A Batch Solution to the Death Date Problem

A Case Study

Hsianghui Liu-Spencer & Tom Lamb
Carleton College
Northfield, Minn.
IUG Conference, San Francisco
April 25, 2013

LC Announcement in 2005/2006

Revised LCRI 22.17

The Library of Congress announced a change in policy about adding death dates to personal name headings and a revision to the LCRI for AACR2 rule 22.17 that allows the option for catalogers to add death dates to personal name headings with open dates.

010 no2011144865

100 1 Sagan, Françoise, |**d** 1935-
|tGarde du coeur.|lEnglish

010 n 79021390

100 1 Sagan, Françoise, |**d** 1935-

400 1 Quoirez, Françoise, |**d** 1935-

One Bibliographic Record

100 1	**Sagan, Françoise,|d1935-**
240 10	**Garde du coeur.|lEnglish**
245 14	**The heart-keeper /|c[by] Françoise Sagan.**
260	**New York :|bDutton,|c1968.**
300	**128 p. ;|c21 cm.**
600 10	**Sagan, Françoise,|d1935-**
600 10	**Sagan, Françoise,|d1935-|vA[illegible]cdotes.**

010 n 79021390

100 1 Sagan, Françoise, |**d** 1935-

400 1 Quoirez, Françoise, |**d** 1935-

010 no2011144865

100 1 Sagan, Françoise, |**d** 1935-2004.
|tGarde du coeur.|lEnglish

010 n 79021390

100 1 Sagan, Françoise, |**d** 1935-2004

400 1 Quoirez, Françoise, |**d** 1935-2004

One Bibliographic Record

100 1	**Sagan, Françoise,|d1935-2004.**
240 10	**Garde du coeur.|lEnglish**
245 14	**The heart-keeper /|c[by] Françoise Sagan.**
260	**New York :|bDutton,|c1968.**
300	**128 p. ;|c21 cm.**
600 10	**Sagan, Françoise,|d1935-2004.**
600 10	**Sagan, Françoise,|d1935-2004|v[illegible]ecdotes.**

010 n 79021390

100 1 Sagan, Françoise, |**d** 1935-2004

400 1 Quoirez, Françoise, |**d** 1935-2004

Authority records for 20060201
Some entries require the ALA BT Courier font to dis
Explorer for best results when viewing this font.
Authority records for 20100606
Authority Records updates posted after January 19, 2006 require Arial Unicode font to display all extended characters. This font
ships with various software packages or can be purchased separately. (Learn more about Arial Unicode fonts.)
Authority records for 20100725
Authority Records updates posted after January 19, 2006 require Arial
Closed Dates in Authority Records
The Library of Congress recently announced a change in policy about adding dates to pe
revisions to the LCRI for AACR2 rule 22.17 that allows the option for catalogers to add death dates to
with open dates (http://www.loc.gov/aba/cataloging/classify/lcri_22.pdf).
Both the Library of Congress and OCLC are aware that many libraries will be very interested in incorporating thi
their catalogs. Libraries that use authority control services for preparing processing will receive these changes as pa
normal workflows but, for other libraries, OCLC is providing this service to announce headings to which death dates h
added.
Lists for each week since the new policy went into effect on February 1, 2006 are available below and as an RSS feed.
instructions for subscribing to RSS feeds can be found on RSS Feeds and Podcasts. A list of newly changed headings will be
added once each week in the future.
NOTE: Authority Records updates posted after January 19, 2006 require Arial Unicode font to display all extended characters.
This font ships with various software packages or can be purchased separately. (Learn more about Arial Unicode fonts.) Older
entries require the ALA BT Courier font to display properly. Use Internet Explorer for best results when viewing the ALA BT
Courier font.
2010
Authority records for 20100826
Authority records for 20100819
Authority records for 20100812
Authority records for 20100805
Authority records for 20100729
Authority records for 20100722
Authority records for 20100715
Authority records for 20100708
Authority records for 20100701
Authority records for 20100624
Authority records for 20100617
Authority records for 20100610
Authority records for 20100603
Authority records for 20100527
Authority records for 20100520
Authority records for 20100513
Authority records for 20100506
Authority records for 20100429
Authority records for 20100422
Authority records for 20100415
Authority records for 20100408
Authority records for 20100401
Authority records for 20100325
Authority records for 20100318
Authority records for 20100311
Authority records for 20100304
Authority records for 20100225
Authority records for 20100218
Authority records for 20100211
Authority records for 20100204
Authority records for 20100128
Authority records for 20100121
Authority records for 20100114
Authority records for 20100107
10/1/2010 5:12 PM
10/15/2010 4:41 PM

A Name Change in LC Weekly List 2010-05-09

The Library of Congress >> Go to Library of Congress Online Catalog

LIBRARY OF CONGRESS AUTHORITIES

Help | New Search | Search History | Headings List | Start Over

◀ Previous Next ▶

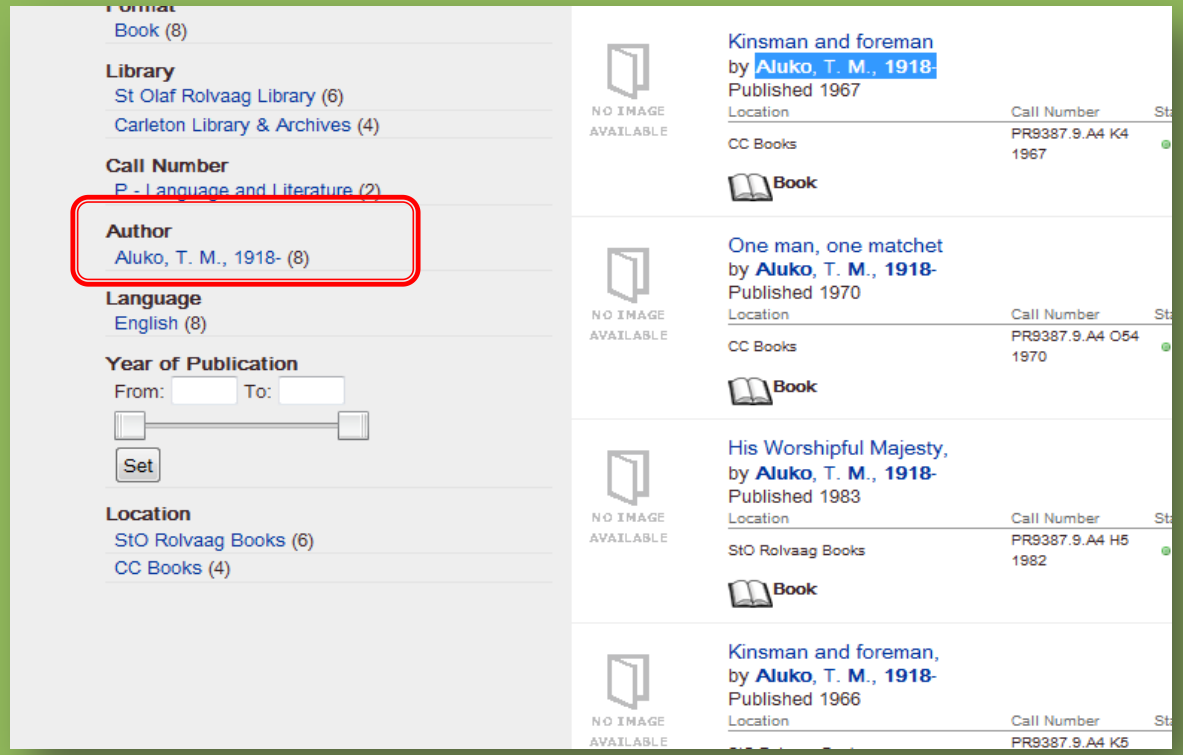
MARC Display | Labelled Display

LC control no.: n 83239259
LCCN permalink: http://lccn.loc.gov/n83239259
HEADING: Aluko, T. M., 1918-2010

000 00880cz a2200181n 450
001 1500399
005 20100503164753.0
008 831103n| acannaabn |a aaa
010 __ |**a** n 83239259
035 __ |**a** (OCoLC)oca01039232
040 __ |**a** DLC |**b** eng |**c** DLC |**d** DLC |**d** InU |**d** OCoLC |**d** DLC
053 _0 |**a** PR9387.9.A4
100 1_ |**a** Aluko, T. M., |**d** 1918-2010
400 1_ |**a** Aluko, Timothy Mofolorunso, |**d** 1918-2010
670 __ |**a** His Wrong ones in the dock, 1982: |**b** t.p. (T. M. Aluko) p. 4 of cover (b. 1918)
670 __ |**a** My years of service, 1994: |**b** t.p. (T.M. Aluko) p. vii (Dr. Timothy Mofolorunso Aluko; one of the forerunners of engineering and town planning professions in Nigeria; also author of popular novels) p. 3 (b. 6-14-18)
670 __ |**a** Wikipedia WWW site, May 3, 2010 |**b** (under T.M. Aluko: Timothy Mofolorunso Aluko; b. June 14, 1918, Ilesha; d. May 1, 2010, Lagos; Nigerian writer)
953 __ |**a** ea08 |**b** tc04

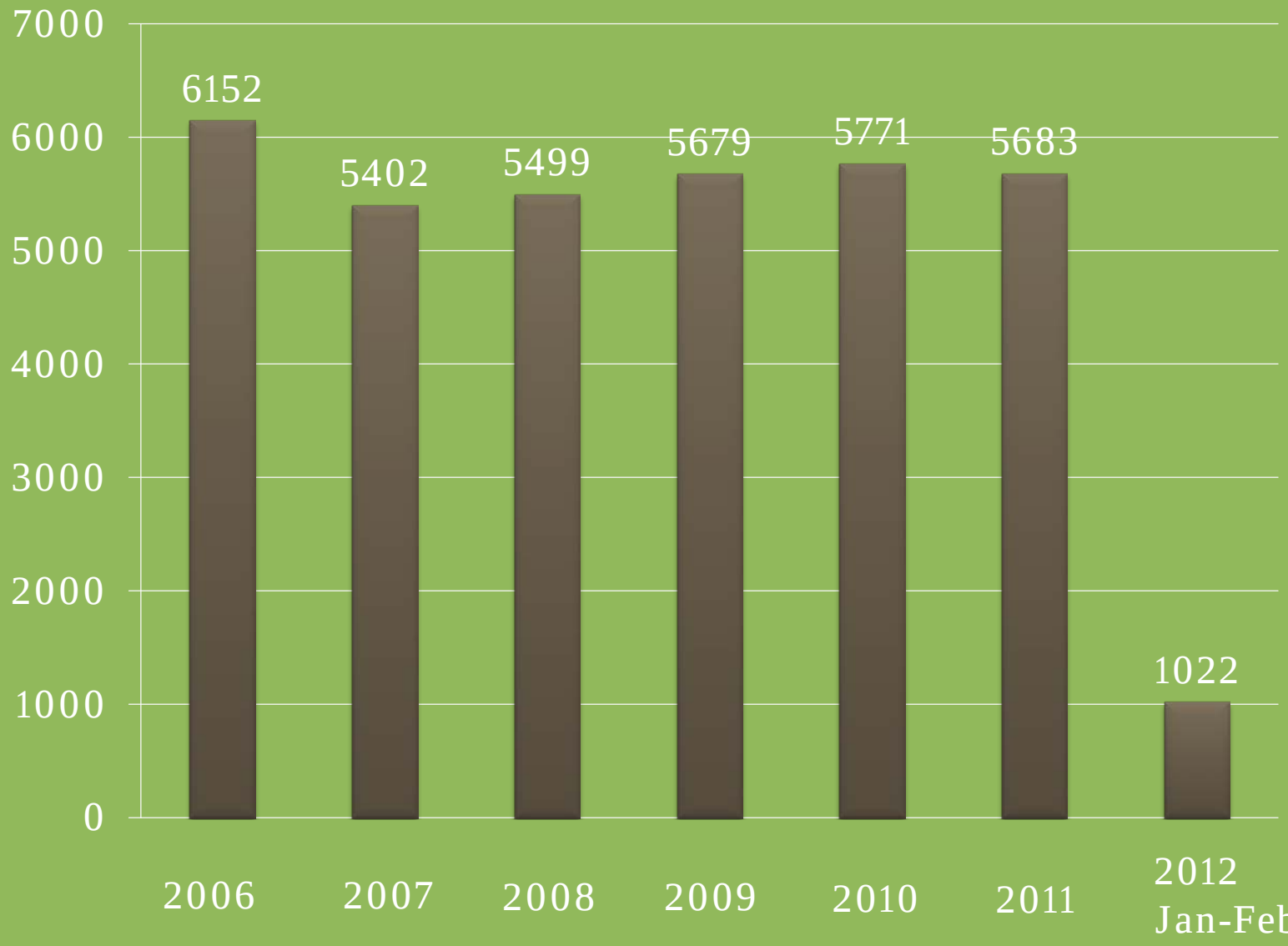
How Does This Impact a Local Database?

