

2018-10-10 VAOH session

Presentation summary

Cynthia Whitacre presented on the topic of language of cataloging and parallel language records. The language of cataloging is characterized using a single language in cataloger supplied elements. The code for language of cataloging may be found in [field 040](#) subfield \$b in the record. If subfield \$b is absent from field 040, the implied language of cataloging is English. Language of cataloging (Field 040 subfield \$b) relates the descriptive cataloging, while the language of the item ([Lang](#) fixed field) relates to the language of the content in the resource itself. Subject headings do not determine the language of cataloging in a bibliographic record and may be in any language on any record. DataSync collections may be set up with a specific language of cataloging code within the collection profile setup, which provides a crucial match point.

Parallel records are multiple records for the same manifestation that describe that manifestation in different languages of cataloging. Each language of cataloging is allowed one record per manifestation. Do not report parallel language records as duplicates to WorldCat Metadata Quality staff; however, do report duplicate records within the same language of cataloging. For example, report two German language records that represent the same manifestation, but do not report a German language record and a French language record representing the same manifestation. Elements to look for when determining language of cataloging include: the authority file used for non-subject heading access points, relator terms used in access point fields, words and abbreviations in subfield \$a and subfield \$b in field 300, and terms used in 33x and 34x fields.

If, after a thorough search of WorldCat, a record does not exist in your language of cataloging for a resource, please create a parallel record in your language of cataloging. You may derive a record from a parallel language record to create a record in your language of cataloging but be sure to change all appropriate fields to your language of cataloging. When correcting records, if the language of cataloging is coded incorrectly or appears to be a hybrid, you are welcome to make the correction yourself or report it to bibchange@oclc.org to be corrected. When correcting hybrid records, consider the intent of the inputting library, the language of the descriptive fields in the record, and the language of the majority of libraries who have attached holdings. Some non-English libraries do accept and input record with English language of cataloging and some libraries base their language of cataloging on the language of the item. When in doubt, sent hybrid record requests to askqc@oclc.org.

For more information on language of cataloging and parallel records see BFAS [3.10 Parallel Records for Language of Cataloging](#). Note that these instructions may eventually be moved to BFAS Chapter 2.

Member questions

What is GLIMIR clustering and what does it have to do with parallel records?

Answer: GLIMIR stands for GLoBal Library Manifestation Identifier. GLIMIR is an effort by OCLC to bring together holdings for the same manifestation that are distributed across multiple parallel language records. GLIMIR began as a project in 2009 and was fully implemented around 2012. When WorldCat was GLIMIR-ized, the same or similar records were clustered together to improve end user searching. Since the purpose of this tool was end user searching, it is not very useful in a cataloging context. It is much more useful in WorldCat Discovery or WorldCat.org since it brings together all of the parallel records as well as print and microform records representing the same manifestation. For catalogers, make sure the GLIMIR box is not checked when you search in a catalog interface.

Why is GLIMIR clustering turned on by default in Connexion if it is not really useful for that interface?

Answer: GLIMIR clustering can be disabled. Once it has been disabled, the system will remember that preference. If it is not, make sure that only one instance of Connexion Client is open, open the search dialog box, and uncheck the "Display using GLIMIR clustering" option. Search WorldCat, then exit out of Connexion Client. The Client will save the last preferences that you chose before exiting the Client. When you open the Client back up, the "Display using GLIMIR clustering" should no longer be selected.

There is a source in Japan (TRCLS) that codes 040 as eng but generally does not follow English language rules on most of their records, e.g. the author is in a 700 instead of 100.

Answer: TRCLS is a vendor institution in Japan and their intent is to catalog in English. While their intent is to create an English language record, they do not catalog the same way we would catalog the resource. These should remain, English language of cataloging records, so please feel free to correct these records as needed.

What is the national French library code, not Canadian French, but French French?

Answer: The OCLC symbol representing the Bibliothèque nationale de France is BDF. To identify the OCLC symbol of a particular library, go to the [Directory of OCLC Members](#) and search by the institution's name. As far as the language code to put in field 040 subfield \$b, the code would be "fre" whether the language of cataloging is French from Canada or French from France.

