IUG
2018
ORLANDO

THE “X” (linked) and “Y” FACTORS

an outline of CU Boulder’s participation in the GPO Preservation Stewardship program, focusing on the challenges of exporting data for bound-with items

Laura Wright
Serials Cataloging Manager
University of Colorado Boulder

Outline

- Overview of the GPO Preservation Stewardship program
  - the commitment
  - the collections
  - the technical requirements
- Workflow challenges
- Data export challenges
  - item-level data
  - messy bibliographic data
  - lots of x-links
- Workarounds, failures, and tentative solutions

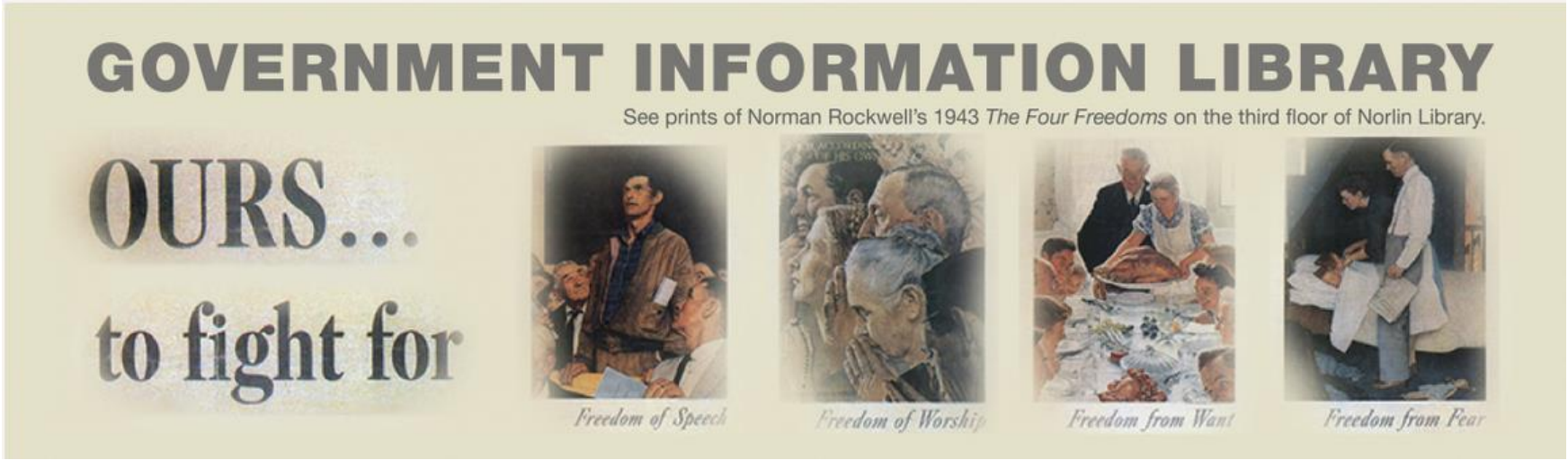
Thanks to Kate Tallman, Leanne Walther, Phil Wrede, Sheila Zoller, Carole Griffin, Megan Lambert, Lauren Stapleton, and numerous student employees

Regional Depositories
Shall retain Government publications permanently in either printed form or in microfacsimile form, except superseded publications or those issued later in bound form which may be discarded as authorized by the Superintendent of Documents (§1911).

...

Depository libraries must make Federal Depository Library Program content available for free use by the general public including all people in their relevant region and Congressional District.

FDLP
Federal Depository Library Program



IUG 2018 ORLANDO

GPO* Preservation Stewardship Program

Preservation Stewards make a commitment to retain specified depository resources for the length of the agreement. They also take on additional responsibilities for preserving depository materials that are preservation copies of record. This includes preventive preservation, or activities that prevent (or delay) materials from degrading or becoming damaged.

Four geographically dispersed Preservation Stewards are required for regional discard.

*U.S. GOVERNMENT PUBLISHING OFFICE

As a Preservation Steward, CU Boulder made a commitment to retain its **Congressional Hearings** collection, **U.S. Serial Set**, and **Bound Congressional Record**

Or from professions, vocations, businesses, trade, commerce, or sales or dealings in property, whether real or personal, growing out of the ownership or use of or interest in real or personal property.