Heading in Bridge Catalog (Carleton & St. Olaf)



Split headings in other library catalog

LC Name Changes from 02/2006 to 02/2012



Source: LC Weekly Lists website

A total of 35,208 names over six-year period

Name Changes in Bridge Catalog

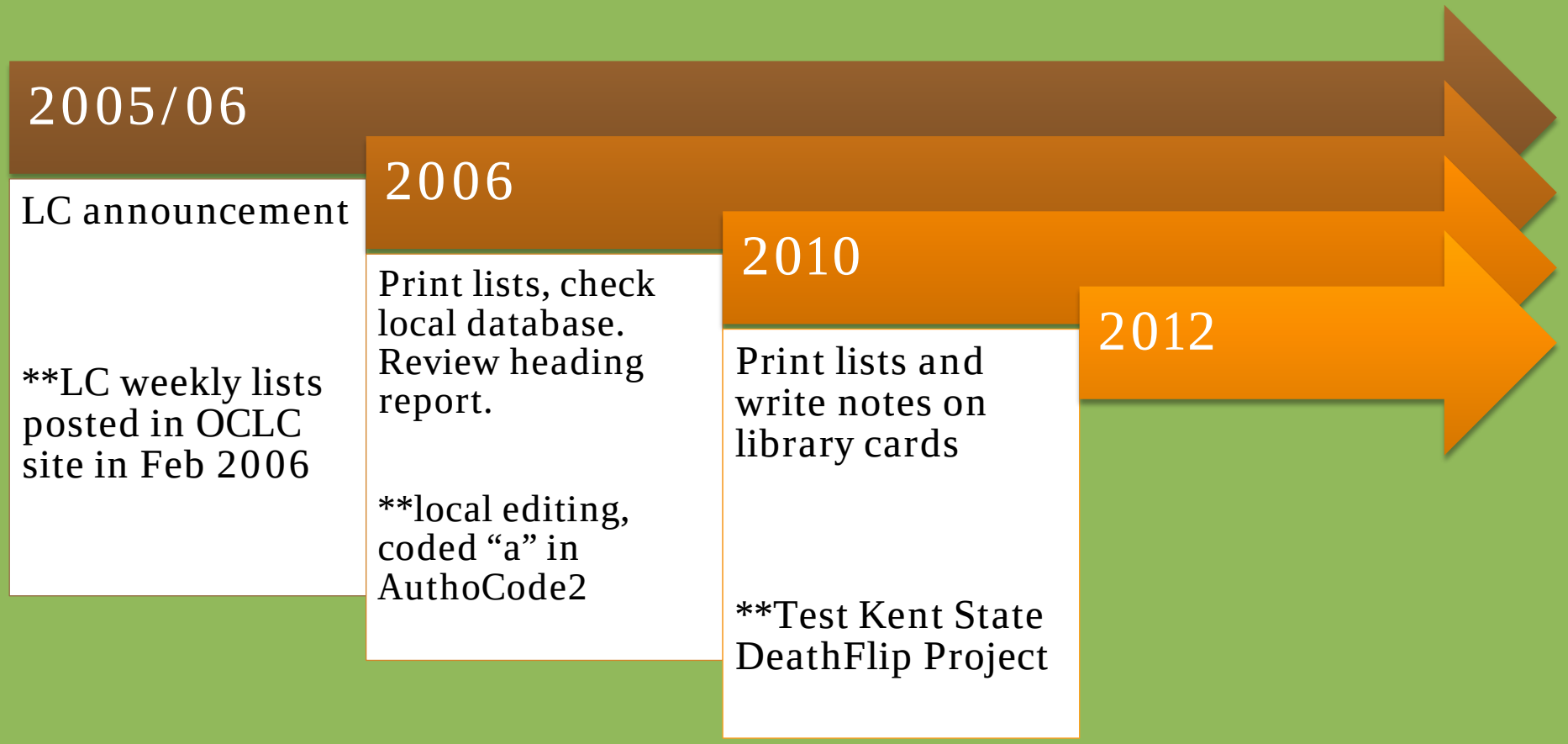
A total of 14,143 names over six-year period

What do you do?

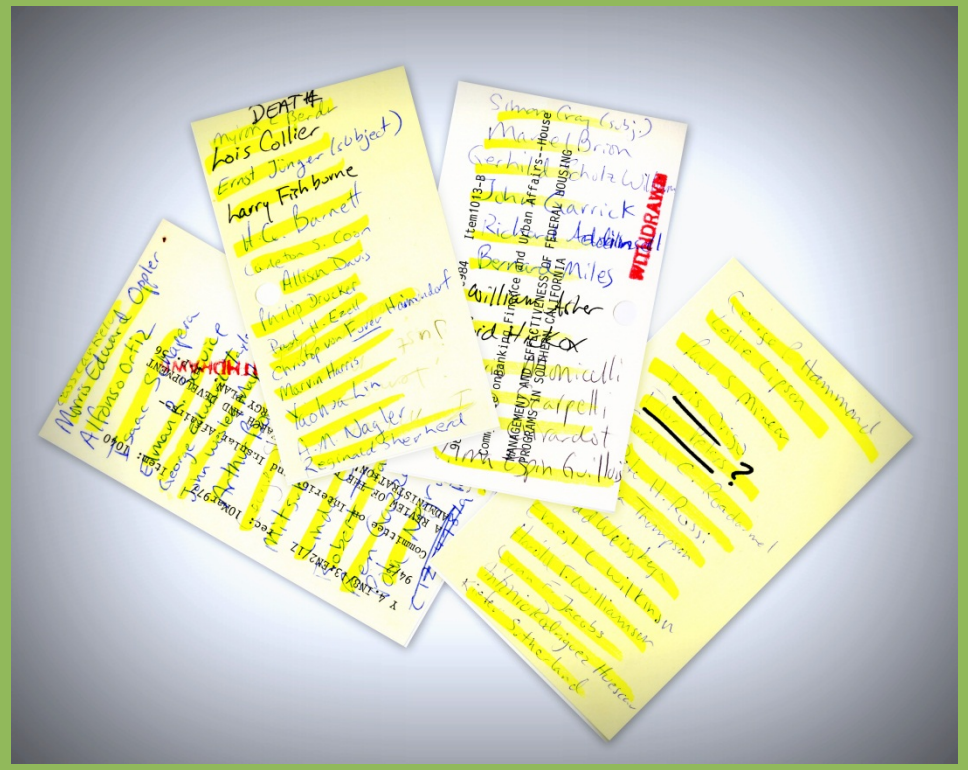
Or do you care?

OCLC-CAT Listserv in Feb, 2012
**13 Responses

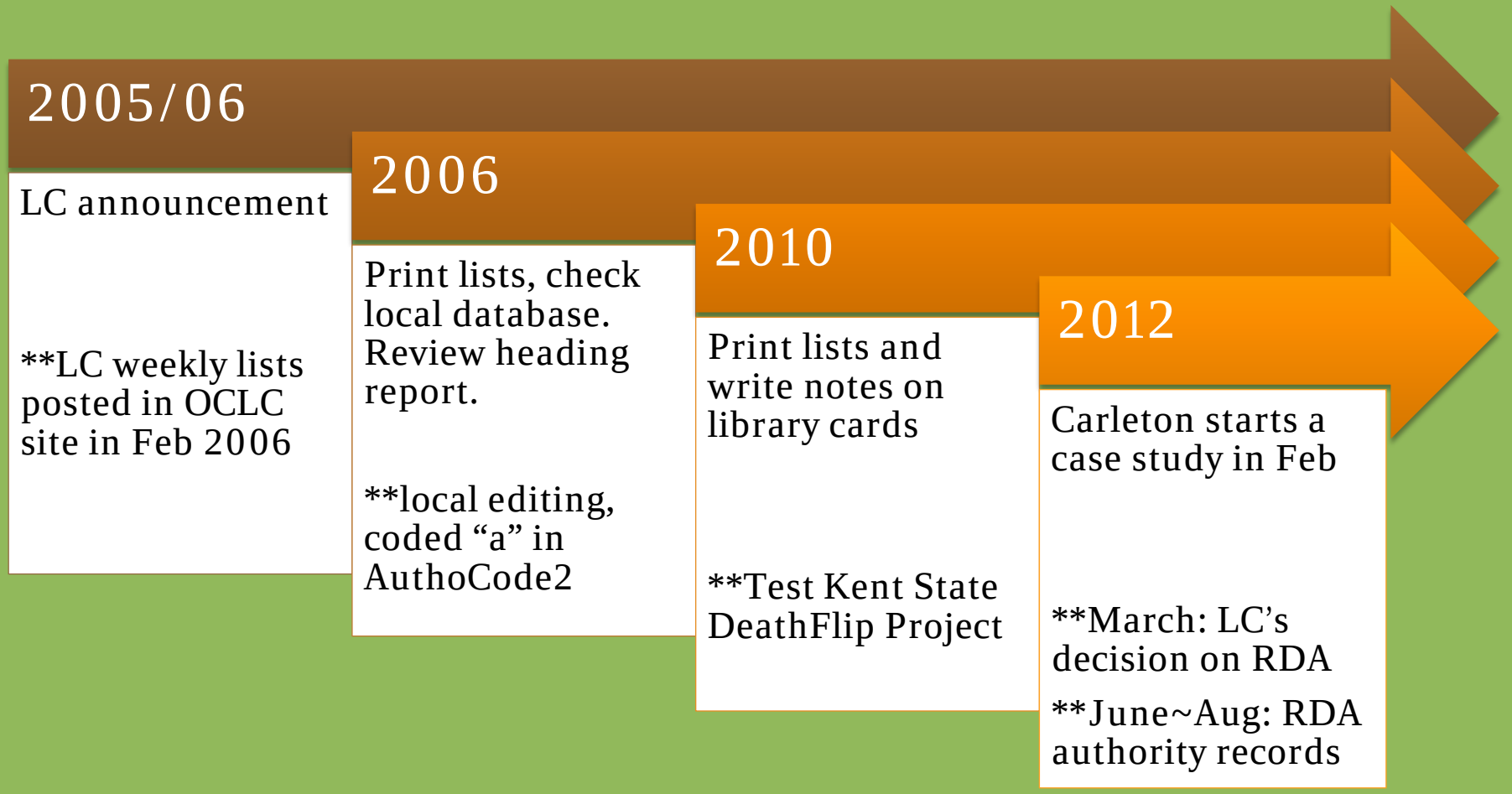
Bridge Consortium Reacts Over the Years



The Old Ways We Tried to Keep Up...



