## Midwinter Online Conference 2023

**Questions from Session 10: OpenRefine for Collections Data, by Karen Kohn (Temple University)**

**Q1: With OpenRefine, does it tend to konk out after 10,000 records? I run into that problem with my inventory reports after 10,000 records and I am seeking a better tool when I need to look at my inventory reports.**

It does not! That is one of the reasons I started using OpenRefine. I've used it with 150,000 rows. As I said in the session, it's not a good tool for entering data into a spreadsheet, as it's meant for data clean-up, but you don't need to be doing complicated clean-up for it to be worth using. If you want to look at your file, sort it, filter, move columns around, it can be faster to do these things in OpenRefine, and it won't crash.

**Q2: How do you get the course isbn listing?**

We are lucky to have a good relationship with our bookstore, and they send us a list of textbook adoptions each semester.

**Q3: Does it matter what order you merge rows back?**

I'm not sure if this question is about what order the rows are in or what order I do the steps in. It definitely matters what order the rows are in. You need to have your unique identifier (in my case it was the MMS ID) in the left column, sort by that, then click on 'Reorder rows permanently." In terms of whether it matters if you merge the rows earlier or later, it might not, but I don't think I can give a blanket answer.

**Q4: Can/would/ should you export back to MARC for reupload to an ILS?**

I don't work with MARC records personally, but I found some instructions here: https://blog.reeset.net/archives/1873.

**Q5: Is OpenRefine software you have to download to use it? Or does it live online?**

You do need to download. It runs in a browser window, but it isn't really "online." When you download, it creates a mini server on your computer, so the software and the files are stored on your computer.

**Resources/insights shared by attendees in chat:**

- Library Carpentry: OpenRefine: https://librarycarpentry.org/lc-open-refine/
- At Harvard we export from Open Refine as a tab-delimited file, and from there use MarcEdit to create .mrc files that can be imported to the library system (e.g. Alma).
- Cataloguers use OpenRefine to clean up e-resource records and upload the MARC. That's how I learned about it - from a cataloguer using to edit ebook records.
- For working with larger files in : https://openrefine.org/docs/manual/installing#increasing-memory-allocation
- Been a while since I used it, but I seem to remember that Open Refine has no native translation to MARC, but you can pull MARC data using other tools, clean it up, and put it back into MARC, again, using something like MARCEDIT
- This may answer the question about exporting to MARC. https://www.carli.illinois.edu/sites/files/i-share/documentation/TSC2016_Pastva_OpenRefine.pdf