



FROM
KNOWLEDGE
TO
ACTION

111TH AALL ANNUAL MEETING
& CONFERENCE

JULY 14-17, 2018
BALTIMORE, MD

W1: The Magic of MarcEdit

July 14, 2018

PRESENTED BY

Yael Mandelstam

Head of Cataloging & Metadata Services
Fordham University School of Law
ymandelstam@law.fordham.edu



Program Outline

- **Introduction to MarcEdit**
 - General Information
 - MarcEdit Main Window
 - Setting MarcEdit Preferences
 - MarcEdit Formats
 - Opening a File of Records in MARC and Other Formats
 - Working With Multiple MARC Files
 - Other Useful MarcEdit Tools
 - The MarcEditor

- **Exploring MarcEdit Functions Via Five Workflows**
 - **Workflow 1:** Analyze File of MARC Records
 - **Workflow 2:** Batch-edit MARC Records
 - **Workflow 3:** Create Task List
 - **Workflow 4:** Enhance Records with Library of Congress Subject Headings From OCLC
 - **Workflow 5:** Convert KBART to MARC

Introduction to MarcEdit

General Information

- **What is MarcEdit?**

MarcEdit is a free universal library metadata tool, developed and owned by Terry Reese

- **System requirements for MarcEdit 6.x**

Any version of Windows that supports Windows .NET Framework 4.0+

- **System requirements for MarcEdit 7.x**

Windows 7+ (XP no longer supported)

Windows .NET Framework 4.6+

- **Installing MarcEdit**

Downloads available at <http://marcedit.reese.net/downloads>

To find out if you need the 32-bit or 64-bit download, click the **Start** button on your computer and click **System** (if it's not listed, it should appear when you type 'system' in the search box). In the box that opens with information about the system, you should find your **system type**.

- **Help/Support**

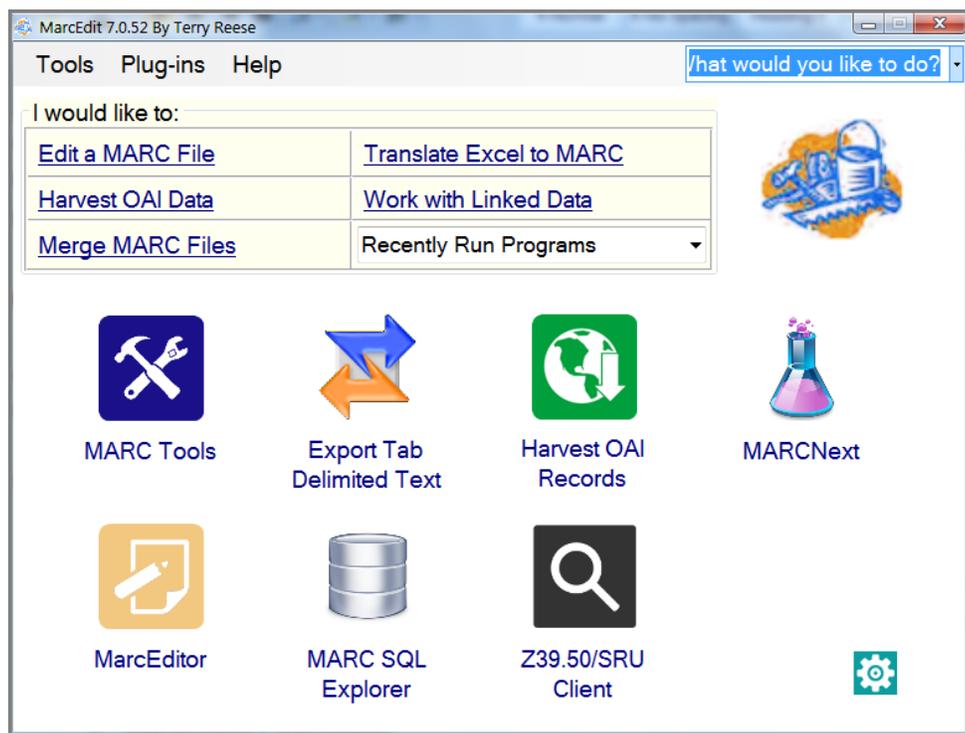
MarcEdit Listserv (subscribe and search list archives): <https://listserv.gmu.edu/cgi-bin/wa?A0=marcedit-l>

Post to MarcEdit Listserv: MARCEDIT-L@listserv.gmu.edu

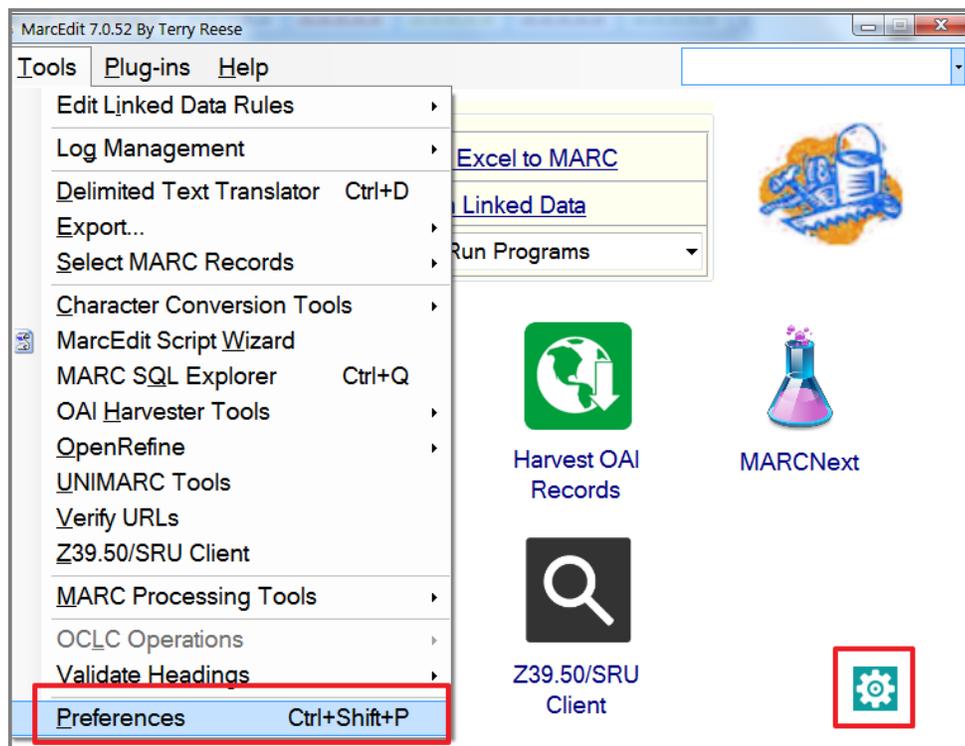
MarcEdit Tutorials on YouTube: <https://www.youtube.com/user/tpreese/playlists>

Terry Reese reese@gmail.com

MarcEdit Main Window



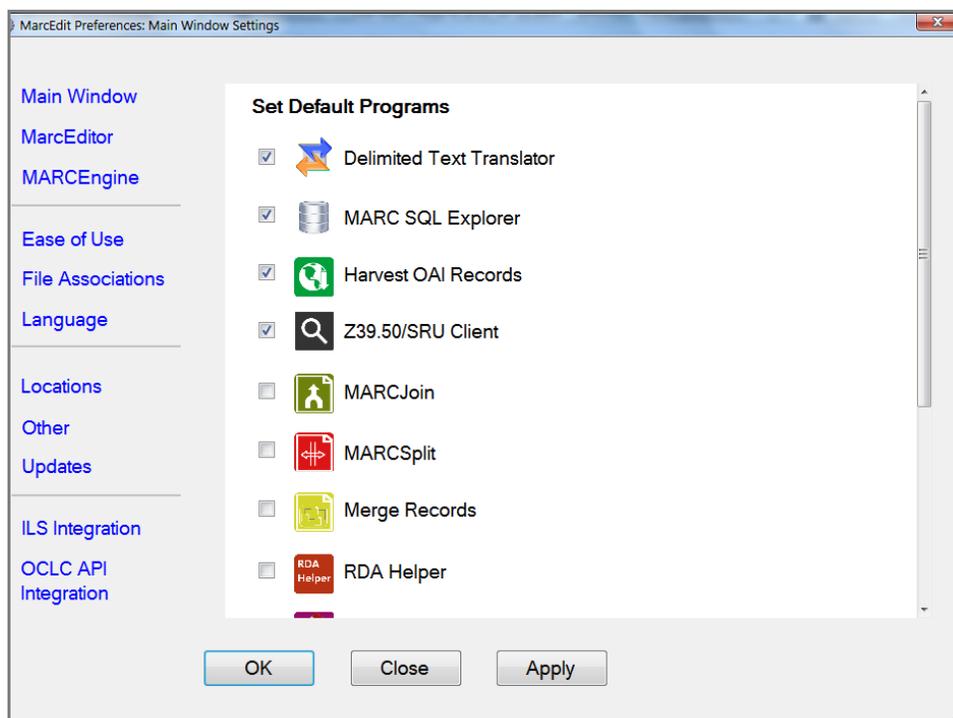
Setting MarcEdit Preferences



Tools > Preferences

OR

Click the Gear icon



Highlights

- **Main Window:** select which icons to display (limit of four) via Main Window
- **MarcEditor:** select font of records in the Editor, number of records to display, and more
- **File Associations:** add file extensions to existing .mrc, .mrk, and .mrk8
- **Language:** select language and font for the user interface (menus, window bars, etc.)
- **Updates:** enable/disable automatic checking for updates

About MarcEdit File Formats

MarcEdit saves MARC21 files in the following formats:

- ***.mrc** for records in raw MARC21 format (binary)
- ***.mrk** for records in human-friendly mnemonic format

To convert records from .mrc to .mrk use **MarcBreaker**

To convert records from .mrk to .mrc use **MarcMaker**

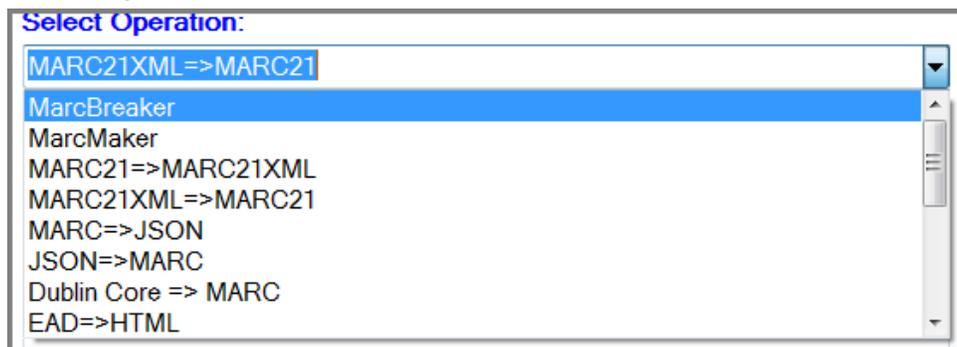
MarcEdit can also read other file types such as Innovative Interfaces **.out** files, OCLC **.dat** files, vendor **.001**, **.marc**, and **.bin** files. To use these files in MarcEdit, change individual file extension via **right-click > Rename** or **F2**. If you regularly need to handle MARC files with specific extensions, you may opt to add these extensions via **Tools > Preferences > File Associations**.

Opening a File of MARC Records

There are various ways to open a file of .mrc records:

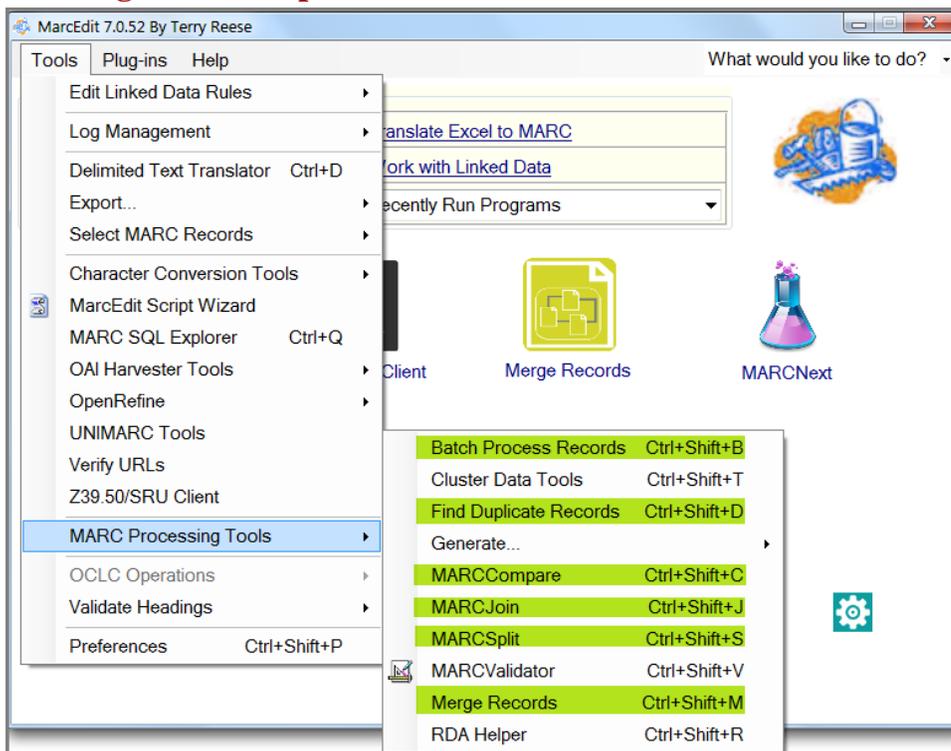
1. Directly from file: Double-click on **filename > Execute > Edit Records**
2. Via MARC Tools
 - Open MarcEdit Main Window
 - Click on **MARC Tools** icon
 - Select Operation: **MarcBreaker**
 - Select Data to Process: Open...: select your binary (.mrc) file
 - Save As...: select location and name mnemonic (.mrk) file
 - Click **Execute > Edit Records**
3. Via MarcEditor
 - Open MarcEdit Main Window
 - Click on **MarcEditor** icon
 - Select **File > Open**
 - At the bottom of the dialog box select **MARC Files (*.mrc)**
 - Locate your file and click **Open**

And Keep in mind ...