Bridge Consortium Reacts Over the Years



Our Goal

- Identify headings in our system that appear in the LC Weekly Closed Dates list and update those headings as needed.

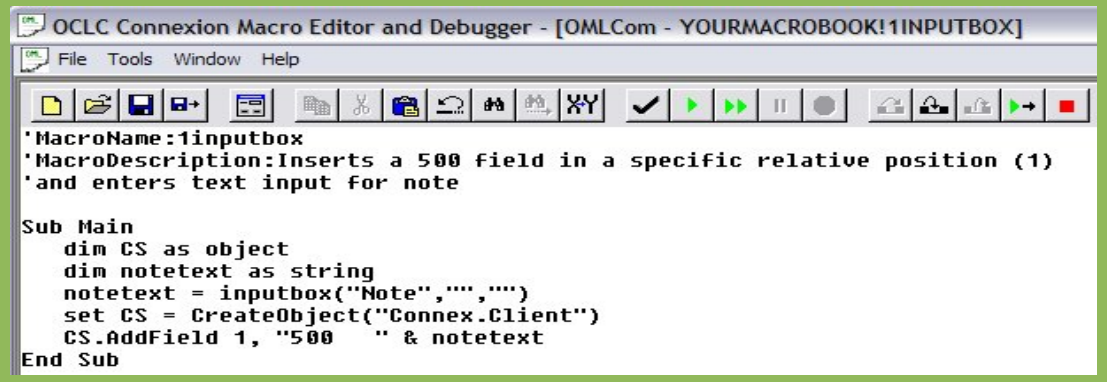
We Need to Be Aware:

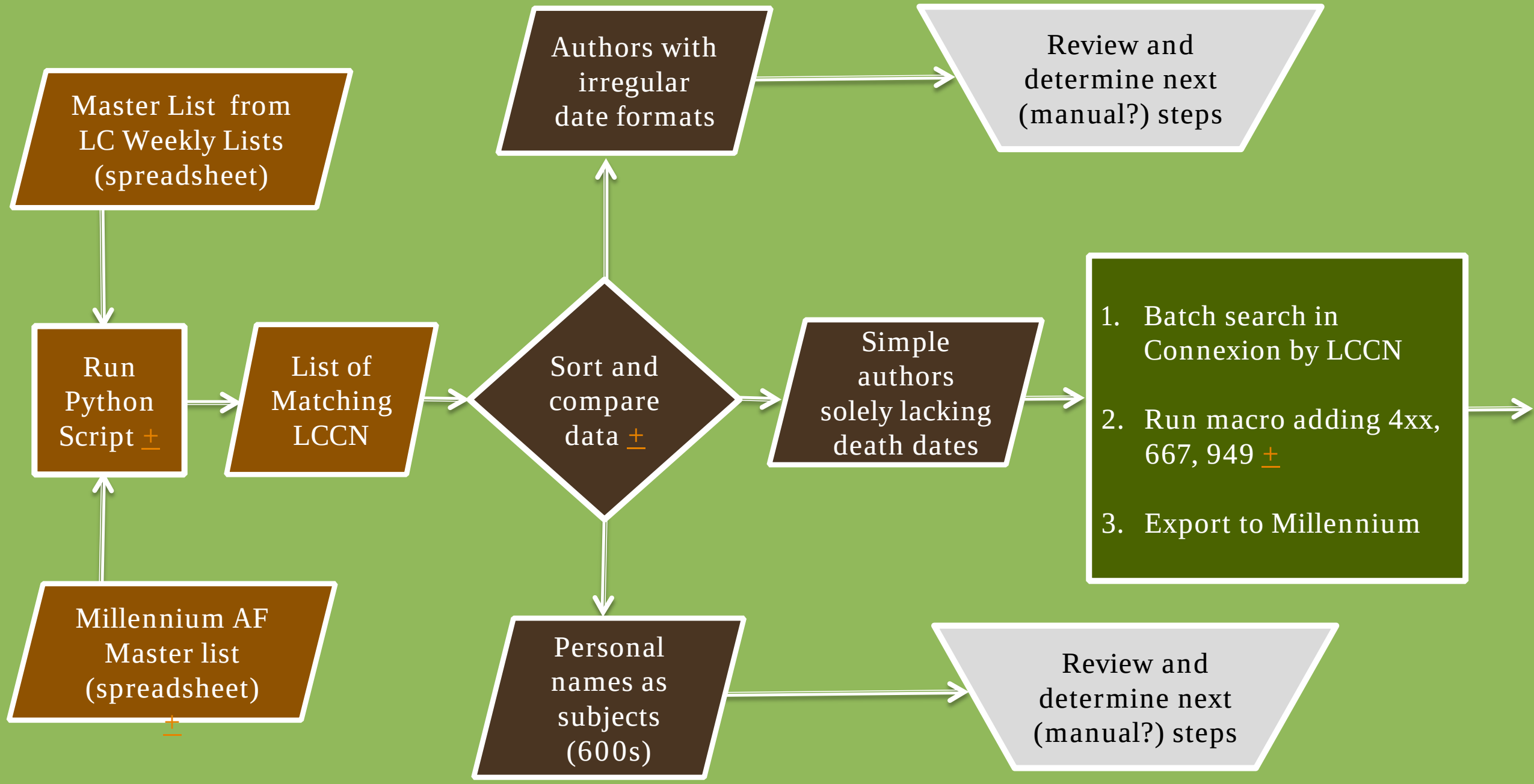
1. Not all identified headings require updating.
2. Identified headings may present unique challenges requiring unique approaches:
    - Heading for 1xx and 7xx
    - Heading for 6xx subject
    - Author/Title added entry & Uniform Title (240)

Tools That We Used

DataTool

http://datatoolbar.com/




Master List from LC Weekly Lists (spreadsheet)
Run Python Script ±
List of Matching LCCN
Sort and compare data ±
Authors with irregular date formats
Review and determine next (manual?) steps
Simple authors solely lacking death dates
1. Batch search in Connexion by LCCN
2. Run macro adding 4xx, 667, 949 ±
3. Export to Millennium
Millennium AF Master list (spreadsheet) ±
Personal names as subjects (600s)
Review and determine next (manual?) steps

Day One

Observe blind references and let AACP work its magic

Day Two

Check heading report in Millennium

Blind references – mostly need to create a different 4xx. So, it will flip the following day

±

Check *Updated Bib Heading* as well as near-match, duplicate AF, etc.

Day Three

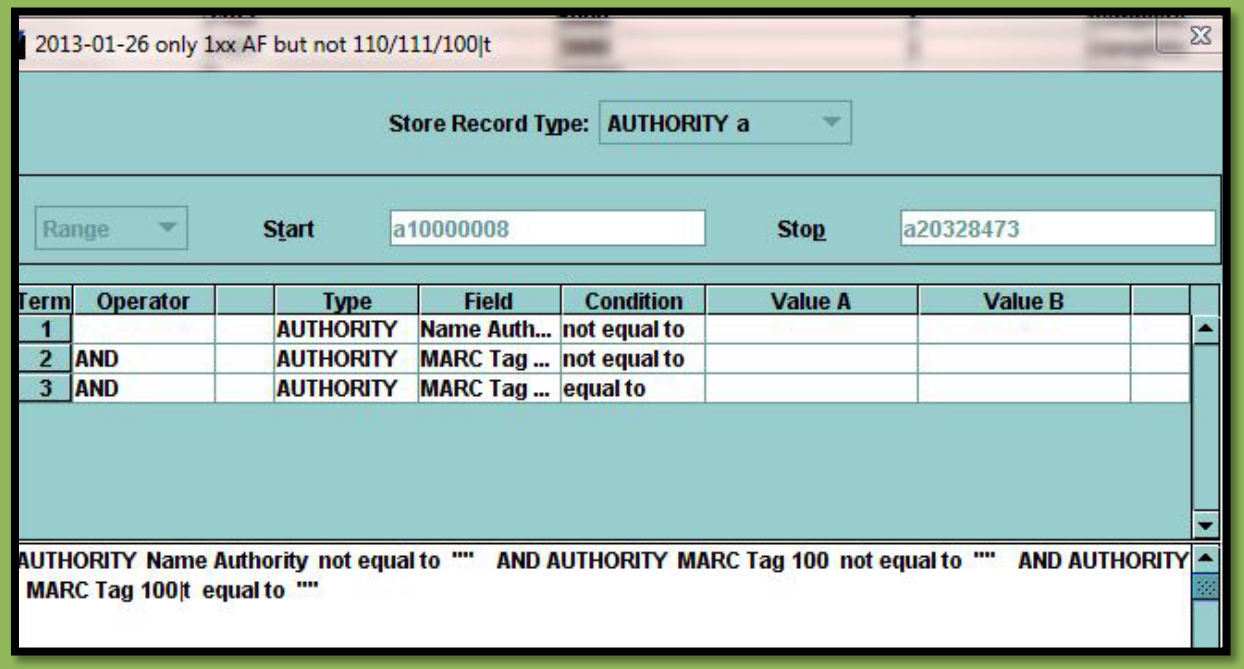
Make sure no more blind references

Don't forget subjects (6xx) and irregular date formats (1xx, 7xx)

