You May Own It…But Can They Find It?
A Panel Discussion

Paul Moeller
University of Colorado Boulder

Yukari Sugiyama
Yale University

Jeff Siemon
Anderson University

Andrew Senior
McGill University
University of Colorado Boulder’s Use of Collection Manager

Charleston Conference 2017

Paul Moeller Director of Metadata Services
University of Colorado Boulder Libraries
CU Boulder Libraries

• Regional Federal Depository

• PCC (NACO, BIBCO, CONSER)

• ECIP, PCC ISNI Pilot, FOLIO, Colorado Alliance of Research Libraries Shared Print Trust, GPO Preservation Steward

• Innovative Interfaces Sierra and Summon
Collections

- 6,401,489 bibliographic records (all formats)
- 875,111 e-book records (excluding gov docs)
- Approximately 150 e-book packages (and growing)
- 463 collections selected in Collection Manager (264 ebooks)

Bib records/access

- Serials Solutions (e-journals, databases, fluctuating and low priority ebooks), vendor/publisher records, one-off cataloging

Consolidating collections (mostly ebooks) with Collection Manager
Why Consolidate?

- DDA + firm and approval orders (shared with CU System libraries), packages, databases
- Moving toward more normalized approach to ebook records
- Standard edits for all collections received via OCLC (we used to have a set of custom edits for each separate collection); this is a time-saver (and helps us be more consistent)
- More accurate holdings on OCLC (ILL)
- Better records and more accurate access
- Better way to share records with partner libraries
- Contribution to the membership
Discrete Collections

English Historical Documents

• 12 titles in collection
• Quality records with content notes
• Updates
Purchased Collections (shared and growing)

Setting up shared collections (DDA, approval, one-off purchases) for CU System libraries.

- CU Boulder, CU Colorado Springs, CU Denver (Auraria and Health Sciences)
- CU Boulder sets up collections
- Partner libraries provide access to selected/all titles
<table>
<thead>
<tr>
<th>Collection ID</th>
<th>Type</th>
<th>ISBN</th>
<th>Date</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>netlibrary.ebooks</td>
<td>Knowledge base collection</td>
<td>-</td>
<td>10/31/2017</td>
<td>eBooks on EBSCOhost, Selected by my library, Titles: 31,664/1,239,674</td>
</tr>
<tr>
<td>customer.462.28</td>
<td>Knowledge base collection</td>
<td>-</td>
<td>10/31/2017</td>
<td>CU Boulder, Selected by my library, Titles: 2,899/2,699</td>
</tr>
<tr>
<td>customer.462.30</td>
<td>Knowledge base collection</td>
<td>-</td>
<td>11/06/2017</td>
<td>CU Boulder, Selected by my library, Titles: 868/858</td>
</tr>
<tr>
<td>customer.462.29</td>
<td>Knowledge base collection</td>
<td>-</td>
<td>11/02/2017</td>
<td>CU Boulder, Selected by my library, Titles: 386/386</td>
</tr>
</tbody>
</table>

Be Boulder.
Going forward

Discrete collections
• If clear when new items are added to collection
• If manageable workload

Shared purchased (for CU System)

Ebook Packages (cooperative)
• Systematic approach to improving records based on encoding level
• Systematic approach to improving collections by making sure all titles are represented by a record
Cooperative cataloging project to improve access to the JapanKnowledge collection

Yukari Sugiyama
Yale University
November 9, 2017
What is JapanKnowledge?
Why catalog JapanKnowledge contents?

<table>
<thead>
<tr>
<th>Uniform Title</th>
<th>誰でも読める日本史年表 (Online)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>誰でも読める日本史年表 [electronic resource].</td>
</tr>
<tr>
<td>Published</td>
<td>[S.I.] : [s.n.]</td>
</tr>
<tr>
<td>Local Notes</td>
<td>Access is available to the Yale comm</td>
</tr>
<tr>
<td>Notes</td>
<td>Title from content provider.</td>
</tr>
<tr>
<td>Access and use</td>
<td>Access restricted by licensing agreement</td>
</tr>
<tr>
<td>Variant and related titles</td>
<td>誰でも読める日本史年表</td>
</tr>
<tr>
<td>Format</td>
<td>Journals &amp; Newspapers / Online</td>
</tr>
<tr>
<td>Language</td>
<td>Unknown</td>
</tr>
<tr>
<td>Added to Catalog</td>
<td>February 26, 2016</td>
</tr>
</tbody>
</table>

Title in the vernacular script is the only access point.
Create and share collection via Collection Manager