MARC Tools can convert other data to and from MARC: JSON, Dublin Core, EAD, MODS, and more

Working with Multiple MARC Files



- **Batch Process Records:** Process multiple MARC files in a folder from MARC to mnemonic .mrk
- **Find duplicate Records:** Identify duplicate records in files. Options include removing duplicates or outputting unique items
- **MARCcompare:** Find differences between two versions of a MARC file
- **MARCJoin:** Combine multiple .mrc or .mrk files into a single file
- **MARCSplit:** Divide large file into smaller files
- **Merge Records:** Merge data from one file into another file

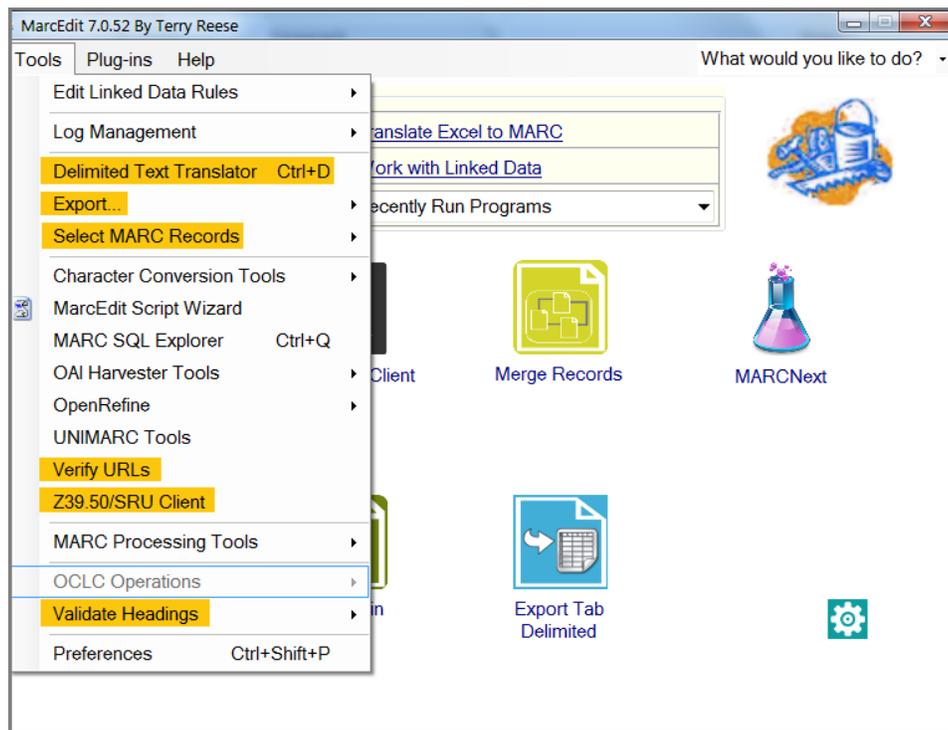
Exercise 1: Preferences

- 1a. Select preferred icons to display on the MarcEdit Main Window
- 1b. Change font type and size for the user interface
- 1c. Change font type and size for records in the MarcEditor
- 1d. Add file extension .dat

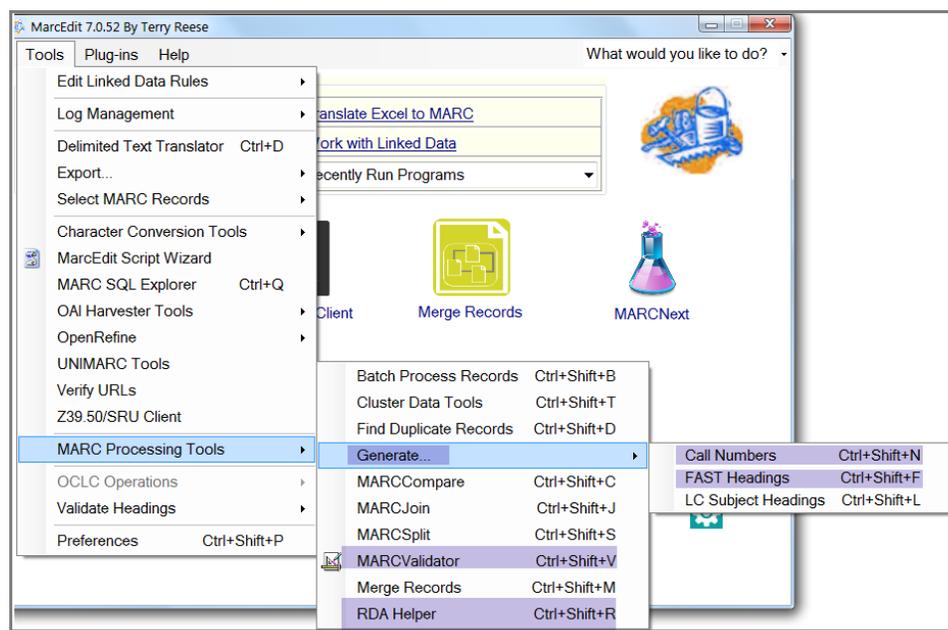
Exercise 2: Working with MARC files

- 2a. Change extension of **exercise2a.out** to **exercise2a.mrc**
- 2b. Combine files **exercise2a.mrc** & **exercise2b.mrc** to create **exercise2c.mrc**
- 2c. Convert **exercise2c.mrc** to .mrk via MarcBreaker and open it in the MARCEditor
- 2d. Compile .mrk file back into .mrc

Other Useful MARC Tools



- **Delimited Text Translator:** Map data in delimited format (e.g., Excel, access) to MARC
- **Export > Tab Delimited Records:** Export select MARC data to a spreadsheet
- **Select MARC Records:** Delete or extract selected records from a MARC file
- **Verify URLs:** Generate report of URL errors
- **Z39.50/SRU Client:** search and retrieve records from remote databases via Z39.50
- **Validate Headings:** Validate 1xx, 6xx, and 7xx headings against LCSH and NAF authorities



- **Generate Call Numbers/FAST Headings:** Interface with OCLC Classify Web Service and FAST Web Service
Note: "Generate LC Subject Headings" is a placeholder for future development and is currently not functional
- **MARCValidator:** Validate records against MARC Rules File
- **RDA Helper:** Translate AACR2 records to RDA

The MarcEditor

The MarcEditor supports many global editing functions, including Find & Replace, Add/Delete MARC fields, Edit Subfield Data, Add/Remove Field Data, Edit Indicator Data, Swap Field Data, Record Deduplication, Record Sorting, and the creation of Tasks. Other functions available via the main MarcEdit Window are also available via the MarcEditor, such as Generate Call Numbers, Record Deduplication, RDA Helper, and more.

IMPORTANT: When editing records in the MarcEditor, resave file with different filenames at various stages of the editing process. This way, if something goes wrong, you'll be able to go back to a previous version of the edited file rather than start from scratch.

And always, ALWAYS keep an unedited copy of your original file!!!

Oops!

Need to undo your previous edit?

Option 1: Use the Windows-wide **Ctrl+Z**. Use **CTRL+Y** to reverse your last Undo

Option 2: Select **Edit > Undo (F2)**

To undo your previous **GLOBAL** edit select **Edit > Special Undo (Ctrl+Alt+F2)**

Workflow 1: Analyze File of MARC Records

Description: We receive record sets from various vendors, some requiring revisions before they are loaded into our ILS. The first step is to familiarize ourselves with each record set and the particular issues it presents.

For this workflow, we'll use file **workflow1.mrc** to perform the following:

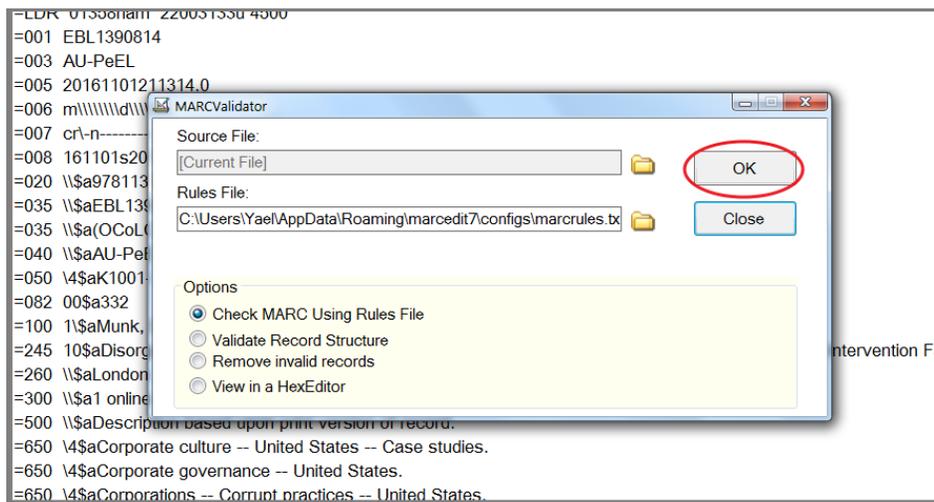
1. Validate records
2. Examine a few individual fields
3. Count fields
4. View fields in isolation
5. Find missing fields
6. Find multiple occurrences of the same field

MarcEdit functions covered

- Tools > Validate MARC Records
- Reports > Field Count
- Edit > Find (Ctrl+F)
- File > Select Records for Edit

Step-by-step Guidelines

1. Validate Records



Open file **workflow1.mrc**

Tools > Validate MARC Records

Select **Check MARC Using Rules File > OK**

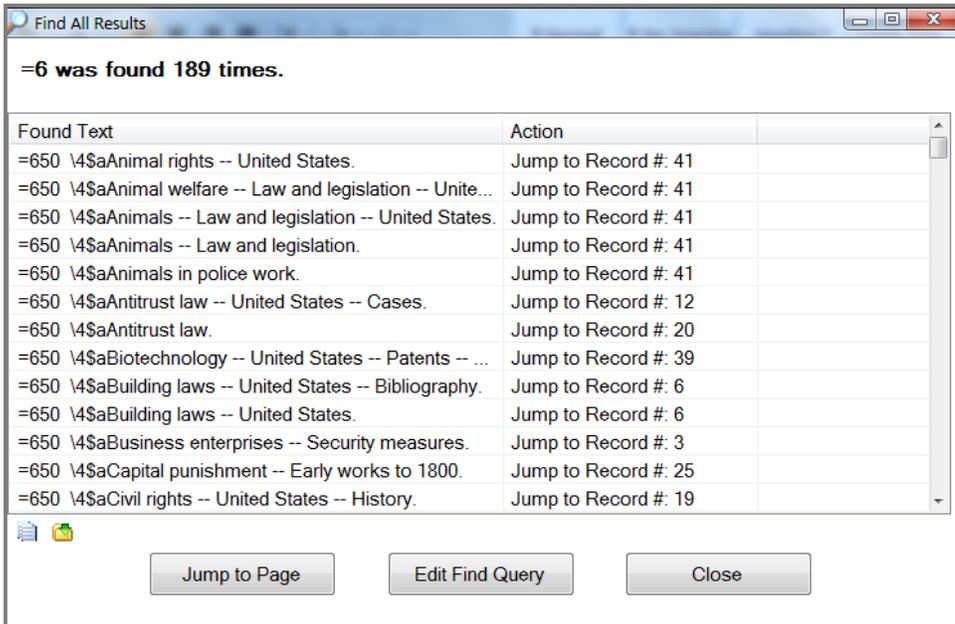
Check errors. If report is long, print report using the Print icon or copy & paste using the Copy icon and pasting in Notepad, Word, or similar.

To correct an error in a record specified in report, select **Edit > Jump to > Jump to...Record**, enter record number, and click **OK**

2. Examine a few individual records

3. Count fields in file: **Reports > Field Count**

4. View fields in isolation

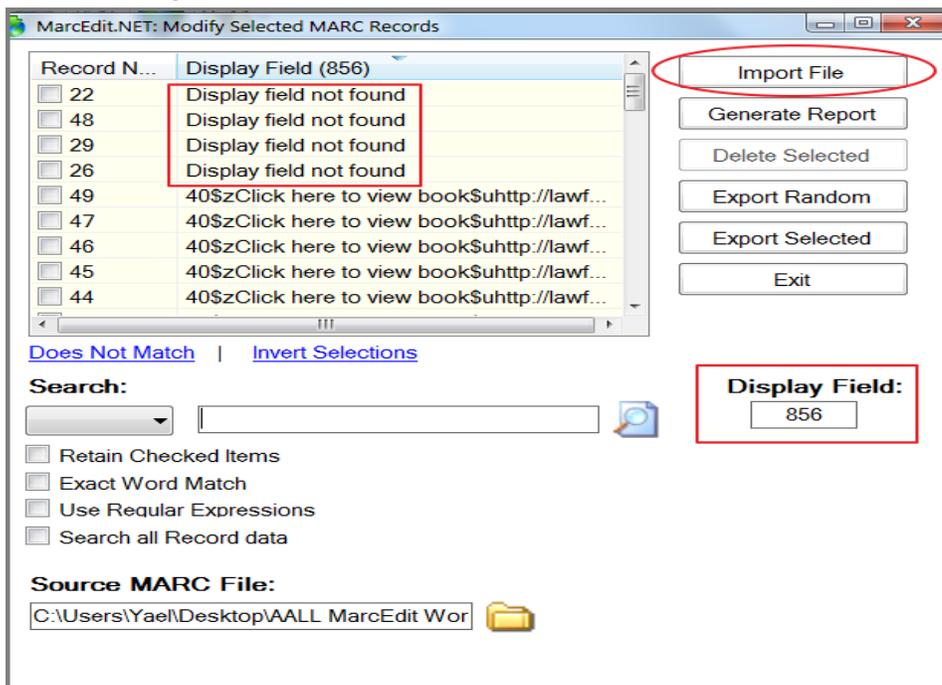


Edit > Find = [field] and click **Find All** (Alt+A)

To display a range of fields such as 6xx or 7xx, enter first digit only (e.g., =6 or =7)

Click **Found Text** to sort 6xx fields

5. Find missing fields



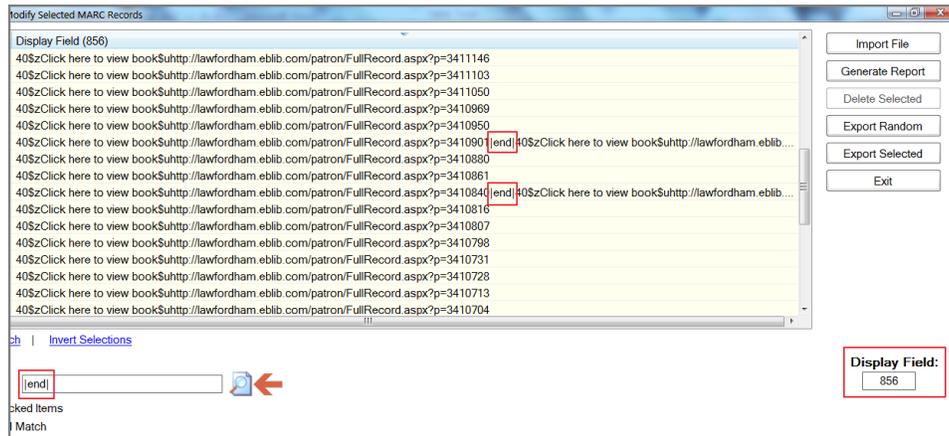
File > Select Records to Edit

Display Field [enter field, e.g., 856]

Click **Import File**

If field not found in record, it will show as **Display field not found** (Click on Display Field to sort results)

6. Find multiple fields



If two or more occurrences of the same field exist in a record, they will be separated by **|end|**. To identify these fields, enter **|end|** in search box and click search icon. The records that fit your criteria will now be selected.

To further edit the selected subset of records, click **Export Selected** and save new file

Exercise 3: Validate MARC Records

- 3a. Open file **exercise3.mrk** in the MarcEditor
- 3b. Find Problems with MARC fields using **Tools > Validate MARC Records > Check MARC Using Rules File**. Can you fix them?

TIP: If you are unable to figure out the problem with record #2, it may be a structural issue that is better described in **Tools > Validate MARC Records > Validate Record Structure**

Exercise 4: Examine Call Numbers

- 4a. Open file **exercise4.mrk** in the MarcEditor
- 4b. View field 050 in isolation: can you detect problems with the call numbers?
- 4c. Are any records missing call numbers?
- 4d. Do any records have more than one call number field?

Workflow 2: Batch-Edit MARC Records

Description: After analyzing the records and identifying the issues with the record set, we can start editing the records using MarcEdit batch editing tools.

For this workflow, we'll use file **workflow2.mrc** to perform the following:

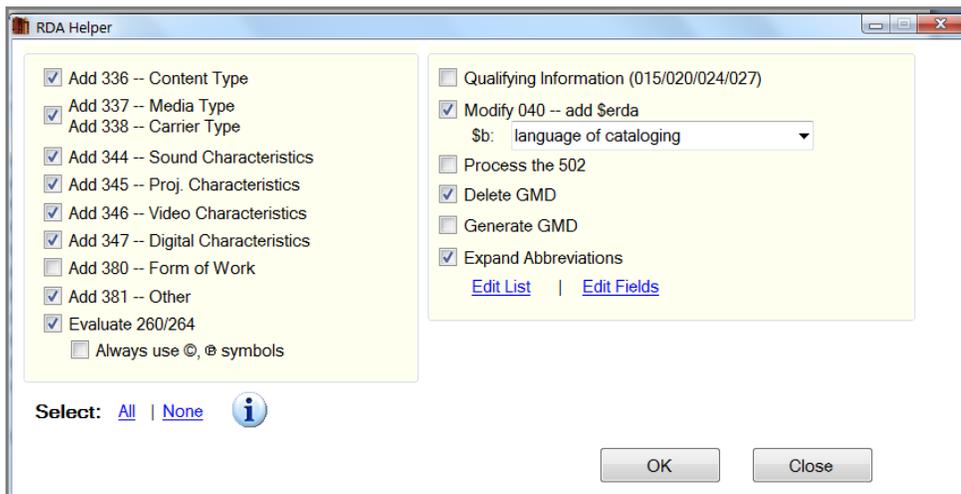
1. Convert non-RDA records to RDA
2. Add suffix to number in 001 field
3. Customize 856\$z
4. Remove undesirable 856 fields (e.g., resources for which we don't have access and URLs for free resources that may not link to full text) two options: either specify domains to remove or specify domains to keep
5. Remove obsolete fields 506, 533, 538, 583 UNLESS they have information for digital preservation projects

MarcEdit functions covered

- Tools > RDA Helper
- Tools > Edit Subfield Data (F9)
- Tools > Edit Field Data (Ctrl+Shift+F3)
- Tools > Add/Delete Field (F7)

Step-by-step Guidelines

1. Convert non-RDA records to RDA



Open file **workflow2.mrc**

Tools > RDA Helper

Select functions from dialog box and click **OK**

2. Add suffix to number in 001 field

Tools > Edit Subfield Data

Field: 001

Find: ^e (add to end of field)

Replace: [text]

Click **Replace Text**

(Prefix: to add to beginning of field, use ^b)

3. Customize 856 field

Tools > Edit Subfield Data

Field: 856 Subfield: 3

Field Data: [unique element in current text] (e.g., llmc or heinonline)

Uncheck **Match case** and check **Delete Subfield**

Click **Remove Text**

Tools > Edit Field Data

Field: 856

Find: \$u[URL]

Replace: \$z[custom text]\$u[URL]

Click **Process**

4. Remove undesirable links in field 856

Option 1: remove URLs with specified domains

Option 2: remove URLs that do not match specified domains using regular expression

About Regular Expressions (regex)

Regex is Find & Replace on steroids. It uses a text string for describing a search PATTERN. Regex is used in MarcEdit and other systems when regular Find & Replace using a literal text string cannot accomplish what we want to do (e.g., add a / (slash) before 245\$c ONLY IF it is missing:

Find: (=245.*[^\s])(\s\$c) Replace: \$1 /\$2

Characters in regex can either be treated as regular characters and are thus called “literals” (e.g., a search for **bio** will find **biography** and **antibiotics**), or have special meaning (e.g., searching for **cat | dog** will find **cat OR dog**, since the pipe character means OR; and the search for **d[!o]g** will find **dig, dug**, but **NOT dog**, since the caret within square brackets means NOT). Characters with special meaning in regex are called “metacharacters” or “special characters”.

