Automating Authority Control Procedures

Written by Stacey Wolf (stacey.wolf@unt.edu)

# Batch loading Name authority records from the Headings Reports or vendor files

\*\*This could also be done with subjects, but pay closer attention to the MARC tags and select 6XX instead of 1XX/7XX when Validating Headings.

## 1. Extract new names from your Headings Reports

a. Highlight and copy the first time Author names off the heading report and paste into Excel (Sheet1).

 DO NOT CLEAR THE HEADING REPORT YET!!

b. Clean up so only the name remains.

i. Deleting columns E-I, A-C

ii. Sort remaining column alphabetically

iii. Delete everything that doesn’t start with “Field: a1…” or “Field: b7…”

iv. If there are Corporate Names in the a110 or b710, copy these to a separate sheet and repeat the process using 110$a as the Map to: in step 2.b.4.b.

v. Remove the “Field: a100 10$a”using Text to Columns,

1. Select “fixed width” and then hit “next”.

2. Use the line arrows to place it just between the “a” and the beginning of the names. Select “next”.

3. And then select “finish”.

4. Delete the column of “field…”.

vi. Remove delimiters using Text to Columns again.

1. Use the “delimited” selection and hit “next”.

2. Uncheck any of the boxes under the delimiters, except “other”

a. Enter | (or the delimiter symbol from Sierra).

b. Click “next” and then “finish”.

## 2. Get the Library of Congress control number (LCCN) for authorized names using MarcEdit to Validate Headings

a. FOR A VENDOR FILE OF RECORDS:

i. (Don’t save these changes so save before hand)

ii. Use Reports -> Validate Headings -> Validate Headings

iii. Select “Authorize 1xx/7xx fields” and check “embed URIs when Possible”.

iv. Click process.

v. Close out of the report.

b. FROM THE HEADINGS REPORTS NAMES:

i. Save your Excel file.

ii. Open MarcEdit and click on "Delimited Text Translator"

1. Source file is your Excel spreadsheet

a. This will populate the drop down box next to "Excel Sheet Name"

2. Output file is whatever name you want to name it

3. Select the appropriate sheet from your Excel file, probably Sheet1

4. On the next screen, add an argument to create 100 fields

a. Select: Field 0

b. Map to: 100$a

c. Indicators: \\

d. Term. Punctuation: blank

e. Click "Add Argument"

5. Click "Finish"

iii. Open your new MARC file in MarcEdit

iv. Select “Authorize 1xx/7xx fields” and check “embed URIs when Possible”.

v. Click process.

vi. Close out of the report.

c. Find all of these to copy to Excel

=.00.\*\$0

(Check use regular expressions).

1. Copy the results to another tab in Excel (Sheet2).

2. Close MarcEdit without saving.

d. Clean up the names so that only the last name, first name and the LCCN remains.

i. Delete Column B is the “Jump to Record #” information.

ii. Use Text to Columns to remove “=\*00 1\$a”

1. Select “fixed width” and then hit “next”.

2. Use the line arrows to place it just between the “a” and the beginning of the names. Select “next”

a. Be sure there are no other arrows dividing the field

3. And then select “finish”.

4. Delete column A

iii. Isolate the LCCN.

1. Replace $0 with ^

2. Use Text to Columns

a. Use the “delimited” selection and hit “next”.

b. Uncheck any of the boxes under the delimiters, except “other”

i. Enter ^

ii. Click “next” and then “finish”.

3. Cut column B with the URL and insert before column a.

iv. Remove the rest of the delimiters using Text to Columns again.

1. Use the “delimited” selection and hit “next”.

2. Uncheck any of the boxes under the delimiters, except “other”

a. Enter $

b. Click “next” and then “finish”.

