

Local Authors: from Medline to Koha

Create list of citations.

Each week, Ovid sends an autoalert of newly-indexed items. Each autoalert has a link allowing you to recreate the search. Click the link, then export the following fields in text format:

```
fa: Authors Full Name
lm: Local Messages
so: Source
ti: Title
ui: Unique Identifier
```

Remember to click the “select all” option; if you don’t, you’ll export only the number of citations on the screen. If there are multiple autoalert, combine all the files into one Notepad file. Save the file as `yyyymmdd_1_citations.txt`, and as `data.txt`.

Sort citations by author.

Open a DOS window (Start | Run | type “cmd” into the run box. Then change to the directory you’re using.). Type “findem > yymmdd_2.txt”. The findem batch file will look for every instance of a name in the file `names.txt`, compare it to every name in the file `data.txt`, and group the names and matches together. The output will look something like this:

```
C:\Documents and Settings\FXKC\Desktop\arthurs>find "Bal," data.txt
----- DATA.TXT
C:\Documents and Settings\FXKC\Desktop\arthurs>find "Bank," data.txt
----- DATA.TXT
C:\Documents and Settings\FXKC\Desktop\arthurs>find "Barac," data.txt
----- DATA.TXT
de Simone, G. Devereux, R B. Roman, M J.
```

Open the file `yyyymmdd_2.txt` in Notepad and use Notepad’s search and replace to remove file path and other extra text. When you’re finished, the file will look something like this:

```
"Bal,"
-----
"Bank,"
-----
"Barac,"
-----
de Simone, G. Devereux, R B. Roman, M J.
```

This makes the file much, much easier to read.

The First Cull

Both the autoalert and the findem program use fairly loose matching criteria to catch the widest possible name variations, so the next step is more labor-intensive. Open the file data.txt in Firefox (Firefox's "find" feature is easier to use for this kind of searching than IE's), open the file yyyyymmdd_2.txt in another window, and open a blank Notepad window in the middle. Name this file yyyyymmdd_3_possible_hits_ui.txt.

Go through each of the names in yyyyymmdd_2 and compare it to your list of names. If there's a possible match, look it up in data.txt to see the full citation. If it still looks like a good match, copy the PMID/UI into the center window of possible hits. Keep going until you're through the entire alphabet. Don't worry about duplicating numbers—if there are any duplicates, they'll be edited out later.

The Second Cull

After you've gone through the list, your possible hits file should have a list of PMIDs considerably smaller than the collection you started out with. Search for these in Ovid using "or." The easiest way is to put them together in one long string (up to a dozen or so at a time) like this:

(21940153 or 23332778 or 23648700).ui.

Then export the citations to a Reference Manager file with the following fields"

as: Abbreviated Source
ab: Abstract
au: Authors
fa: Authors Full Name
do: Digital Object Identifier
il: ISSN Linking
lg: Language
lm: Local Messages
sh: MeSH Subject Headings
nj: NLM Journal Name
pt: Publication Type
so: Source
ti: Title
ui: Unique Identifier

DO NOT export the citations with the boxes "Link to External Resolver" or "Include URL" checked. This will export URLs that don't work. Do, however, make sure you export all the citations. Ovid will name the file refmanager.ovd. Open it in Notepad, save it as yyyyymmdd_4_possible_hits.

Next comes another labor-intensive step. Look up each citation, check the authors' affiliations to see if they are local authors. Most journals will let you look at the authors' affiliations even if you can't view the article.

There are two different author fields in each citation. The AU field is the abbreviated field as it appears in the basic Medline citation. The FA is the full name of the author. If there aren't any local author matches, delete the citation. If you find one or more matches, delete all the FA fields that are NOT local authors. Keep all the AU fields. If the full text of the article is available online, copy the URL and paste it into the citation; add the tag "XL - " in front of it. (It may be easier to add the XL - tag later.) If there's already an XL - tag at the bottom, that means that you accidentally exported the URL from Ovid. Export the list again, but this time uncheck the "Link to External Resolver" and "Include URL" boxes. Save the file as `yyyymmdd_5_hits.txt`.

