



University of Massachusetts Amherst

From the Selected Works of Jennifer Eustis

May 13, 2022

MarcEdit 101 An Introduction.pdf

Jennifer Eustis, *University of Massachusetts Amherst*



This work is licensed under a [Creative Commons CC_BY-NC-SA International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).



Available at: <https://works.bepress.com/jennifer-eustis/16/>



MARCEDIT 101

Jennifer Eustis
University of Massachusetts Amherst
May 13, 2022

WHO ARE YOU?

HAVE YOU USED MARCEDIT BEFORE?

WHAT IS ONE THING ABOUT
MARCEDIT YOU WANT TO LEARN
ABOUT TODAY?

TODAY'S TOPICS

We are going to see the basics of MarcEdit, a tool created and maintained by Terry Reese. This includes:

- Brief introduction to MarcEdit
- Understanding settings, preferences, and main level menu
- Working in the MarcEditor and using its features to manipulate MARC data
- Working with tasks lists to create Tasks
- Understanding integrations and regular expressions

By the end of this, I hope that you will come away with:

- Understand how to download MarcEdit and where to find help
- How to update your preferences in Settings
- Use the basic top level features of the Main Menu
- Use the MarcEditor to manipulate and profile MARC records and metadata
- Learn about more advanced features

WHAT IS MARCEDIT?

MarcEdit is a free suite of tools created and maintained by Terry Reese. These tools are meant to be used with library data, in particular MARC data. As we know, library data comes in different formats and has varying levels of quality. MarcEdit allows you to make edits to one or more records, automate common practices, convert data, or enhance records.

FEATURES & DOWNLOADS

List of all MarcEdit's features

Download instructions for PC and MAC

Sign up for the MarcEdit listserv

Consult the MarcEdit knowledge base

Check out video tutorials

This presentation uses MarcEdit version 7
for PC (single user installation)

THE SETUP WIZARD

When you download and open MarcEdit for the first time, you will be prompted to go through the Setup Wizard. Don't skip this step as it could cause issues later on.

Fill in the MARC format, for the US this will be MARC21.

The default record encoding will be usually UTF8. If you work with primarily MARC8 encoded records, don't worry. You can always update this in your preferences.

You can leave the default for XML processing engine.

MarcEdit 7 Startup Wizard: Page 3

MarcEdit Data Settings

MarcEdit is MARC agnostic, so let's tell it a little about your data.

My MARC Format is:
MARC21

Default Record Encoding:
MARC8/UTF8

☒ Enforce Unicode Normalization

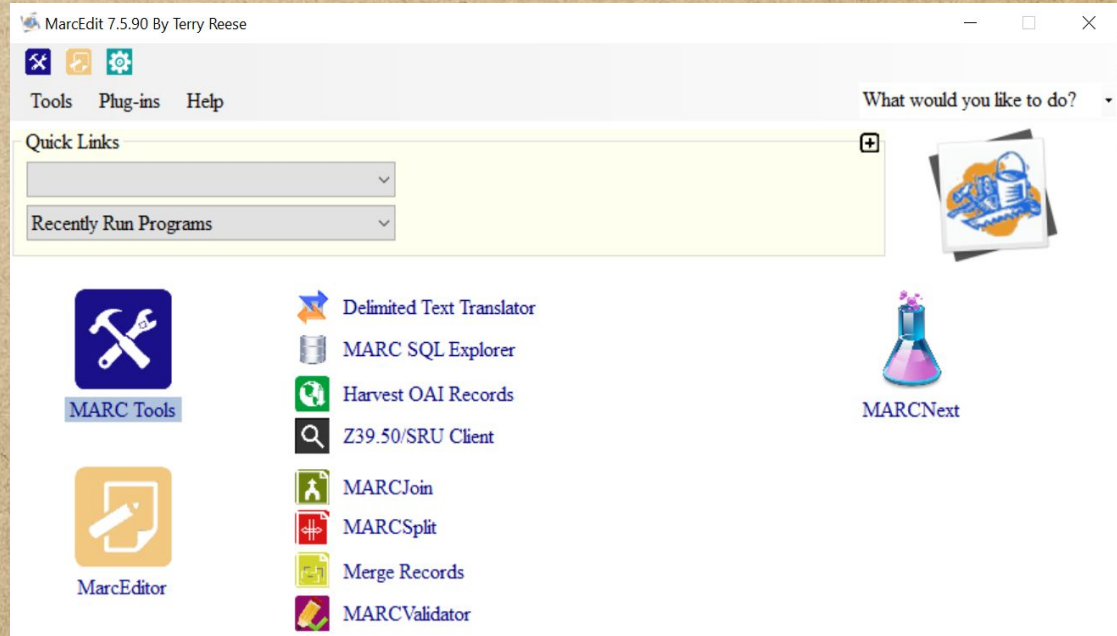
XML Processing Engine:
Choose For Me

We're Almost Done! Squirrel

MAIN MENU & FEATURES

When you open MarcEdit, you reach the main level menu or the main window. One thing about MarcEdit is that often there is more than one place to get to a feature. Notice that here MARCJoin is in my shortcuts. You can also access that feature by going to Tools and then MARC Processing tools.

