Loading MARCIVE Gov Docs Records

Weekly: Alma automatically retrieves and imports file of new SLS records with profile **MARCIVE SLS import**. (filename prefix: @MINW). No action is necessary as long as weekly loads are successful.

Monthly: MARCIVE provides files of new and updated physical (@MINP) and online (@MINO) records. These must be manually retrieved from website or FTP and processed before loading.

Physical (@MINP) Files

1. Retrieve the file from MARCive's FTP server.
2. Convert the file to UTF8 encoding:
    a. In MarcEdit, choose Tools > Batch Process Records. Then select Character Conversions.
    b. Select source file and destination file.
    c. In Original Encoding, MARC8 should be selected. In Final Encoding, select UTF-8.
    d. Click Process.
3. Add 909 fields to the records:
    a. Double-click the file to open the MARCbreaker tool in MarcEdit.
    b. Click Execute.
    c. Click Edit Records.
    d. Choose Tools > Add/Delete Field. Enter "909" in the Field: text box, and "\\$abcat$bMNU$c[date]$dmarcive" in the Field Data: text box. *[You can also create a Task List in MarcEdit to do this step automatically.]*
    e. Click Add Field.
    f. Choose File > Compile File into MARC. Name the file MINP[date]_909, and Save as type: "UTF-8 MARC File."
    g. Click Save.
    h. Close the file.
4. In Alma, run import profile **MARCIVE overlay records for physical or online titles.** This profile overlays existing records if there is a single match based on an 035 field (could be either a MARCIVE ID or an OCLC number) but does not create inventory.
5. When the import process is finished, find it in Monitor and View Imports. Click Actions > Report.

6. In the report, click the plus sign to open the Records Not Imported section. On line 2, “Records not imported upon no match,” click Actions > Download Sources (binary).
7. Unzip the file and save the MARC file as @MINP[date]_nomatch.
8. In the same section of the report, find line 7, “Records not added (automatic multi-match). Click Actions > Download Sources (binary).
9. Unzip the file and save the MARC file as @MINP[date]_multimatch
10. In Alma, run import on nomatch file with profile **MARCIVE physical full records -- no match in overlay.** This profile will create new bib records without inventory. Records will be suppressed for Gov Pubs staff to process manually.
11. Create an itemized set of the records just added via this process.
12. Notify Clarice that a suppressed file of physical items is ready for her to process.

Online (@MINO) Files

1. Retrieve the file from MARCIVE’s FTP server.
2. Convert the file to UTF8 encoding, following steps 2a through 2d in the “Physical (@MINP) Files” section of this procedures document.
3. Split the file into separate files for each material type using either a Python script OR Extract Selected Records tool in MarcEdit:
    a. To run the Python script “**gpo_fmt_campus_split_909.py**”:
        i. First place a copy of the script file in the same file directory as the MINO file you wish to process.
        ii. With either IDLE or Spyder, launch the script in the console.
        iii. When prompted, enter the filepath of the MINO files to be split:
            1. Shift click on the file in any file explorer window. Select “Copy as Path”.
            2. Paste the path into any text document. Find/Replace any backslashes (“\”) with forward slashes (“/”). Remove any quotation marks around the filepath.
            3. Press the Enter key, and the file should execute.
            Note: *If you get an error at this point* – see the “Troubleshooting the Python script” instructions at the end of this document.
            4. You should now have separate files for each material type, for each campus.
    b. To split the file using MarcEdit:

i. From the MarcEdit main menu, choose Tools > Select MARC Records > Extract Selected Records.
ii. In Source MARC File, browse to select the UTF8 encoded version of the @MINO file.
iii. In Display Field, type LDR. Then click Import File.
iv. Check “Use Regular Expressions.”
v. In Search, type **^.{6}e.** Then click the magnifying glass icon.
vi. When the search completes, click Export Selected. When prompted “Would you like to delete records from the source file?” click No. Supply a filename for the extracted records (@MINO[date]_map).
vii. In the Extract Selected Records window, change Search to **^.{6}k.** Then click the magnifying glass icon.
viii. When the search completes, click Export Selected. When prompted “Would you like to delete records from the source file?” click No. Supply a filename for the extracted records (@MINO[date]_vis).
ix. In the Extract Selected Records window, change Search to **^.{6}as|ai.** Then click the magnifying glass icon.
x. When the search completes, click Export Selected. When prompted “Would you like to delete records from the source file?” click No. Supply a filename for the extracted records (@MINO[date]_ser).
xi. In the Extract Selected Records window, change Search to **^.{6}am|cm.** Then click the magnifying glass icon.
xii. When the search completes, click Export Selected. When prompted “Would you like to delete records from the source file?” click No. Supply a filename for the extracted records (@MINO[date]_bks).

4. In Alma, open the **MARCIVE TEST Twin Cities online record import and overlay** profile for editing. Double-check that on the Inventory Information tab, the Material type is set to “Book.”
5. In Alma, run import on @MINO[date]_bks file with the **MARCIVE TEST Twin Cities online record import and overlay** profile. This profile overlays existing records if there is a single match based on an 035 field (could be either a MARCIVE ID or an OCLC number). If there is no match, it will import new (nomatch) records. In either case, it will process inventory by either creating new portfolios or updating old portfolios in the collection, setting “Books” as the e-material type.
6. When the import job completes, find it in Monitor and View Imports. Click Actions > Report.

