19 June 2023

OCLC AJL Update

Nathan Putnam
Director, Data Quality & Governance
OCLC

Cynthia Whitacre
Senior Metadata Operations Manager
OCLC
OCLC apologizes to the community for the antisemitic hate speech that was posted to the OCLC-CAT distribution list in early March.
"We know emotional harm is real, and there was real harm resulting from this incident."
Our current plans

• To establish a code of conduct for all OCLC virtual spaces and share with this community.
• To clarify incident reporting procedures to ensure information on how to report harmful communications is readily accessible.
• To review and update listserv management policies, ensuring OCLC staff members are prepared to step in when needed to enforce the code of conduct.
546 million records

3.3 billion holdings

WorldCat gets one new record every second

Source: OCLC, May 2023
WorldCat languages

490 languages represented in WorldCat

Source: OCLC, May 2023
WorldCat scripts

35 non-Latin scripts used in 10 or more records

Source: OCLC, May 2023
A modern, user-friendly interface accessible by a web browser

Available to all Cataloging subscribers, including WMS libraries

MARC 21 editor and Text View editor to create new or enrich existing records in WorldCat

A saved in-progress file that is shared with the Connexion save file

Integrated management of staff users and permissions
Improved full search results with filtering & faceting – enhanced view

Support for numerous authority files and controlled headings
Authorities

Art & Architecture Thesaurus (AAT)
The Getty Research Institute

Biblioteca Nacional de España (BNE) Authority File
Biblioteca Nacional de España (National Library of Spain)

Canadiana (Autorités de noms Canadiana en français)
Bibliothèque et Archives Canada (Library and Archives Canada)

Canadian Subject Headings (CSH)
Bibliothèque et Archives Canada (Library and Archives Canada)

Faceted Application of Subject Terminology (FAST)
OCLC – coming soon!

Integrated Authority File (GND)
Deutsche Nationalbibliothek (German National Library)

LC/NACO Name Authority File (LCNAF) & Library of Congress Subject Headings (LCSH)
Library of Congress

Māori Subject Headings File
Ngā Upoko Tukutuku

Medical Subject Headings (MeSH)
U.S. National Library of Medicine

Nederlandse Thesaurus van Auteursnamen (NTA Names)
Koninklijke Bibliotheek (National Library of the Netherlands)

Répertoire de vedettes-matière (RVM)
Université Laval
Merging in Record Manager

• Merging functionality coming soon!
• Watch for announcements in the Record Manager and Member Merge Community Centers
• Details will be included in Record Manager release notes once functionality is available
Shaded areas show missing fields in the compared record.
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Becker, Aaron.</td>
</tr>
<tr>
<td>245</td>
<td>0 Journey $h$ (BOOK) / $c$ Aaron Becker.</td>
</tr>
<tr>
<td>250</td>
<td>1st ed.</td>
</tr>
<tr>
<td>264</td>
<td>1 Somerville, MA. $b$ Candlewick Press, $c$ 2013.</td>
</tr>
<tr>
<td>300</td>
<td>Unpaged.</td>
</tr>
<tr>
<td>336</td>
<td>text $b$ bt $d$ rdacontent</td>
</tr>
<tr>
<td>337</td>
<td>unmediated $b$ n $d$ rdmedia</td>
</tr>
<tr>
<td>338</td>
<td>volume $b$ nc $d$ rdacarrier</td>
</tr>
<tr>
<td>245</td>
<td>0 Journey / $c$ Aaron Becker.</td>
</tr>
<tr>
<td>250</td>
<td>First edition.</td>
</tr>
<tr>
<td>264</td>
<td>1 Somerville, Massachusetts : $b$ Candlewick Press, $c$ 2013.</td>
</tr>
<tr>
<td>264</td>
<td>4 $c$ ©2013</td>
</tr>
<tr>
<td>300</td>
<td>40 unnumbered pages : $b$ color illustrations ; $c$ 25 x 28 cm.</td>
</tr>
<tr>
<td>336</td>
<td>text $b$ bt $d$ rdacontent</td>
</tr>
<tr>
<td>337</td>
<td>unmediated $b$ n $d$ rdmedia</td>
</tr>
<tr>
<td>338</td>
<td>volume $b$ nc $d$ rdacarrier</td>
</tr>
</tbody>
</table>

Field level alignment to make for easy comparison.
Provides access to the WorldCat knowledge base to simplify and automate your e-workflows