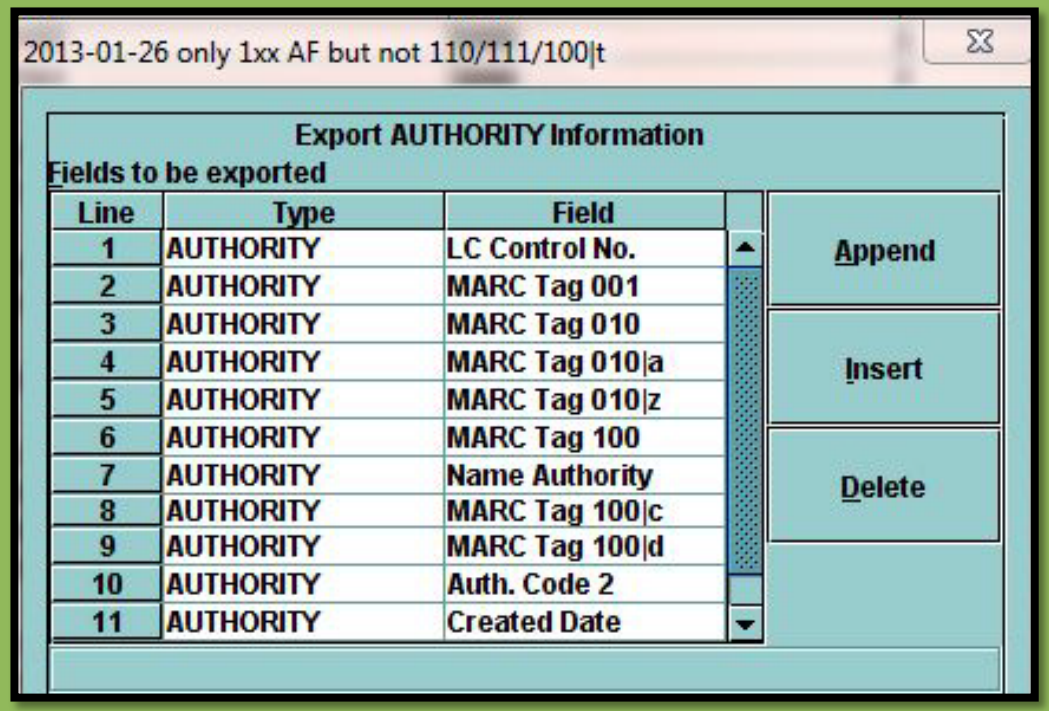
->

Review List for 1XX Name Heading

Search strategy:
contains 1XX but not 110/111 or 100|t

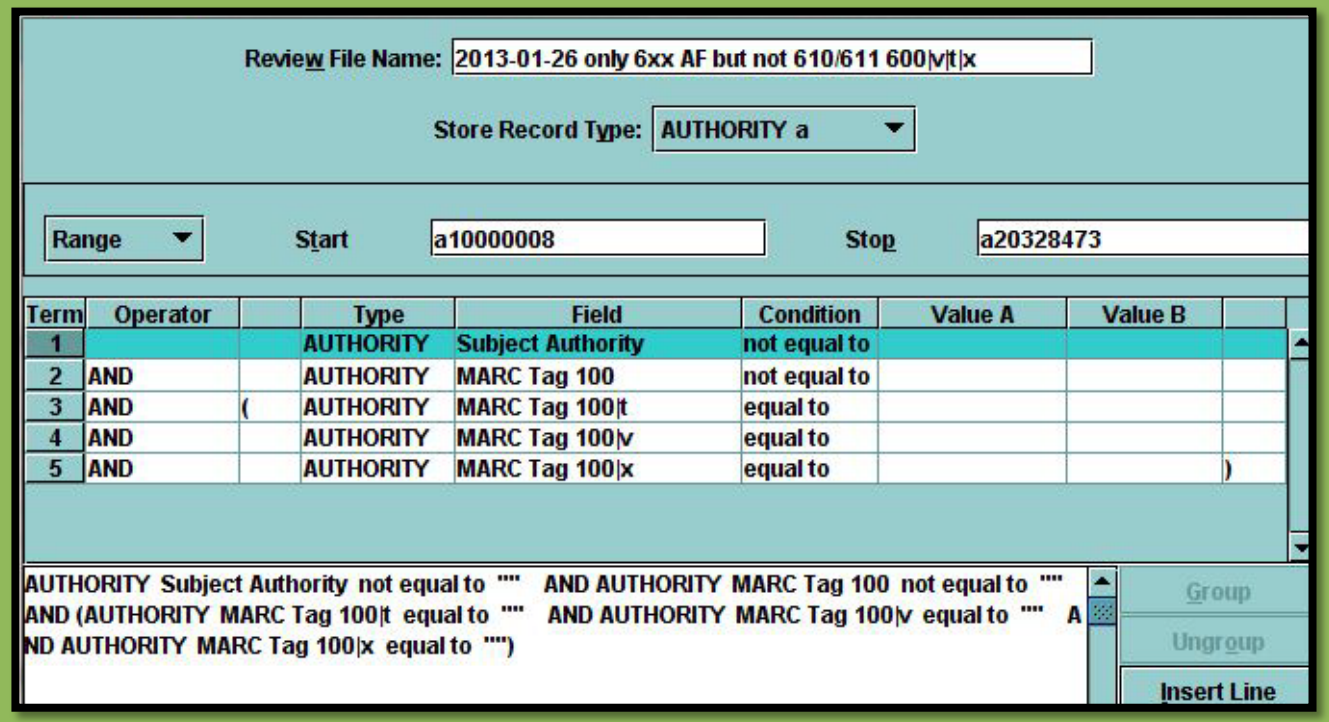


Export 12 data elements to spreadsheet

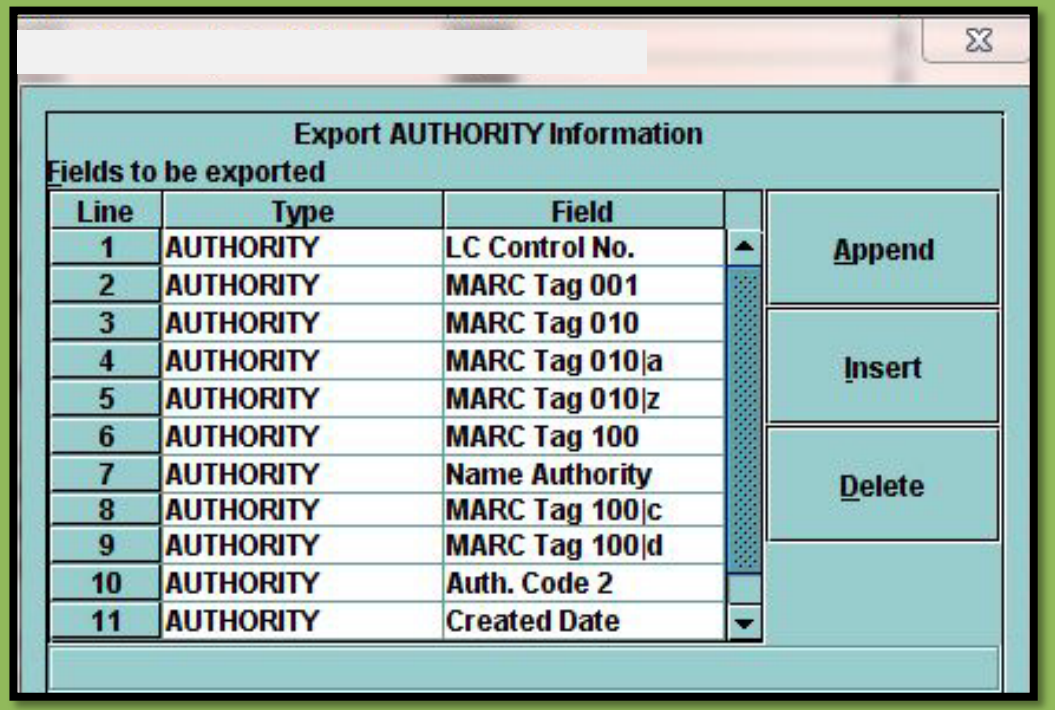


Review List for 6xx Name Heading as Subject

Search strategy:
6xx but not 610/611; 600|t, |v, or |x

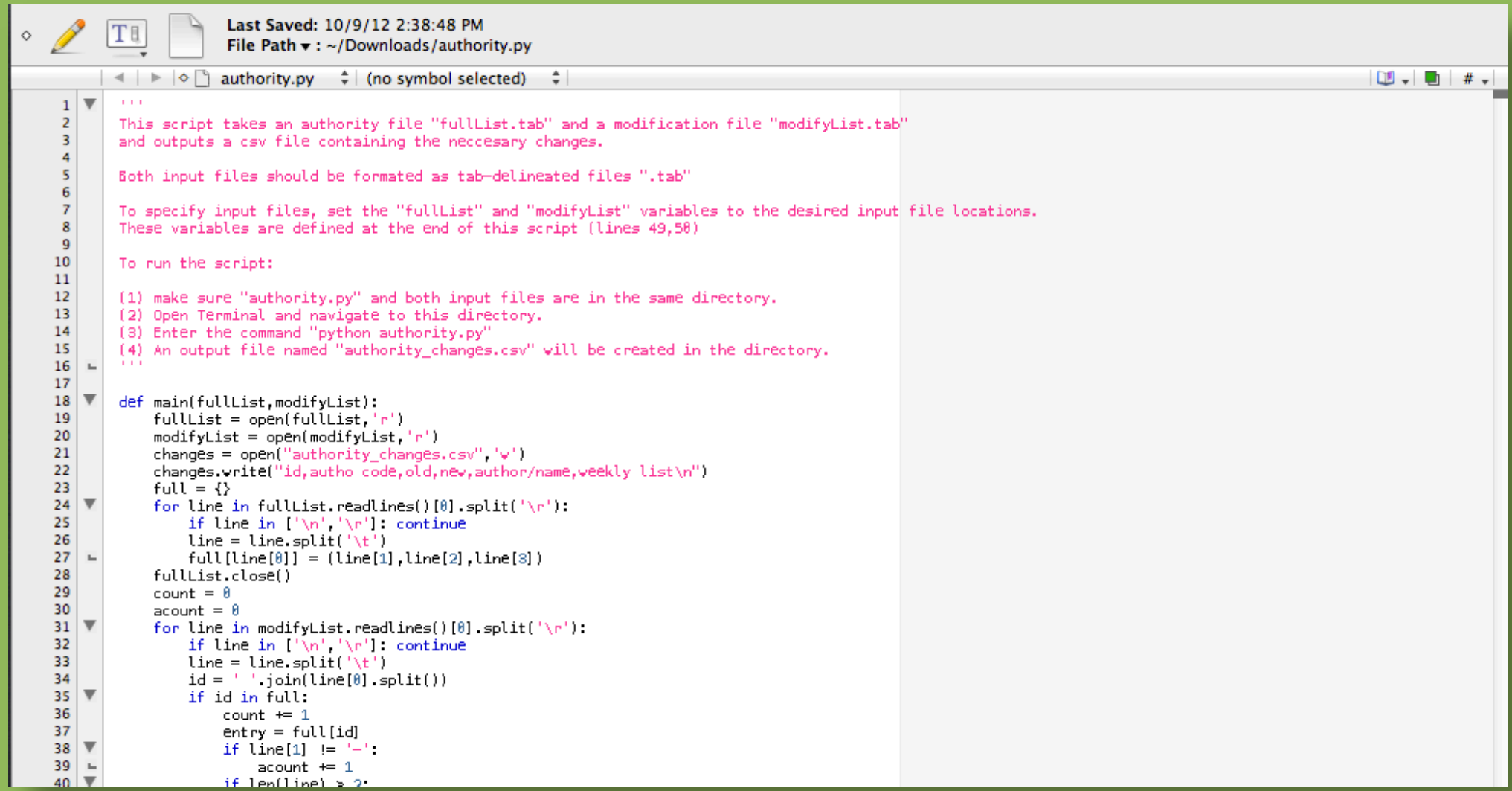


Export 12 data elements to a spreadsheet



<-

Python Script (a snippet)



Excel Spreadsheets

Formulae:

- Determine final digit - =RIGHT(E6,1)
- Remove period - =LEFT(E5,LEN(E5)-1)
- Compare cells - =(N5=P5) returns TRUE or FALSE
- Sort spreadsheet by cell color