I was led to offer this amendment largely on account of a colloquy I had with the Senator from Mississippi [Mr. WILLIAMS] the other day, who seems to have become indifferent and who does not regard the matter as worthy of his attention or presence. I recall, however, the Senate to the colloquy that I mentioned a moment ago. I asked this question:

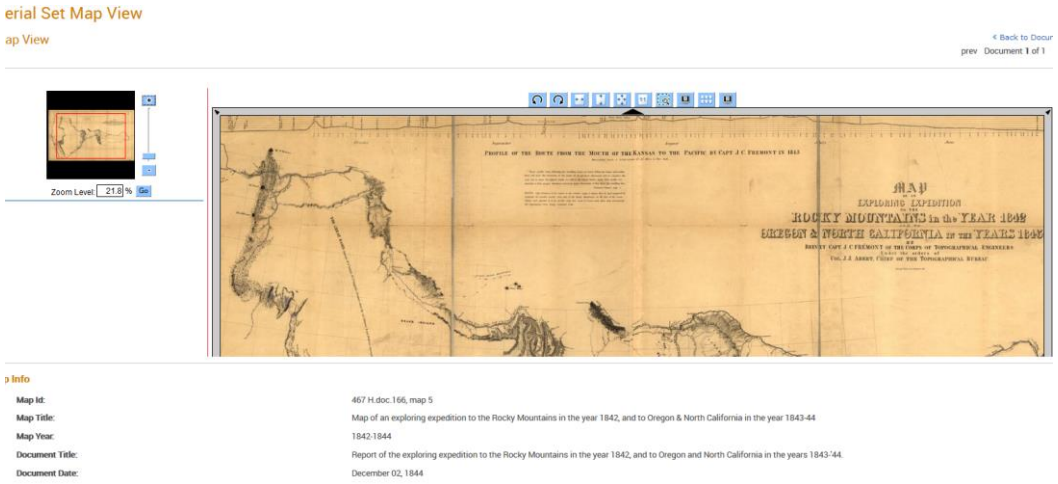
The Senator from Mississippi must certainly understand what I am trying to say. If applied to a general business, in which purchases and sales take place and gains and profits are reckoned. I can very well understand that the Senator from Mississippi is right, under the language of this bill. But suppose 10 years ago I had bought a horse for $1000, and this year I had sold him for $1,000, what would I do in the way of making a return?

Mr. WILLIAMS. I will tell the Senator precisely what he would do.

Mr. CUMMINS. I mean, what would other men do?

Mr. WILLIAMS. I know; but what I mean is precisely what the Senator would do, or precisely what he ought to do. He bought the horse 10 years ago and sold him this year for a thousand dollars. That thousand dollars is a part of the Senator's receipts for this year, and being a part of his receipts, that much will go in as part of his receipts, and from it would be deducted his disbursements and his exemptions and various other things.

Mr. CUMMINS. Would the price I paid for the horse originally be deducted?



And we got some nice press coverage

University of Colorado, Boulder: GPO's First Preservation Steward

Details
Last Updated: December 12 2016
Published: December 07 2016

On October 18, during the 2016 Depository Library Council Meeting & Federal Depository Library Conference, Laurie Hall, the U.S. Government Publishing Office's (GPO's) Acting Superintendent of Documents, signed GPO's first Preservation Steward agreement with the University of Colorado, Boulder (CU, Boulder). Depository Coordinator Kate Tallman represented CU, Boulder at the signing. As a Preservation Steward, CU, Boulder made a commitment to retain its Congressional Hearings collection, U.S. Serial Set, and Bound Congressional Record. The library has also taken on the additional responsibilities for preserving that material.

Among the highlights in the CU, Boulder Preservation Steward Collection are the set of Western exploration surveys issued as part of the Serial Set. Volume 2 of the Report on the United States and Mexican Boundary (1852) contains color illustrations that Kate Tallman says, "look like they were printed just yesterday." The collection also contains the reports of the Hayden, King, and Powell Surveys along with the accompanying drawings of flora, fauna, and landscapes. The drawings from the Powell Survey are some of the earliest illustrations of the Grand Canyon.

View full archive (2010-present)

Subscribe to the FDLP Connection

Enter your name and email address to subscribe to our newsletter.