If we want to use a special character as a literal, we can do so by using a \ (backslash). In regex this is called “escaping” a special character. For example, since \$ (dollar sign) is a special character in regex, if we want to search for \$a in a MARC record, we need to express it as \a.

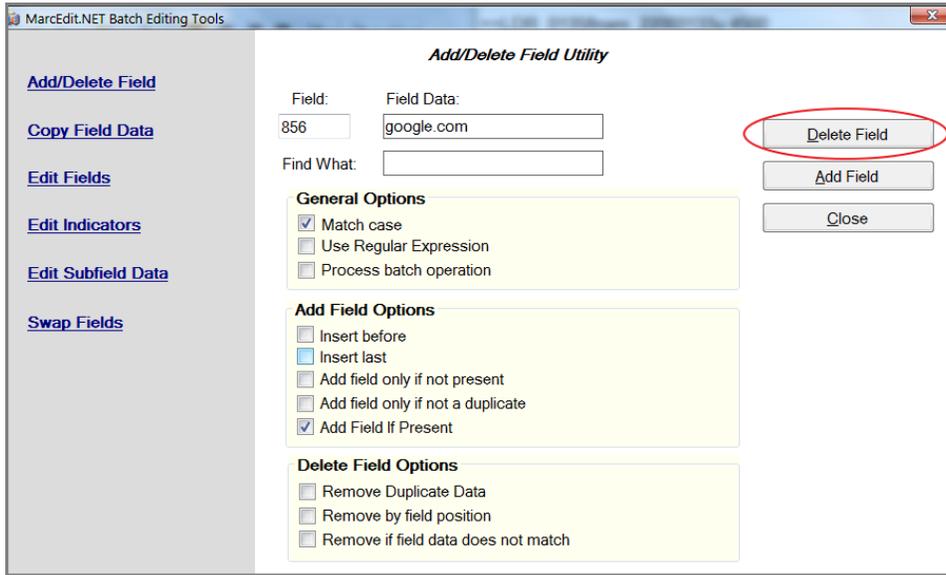
For a list of regex characters commonly used in MarcEdit, see the accompanied **chart at the end of this handout**.

For a complete regex reference see [https://msdn.microsoft.com/en-us/library/az24scfc\(v=vs.110\).aspx](https://msdn.microsoft.com/en-us/library/az24scfc(v=vs.110).aspx).

To test your expression, try <https://regex101.com/>

If you need help constructing your expression, consider posting your query on the MarcEdit Listserv. It has several regex experts who will not only help you construct your expression, but will often add a detailed explanation that can be very educational.

Option 1: remove URLs with specified domain

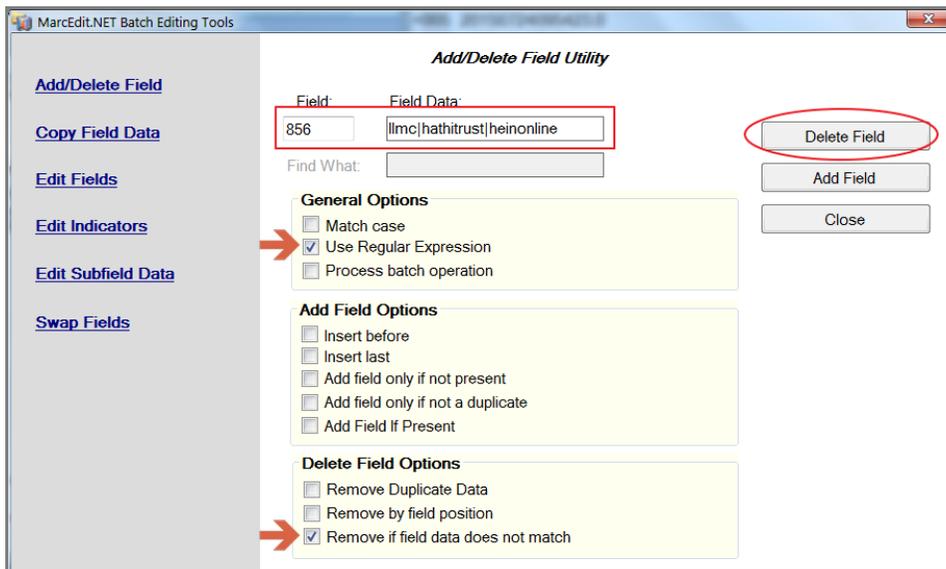


Tools > Add/Delete Field (F7)

Field Data: [domain]

Click **Delete Fields**

Option 2: remove URLs that do not match specified domains
(Using regex syntax for **OR**)



Tools > Add/Delete Field (F7)

Field Data: [domain1] |[domain2] |[domain3] (The pipe/vertical bar is a regex metacharacter that means **OR**)

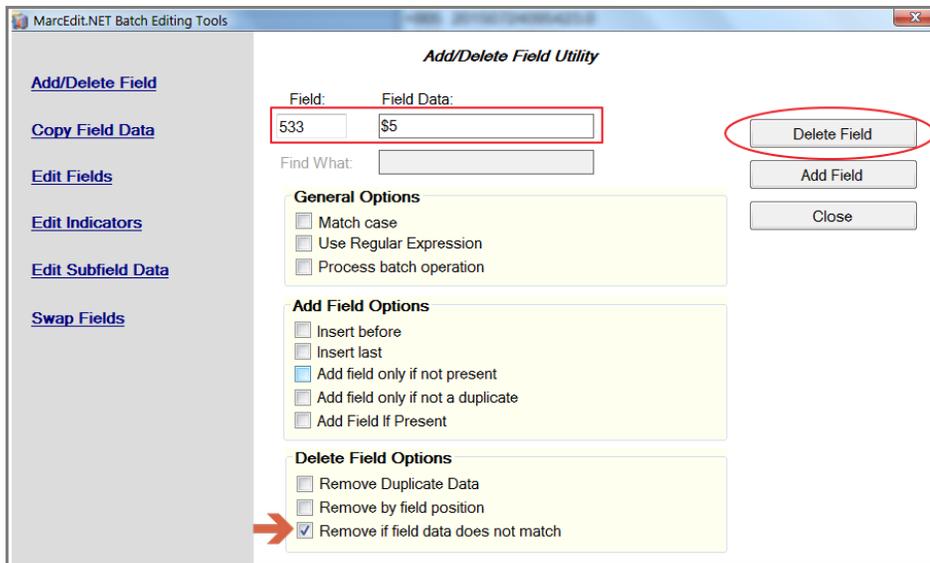
Check **Use Regular Expression**

Check **Remove if field data does not match**

Click **Delete Fields**

5. Remove obsolete fields 506, 533, 538, 583

(UNLESS they have information for digital preservation projects. The latter can be identified by the presence of \$5, which contains an institutional code)



Tools > Add/Delete Field

Field: 533, etc.

Field Data: \$5

Check **Remove if field data does not match**

Click **Delete Field**

[Exercise 5: Edit MARC Records](#)

- 5a. Open file **exercise5.mrk** in the MarcEditor
- 5b. Convert records to RDA
- 5c. Change current text in 856 \$z or \$3 to your library's customized text (e.g., \$zOnline Resource -- access restricted to [institution's name])
- 5d. Add suffix (**e-book**) to call numbers in field 050

Workflow 3: Create Task List

Description: Once we become familiar with a given record set and the recurring issues it presents, the edits for that set can be compiled into a single MarcEdit TASK for future batch edits.

For this workflow, we'll use file **workflow3.mrc** to perform the following:

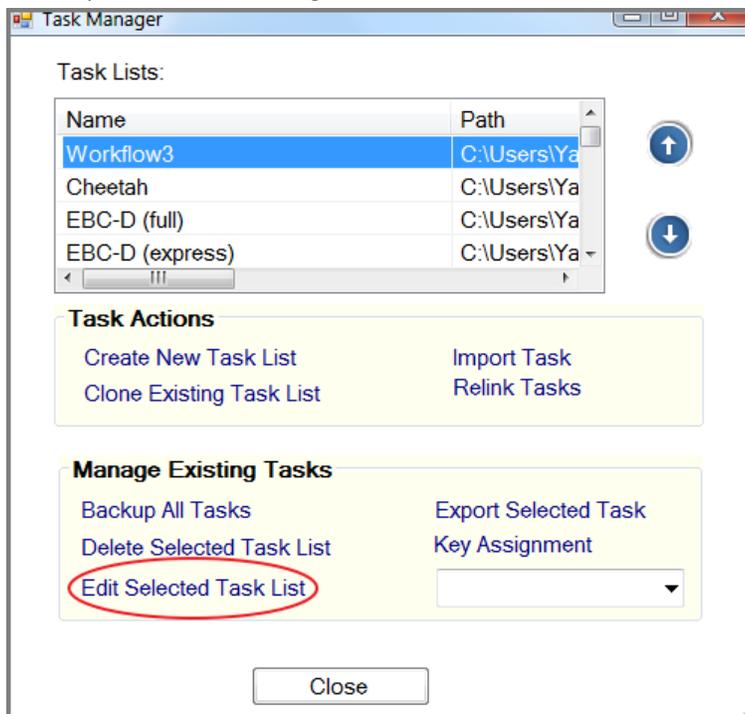
1. Create Task List
2. Edit Task
3. Run Task
4. Export Individual Task List
5. Export all Task Lists and other configurations
6. Import configuration/individual tasks

MarcEdit functions covered

- Tools > Manage Tasks
- Task Manager: Create New Task List
- Task Manager: Edit Selected Task List
- Tools > Assigned Tasks > [select task]
- Task Manager: Export/Import Selected Task
- MarcEdit Main Window: Tools > Share Configuration Settings > Export/Import Settings

Step-by-step Guidelines

1. Compile edits into a single task list



Open file **workflow3.mrc**

Tools > Manage Tasks

Task Actions > Create New Task List

Enter Task Name

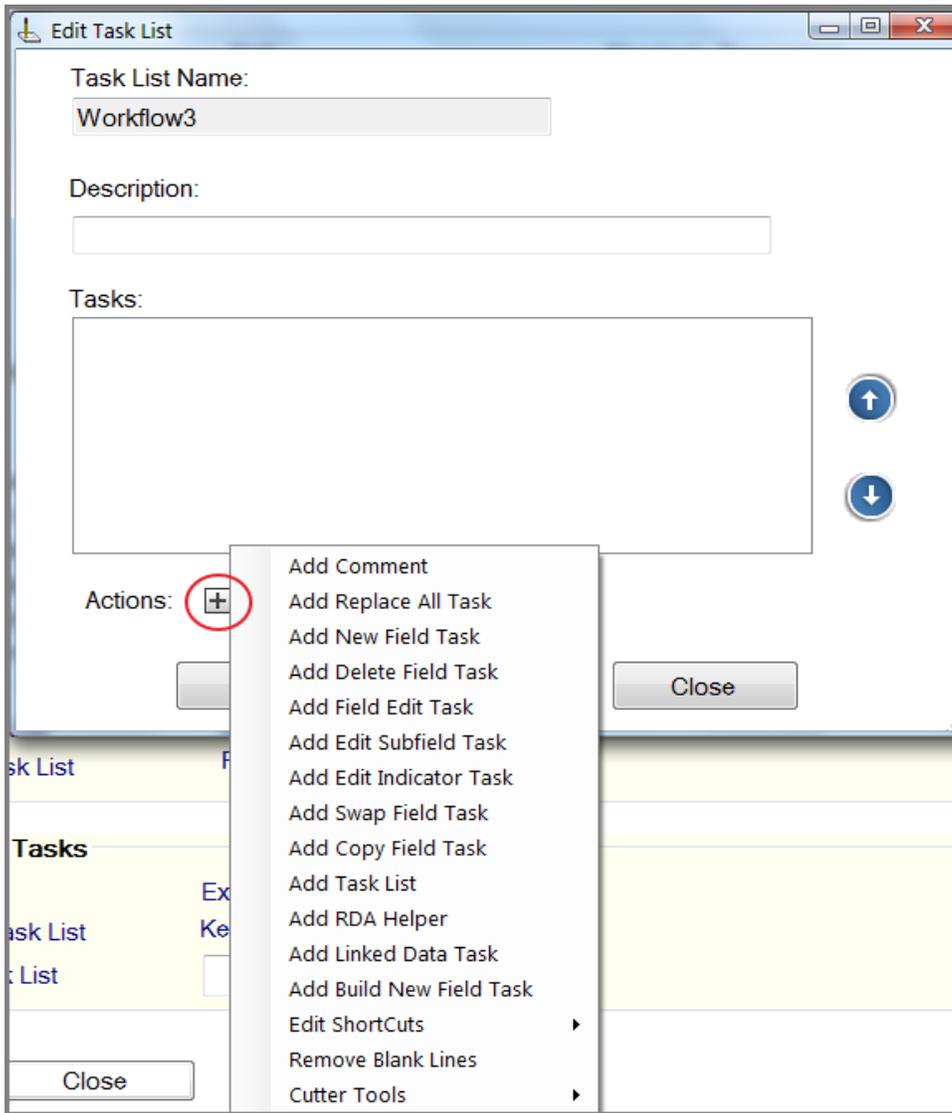
Optional: create shortcut by using **Key Assignment:**
 Highlight task
 Select desired key (e.g., Shift+F1)
 Click **Key Assignment**

Compile tasks:

Highlight task

Manage Existing Tasks > Edit Selected Task List

2. Edit Task List

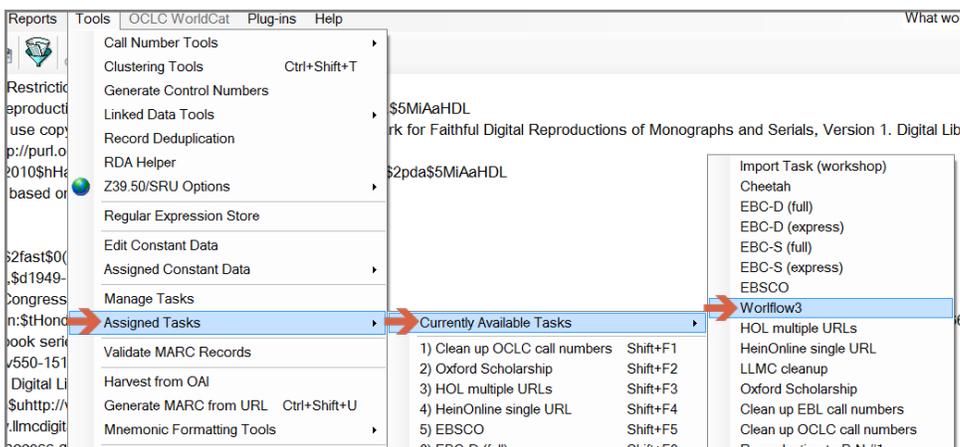


Actions: +

Select desired task (we'll use some of the edits described in Workflow 2)