A sample of the resulting spreadsheet

id	autho code	old	new	author/name	period removed	weekly list	final digit	final digit of correct heading	STATUS				
n 81083810	-	Marshall, E. G., 1910-	Marshall, E. G., 1910-1998	Marshall, E. G., 1914-1998		20060322	8		NOT FIXED	1910-1998	1914-1998	1910-1998	FALSE
n 85378827	-	Fessier, Michael, 1907-	Fessier, Michael, 1907-1989	Fessier, Michael, 1907-1988.	Fessier, Michael, 1907-1988	20060301	8		NOT FIXED	1907-1989	1907-1988	1907-1989	FALSE
n 80094212	-	Mayne, John Dawson, 1828-1917	Mayne, John D. (John Dawson), 1828-1917.	Mayne, John Dawson, 1828-1917		20101212	7		NOT FIXED	828-1917.	1828-1917	828-1917.	FALSE
n 86869628	a	Herbert, F. Hugh (Frederick Hugh), 1897-	Herbert, F. Hugh (Frederick Hugh), 1897-1952	Herbert, F. Hugh (Frederick Hugh), 1897-1958.	Herbert, F. Hugh (Frederick Hugh), 1897	20080203	8		NOT FIXED	1897-1952	1897-1958	1897-1952	FALSE
n 85183854	a	Helm, Brigitte, 1906-	Helm, Brigitte, 1906-1996	Helm, Brigitte, 1908-1996.	Helm, Brigitte, 1908-1996	20070808	6		NOT FIXED	1906-1996	1908-1996	1906-1996	FALSE
n 85017840	-	Ambrose, Thomas, 17/18th cent.	Ambrose, Thomas, 17th/18th cent.	Ambrose, Thomas, 17th/18th cent.	Ambrose, Thomas, 17th/18th cent	20080706	t		OK				
n 85099878	-	Culpepper, Nathanael, 17/18th cent.	Culpepper, Nathanael, 17th/18th cent.	Culpepper, Nathanael, 17th/18th cent.	Culpepper, Nathanael, 17th/18th cent	20080706	t		OK				
n 82137476	-	Wang, Yuanliang, 13th cent.	Wang, Yuanliang, 1241-ca. 1317	Wang, Yuanliang, 13th cent.	Wang, Yuanliang, 13th cent	20111204	t		OK				
n 82048664	a	Wallace, George C. (George Corley), 1919-	Wallace, George C. (George Corley), 1919-1998	Wallace, George C. (George Corley), 1919-1998.	Wallace, George C. (George Corley), 191	20060329	8		OK	1919-1998	1919-1998	1919-1998	TRUE
n 50011902	a	Thurmond, Strom, 1902-	Thurmond, Strom, 1902-2003	Thurmond, Strom, 1902-2003		20060329	3		OK	1902-2003	1902-2003	1902-2003	TRUE
n 50032462	a	Carmichael, Hoagy, 1899-	Carmichael, Hoagy, 1899-1981	Carmichael, Hoagy, 1899-1981.	Carmichael, Hoagy, 1899-1981	20060308	1		OK	1899-1981	1899-1981	1899-1981	TRUE
n 92106624	a	Culp, Robert, 1930-	Culp, Robert, 1930-2010	Culp, Robert, 1930-2010.;"Culp	Culp, Robert, 1930-2010.;"Culp	20100328	p		REVIEW				
n 00006682	-	Poulsson, Magnus, 1881-	Poulsson, Magnus, 1881-1958			20071107			SUBJECT?				
n 00030853	-	Koppel, Ted, 1940-	Koppel, Ted, 1940 Feb. 8-			20080525			SUBJECT?				
n 00032732	-	Jackson, David, 1922-	Jackson, David, 1922-2001			20100516			SUBJECT?				

<-

Connexion macro (a snippet)

OCLC Connexion Macro Editor and Debugger - [Tom!Add4xxAuth.sbl]

File Tools Window Help

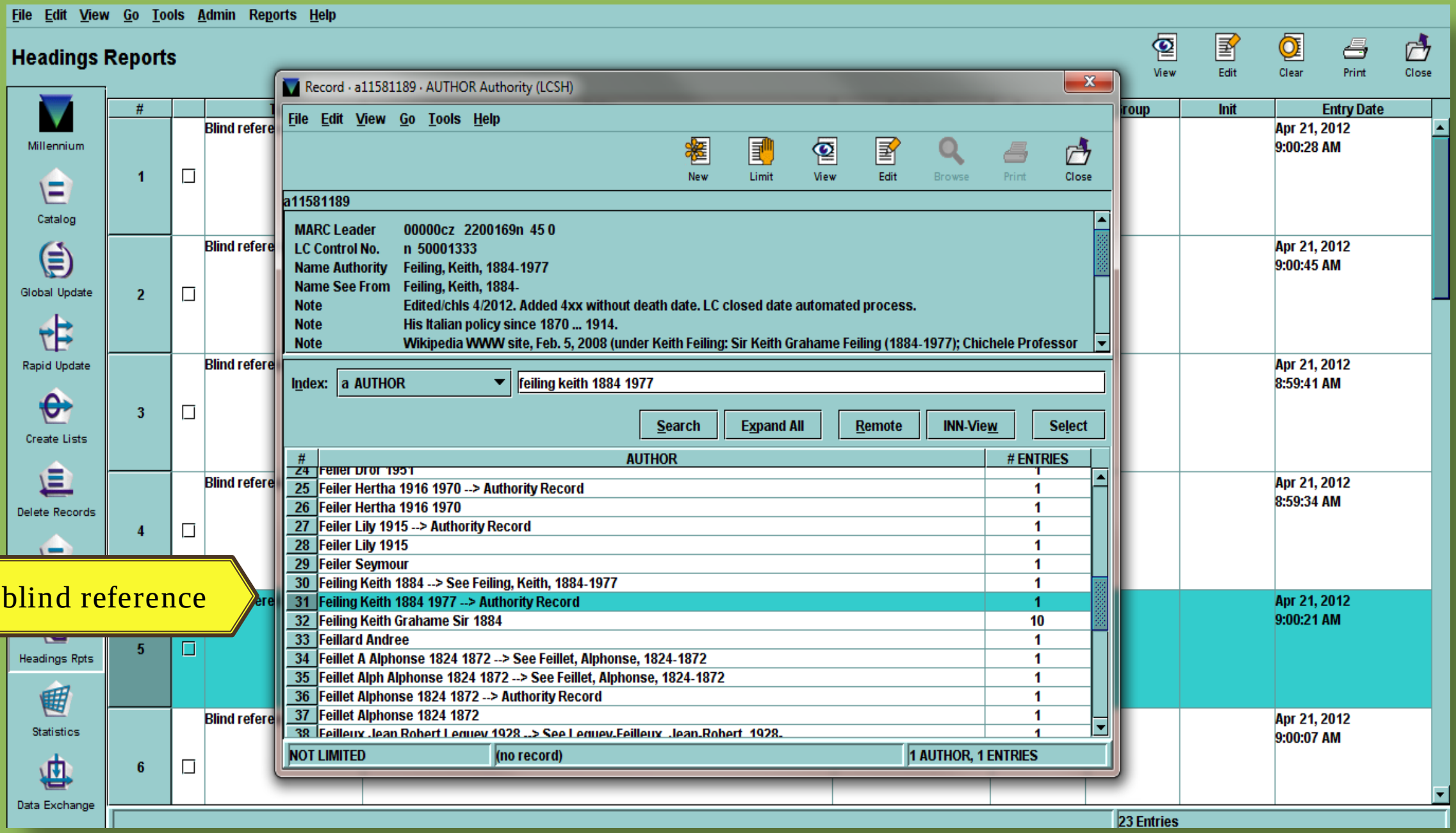
```
q = Val(Right(h,4))
If q = 0 Then
   MsgBox "The heading does not end with a date. Moving to the next record...", 64, "No Date"
   Goto PickNextRec
End If

' If the fifth character from the end is a hyphen, continue
' editing; otherwise, display a message (replacing delimiters with
' dollar signs) and move on to the next record

If Left(h, 1) <> "-" Then
   j = Trim(field100$)
   n% = Len(field100$)
   start% = 1
   Do
      finddlmt = Mid$(j, start%, 1)
      If finddlmt = delimiter$ Then
         Mid$(j, start%, 1) = "$"
      Else
         start% = start% + 1
      End If
   Loop Until start% = n% + 1
   MsgBox "The heading " & Chr(34) & j & Chr(34) & " does not use the YYYY-YYYY death date fo
   Goto PickNextRec
End If


' Remove the last four characters (death date) from the
' 100 field; copy and convert the abridged heading to a
' new 400 field

i = Len(field100$)
newfield400$ = Trim(Left(field100$, i-4))


' Replace the 100 tag with the 400 tag in new field

Mid$(newfield400$, 1, 3) = "400"


' Replace the 400 second indicator with a 9; the first indicator is ignored

Mid$(newfield400$, 5, 1) = "9"


' Add the 400 field to the record; if one or more 400 fields
' already exist, add the new field as the first one

bool400 = CS.AddField(1, newfield400$)
```

400 _9 Last, First, |dYYYY-

667 19 Edited/XXXX . Added 4xx without death date. LC closed date automated process.

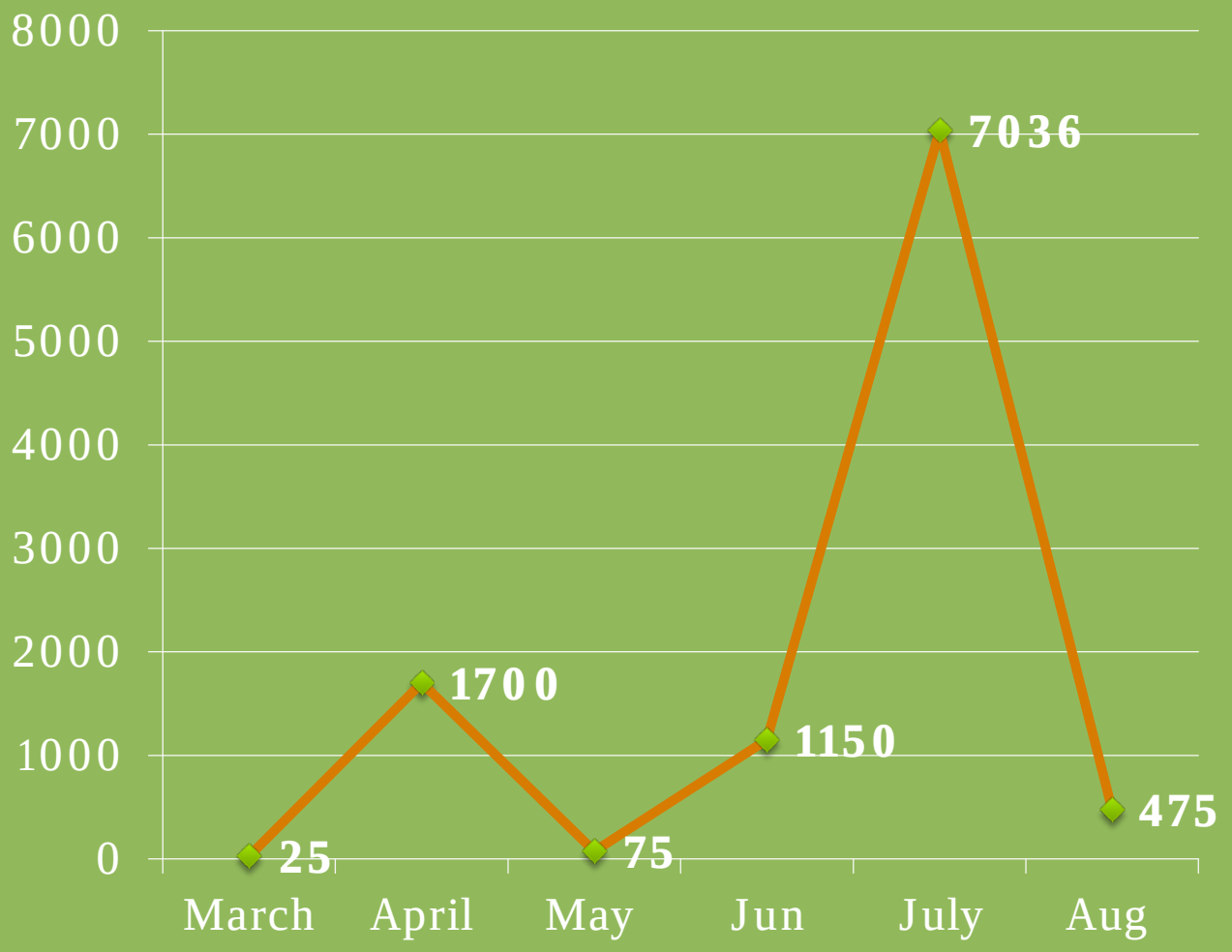
949 *atab=anam; ins=XXXX OR
949 *atab=asub; ins=XXXX

