# Demystifying Batchload Analysis

Yael Mandelstam
Fordham Law Library
ymandelstam@law.fordham.edu

AALL 2009 Annual Meeting

## Sample Record Sets for Law Libraries

- Making of Modern Law (MOML)
- Serials Solutions
- LLMC-Digital
- BNA
- CALI
- HeinOnline Legal Classics
- HeinOnline World Trials
- LexisNexis/Westlaw Cassidy Collections

## Analyzing Batches of Records

- step 1: examine several individual records
- step 2: count fields in file
- step 3: view isolated fields

## Before We Start …

Keep in mind that we will be **viewing** records and isolated fields. This technique is not meant to replace proper authority control, use of URL checkers, etc.

## Readily-Available Tools

- MarcEdit
  - free editing utility available for download at http://oregonstate.edu/~reeset/marcedit/html/
  - always keep a copy of the original file

- Microsoft Excel
  - for sorting purposes, keep in mind that:
    - version 2003 has 65,536 rows and can accommodate approx. 3000 full MARC records
    - version 2007 has one million rows and can accommodate approx. 45,000 full MARC records

## Step 1: Examining Individual Records

- look for
  - level of cataloging (full, core, minimal)
  - adherence to cataloging standards
  - quality of subject headings and classification
  - decisions made by cataloging agency (e.g. treatment of aggregator in 260/533 and 710 fields, resource cataloged as monograph, serial, or integrating)
  - errors and inconsistencies

## Step 1: Examining Individual Records
### Example: Print to Electronic Conversion

- electronic record derived from print, but no changes made to fields 130 and 222

```
050  14$aQ180.55.M3$bJ687
082  10$a001$213
130  0\$aJournal of research administration (Print)
210  0\$aJ. res. adm.$b(Print)
222  \4$aThe journal of research
administration$b(Print)
245  14$aThe journal of research
administration$h[electronic resource].
260  \\$aWashington, DC :$bSociety of Research
Administrators,$cc2000-
300  \\$av   :$c26 cm.
```

## Step 1: Examining Individual Records
### Example: Decision Regarding 260 Field

- digitized e-book: original publisher in 260 …

```
260  \\$aWashington, D.C. :$bO.H.
Oldroyd,$cc1901.
300  \\$axviii, 305 p. :$bill., maps ;$c20 cm.
533  \\$aElectronic reproduction.$b[Buffalo, NY]
:$cW.S. Hein,$d[2006].$f(Hein online).$nMode of
access: Internet.  Address:
http://www.heinonline.org.$nSubscription required
for access.$7s2006    nyun s
600  10$aLincoln,
Abraham,$d1809-1865$xAssassination.
```

## Step 1: Examining Individual Records
### Example: Decision Regarding 260 field

- … and the same resource with digitizer in 260

```
Harris.
260  \\$a[Buffalo, NY] :$bWilliam S. Hein &
Co.,$c[2007]
300  \\$a1 electronic text (xviii, 305 p. : ill.,
facsims., map, plans, port.) :$bPDF file.
500  \\$aTitle from PDF title page, viewed
December 7, 2007.
500  \\$aPrint version published: Union, N.J.:
Lawbook Exchange, 2001. Originally published:
Washington, D.C.: O. H. Oldroyd, 1901.
538  \\$aMode of access: World Wide Web
```

## Step 2: Counting Fields

- make sure all records have all required fields
- examine report for inconsistencies
- be mindful of locally-assigned fields that may cause problems when loading records into your ILS

---

## Step 2: Counting Fields in MarcEdit

- select Reports, Field Count



---
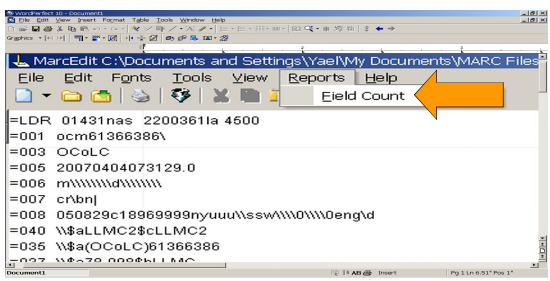
## Step 2: Counting Fields
### What's Wrong With This Picture? Example #1