- Record Manager
- Collection Manager
- My Files

Search Collections with """"
Project process in a nutshell

1. Record assessment (search for e-records, provider-neutral record, print equivalents)

2. Cataloging (524 titles by a bulk process; 348 titles by project team members)

3. Create a KBART file for title data

4. Create JapanKnowledge collection in Collection Manager
## Collection: JapanKnowledge (Customizable)

<table>
<thead>
<tr>
<th>Title</th>
<th>Content Availability</th>
<th>Identifiers</th>
<th>Selections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abudurra monogatari</td>
<td>E-book/Proceeding</td>
<td>OCN: 786191562</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Afurika no minzoku to bunka</td>
<td>E-book/Proceeding</td>
<td>OCN: 932129316</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Airurando</td>
<td>E-book/Proceeding</td>
<td>OCN: 929688483</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Aisho shumi</td>
<td>E-book/Proceeding</td>
<td>OCN: 929029037</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Akamatsu Noriyoshi hanteidan</td>
<td>E-book/Proceeding</td>
<td>OCN: 639295683</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Akiyama kiko</td>
<td>E-book/Proceeding</td>
<td>OCN: 645388526</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Akuma no bunkashi</td>
<td>E-book/Proceeding</td>
<td>OCN: 929030746</td>
<td>Held by my library</td>
</tr>
<tr>
<td>Amerika Goshūkoku no chiri</td>
<td>E-book/Proceeding</td>
<td>OCN: 929030461</td>
<td>Held by my library</td>
</tr>
</tbody>
</table>
Outcome – record example

**Dare de mo yomeru Nihon shi nenpyō**

**Title**
Dare de mo yomeru Nihon shi nenpyō / Yoshikawa Kōbunkan.
誰でも読める日本史年表 / 吉川弘文館.

**Publication**

**Physical Description**
1 online resource

**Local Notes**
Access is available to the Yale community.

**Access and use**
Access restricted by licensing agreement.

**Format**
Books / Online

**Language**
Japanese

**Added to Catalog**
April 6, 2017

**Subjects**
Japan > History.
Japan.

**Also listed under**
Yoshikawa Kōbunkan.
吉川弘文館
Challenges and future concerns

• Record assessment
• Quality control over catalog records created by project members
• Promoting of the adoption of Collection Manager
• Ongoing collection maintenance
For more details about our project:

http://scholarsarchive.byu.edu/jeal/vol2016/iss163/4

Special thanks to the project participants and the content provider for supporting the project.

• Fabiano Rocha and Marlene Van Ballegooie (Toronto)
• Eva Bolkovac (Yale)
• Keiko Suzuki (New School)
• Mieko Mazza (Stanford)
• Ryuta Komaki (Washington, St. Louis)
• NetAdvance, Inc.
Collection-Level-Cooperative-Cataloging

JEFF SIEMON,
ANDERSON UNIVERSITY (ANDERSON, INDIANA)
NOVEMBER 9, 2017
Problem:
Patrons cannot find resources we provide (and have paid for) because many e-Resources have Poor Metadata.

“E-resource discovery is almost entirely dependent upon metadata that is supplied by parties outside of the library. ... libraries now work in an environment where there are more interdependencies than ever; the successful provision of access to electronic resources is fully dependent on the transmission of high quality metadata throughout the e-resource supply chain.”

http://scholarsarchive.byu.edu/jeal/vol2016/iss163/4
Problem: Patrons cannot find resources we provide (and have paid for) because many e-Resources have Poor Metadata.

Collection metadata has a complex supply chain, with few librarians reviewing or improving this data.

Vendor supplies:

- **Title list for purchase**
- **Title list metadata** (often less complete, fewer titles, coverage mistakes, missing ISSN/ISBNs)
- **Title list MARC records** (often poor subject access, non-standard author names, etc.)

OCLC supplies MARC records, OCLC numbers for ILL and other integration.

Local library turns on WorldCat knowledge base and/or downloads discovery data

WorldCat Discovery or another discovery layer provides indexing, discovery and delivery.
Plea for Catalogers to add KBart, Excel, MarcEdit, Collection Manager to their skill set.

• Catalogers have been active in item level (MARC) cooperative cataloging for decades.

• It’s time for some catalogers make the move to Collection Level Cooperative Cataloging.
Plea for Catalogers to add KBart, Excel, MarcEdit, Collection Manager to their skill set.

WHY make the move to collection level cataloging?

• Catalogers traditionally focus on Discovery, and e-Resources collections need improved metadata for better Discovery.

• Catalogers have derived meaning from their work by knowing that their improvements are being shared by thousands of other libraries.

• Collection Manager is the only knowledge base that encourages the cooperative improvement of the collection level metadata.

• Catalogers are good with detail, and knowledge base collections supplied by vendors need librarians to improve the details of this metadata.