Plugins are added features. This presentation doesn't cover this. More information can be found on Terry Reese's [page on plugins](#).





MARCEDIT PREFERENCES: MAIN WINDOW SETTINGS

To access your preferences, you can click on the gearbox or go to Tools/Preferences (Ctrl+Shift+P)

In the main window settings, you can:

Select Up to 4 items default programs that display on the main window

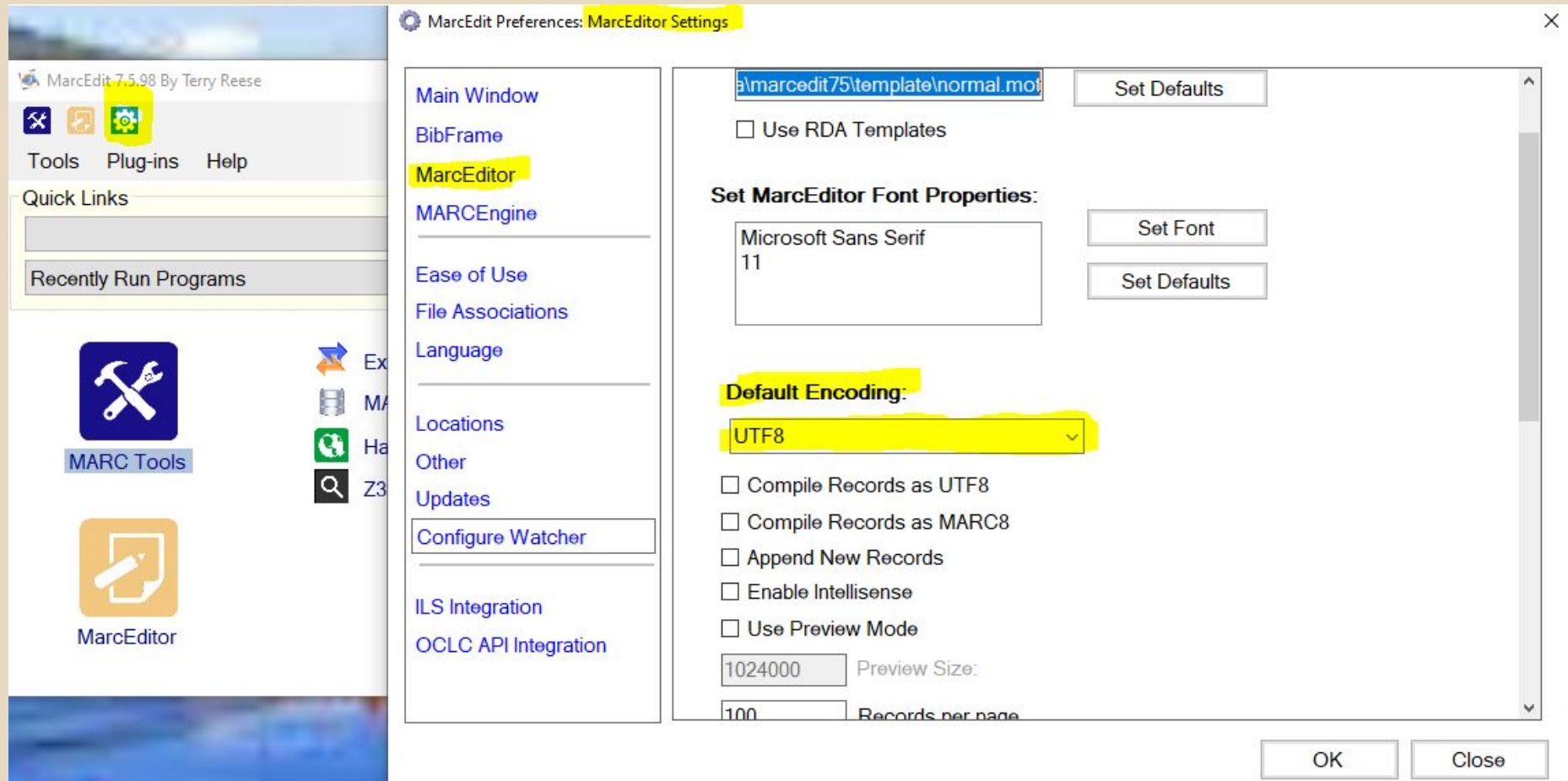
Select your default character encoding for the MarcEditor

Set up integrations to an ILS or OCLC

Set a default font and size for your records in the MarcEditor

HOW DO YOU CHANGE YOUR
DEFAULT ENCODING?