I have seen a lot of 520 in foreign languages, quoted directly from the book jacket of a novel. These are done by English-language cataloging agencies. Shouldn't the quoted foreign-lg note be in a 500? Is 520 supposed to be in English if your cataloging lg is Eng?

Answer: If the 520 is a quoted note, then that is perfectly fine to keep in the record no matter what the language of cataloging. If your library has the language knowledge to add a translated summary note in the language of cataloging of that record, you may do so and replace the quoted summary in the other language. If you are cataloging in English then, yes, the summary note in the 520 field is supposed to be in English as well. Allowances are made for libraries serving multiple communities, for example a library who catalogs in English but also serves a Spanish language community, may keep the quoted Spanish language summary statement along with the English language summary note. If you see these translated summary notes in a record, you may fix it to represent the language of cataloging of that record.

When searching OCLC Connexion, how do we search English records? Which option do we pick, "Apply language of cataloging limiter" or language? There are two choices.

Answer: To limit records to a particular language of cataloging, use the "Apply language of cataloging limiter" option. You may also use ll: if searching in the command line search. For example, ll:eng would limit your command line search to only English language of cataloging records.

Should we check to see whether fixing a hybrid record will result in a duplicate for that language of cataloging?

Answer: If you notice that correcting the hybrid record will result in a duplicate record for that language of cataloging, we encourage you to report them as duplicates to bibchange@oclc.org. In many cases, DDR will come along a few days later and find and merge these records. For more information on DDR, see [Defending Differences from Duplicate Detection](#) and BFAS [5.1 OCLC Member Quality Assurance](#).

Is it corrected to say records in traditional Chinese and simplified are considered as the same record and get merged?

Answer: The records may be merged but Metadata Quality staff would first consult with experts in the language before making a final determination on the matter.

When deriving a new parallel record, should the same 010 be given in the new record?

Answer: The same field 010 may be used in a parallel record since the identifier in that 010 field represents the same resource.

I've seen a few records created by Dutch language institutions that have 040s with eng as language of cataloging but the 700 fields with (NL-LeOCL). Do you think the intent was English language cataloging or was it Dutch? Sometimes the 3xx fields may only be partially filled in.

Answer: It is common practice for some Dutch academic institutions to catalog certain records in English. If all of the fields except for the access points are cataloged in English, then most likely the intent of the institution inputting the record was to catalog it in English even if they have used the Dutch authority file. Verify that field 300, the 33x fields, non-quoted notes, and other elements are English language of cataloging then you may correct the form of the access point to conform to the English language of cataloging practice. It's also worth noting that a number of Dutch academic libraries are planning on joining NACO and will start creating English language authority records. This should make a difference for records contributed by Dutch academic libraries in the future.

What is the status of the "PR" notes in 936 fields? Can you explain what happened to field 936?

Answer: When parallel records were introduced in 2003, OCLC also introduced "PR" notes in field 936. This note contained a list of OCLC numbers representing parallel language records for that same manifestation. Because this field was not always used as intended, OCLC stopped using the field altogether a number of years ago and deleted them from WorldCat.

Do we add the language of cataloging or the language of the item to the 3XX fields? Is it mandatory?

Answer: When entering the codes into the 3xx if your institution enters English language records and is following RDA guidelines, you do not need to add the language code after the code in subfield \$2. For example, If you were cataloging an item in English under RDA guidelines, you would expect to have English language terms in the 33x fields subfield \$a with rdamedia, rdacontent, or rdacarrier in subfield \$2. Under RDA guidelines, there is no need to include the slash "eng" (/eng) after the code in the subfield \$2. If you were cataloging in a language other than English, though, you would include the appropriate term in subfield \$a for that language and then follow the source code in the subfield \$2 with a slash and the appropriate language code.

A lot of records from foreign libraries are batchloaded and I see the "language of resource" in 008 coded to be the same as language of cataloging when in fact it's a different language. Do you encourage OCLC members to correct such errors?

Answer: Yes, please correct errors that you find in these records. If you are unsure or see a pattern of errors, please send these to bibchange@oclc.org for WorldCat Metadata Quality staff to review and correct and needed. Metadata Quality staff correct these errors on a regular basis, but member libraries are also encouraged to correct them as well.

