OCLC Updates
Session: OCLC and Linked Data
Shanna Griffith
Database Specialist II

Introduction
OCLC and Linked Data

John Chapman
A decade with Linked Data

- Publish linked data - FAST, VIAF, WorldCat (2009 - )
- EntityJS Research Project (2013)
- Person Entity Lookup Pilot (2014)
- CONTENTdm Metadata Refinery (2015-16)
- Project Passage (2017-18)
- CONTENTdm Linked Data Pilot (2019-20)
- Shared Entity Management Infrastructure (2020-21)
Event announcement

• “OCLC and Linked Data – the transition to contextual metadata”
• 2 March 2021; 11 AM – 12 PM US Eastern Standard Time
• Speakers:
  • John Chapman
  • Shane Huddleston
  • Nathan Putnam
• Significant question/answer time included
Have a question or want to get involved? LinkedData@oclc.org
And the winner is....

Corinne Jacox
Creighton University Law Library
Trivia time!

• WebEx polling will begin

• Please answer swiftly

• Click on your answer
• Press “Submit”
Trivia #2 Klingon in WorldCat

How many records within WorldCat contain the Klingon language *tlhIngan Hol*, MARC language code *tlh*?

- 1
- 10
- 245
- 1000

Source:
https://commons.wikimedia.org/wiki/File:Alfabeto_klingon_de_Skybox.png
Trivia #2 Klingon in WorldCat

• How many records within WorldCat contain the Klingon language _tlhIngan Hol_, MARC language code _tlh_?
  • 1
  • 10
  • 245
  • 1000

Letra klingon  Valor en inglés (y español)  Valor en klingon

\[\begin{array}{lll}
\text{e, c} & \text{e, ch} & \text{CH} \\
\text{f, p} & \text{p, v} & \text{EH} \\
\text{g, d, o, 'D, o, '} & \text{l, v, y, a} & \text{EH} \\
\text{h, r} & \text{H, r} & \text{ng} \\
\text{i, s} & \text{l, S} & \text{gh} \\
\end{array}\]

* ' (inicio de palabra)

Source:
[https://commons.wikimedia.org/wiki/File:Alfabeto_klingon_de_Skybox.png](https://commons.wikimedia.org/wiki/File:Alfabeto_klingon_de_Skybox.png)
Session: OCLC Cataloging Update
Shanna Griffith

Database Specialist II

Introduction
OCLC Cataloging Updates

Metadata Quality staff
Upcoming OCLC-MARC Validation installation

• Includes MARC 21 Bibliographic, Authority, and Holdings Updates No. 31 (December 2020), MARC Codes from four LC Technical Notices (October-December 2020), several bug fixes

• Details will be in WorldCat Validation Release Notes soon at oc.lc/validation-release-notes

• Currently scheduled for late February, but always subject to change
REALM Project

• **RE**opening **A**rchives, **L**ibraries, and **M**useums (REALM) research project
• Conducted by OCLC, the Institute of Museum and Library Services (IMLS), and Battelle
• Producing science-based information focusing on material handling to mitigate COVID-19 exposure in archives, libraries, and museums
• Latest news and all the research: [https://www.oclc.org/realm/home.html](https://www.oclc.org/realm/home.html)
Member Merge Project (MMP)

• Pilot project began in 2013 with four participants

• 53 libraries currently involved (44 Academic Libraries, 4 State Libraries, 2 National Libraries, 2 Public Libraries, 1 Other (CRL))

• Over 66,000 sets of duplicates merged by MMP participants since 2012 · Interested in joining? Must be PCC participant. Email us at askqc@oclc.org
Virtual AskQC Office Hours (VAOH)

• VAOH sessions mostly monthly since January 2018, now twice a month
• AskQC@oclc.org is email address for catalogers to ask questions about policies, standards, practices
• Recordings, slides, Q&A from past sessions, as well as registration for current and future sessions at oc.lc/askqc
OCLC Annual Report


• Accomplishments, enhancements, research, statistics, governance, and financials for Fiscal Year 2020
Encoding Levels

• Goal – Eliminate OCLC codes I, J, K, and M in favor of MARC 21 codes
  • Eliminates redundant codes (e.g., K and 7 both mean minimal-level)
  • Conforms to MARC 21 standards
  • Simplifies documentation and training

• Discussed changes with libraries (PCC and Expert Community Session at ALA) as well as conducting focus groups with different types of libraries

• Made system changes to allow record creation and upgrades using MARC 21 encoding levels for all cataloging authorizations