3. Deleted the rest of the columns, keeping only A and B with the URLs and the names.

v. Find/replace all:

http://id.loc.gov/authorities/names/

[blank]

e. Add a filter to the columns.

i. Insert a blank row above the first row

ii. Highlight columns A and B.

iii. In the Data tab, click on “Filter”

f. Use the names from Sierra (Sheet1) to limit the names with LCCNs (Sheet2) to just those that the heading report said we need.

i. Copy the columns of names from Sheet1 to the bottom of column B with the names on Sheet2.

ii. Highlight duplicate values

1. Select columns A and B.

2. Use the “Conditional formatting—Highlight cell rules—Duplicate values”

3. Remove any duplicates found in Column A, leaving just one entry per duplicate.

iii. Sort Column B based on cell color using the drop down in the filter

iv. Delete all of the cells in column A and B that weren’t highlighted in the previous steps.

g. Save the LCCNs as a text file

i. Copy the LCCNs for duplicate names to a new tab (Sheet3)

ii. Save As just Sheet3 as a .txt file or text (tab delimited)

1. Name it whatever makes sense to you

2. Click ok to save just the one sheet

3. Click yes to continue saving

## 3. In OCLC, batch load the LCCNs to save these records to the local file.

a. Use the Batch tab to find “Enter Authority Search Keys”

i. Change the default index to “LCCN (ln:)”

ii. Click on “import” to find your text file.

iii. Don’t delete your file (in case it doesn’t work).

1. If you can’t get this file to load, make a copy of your text file and try loading the copy.

iv. Click “save” and then “close”.

b. Use Batch tab to “Process Batch”

i. Under Path, check the second option (File name DefaultAuth.auth.db).

ii. Under process, check the “online searches” box.

iii. Click ok.

iv. Notice any entries that had more than one record found for it.

## 4. Open the authority records in the local file

a. Authorities --> “Show” --> “By local save file status”

i. Check the “In Process” box under “Workflow Status”

ii. Click “Show Records”

b. Investigate any duplicate entries to see which record you want to keep or remove them all.

c. Select all of the records

## 5. Add a 949 field to add fixed fields during the import process into Sierra

a. Edit --> Constant Data --> Online --> Apply from List --> Hit “search” without entering any information

i. Highlight “\*atab=anam”

ii. Click “apply selected”

iii. Close out

iv. Open a few records from your list to see that the 949 field is there

## 6-1. Export these records to Sierra individually.

a. Select all of the records you need to export and hit F5 or "Export".

b. You can overlay each record by using 949 \*atab=asub;ov=.a####### in OCLC

**Or**

## 6-2. Export these records to Sierra as a batch.

a. Your export setting in OCLC should have the destination set as "File:(prompt for filename)"

b. Select all of the records you need to export and hit F5 or "Export"

c. Save the file under whatever name and in whatever location you want

d. Load into Sierra using "Data Exchange"

i. Select the "Load MARC NAME authority records from tape or FTS"

ii. "Upload files from PC" using the "sfts" type when prompted

iii. Select the newly uploaded file and "preprocess" it

iv. Click on the newly created "sub" type line

v. Select load

vi. Check "use review files" and test the file

vii. If there are no errors, then load the records

## 7. Check in Sierra to ensure they imported

a. Close out of the heading report and reopen it. The first time names should be gone.

b. Finish processing the rest of the report.

## 8. Delete your list in OCLC.

a. Select all records.

b. Hit the ActionDeleteRecord image (record with red arrow into trash can) or (Ctrl+Alt+D)

i. Click “yes” to delete all of the records

c. Close out of your report too.

## 9. Delete/archive your Excel and/or NotePad files.

# Batch Update Subject Headings

https://www.loc.gov/aba/cataloging/subject/weeklylists/​

## 1. Extract the LCSH record number from the list of changed headings​

a. Copy the subject change list into Excel​

b. Highlight everything and unmerge cells​

c. Use (ctrl+f) Find/replace to replace [sp with ^​

d. Use (ctrl+f) Find/replace to replace ] with nothing​

e. Use the Text to columns feature to separate ^ to its own column

i. Delimited --> ​

ii. Other: ^ --> ​

iii. Finish

## ​2. Paste the copied numbers from step 1 into a new tab in Notepad ++​ and format these to form the search query

a. Each number should be on its own line​

b. Remove any empty lines (EDIT--LINE OPERATIONS--REMOVE EMPTY LINES)​

c. Use Search -> Replace (Ctrl-H) and Regular Expressions to add the necessary text around the numbers​

