

Presentation summary

Jay Weitz and Bryan Baldus presented on updates to record validation and the OCLC MARC update. The topic included the principles behind validation and some real life examples of some of the common messages you may see.

Member questions (2020-04-07)

Is there any place that shows differences between authorization levels?

Answer: There is a [table](#) in Bibliographic Formats and Standards, section 5.2, that outlines the different authorization levels and capabilities for Connexion. You can also find information on the support site at this

link: <https://www.oclc.org/content/dam/support/connexion/documentation/client/catalogingauthorizationlevels.pdf>

Why are so many validation errors allowed in group batch processing?

Answer: When working on a single record, it is relatively easy to see what the validation errors are in order to fix those. When we are talking about large numbers of records, there is no way for a cataloger to individually track each record that has a validation error. There are reports that give the validation errors after processing is completed in DataSync that the institution may choose to follow up on and correct the validation errors in the records in WorldCat.

How do staging records play into validation?

Answer: There is validation that goes into DataSync but it's not as strict as online validation, so more validation errors may be present in new records that are added to WorldCat. Incoming records are validated during processing and if they have validation errors, there is a level assigned from 1-3, with 3 being the most severe and will prevent records from being added to WorldCat as fully indexed without the records being corrected. Level 1 errors are minor and generally very easy for users to correct if they encounter them when working with records.

Are these validation checks only run on single records when they are saved, or are they run on batch loads too?

Answer: There are validation checks that are run on single records on certain points in the workflow including when you try to replace a record. There are also validation checks on batch-loaded records as well, as described above.

I occasionally run into records that have validation problems already in WorldCat. Are records validated before batch loading?

Answer: Yes, but there are less serious validation errors that may make it into WorldCat in order to get the records added to WorldCat.

How is verifying MARC-8 characters related to this process?

Answer: Since we removed checking for validation for Unicode characters, allowing pretty much any characters to be used in authority records, we don't provide feedback on when characters outside of the MARC 8 rule set are present. We rely on our users to enter the correct characters.

Member questions (2020-04-16)

Is there a list of validation errors?

Answer: No. The variation in the possibilities of validation messages would be huge. It's not possible to list all validation errors.

After successfully uploading individual Level I English language records of Hebrew print materials, after some time I find they have been changed to Level M. How is this possible?

Answer: An example would really be helpful but it's possible that the record is being overlaid by the same institution when the record is being sent back through their DataSync project and overlaying itself. If you encounter this situation, report that to bibchange@oclc.org so we can look into it.

Are those potential corrections only up to the institution or does Quality Control get involved?

Answer: Metadata Quality gets involved when made aware of any quality issues, including validation errors. Report those to bibchange@oclc.org when you encounter them.

When authority records are added to WMS, are there any authority records being put to staging?

Answer: No. NACO authority records automatically get put into the queue to be sent to LC. Staging only applies to bibliographic records.

I frequently come across Encoding Level M records that are completely full, and have always assumed that these are batch-loaded records. Since we use automated algorithms to find copy cataloging records, these M records are rejected. Is there a way that these records can be more correctly coded?

Answer: Any record that is coded M by definition has been batch-loaded. As you may or may not have heard, we are in the process of trying to slowly do away with the OCLC defined Encoding levels, including I, K, and M, and making the transition to using only the MARC defined codes which is blank and numeric codes. The thing about Encoding level M is that it represents two things that should not be combined into one code. It does mean that the record was batch-loaded, but it also used to mean that it was a minimal level record, but that is absolutely not true in many cases, they can be of any fullness. If your algorithms are rejecting M as minimal, that is something you should look into changing. As part of our eventual transition of Encoding levels to using exclusively MARC 21 codes, we are going to try to assign more accurately encoding levels when we convert them from I, K, or M to a numeric codes and make them more reflective of the bibliographic records themselves.

Our VAOH in June is going to be on the subject of encoding levels, so you might want to tune in.

I was hoping that there was a list with solutions. I received a 1e error in Connexion (might not be the exact error message, since that happened last week).

Answer: You can send the OCLC number bibchange@oclc.org and we could investigate to see what is going on with the record.

I often see validation errors in the 084 field for other classification number. Figuring out what's wrong with them can be challenging if you don't use that classification scheme or if they have been incorrectly placed in the 084 field.

Answer: We would encourage you to report that to us and we will try and figure it out.

When I see "049 too short" I just delete the field, but I don't really understand what it is about.

Answer: The 049 code has a requirement of 4 digits. It isn't retained in the master record, so if you delete it, it's not hurting anything.

Are validation scans being continuously run, like DDR or other quality control routines?

Answer: Only on authority records as they come in from being distributed back from LC. We get a report of validation errors in NACO records and we correct those.

Can you go over again what happens when an authority record is submitted with non-allowed diacritics? In that scenario, how long does the authority record usually stay in distribution until it is corrected?

Answer: When a record is sent off to LC, we will get a report back, usually the next day, that the record is being rejected because of incorrect characters. Metadata Quality staff go through that report usually on a daily basis and correct the records. It may take several days for the record to go back through the distribution. Our turnaround for authority records that are stuck in distribution is pretty quick. If you find a record that has been in distribution for more than a week, send an email to authfile@oclc.org.

