OCLC Bibliographic Data Sync processing - collection configuration, test files and report output

- Overview - Data Sync collections for Reclamation phase
- Bibliographic Data Sync Processing Log - Reclamation
- Bibliographic Data Sync Processing Log - Test Files
- Error report details

Overview - Data Sync collections for Reclamation phase

For each symbol, files will be exported for four Data Sync collections: Manuscript, Reproduction, No trusted OCN, Trusted OCN. Files of microform and electronic reproductions are extracted in separate files for the Reproduction collection. For "No Trusted OCN" separate files are pulled for records with and without vernacular. Files for each collection are delivered to OCLC for Data Sync processing. Criteria for extracting records ensures that if a record meets the criteria for one collection, it will not be pulled again as part of another collection for that symbol.

<table>
<thead>
<tr>
<th>Data Sync Collection</th>
<th>Alma criteria (suppressed and provisional holdings excluded)</th>
<th>Data Sync Matching</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript</td>
<td>(LDR/06 = d or t) OR (LDR/06 = a) OR (040 $e = amim or appm or dacs or dcrg or dcrrm or dcrrmm)</td>
<td>Trusted partner matching.</td>
<td>Manuscript collection excludes reproductions. Holdings may or may not have 079, or vernacular</td>
</tr>
<tr>
<td>Reproduction (matching OCN loaded to holding only)</td>
<td>843 $a contains Microfilm or Microfiche</td>
<td>All matching but limited to English language of cataloging only.</td>
<td>Reproduction collection excludes manuscripts. Holdings may or may not have 079, or vernacular.</td>
</tr>
<tr>
<td>No trusted OCN</td>
<td>(no 079 field) AND (no 880 field)</td>
<td>All matching but English language of cataloging only.</td>
<td>No Trusted OCN means holdings have not been sent to OCLC previously. Collection excludes manuscripts and reproductions.</td>
</tr>
<tr>
<td>Trusted OCN</td>
<td>079 field</td>
<td>OCN match only. No validation, no enhancement of master records.</td>
<td>Holding has been sent to OCLC before. Sent to Data Sync with OCN from 079 only. Excludes manuscripts and reproductions.</td>
</tr>
</tbody>
</table>