<-


File Edit View Go Tools Admin Reports Help
Headings Reports
View
Edit
Clear
Print
Close
Millennium
Catalog
Global Update
Rapid Update
Create Lists
Delete Records
Headings Rpts
Statistics
Data Exchange
Record · a11581189 · AUTHOR Authority (LCSH)
File Edit View Go Tools Help
New
Limit
View
Edit
Browse
Print
Close
a11581189
MARC Leader 00000cz 2200169n 45 0
LC Control No. n 50001333
Name Authority Feiling, Keith, 1884-1977
Name See From Feiling, Keith, 1884-
Note Edited/chls 4/2012. Added 4xx without death date. LC closed date automated process.
Note His Italian policy since 1870 ... 1914.
Note Wikipedia WWW site, Feb. 5, 2008 (under Keith Feiling: Sir Keith Grahame Feiling (1884-1977); Chichele Professor
Index: a AUTHOR
feiling keith 1884 1977
Search
Expand All
Remote
INN-View
Select
AUTHOR # ENTRIES
25 Feiler Hertha 1916 1970 --> Authority Record 1
26 Feiler Hertha 1916 1970 1
27 Feiler Lily 1915 --> Authority Record 1
28 Feiler Lily 1915 1
29 Feiler Seymour 1
30 Feiling Keith 1884 --> See Feiling, Keith, 1884-1977 1
31 Feiling Keith 1884 1977 --> Authority Record 1
32 Feiling Keith Grahame Sir 1884 10
33 Feillard Andree 1
34 Feillet A Alphonse 1824 1872 --> See Feillet, Alphonse, 1824-1872 1
35 Feillet Alph Alphonse 1824 1872 --> See Feillet, Alphonse, 1824-1872 1
36 Feillet Alphonse 1824 1872 --> Authority Record 1
37 Feillet Alphonse 1824 1872 1
NOT LIMITED
(no record)
1 AUTHOR, 1 ENTRIES
Init
Entry Date
Apr 21, 2012 9:00:28 AM
Apr 21, 2012 9:00:45 AM
Apr 21, 2012 8:59:41 AM
Apr 21, 2012 8:59:34 AM
Apr 21, 2012 9:00:21 AM
Apr 21, 2012 9:00:07 AM
23 Entries
A blind reference

The Numbers

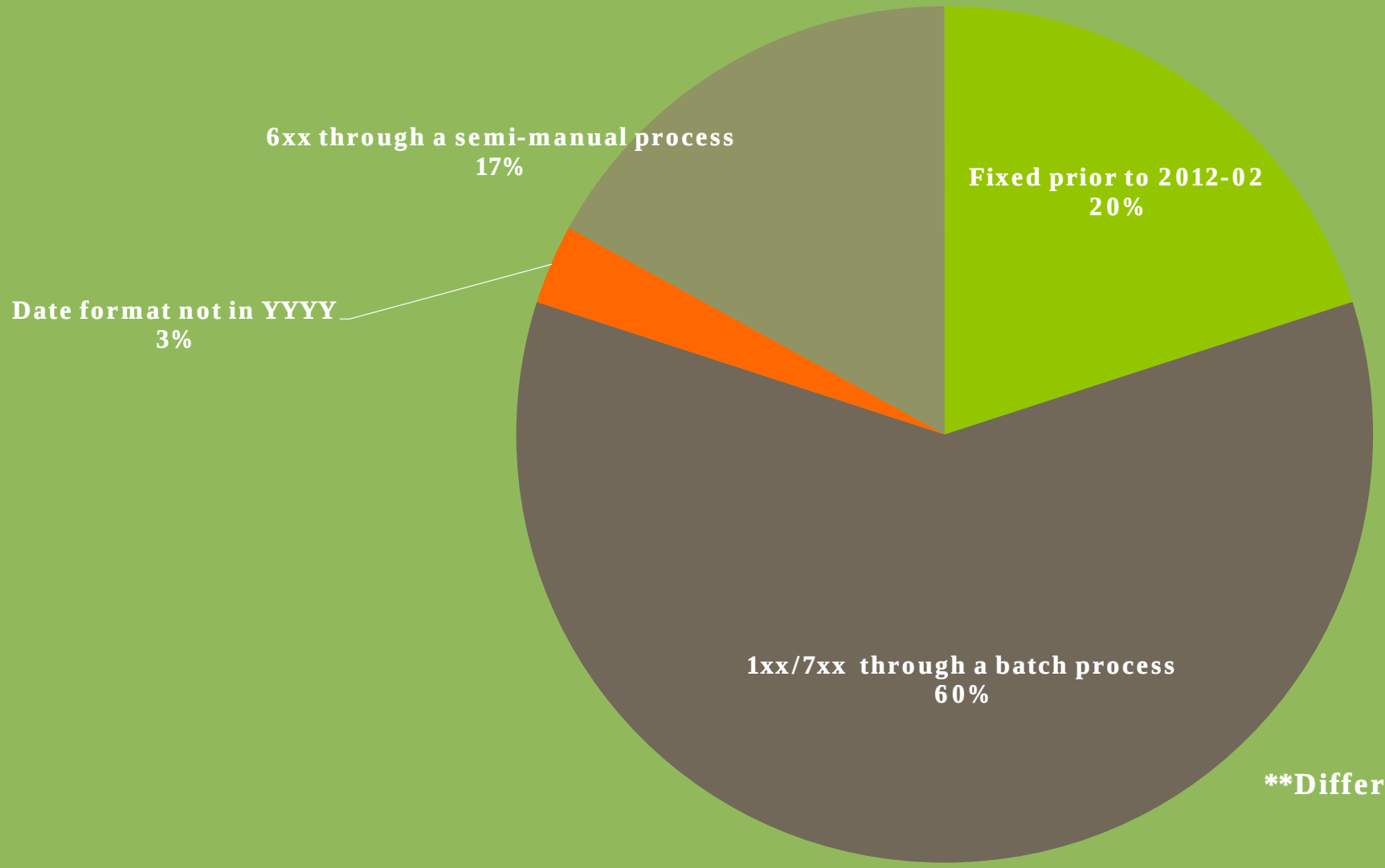
Data Sources	Numbers of Headings	Percentage
LC weekly list (02/2006 ~ 02/2012)	35,208	
Millennium AF list up to Feb. 2012 (including everything)	636,954+	
Matches between LC and our database generated by script, March 2012	16,859	
Identify duplicate headings from the spreadsheet (main heading and subject)	2,716	
LC updated headings in our database (16,859-2,716)	14,134	14,134/35,208=**40%**
An average cataloger at Carleton would have caught and manually updated, ***approximately***	600	600/14,134=**4.3%**



Name headings batch-loaded from Connexion to Millennium 60%

- Blind references on Day 2
  13% to fix
- Total bib updated through AACP:
  42,929+
- Average number of bibs per heading:
  4~6 bibs per AF
  (1~110 bibs per AF)

A Statistics View (16,859 matching names)



**Differences – 0.06%, 111 headings

Are we there yet?

Cycle incomplete. We addressed 1xx/7xx and 6xx.
Author/title added entry (240) not yet addressed.

010 no2011144865

100 1 Sagan, Françoise, |d 1935-
|tGarde du coeur.|lEnglish

010 n 79021390

100 1 Sagan, Françoise, |d 1935-2004

400 19 Sagan, Françoise, |d 1935-

One Bibliographic Record

100 1	Sagan, Françoise,|d1935-2004.
240 10	Garde du coeur.|lEnglish
245 14	The heart-keeper /|c[by] Françoise Sagan.
260	New York :|bDutton,|c1968.
300	128 p. ;|c21 cm.
600 10	Sagan, Françoise,|d1935-2004.
600 10	Sagan, Françoise,|d1935-2004|v[illegible]ecdotes.

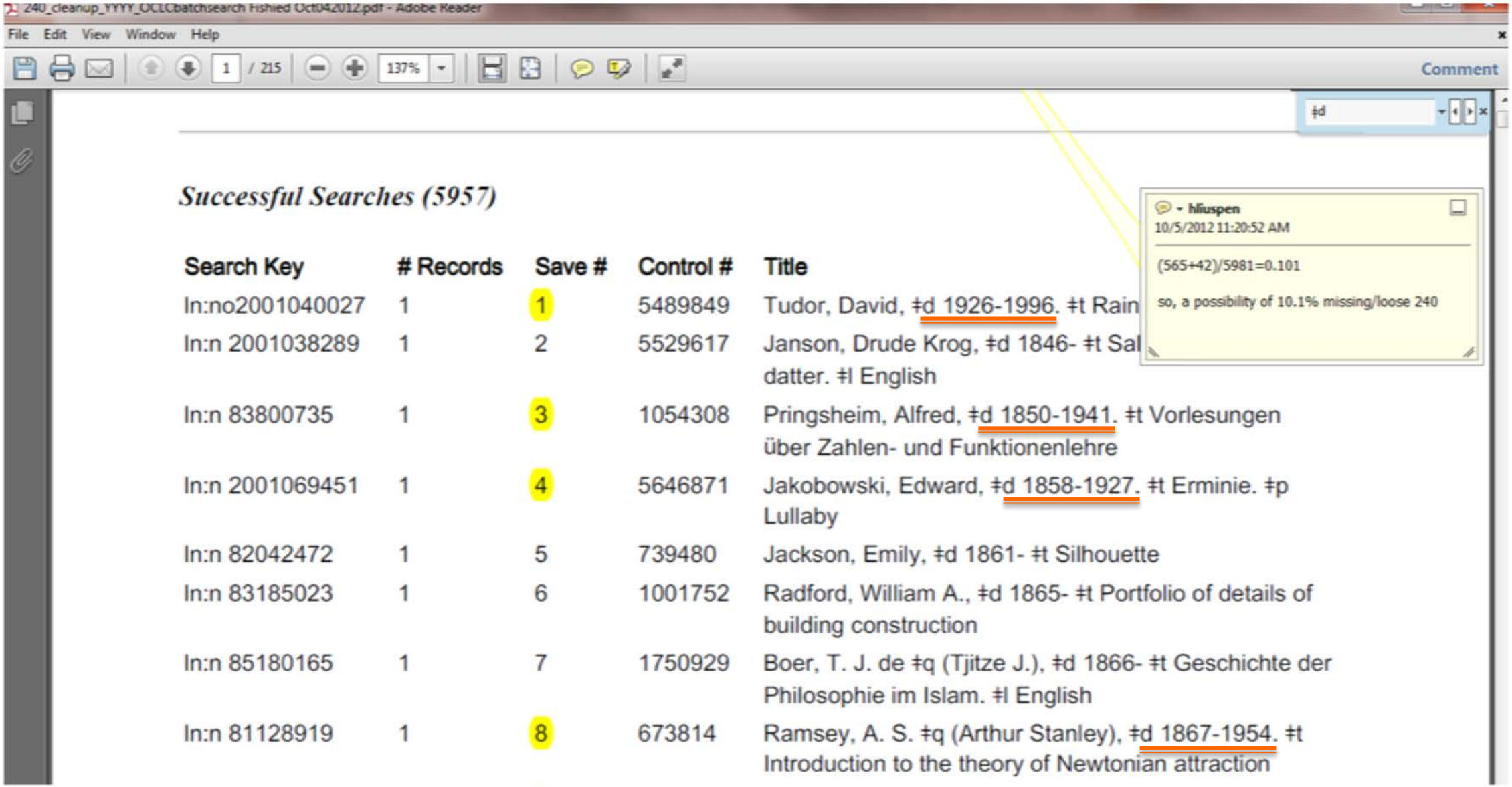
010 n 79021390

100 1 Sagan, Françoise, |d 1935-2004

400 19 Sagan, Françoise, |d 1935-

Author/Title Added Entry (100/240)