- file has 9111 records for e-books …

## Step 2: Counting Fields
### What's Wrong With This Picture? Example #1

- … but only 64 subject headings!

```
555        1
580        1
600        19
610        4
650        38
655        3
700        2099
710        176
730        476
740        13
752        1
775        2
850        16
856        9111
```

## Step 2: Counting Fields
### What's Wrong With This Picture? Example #2

- file has 221 records for e-journals …

```
Label     No. in Fields    No. in
Subfields
LDR       221
001       221
003       120
005       221
006       2
007       2
008       221
010       120
010$a                120
010$z                41
016       6
016$2                6
```

## Step 2: Counting Fields
### What's Wrong With This Picture? Example #2

- … but only 31 records have URLs!

```
853$v                15
853$w                16
853$x                11
853$y                3
856       31
856$2                2
856$3                1
856$u                33
856$x                3
856$y                3
856$z                14
863       18
863$8                18
863$a                18
```

## Step 3: Isolating Fields for Analysis

- in MarcEdit
  - copy all records (Ctrl+A, Ctrl+C)

- in Excel
  - paste records (Ctrl+V)
  - select all records (Ctrl+A)
  - sort all fields by selecting Data, Sort, Column A

## Step 3: Isolating Fields for Analysis
### The Power of Eyeballing

- scroll down to view individual fields
  - look for errors and inconsistencies
  - target words/strings in specific fields

- to search a specific field, copy all its occurrences and paste in separate tab
  - search for specific elements using the Find function (Ctrl+F)
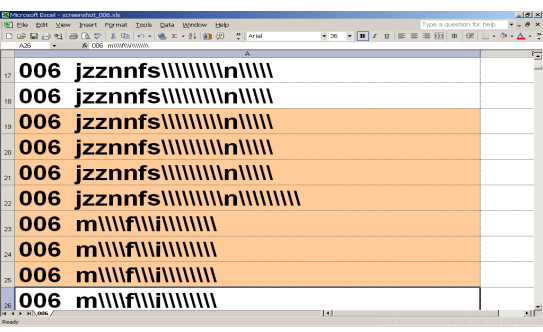
## Step 3: Isolating Fields for Analysis
### The Power of Eyeballing

- keep your eyes open for
  - coding in Leader, 006, 007, and 008 fields
  - errors, inconsistencies, and obsolete indicators/subfields in variable-length fields
  - 856 $z inconsistent with your local practice
  - 1xx and 7xx inconsistencies that cannot be detected by automatic authority control
  - incorrect 6xx subdivisions that cannot be detected by automatic authority control

# Step 3: Isolating Fields for Analysis
## The Power of Eyeballing

- □ examples of problematic subject headings
  - obsolete headings (e.g. Afro-Americans, Trade-unions)
  - obsolete subdivisions (e.g. Constitutional law, Constitutional history)
  - incorrect Law and legislation subdivisions
  - incorrect application of subdivisions authorized for use with legal headings (e.g. Cases, Digest)

# Step 3: Isolating Fields for Analysis
## Example: Inconsistent Coding in Field 006



# Step 3: Isolating Fields for Analysis
## Example: Inconsistent Form of Name

## Step 3: Isolating Fields for Analysis
### Example: Annuals Cataloged as Monographs

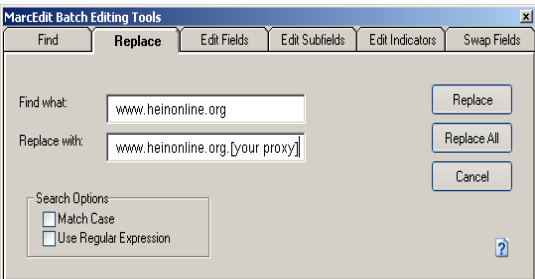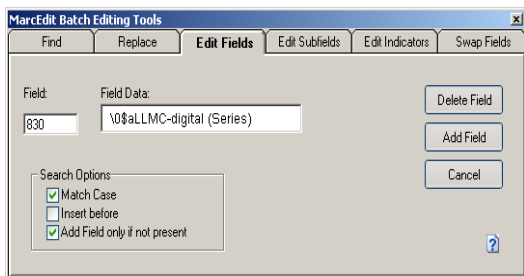| | |
|---|---|
| 96 | 245 00$aCFR 1999$h[electronic resource]. |
| 97 | 245 00$aCFR 2000$h[electronic resource]. |
| 98 | 245 00$aCFR 2001$h[electronic resource]. |
| 99 | 245 00$aCFR 2002$h[electronic resource]. |
| 100 | 245 00$aCFR 2003$h[electronic resource]. |
| 101 | 245 00$aCFR 2004$h[electronic resource]. |
| 102 | 245 00$aCFR 2005$h[electronic resource]. |
| 103 | 245 00$aCFR 2006$h[electronic resource]. |
| 104 | 245 00$aCFR, 1981$h[electronic resource]. |
| 105 | 245 00$aCFR, 1982$h[electronic resource]. |
| 106 | 245 00$aCFR, 1983$h[electronic resource]. |
| 107 | 245 00$aCFR, 1984$h[electronic resource]. |
| 108 | 245 00$aCFR, 1985$h[electronic resource]. |