IF YOU HAVE A FILE WITH ANOTHER
CHARACTER ENCODING, DO YOU NEED TO
UPDATE THE DEFAULT ENCODING AND
THEN CHANGE IT BACK?



MAIN WINDOW FEATURES

From the main window, you can access:

MarcEdit Tools

MarcEditor

MARC Processing Tools (Joining or Splitting Marc files for instance)

Character Conversion Tools

Cuttering Tools

OAI-PMH Tools

Export

Generate Reports (field count, count by material type)

Verify URLs

Regular Expression Store

OCLC API Operations

MAIN WINDOW: MARCEDIT TOOLS



From the main window, you can select:

- **MarcBreaker:** This "breaks" the MARC binary file into a readable format that can be edited in the MarcEditor.
- **MarcMaker:** This takes the readable format MARC data file and creates the MARC binary file.
- **MARC21 to MARC21XML:** This converts a MARC21 file to MARC21XML.
- **MARC21XML to MARC21:** This converts a MARC21XML file to a MARC21 binary file.
- **MARC to JSON:** This converts a MARC21 file to a JSON file.
- **JSON to MARC:** This converts a JSON file to a MARC21 file.
- **JSON to XML:** This converts a JSON file to XML.
- **XML to JSON:** This converts an XML file to JSON. Options include changing the default character encoding only for the tool that you select, translating the MARC data to MARC-8 or UTF8. To change the default character encoding, you will need to go to Settings.



1.

USING THE MARCEDITOR

*Breaking MARC (mrc) to work
with MARC (mrk)*

03926cam a2200601Ki
 450000100100000000300060001000500170001600600190003300700
 150028504300120040005000220041207200250043407200230045908
 700190077949000290079850400570082758800260088450507160091
 065000600270465000540276465100380281864800200285665500220
 106.1RSm o d RScru cnu---unuuuRS151104s1976
 USa1400878764USq(electronic bk.)RS USa(OCOLC)ocn9274428
 USa22573/ctt182d5zdUSbJSTORRS
 USAN\$TUSbengUSerdaUSepnUScN\$TUSdIDEBKUSdYDXCPUSdJSTORUS
 USaa-tu---RS 4USaDR569USb.D3 1976ebRS 7USaHISUSx026000U
 H.RS10USaReform in the Ottoman Empire, 1856-1876 /UScby I
 online resource (xviii, 479 pages)RS USatextUSbtxtUS2rd
 fileUS2rdaRS1 USaPrinceton Legacy LibraryRS USaIncludes
 --UStPreface --UStNote on Transliteration, Names, Dates,
 Hatt-ı Hümayun of 1856 and the Climate of Its Reception -
 Reorganization of the Non-Muslim Millets, 1860-1865 --US
 Political Agitation: the New Ottomans --USVII. The Final
 1876-the Year of the Three Sultans --USIX. The Constituti
 examines in detail the Tanzimat reforms, focusing on the
 published in 1963.The Princeton Legacy Library uses the l
 the distinguished backlist of Princeton University Press.
 durable paperback and hardcover editions. The goal of the
 the thousands of books published by Princeton University
 USaHampshire College patrons have access.US5MAHRS1 USaMo
 access.US5MNSRS1 USaUMass Amherst patrons have access.US
 governmentUSy1829-1878.US0http://id.loc.gov/authorities/s
 EastUSxTurkey & Ottoman Empire.US2bisacshRS 7USaPolitics
 7USa1829-1878US2fastRS 4USaElectronic books.RS08USiPrint
 0USaPrinceton legacy library.US0http://id.loc.gov/authori
 1976ebUSkE-BookRS40USuhttp://5colauthen.library.umass.ed
 resourceRS USa947725502USa959919012USa992905657USa99935