I wonder if the OCLC team knows if it's possible to link fields after they've been exported to an ILS?

Answer: No. You will need to contact your specific ILS provider and work with them.

You mentioned subject headings can be in any language. Does that include allowing non-English uncontrolled (653, 650_4) SHs without \$2 in English-language records?

Answer: Yes.

We put 520s in Spanish for our Spanish-language patrons and use a second one for the English version. I have seen some folks say we should use only 1 520 with both languages in it. Which is preferred?

Answer: OCLC prefers that summary notes added to field 520 match the language of cataloging of the record. While it's understandable that a library would want to include summary notes in other languages, its preferred that libraries treat these other language summary notes as local notes and either add it to your record locally or add it to an LBD record.

All simplified characters vs All traditional characters should not be merged because it's likely it's a different publisher.

Answer: We urge you to look carefully at all elements in the records but if you do feel like they are duplicates, please report them to bibchange@oclc.org. We will make a determination whether to merge them or not, consulting with our language experts if needed.

From what you said, you can submit duplicates for other languages of cataloging in \$b of 040. Is it allow to update those records if there are wrongly displayed diacritics?

Answer: Yes, please feel free to correct errors if you see them.

Can you derive a new record from a parallel record?

Answer: Yes, you may derive a new record from a parallel record. When you do, be careful not to enter a hybrid record. Make sure to verify the non-transcribed elements to make sure that the language of cataloging matches the language of cataloging of your institution.

Most CJK records I see have 520s only in the vernacular and a transliteration. So, all of these should only have been in English?

Answer: It depends on the language of cataloging of the record. Look at the whole record to determine what the intended language of cataloging was before deciding what to do. If the language of cataloging is determined to be Chinese and the summary notes in the 520 fields are in the vernacular with a transliteration, then do not change this record. If the language of cataloging is determined to be English and the summary notes in the 520 fields are in the vernacular with a transliteration, then you may correct the 520 field. Yes, OCLC does allow the practice of adding a quoted summary statement in CJK along with an English summary note in an English language record.

How long does it takes after you report a duplicate record via the OCLC form on the web for the duplicates to be processed?

Answer: When you send duplicate requests using the OCLC form on the web, the request is sent directly to bibchange@oclc.org. Once its sent to Bibchange staff, its placed into the duplicate workflow. In general, staff trained in each format process duplicate requests on a first in, first out basis. Be aware that there is currently a backlog in duplicate requests. Some formats have a bigger backlog than others, but your requests sent through the online form will be sent to bibchange@oclc.org and processed when staff are able to get to them. For more information on Bibchange Staff workflow, please see [Processing change requests](#), the Virtual AskQC Office Hours presentation given on March 28, 2018.

There are a lot of pre-2003 records that are hybrid records (040 says eng but the notes are in French or Spanish, for example). In general, is it better to fix them or to derive a new one? What happens to all the holdings that were attached to the old record if you derive a new one? Will they move or stay?

Answer: When you derive a new record, you are creating a new record, so the holdings will not move. Holdings only move if records are merged together. It's a judgement call whether to fix the record or just derive a new one.

You should determine the language of cataloging based on the intent of the cataloging agency who input the record, the language of the libraries who have attached holdings, and the language of the descriptive cataloging elements in the record. Deciding what language of cataloging should be used on a record needs to be determined on a case by case basis. It is worth looking at the hybrid record in

WorldCat to see if you can resolve the problem since incorrect coding in field 040 subfield \$b may lead to an incorrect merge via DDR. Also, you may have a case where considering the holdings will inform you on the best course of action. If a record was intended to be a French language of cataloging record but only 2 out of the 102 holdings are French libraries, while the other 100 libraries are English language libraries, then it may be that it should remain English and be corrected to reflect English language of cataloging instead. If you can do this, then feel free to modify the record as needed to correct the hybrid record, otherwise you may report the record to bibchange@oclc.org.

Does evidence that names are controlled to a non-English authority file, does that count in deciding what the intended language of cataloging was?