Replace Text: First Pass

Once you have the list pared down to local authors, it's time to add the department affiliations. Open Replace Text and add the file `yyyymmdd_5_hits.txt` to the Replace Group "Local Authors Step 1 Name Search part 1." Select the Destination tab, uncheck "Save processed file(s) to original location," and change the destination file to `yyyymmdd_6_fa.txt`. Then run the replace function.

Replace Text: Second Pass

The first pass with Replace Text replaced all the FA - tags with FA - NOT FOUND. Now it's time to add the department affiliations. The second pass with Replace Text will replace NOT FOUND with the author's department.

Remove the file `yyyymmdd_5_hits.txt` from the Replace Group "Local Authors Step 1 Name Search part 1" and add `yyyymmdd_6_fa.txt` to the Replace Group "Local Authors Step 2 Name Search part 2." Make the destination file "`yyyymmdd_7_depts.txt`."

Edit the Names and Departments

The accuracy of Replace Text depends on the completeness of the authors in the Name Search part 2 file. Open file `yyyymmdd_7_depts.txt` in Notepad and look for NOT FOUND. You will either need to add a new name to the Replace Text file, or add a variation to an existing name. Edit all the NOT FOUND entries. The good news is that you'll have fewer NOT FOUND names as you build the Replace Text file.

When you have replaced all the NOT FOUND entries with a vocabulary-controlled name and department, save the file as `yyyymmdd_8_refman_upload.txt`

Import the Citations into Reference Manager

You're in the home stretch! Open Reference Manager, then open the Local Authors database. Click File, then Import Text File, then select `yyyymmdd_8_refman_upload.txt`. Import it in the Medline style.

Print the Available Articles

Your new citations will be saved in the tab "Imported," which gives you a chance to do the final editing and polishing if needed. Go through the new citations and look for a URL in the "Related Links" field. If there is one, click on it to ensure that it's the correct link to the article. If it isn't, change it in Reference Manager and in `yyyymmdd_8_refman_upload.txt`. If it is, print the article and write the PMID in the top right corner. Change the Reprint field to "In File."

Convert Reference Manager Tags to MARC Tags

Go back to Replace Text. Remove `yyyymmdd_7_depts` from the "Name Search part 2" group, then add `yyyymmdd_8_refman_upload.txt` to the "Add MARC Tags" group. Make the destination file `yyyymmdd_9_marc.txt`.

Remove Duplicate Departments, add Upload Data

Although Reference Manager doesn't have problems with records with duplicate lines, Koha does. If you upload a file with ten authors, all affiliated with MHRI, the "Departments" field (MARC 656) will have ten instances of MHRI in the OPAC display screen. To prevent this from happening, it is necessary to remove duplicate 656 fields from the MARC records.

[There may be a better way to do this; until then, try this method.]

Open `yyyymmdd_9_marc.txt` in Notepad and in Firefox. Search for "`=656 \\`" in Firefox, then click the Highlight All button. Look through each record for the highlighted tags, then delete the duplicates in Notepad. Be sure to leave one copy of each department.

You also need to add the current date to the 008 field. Replace CHANGE with the date in format `yymmdd`. Save the file (use the same name this time) , then reload it in Firefox and double-check the departments.

Remove the Citation Number Tags

Although the citation numbers in brackets have been handy for keeping the citations separate up to now, they'll choke the system when you compile a file to upload to Koha. Go back to Replace Text, remove `yyyymmdd_8_refman_upload.txt` from the Add MARC Tags group, then add `yyyymmdd_9_marc.txt` to the "To remove the citation number tags" group. Change the destination file to `yyyymmdd.mrk`.

Compile the Citations into an Upload File

Open `yyyymmdd.mrk` in MarcEdit. Does everything look OK? If so, select File | Compile File into MARC. Name the file `yyyymmdd`; MarcEdit will compile into a `.mrc` file ready to stage and import into Koha.