MarcEditor: exampleOfMrk.mrk

File Edit Fonts Reports Tools OCLC WorldCat Plug-ins

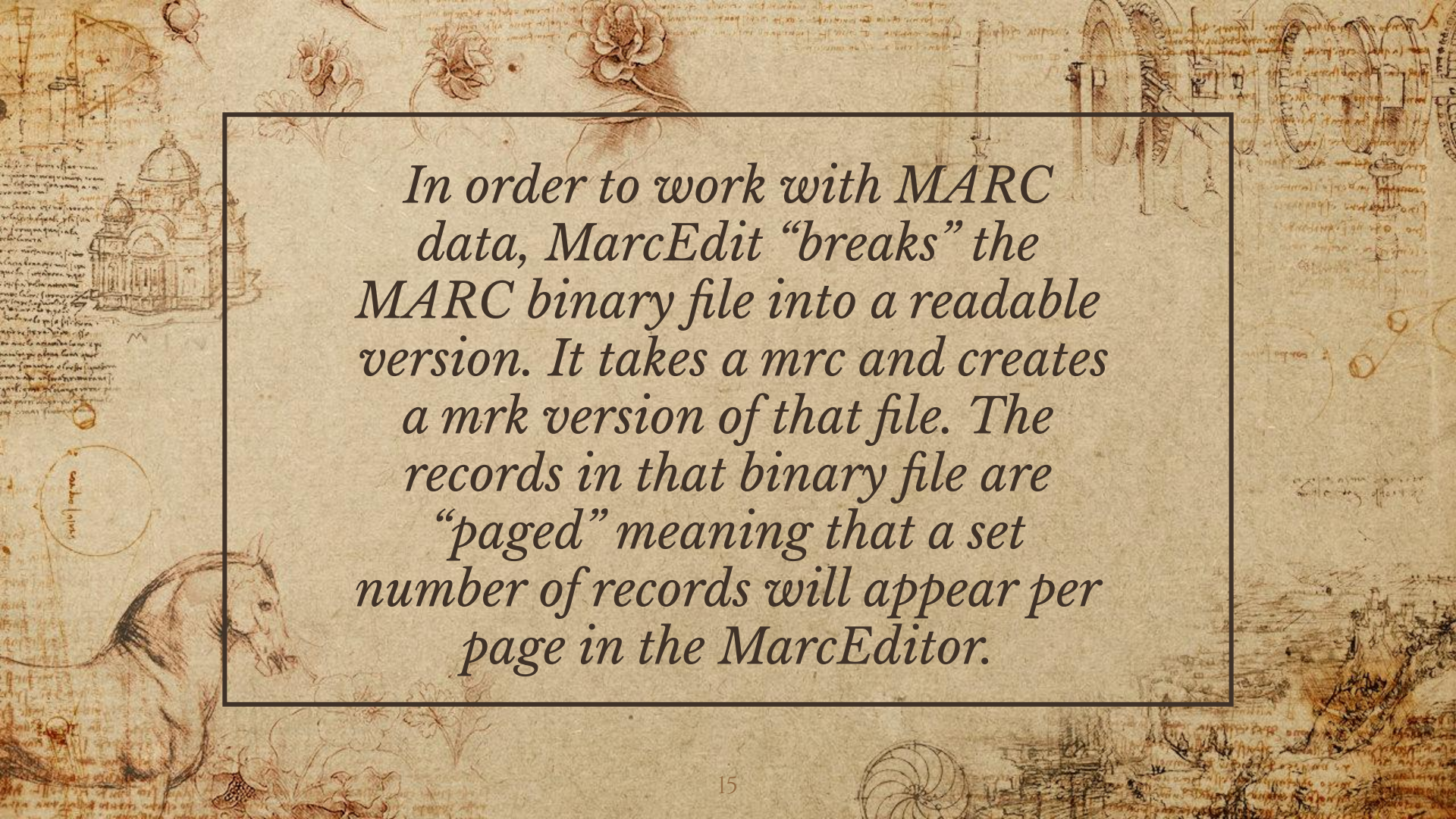


```

=LDR 03922cam a2200601Ki 4500
=001 014978780
=003 Aleph
=005 20200725013106.1
=006 m|||||o||d|||||
=007 cr\cnu---unuuu
=008 151104s1976|||||nju|||||ob|||||000\0\eng\ld
=020 \\\$a9781400878765$q(electronic bk.)
=020 \\\$a1400878764$q(electronic bk.)
=035 \\\$a(OCOLC)ocn927442845
=035 9\|a(FCJSP)10.2307/j.ctt183q0cn
=035 9\|a(JSTOR)ocn927442845
=037 \\\$a22573/ctt182d5zd$bJSTOR
