Editing Records with the MarcEditor

TERRY REESE
HEAD OF DIGITAL INITIATIVES
THE OHIO STATE UNIVERSITY
REESE.2179@OSU.EDU
Let’s work through these examples together
Keypoints

MarcEditor
- What is it?
- What do the properties mean
- Preview mode?
- Paging mode?

Editing Functions

Field Count

Task Automation

Validation

OAI Harvesting
Editing MARC

MarcEditor

- Specialized TextPad designed specifically for MARC records.

- Is UTF8 aware – can be used to generate records in MARC8 (though mnemonics) or UTF8 charactersets.
MarcEditor Properties

Templates
Fonts
Encodings
Preview Settings
MarcEdit Templates

Templates work much like Microsoft Word Templates
- Define a set of default data that will appear on a screen
- Templates exist for all material formats
- Can be customized to suit your needs.
MarcEdit’s Preview Mode

One of the most confusing features

- Allows MarcEdit’s MarcEditor to address files over the allowed 2 GB Windows page file limit (though practical limits are closer to 300 MB)
- Reads a small snippet of the file into the editor – but edits are done to the entire file.
- Can be turned off.
Configuring Preview Mode

MarcEdit Preferences: MarcEditor Settings

- **Set MarcEditor Template:**
  - Current Template: C:\nei_marcedit\GRMProgram1
  - Use RDA Templates

- **Set MarcEditor Font Properties:**
  - Arial Unicode MS 10

- **Default Encoding:**
  - UTF8

- **Compile:**
  - Records as UTF8
  - Records as MARC8

- **Append:**
  - New Records

- **Enable Intellisense**

Controls:
- OK
- Close
- Apply
MarcEdit Preview Mode

```
LDR 00634nm a2200145Ka 4500
=042 $0
dc
=245 00$aThe Boy Chums in the Forest$[electronic resource]
=260 $aSalt Lake City :$bProject Gutenberg Literary Archive Foundation, $c2008.
=500 $aFrom "or Hunting for Plume Birds in the Florida Everglades".
=500 $aRecords generated from Project Gutenberg RDF data.
=500 $aISO 639-2 language code: en
=640 $aApplicable license: http://www.gutenberg.org/license
=690 $aWorld politics -- Handbooks, manuals, etc.
=720 $aEly, Wilmer M. (Wilmer Mateo)$bauthor
=856 4$uhttp://www.gutenberg.org/etext/27561

LDR 00741nm a2200169Ka 4500
=042 $0
dc
=245 00$aThe 2006 CIA World Factbook$[electronic resource]
=260 $aSalt Lake City :$bProject Gutenberg Literary Archive Foundation, $c2008.
=500 $aRecords generated from Project Gutenberg RDF data.
=500 $aISO 639-2 language code: en
=640 $aApplicable license: http://www.gutenberg.org/license
=690 $aWorld geography -- Handbooks, manuals, etc.
=690 $aWorld politics -- Handbooks, manuals, etc.
```
Configuring New Paging

Set in the Options dialog
MarcEdit Paging

Paging Change Notes

◦ The preview page functionality is still present, but full page now defaults to the new paging functionality.

◦ Preview functionality – on load – the application reads the entire file to prep – this is where most of the loading time takes place. After that, pages are addressed directly.
Paging Example

Preview Mode Still Exists
Paging Example

If you load the full file, or turn the preview mode off
So when would you use paging versus preview?

Default and most common setting should be the paging mode.

Preview mode should likely be used when:

1. Using the MarcEditor to edit files over 300 MB
   - Reduces load times
   - Allows MarcEditor to process data more quickly (no need to create file page map)
Editing MARC

MarcEditor

- Supports a number of global editing functions:
  - Find/Replace functionality
  - Globally Add/Delete MARC fields
  - Globally Edit Subfield data
    - Conditionally add/remove field data
  - Globally Edit Indicator data
  - Globally Swap field data
  - Record Deduplication
  - Record Sorting
  - Call Number Generator
  - Macros
  - Z39.50 Cataloging
MarcEdit Regular Expression Support

When processing regular expressions with MarcEdit, MarcEdit makes entire fields or subfields available for processing

- i.e., when processing a delete field function – all data from =[field number] are part of the field that can be queried.

MarcEdit’s regular expression by default deals with one field at a time (i.e., regular expressions do not allow you to find data across fields by default)

MarcEdit’s Regular Expression Support Pre-5.x was a custom regular expression engine.

MarcEdit’s Regular Expression Support 5.x+ is defined by Microsoft .NET’s Regular Expression object

- This object uses a syntax that looks Perl-like, but has some differences.
MarcEdit Regular Expression Support

When working with regular expressions with the Replace Function, MarcEdit will remember the last 10 replacements. This should help with trial and error.

When dealing with Regular Expressions or any global replacements, MarcEdit has a Special Undo function that will undo your last global update.
Editing Subsets of Data

- Within the MarcEditor, you can extract subsets of data for processing
Edit Shortcuts

- Edit shortcuts wrap common regular expressions for ease of use
Editing MARC – Find/Replace

Works like a normal Find/Replace in most Textpad utilities.

Unlike most Textpads, Replace supports UTF-8 (when working with UTF-8 files) and regular expressions.
Conditional Replacements
Find all function was designed for use with the Paging mode

Allows users to find any text across all pages

Generates a jump list that can be used to find individual records for edit
Miscellaneous Tools

Generate Cutters (sandborn and other)
Generate Record from URL (one a website, one from LC)
Editing MARC – Global Add/Delete Field

Globally add fields to all MARC records
- Allows users to set insertion position.