Non authorized MARC fields and subfields that have not been implemented by LC should not be used in authority records. That can also cause records to get stuck in distribution.

In our weekly OCLC error report for batch-loaded records, sometimes we see "Sparse Check failed. Sparse record" What does it usually mean?

Answer: Please see Sparse records information at this [link](#).

Are local holdings records validated?

Answer: Yes, and for the most part they follow the MARC 21 holdings format.

How is it that some records get into the system with validation errors and then when we try to set holdings, we are unable because of these errors. Are these not caught by machine?

Answer: There are various validation levels for records that come in through DataSync and some of the less severe errors are let in that can be corrected either by the institution or by OCLC. Batch-loaded records may have validation errors that would not occur if a record is added online.

I use LC classification, but often get validation errors for 084 fields. I don't know how to correct them so I can replace the master record.

Answer: If you are not sure how to correct an error, you can report those to bibchange@oclc.org and we will correct if for you. It's also useful to have patterns of errors reported to us so that we can relay the information back to the contributing institution to have them correct their records in future loads. We are also working on a solution to prevent validation errors on this particular field from entering WorldCat.

Can the last validation error mentioned by Bryan be fixed? I have produced NACO authority records only to find out weeks later that the record was never produced because it contained a pre-composed character with a diacritic, which I didn't know about when I produced the record. Can't validation catch this?

Answer: We've been looking not how records with precomposed characters get into the process and somehow are exported with the pre-composed characters. It seems that there may be something within the Client that de-composes the characters upon display, but still sends them to LC as pre-composed. Nothing is transforming them as they should be and we are looking into it to try and figure out where it's not going as expected.

When getting an invalid character message, it is sometimes very difficult to determine exactly where/what the error is, it's necessary to count the characters. Would it be possible for the error message to display what the invalid character is?

Answer: Unfortunately no. Validation can't be that specific because of the way that the messages are used from templates. They can't be as specific as "this particular character" is incorrect.

(OCoLC)01837402 has multiple 505 fields with pagination - I just have not seen multiple 505 fields created in that fashion. I suppose since this is variable, there isn't any kind of check on this type of field?

Answer: These field transferred during a merge. The record will be corrected.

Another instance I have found when no error message is displayed, but the record will not validate, is when 336 \$c is used erroneously instead of \$2. Why is this?

Answer: That is a mystery, so please report those in the future if you encounter it again to bibchange@oclc.org. All of the validation errors are created by templates and they have to be manually formulated by relationships that are checked by validation. So if there is an error that doesn't generate an error message, it must be that we missed that.

Sometimes I want to change a monographic record my library created and my library is the only holder to a continuing resources record, but after I make all the changes it says I am not authorized to change the Type and I lose all my edits as well as the saved record if I saved it. Is there a way to avoid losing the edits?

Answer: If you are the only holding library, you should be able to make changes to Type and Blvl. Report that problem to bibchange@oclc.org to be investigated if you encounter this problem again.

How do we deal with these validation errors that are outside our knowledge areas? e.g. Dewey, etc.

Answer: You can report these to bibchange@oclc.org.

I thought you used to be able to catch precomposed characters in Connexion with Edit -> MARC-8 Characters -> Verify, but it hasn't worked for me lately -- did this change?

Answer: This did change when OCLC implemented Unicode, so now all characters are valid within Connexion. So this particular verification is no longer able to be used.

Why not just switch the Marc-8 characters to Unicode in authority records?

Answer: LC's system does not yet accept characters that are outside of MARC-8.

Is there a way to find out the validation rules for local holdings records?

Answer: The use of the local holdings record is described [here](#).

Are records displayed in worldcat.org composed from multiple records and is there a way to find out what those are? For example this record has a content note from a different book <https://www.worldcat.org/title/land-beyond-the-mountains/oclc/32703514> but I don't see it in the OCLC records. Why?

Answer: In WorldCat.org fields can be populated into a representative record from other bibliographic records or imported from a 3rd party provider. These include summary notes, abstracts, and contents notes. Errors should be reported to bibchange@oclc.org and if the data is not from a bib record, we will forward the request on to OCLC Customer Support to be removed.

Also related to worldcat.org -- A coworker told me that holdings used to appear as soon as he added the library's holdings symbol via Connexion Client but now that does not happen. Was there a change?

Answer: Holdings may not show up right away in WorldCat.org depending on browser settings. It is possible a cached version of the page is being displayed, to see the immediate change in holdings you may need to adjust your browser settings.

Other reasons they might not appear, the member has a cataloging subscription, but no subscription to WorldCat Discovery/FirstSearch. You need to have both subscriptions to see your WorldCat holdings in WorldCat.org.

Here is a link to our Help documentation on "Why aren't my library's holdings displaying in WorldCat.org?"

https://help.oclc.org/Discovery_and_Reference/WorldCat-org/Troubleshooting/Why_arent_my_librarys_holdings_displaying_in_WorldCat_org