Notes from the OCLC Cataloging Community Meeting on 28 January 2022

Access the recordings and presentation slides

Session: WorldCat Quality Update

Presenters

- Robert Bremer, Senior Consulting Database Specialist, OCLC
- Nathan Putnam, Director, Metadata Quality, OCLC
- Jay Weitz, Senior Consulting Database Specialist, OCLC
- Charlene Morrison, Database Specialist, OCLC
- Shanna Griffith, Database Specialist, OCLC
- Hayley Moreno, Database Specialist, OCLC

Presentation summary

Robert Bremer presented on controlled headings changes, specifically regarding the update to "Illegal aliens" in WorldCat®. The Library of Congress announced in November 2021 that this heading would be canceled and replaced by two new subject headings, “Noncitizens” and “Illegal immigration.” OCLC received these updates in late November and began updating WorldCat immediately. Special processing was necessary due to the single term being split into two new terms. As with any large-scale conversion, there was also manual cleanup needed to fix typos, incorrect usage of subject headings, incorrect subfield coding, and more. OCLC changed headings in about 44,000 records.

Nathan Putnam spoke about the controlling project, which includes the WorldShare® Record Manager functionality of “Add Equivalent Subject Headings” to automatically enrich a record with additional headings from Library of Congress Subject Headings (LCSH), Art & Architecture Thesaurus (AAT), Medical Subject Headings (MeSH), and Répertoire de vedettes-matière (RVM). OCLC is planning a retrospective controlling effort that will control all headings that are in the system to their respective authority files and apply heading enrichment. This will update about three quarters of WorldCat records. The plan is for it to begin in February 2022; there will be further updates.

Jay Weitz presented on recent updates and quality control issues. A WorldCat Validation Installation is planned for February 2022, which will include validation of all MARC codes announced November/December 2021 as well as bibliographic and authority updates recently announced. There was a WorldCat Discovery installation in January 2022 with new features and enhancements. A WorldShare Circulation installation and WorldShare Reports and Report Designer updates took place in January 2022 as well. Full details of all these updates can be found in the release notes (URLs below). Jay recommended subscribing to the OCLC Research blog Hanging Together, and he noted that OCLC is currently seeking an Accessibility Strategist. Additionally, OCLC’s suite of Capira library services was recognized with the LibraryWorks 2022 Modern Library Award.
Charlene Morrison walked through some statistics showing Data Quality activity. OCLC and Metadata Quality staff have worked hard to maintain WorldCat’s quality by replacing, merging, and deleting WorldCat records. As of December 2021, Metadata Quality staff enhanced and replaced approximately 2 million WorldCat records in 2021. Over the last year, almost 60,000 duplicate records were removed through staff efforts.

Shanna Griffith spoke about AskQC and OCLC’s virtual AskQC office hour (VAOH). AskQC is a member-outreach program providing access to subject-matter experts. Catalogers can submit questions to the email address askqc@oclc.org. During office hours, OCLC Metadata Quality staff answer questions about cataloging policy, standards, and practices. Since January 2018, Virtual AskQC Office Hours extended this outreach to a live question and answer session, where staff give a brief presentation on a topic of interest to catalogers, followed by an open Q&A. These sessions are archived, and recently the AskQC page was improved for ease of use.

Hayley Moreno spoke on Bibliographic Formats and Standards. Metadata Quality staff continue to update the documentation as needed. Since January 2021, we have created new chapter and page sections and incorporated MARC updates. We have added an 007 General Information page, which provides an overview of how the field may be used, how it is displayed, and the underlying structure. There is also now a new section in Chapter 5, Field Transfer, which provides a highly detailed table of what conditions are considered for field transfer in merging. Bib Formats and Standards has also been updated to reflect the changes from MARC Updates 31 and 32.