Globally delete fields
- Allows global delete
- Allows conditional delete

Supports Regular Expressions
Editing MARC – Modifying subfield data

Allows for the modification of variable MARC field subfield data (MARC fields >10)

Allows for the modification of control field data by position or range of positions

Allows users to prepend and append data to subfields.

Allows users to change subfield tagging.
Editing MARC – Modifying subfield data

Allows users to insert new subfields and define subfield placement.

Allows users to move field data from one field to another.

Supports:
- UTF-8 with UTF-8 files
- Regular Expressions
- Adding new subfields.
Editing MARC – Modifying subfield data
Swapping Fields

- Define field, indicator and subfields to move.
- Can move field data and delete the original field or clone the field data and move the clone to the new location.
- Can add data to an existing field.
Build New Field Tool

This method works well when needing to build complex proxy statements.
Build New Field

Modifier functions

- Find
- replace
- trim
- trimstart
- trimend
- substring
- Regex

All functions can be stacked except the find function. If you use this function, it must be the first in the chain.

Example:  =856 \$a{090.find(“G42”).trim()}$zConnect to this resource online.
Example:  =901 \$a{LDR.substring(6,1)}$b{904$a}$mDate: {260$c.trim(“ .”)}
Character Conversions within the MarcEditor

MarcEditor allows users to convert character data between different character sets.
Fixing Boo-foos

MarcEdit’s Special Undo
- Allows you to step back one global change.
Sorting Fields

MarcEdit provides multiple sorting types:

- Control Number
  - Sorts record position within the file
- Title
  - Sorts record position within the file
- Author
  - Sorts record position within the file
- Call Number
  - Sorts record position within the file
- 0xx Fields
  - Sorts the 0xx fields within individual records (does not change record position within a file)
- All Fields
  - Sorts all fields within individual records (does not change record position within a file)
- Custom Sort
  - Sorts all defined fields within individual records (does not change record position within a file)
MarcEdit includes a number of validation/information reports:

- Field Count Reports
  - Including information related to indicator/subfield usage

- Material Type Reports
  - Generate Reports and pull together records for each material type

- Validate ISBN/ISSN Reports
  - Mathematically validates check digit

- Validate Headings *NEW*
  - Validates all access points against LCSH and LC NAF
Field Counts

Field Count
- Provides a quick count of fields
- Report of subfields used within a particular field
- Detailed reports of all fields/subfields used within a fileset.
Material Type Report

- Reports number of records by material type
- Breaks down material type by sub-types
- Utilizes the Leader, 008 and GMD to determine format types

Results:
- Books: 66
  - Large Print: 2
- Continuing Resources: 2
- Visual Materials: 1
  - Videorecording: 1
Validate ISSN/ISBN/Headings
In-Line Validation

MarcValidator-lite
- Can access MarcValidator for quick validation of data elements found in the file set
- Validation can use any defined rules set.
Generating Records from URLs

From the MarcEditor – you have the ability to generate MARC records directly from a URL or directly from the Library of Congress.
Task Automation Tool

New to MarcEdit 5.2, Task Automations

- Task automation provides a way for non-programmers to create defined task lists that can then be executed automatically.
- The difference between a task and a macro is that MarcEdit tasks essentially function like the user was calling specific functions within MarcEdit.
- Anything that you can do in the MarcEditor, you can automate as a task.
Task Automation

Managing Tasks
- Task management works like macro management
- You can
  - Create new tasks
  - Clone tasks
  - Rename tasks
  - Delete tasks
  - Edit tasks
Task Automation Demo

Additional Information:

- **Youtube:**
  - Introduction to task automation: [http://www.youtube.com/watch?v=gmqTGfTubU4](http://www.youtube.com/watch?v=gmqTGfTubU4)
  - Introduction to new task automation functions: [http://www.youtube.com/watch?v=fnorN0MFFN0](http://www.youtube.com/watch?v=fnorN0MFFN0)
Harvesting Metadata

MarcEdit includes a built-in OAI harvester

Allows for direct XML=>MARC translations

Allows for custom modification of XSLT translation tables.
ILS Integration currently supports direct integration with Koha, Alma, and a local option.

Are other integrations possible?

- [http://blog.reeset.net/archives/2133](http://blog.reeset.net/archives/2133)
Working with OCLC Connexion

https://youtu.be/a7Cen0gxFCw?list=PLrHRsJ91nVFScJLS91SWR5awtFfpewMWg
Working with OCLC’s Metadata API

MarcEdit can work directly with WorldCat via the Metadata API.
MarcEdit: Batch WorldCat Holdings Management

Update WorldCat Holdings

Update your OCLC WorldCat Holdings

Institution Code: OCPSB

Classification Schema: Library Of Congress

Options
- Add Holdings
- Delete Holdings

Print Debugging Info
MarcEdit: Batch Bibliographic Record Upload

Upload Records to WorldCat

Bibliographic records with OCLC numbers will be updated in WorldCat. Bibliographic records without OCLC numbers will be created in WorldCat.

Process

Process as Local Bibliographic Records

Close

Institution Code: OCPSB

Holding Code: MAIN Get Codes

Classification Schema: Library Of Congress
More Information

OCLC’s Developer Network:
- [http://oclc.org/developer/](http://oclc.org/developer/)

OCLC Metadata API Documentation:
- [http://oclc.org/developer/services/worldcat-metadata-api](http://oclc.org/developer/services/worldcat-metadata-api)

Notes on MarcEdit Integration:
- [http://blog.reeset.net/archives/1245](http://blog.reeset.net/archives/1245)

C# OCLC API Library
- [https://github.com/reeset/oclc_api](https://github.com/reeset/oclc_api)
Questions