i. Find what: ​(^.\*)​

ii. Replace with: ​{"target": {"record": {"type": "authority"},"field": {"marcTag": "010"}},"expr": {"op": "regex","operands": ["sh\s\*$1"]}},"or",​

iii. Check “Regular expression”​

iv. Click “Replace All”​

v. Note: Sierra has the record numbers stored in the 010 field with a prefix of sh and some number of spaces. In order to avoid searching for name record numbers also, change the sp from the original list to sh and allow for any number of spacing in the search.​

d. Open a new tab in Notepad ++​

e. Copy the following JSON code into that new tab​

{​"queries":[​

ENTER FORMATTED SEARCH TERMS HERE​

]​}​

## 3. Run the search query through Sierra

a. Name your file​

b. Store record type =​ a authority​

c. Select JSON​

d. Highlight the {} and paste your search from Notepad++​

e. Click Search​

## 4. Use the results of the search to find these records in OCLC

a. Export the 010 fields

b. Use the list of 010 fields you exported from Sierra and save them as a text file.

c. In OCLC, under "Batch" select "Enter Authority Batch Search Keys"

i. Use the default index of "LCCN (ln:)"

ii. Hit import

iii. Find your file location and select it

\*I always click no when it asks me if I want to delete my original import file.

iv. The dialog box under the default search index should have populated with your search terms

v. Hit save and then close

d. Under "Batch" select "Process Batch"​

i. Check the box next to the default auth file name (as opposed to bib)​

ii. Click online searches and then click ok

e. You should get a "Batch Search Report"​

f. View the records you just saved to the local file

i. Use the "Authorities" – "Show" - "By local save file status"

ii. Check the "in Process" box under "Workflow status"

iii. Click "show records"

## 5-1. Export these records to Sierra individually.

a. Select all of the records you need to export and hit F5 or "Export".

b. You can overlay each record by using 949 \*atab=asub;ov=.a####### in OCLC

**Or**

## 5-2. Export these records to Sierra as a batch.

a. Your export setting in OCLC should have the destination set as "File:(prompt for filename)"

b. Select all of the records you need to export and hit F5 or "Export"

c. Save the file under whatever name and in whatever location you want

d. Load into Sierra using "Data Exchange"

i. Select the "Load MARC SUBJECT authority records from tape or FTS"

ii. "Upload files from PC" using the "sfts" type when prompted

iii. Select the newly uploaded file and "preprocess" it

iv. Click on the newly created "sub" type line

v. Select load

vi. Check "use review files" and test the file

vii. If there are no errors, then load the records

viii. If all of the records are added as new and not overlaid, you will want to delete the old records, which can easily be done with the original create list file from step 3

## 6. Check your headings reports for duplicates and blind references

# Batch Update Closed-Dates and Updated Name Headings Using JUST RECORD NUMBERS

https://www.oclc.org/authority-records.html​

## 1. Build the JSON query in Notepad++

a. Copy and paste the list from the OCLC website into Notepad++​

b. Delete the header text and remove any empty lines (EDIT--LINE OPERATIONS--REMOVE EMPTY LINES)​

c. Sort list so that the record numbers are all together at the end of the document. (EDIT—LINE OPERATIONS—SORT LINES LEXICOGRAPHICALLY ASCENDING)

d. Delete everything except the record numbers

e. Use Search -> Replace (Ctrl-H) and Regular Expressions to strip out everything except the numbers of the record number

i. Find what: ^[^0-9]+([0-9]\*)

ii. Replace with: $1​

iii. Check “Regular expression”​

iv. Click “Replace All”​

f. Use Search -> Replace (Ctrl-H) and Regular Expressions to add the necessary text around the numbers​

