s into relevant columns
  - Note: often in wrong column due to export issue
Compare Shelflist

Initial Shelf Inventory
◇ Verify status of record
◇ Check the data matches the item in hand
◇ Update, as needed or add new record
◇ Barcoded the item
◇ Flag for cataloger review, as needed

Quality Control
◇ Review the work in initial pass-through to make sure it was correct
◇ Ensure new items weren’t added to the collection in the interim, add if needed
◇ Assign call number

Thesis Data Pull

<table>
<thead>
<tr>
<th>#</th>
<th>Thesis</th>
<th>Status</th>
<th>Barcode</th>
<th>Author Last</th>
<th>Author First</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>129</td>
<td>1967_Gulla</td>
<td>Found</td>
<td>3906002047098</td>
<td>Gulla</td>
<td>Amin Ismail</td>
<td>Factors influencing population growth of Tribolium brachysperm in the laboratory</td>
</tr>
<tr>
<td>130</td>
<td>1967_Haggerty</td>
<td>Found</td>
<td>39060020471024</td>
<td>Haggerty</td>
<td>Samuel</td>
<td>Effects of physical stress on gross motor performance: study of selected adaptive physical education programs</td>
</tr>
<tr>
<td>131</td>
<td>1967_Hambien</td>
<td>Found</td>
<td>39060020471081</td>
<td>Hambien</td>
<td>Harold Edward</td>
<td>An application of a direct labor control system in the research and development environment of the aerospace i...</td>
</tr>
<tr>
<td>132</td>
<td>1967_Hammond</td>
<td>Found</td>
<td>39060020471149</td>
<td>Hammond</td>
<td>LeAngia</td>
<td>The high school girl's interest in home sewing in relation to her values and creativity</td>
</tr>
<tr>
<td>133</td>
<td>1967_Hancock</td>
<td>Found</td>
<td>39060020471206</td>
<td>Hancock</td>
<td>Dennis Howard</td>
<td>A follow-up study of the high school graduates from the Cache County school district from 1956 thru 1965</td>
</tr>
<tr>
<td>134</td>
<td>1967_Hanni</td>
<td>Found</td>
<td>39060020471263</td>
<td>Hanni</td>
<td>Kenneth Max</td>
<td>Economic status, a factor that influences children’s educational achievement in the elementary grades</td>
</tr>
<tr>
<td>135</td>
<td>1967_Harris</td>
<td>Not Found</td>
<td></td>
<td>Harris</td>
<td>John Henry</td>
<td>The development of a prediction system for the occurrence of law violations on the Ogden Ranger District. Web...</td>
</tr>
<tr>
<td>136</td>
<td>1967_Harris</td>
<td>Added</td>
<td>39060020471321</td>
<td>Harris</td>
<td>Richard Wayne</td>
<td>Numerical restoration of optical objects obscured by diffraction and noise</td>
</tr>
<tr>
<td>137</td>
<td>1967_Hart</td>
<td>Found</td>
<td>39060020471388</td>
<td>Hart</td>
<td>Camile B.</td>
<td>The role of motivation in remedial reading</td>
</tr>
<tr>
<td>138</td>
<td>1967_Hart</td>
<td>Boundwith</td>
<td>39060020471446</td>
<td>Hart</td>
<td>David F.</td>
<td>How physical education helps build character; Methods of achieving a game situation in foul shot practice</td>
</tr>
<tr>
<td>139</td>
<td>1967_Hart</td>
<td>Added</td>
<td>39060020471503</td>
<td>Hart</td>
<td>Eugene Blake</td>
<td>Food-related movements and incidental observations of the cliff shipwreck. Eutamias dorsalis</td>
</tr>
</tbody>
</table>
Batch Assignment

Merge
◇ Correct format record exists in OCLC
◇ Only incorrect format record exists in catalog

Overlay
◇ Correct format record exists in OCLC
◇ Correct format record exists in catalog

Original
◇ Correct format record does not exist in OCLC
◇ Correct format record does not exist in catalog

Additional action
◇ Add an item record with barcode and call number
Batch Processing

**MERGE process**
- **Airtable**: export CSV
- **MarcEdit**: map data to MARC records
  - Add constant data & save
- **OCLC**: batch-search using 035 numbers
  - Delete 502
  - Export new file
- **MarcEdit**: merge OCLC with Brief file
- **Sierra**: import merged file into local catalog:
  - 949 creates item records

**OVERLAY process**
- **Airtable**: export CSV
- **MarcEdit**: map data to MARC records
  - (Include 907)
  - Add constant data & save
- **OCLC**: batch-search using 035
  - Delete 502
  - Export file
- **MarcEdit**: merge OCLC with Brief file
- **Sierra**: import merged file
  - 907 overlays record
  - 949 creates items

**ORIGINAL process**
- **Airtable**: export CSV
- **MarcEdit**: map data to MARC
  - Add constant data & save
  - Edit 008 & Leader
  - Troubleshoot & validate
- **OCLC**: upload to local save file & validate
  - Update & receive OCLC numbers
- **Sierra**: import file of new records into catalog
  - 949 creates items
Physical Processing

- Versatile timing
  - During QC
  - After batch processing
- Apply labels and RFID tags
- Final QC process
Common Problems
Common Problems

- Theses from other universities
- Bound with theses (2 titles bound together)
- Early theses lacking consistent title page layout
- Multiple volumes or copies
- Dual authored theses
- Non-standard theses (senior reports, honors reports, etc.)
- Cataloging records containing a different date than the one listed in the item
- Items added into collection after initial temporary organization and numbering
- Limited access to collection because of pandemic
- Student technician turnover and training of new hires
5

Pros & Cons of Chosen Model
Pros

- Much faster
  - Estimated 4.5 years to catalog if traditionally cataloged at a title-by-title level
  - This model would take about 1 year to complete
- Allows updating of all records to current standards
- Saves hand keying most of the fields
- Provides inventory for next phase of process
- Uncovered errors with other formats that can be fixed
  - Duplicate electronic records for one title
  - Microfilm/microfiche attached to print records
- Useful to have intellectual control of the collection (particularly as it is a highly used collection for digitization and ILL)
Cons

- Lots of unexpected inconsistencies
- A lot of hands = more need to re-train
  - Student labor fluctuations affects early process
- More difficult to track statistics on the process
  - Who owns the statistics for the final numbers?
- Communicating progress is harder
Lessons Learned & Next Steps
Staffing needs

- **Project team:**
  - Catalogers (Batch and Individual): 4
  - Student techs: 10 (part and/or full time at different times)
  - Data prep: 3

- Time investment comes in waves

- Planning was time intensive and involved a lot of parties (SCA, LIT, CMS) - relied on heavy knowledge of the collection from previous work with theses
Next Steps

◆ Finish the current processes

◆ Once SCA barcoding is complete, will move onto the cleanup of other physical copies in stacks, as well as microform, and electronic formats
Resources

- Step-by-step process: https://usulibrary.atlassian.net/l/c/Fv5adhog

Acknowledgements

- Abigale Rodabough
- Alisha Grant
- Anastasia Davis
- Bryn Larsen
- Isabel Hess
- Josee Butler
- Maddie Gardner
- Maddie Young
- MaKayla Roundy
- Rachel Olsen
- Sabrina Leatham
- Victoria Mendoza
Thanks!

Any questions?