• Libraries are purchasing fewer individual books, and fewer print books, and more collections of materials, especially eBooks. It makes sense for some catalogers to learn new skills and focus on this collection level metadata. (job security)
Plea for Electronic Resources team leaders and Library Directors to add metadata specialists (catalogers) to their team.

• Libraries are purchasing more collections of e-Journals and eBook but think that vendors will supply flawless discovery and access, which, of course, is fallacious.

• Many “Electronic Resources Librarians” came from Serials or Acquisitions, and focus on negotiating collection prices and licenses, etc. yet don’t have deep metadata skills.
Plea for Electronic Resources team leaders and Library Directors to add metadata specialists (catalogers) to their team.

• Electronic Resource team leaders: Why not add to your team librarians with collection level metadata skills? (hire or move from other library departments)

• Library Directors: Can you catch a new vision for strategically reassigning catalogers?
  – To date, catalogers are mostly being reassigned to unique collections and archives.
  – Catalogers also can be reassigned to improve poor metadata supplied by vendors for databases, journals and eBook collections.

• Also collection level cooperative cataloging can improve Discovery for open access collections.
Some reasons why you’d want to do cooperative collection level cataloging:

When one library creates this data, then we all can share it.

- The vendor may not supply metadata (yet)
- The vendor may not supply adequate metadata.
- The Vendor data may need to be improved with ISSN, ISBN, Complete titles, correct Romanization, OCLC numbers for best MARC record
- The Vendor may be late in sending metadata to discovery services.
- The Vendor may use only the current journal title in their metadata, even though they have full-text for preceding titles of the "same" journal.
- The Vendor may be reluctant to release their data.
- The Vendors may not separate out open access data.
When we make collection level data better …

Then these titles can be Discovered by and Delivered to your patrons!

This is why you’d want to do cooperative collection level cataloging.
Detailed Examples to help you learn and act upon:

Examples of Collection level Cataloging
https://www.slideshare.net/jsiemon/adding-oclc-numbers-iss-ns-and-issns-to-the-knowledge-base

How to add OCLC numbers to KBart data
https://drive.google.com/open?id=0B2sHKamxnI-dQmFkcUVmS1d4RTQ

How to enhance collections
https://vimeo.com/237468207/00de177b16
Enhancing and Curating KBART Metadata

Andrew Senior
McGill University Library
November 9, 2017
KBART Metadata and the Quantity-Quality Challenge

• Range of KBART data sources
• McGill BANQ publications, McGill hosted journals, open access collections
• Change is a constant
• Collaborative solutions “crowd sourcing” are invaluable
• Although one-by-one solutions are time-consuming, human intervention is still necessary
Types of KBART Data Enhancements

• Title metadata
• Title IDs, document status
• Dates
• URLs
• OCLC numbers
• Different approaches according to the problem
Strategies

- For open access titles, dedupe against publishers collections
- Article level linking more likely to be available for journal collections
- Improve existing metadata and contribute back to community
- Batch processes and automated solutions for some tasks
- Run targeted projects to clean or enhance data
- Empower staff to contribute changes
For Each Task Its Tool – OCLC Connexion

- Batch searching in Connexion to find problems
- Identify wrong formats or cataloging language
- “no:49108053 NOT mt:ebk”
For Each Task Its Tool – Excel and Text Editors
For Each Task Its Tool - OpenRefine

OpenRefine: http://openrefine.org/
For Each Task Its Tool - URL Curation

- Locating broken links in McGill OA collections
- Used W3C Link Checker to programmatically test each KBART URL

W3C Link Checker: [https://validator.w3.org/checklink](https://validator.w3.org/checklink)
For Each Task Its Tool - URL Curation

- W3C Program bot searches URLs and returns a list of HTTP codes
- Very good at identifying problematic links
  - Broken URLs
  - Redirecting URLs
  - Client/Server errors
  - (403 “Forbidden”, 404, “Not found”, 504 “Gateway Timeout” etc.)
- Enabled identification of over 1000 problematic URLs in our open access collections.
- Working through these URLs to fix or remove from collections.
Conclusion

• Collaborative contributions help enhance metadata

• Programmatic and batch processes are worth investing time in to automate repetitive tasks

• Greater contributions to global collections result in improved user experience, confidence and discoverability
Thank You! Any Questions?

Paul Moeller
University of Colorado Boulder
Paul.Moeller@Colorado.EDU

Yukari Sugiyama
Yale University
yukari.sugiyama@yale.edu

Jeff Siemon
Anderson University
josiemon@anderson.edu

Andrew Senior
McGill University
andrew.senior@mcgill.ca

Charleston Conference™
ISSUES IN BOOK AND SERIAL ACQUISITION
November 6 - 10, 2017
"WHAT'S PAST IS PROLOGUE"