Answer: If the names in a record are controlled to a non-English authority file, then the language of cataloging ought not to be English. Records cataloged with English language of cataloging, are controlled to the NACO authority file. Records cataloged with Dutch language of cataloging are controlled to the Dutch language authority file. The same with German, French, etc. It's unlikely that you will see an English language record controlled to another authority file other than the NACO authority file, because of how OCLC controls authorized access points. There may be cases where a record coded as English language of cataloging has access points with subfield \$0s that link to a non-English authority file. If you encounter this, carefully review the record to determine which access points to retain and edit the record accordingly.

What if there are two authority files in the same language? e.g., French from Canada, and French from Europe?

Answer: This is a great question and is something that we have been wondering what to do about. So far, we have the French-Canadian authority file that we will be using to control access points in French language records. We currently do not have a French authority file from Europe yet. We are thinking about it and carefully considering what we will do in the future but in the short term there is no conflicting French authority file to consider in the short term.

Now that credits are long gone what is the incentive to use different encoding levels? Such a K vs I ?

Answer: While there may no longer be a credit incentive as there was in the past, encoding level K and I communicate to other catalogers the record's level of completeness. Updating the ELvl fixed field when upgrading the record will assist others in identifying what your intent was when cataloging it. Please feel free to upgrade minimal level records to full level records if you come across them.

I am seeing a lot of 856 fields with "Table of contents" and " |z Available to ----affiliated users at |u". If a note indicates that the resource is only available to one institution, should it be deleted?

Answer: Yes. You may delete local URLs from the master record. For more information on URLs in field 856, see [URLs in a shared cataloging environment](#), the Virtual AskQC Office Hours presentation given on April 15, 2018.

How long do vendor records with titles that actually say untitled stay in OCLC? There are some going back to 2017, many are level M.

Answer: Level M means that the record came in through Data Sync, it doesn't indicate the completeness of the record. We do work to clean these up as we see them. The current policy for vendors is to delete vendor contributed records that have existed in WorldCat for about 4 years and do not have library holdings attached. If you can identify one of these records by ISBN, you are welcome to upgrade it to match the item you have in hand. You may also report these to bibchange@oclc.org.

What is datasync?

Answer: This is the new OCLC batch load system and stands for Data Synchronization. Libraries send us files of MARC records and those files are taken in through the Data Sync system so that the records are either matched to records or added as new records in WorldCat in a batch mode.

To the question of 856 fields: If we find a particular institution that is adding these fields, should we report it to OCLC, so you will delete them from the database?

Answer: Yes, feel free to delete these or report them to bibchange@oclc.org if you see a institution's local URLs being added to master records. Staff are continually working to remove these from master records and anticipate changes will be made behind the scenes this fiscal year so that less of those will get added to WorldCat.

Merging records, do you encourage members to report multiple records for merging?

Answer: Yes, we do encourage members to report multiple records for merging. Please send all duplicate requests to bibchange@oclc.org.

Why do so many records, especially level M, have data that can't pass OCLC's own validation program? We do an awful lot of cleanup on fields that have to be fixed or just plain deleted if we can't fix the problem. Why doesn't your data ingest program validate records first?

Answer: The Data Sync process does make use of validation; however, it's handled differently than online inputs. The reason for this is that when you catalog online and receive a validation error, you are able to fix the validation problem at that time, but when records arrive in files through Data Sync, they are run through validation and separated into significant errors and minor errors. The significant errors include a bad tag or an incorrect record structure. These records are set aside and not loaded into WorldCat. The minor errors include relationship errors, such as if you have one of these then you must have one of those. These records are loaded into WorldCat. If these records with minor errors were not added the copy would not be made available for other libraries to use and the library's holdings would not be added as well. We do realize that validation errors are a problem and have change instructions up front where we try to fix some of the errors that are coming in. Metadata Quality staff also have macros and tools to clean up the records as they are loaded as well. If you are seeing a pattern of a problem, please report these to askqc@oclc.org. We may be able to fix the one record you are reporting but also make the same change across the thousands of records that have the same problem.

If we upgrade a master record to PCC, are we responsible for checking the validity of the 856 fields?

