Editing Records with the MarcEditor

TERRY REESE
HEAD OF DIGITAL INITIATIVES
THE OHIO STATE UNIVERSITY

REESE.2179@OSU.EDU
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 Preview Mode

![MarcEditor catalog.mrk](image)

C:\Users\reese\Desktop\Sample Files\catalog.mrk loading in 0.109200 seconds.
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
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
Editing MARC

Since this uses C# -- it makes use of Microsofts RegEx object which are designed to be compatible with PERL 5 regular expressions


Example:

◦ String =100 1\$aPauling, Linus, $d1908-

Find Text: <?<1>=100)<?2>.*)(?<3>[^
].\])$

Replace with: ${1}${2}-1999.

Result: =100 1\$aPauling, Linus, $d1908-1999.
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.
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.
Jump List

Find All
Jump List

Jump List Example

<table>
<thead>
<tr>
<th>Found Text</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>=245 00SaThe 2005 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 2</td>
</tr>
<tr>
<td>=245 00SaThe 2004 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 3</td>
</tr>
<tr>
<td>=245 00SaThe 2003 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 4</td>
</tr>
<tr>
<td>=245 00SaThe 2001 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 611</td>
</tr>
<tr>
<td>=246 00SaThe 2002 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 1939</td>
</tr>
<tr>
<td>=245 00SaThe 1999 CIA FactbookSh[internet resource]</td>
<td>Jump to Record # 2303</td>
</tr>
<tr>
<td>=245 00SaThe 1996 CIA FactbookSh[internet resource]</td>
<td>Jump to Record # 2384</td>
</tr>
<tr>
<td>=245 00SaThe 1997 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 9969</td>
</tr>
<tr>
<td>=245 00SaThe 1996 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 15639</td>
</tr>
<tr>
<td>=245 00SaThe 2007 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 20673</td>
</tr>
<tr>
<td>=245 00SaThe 2006 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 22590</td>
</tr>
<tr>
<td>=245 00SaThe 2000 CIA World FactbookSh[internet resource]</td>
<td>Jump to Record # 26562</td>
</tr>
</tbody>
</table>
Jump List

When using the jump list:

◦ Will jump to the page and record within the set
◦ Will save (temporarily) any items modified or pages automatically (though to set saved items, you need to actually save the page)
Jump to

Jump to...record:
  ◦ Allows you to jump to any records

Jump to...page:
  ◦ Allows you to jump to any page
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
Editing MARC – Swapping Fields

Swap parts of MARC Fields or entire MARC 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.
Character Conversions within the MarcEditor

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

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 provides a simple dedup tool that can:

- Dedup on a defined control field (any field)
- Dedup on a transaction field (or using an additional transaction field)

**Output**

- Removes all duplications and saves the duplications to a file
- Prints just unique items within the file (i.e., those without a duplicate pair)
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
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.
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=gmqTGuUbU4](http://www.youtube.com/watch?v=gmqTGuUbU4)
  - 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.
Harvesting Metadata

Harvesting Demo
- Current OSU ETD workflow
  - Moving from Dspace to OCLC
Working with OCLC Connexion
Working with OCLC’s Metadata API

MarcEdit can work directly with WorldCat via the Metadata API.
Available Operations

- Create/Read/Update Bibliographic Records
- Update/Delete Institutional Holdings
- Retrieve Holding Code information about an Institution (Why?)
- Create/Read/Update Local Bibliographic Data
Challenges: Authentication

Challenges: Documentation

Challenges: No Sandbox

But this is really cool because we can…

- Further automate traditional technical services processes
  - Specifically holdings management
  - Batch record ingestion
- Build pipelines between our repository systems and WorldCat
- Develop localized interfaces for metadata entry outside the library
- Opens up the opportunity for tool builders to interact with the OCLC member community
MarcEdit

MARC Tools
Delimited Text Translator
MarcEditor
Harvest OAI Records
MARC SQL Explorer
Z39.50/SRU Client
Kittens riding unicorns

Onward, to victory!

http://kulworlddailydrivel.files.wordpress.com/2013/02/kittens-riding-unicorns.jpg?w=492&h=369
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.

[Field for file]

Process

[Box to select] Process as Local Bibliographic Records

Process

Institution Code: OCPSB

Holding Code: MAIN

Classification Schema: Library Of Congress

Get Codes
MarcEdit Integration

Just don’t forget changes are made live…
Conspicuously Absent Operations

- Record Validation
- Anything to do with authority data
- Record Locking (for record editing)
- Service Status
- User Validation (for permission validation)
What more information?

- OCLC’s Developer Network:
  - http://oclc.org/developer/

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

- Notes on MarcEdit Integration:
  - http://blog.reeset.net/archives/1245

- C# OCLC API Library
  - https://github.com/reeset/oclc_api