

Marmot Union Cataloging Committee
Wednesday, May 23, 2018
Minutes

X	Jamie Walker [Chair]	CMU	X	Betsey Dick	MCPLD
X	Mary Walsh	Adams State	X	Shelly Fratzke	SD#51
	Gail Owens	Basalt		Wendi Weinstein	Pine River
	Nina Darlington	Bud Werner	X	Carol McArdell	PCL
X	Kim Grindrod, Jan Quick	BVPL		Debbie Barnes	Rampart LD
	Oliver Schulz	CCU		Becky Nelson & Kristin Trombo	Salida
	Amy Currier	CMC Glenwood	X	Sarah Nordholm	Summit Co
	Chris Bockstiegel	CMC Leadville		Jo Norris	VPL
	Jonathan Beam	CMC Steamboat		Nicole Becwar	WSCU
X	Diane Levin & Jolanda Durbin	EVLD	X	Amy Sieving, Alison Farnham	WPL
X	Selene Gardner	EPL	X	Tammy Poquette	Marmot
	Martha Talman, Lara Aase, Michael Bovee	FLC	X	Lloyd Chittenden	Marmot
	Amy Shipley	Garfield Co		Brandon Cole	Marmot
X	Anna Szczepanski	Grand Co		Sean Hanson	Marmot
	Jennifer Hillebrandt	Gunnison			

Completed action items

Action item	Report
Experiment with moving 690/691 into field group 'v'	Putting 690/91 into FGT v does cause them to move with attached records, but it will also take them out of the subject index. We will bring this up in the reindexing project. We will see if we can have both.
Ask Marcive about 34x fields without 0	Marcive is willing to add the 34x fields without the 0 if we want. The 34x fields might be useful to simplify Pika's format facet logic. This is an option to keep in mind, but R&D is not in a position to implement any such overhaul.
Meeting with Telluride and Grand to set up Marquis dup checking	This meeting took place and Grand is working on making use of the dup checking in Marquis.
902 fields deleted	Lloyd finished deleting all the 902 fields.

Announcements

- Marmot staffing changes
 - Mark and Jordan have left Marmot. Replacement hiring process is on hold until the new executive director starts.
- LoC sample non-ISBD records
 - We have had time to take a look at these sample records, and ***Action Item:** Lloyd will delete them.
- Deleting 938s
 - Lloyd is still working on deleting the 938 fields. These were vendor data that was cluttering up our records. Most of them have been deleted already, but there are a few that appear to be something from CMC, so they may have local value. ***Action Item:** Lloyd will check with CMC about them.
 - The load profiles have been changed to not load the 938s. So you don't have to continue to delete them from OCLC. However, if anyone does need these to load, we can change them back.

Reports

- Duplicate Committee Report
 - Duplicate committee continues to work with Brandon on the new duplicate checking utility with Tableau. The most current version can be accessed on dev.marmot.org. Contact User Services for the login. Brandon tried to add a location limiter, however the current behavior is that it won't see a duplicate unless you have selected both libraries that have one of the records. That way you would have to select all libraries to see all of your duplicates, but then you would see everyone's duplicates. We are continuing to refine this.

- Reindexing Report
 - The committee did not meet this month because we were waiting to formal the arraignment with III. The agreement is now finalized, and we expect to have the III data analyst join our reindexing meeting in June and move ahead with the indexing project in June.

Discussion Topics

- IUG report – Instead of a mini-IUG this year Brandon and Lloyd will report what we learned to each of our committees.
 - B&T implements acquisitions API in August
 - Many book and media vendors were present in the vendor room. Lloyd found out that Baker & Taylor is planning to implement III's acquisitions API in August. Gobi (formerly YBP) has been a beta tester for a while, but none of our members use Gobi as a major supplier. However, several of our members use B&T, so we could try to implement the new system this year if any members want to try.
 - Jamie says that CMU has been considering starting to purchase from an academic book vendor, like Gobi, instead of Amazon, and they might be interested in trying this out with them.
 - We don't know how this new process would affect our duplicates problem. We will have to experiment with how to make it work best.
 - Will the API create a duplicate with every order? Can it attach orders to existing records? Based on what match point? Do order records need 020 fields? These are questions to figure out when setting API acquisitions. We want to avoid setting up a system that will really be more work because it creates duplicates.
 - III Context Engine
 - This is what III is calling their new product. They claim it will be a revolutionary implementation of linked data. We'll see. It sounds like they hope that it will eventually replace Encore, Sierra and Polaris. III seems to be devoting a lot of resources toward it.
 - Regex in field protection
 - In the load profile forum users talked about how they discovered that load profiles can protect fields, or not, based on the content of the target field. For example, maybe we could protect our 590 fields if and only if they had the code 'MLN' in the field.
 - It seems that this is not a feature that III intended to build into the system or was even aware of. It could be very powerful, however as an unsupported feature its continued functioning from version to version is not guaranteed.
 - OpenRefine
 - Lloyd presented at IUG about a few topics. Some of them Marmot people are already aware of. However, the part about using OpenRefine may be new to some of you. He reprises that part of the presentation quickly for this group. The whole Powerpoint will be posted on the UCC page.
 - You could also learn about OpenRefine here: <https://data-lessons.github.io/library-openrefine/>
- Remaining 938 fields
 - Discussed in announcements.
- Record numbers in bib records. 908 .o, 909 .c
 - Lloyd realized that we have record numbers for attached records in our bib records. There are order record numbers in 908 fields and a few checkin record numbers in 909 fields.
 - This is poor practice because the attached records can move to different bibs or even be deleted, but the MARC fields are static.
 - When we export records the export profile populates these same fields with the numbers from the currently attached records. If numbers live in the bib records those would be exported as well. The numbers in the export would be duplicated or worse, the exported numbers could be obsolete because those records aren't actually attached to this bib any longer.
 - For this reason Lloyd would like to delete these fields from the bib records and change the loaders to exclude these fields from loading in the future. If anyone is actually using these data please contact Lloyd.
 - Nobody at the meeting reports needing these numbers in the bib records.