Bibliographic Data Sync Processing Log - Reclamation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Data Sync Collection</th>
<th>Sent / Processed / Staging</th>
<th>Match / Add</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berenson (HTV)</td>
<td>Reproduction - microform</td>
<td>12104 / 12083 / 20</td>
<td>9138 / 2945 BER MFRM holdings data - Reclamation</td>
</tr>
<tr>
<td></td>
<td>Reproduction - electronic</td>
<td>179 / 178 / 1</td>
<td>13 / 165 BER CDROM holdings data - Reclamation</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>9170 / 8940 / 229</td>
<td>7890 / 1050 Bibs that lack a trusted OCN - Reclamation</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>32 / 31 / 1</td>
<td>31 / 0</td>
</tr>
<tr>
<td></td>
<td>Trusted OCN</td>
<td>99997 / 99996 / None 39893 / 39892 / None</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports HTV (Excel) Scan / Delete report (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine Arts (HFL)</td>
<td>Trusted OCN</td>
<td>429547 / 429533 / 3</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>37659 / 35838 / 1819</td>
<td>31316 / 4522</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>3933 / 3854 / 77</td>
<td>3289 / 585</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>5022 / 5017 / 5</td>
<td>4854 / 163</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports HFL (Excel) Scan / Delete report (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cabot (CLS)</td>
<td>Trusted OCN</td>
<td>195290 / 195260</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>4794 / 4667 / 127</td>
<td>3920 / 747</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>22 / 22 / 0</td>
<td>20 / 2</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>222 / 222 / 0</td>
<td>174 / 48</td>
</tr>
<tr>
<td>Reproduction - electronic</td>
<td>519 / 519 / 0</td>
<td>501 / 18</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scan / Delete</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hellenic Studies (HCH)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/2/20</td>
<td>Trusted OCN</td>
<td>44235 / 44231</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>1797 / 1796 / 1</td>
<td>1653 / 153</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>31 / 31 / 0</td>
<td>30 / 1</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>54 / 54 / 0</td>
<td>38 / 16</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports HCH (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dumbarton Oaks (DDO)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/15/20</td>
<td>Trusted OCN</td>
<td>163919 / 163864 / 18</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>14584 / 14526 / 58</td>
<td>12556 / 1970</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>933 / 932 / 1</td>
<td>852 / 80</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>383 / 382 / 1</td>
<td>258 / 124</td>
</tr>
<tr>
<td></td>
<td>Reproduction - electronic</td>
<td>17 / 17 / 0</td>
<td>3 / 14</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divinity (BHA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/19/20</td>
<td>Trusted</td>
<td>389569 / 389524 / 37</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>17864 / 17822 / 41</td>
<td>15551 / 2271</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>56 / 56 / 0</td>
<td>51 / 5</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>45142 / 45054 / 88</td>
<td>29789 / 15265</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gutman (HMG)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/5/20</td>
<td>Trusted</td>
<td>182094 / 182068 / 1</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>6996 / 6632 / 364</td>
<td>5274 / 1358</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>26 / 26 / 0</td>
<td>26 / 0</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>11954 / 11925 / 29</td>
<td>6330 / 5595</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kennedy (KSG)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/23/20</td>
<td>Trusted</td>
<td>47254 / 47254 / 0</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>3175 / 3163 / 12</td>
<td>2062 / 101</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>26 / 26 / 0</td>
<td>26 / 0</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>1 / 1 / 0</td>
<td>1 / 0</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baker (HBS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/28/20</td>
<td>Trusted</td>
<td>238486 / 238382 / 1</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>27499 / 27491 / 8</td>
<td>2107 / 6484</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN - vernacular</td>
<td>32 / 32 / 0</td>
<td>31 / 1</td>
</tr>
<tr>
<td></td>
<td>Reproduction - microform</td>
<td>70367 / 70361 / 6</td>
<td>68850 / 1511</td>
</tr>
<tr>
<td></td>
<td>Reproduction - electronic</td>
<td>1 / 1 / 0</td>
<td>1 / 0</td>
</tr>
<tr>
<td></td>
<td>Collected Bib Detail Error reports (Excel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scan / Delete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Countway (HMS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8/2/20</td>
<td>Trusted</td>
<td>378785 / 378748 / 1</td>
<td>All matched</td>
</tr>
<tr>
<td></td>
<td>No trusted OCN</td>
<td>4700 / 4679 / 21</td>
<td>4519 / 160</td>
</tr>
<tr>
<td>Location</td>
<td>Status</td>
<td>All matched</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Yenching (HMY)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>8/26/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master Microforms (HMM)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>8/27/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widener (HLS)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>9/8/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFR, CEL, CHI, DOC, FUN, GIB, HIS, HSL, LAM, LIN, NEL, PAL, PHI, SAN, SMY, URI (HLS)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>9/8/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microforms (Lamont) (HLS)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>9/8/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map (HLS)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>9/8/20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Networked Resource - FIG (HLS)</td>
<td>Trusted</td>
<td>All matched</td>
<td>No trusted OCN</td>
</tr>
<tr>
<td>9/8/20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bibliographic Data Sync Processing Log - Test Files

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Data Sync Collection</th>
<th>Sent/Processed/Rejected</th>
<th>Error detail report</th>
<th>Review document</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTV</td>
<td>Manuscript</td>
<td>22/20/2</td>
<td>metacoll.SYS.datsync.1034251.D20191225. bibdetailexcpt.1.txt</td>
<td>BER manuscripts test file 20191220 - review HU and OCLC</td>
<td>On hold while OCLC investigates bibs that did not match.</td>
</tr>
<tr>
<td>HTV</td>
<td>Reproductions - Microform</td>
<td>26/26/0</td>
<td>None</td>
<td>BER MFRM holdings data</td>
<td>Complete</td>
</tr>
<tr>
<td>HTV</td>
<td>Reproductions - Electronic</td>
<td>15/15/0</td>
<td>None</td>
<td>BER CDROM holdings data</td>
<td>Complete</td>
</tr>
<tr>
<td>HTV</td>
<td>No trusted OCN</td>
<td>40/31/9</td>
<td>metacoll.SYS.datsync.1034633.D20200211. T214059000.bibdetailexcpt.1.mrc.txt</td>
<td>Bbs that lack a trusted OCN</td>
<td>Complete</td>
</tr>
<tr>
<td>HTV</td>
<td>No trusted OCN - vernacular</td>
<td>29/29/0</td>
<td>None</td>
<td>Bbs that lack a trusted OCN - vernacular</td>
<td>Complete</td>
</tr>
<tr>
<td>HTV</td>
<td>Trusted OCN</td>
<td>7/7/0</td>
<td>None</td>
<td>Bbs with trusted OCN</td>
<td>Complete</td>
</tr>
</tbody>
</table>

Data Sync enhancements - June 25, 2020

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Data Collection</th>
<th>Sent/Processed/Rejected</th>
<th>Error detail report</th>
<th>Review document</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTV</td>
<td>Manuscript</td>
<td>22/20/2</td>
<td>Cross-reference report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unresolved cross-reference reports</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exception detail report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bib processing report</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bib summary report</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Error report details