When done compiling your tasks, click **Save**

3. Run Task

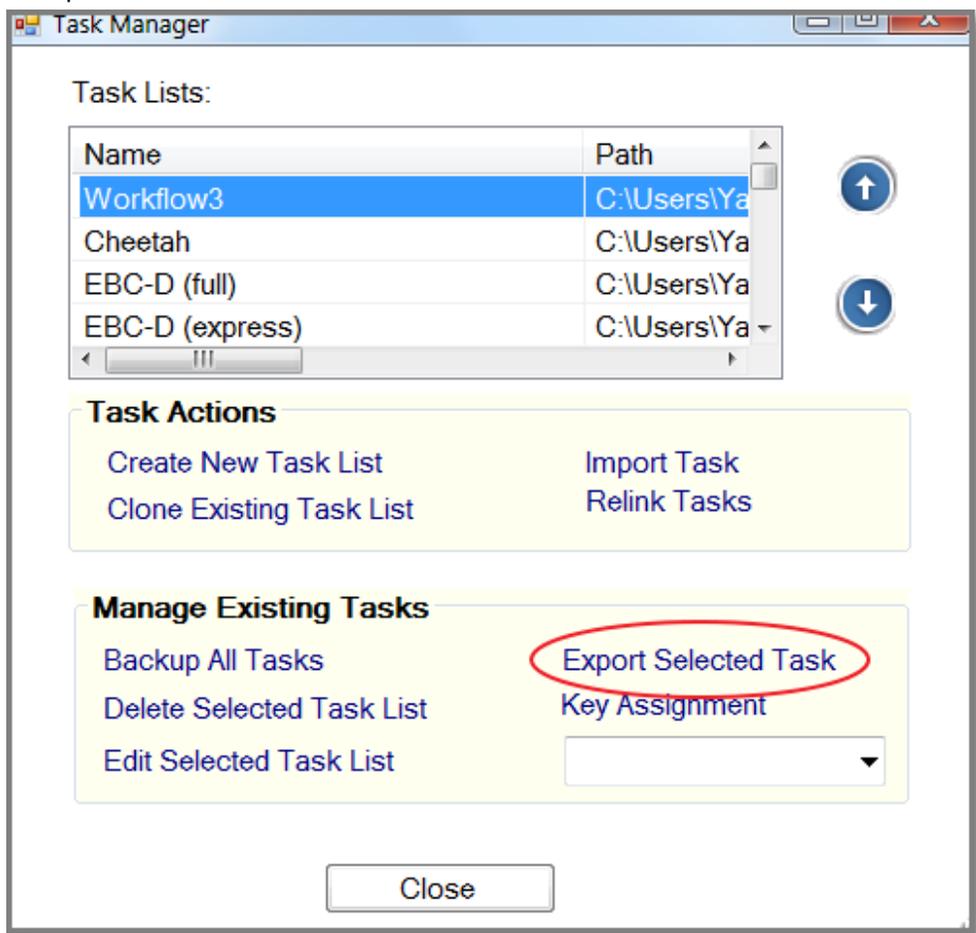


Tools > Assigned Tasks > Currently Available Tasks > [select your task]

OR

If you assigned a shortcut to your task, press the shortcut keys (e.g., Shift+Ctrl+F8)

4. Export individual task list



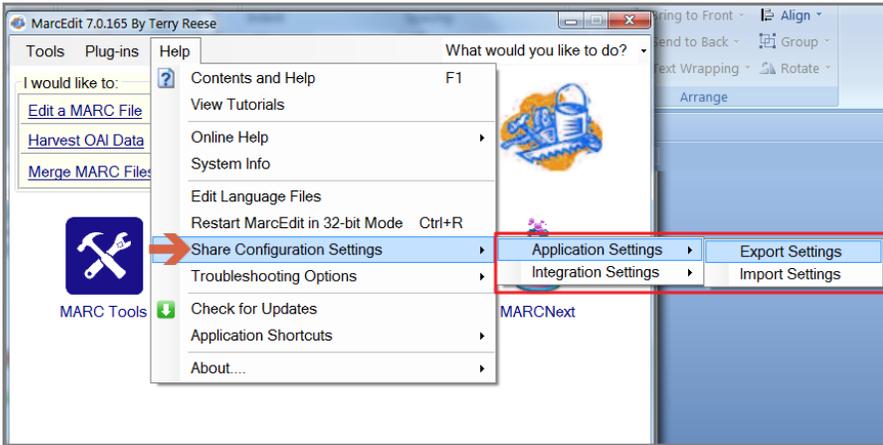
Tools > Manage Tasks

Highlight task

Click **Export Selected Task**

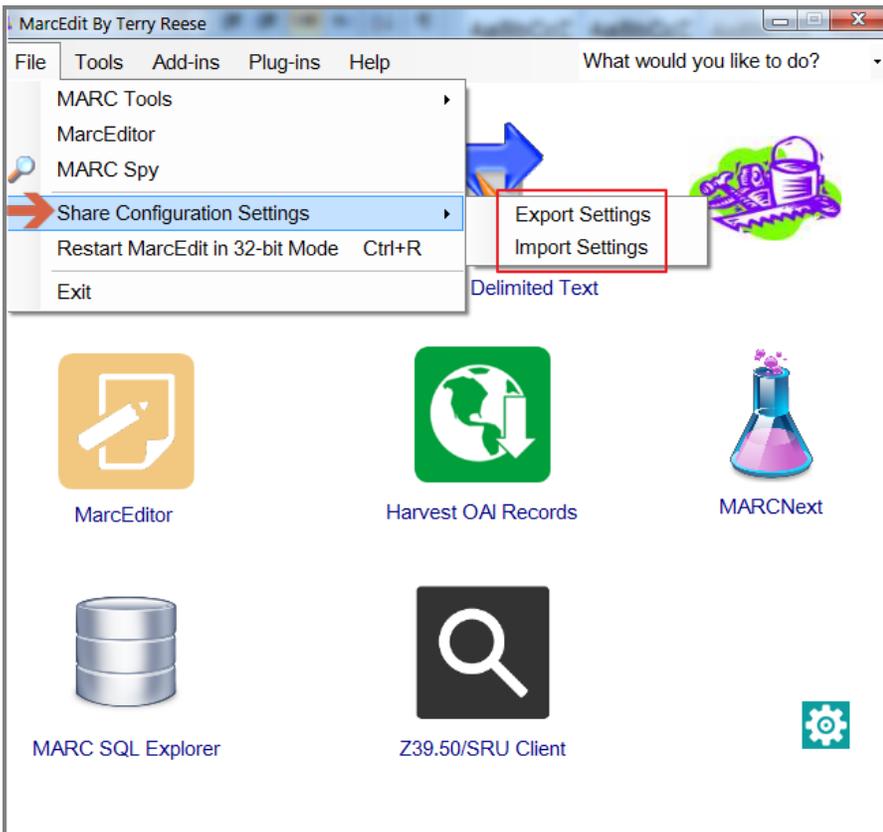
Select file location, name file, and click **Save**

5. Export multiple task lists and other configurations



In MarcEdit 7

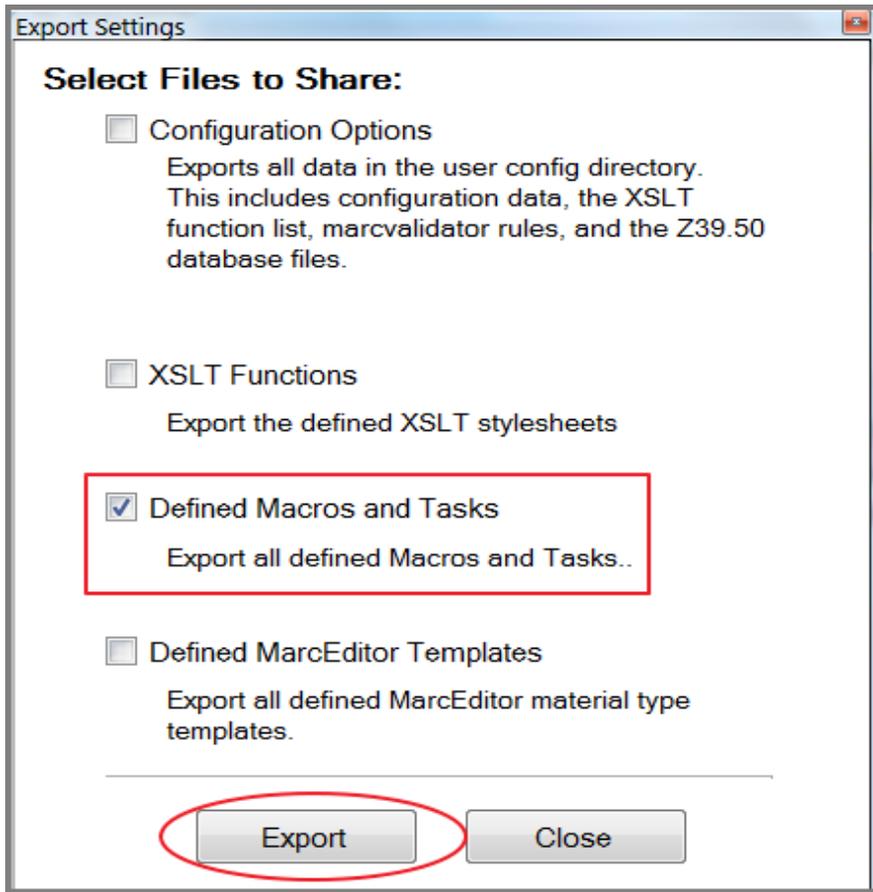
Help > Share Configuration Settings > Application/Integration Settings > Export/Import Settings



In MarcEdit 6

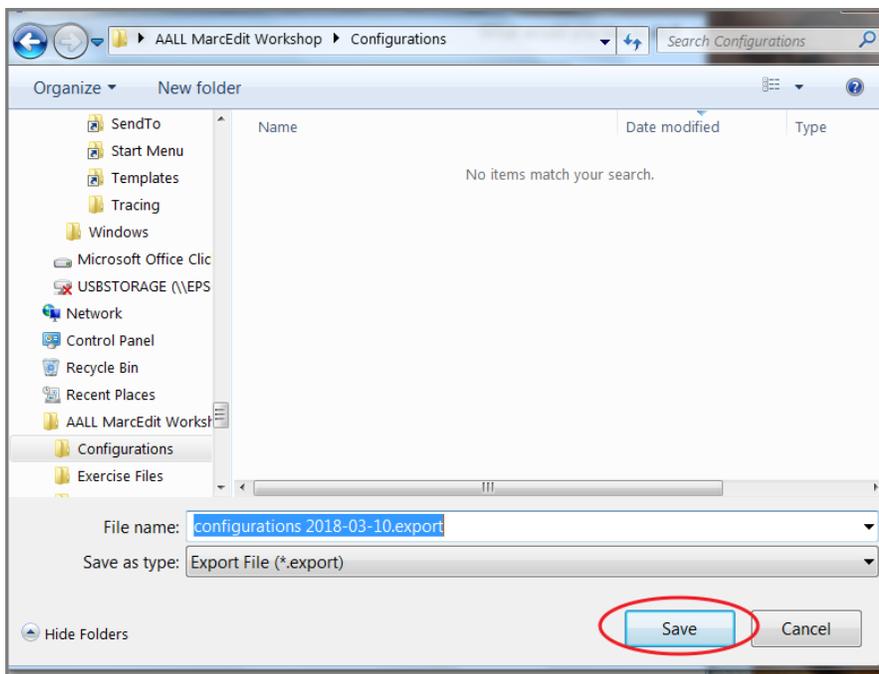
File > Share Configuration Setting > Export/Import Settings

Dialog box in both MarcEdit versions



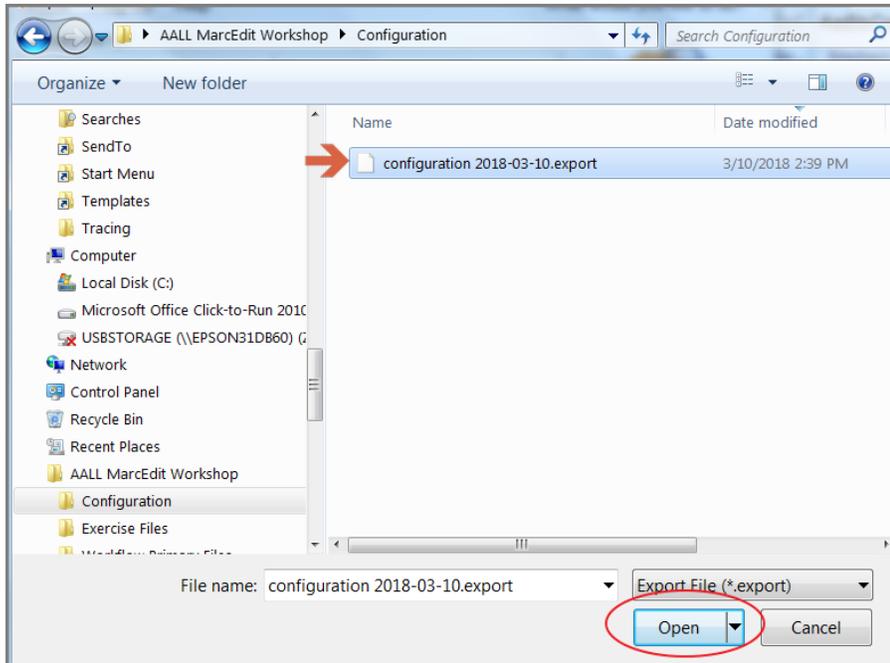
Check **Defined Macros and Tasks**, **Export all defined Macros and Tasks**

Click **Export**, select your file path, and click **Save**



Select your file location, name your file, and click **Save**

6. Import configurations



In MarcEdit Main Window
Select **Share Configuration Settings** (see screenshots above for location of this function in MarcEdit 6 and MarcEdit 7)

Select **Import Settings**

Select your desired
configuration file and click
Open (see warning below)

WARNING

The imported configuration settings will **REPLACE** your current settings, so you may want to export and save your current configuration file **BEFORE** importing a new file.

If you wish to import a single task without affecting the rest of your configurations, you can do so via the MarcEditor **Tools > Manage Tasks > Import Task**

Exercise 6: Import Single Task

Import **exercise6.task** directly into the Task Manager

About Merge Records

This tool enables us to compare two different files, and when matched records are found in both files, selected fields from **File B (Merge File)** are merged into the matching records in **File A (Source File)**.

The matching is done by comparing and matching unique numerical identifiers from one of the following fields:

001
010\$a
020\$a
022\$a
035\$a

Another option, listed as **MARC21**, uses a complex algorithm comparing and weighing multiple fixed-length and variable-length fields in both files in an attempt to find matching records. This option is useful when no matching unique numerical identifiers are available in both files (e.g., two files of pre-ISBN records from two different vendors, at least one of which has no OCLC numbers in fields 001/035 or LC Control Numbers in field 010).

Group Brainstorming

- What are some scenarios in which this tool can be used?
- What would be the potential challenges in each scenario?

Workflow 4: Enhance Records with LCSH from OCLC

Description: We have a record set with many records that lack Library of Congress Subject Headings. If the records have ISBN/ISSN, we'll try and enhance them with LCSH from OCLC.