=040 \\\$aN{$}T$beng$erda$epn$cN{$}T$dIDEBK$dYDXC
=043 \\\$aa-tu---
=050 \\\$aDR569$b.D3 1976eb
=072 \\\$aHIS$x026000$2bisacsh
=072 \\\$aHIS055000$2bisacsh
=082 04$a956.101$222
=100 1\|aDavison, Roderic H.
=245 10$aReform in the Ottoman Empire, 1856-1876 /$cby Roder
=264 1\|aPrinceton, N.J. :$bPrinceton University Press,$c1963.
=300 \\\$a1 online resource (xviii, 479 pages)
=336 \\\$atext$btxt$2rdacontent
=337 \\\$acomputer$bcb$2rdamedia

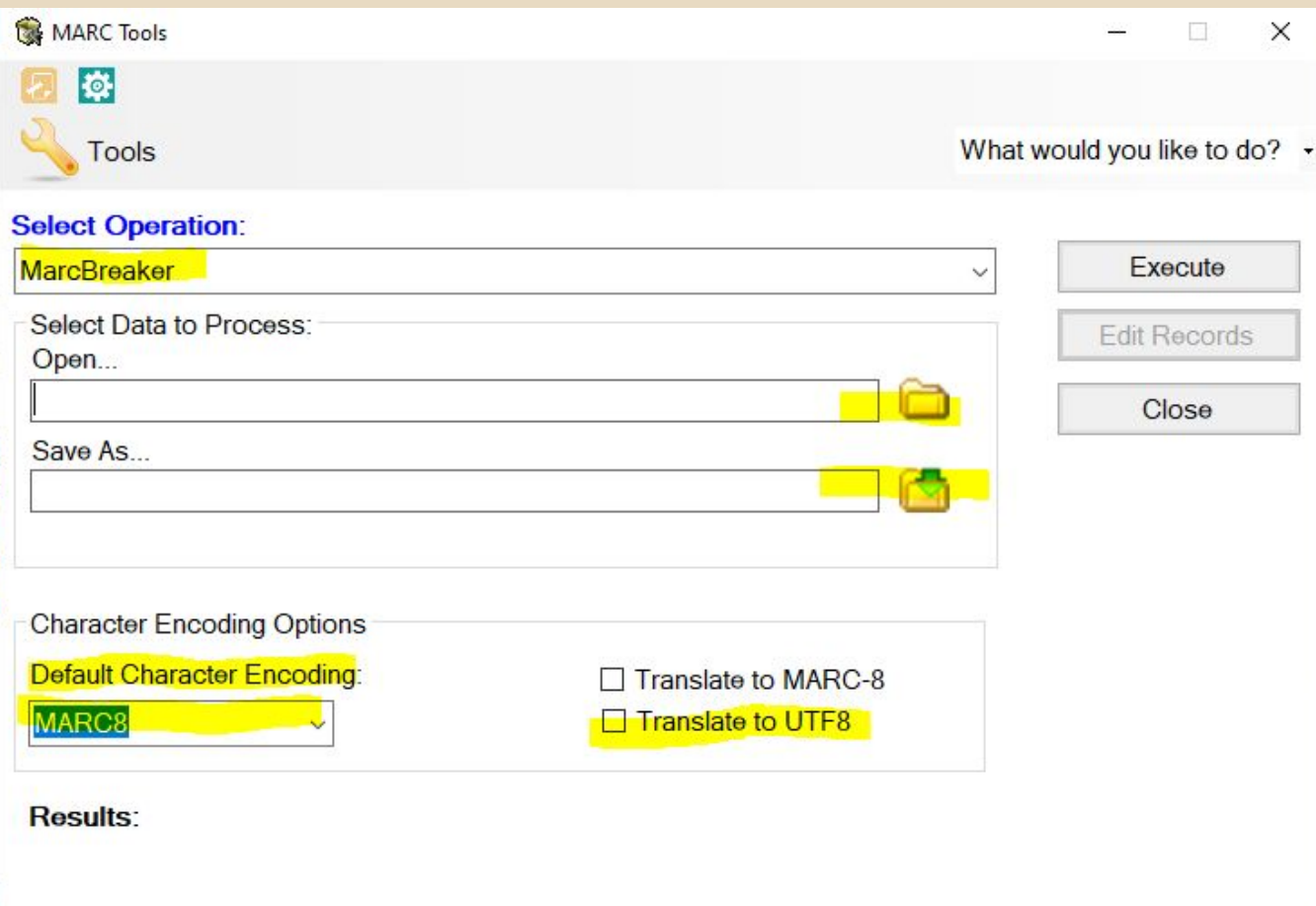
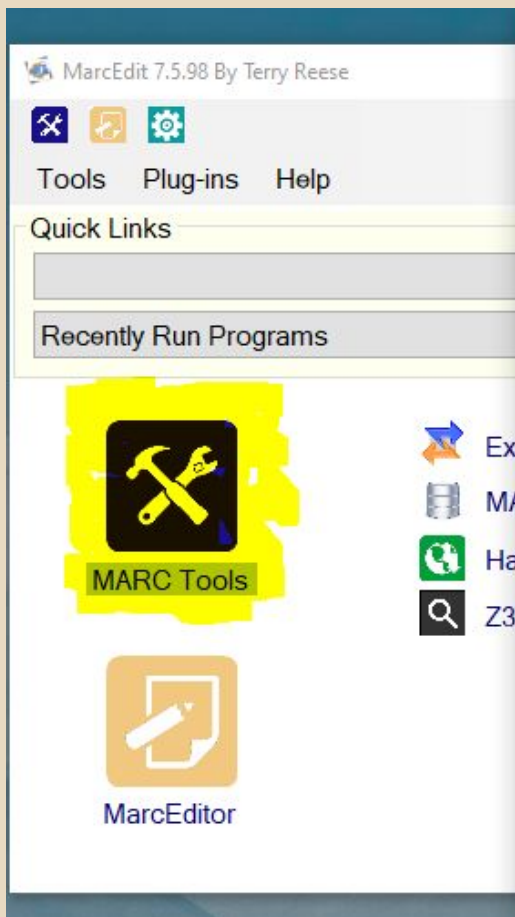
```