Automatically maintains WorldCat holdings

Improves the quality of your catalog with optional WorldCat MARC record delivery

Makes your e-resources easier to find in WorldCat.org and Google Scholar

Available to all Cataloging subscribers, including WMS libraries
WorldCat knowledge base

Over 23,000 collections representing 65 million items from nearly 700 providers

- Subscription packages
- Title by title purchases
- Open Access content
- Demand Driven Acquisitions
- Consortial collections
- Cooperatively created and managed
- Locally defined

For a complete list, visit oc.lc/kbcollections
Automated holdings management

For more information, visit oc.lc/autoload
New staff

Kate James
Program Coordinator, Metadata Engagement

Morris Levy
Consulting Database Specialist

Jason Bourg
Lead Data Ontologist
Connexion Browser: end of life

- 2024 April 30
  - end of life date for Connexion Browser
- For a more modern web interface
  - use Record Manager
- Support continues for Connexion Client

NO end of life date for Connexion Client

Photo credit: Cynthia Whitacre
OCLC linked data ecosystem

OCLC improves access to bibliographic metadata by working with the global library community to create and nurture a linked data ecosystem encompassing infrastructure, data, applications, APIs, and a global network.
Solving for a variety of challenges

• Increasingly complex and messy hybrid environments
• **Disparate URIs** make authority work difficult and time-consuming
• **Lack of tools** supporting a unified cataloging environment prohibits data connections
• **Inability for data to serve multiple workflows** leads to duplicative creation and curation
• **Absence of an ecosystem** that enables shared, collaborative linked data management at scale promotes disparate workflows which are ill-suited and unsustainable
• **Increasing need for metadata work to support DEI initiatives**
Building a comprehensive and sustainable linked data ecosystem

• Create a global identifier that can be used across workflows by adding to MARC data

• Release tools that bridge the gap between traditional records and linked data frameworks to enable seamless creation and curation of data

• Add entities and entity types (including Persons, Works, Concepts, Events, and Organizations) to enrich data set

• Enhance our data model/ontology designed to describe entity types in the library domain and maximize their relational impacts
WorldCat Entities and OCLC Meridian

- WorldCat Entities launched May 2022
- WorldCat Entities & OCLC Meridian development partner phase
  - Partners in North America and Europe
  - Varied workflows and library types
    - Partner-planned projects
  - Worked 1:1 with OCLC staff
  - Lasted 2-3 months

Image from Pixabay – no attribution required
WorldCat Entities and OCLC Meridian

Outcomes

- Tested and validated linked data workflows
- Helped prioritize OCLC data quality enhancements
- Identified additional properties to provide a fuller description of Person and Work entities
- Tested overall usability, i.e., multi-lingual support

Currently incorporating feedback into WorldCat Entities and OCLC Meridian
WorldCat Entities and OCLC Meridian

• Interested in WorldCat Entities or OCLC Meridian?
• Want to be involved with future testing and feedback opportunities?

Contact:

Jeff Mixter, Senior Product Manager (mixterj@oclc.org)
Anne Washington, Product Analyst (washinga@oclc.org)

Or talk in person at ALA 2023, booth #4011
Updates to the authority file

• Updates coming soon
  – Removing field 375 – Gender
  – Recoding family names in the authority file as valid for subject use

• Current information
  – Waiting on authorization from the Library of Congress to start the project
  – 1.3 million records with filed 375
  – 1200 records with family names
880 fields (Non-Latin script)

• 880 fields may not be added or edited on PCC records
  – Do not need a PCC authorization (Full-level Connexion Authorization or Cataloging Full Record Manager Role)

• Updates to Bibliographic Formats and Standards coming soon
  – 5.2.3 Editing capabilities for PCC records
DDC enrichment

Goal: add DDC numbers to bib records that lack them

- Class number comes from other records in a work cluster
- Formatting and conditions set by Alex Kyrios, Senior Dewey Editor
- Treat as any other class number found in a record