Name

E-mail

Subscribe

Division of labor

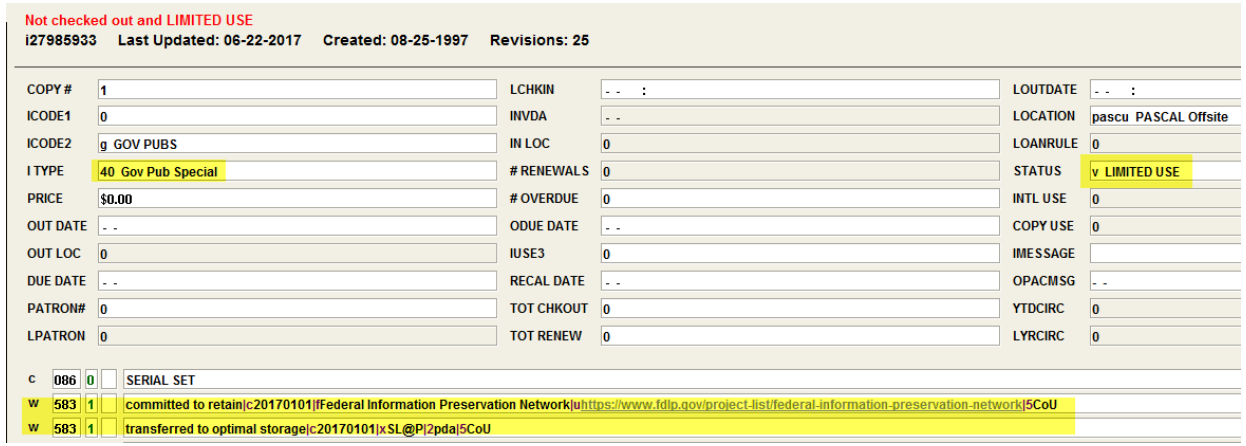
Government Information: identify materials, process and route materials, update individual bibliographic and item records (as new material is received)

Preservation: evaluate materials, create enclosures, take and document actions according to PDA* standards

Metadata Services: globally update item and bibliographic records, create item-level holdings list for GPO

* MARC 583 $2 Preservation and Digitization Actions

Item records



Not checked out and LIMITED USE

i27985933 Last Updated: 06-22-2017 Created: 08-25-1997 Revisions: 25

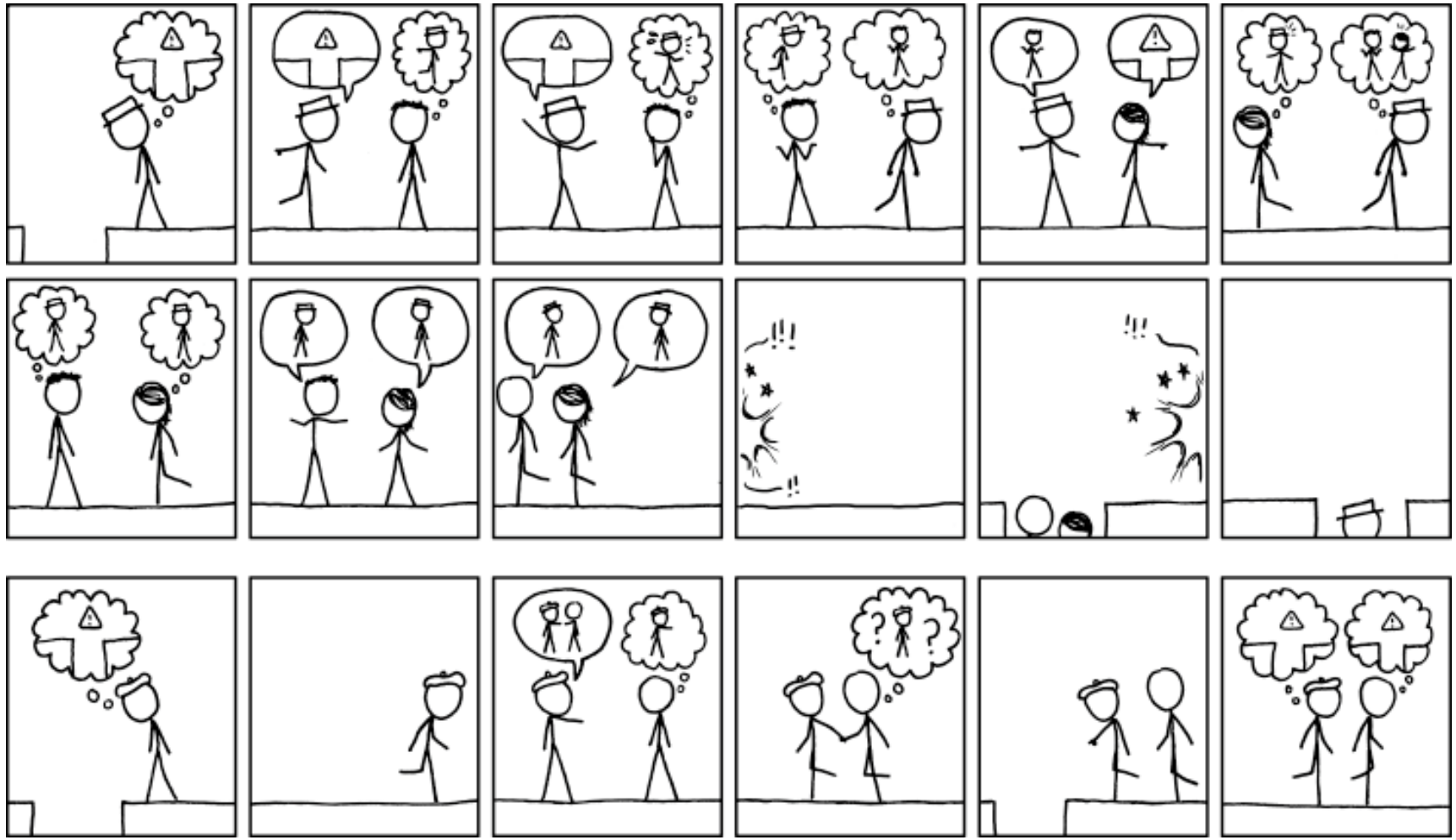
COPY #	1	LCHKIN	- - :	LOUTDATE	- - :
ICODE1	0	INVDA	- -	LOCATION	pascu PASCAL Offsite
ICODE2	g GOV PUBS	IN LOC	0	LOANRULE	0
I TYPE	40 Gov Pub Special	# RENEWALS	0	STATUS	v LIMITED USE
PRICE	$0.00	# OVERDUE	0	INTL USE	0
OUT DATE	- -	ODUE DATE	- -	COPY USE	0
OUT LOC	0	IUSE3	0	IMESSAGE	
DUE DATE	- -	RECAL DATE	- -	OPACMSG	- -
PATRON#	0	TOT CHKOUT	0	YTDCIRC	0
LPATRON	0	TOT RENEW	0	LYRCIRC	0