1 records processed in 0.147333 seconds.



In order to work with MARC data, MarcEdit “breaks” the MARC binary file into a readable version. It takes a mrc and creates a mrk version of that file. The records in that binary file are “paged” meaning that a set number of records will appear per page in the MarcEditor.

HOW DO YOU “BREAK” A MARC BINARY
FILE TO WORK WITH IT IN THE
MARCEDITOR?





PROFILING

MarcEdit Features that help to Profile MARC data

UNDERSTANDING YOUR DATA BEFORE CHANGING IT



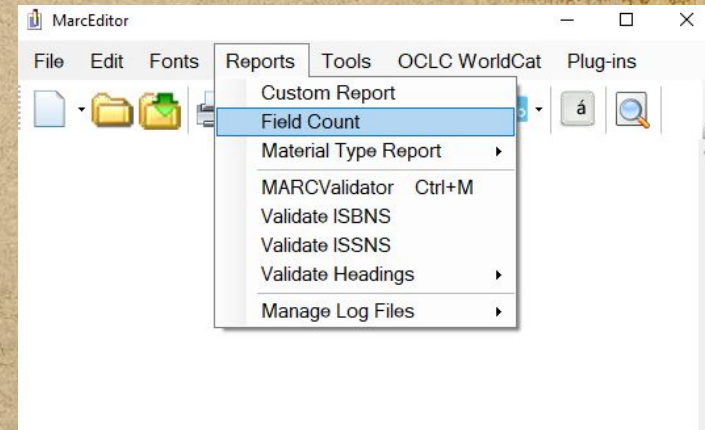
MARC Reports

- ✧ Field Count
- ✧ Material Type Count
- ✧ Custom



MARC Validation

- ✧ Validate by rules
- ✧ Validate by record structure



CAN YOUR MARC FILE HAVE A VALID
RECORD STRUCTURE AND NOT FOLLOW
THE RULES?

A MARC record is composed of three elements: **the record structure, the content designation, and the data content of the record.** The MARC Validator can check the record structure or the content designation here referred to as rules. Record structure consists of having the minimum number of fields in a MARC record and respecting the format of how to construct fields, subfields, fixed fields, or the leader. For more information on record structure go to <https://www.loc.gov/marc/specifications/specrecstruc.html>. The content designation or rules are outlined in a file that is downloaded when you install MarcEdit. For more information, you can go to <https://github.com/cmharlow/ShareFest15MetadataQA/blob/master/examples/MARCEdit/MARCvalidation.md>.

MARC VALIDATOR

MARCValidator

Source File:

Rules File:

C:\Users\Jennifer\AppData\Roaming\marcedit7\configs\marc

[Edit the Rules File](#)

Options

- ☒ Check MARC Using Rules File
 - ☐ Group results by error
- ☐ Validate Record Structure
- ☐ Remove invalid records
- ☐ View in a HexEditor

OK

Close




MANIPULATING

Most Common MARC Manipulation Features in the MarcEditor

BASIC MANIPULATION FEATURES

- ✦ MARC Fields: Add, Delete, Copy, Build, Find/Replace
- ✦ MARC Subfields: Add, Delete, Find/Replace
- ✦ MARC Indicators: Change, Find/Replace
- ✦ Special Undo

Tools	OCLC WorldCat	Plug-ins
Call Number Tools		▶
Clustering Tools		Ctrl+Shift+T
Generate Control Numbers		
Linked Data Tools		▶
Record Deduplication		
RDA Helper		
Transliterate		▶
 Z39.50/SRU Options		▶
Regular Expression Store		
Edit Constant Data		
Assigned Constant Data		▶
Manage Tasks		
Assigned Tasks		▶
Validate MARC Records		
Harvest from OAI		
Generate MARC from URL		Ctrl+Shift+U
Mnemonic Formatting Tools		▶
Add/Delete Field		F7
Build New Field		Ctrl+Shift+F1
Copy Field		Ctrl+Shift+F2
Edit Field Data		Ctrl+Shift+F3
Edit Indicator Data		F8
Edit Subfield Data		F9
Sort by...		▶
Swap Field Data		F11

HOW DO YOU DELETE ALL 900 AND ANY
VARIATIONS IN YOUR MARC FILE?