Answer: No, if you are upgrading a master record to PCC, you are not responsible for checking the validity of the 856 fields. That being said, we encourage you to look at them and delete any URLs that are obviously local to an institution. You may not be able to check all of the URLs because your institution may not have access to the providers.

I was told that enhancing an OCLC bib will not receive monetary credit whereas creating/deriving a new bib will receive monetary credit. Is that true?

Answer: No, that is not true. In the past there was a system of credits, but the credits have been discontinued for quite a few years now. So, you are not going to receive a monetary credit for creating or enhancing bibliographic records. For guidelines on when to input a new record or use a record already in WorldCat, see BFAS Chapter 4, [When to input a new record](#).

Is there any update on fixing the corrupted copyright symbols in the 264 field or corrupted diacritics? I'm seeing a lot on recent DLC records.

Answer: This is similar to the issue where more 856 fields than we would like are transferring to the WorldCat record. There is an effort underway to address the corrupted copyright symbol problem so that we don't transfer in these fields as often as we have in the past. The problem that we have with some of these corrupted diacritics is that they turn into a character that, while a valid Unicode character, is one that we do not want in the master record. They transfer in because it looks to the system like it's a valid non-Latin script. We are currently working to resolve this problem. Once we get the point where we are no longer transferring them to the extent that we currently do, then we will start the cleanup process in WorldCat. Unfortunately, as we have cleaned up these records, we have seen the same errors reappear on the record from a different library's Data Sync load on the same day. Because of this, we are focusing efforts on resolving the underlying source of the problem before going in and cleaning up the WorldCat records that are currently affected by this problem.

I have seen records with multiple merges and after researching the merges have found (for example) DVD records that have 2 discs where the current record displays only 1 disc. Are these records checked for number of discs before merging? Two of the 4 OCLC records merged had differing numbers of discs.

Answer: Yes, that is something we do check for in both DDR and manual merging. If you suspect that records have been incorrectly merged, please report them to bibchange@oclc.org and we will look into the records through OCLC's Journal History and if appropriate, we will recover the records.

If you're looking for common validation problems, one that we find frequently that stops validation is in 775 or 776 fields that have one extra space in a \$w.

Answer: This is good to know. Staff have gone through and fixed this spacing error before, but we appreciate knowing that these are coming back. We'll work on cleaning these up again.

Is there any way you can get certain institutions to update their records for online resources to reflect the publication vs. their digital imaging data in the 260/264 and fixed fields? They will often have duplicate records coded, one coded mixed material records and the other coded as an online monograph. These are not always Dublin Core vs. MARC.

Answer: Most likely these are [digital gateway records](#), which maps Dublin Core data to MARC 21. This often results in a mixed materials type, so it would not be added as a book. Because they are not necessarily constructed according to the same cataloging rules that we would use for other materials, we do not merge these records at this point. If there is a specific institution, please contact askqc@oclc.org so staff can look into the problem to see how significant it is.

I am not sure if it is an appropriate question but which language should be used in the fixed fields in bilingual DVDs?

Answer: The [Lang](#) fixed field contains the language of the resource itself. If the resource is bilingual then you would use both the Lang fixed field and [field 041](#) to code the languages of the resource. For example, if the DVD of a German film included dialog in both German and French with subtitles in English and Spanish, the Lang fixed field would contain "ger" while the 041 field would contain all of the languages included in the appropriate subfields. For example:

Lang: ger

041 1 ger \$a fre \$j eng \$j spa

Was that 20-minute presentation the topic in its entirety?

Answer: The presentation given today was an overview of the topic. Please refer to the guidelines for more information and further examples. Metadata Quality staff are currently working to clarify the guidelines for Parallel language records, which is currently located in BFAS 3.10, [Parallel Records for Language of Cataloging](#). These guidelines, once refined, most likely will be moved to BFAS Chapter 2. And, if you have additional questions, please send them to askqc@oclc.org.

Can we report duplicate level M records by sending the ISBN to Bibchange or do you always need the OCLC numbers?

Answer: Yes, you may send ISBNs as duplicate requests to bibchange@oclc.org.