i. Find what: ​(^.\*)​

ii. Replace with: ​{"target": {"record": {"type": "authority"},"field": {"marcTag": "010"}},"expr": {"op": "regex","operands": ["n.\*$1"]}},"or",​

iii. Check “Regular expression”​

iv. Click “Replace All”​

g. Open a new tab in Notepad ++​

h. Copy the following JSON code into that new tab​

{​"queries":[​

ENTER FORMATTED SEARCH TERMS HERE​

]​}​

## 2. Run the search query through Sierra

a. Name your file​

b. Store record type =​ a authority​

c. Select JSON​

d. Highlight the {} and paste your search from Notepad++​

e. Click Search​

## 3. Use the results of the search to find these records in OCLC

a. Export the 010 fields

b. Use the list of 010 fields you exported from Sierra and save them as a text file.

c. In OCLC, under "Batch" select "Enter Authority Batch Search Keys"

i. Use the default index of "LCCN (ln:)"

ii. Hit import

iii. Find your file location and select it

\*I always click no when it asks me if I want to delete my original import file.

iv. The dialog box under the default search index should have populated with your search terms

v. Hit save and then close

d. Under "Batch" select "Process Batch"​

i. Check the box next to the default auth file name (as opposed to bib)​

ii. Click online searches and then click ok

e. You should get a "Batch Search Report"​

f. View the records you just saved to the local file

i. Use the "Authorities" – "Show" - "By local save file status"

ii. Check the "in Process" box under "Workflow status"

iii. Click "show records"

## 4-1. Export these records to Sierra individually.

a. Select all of the records you need to export and hit F5 or "Export".

b. You can overlay each record by using 949 \*atab=anam;ov=.a####### in OCLC

**OR**

## 4-2. Export these records to Sierra as a batch.

a. Your export setting in OCLC should have the destination set as "File:(prompt for filename)"

b. Select all of the records you need to export and hit F5 or "Export"

c. Save the file under whatever name and in whatever location you want

d. Load into Sierra using "Data Exchange"

i. Select the "Load MARC NAME authority records from tape or FTS"

ii. "Upload files from PC" using the "nfts" type when prompted

iii. Select the newly uploaded file and "preprocess" it

iv. Click on the newly created "nam" type line

v. Select load

vi. Check "use review files" and test the file

vii. If there are no errors, then load the records

viii. If all of the records are added as new and not overlaid, you will want to delete the old records, which can easily be done with the original create list file from step 3

## 5. Check your headings reports for duplicates and blind references

6. Update names in bibliographic records now or wait to see what bibliographic maintenance doesn't fix overnight.

# ​Batch Update Closed-Dates and Updated Name Headings Using OLD NAME FORMATS

https://www.oclc.org/authority-records.html​

## 1. Build the JSON query in Notepad++

a. Copy and paste the list from the OCLC website into Notepad++​

b. Delete the header text and remove any empty lines (EDIT--LINE OPERATIONS--REMOVE EMPTY LINES)​

c. Sort list so that the record numbers are all together at the end of the document. (EDIT—LINE OPERATIONS—SORT LINES LEXICOGRAPHICALLY ASCENDING)

d. Delete everything except old names

e. Use Search -> Replace (Ctrl-H) and Regular Expressions to remove the "Old: 100 1" information

i. Find what: ...: 100 1.

ii. Replace with:

iii. Check “Regular expression”​

iv. Click “Replace All”​

f. Use Search -> Replace (Ctrl-H) and Regular Expressions to remove everything after the last comma or period, leaving just subfield a

i. Find what: (.\*,.\*)[,\.].\*​

ii. Replace with: $1

iii. Check “Regular expression”​

iv. Click “Replace All”​

v. Repeat as necessary until all of the subfields except a are removed

vi. Search for the delimiter symbol to see if there are any that were missed and take care of those manually

g. Use Search -> Replace (Ctrl-H) and Regular Expressions to add the necessary text around the old names

i. Copy and paste the all of the old names twice so that there are three sets of names