For this workflow, we'll use file **workflow4.mrc** to perform the following:

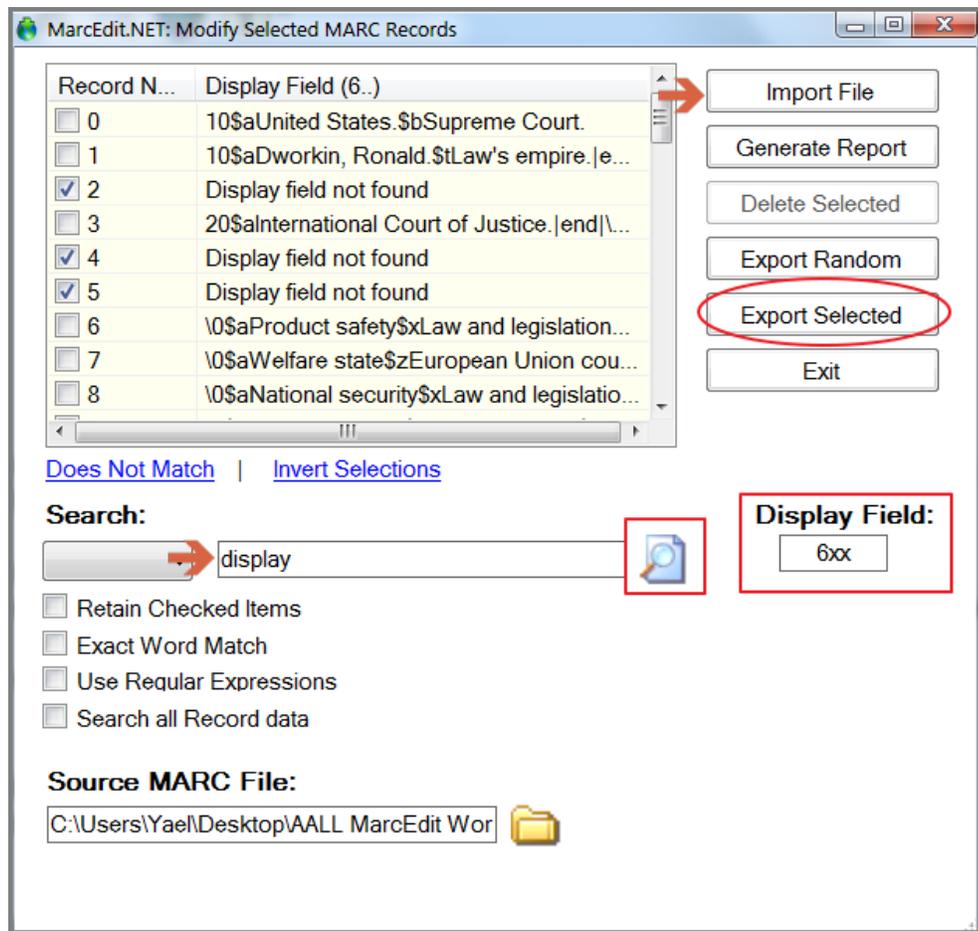
1. Extract records without LCSH and save new file as **local_source.mrk**
2. Copy ISBNs in 020\$z and 776\$z to 020 with "fake" indicators (to be deleted later).
3. Move non-numerical data in 020\$a to 020\$q
4. Export 020\$a from **local_source.mrk** to .csv file and open in Excel. If more than one ISBN in column, use Data > Text to column. Copy & Paste to have all ISBN in a single column and save as **isbn.txt**.
5. Batch-search OCLC via MarcEdit Z39.50 using **isbn.txt** file. Save OCLC records as **oclc.mrc**
6. Open **oclc.mrc** and clean up 6xx
7. Extract records with 6xx. Save as **oclc_merge.mrk**
8. Optional: Dedup records matching on ISBN
9. Merge 6xx in **oclc_merge.mrk** into **local_source.mrk**, matching on 020\$a
10. Dedup 6xx
11. Delete 020 with "fake" indicators
12. Compile file into MARC and load back into ILS

MarcEdit functions covered

- File > Select Records to Edit (extract records)
- Tools > Swap Field Data (F11)
- Tools > Edit Subfield Data (F9)
- (From MarcEdit Main Window) Tools > Export > Export Tab Delimited Records
- (From MarcEdit Main Window) Tools > Z39.50/SRU Client
- Tools > Edit Field Data
- Tools > Record Deduplication
- (From MarcEdit Main Window) Tools > MARC Processing Tools > Merge Records
- Tools > Add/Delete Field (F7)
- File > Compile File into MARC

Step-by-step Guidelines

1. Extract records without LCSH



Open file **workflow4.mrc**

File > Select Records to Edit

Display Field: [field number (e.g., 650) or field group such as 6xx

Click **Import File**

Search: 'display field not found' (or just 'display') and click on the search icon

Click **Export Selected**

We have left our previous file and are in the file of the extracted records. Save this file as **local_source.mrk**

Prepare **local_source.mrk** file for later merge:

2. Copy ISBNs in 020\$z and 776\$z to 020 with “fake” indicators (to be deleted later).

Tools > Swap Field Data

Enter fields/subfields
we want copied

Check **Copy Source** and
Sort Modified Field

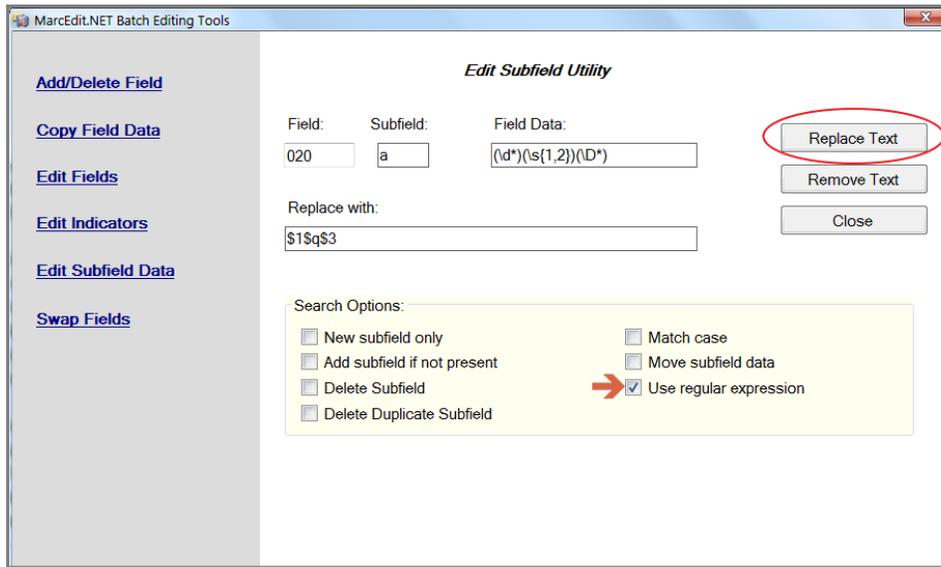
Click **Process**

Explanation

Since it doesn't really matter if our LC Subject Headings come from records for print or other formats, we want to cast the widest net of ISBNs for the OCLC search, be it from 020\$z or from other formats in 776\$z.

Later, when we merge the LC headings from OCLC into our local file, we'll use the ISBN as the match point, and this can only be done in MarcEdit with numbers in 020\$a. Therefore, we want to make sure that all our ISBNs are copied into 020\$a. When the merge is done, we'll want to delete these fields, so we need to add something that will help us distinguish them from the original ISBNs. One solution is to use “fake” indicators such as '99', which will be easy to identify later on.

3. Move non-numerical data from 020\$a to 020\$q (e.g., move **(ebook)** following the ISBN number to \$q)



Tools > Edit Subfield Data

Field: 020

Subfield: a

Find: **(\d*)(\s{1,2})(\D*)**

Replace: **\$1\$q\$3**

Check **Use Regular Expressions**

Click **Replace Text** and save file

Explanation

Parentheses are used in regex for grouping elements in FIND that can later be referenced in REPLACE. Here is an example of grouping elements in 020\$a:

020\$a: **9780191827716**[space]**(ebook)**

Group 1: **\$a9780191827716** = expressed in regex pattern **(\d*)**

Group 2: One or two **spaces** after the ISBN = expressed in regex pattern **(\s{1,2})**

Group 3: **(ebook)** = expressed in regex pattern **(\D*)**

Explanation of Group 1 regex

() – open and close parentheses to create a group

\d* – all digits

Explanation of Group 2 regex

() – open and close parentheses to create a group

(\s{1,2}) – space **\s** appearing 1 or 2 times **{1,2}** (this will catch a potential extra space)

Explanation of Group 3 regex

() – open and close parentheses to create a group

\D* – all non-digit characters

In the REPLACE box these three groups can be referenced as:

Group 1 = **\$1**

Group 2 = **\$2**

Group 3 = **\$3**

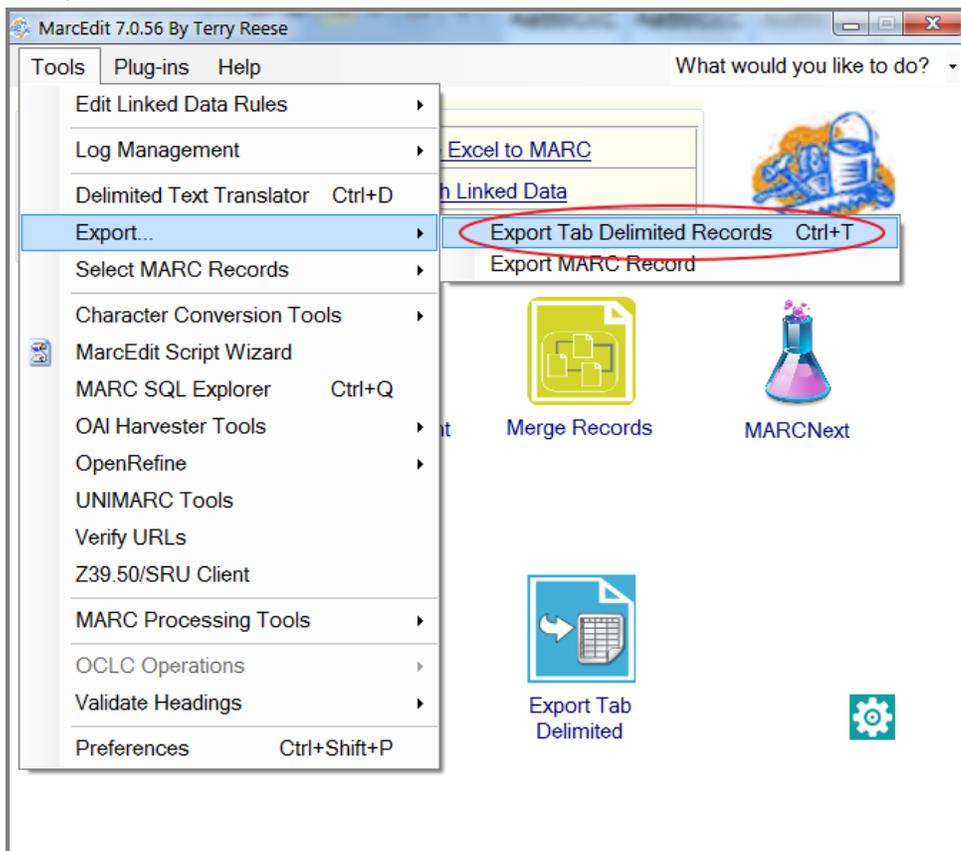
Since we want to retain the first group, eliminate the space(s) in the second group, and move the third group to \$q, our Find & Replace will ignore group 2 and will insert \$q between group 1 and group 3:

Find: **(\d*)(\s{1,2})(\D*)**

Replace: **\$1\$q\$3**

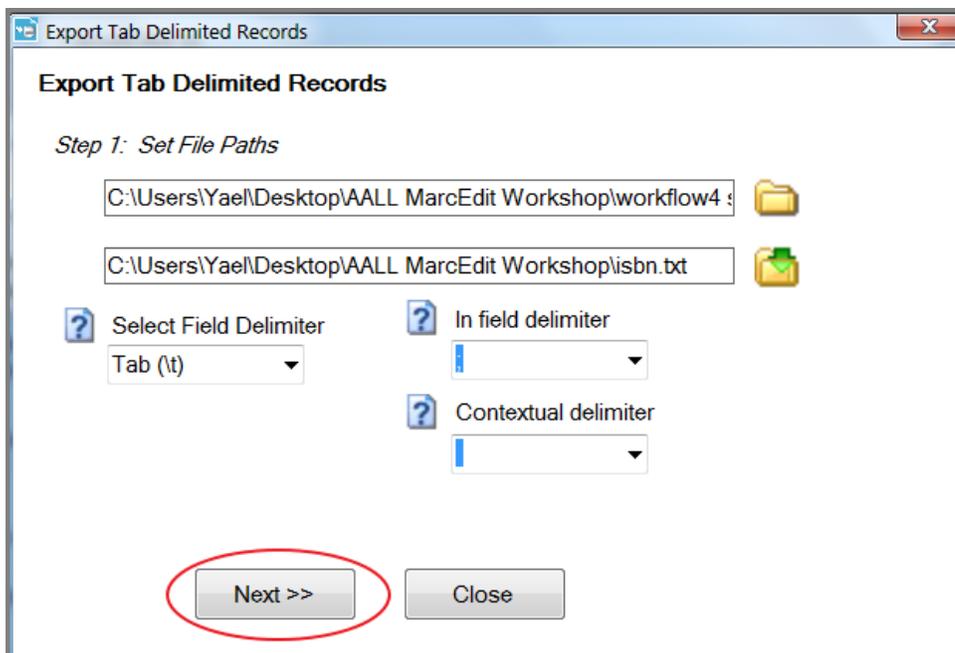
Using the example above, our 020 will now look like this: **\$a9780191827716\$q(ebook)**

4. Export ISBNs into a tab delimited file



Open the MarcEdit Main Window

Tools > Export > Export Tab Delimited Records



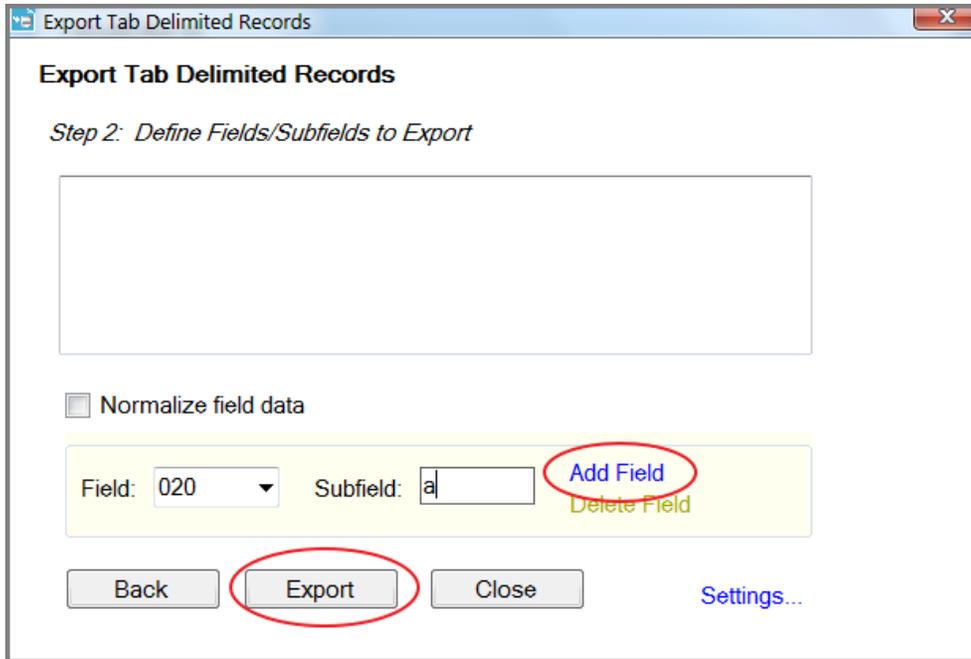
In the top box of the dialog screen select the file of extracted records (**local_source.mrk**)

In the bottom box of the dialog screen select the location for your .txt file and name it **isbn.txt**

Select Field Delimiter: Tab (t)

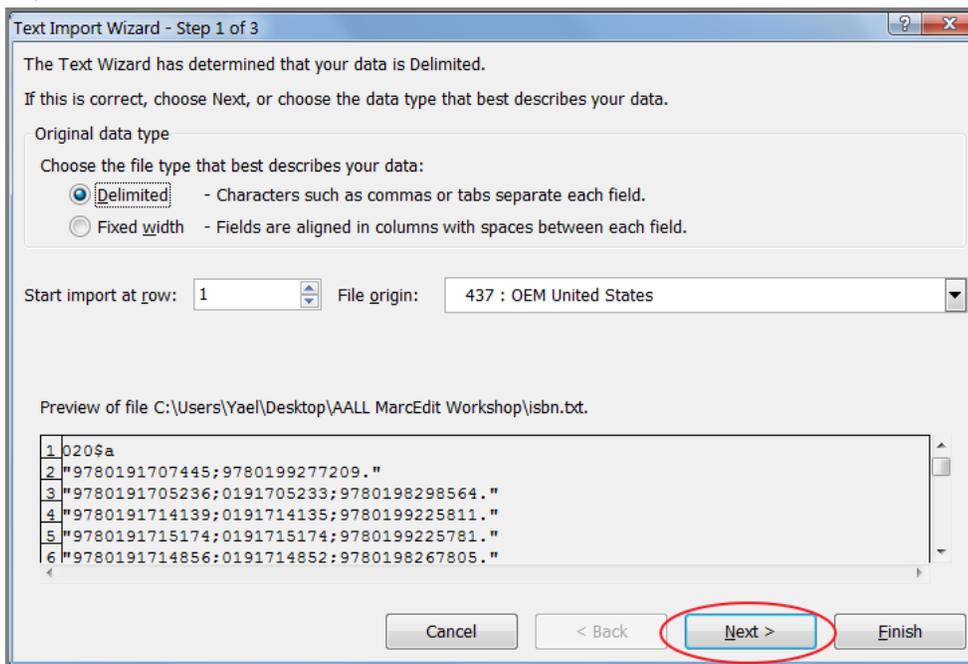
In Field Delimiter: ;