Add/Delete Field

Copy Field Data

Edit Fields

Edit Indicators

Edit Subfield Data

Swap Fields

Add/Delete Field Utility

Field:

9XX|

Field Data:

Find What:

[Copy Add Task](#)

[Copy Delete Task](#)

Delete Field



Add Field



Close

General Options

- ☒ Match case
- ☐ Use Regular Expression
- ☐ Process batch operation

Add Field Options

- ☒ Insert before
- ☐ Insert last
- ☐ Add field only if not present
- ☐ Add field only if not a duplicate
- ☐ Add Field If Present/NOT Present

Delete Field Options

- ☐ Remove Duplicate Data
- ☐ Remove by field position
- ☐ Remove if field data does not match
- ☐ Remove if invalid UTF-8 characters

MORE MANIPULATION FEATURES

- ❖ Build a Field
- ❖ RDA Helper
- ❖ Edit Shortcuts
 - ✖ Change case
 - ✖ Field Edits
 - Find missing or duplicate MARC fields
 - Find missing words in a MARC field
- ❖ Sort By
- ❖ Select Records for Edit

CAN YOU ADD A GMD INTO YOUR
RECORDS?



- ☒ Add 100 See Relator Field
- ☒ Add 336 -- Content Type
- ☒ Add 337 -- Media Type
- ☒ Add 338 -- Carrier Type
- ☐ Add 344 -- Sound Characteristics
- ☐ Add 345 -- Proj. Characteristics
- ☐ Add 346 -- Video Characteristics
- ☐ Add 347 -- Digital Characteristics
- ☒ Add 380 -- Form of Work
- ☒ Add 381 -- Other
- ☒ Evaluate 260/264
 - ☒ Always use ©, ® symbols

☐ Qualifying Information (015/020/024/027)

☒ Modify 040 -- add Serda

\$b:

☐ Process the 502

☒ Delete GMD

☐ Generate GMD

☒ Expand Abbreviations

[Edit List](#)

| [Edit Fields](#)

Select: [All](#) | [None](#)



[Copy Task](#)

OK

Close



TASKS

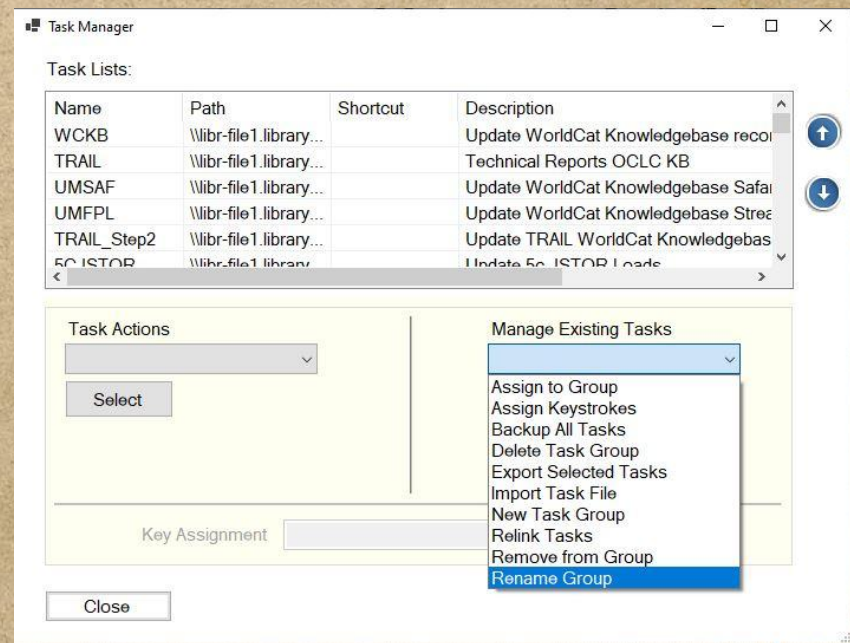
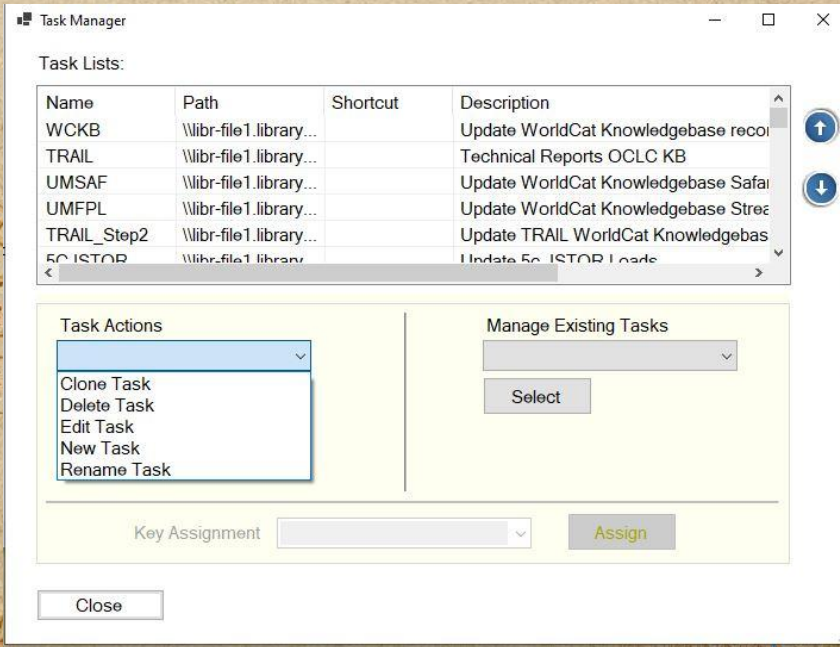
Automating common practices

