

2019-05-08 VAOH Session

Presentation summary

Robert Bremer and Bryan Baldus explain how small errors in bibliographic records can have a big impact on indexing and retrievability, identification of the resource, copy cataloging workflows, and local system processing. These small errors include incorrectly coded non-filing indicators, incorrect punctuation in authorized access points, incorrectly used characters, inconsistent terminology, incorrect relator terms and codes, and other incorrect MARC coding.

URLs mentioned during the presentation:

Cataloging defensively: <https://www.oclc.org/en/events/cataloging-defensively.html>

WorldShare Record Manager search:

https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Search

PasteUnformattedPlus macro: <https://github.com/wnickeson/WaltsMacros>

Member questions

Are mixtures of Latin and non-Latin characters usually human introduced, for example, someone with a different keyboard encoding manually typing one or a few characters in a field, or...? What is the best prevention?

Answer: It is very likely that they would be human introduced especially if the cataloger wasn't typing on a multilingual keyboard or wasn't used to entering text in a particular language. On the other hand, the cataloger may have copied characters from a Web page and pasted them into the record.

There is also the scenario that script, such as Cyrillic was entered by the cataloger, but the cataloger forgot a letter. Instead of adding the correct Cyrillic letter, though, the cataloger added the Latin form of the letter. This is similar to the transition from typewriters to computers, where the letter "l" was no longer used interchangeably with the number "1". When using computers, these are different characters. So, the same thing is true with the Cyrillic vs. Latin letters. They are different characters and should not be used interchangeably.

Sometimes pasting summaries or abstracts into the 520 introduces character errors.

Answer: The problem is most likely the smart characters when picking up information from a webpage and pasting it in a record.

How do you prevent the smart characters from getting into the record?

Answer: Convert the smart characters before pasting them into the record. One way to do this is using find and replace within a Word processor. Another way would be to create a macro to convert all of the characters automatically. Macros can be created in both Connexion and other programs such as Notepad++. There are also Connexion macros available in repositories such as GitHub. An example

provided by a library in the text chat suggested PasteUnformattedPlus, which can be used when copying text from other sources and can be downloaded from <https://github.com/wnickeson/WaltsMacros>.

The 'Cataloguing defensively' instruction tells us not to correct minor errors. Do the examples given count as major errors or are we now encouraged instead to do these minor corrections?

Answer: You are welcome and encouraged to correct errors in cataloging if you wish. This is a hallmark of the cooperative cataloging environment in WorldCat. We will review our instructions on cataloging defensively to see if any corrections need to be made.

It may also be the case that, the statement to not correct minor errors in the cataloging defensively instruction was a cautionary statement. If the situation is that the difference is minor, then do not change the master record. For example, if the pagination in the bibliographic records is "330 pages" but the resource you have in hand is "331 page" then use that master record but only edit your local copy if you wish to change the pagination. These would be considered copies of the same edition.

However, the small functional errors discussed in this presentation, like a comma missing from a personal name, can have an effect on that name being found, or if the non-filing indicator is incorrect, that has an impact on that title being found. These are the minor errors that are worth fixing.

I find a number mistakes with the FAST headings. Is it a good idea to just delete FAST headings?

Answer: The answer depends on the context. If the FAST heading is inappropriate for the item, then since they were generated from the LCSH, you are encouraged to fix both the FAST and LCSH in the master record. If the error involves a FAST heading that was mapped incorrectly from the LCSH, then report these errors to askqc@oclc.org so that staff can fix the mapping to prevent the problem from being repeated on other records. If the FAST heading was manually added to the bibliographic record and a typo was introduced, then you are welcome to correct the heading or report it to bibchange@oclc.org for staff to review and correct.

How often do you do this correction work? Is it constant or project-based?

Answer: For the things previously mentioned, we sometimes rely on people spotting and fixing the error themselves when possible. There are certain errors that we can go looking for because they are readily found. In some of these cases, the number of occurrences is overwhelming so that we are constantly trying to catch up.

One example of this is the incorrect non-filing indicator. Entering the correct non-filing indicator often does involve a detailed knowledge of the language of the resource. It is not as simple as setting the non-filing indicators to a certain value if you see a particular word. For example, the word "the" in an initial article in English but if there are diacritics such as in Vietnamese or French, then it should not be coded as an initial article as these are different words altogether. Another example is the letter "A" in English. Titles for children's works might begin with "A is for... ". In this case, we want to file on that "A" since this is talking about the letter "A" and not the initial article.

Some of these errors, while small, do take a lot of effort to correct, although some so lend themselves to macro corrections. For example, spacing problems related to the comma used in a personal name access point. This would affect how the name would appear in a browse list. It is difficult to insert a comma into

a personal name that is missing one. To insert the comma, you need to determine where the surname ends and the forename begins as well as how many words the surname includes.

In your slides you showed 'WorldCat browse list', for titles, for Personal names. We use RecordManager, and haven't found such functionality. Does it exist in RM? We could really do with such browse list in order to check Authorized Access Points!