Click **Next**

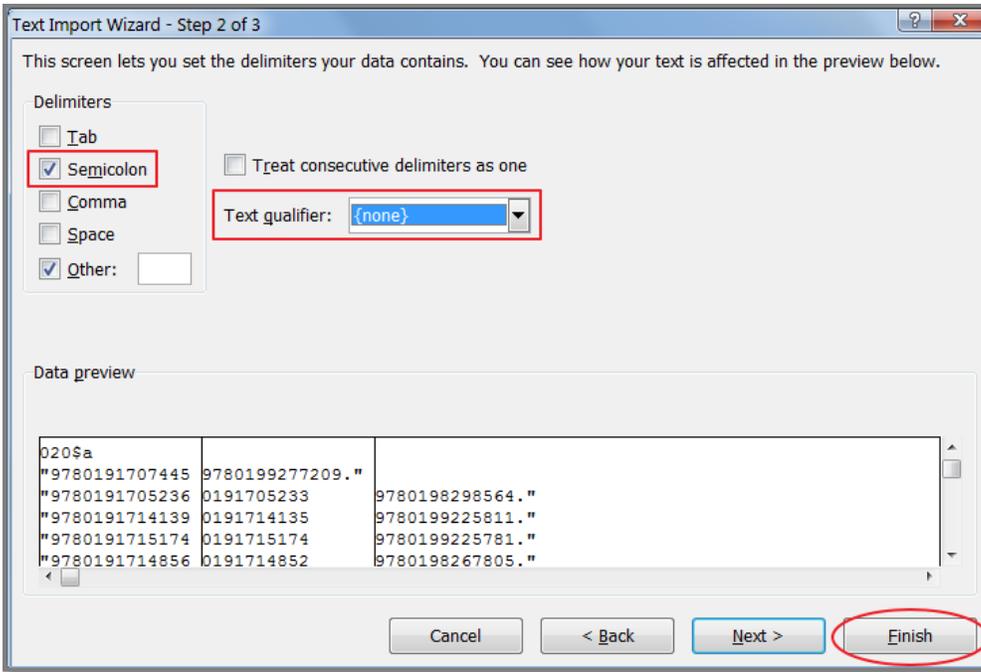


Field: 020
 Subfield: a
 Click **Add field**
 Click **Export**

Open file in Excel



Open Excel
 Click **Office button** on upper left corner > **Open**
 In dialog box change the drop-down menu to All Files *.*
 Locate your **isbn.txt** file and click **Open**
 In the dialog box select **Delimited** and click **Next**

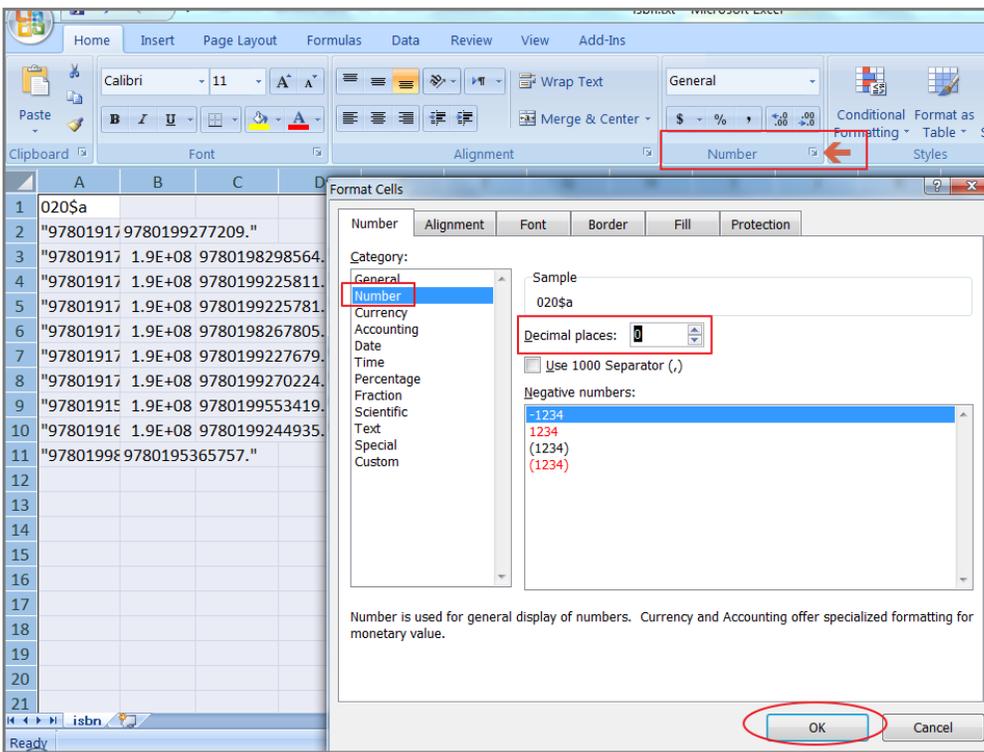


Select **Semicolon**

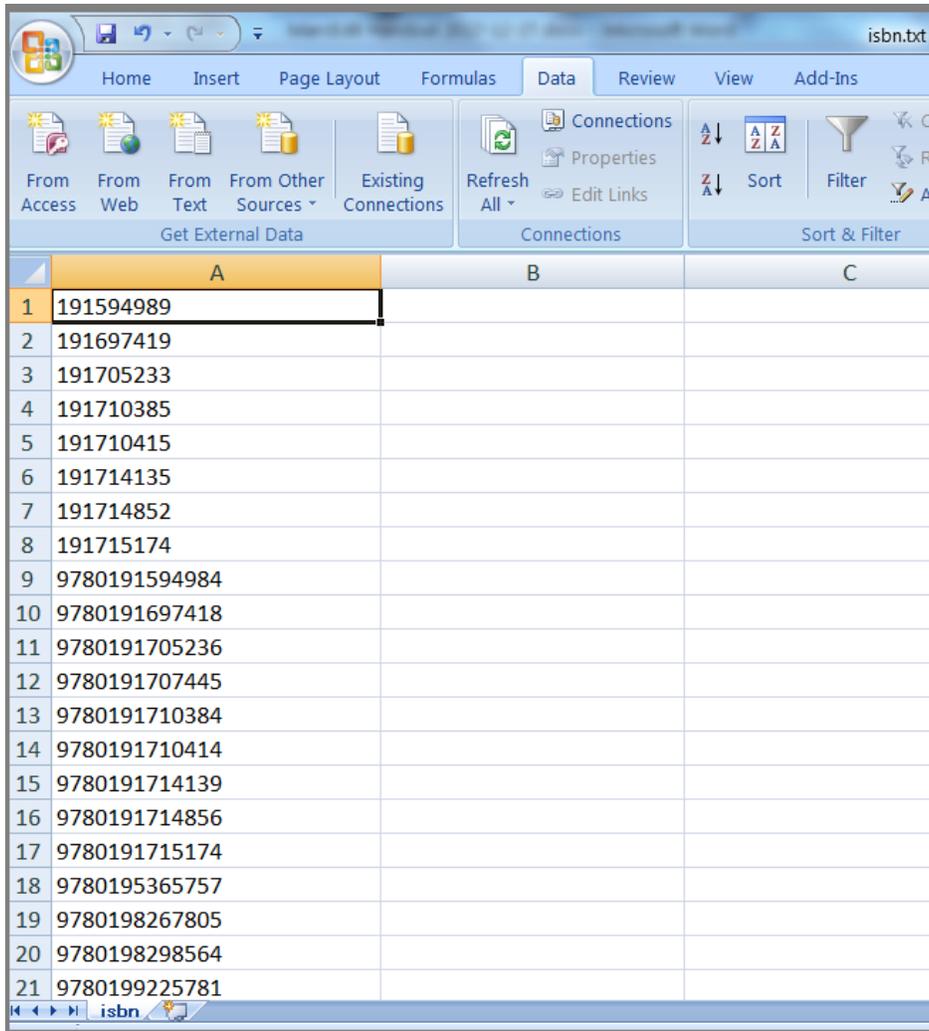
Text Qualifier: **none**

Click **Finish**

Note: If for any reason you are not prompted to reformat the text when opening the .txt file, the same dialog box can be found when you open the file, highlight the first column, and select **Data > Text to Column**



If numbers don't display properly even after you increase column size, highlight column(s), click **Number**, and in the dialog box select **Number** and change **Decimal places to 0**.



Cut data from second and third columns and paste at the end of the first column to create a single column with all the ISBNs

Delete first row with 020\$a

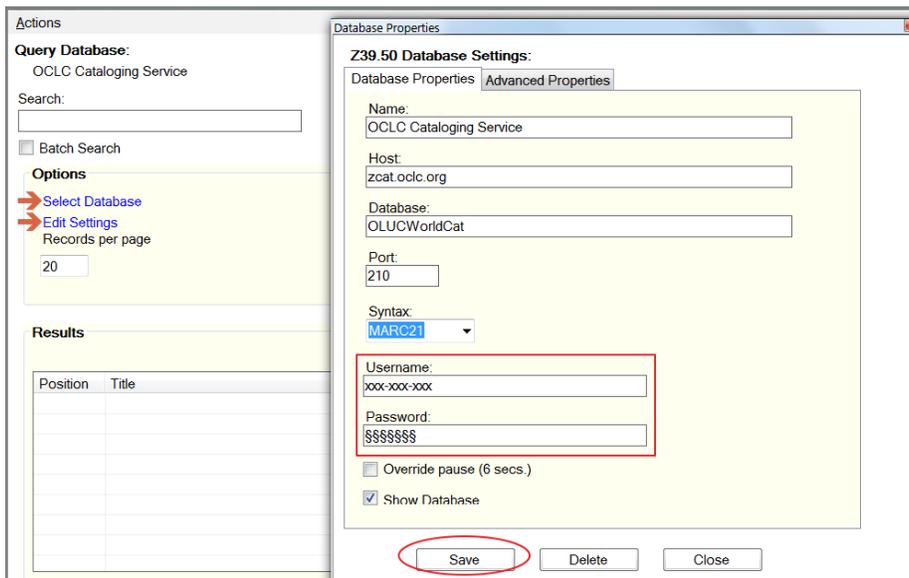
Using **Find & Replace** (Ctrl+H), remove quotation marks and periods

Sort numbers via **Data > Sort**

If file has duplicate numbers, Dedup them via **Data > Remove Duplicates**

Save file as **isbn_oclc.txt**

5. Batch-search OCLC via MarcEdit Z39.50



In MarcEdit Main Window select **Tools > MarcEdit Z39.50/SRU Client**

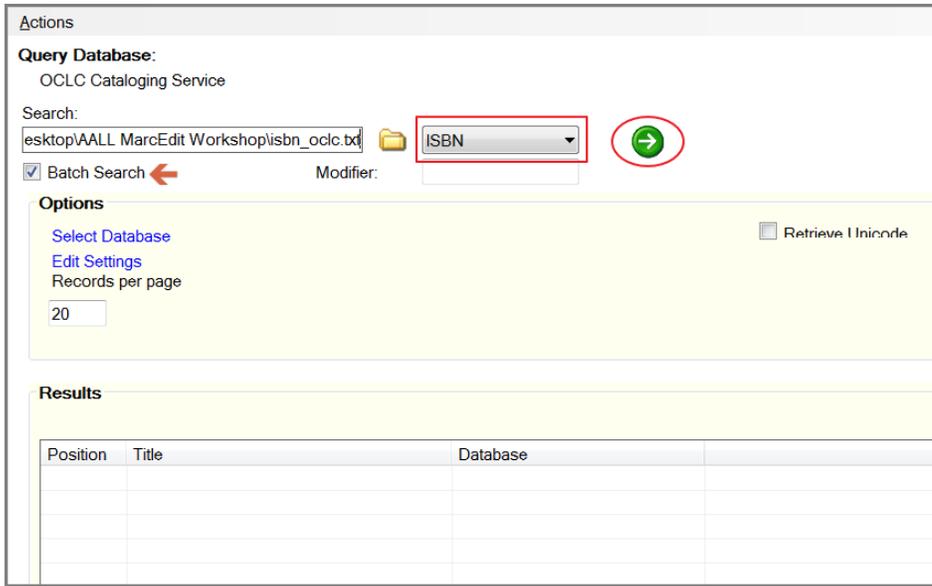
The first time you use this tool you'll need to set up your OCLC connection:

Expand Options and click **Select Database**

Click **OCLC Cataloging Service** and click **Select**

Click **Edit Settings**

Insert your OCLC **Username** and **Password** and **Save**



Select **Batch Search**

Click on file icon to the right of the search box and select your saved file **isbn_oclc.txt**

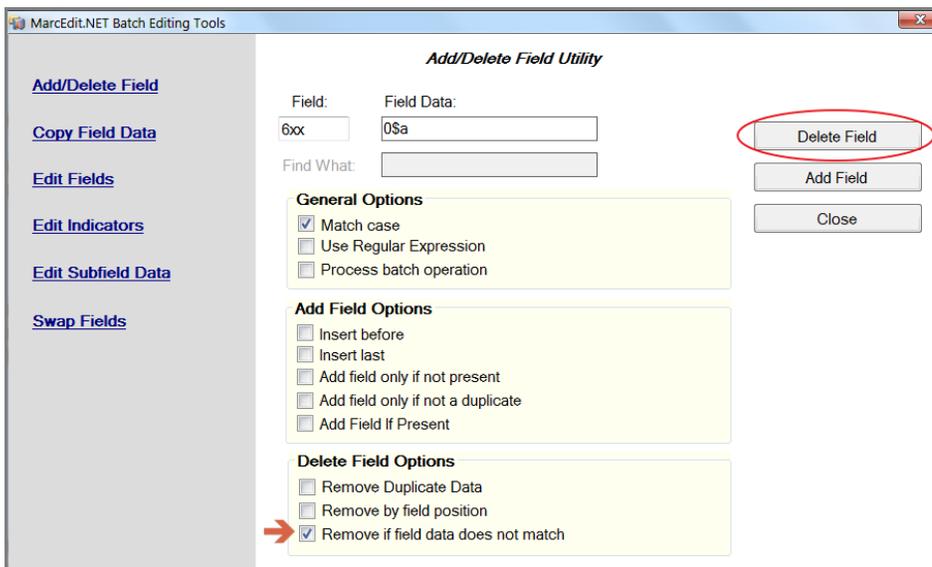
Select **ISBN**

Click the **green arrow**

In the dialog box that opens select where you want the OCLC file to be saved and name that file **oclc.mrc**

The ISBNs are now being searched and matched records are saved in the **oclc.mrc** file. Note that each ISBN search often retrieves multiple records, and that batch searching big files may take a long time

6. Open **oclc.mrc** and remove non-LC 6xx



Option 1: Remove all 6xx without 2nd indicator "0" (655 #7 will be removed as well. If file includes 655 with \$2lcf, use **Option 2** below)

Note that both options will not remove 653 #0 (uncontrolled topical term), so this field will have to be removed separately

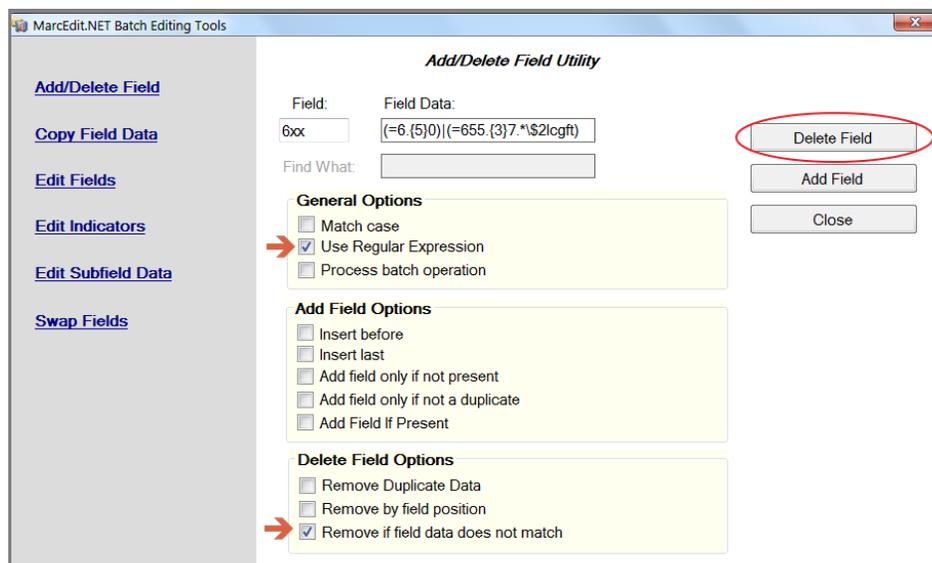
Tools> Add/Delete Field

Field: 6xx (covers 600-699)

Field Data: 0\$a (2nd indicator "0" for LCSH followed by \$a)

Select **Remove if field data does not match**

Click **Delete Field**



Option 2: Remove 6xx without 2nd indicator “0” and 655 #7 without \$2lcgft

Field: 6xx

Field Data:
(=6.{5}0)|(=655.{3}7.*\\$2lcgft)