- Jamie suggests that Lloyd should investigate where these numbers are coming from before we delete them. ***Action Item:** Lloyd figure out whose record numbers are in the 908s and 909s.
- Create Lists files reorg
 - We started doing a nightly export of records for CMU. This is using review file number 1 in Create Lists. This was our largest review file (besides the one for the Pika export).
 - The nightly process will over write whatever is in file 1 and rename it CMU Export – Do not use.
 - CMU does not use the file during the day, so people could actually use it, but they need to be aware that it will be cleared every night.
 - Lloyd moved files around so review file 1 is only as large as it needs to be for CMU’s export. File 3 is now our largest available list at 724,516 records.
 - Also, the 400,000 and 500,000 record lists have been eliminated because Lloyd realized that you can’t do a global update on more than 300,000 records at a time, so they were not actually very useful.
 - There is now an additional 200,000 record list instead.
 - If you do need an additional very large file, contact Lloyd and we can move files around to make it possible.
 - Please take care not to use these very large lists for small files that could fit in our much more numerous smaller lists.
 - We are anticipating that Sierra 4.0 will allow us to use APIs for record exporting. Theoretically this will mean we don’t need any review files for the Pika or CMU bib record exports. This would free up 2.8 million records that we could spread around our review files. If that pans out we can look at a complete reorganization of our review files again.
- Sierra 4.0
 - There is a problem with Prospector that prevents us from upgrading at this point. It will not display item status correctly in the Encore version if we upgrade to Sierra 4.0. It would only display the status code instead of the label. That is very unhelpful for Prospector users who would not be able to see if items are checked out or available unless they knew the meaning of our internal codes. This problem should be fixed when Prospector upgrades to the newer version of Inn-Reach later this summer.
- Can we get rid of these load profiles?
 - m2btab.msgovn G > Load CMU, COMN file [m2btab.msgovn]
 - m2btab.netlib 8 > Load shared elec. res. match on o, overlay, URL in bib [m2btab.netlib]
 - m2btab.netlib2 W > Load non-shared elec. res. match on o, overlay, URL in item [m2btab.netlib2]
 - m2btab.elec3 9 > Load mdl state docs SLOW [m2btab.elec3]
 - m2btab.acq
 - m2btab.asgov
 - m2btab.bo
 - m2btab.bprism
 - m2btab.bta
 - m2btab.bta.gov
 - m2btab.bta.lc
 - m2btab.bvorder
 - m2btab.clickbackup
 - m2btab.cmgov
 - m2btab.eleres
 - m2btab.fse
 - m2btab.fse.backup
 - m2btab.fse.c
 - m2btab.fse.i
 - m2btab.gov
 - m2btab.grand
 - m2btab.mpgov
 - m2btab.msdea
 - m2btab.ser
 - m2btab.sercmc
 - m2btab.wsc
 - m2btab.wsgd

- m2btabs.wsgdtest
 - Lloyd used the notes we've been creating in 995 fields to find these load profiles that are no longer seem to be getting used, and he would like to have III delete them unless they are actually being used.
 - Jamie says that m2btabs.msgov is one they use regularly so we should keep that one.
 - That means that Lloyd's process to check use of these may have been in error. He will double check them all before having III delete any of them.
 - Jamie points out that they can be saved as text files and easily recreated, so deleting them is not actually that big a deal.
 - We will wait until after next month's UCC meeting before deleting any. Contact Lloyd if you need one of them kept.

Ongoing action items

Action	Responsible parties	Report
Figure out how to copy Garfield birthdates to variable field	Lloyd	Exporting the patron records in a format that they can be reloaded is proving a little difficult. Brandon may be able to help by using SQL to extract the information for reload.
Lloyd will look into creating an exporter that excludes 0 for CMU.	Lloyd	Jamie asks if we remove the 0 fields that we have now in authorized fields, would Marcive add those back for free in the future if we were to make use of them in Pika? Yes, they will send us the 0 again, but Lloyd is concerned that Pika never will make use of them if we don't have them. *Action Item: Lloyd will check with Marcive to see if returning the information would be free. We might want to wait at least until the new executive director starts before making any major changes.
New export profile for 538 field for CMU last copy project	Lloyd/Jamie	No update
Update cataloging standards document, bring updated version of it back to the UCC, and get it up on the web page	Lloyd	No update

- Tammy proposes that we should coordinate UCC and R&D to determine how we can improve Pika grouping. She reports finding a lot of bad grouping. We think it is mostly due to bad records from Overdrive and Hoopla, but there could also be problems with Sierra records. We can't change the bad records from vendors, but we can look into how we might change Pika to work anyway. ***Action Item:** Set up meeting with duplicates committee and R&D (when they have time) to look at improving automatic grouping.

New Action Items

Delete sample non-ISBD records from Library of Congress	Lloyd
Figure out whose .o and .c numbers are in bib records, and see if we can find a pattern. (with OpenRefine maybe)	Lloyd
Check with Marcive about whether we could reload our 0 fields for free if we decided to delete them.	Lloyd
Grouping meeting with R&D (when they have time)	Duplicates Committee/R&D

Duplicates Sub-committee meeting: June 13, 10 am MT
 Reindexing Sub-committee meeting: June 13, 11 am MT
 Next UCC meeting: June 27, 2018, 9-11 am MT