7. In the report, click the plus sign to open the Records Not Imported section. On line 7, “Records not added (automatic multi-match)”, click Actions > Download Sources (binary).
8. Unzip the file and save the MARC file as @MINO[date]_bks_multimatch.
9. In Alma, open the **MARCIVE TEST Twin Cities online record import and overlay** profile for editing. On the Inventory Information tab, change Material type to “Map,” then repeat steps 5-8, using appropriate filenames in steps 7 and 8.
10. Repeat step 9 for serials (changing material type to “Journal”) and then for visual materials (changing material type to “Other Visual Material”).
11. Identify any remaining material types and load the records:
    a. Open the MINO[date]_other file output by the “gpo_fmt_campus_split_909.py” script.
    b. Determine an appropriate Alma material type for the items per the “Material Types in Alma Item Records, Portfolio Records, and Purchase Order Lines” policy. If there is more than one material type applicable to records in the file, split the file by material type for upload.
    c. Repeat step 9 to import the record batch(es).
12. When the last file has been run through **MARCIVE TEST Twin Cities online record import and overlay,** open the profile for editing and change material type back to “Book.”

Loading multi-match files

1. Place all multi-match files (from both MINP and MINO sets) in a folder.
2. In MARCEdit, choose Tools > MARCJoin. In Destination, provide a path and filename for the joined multimatch file (@MIN[date]_multimatch_all).
3. In File(s) to: browse and select all of the @MINP[date]_multimatch and @MINO[date]_[type]_multimatch files.
4. Click Process.
5. In Alma, run import on @MIN[date]_multimatch_all file with profile **MARCIVE multi-match import**. This sets up multimatch records for manual handling.
6. Email Clarice Ostman when the multi-match import job is running.
7. Clarice will notify when multimatch cleanup is complete so the next batch can be run.

Loading online records for Duluth and Morris

Note: Do not begin this process until steps above are complete.

1. Identify the records to be loaded for the coordinate campuses:
    a. If you used the Python script:
        i. Duluth and Morris records can be identified as files output by the "gpo_fmt_campus_split_909.py" script with either "_MNGE" (for Duluth) or "_MNXN" (for Morris) at the end of the filename.
        ii. Any file which appears as having size of 0 KB in the file explorer window is likely empty (no records met the criteria). These may be deleted (double-check they're empty if you like.)
        iii. Identify which material types you will need to load for each coordinate campus.
    b. If you used MarcEdit:
        i. Parse each @MINO material-type-specific file to extract one file per campus per material type:
            1. From the MarcEdit main menu, choose Tools > Select MARC Records > Extract Selected Records.
                a. In Source MARC File, browse to select the **@MINO[date]_bks** file.
            2. In Display Field, type 049. Then click Import File.
            3. In Search, type **MNGE.** Then click the magnifying glass icon.
            4. When the search completes, click Export Selected. When prompted "Would you like to delete records from the source file?" click No. Supply a filename for the extracted records (@MINO[date]_bks_duluth).
            5. In the Extract Selected Records window, change Search to **MNXN.** Then click the magnifying glass icon.
            6. When the search completes, click Export Selected. When prompted "Would you like to delete records from the source file?" click No. Supply a filename for the extracted records (@MINO[date]_bks_morris).
            7. Repeat steps 1-6. for each of the other format-specific files. Note that for some formats, either Duluth or Morris may not have any selections. If the search finds no records with either campus's 049 code, there's no need to extract a file for that format for that campus.
2. In Alma, open the **MARCIVE TEST Duluth online record import and overlay** profile for editing. Double-check that on the Inventory Information tab, the Material type is set to "Book," "Map," "Journal," "Other visual material," etc. as necessary for the material type of the file you are importing.

3. In Alma, run import on @MINO[date]_[material type]_MNGE file with the **MARCIVE TEST Duluth online record import and overlay** profile. This will update or create portfolios on the already-loaded records with the Duluth group setting and whatever was selected in step 2 as the e-material type.
4. When the import job completes, find it in Monitor and View Imports. Click Actions > Report. If any records were not added because they were multi-matches, click Actions > Download Sources (binary). Unzip the file and save as @MINO[date]_[material type]_MNGE_multimatch.
5. Repeat steps 2 through 4 for each material type represented in the files ("Map" for maps, "Journal" for serials, "Other Visual Material" for visual material, other material types available per policy and presence of records in any "other" files).
6. When the last file has been run through **MARCIVE TEST Duluth online record import and overlay**, open the profile for editing and change material type back to "Book."
7. Repeat steps 2-6 using the **@MINO[date]_[format]_MNXN** files with the **MARCIVE TEST Morris online record import** profile. This will update or create portfolios with the Morris group setting.
8. If any multi-matches turned up in step 4, follow the instructions in the "Loading multi-match files" section of this document to process them.

Troubleshooting the Python script

1. If you get an error when you are running the "gpo_fmt_campus_split_909.py" script, first make sure you are running Python verson 3.0 or greater.
2. Next, use the "id_marc_problem.py" script to identify the fail point in the record batch:
    a. Launch the script in either IDLE or Spyder console.
    b. Enter the filepath of your record batch as described in step 3c in the "online (@MINO) Files" section of this procedures document.
    c. Look at the output file the script generates. The last record listed in this output file is the record *before* the record with the error in the original MINO file (likely a character set problem).
    d. Use MarcEdit to open the MINO file (you will need to convert it to .mrk as described in steps 3a through 3c in the "Physical (@MINP) Files" section of this procedures document.)
    e. In MarcEdit, press CTRL + f. Enter the OCLC number of the last record in the output file. Scroll down in the file to examine the next record. This is the [first instance of a] problem record. If you can find a visible character

error, correct it and re-run the “gpo_fmt_campus_split_909.py” script. If not, remove the record from the batch and re-run the “gpo_fmt_campus_split_909.py” script, then manually process the problem record.

f. Repeat these steps as many times as necessary to identify any/all problem records in the set and remediate or remove them.

Version history:

Revised: 9/23/2016 by Stacie Traill

Revised: 2/9/2017 by Kelly Thompson in consultation with Stacie Traill.