- Output authority record (240) from Millennium to a spreadsheet
- Sort the data elements by 1) YYYY-YYYY, 2) no date and 3) YYYY-
- Low-hanging fruit – names with the format of YYYY- and batch search in Connexion
- From the result list in Connexion, identify the date YYYY-YYYY (a closed date)



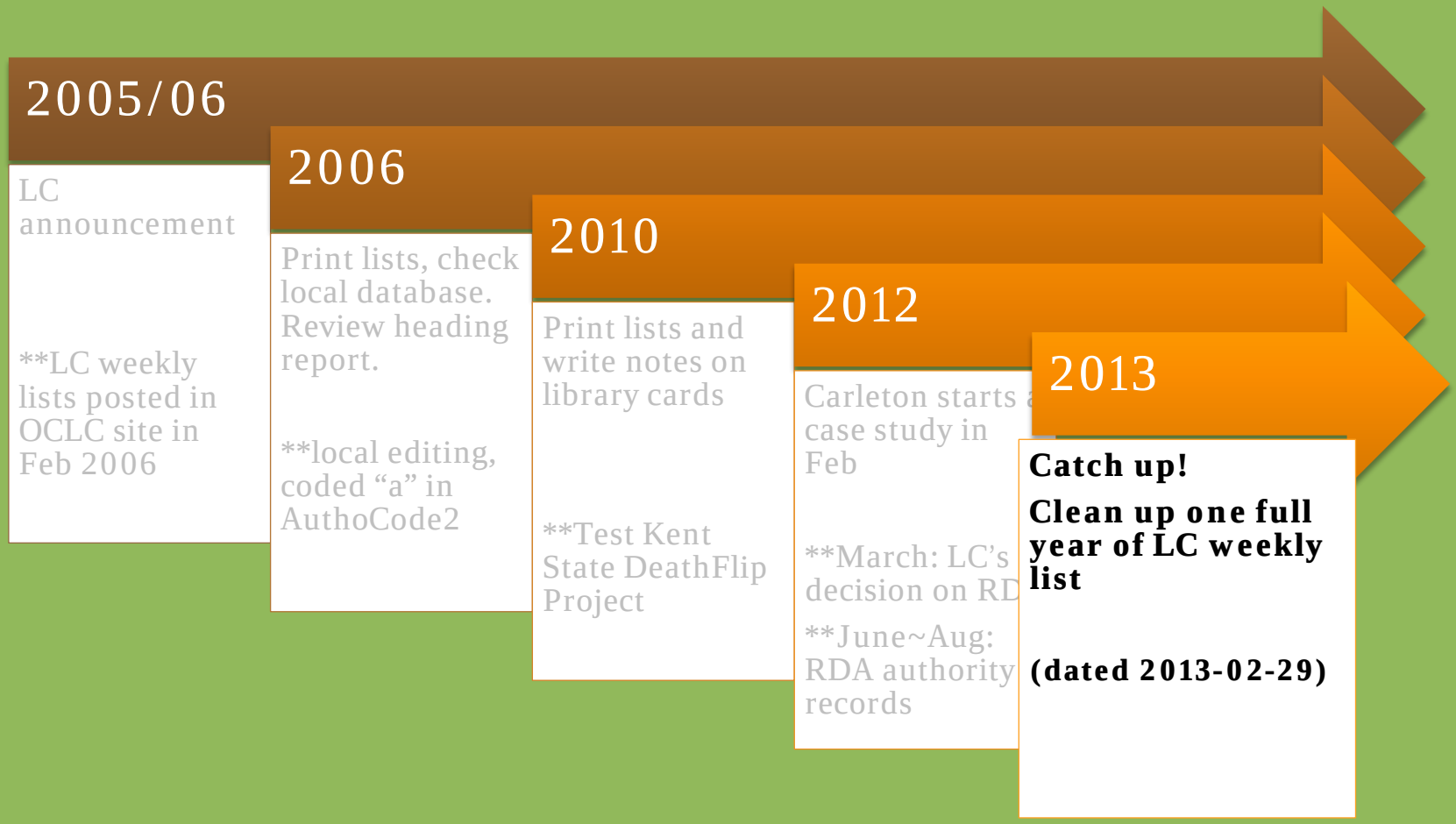
Successful Searches (5957)

Search Key	# Records	Save #	Control #	Title
In:no2001040027	1	1	5489849	Tudor, David, ‡d 1926-1996. ‡t Rain
In:n 2001038289	1	2	5529617	Janson, Drude Krog, ‡d 1846- ‡t Sal datter. ‡l English
In:n 83800735	1	3	1054308	Pringsheim, Alfred, ‡d 1850-1941. ‡t Vorlesungen über Zahlen- und Funktionenlehre
In:n 2001069451	1	4	5646871	Jakobowski, Edward, ‡d 1858-1927. ‡t Erminie. ‡p Lullaby
In:n 82042472	1	5	739480	Jackson, Emily, ‡d 1861- ‡t Silhouette
In:n 83185023	1	6	1001752	Radford, William A., ‡d 1865- ‡t Portfolio of details of building construction
In:n 85180165	1	7	1750929	Boer, T. J. de ‡q (Tjitze J.), ‡d 1866- ‡t Geschichte der Philosophie im Islam. ‡l English
In:n 81128919	1	8	673814	Ramsey, A. S. ‡q (Arthur Stanley), ‡d 1867-1954. ‡t Introduction to the theory of Newtonian attraction

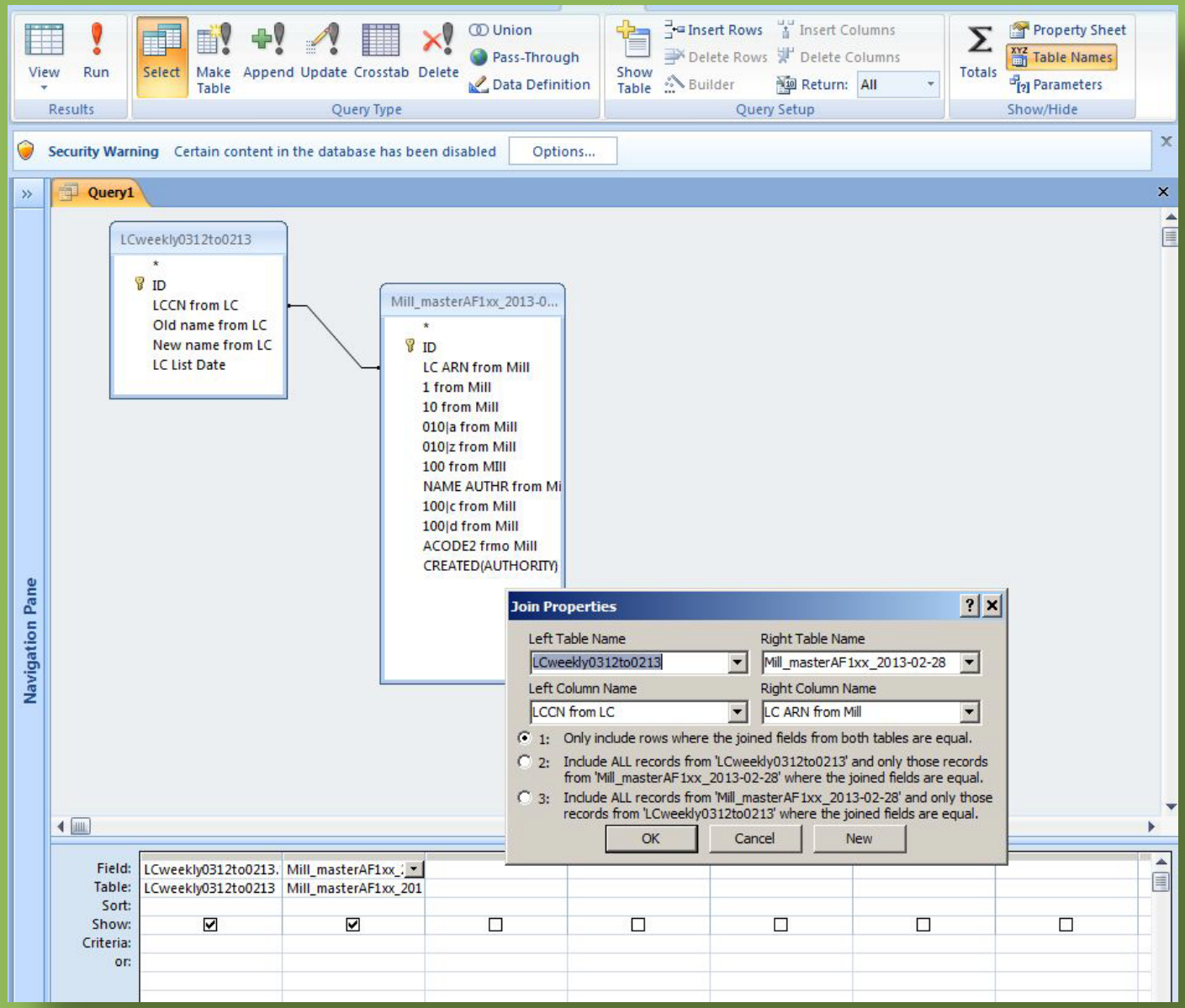
Keeping up – the basic steps

1. Scrape data from LC site (spreadsheet vs. RSS)
2. Output auth records from Millennium
3. Generate a list of matching LCCN
4. Review list to sort headings into categories
5. Batch export from Connexion into Millennium
6. Attend to next day Headings Reports
7. Address headings requiring manual attention
8. Address author/title headings

