Splitting subject headings (LTS procedure #175)

Scope: *When an LC Subject heading is changed or split into multiple headings, a large number of bibliographic records may need review to determine what the appropriate new heading is. Rather than examine each record individually in FOLIO, a spreadsheet with relevant bibliographic fields can assist with determining the correct heading and allow for batch editing of these records.*

Contacts: Mary Campany

Unit: Metadata Design and Operations

Date last updated: June 2023

Date of next review: June 2024

Creating a Spreadsheet to Review Changed/Split Subject Headings

1. Use steps 7-13 in "Batch edit workflow" to retrieve the relevant bibliographic records as a .mrc/.mrk file. (Either a .mrc or a .mrk may be used for this process.)
2. In MarcEdit, select Export > Export Tab Delimited Records from the Tools drop down menu.
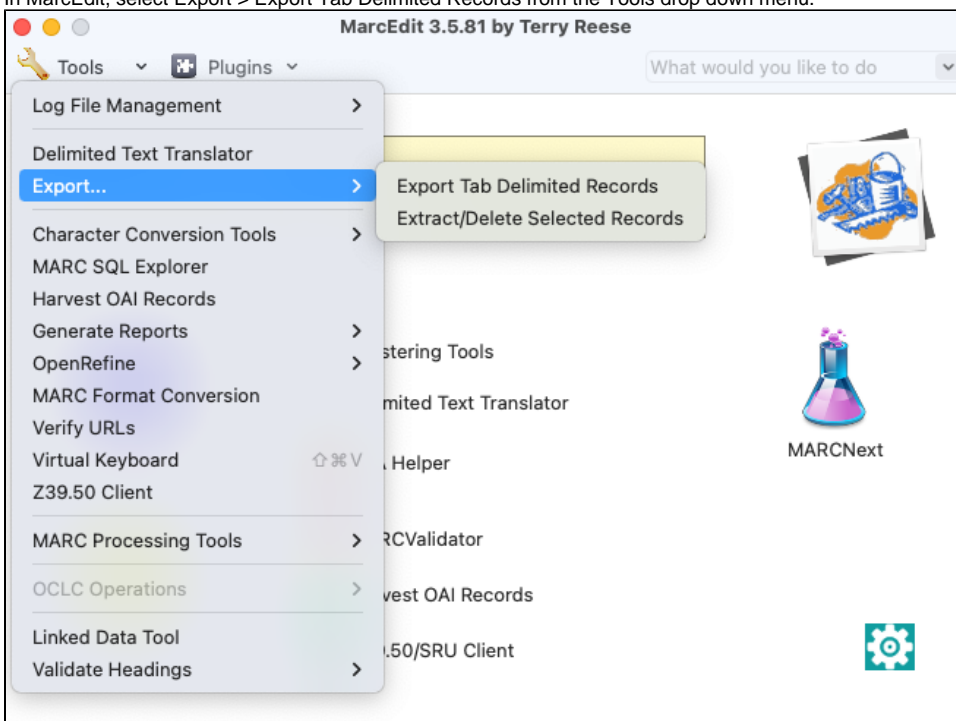
3. In the MARC Utilities window, do the following:
    a. MARC File: select the folder icon to select the source file containing the bibliographic records retrieved from FOLIO.
    b. Save File: select the folder icon to select the location of where you would like to save the csv file, and name the csv file. Make sure you use the .csv extension in your filename.
    c. Delimiter: Select Comma(,)
    d. In Field Delimiter: This will be used to separate values for fields that have multiple entries. For example, if a record contains more than one 650, each 650 will appear in the column separated by the indicated delimiter. You may use whatever you'd like, but default is a semi-colon (;).
    (If you are using Windows, select "Next" to move on to the next section)
    e. Field: Enter the MARC values for the fields you wish to add to the csv, and select the "+" button to add them.
        i. The fields that are generally useful for subject changes/splits are the 999 $i (for batch editing), 245, 246, 500, 505, 520, 650, and 651.
        ii. If you do not specify a subfield, the csv will include all subfields.
        iii. Checking the "Normalize Data" box will replace MARC subfield letters with normalized spaces/punctuation (e.g. an em dash instead of a $x in the 650, or a space instead of a $b in a 245).

4. When all fields are added, select "Process" ("Export" on Windows). You should see a message that says "Items have been exported to [file location/filename.csv]."

5. A CSV file should appear in your destination folder. You can open this file in Excel or OpenRefine.

Determining which newly-split heading each record needs

1. Open the CSV in OpenRefine.
2. Open the new LC authority records in your preferred tool (OCLC Connexion, id.loc.gov, Classweb). Read the new LC authority records carefully, paying close attention to scope notes, broader/narrower terms and alternate terms. Create a list of key words that may apply to one heading, but not the other.
   a. For example, "Primitive societies" was split into "Anthropology" and "Prehistoric peoples" in 2022. Some words that suggest that an item may not actually be about prehistoric peoples are "Indigenous" or "Indians".
   b. Look at the columns that include the note fields for additional context and ideas for more keywords.
3. Select a column and hover over "Facet". Select "Custom text facet" from the list. Use "Or value" expression to filter out phrases that apply to one of the newly-split headings.
   a. An "Or value" expression is: or(value.contains("keyword 1"), value.contains("keyword 2"))

4. Select the "OK" button to run the expression.
5. Use the Facet/Filter panel on the left to exclude records that contain the keywords.

a. Select "Include" on the "false" line to see only results that don't match your keywords. This will make it easier to identify additional keywords.

6. Examine the remaining records. If there are additional words that could be added to the expression, repeat steps 2-4 with the new keywords.
7. It will usually not be possible to fully separate concepts. In that case, change the 650 filter to include only the key terms you've filtered. This will show only those that are a clear match with one of the concepts.
8. Add a star to all of the rows by selecting the dropdown menu at the top of the "all" column, hovering over "edit rows", and selecting "star rows".
9. Select the "export" dropdown menu in the top right section of the screen. Select "Comma-separated value" to export as a CSV file.
10. Delete everything except the values in the 999 $i. Use these to export records and batch edit according to the "Batch edit workflow" document.
11. Filter out all rows that were marked with a star in step 8. These items were resolved as part of step 10.
12. Repeat steps 2-10 with keywords related to the second split term.
13. Correct any remaining results (still viewable in Blacklight) manually in Folio.
    a. Consider looking at the record in OCLC in case it has been updated to include a notes field or the subject itself has been updated.
    b. Sometimes a trip into the stacks to look at the item will be the best way to confirm the appropriate subject for it.