• Changed BFAS to encourage use of MARC 21 codes
Encoding Levels

• Will begin conversion of encoding level K records in January
• Will convert 25.9 million records via a macro
• Macro will evaluate the record content
  • Some will be upgraded to full (i.e., encoding level blank)
  • Others will remain minimal (i.e., encoding level 7)
  • Few will be changed to abbreviated (i.e., encoding level 3)
• May have some impact on output through Collection Manager
• Currently, no estimated date of completion
• For more information, review VAOH session for June 2020 @ oc.lc/askqc
Updates to Bibliographic Formats & Standards

- Chapter 5.2
- Replaced (Field 005)
- Field 042
- Field 045
- 050-099 Introduction to Call Numbers
- Field 688 (NEW)
- 6xx Fields
- 7xx Fields
- MARC 21 links
Authority-related updates

• Derive authority record from a bibliographic record (June)
• Add equivalent subject headings to a bibliographic record (June)
• Biblioteca Nacional de España (BNE) (December)
<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP retained records</td>
<td>1,769</td>
<td>1,783</td>
<td>2,879</td>
<td>3,126</td>
<td>3,076</td>
<td>3,828</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15,921</td>
</tr>
<tr>
<td>MMP deleted records</td>
<td>3,288</td>
<td>2,844</td>
<td>3,932</td>
<td>5,381</td>
<td>5,574</td>
<td>5,885</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26,854</td>
</tr>
<tr>
<td>MMP processed records</td>
<td>5,037</td>
<td>4,597</td>
<td>6,311</td>
<td>8,507</td>
<td>8,450</td>
<td>9,683</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>42,585</td>
</tr>
</tbody>
</table>

Retained = The count of records retained after duplicates were merged into it (a.k.a. merge transactions).
Deleted = The count of records that were merged into a retained record.
Processed = The count of all reviewed records that were processed, i.e. Retained records + Deleted records.

All-time count of merge transactions (from 2014 to present) 73,499
FY21 WorldCat community enrichments

### Bib enrich counts

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>Q3</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert</td>
<td>26,052</td>
<td>26,614</td>
<td>26,138</td>
<td>26,332</td>
<td>22,419</td>
<td>19,622</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>146,167</td>
</tr>
<tr>
<td>DB enrichment</td>
<td>17,032</td>
<td>16,807</td>
<td>26,002</td>
<td>23,375</td>
<td>21,064</td>
<td>20,906</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>127,186</td>
</tr>
<tr>
<td>Regular enhance</td>
<td>8,786</td>
<td>8,077</td>
<td>7,152</td>
<td>7,438</td>
<td>5,906</td>
<td>5,957</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>43,316</td>
</tr>
<tr>
<td>National enhance</td>
<td>4,339</td>
<td>4,587</td>
<td>4,452</td>
<td>4,932</td>
<td>4,472</td>
<td>5,407</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28,189</td>
</tr>
<tr>
<td>Minimal-level</td>
<td>13,359</td>
<td>14,250</td>
<td>17,207</td>
<td>18,250</td>
<td>24,350</td>
<td>13,575</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91,041</td>
</tr>
<tr>
<td>CONSER authentic</td>
<td>1,028</td>
<td>1,029</td>
<td>709</td>
<td>556</td>
<td>433</td>
<td>552</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4,287</td>
</tr>
<tr>
<td>CONSER maintai</td>
<td>3,507</td>
<td>3,258</td>
<td>2,882</td>
<td>2,783</td>
<td>2,664</td>
<td>2,772</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17,846</td>
</tr>
<tr>
<td>Community enrich</td>
<td>73,133</td>
<td>74,632</td>
<td>84,542</td>
<td>85,606</td>
<td>71,348</td>
<td>68,771</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>458,032</td>
</tr>
</tbody>
</table>

### Bib enrich activities

![Bar chart showing bib enrich activities by quarter and month.]