ii. Highlight the first set of the names

a. Find what: ​(^.\*)​

b. Replace with: ​{"target": {"record": {"type": "authority"},"field": {"marcTag": "100"}},"expr": {"op": "regex","operands": ["n.\*$1"]}},"or",​

c. Check “Regular expression”​

d. Click “Replace All”​

iii. Highlight the second set of the names

a. Find what: ​(^.\*)​

b. Replace with: ​{"target": {"record": {"type": "authority"},"field": {"marcTag": "700"}},"expr": {"op": "regex","operands": ["n.\*$1"]}},"or",​

c. Check “Regular expression”​

d. Click “Replace All”​

iv. Highlight the third set of the names

a. Find what: ​(^.\*)​

b. Replace with: ​{"target": {"record": {"type": "authority"},"field": {"marcTag": "600"}},"expr": {"op": "regex","operands": ["n.\*$1"]}},"or",​

c. Check “Regular expression”​

d. Click “Replace All”​

h. Open a new tab in Notepad ++​

i. Copy the following JSON code into that new tab​

{​"queries":[​

ENTER FORMATTED SEARCH TERMS HERE​

]​}​

## 2. Run the search query through Sierra

a. Name your file​

b. Store record type =​ b bibliographic​

c. Select JSON​

d. Highlight the {} and paste your search from Notepad++​

e. Click Search​

## 3. Look through the results and compare them to the original list of names

a. Export the 100, 700, and 600 fields from Sierra

b. Copy the list to Excel and alphabetize the list (or you can use Notepad++)

c. Remove duplicates from the list

i. If using Notepad++, remove duplicates with the following Replace

ii. Find what: ^(.\*?)$\s+?^(?=.\*^\1$)​

iii. Replace with:

iv. Check "regular expressions"

d. Compare the names on the list to the list from OCLC

e. Copy the record number from the OCLC list of matched names and use them to run a batch search in OCLC

f. You may need to investigate some names to make sure that a name without a date in your ILS isn't the same person as the updated name. You can search by names instead of LCCN in OCLC to help gather the authority records quicker, but it will still take alot of manual investigation.

## 4. Batch search for the record numbers in OCLC

a. In OCLC, under "Batch" select "Enter Authority Batch Search Keys"

i. Use the default index of "LCCN (ln:)" (or "Personal Names (pn:) if you want to search by names, but this will result in more records than you need)

ii. Hit import

iii. Find your file location and select it

\*I always click no when it asks me if I want to delete my original import file.

iv. The dialog box under the default search index should have populated with your search terms

v. Hit save and then close

b. Under "Batch" select "Process Batch"​

i. Check the box next to the default auth file name (as opposed to bib)​

ii. Click online searches and then click ok

c. You should get a "Batch Search Report"​

d. View the records you just saved to the local file

i. Use the "Authorities" – "Show" - "By local save file status"

ii. Check the "in Process" box under "Workflow status"

iii. Click "show records"

## 5-1. Export these records to Sierra individually.

a. Select all of the records you need to export and hit F5 or "Export"

b. You can overlay each record by using 949 \*atab=anam;ov=.a####### in OCLC

c. See the subject section for names instructions for names used as subjects

**OR**

## 5-2. Export these records to Sierra as a batch.

a. Your export setting in OCLC should have the destination set as "File:(prompt for filename)"

b. Select all of the records you need to export and hit F5 or "Export"

c. Save the file under whatever name and in whatever location you want

d. Load into Sierra using "Data Exchange"

i. Select the "Load MARC NAME authority records from tape or FTS"

ii. "Upload files from PC" using the "nfts" type when prompted

iii. Select the newly uploaded file and "preprocess" it

iv. Click on the newly created "nam" type line

v. Select load

vi. Check "use review files" and test the file

vii. If there are no errors, then load the records

viii. If all of the records are added as new and not overlaid, you will want to delete the old records, which can easily be done with the original create list file from step 3

## 6. Check your headings reports for duplicates and blind references

## 7. Update names in bibliographic records now or wait to see what bibliographic maintenance doesn't fix overnight