URLs mentioned during the event

- WorldCat Discovery release notes, January 2022: [https://help.oclc.org/Discovery_and_Reference/WorldCat_Discovery/Release_notes/2022_release_notes/100WorldCat_Discovery_release_notes_January_2022](https://help.oclc.org/Discovery_and_Reference/WorldCat_Discovery/Release_notes/2022_release_notes/100WorldCat_Discovery_release_notes_January_2022)
- Subscribe to Hanging Together: [https://hangingtogether.org/subscribe-to-hanging-together/](https://hangingtogether.org/subscribe-to-hanging-together/)
- AskQC: [https://help.oclc.org/WorldCat/Metadata_Quality/AskQC](https://help.oclc.org/WorldCat/Metadata_Quality/AskQC)
- Control headings, add equivalent subject headings: [https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Bibliographic_records/MARC_21_view/Control_headings#Add_equivalent_subject_headings](https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Bibliographic_records/MARC_21_view/Control_headings#Add_equivalent_subject_headings)
Audience questions

When Robert speaks of “the client macro” does that mean literally a macro from the Tools > Macros dropdown menu? Or is that just the function that happens when someone invokes Control heading/Control all headings?

It is a macro that OCLC has programmed in the same way that any user of the Connexion® client could use the OCLC macro language to create a macro for their workflow. In OCLC processing, we have a macro that has many improvements that we routinely apply to records. This seemed to be the logical place for us to put in the criteria for dealing with “illegal aliens,” because we continue to load records through batchload processes. Even though we clean up the headings so that “illegal aliens” is gone from the database, it will return. We need to make use of that same logic in a macro over and over.

Regarding controlling in the retrospective, how are things like dissertations not based on subject headings handled?

The “Add equivalent subject headings” feature is only for headings that we have been able to control, so records that do not have any of those four vocabulary strings (RVM, AAT, LCSH, or MeSH) will not go through this process. If we’re unable to control a heading for whatever reason, those will not result in any equivalency additions.

What mapping file(s) are you using for the subject heading enrichment? How are the mappings maintained?

It is coming from 7xx fields in the RVM authority file, so those who create and monitor the RVM authorities are the ones who have added these equivalencies to the other vocabulary systems within their files.

Has the ability to access OCLC’s authority record “cheat sheet” changed recently? I used to be able to hit the ARN link in OCLC and immediately get a searchable MARC field list. Now I get a general page which is not helpful.

The page that the fixed field labels point to is correct. In the past, we reverted back to /support/services because the redirect didn’t function in version 2.63. Although there are no plans to revert again, the team is aware of the inadequacies and are looking for a more workable solution. Currently, the fixed field labels point to the table, with the idea that users can search specific fields via CTRL.

Would the Homosaurus ever become a controllable vocabulary for WorldCat records? And if not, why?

Probably, and there are several other vocabularies that we would like to add as well. We need to make sure we have the right resources in place, and we need to work with the Record Manager team to make sure we get those as well. We are shifting from MARC data to graph data, and there will be additional things we can do with other vocabularies then.

For the retrospective controlling, would it be possible to make sure that 655 _0 headings do NOT get converted to LCGFT (655 _7 etc.) headings?

If the 655 _0 matches a LCGFT heading or reference, then it will automatically control to 655 _7 $2lcgft. If no match is found or if more than one match is found, then the 655 _0 will remain the same. So, if
control-all results in the 655 _0 being controlled to 655 _7 $2lcgft, the same will be true during the retrospective.

We are a WorldShare Management Services (WMS) library, and we've been having a problem with some records in WorldCat created by other institutions with Digital Collection Gateway. The 856 fields in these records often include outdated links that display in WorldCat Discovery, leading our patrons to believe they have access to materials that they can't access. We've been told that OCLC can't remove the links from these records. There are too many for us to edit. Does OCLC have any plans to assess and maintain records like this?

We know this is a problem, and we continue to work on it. We're developing a plan to make sure the kind of things that you're seeing won't continue to be there. It won't be a fix tomorrow, but we appreciate you and others who have reported these. Thank you for your patience as we continue to work on these.