## Step 3: Isolating Fields for Analysis
### Example: Incorrect "Cases" Subdivision

| | |
|---|---|
| 1 | 650 0|aCarriers|xLaw and legislation|zUnited States.@ |
| 2 | 650 0|aCarriers|xLaw and legislation|zUnited States|xCases. |
| 3 | 650 0|aCarriers|zUnited States.@ |
| 4 | 650 0|aCarriers|zUnited States.@ |
| 5 | 650 0|aCarriers|zUnited States|xCases.@ |
| 6 | 650 0|aCarriers|zUnited States|xCases.@ |
| 7 | 650 0|aCarriers|zUnited States|xCases.@ |
| 8 | 650 0|aCarriers|zUnited States|xCases.@ |
| 9 | 650 0|aCarriers|zUnited States|xCases.@ |
| 10 | 650 0|aCarriers|zUnited States|xCases.@ |
| 11 | 650 0|aCarriers|zUnited States|xCases.@ |
| 12 | 650 0|aCarriers|zUnited States|xCases.@ |
| | 650 0|aCartels.@ |

## Step 3: Isolating Fields for Analysis
### Example: Obsolete $x Constitutional law/history

| | |
|---|---|
| 70 | 651 0|aUnited States|xConstitutional history.@ |
| 71 | 651 0|aUnited States|xConstitutional history.@ |
| 72 | 651 0|aUnited States|xConstitutional history.@ |
| 73 | 651 0|aUnited States|xConstitutional history.@ |
| 74 | 651 0|aUnited States|xConstitutional history.@ |
| 75 | 651 0|aUnited States|xConstitutional history.@ |
| 76 | 651 0|aUnited States|xConstitutional history.@ |
| 77 | 651 0|aUnited States|xConstitutional law.@ |
| 78 | 651 0|aUnited States|xConstitutional law.@ |
| 79 | 651 0|aUnited States|xConstitutional law.@ |
| 80 | 651 0|aUnited States|xConstitutional law.@ |
| 81 | 651 0|aUnited States|xConstitutional law.@ |

## Cleanup Time!
## Global Update With MarcEdit

▫ find and replace data (e.g. add proxy to 856 $u)

**MarcEdit Batch Editing Tools**

| Find | Replace | Edit Fields | Edit Subfields | Edit Indicators | Swap Fields |

Find what: www.heinonline.org

Replace with: www.heinonline.org.[your proxy]

Replace
Replace All
Cancel

Search Options
☐ Match Case
☐ Use Regular Expression

## Global Update With MarcEdit

▫ add/delete fields (e.g. add missing series)

**MarcEdit Batch Editing Tools**

| Find | Replace | **Edit Fields** | Edit Subfields | Edit Indicators | Swap Fields |

Field: 830

Field Data: \0$aLLMC-digital (Series)

Delete Field
Add Field
Cancel

Search Options
☑ Match Case
☐ Insert before
☑ Add Field only if not present

## Global Update With MarcEdit

▫ edit subfield data (e.g. change 650 $x to $v)

**MarcEdit Batch Editing Tools**

| Find | Replace | Edit Fields | **Edit Subfields** | Edit Indicators | Swap Fields |