Answer: While there is browse functionality in Record Manager, it is not exactly the same. A library suggested in chat that one way to browse in Record Manager is to change the scope to "All WorldCat Browse" then use an index such as "Personal Name - Whole Phrase". More information about searching using [browse](#) in Record Manager can be found on the [WorldShare Record Manager Search](#) page. If you have further questions about the functionality, please contact [OCLC Support](#).

I see a lot of subject headings coded as LCSH when they are not valid LCSH headings and more likely something used locally. What is the best way to handle this situation?

Answer: You can re-code any non-LCSH to show that it is a local subject heading by using second indicator 4. You are also welcome to report it to askqc@oclc.org so staff can investigate it further especially if you are seeing a pattern from a specific institution.

Going back to correcting minor errors, sometimes I have wondered if some corrections were based on opinions rather than correct vs incorrect. For example: For a week, I got the same record with 100 field corrections DR. then Dr then DR. again and back and forth.

Answer: The name should be based on the author's usage and you would normally apply rules for capitalization, which would mean that "Dr." would be the chosen form. You are welcome to send this particular record number to askqc@oclc.org and staff can take a look to see what is happening with this record.

Will OCLC allow 020 added for reprint periodicals? Here is the email I got from Les Hawkins, Coordinator of CONSER: "publishers will frequently issue an ISBN for a complete reprint, possibly of a ceased serial or as your example shows an ISBN for each of the 6 volumes of a ceased serial., This might be a good case for OCLC changing the validation of the 020 fields if there is a way to distinguish when the someone is recording ISBN for reprints vs. when they are recording ISBN for each issue of a (non-reprint) serial."

Answer: This is an issue that has come up now and then. Back in the days of AACR2, instructions in Chapter 12 told catalogers to record an ISSN in the record but there was nothing that said to record an ISBN even though we treated serial reprints that came out all at once as serials and cataloged them according to Chapter 12. Reprints of serials will sometimes have ISBNs printed on them and so they do seem valid to include them. In RDA, instructions tell catalogers to record standard numbers, which leads to the question of including ISBNs in serial records. This situation was brought up a few months back on the CONSER list. While adding ISBNs to serial records would be nice to do, it is technically difficult for us to do at this point. In the MARC environment, you don't want to have the situation where all of the ISBNs relating to individual issues are listed in separate 020 fields. If you did this, you might end up with 100's of 020 fields in a serial MARC record, with each 13 digit version being duplicated by the 10 digit version. The consensus in CONSER was that would be nice to do for reprints at least, but not at this time.

When we move off of the MARC format, this would be much easier to implement. For the time being, we suggest that you add these ISBNs to your local copy of the record.

Are AU@ Libraries Australia MARC records merged with other eng cataloging rules MARC records (say PCC or DLC) or are AU@ MARC records retained as parallel records?

Answer: All records cataloged in English are merged to any other records cataloged in English if they are duplicates, including AU@ record.

A record for Stephen King's original Pet Sematary has the ISBNs for multiple editions, including the new movie tie-in. This was very confusing. The Stephen King Pet Sematary is #14046555 or #9576009. published in 1983. None of the incorrect ones are in subfield z.

Answer: The ISBNs could have been matched to an existing record in WorldCat representing the book from some years back, however, the ISBNs for the newer editions and the movie tie-ins don't belong on the master record. For ISBNs, you would record all ISBNs that appear on the item in hand in field 020. The ISBN representing the resource itself would be recorded in field 020, subfield \$a and ISBNs representing related resources would be recorded in field 020, subfield \$z. If ISBNs don't apply to the item that the record describes, such as a movie tie-in or a later edition, and they don't appear on the item that the record describes, then those ISBNs should not be added to the master record. For this particular example, we've cleaned up the records you've mentioned along with some related records based on available online proof and evidence found in OCLC's Journal History.

If you do come across errors that you are unsure whether to submit or correct, please feel free to reach out to askqc@oclc.org or bibchange@oclc.org and staff would be happy to investigate further. If a large problem is identified, then staff can perform scans to compile the records for correction.

I'm cleaning up some old records and matching them to OCLC records. I find a number of these older titles have multiple mergers in the 019, and there are fields for a narrator on a print record, a note for the artistic qualities of the illustrations for a presumably special edition.

Answer: This is a good example of errors to report to askqc@oclc.org for further investigation. Staff will be able to determine if an incorrect merge occurred or if there is something else at play.

I semi-regularly see ISBNs on records for early books that couldn't have ISBNs.

Answer: It is possible that there was a reprint that had an ISBN that was added later. In this case, it would be appropriate to keep this ISBN on the master records since it represents the same manifestation.

In Record Manager, there is a filter to search only OCLC Publications New Zealand records in Advanced Search. Is this filter based on the holdings of NZ or by some codes in the records?

Answer: The filter option for "OCLC Publications New Zealand records" in Record Manager limits the search to records with the code "nznb" in field 042. This would retrieve results similar to if you limited a search by [authentication code](#) in Connexion, FirstSearch, or Discovery.