c	086	0		SERIAL SET
w	583	1		committed to retain|c20170101|fFederal Information Preservation Network|uhttps://www.fdlp.gov/project-list/federal-information-preservation-network|5CoU
w	583	1		transferred to optimal storage|c20170101|xSL@P|2pda|5CoU

Bibliographic records (local note)

|3Documents in PASCAL: |aFIPNet Preservation copy of record in partnership with the US Government Publishing Office. Electronic version is available, email govpubs@colorado.edu for help locating a digital copy

Workflow challenges

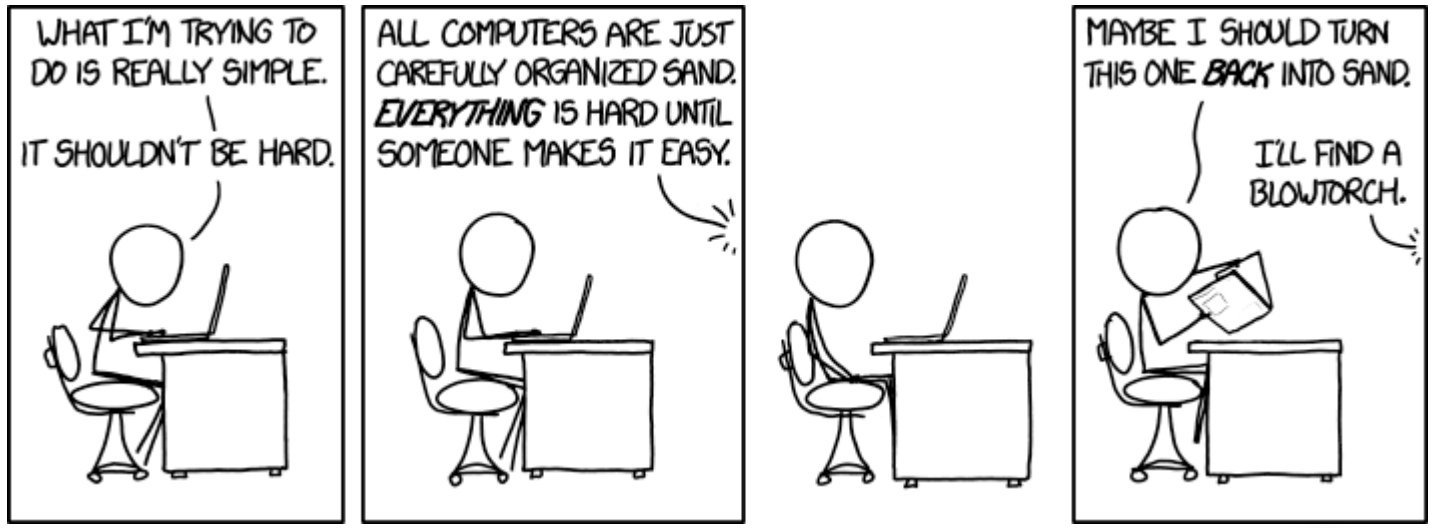


https://xkcd.com/1028/

Item level data

GPO needs **item level** information about our materials (as illustrated on the right). This information comes from both bibliographic and item records.