Bridge Consortium Reacts Over the Years



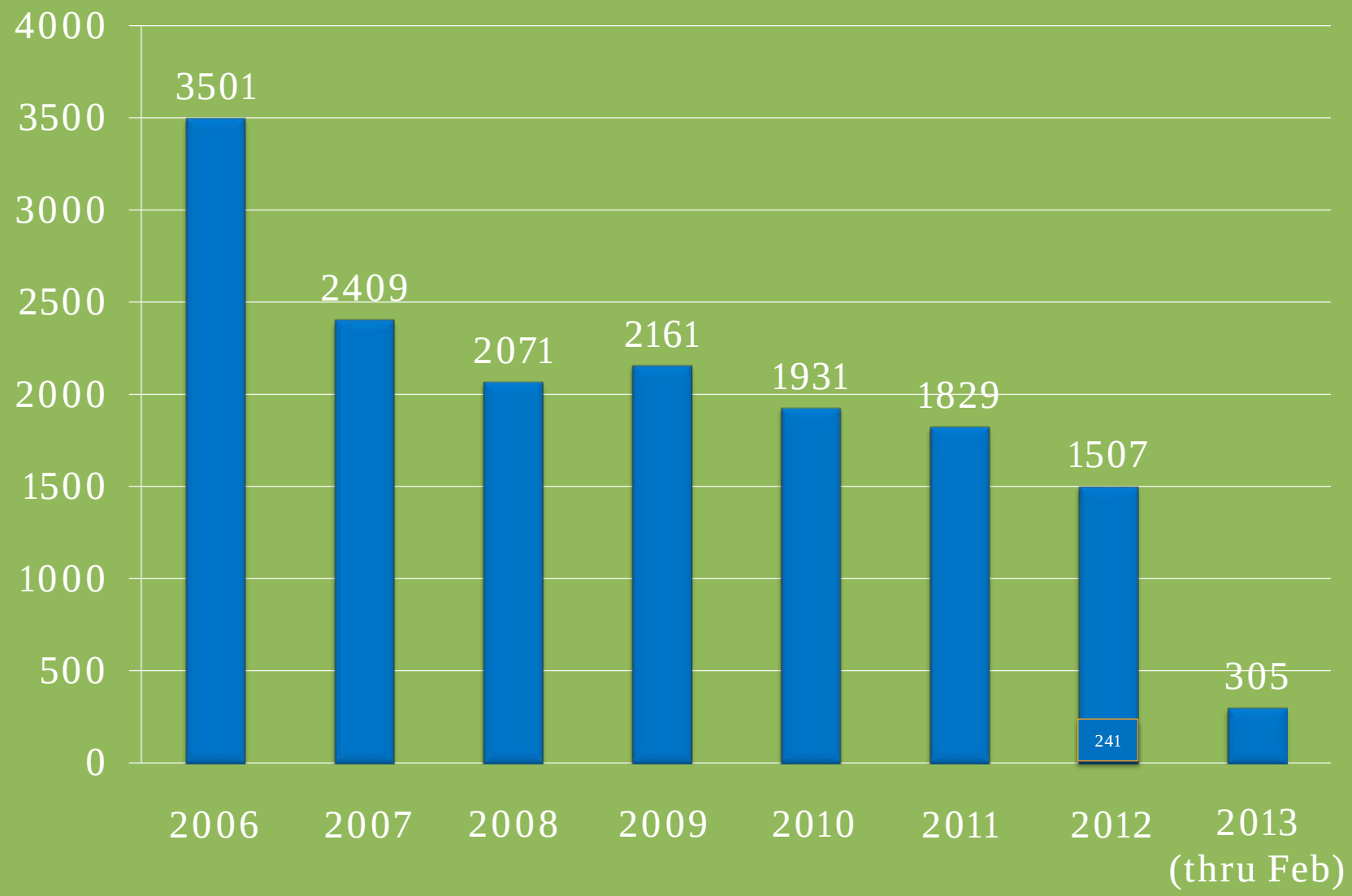
Data Comparison with Access



Numbers for the 7th year 03/2012 to 12/2013

Data Sources	Numbers of Headings	Percentage
LC weekly list (03/2012 ~ 02/2013)	4,503	**note: LC didn't publish any lists in August 2012
Millennium AF list up to Feb. 2013 (only contains 1xx and 6xx with no subfields x,v,t)	656,898+	
Matches between LC and our database generated by MS Access, March 2013	1,571	
LC updated headings in our database	1,571+	1,571/4,503=**34.88%**

Name Changes in Bridge Catalog



Lessons Learned

- Saving time?
- Taking advantage of OCLC batch searching and exporting
- Understanding the process of AACP in Millennium
- Subject template in Millennium load table: 949 *atab=asub
- Global update in 400 19 |wnnea Name
- Making sure to keep up with heading reports and clear up the space for loading more records
- Check the setting for AACP: Millennium III manual: #107824
    - Set this option to YES to enable the application of name authority record updates to name-title bibliographic headings

Our Goal

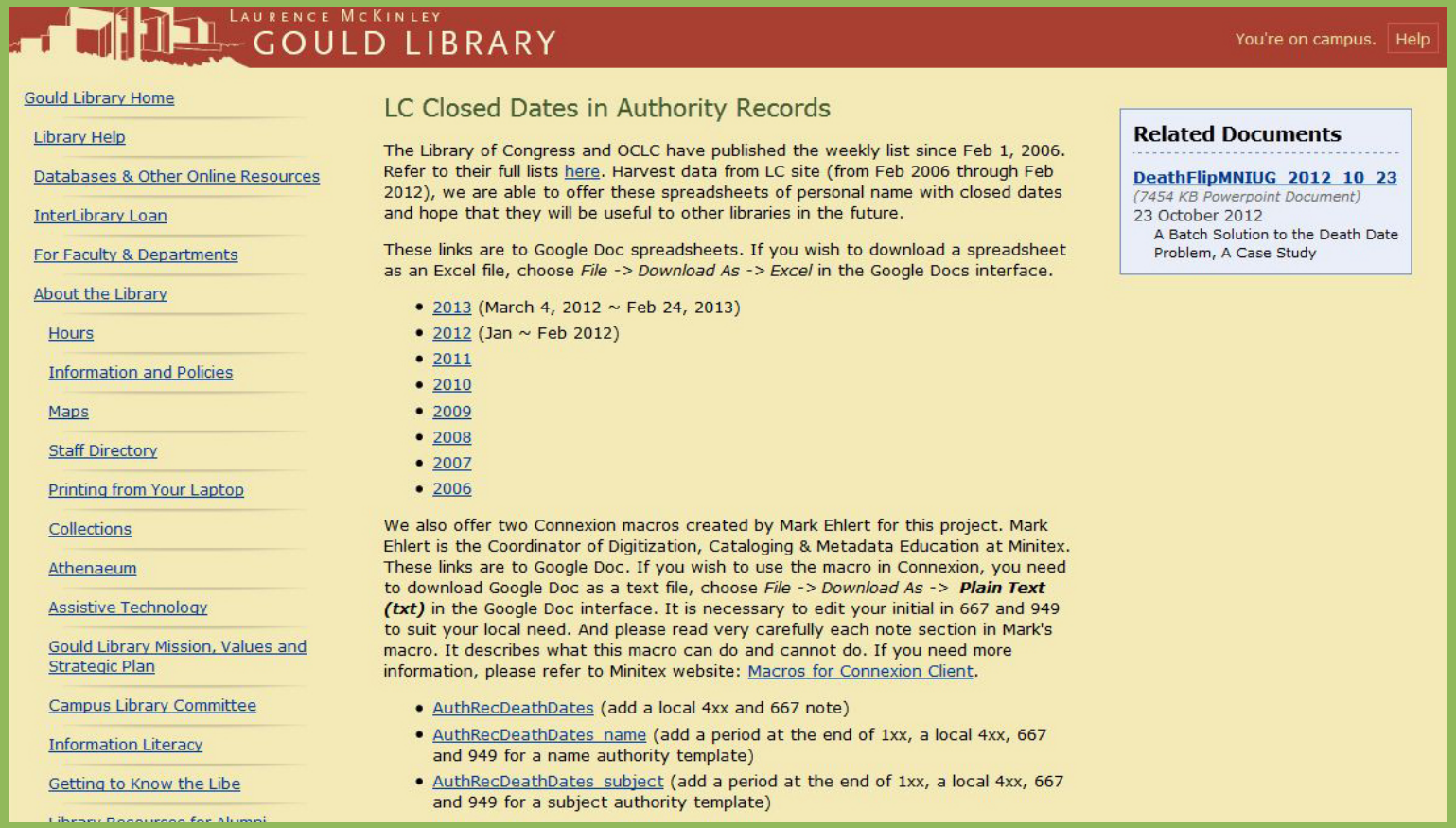
- Identify headings in our system that appear in the LC Weekly Closed Dates list and update those headings as needed.

We Need to Be Aware:

1. Not all identified headings require updating.
2. Identified headings may present unique challenges requiring unique approaches:
    - Heading for 1xx and 7xx
    - Heading for 6xx subject
    - Author/Title added entry & Uniform Title (240)

The Plan to Move Forward

- Keep up with the RDA changes to personal name headings
- Develop our own monthly schedule of Closed Date updating
- Share harvested data from LC Closed Dates Project **http://bit.ly/PQMkBU**



LAURENCE MCKINLEY GOULD LIBRARY

You're on campus. Help

Gould Library Home
Library Help
Databases & Other Online Resources
InterLibrary Loan
For Faculty & Departments
About the Library
Hours
Information and Policies
Maps
Staff Directory
Printing from Your Laptop
Collections
Athenaeum
Assistive Technology
Gould Library Mission, Values and Strategic Plan
Campus Library Committee
Information Literacy
Getting to Know the Libe

LC Closed Dates in Authority Records

The Library of Congress and OCLC have published the weekly list since Feb 1, 2006. Refer to their full lists here. Harvest data from LC site (from Feb 2006 through Feb 2012), we are able to offer these spreadsheets of personal name with closed dates and hope that they will be useful to other libraries in the future.

These links are to Google Doc spreadsheets. If you wish to download a spreadsheet as an Excel file, choose *File -> Download As -> Excel* in the Google Docs interface.

- 2013 (March 4, 2012 ~ Feb 24, 2013)
- 2012 (Jan ~ Feb 2012)
- 2011
- 2010
- 2009
- 2008
- 2007
- 2006

We also offer two Connexion macros created by Mark Ehlert for this project. Mark Ehlert is the Coordinator of Digitization, Cataloging & Metadata Education at Minitex. These links are to Google Doc. If you wish to use the macro in Connexion, you need to download Google Doc as a text file, choose *File -> Download As -> **Plain Text (txt)*** in the Google Doc interface. It is necessary to edit your initial in 667 and 949 to suit your local need. And please read very carefully each note section in Mark's macro. It describes what this macro can do and cannot do. If you need more information, please refer to Minitex website: Macros for Connexion Client.

- AuthRecDeathDates (add a local 4xx and 667 note)
- AuthRecDeathDates_name (add a period at the end of 1xx, a local 4xx, 667 and 949 for a name authority template)
- AuthRecDeathDates_subject (add a period at the end of 1xx, a local 4xx, 667 and 949 for a subject authority template)

Related Documents

DeathFlipMNIUG_2012_10_23
(7454 KB Powerpoint Document)
23 October 2012
A Batch Solution to the Death Date Problem, A Case Study

Thanks Go To

Kathy Blough

St. Olaf Music Cataloger

Jason Cohn

Student Worker in Carleton Archives

Nat Wilson

Digital Archivist & Technology Coordinator, Carleton College

Mark Ehlert

Coordinator of Digitization, Cataloging & Metadata Education, Minitex

Credit Also Goes To

Susanne Nevin

St. Olaf Cataloger

Sue Ims

Carleton College Cataloger

JoEllen LaPrade

St. Olaf Cataloger

Q & A

Thank you!

Tom Lamb tlamb@carleton.edu
Hsianghui Liu-Spencer hliuspen@carleton.edu