Field: 650    Subfield: x    Field Data: ^cPeriodicals

Replace With: v

Replace Text
Remove Text
Cancel

Search Options
☐ Add subfield if not present    ☐ Use Regular Expression
☐ New subfield only
☑ Match Case

## Global Update With MarcEdit

◘ edit subfield data (e.g. add suffix to 001)

**MarcEdit Batch Editing Tools**

| Find | Replace | Edit Fields | **Edit Subfields** | Edit Indicators | Swap Fields |

Field: 001
Subfield:
Field Data: ^e

Replace With: moml

Replace Text
Remove Text
Cancel

Search Options
☐ Add subfield if not present   ☐ Use Regular Expression
☐ New subfield only
☐ Match Case
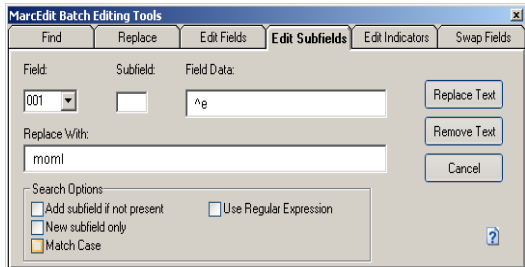
## Global Update With MarcEdit

◘ edit indicator data (e.g. change obsolete 740 2nd indicator "0" to "blank")

**MarcEdit Batch Editing Tools**

| Find | Replace | Edit Fields | Edit Subfields | **Edit Indicators** | Swap Fields |

Field: 740
Indicators: 0
Field Data:

Replace
Cancel

Replace Indicators: \
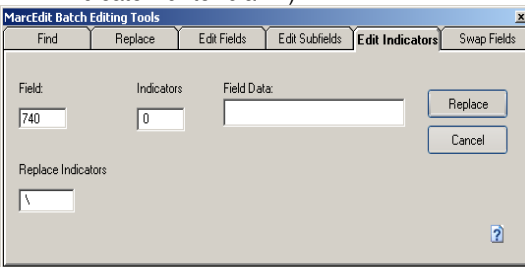
## Global Update With MarcEdit

◘ swap field data (e.g. move data from 035 to 001)

**MarcEdit Batch Editing Tools**

| Find | Replace | Edit Fields | Edit Subfields | Edit Indicators | **Swap Fields** |

**Original Data:**
Field: 035   Indicators:   Subfields:   Find:

**Modified Data:**
Field: 001   Indicators:   Subfields:

OK   Cancel

Search Options
☐ Copy Source Data   ☐ Add to Existing Field   ☐ Override default
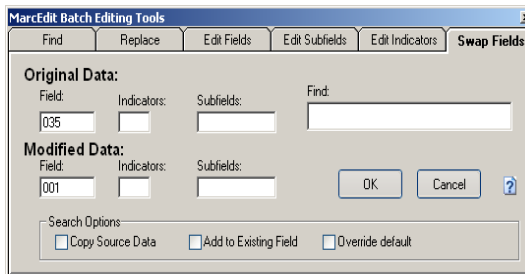
## Eliminating Redundancies

- when finding problems with record sets, we can either clean up files locally

  OR

- work with vendors to improve the quality of the records to the benefit of all

## Too Daunting to Do on Your Own?

- get help from the Task Group on Vendor-Supplied Bibliographic Records

  - official website: http://www.aallnet.org/sis/tssis/committees/cataloging/vendorbibrecords/

  - wiki: http://tsvbr.pbwiki.com/
    find vendor information and evaluation of record sets

## Task Group on Vendor-Supplied Bibliographic Records (VBR)

- offers to assist vendors in preparation of specifications for contract cataloging and recommends changes and revisions as appropriate

- encourages vendors to make arrangements with authority control vendors to supply current authority data for the cataloging produced

- develops and disseminates evaluations/reviews of the cataloging data to TS-SIS members via the group's wiki

## VBR Wiki: Areas Covered

- general information
  - name of vendor
  - contact information
  - description
  - records availability
  - test records availability
  - methodology for creation of records
  - customization

## VBR Wiki: Areas Covered (cont.)

- evaluations
  - review status
  - authority control
  - fixed fields
  - variable-length fields
  - subject headings/classification
  - other
- additional comments

## VBR Wiki:
### View of Record Set Page

_____

_____

_____

_____

_____

_____

_____