### Bib enrich and replace totals

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>Q3</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community enrich total</td>
<td>73,133</td>
<td>74,632</td>
<td>84,542</td>
<td>85,606</td>
<td>71,348</td>
<td>68,771</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>458,032</td>
</tr>
<tr>
<td>Other replace</td>
<td>171,728</td>
<td>169,263</td>
<td>131,392</td>
<td>192,822</td>
<td>169,454</td>
<td>159,159</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>586,785</td>
</tr>
<tr>
<td>Community enrich &amp; replace total</td>
<td>244,861</td>
<td>244,895</td>
<td>216,137</td>
<td>278,448</td>
<td>240,796</td>
<td>227,936</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1,045,817</td>
</tr>
</tbody>
</table>

Source: OCLC
# FY21 Duplicate Detection and Resolution (DDR) stats

## Record activity counts

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Records reviewed</td>
<td>7,238,263</td>
<td>16,952,527</td>
<td>15,142,694</td>
<td>4,055,215</td>
<td>6,296,294</td>
<td>3,757,563</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53,442,656</td>
</tr>
<tr>
<td>Records retained</td>
<td>1,045,041</td>
<td>2,050,775</td>
<td>1,631,028</td>
<td>555,990</td>
<td>897,378</td>
<td>492,944</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6,623,156</td>
</tr>
<tr>
<td>Records deleted</td>
<td>1,534,942</td>
<td>2,360,666</td>
<td>1,722,962</td>
<td>756,798</td>
<td>1,084,576</td>
<td>713,103</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8,153,047</td>
</tr>
<tr>
<td>Records processed</td>
<td>2,559,983</td>
<td>4,411,441</td>
<td>3,353,990</td>
<td>1,292,788</td>
<td>1,951,954</td>
<td>1,206,047</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14,776,203</td>
</tr>
</tbody>
</table>

% processed of rev.. 35.37% 26.02% 22.15% 31.88% 31.00% 32.10%

Reviewed = All records that were to be reviewed by DDR (prior to FY20 this was called processed).
Retained = The count of records retained after duplicates were merged into it (a.k.a. "merge transactions").
Deleted = The count of records that were merged into a retained record.
Processed = The count of all reviewed records that were processed, i.e. Retained records + Deleted records.

## Reviewed record counts

![Graph showing reviewed records by quarter]

Source: OCLC

## History since August 2009

<table>
<thead>
<tr>
<th>Total records rev..</th>
<th>918,901,121</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total records re..</td>
<td>60,735,945</td>
</tr>
<tr>
<td>Total % removed</td>
<td>6.61%</td>
</tr>
</tbody>
</table>
FY21 OCLC replaces and manual merges

Replace activity

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCLCQ replaces</td>
<td>2,249,814</td>
<td>1,621,996</td>
<td>1,696,855</td>
<td>1,944,113</td>
<td>2,954,149</td>
<td>2,423,839</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12,896,796</td>
</tr>
<tr>
<td>OCL Replaces</td>
<td>1,288,100</td>
<td>45,450</td>
<td>964,932</td>
<td>19,998</td>
<td>44,046</td>
<td>325,289</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2,672,814</td>
</tr>
<tr>
<td>Replaces total</td>
<td>3,537,914</td>
<td>1,667,446</td>
<td>2,661,787</td>
<td>2,163,011</td>
<td>3,399,195</td>
<td>5,749,128</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15,569,606</td>
</tr>
</tbody>
</table>

Merges by all OCLC staff

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCLC retained merges total</td>
<td>2,045</td>
<td>121,367</td>
<td>4,574</td>
<td>21,482</td>
<td>71,644</td>
<td>37,639</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>258,771</td>
</tr>
<tr>
<td>OCLC deleted merges total</td>
<td>2,533</td>
<td>152,565</td>
<td>5,667</td>
<td>23,163</td>
<td>73,175</td>
<td>38,712</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>265,814</td>
</tr>
<tr>
<td>OCLC processed merges total</td>
<td>4,578</td>
<td>243,932</td>
<td>10,241</td>
<td>44,646</td>
<td>144,819</td>
<td>76,351</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>524,585</td>
</tr>
</tbody>
</table>

Processed vs. retained merges

Source: OCLC
Submit questions in Chat. Send chat to “Everyone”
Nathan Putnam
Director, Metadata Quality

Closing
Thank you
Final notes...

Recordings/slides/notes

• Watch your email
  • WebEx email
  • Listservs
• Linked on the Community Center
  • Cataloging and Metadata
  • Record Manager
  • Collection Manager
• Linked on the AskQC help site
Final notes...

Recordings/slides/notes

- Watch your email
  - WebEx email
  - Listservs
- Linked on the Community Center
  - Cataloging and Metadata
  - Record Manager
  - Collection Manager
- Linked on the AskQC help site

The conversation continues...

- Cataloging policy
  - Email AskQC@oclc.org
  - AskQC twice monthly webinars
- News, discussions, and more...
  - Cataloging and Metadata community
  - Record Manager community
  - Collection Manager community
Thank You!

Please send questions and comments to AskQC@oclc.org
OCLC Cataloging Community Meeting
15 January 2021