WHAT IS A TASK?

So far we've been looking at individual tasks. We've changed fields, removed fields, etc. These tasks can be added to a list and run. The tasks in the list are run in sequential order.

A list of tasks is called a Task. You can create and manage tasks.


WORKING WITH TASKS




CREATE A NEW TASK

- ✦ Go to Tools and Manage Tasks
- ✦ In Task Actions, Select New Task
- ✦ Enter a name for your Task
 - ✦ The only way to rename a task is through the Task Actions
- ✦ Once you click save, you are brought to the Edit Task List Editor
 - ✦ Add a description
 - ✦ Add tasks using the icons to add, delete, copy, edit

TASK LIST EDITOR

 Edit Task List

Task List Name:



Description:





☐ Override Broker Action (not recommended)

Tasks:

↑

↓

Actions:

Save

Debug

Close



INTEGRATIONS

INTEGRATIONS WITH ILS AND OCLC

MarcEdit can talk to other systems such as your ILS or OCLC. Current supported ILS are Koha and Alma. OCLC integration can occur via Z39.50 or the OCLC Search and Metadata API.

For the OCLC API, you will need to request an API key from OCLC.

WHERE TO FIND HELP

- ❖ How to set up MarcEdit OCLC Integration
- ❖ Getting started with OCLC's Z39.50 server
- ❖ Getting started with MarcEdit Alma Integration

OCLC Z39.50 SETUP

Database Properties

Z39.50 Database Settings:

Database Properties **Advanced Properties**

Name:
OCCL Cataloging Service

Host:
zcat.oclc.org

Database:
OLUCWorldCat

Port:
210

Syntax:
MARC21

Username:

Password:

☐ Override pause (6 secs.)

☒ Show Database

Save Delete Close

- ❖ Username/Password: Your OCLC credentials used to log into Connexion client or browser
- ❖ Advanced Properties allow you to set search parameters for authors, titles, subjects, isbn/issns.
- ❖ Options allow to set the character encoding, to retrieve unicode, and how many records are retrieved per page.



REGULAR EXPRESSIONS

WHAT ARE REGULAR EXPRESSIONS?

Regular expressions are a sequence of characters that define a specific pattern in text or a string of characters. These patterns are used often to find an replace text.

For example, an email is a sequence of characters with a pattern, namely characters preceding and following the @ symbol. Another example are U.S. phone numbers which have 3 digits, hyphen, 3 digits, hyphen, 4 digits.

Search



- ☒ Search Metadata
- ☒ Search Public Store

Results



- 0: Recursive Regular Expression Example
- 1: Using negative assertions within regular ex
- 2: Find Single Word 510a**
- 3: Move Subfield Example (not necessarily F
- 4: Find records with multiple 856\$3 that also
- 5: Separate 041 language codes into own su
- 6: If 504 index statement exists, edit 008 to 1
- 7: Recursive Regular Expression Example

Public ID:

Expression Name

Expression Description

This is an example of how to format a regular expression to use in the Field Data box of an Edit Subfield task to find specific single-word citation forms in the 510 subfield \$a.

In this expression "a" is the subfield tag (this must be included when using a regular expression with Edit Subfield), (?i) makes the rest of the match case-insensitive, "Bristol" is the sample citation data, and \$ requires the match to occur at the end of the subfield. This expression will find "Bristol" or

Expression

☐ Share Expression

Actions



Delete

Close

Close Window

MORE RESOURCES

- ❖ LinkedIn Learning (formerly Lynda.com)
- ❖ Library Carpentry: Introduction to Working with Data (Regular Expressions)
- ❖ University of Illinois Libraries MarcEdit LibGuide
- ❖ University of Illinois Libraries PDF Guide by Patrick Harrington
- ❖ RegexOne: Learn Regular Expressions

THANKS!

Any questions?

jeustis@umass.edu