OCLC documentation on validation errors, sparse records and unresolved records: [https://help.oclc.org/Metadata_Services/WorldShare_Collection_Manager/Understand_record_processing/Data_sync_processing](https://help.oclc.org/Metadata_Services/WorldShare_Collection_Manager/Understand_record_processing/Data_sync_processing)

<table>
<thead>
<tr>
<th>Data Sync Exception Description</th>
<th>Bib Validation Error</th>
<th>Example – DRAFT</th>
<th>Resolution by norm rule – DRAFT</th>
<th>Resolution by analysis – DRAFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$ in 245 occurs too many times.</td>
<td>Severe</td>
<td>If $a$ are identical, remove second</td>
<td>Agree/KCY</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If $a$ are different, flag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date1/Beginning Date of Publication (008/07-10) is too short.</td>
<td>Severe</td>
<td>Pull from 26X?</td>
<td>Agree/KCY</td>
<td></td>
</tr>
<tr>
<td>Date2/Ending Date of Publication (008/11-14) is too short.</td>
<td>Severe</td>
<td>Pull from 26X? Could be more difficult. Or make a rule to read the century of a date in the 26X to autocorrect the 008/11-14</td>
<td>Agree/KCY</td>
<td></td>
</tr>
<tr>
<td>Invalid code in 1st $a$ in 1st 041.</td>
<td>Severe</td>
<td>Rule to correct common errors, but otherwise &gt;&gt;</td>
<td>Analysis required. Is there a default language code we could use? ZXX?</td>
<td></td>
</tr>
</tbody>
</table>
**Invalid relationship** - when Form of Item (008/23) is equal to d, then Type of Record (Leader/06) must not be equal to j.

Severe 008/23 d = large print; LDR/06 j = musical sound recording
Which takes precedent? Change the other
LDR/06=j cannot exist with Form of Item (008/23) =d (see OCLC documentation [https://www.oclc.org/bibformats/en/fixedfield/type.html](https://www.oclc.org/bibformats/en/fixedfield/type.html)/KCY

**Invalid relationship** - when Type of Date/Publication Status (008/06) is equal to c, then Bibliographic Level (Leader/07) must not be equal to m.

Severe 008/06 c = Continuing resource currently published; LDR/07 m = monograph item
Which takes precedent? Change the other
If 008/06=c, change LDR/07= s/KCY

**Invalid relationship** - when Type of Date/Publication Status (008/06) is equal to k, then Bibliographic Level (Leader/07) must not be equal to m.

Severe 008/06 k = Range of years of bulk of collection; LDR/07 m = monograph item
Which takes precedent? Change the other
If 008/06=k, change LDR 07= c/KCY

**Invalid relationship** - when Type of Date/Publication Status (008/06) is equal to m, then Bibliographic Level (Leader/07) must not be equal to s.

Severe 008/06 m = multiple dates; LDR/07 s = serial
Which takes precedent? Change the other
008/06=m (computer file), agree, why is LDR/07 not s?

**Invalid relationship** - when Type of Date/Publication Status (008/06) is equal to s, then Bibliographic Level (Leader/07) must not be equal to m.

Severe 990118026470203941 Change Leader/07 to ‘a’ Agree/KCY

**Invalid subfield 1st $c in 1st 505.**

Severe Should this be read as: If 1st subfield in 1st 505= c? If yes, then change to $a (??)/KCY
Changeable by norm rule, but to what?
Can we see some examples

**Invalid subfield 1st $d in 1st 260.**

Severe Should this be read as: If 1st subfield in 1st 260= c? If yes, then change to $c (??)/KCY
Changeable by norm rule, but to what?
Can we see some examples

**Invalid subfield 1st $d in 1st 510.**

Severe Should this be read as: If 1st subfield in 1st 510= d? If yes, then change to $b(??)/KCY
Changeable by norm rule, but to what?
Can we see some examples

**Invalid tag 779.**

Severe Just remove? Agree/KCY

**1st $6 in 1st 504 is too long.**

Critical $6 is valid, but maybe just remove??? $6=system supplied. Changeable by norm rule, but to what?
Can we see some examples

**Invalid code in 1st $e in 1st 040.**

Critical If ‘RDA’ is entered with typo, list how to flip back to ‘RDA’ Agree/KCY

**Sparse Check failed. Sparse record**

Critical Here we want to break these out with sample error messages we’d like from OCLC. Look at this spreadsheet for examples: [https://docs.google.com/spreadsheets/d/1odQhjLPc4qS6ForJMaEmVgH5s4QibQU_n6iOCAVAXX9U/edit?gid=777884637](https://docs.google.com/spreadsheets/d/1odQhjLPc4qS6ForJMaEmVgH5s4QibQU_n6iOCAVAXX9U/edit?gid=777884637)