Select **Use Regular Expressions**

Select **Remove if field data does not match**

Click **Delete Field**

Explanation

(=6.{5}0) matches =6 + any succeeding 5 characters (2 digits following the 6 + 2 spaces + 1st indicator) + 2nd indicator “0”

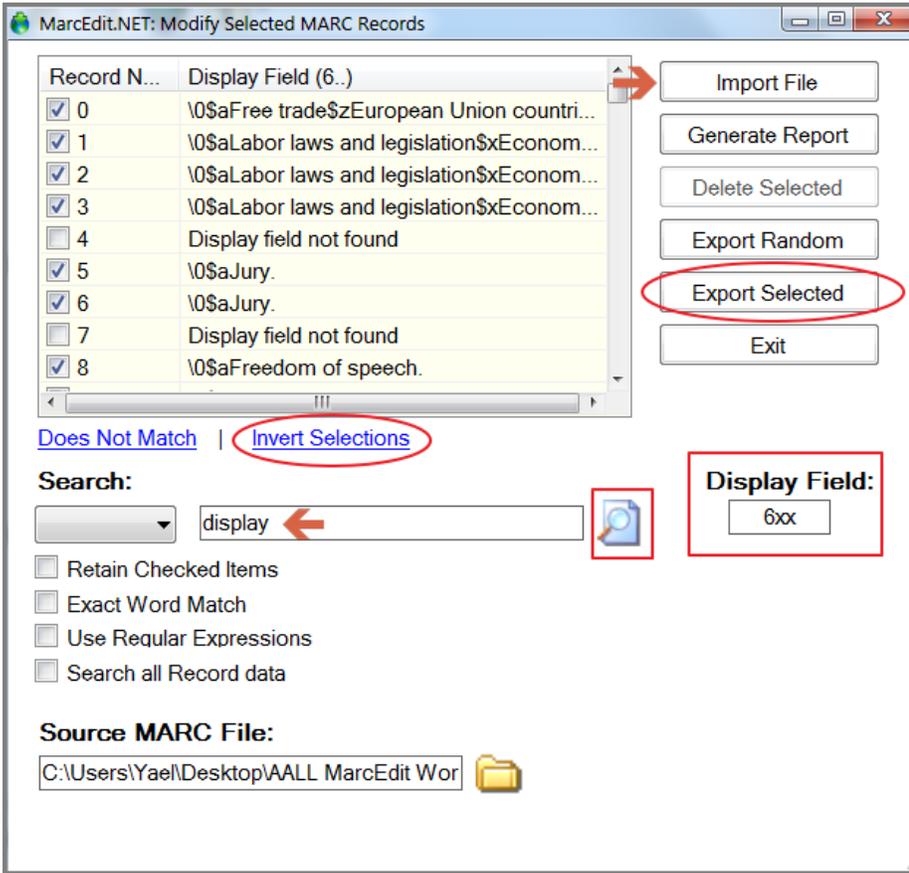
| (pipe) is regex metacharacter for “OR”

(=655.{3}7) matches =655 + any three characters (2 spaces + 1st indicator) + 2nd indicator “7”

.* matches any characters (in this case it will match \$a + the genre/form term)

\\$2lcgft) matches \$2 with “lcgft” (the dollar sign is treated here as a literal, and since in regex it’s a metacharacter, it needs to be escaped with a “\”)

7. Extract records with LCSH



In your cleaned up OCLC file, go to **File > Select Records to Edit**

Display Field: 6xx

Click **Import File**

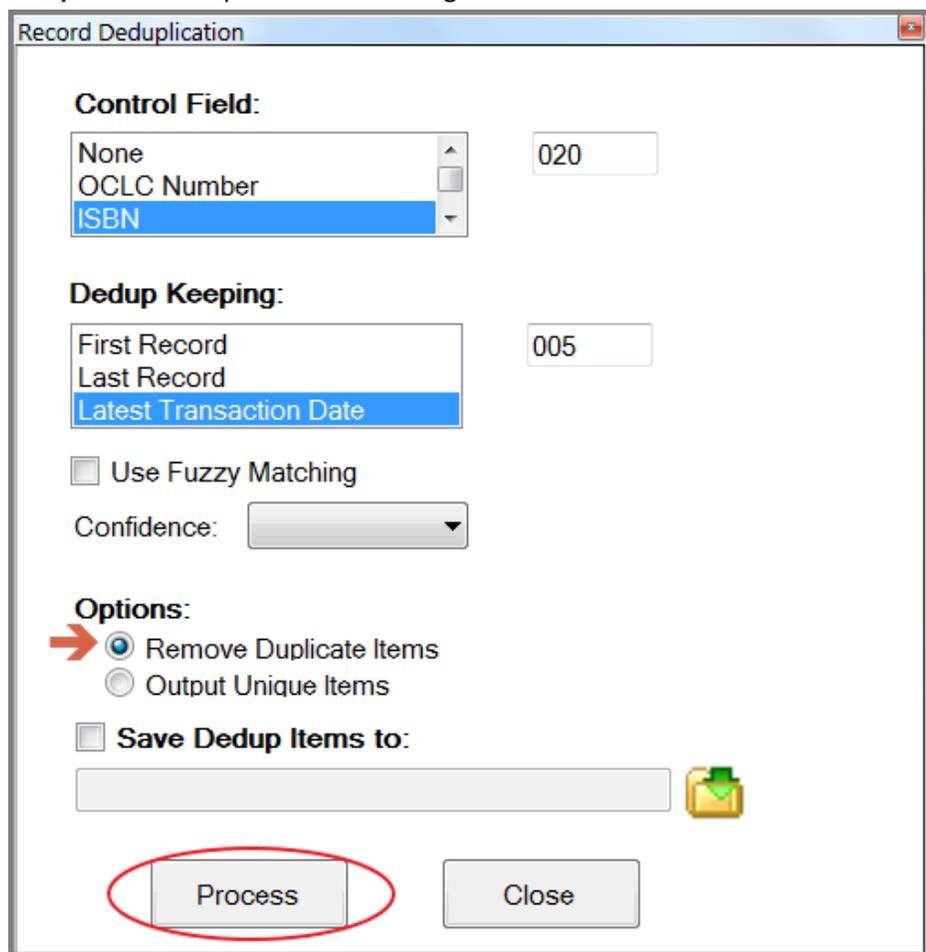
Search: display
This will select all records with "display field not found"

Click **Invert Selections**
Inverting the previous selection will result in records WITH 6xx

Click **Export Selected**

Resave new file as **oclc_merge.mrk**

8. Optional: Dedup Records matching on ISBN



Since the OCLC file is likely to include multiple records per ISBN, we may choose to reduce that number by deduping the records matching on the ISBN.

Pro: reduce number of LCSH merged into each record in our local file (though exact duplicate fields can be deduped after merge)

Con: it's a blunt instrument that may remove some records with better quality headings

Tools > Record deduplication

Control Field: ISBN

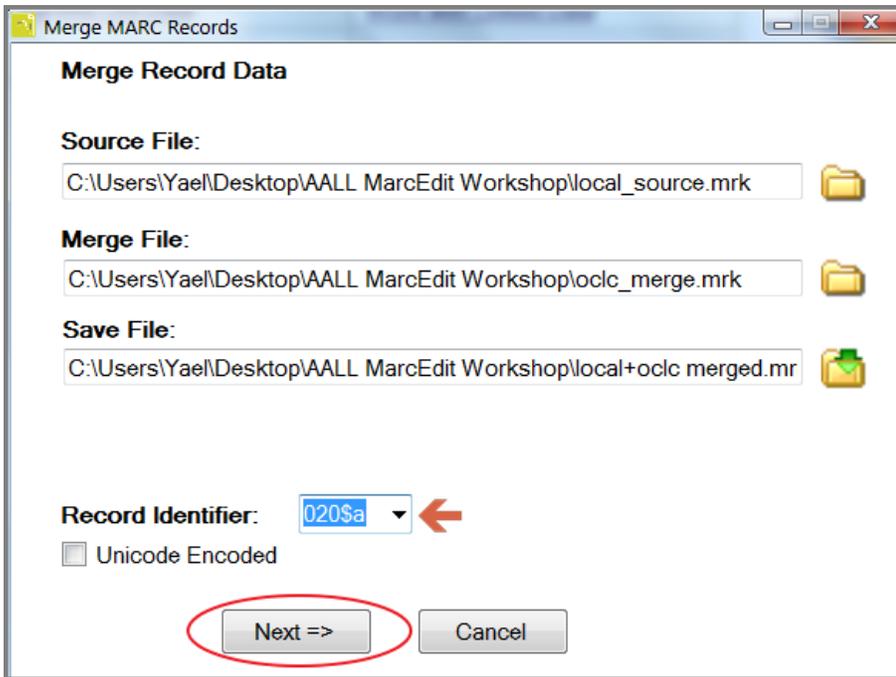
Dedup Keeping: any

Click **Process**

Save and close file

9. Merge LCSH into our local file, matching on ISBN

We will now merge LCSH from **oclc_merge.mrk** into **local_source.mrk**



Open the MarcEdit Main Window

Tools > MARC Processing Tools > Merge Records

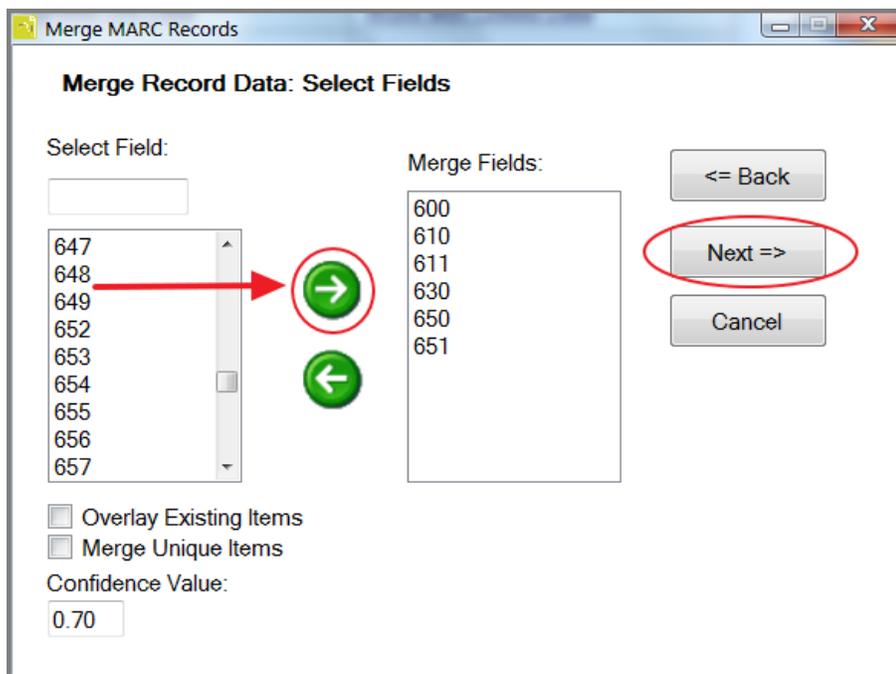
In Source File box retrieve **local_source.mrk**

In Merge File box retrieve **oclc_merge.mrk**

In Save file box select your location and save file as **local+oclc_merged.mrk**

Record Identifier: 020\$a

Click **Next**

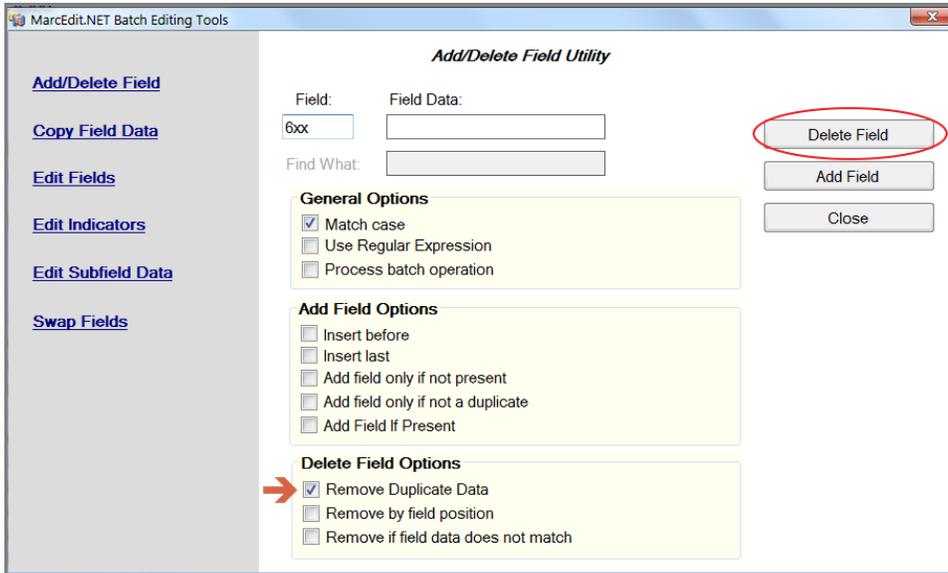


Select fields you wish to merge from the left column and move them to the right column by using the **green right arrow**

When all the fields are selected, click **Next**

10. Dedup 6xx fields

Since we merged 6xx from multiple OCLC records, we are likely to end up with some duplicate data, which should be deduped.



Open file
local+oclc_merged.mrk

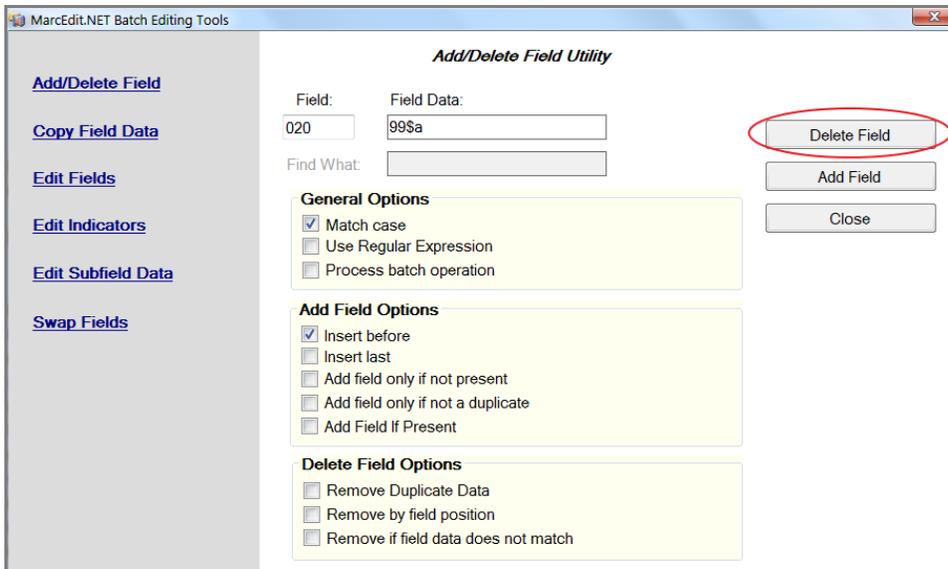
Tools > Add/Delete Field

Field: 6xx

Check **Remove Duplicate Data**

Click **Delete Field**

11. Delete 020 with “fake” 99 indicators



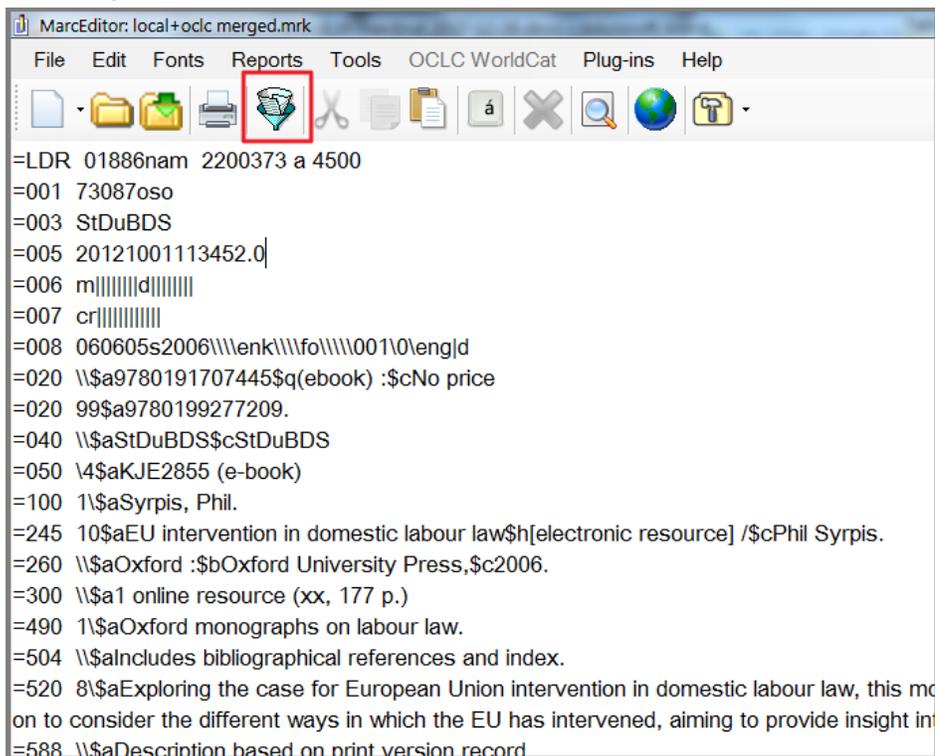
Tools > Add/Delete Field

Field: 020

Field Data: 99\$a

Click **Delete Field**

12. Compile file with enhanced LCSH into MARC via



Click **Compile Records** icon

OR

Select **File > Compile File Into MARC**

Select file location and name file

Load file with enhanced records to ILS

Note: If you plan to do this regularly with files you receive from a specific vendor, **you may want to protect the 6xx fields from overlay** in the load table for records from that vendor, otherwise your LCSH will not be retained during future loading of updated records.