Progress so far

- ~33 million records processed since February of an expected 85M

Example

- 082 04 $a 306.909421 $q OCoLC $2 23/eng/20230216
OCLC labeling project

<table>
<thead>
<tr>
<th>Record 1</th>
<th>Fields</th>
<th>Record 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>cam a M</td>
<td>leader</td>
<td>cam a Ma</td>
</tr>
<tr>
<td>104973480</td>
<td>d001</td>
<td>967864806</td>
</tr>
<tr>
<td>730915x1957</td>
<td>d008</td>
<td>060915x1957</td>
</tr>
<tr>
<td>000 0 eng d</td>
<td>5005912</td>
<td>5005912</td>
</tr>
<tr>
<td>s $a NAM b$eng c$NAM</td>
<td>d010</td>
<td>s $a NAM b$eng c$NAM</td>
</tr>
<tr>
<td>Missing Entry</td>
<td>d050</td>
<td>s $a MacEachem, Malcolm Thomas. $d 1881-1956</td>
</tr>
<tr>
<td>s $a MacEachem, Malcolm Thomas. $d 1881-1956</td>
<td>d100</td>
<td>s $a MacEachem, Malcolm Thomas. $d 1881-1956</td>
</tr>
<tr>
<td>s $a Hospital organization and management / $c by Malcolm T. MacEachem</td>
<td>d246</td>
<td>s $a Hospital organization and management / $c by Malcolm T. MacEachem</td>
</tr>
<tr>
<td>s $a [Rev. 3rd ed.]</td>
<td>d250</td>
<td>s $a Rev. 3rd ed</td>
</tr>
<tr>
<td>Missing Entry</td>
<td>d280</td>
<td>s $a Chicago : s$b Physicians' Record Co. $c 1957</td>
</tr>
<tr>
<td>s $a Chicago : s$b Physicians' Record Co. $c 1957</td>
<td>d284</td>
<td>Missing Entry</td>
</tr>
<tr>
<td>s $a xii, 1,116 pages : s$2 illustrations, color portrait, forms : s$c 25 cm</td>
<td>d300</td>
<td>s $a xii, 1,116 pages. 27 unnumbered leaves of plates : s$2 illustrations (some color) : s$c 25 cm</td>
</tr>
<tr>
<td>s $a text : s$b bd $2 rlncoretant</td>
<td>d326</td>
<td>s $a text : s$b bd $2 rlncoretant</td>
</tr>
<tr>
<td>s $a unmediated : s$b n $2 rlncoretant</td>
<td>d327</td>
<td>s $a unmediated : s$b n $2 rlncoretant</td>
</tr>
<tr>
<td>s $a volume s$b nc $2 rlncoretant</td>
<td>d338</td>
<td>s $a volume s$b nc $2 rlncoretant</td>
</tr>
<tr>
<td>s $a includes bibliographies</td>
<td>d804</td>
<td>s $a includes bibliographies and index</td>
</tr>
</tbody>
</table>
Data labeling project outcomes

• Collected data on 34,356 pairs of potential duplicate records
  – Data used to train the deduplication model
  – Data Quality staff reviewed results from the model
• More information: April 2023 VOAH session at oc.lc/askqc
• Next steps: applying the model to WorldCat
  – Stay tuned for details, coming soon!
<table>
<thead>
<tr>
<th>Month</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>Cataloging rare materials defensively</td>
</tr>
<tr>
<td>July</td>
<td>De-biasing Dewey: righting the past by rewriting the classification</td>
</tr>
<tr>
<td>August</td>
<td>Rapid harm reduction with locally preferred subjects in WorldCat Dis</td>
</tr>
<tr>
<td>September</td>
<td>Get informed about genre/form terms</td>
</tr>
<tr>
<td>November</td>
<td>MARC fields for manuscripts and archival collections</td>
</tr>
</tbody>
</table>

Registration at [oc.lc/askqc/](http://oc.lc/askqc/)
Join us in-person @ ALA

- OCLC Cataloging Community Meeting
  Friday, June 23 at 10:30
  McCormick Place, Rm W179a
- Meet with OCLC staff and colleagues
- Ask cataloging policy questions
- Get the latest and greatest info
- Celebrate Jay!

Jay Weitz, Senior Consulting Database Specialist
Celebrating Jay Weitz!

Congratulations to Jay Weitz, who recently received the MLA Citation at the Music Library Association’s annual conference. The MLA Citation, the Association’s tribute for lifetime achievement, is awarded in recognition of distinguished service to music librarianship over a career. Citation recipients become Honorary Members of the Music Library Association. Jay is a Senior Consulting Database Specialist on the WorldCat Quality Management team in Dublin, and has served since 1989 as OCLC’s liaison to the music catalogers community.
thank you