Workflow 5: Converting KBART to MARC

Description: We are subscribed to a new database and would like to add MARC records to our catalog for the individual titles, but are unable to obtain such records. The only data available to us is the content provider's KBART file, which we can convert to MARC via MarcEdit Delimited Text Translator. The result will be brief records that may later be enhanced with classification and subject headings using some of the processes described above.

About KBART

Knowledge Bases and Related Tools (KBART) is a set of guidelines created by NISO and UKSG to ensure the timely transfer of accurate metadata to knowledge bases and link resolver providers. KBART files are usually available in both .txt and .xlsx formats, with column headers describing the data in particular fields (e.g., title, publisher, URL), and the rows containing data about specific titles.

Although in this workflow we'll be working specifically with a KBART file, MarcEdit Delimited Text Translator can be used for converting any delimited text to MARC

For this workflow, we'll use files **kbart.xlsx** and **workflow5.xlsx** to perform the following:

1. Review and edit KBART file in Excel
2. Convert KBART data to MARC

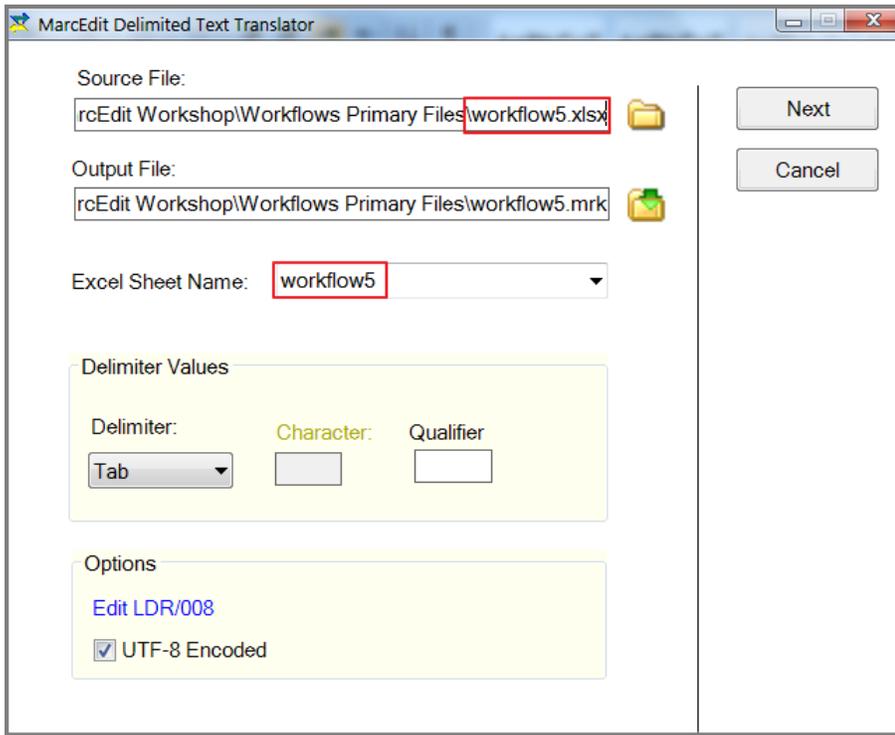
MarcEdit function covered

- MarcEdit Main Window > Tools > Delimited Text Translator (Ctrl+D)

Step-by-step Guidelines

1. Open **kbart.xlsx** in Excel and review available fields
2. To make the mapping process from KBART to MARC less cluttered, decide which columns you want to map to MARC and delete all other columns (e.g., delete columns with no data and fields such as 'access_type' and 'discipline' that you may not need in your MARC records. For this workflow we'll retain nine columns: **publication_title; online_identifier; title_url; first_author; publisher_name; monograph_edition; first_editor; copyright_year; series**). When done, resave file as **workflow5.xlsx**
3. In MarcEdit Main Window select **Tools > Delimited Text Translator**

4. Select your source and input files

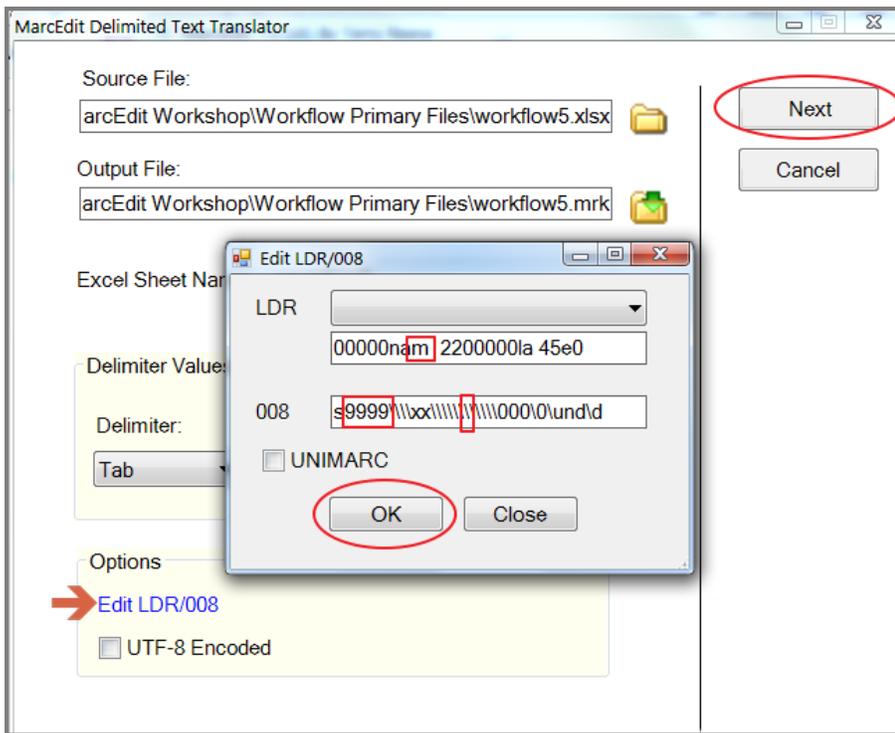


Source File: **workflow5.xlsx**

Output File: [new MARC file]

Note: When working with an Excel file, you also need to select the name of the Excel sheet, which by default is the same as the file name. If you are working with a .txt file, ignore this box

5. Edit Leader and 008



Click **Edit LDR/008**

Check bib level (LDR/07)

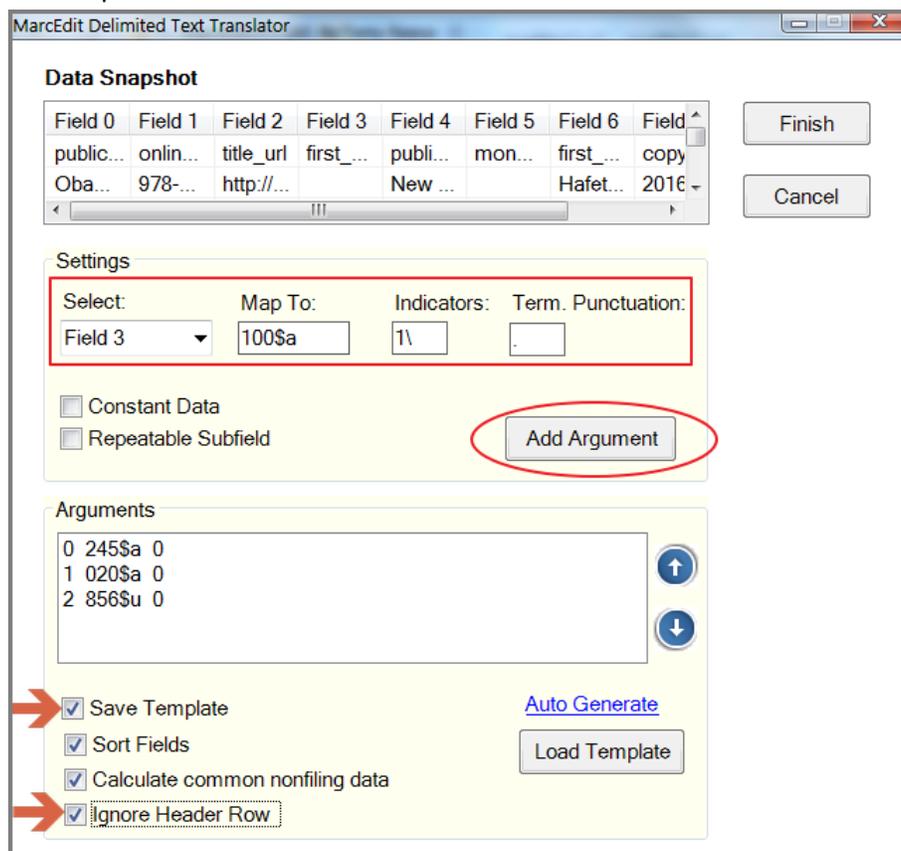
To use the date in 264\$c, change 008/07-10 from **9999** to **{264\$c}**

Change 008/23 from \ to **o**

(This and other Leader and 008 edits can be done later via Edit Field Data)

Click **OK** then Click **Next**

6. Map data to MARC



Select Field #

Map column to MARC field and subfield

Add indicators and end punctuation

Click **Add Argument**

To edit/delete argument, highlight it, right click, and select **Edit** or **Delete**.

If you wish to join two subfields into a single MARC field (e.g., 264\$a & 264\$c), highlight both fields, right click, and select **Join**

Check Ignore Header Row, otherwise the first MARC record will consist of the headers

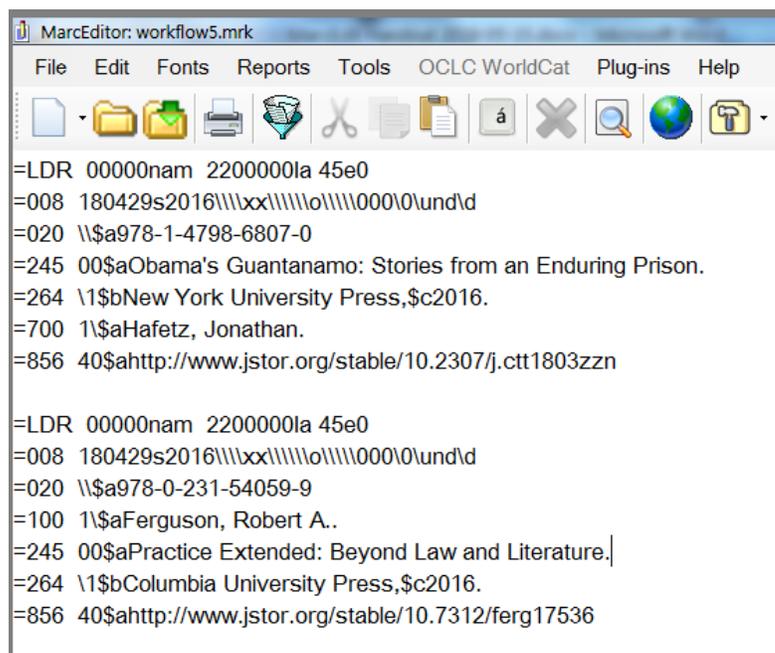
If you wish to save this mapping for future use, select Save Template

When done mapping, click **Finish**

Note that the Arguments box does not display indicators and end punctuations (the zeros on the right mean the arguments are not constant data. Constant data would display as 1)

Mapping workflow5.xlsx to MARC

- Field 0 (publication_title) → 245\$a
- Field 1 (online_identifier) → 020\$a
- Field 2 (title_url) → 856\$u
- Field 3 (first_author) → 100\$a
- Field 4 (publisher_name) → 264\$b
- Field 5 (monograph_edition) → 250\$a
- Field 6 (first_editor) → 700\$a
- Field 7 (copyright_year) → 264\$c
- Field 8 (series) → 490\$a



Homework: Edit & Enhance Your MARC File

LDR/008: Edit via **Tools > Edit Subfield Data**

Example: Change 008/35-37 (language) from **und** to **eng**: Field: **008**, Position: **35:3** (position 35, total of 3 bytes), Find: **und**, Replace: **eng**

020: Remove dashes via **Tools > Edit Subfield Data**

245: Add **\$b** after colon via **Tools > Edit Subfield Data**

264: Add **\$a[Place of publication not identified]** via **Find & Replace**

Control numbers: Generate via **Tools > Generate Control Numbers**

Classification / FAST headings: Add via **Tools > Call Number Tools > Generate Call Numbers**

LC Subject headings: Add following the process described in **Workflow 4**

Extra Challenge: Change **245** indicator **0** to **1** ONLY in records with 1xx fields. Use **Find & Replace** and check the conditional '**Perform Find/Replace if ...**'

(If you need help, see screenshot at <https://snag.gy/K5cSGN.jpg>)

Regular Expressions (regex) Commonly Used in MarcEdit

Metacharacters: . ^ \$ | \ () * + ? { } [] -- These characters need to be escaped with a backslash (\) to match literal characters

Anchors, Escapes, and other Special Characters

^ (caret)	Matches beginning-of-line position	^cat matches cat , catalog , but NOT scatter
\$ (dollar)	Matches end-of-line position	ght\$ matches night , bright , but NOT lightning
\ (backslash)	Specifies the next character as either a special character or a literal	\\$a matches \$a
\n	Matches new line	.\n. matches This is first line; This is second line
. (period)	Matches any single character	b.g matches big , bag , bug
 (pipe)	Or	cat dog matches cat , dog
() (parentheses)	Group. When matching a pattern within parentheses, we later use \$1, \$2, etc. to refer to the previously matched pattern	Reverse [Last name], [First name] to [First name] [Last name]: Find: ([A-Z][a-z]+), ([A-Z][a-z]+) Replace: \$2[space]\$1

Quantifiers

* (asterisk)	Matches preceding expression zero or more times (.* matches any character zero or more times)	il*e matches mile , pier , grille 12*5 matches 15 , 125 , 12225 , etc.
+ (plus)	Matches preceding expression one or more times	ar+ matches are , arrow , carrier but NOT apple 12+5 matches 123 , 1225
? (question mark)	Matches the preceding expression zero or one time	mo?r matches more , ramrod but NOT mooring 12?5 matches 15 , 125
{n}	Matches preceding expression exactly n times	12{3}5 matches 12225
{n,}	Matches preceding expression n or more times	12{3,}5 matches 12225 , 122225 , 1222225 , 12222225 , etc.
{n,m}	Matches preceding expression between n and m times	12{3,5}5 matches 12225 , 122225 , 1222225

Character Classes

Only **^ - \]** need to be escaped inside a character class

[abc] (Same as (a b c))	Character set. Matches any one of the enclosed characters.	s[mc] matches small , scroll
[^abc]	Negative character set. Matches any character NOT enclosed	c[^au]t matches cot but NOT cat or cut
[-] (hyphen with square brackets)	A range of characters	[a-zA-Z] matches any lowercase/uppercase letter [0-9] matches any digit
\d (Same as [0-9])	Matches any digit	1\d3 matches 103 , 113 , 123 , 133 , etc.
\D (Same as [^0-9])	Matches any non-digit	1\D3 1a3 , 1@3 , 1B3
\w (Same as [a-zA-Z0-9])	Matches any alphanumeric	1\w3 123 1a3 1B3
\W (Same as [^a-zA-Z0-9])	Matches any non-alphanumeric	1\W3 1%3 , 1+3 , 1\$3
\s (Same as []*)	Matches whitespace	.\s{2,3}[a-z] matches 1. xyz 2. xyz
\S (Same as [^]*)	Matches non-whitespace	.*\.\S.* matches space missing between period and beginning of new sentence, e.g., abc.Xyz