Title	SuDoc Number	Publication Date(s)	Volume, Issue, or Date	Format	Condition	Storage Location	Access	OCLC Number	Local System Number	Notes
b	b	b	i					b	i	
MARC 245	MARC 086 $a	008	Volume (v)					MARC 001	Record number	

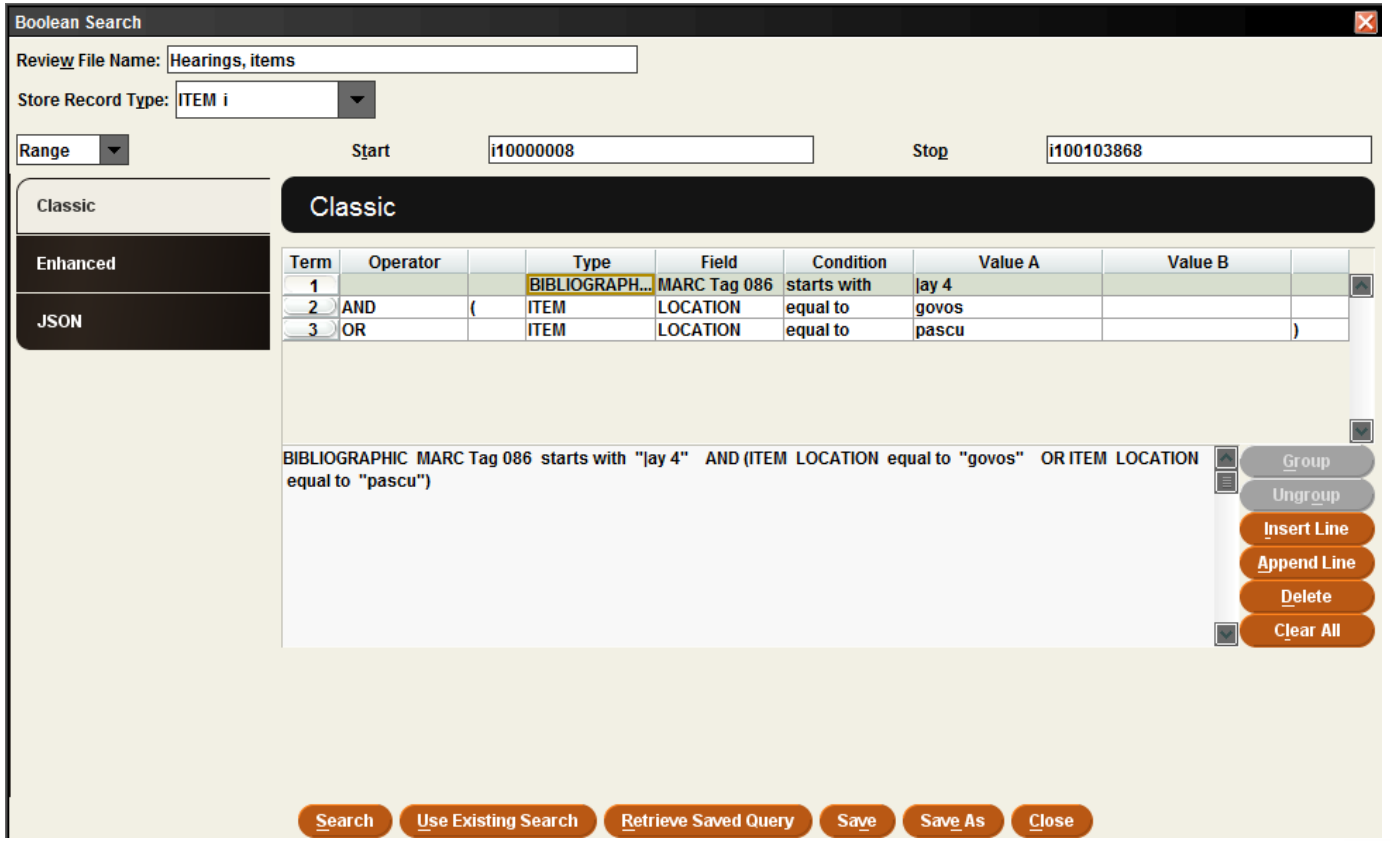


https://xkcd.com/1349/

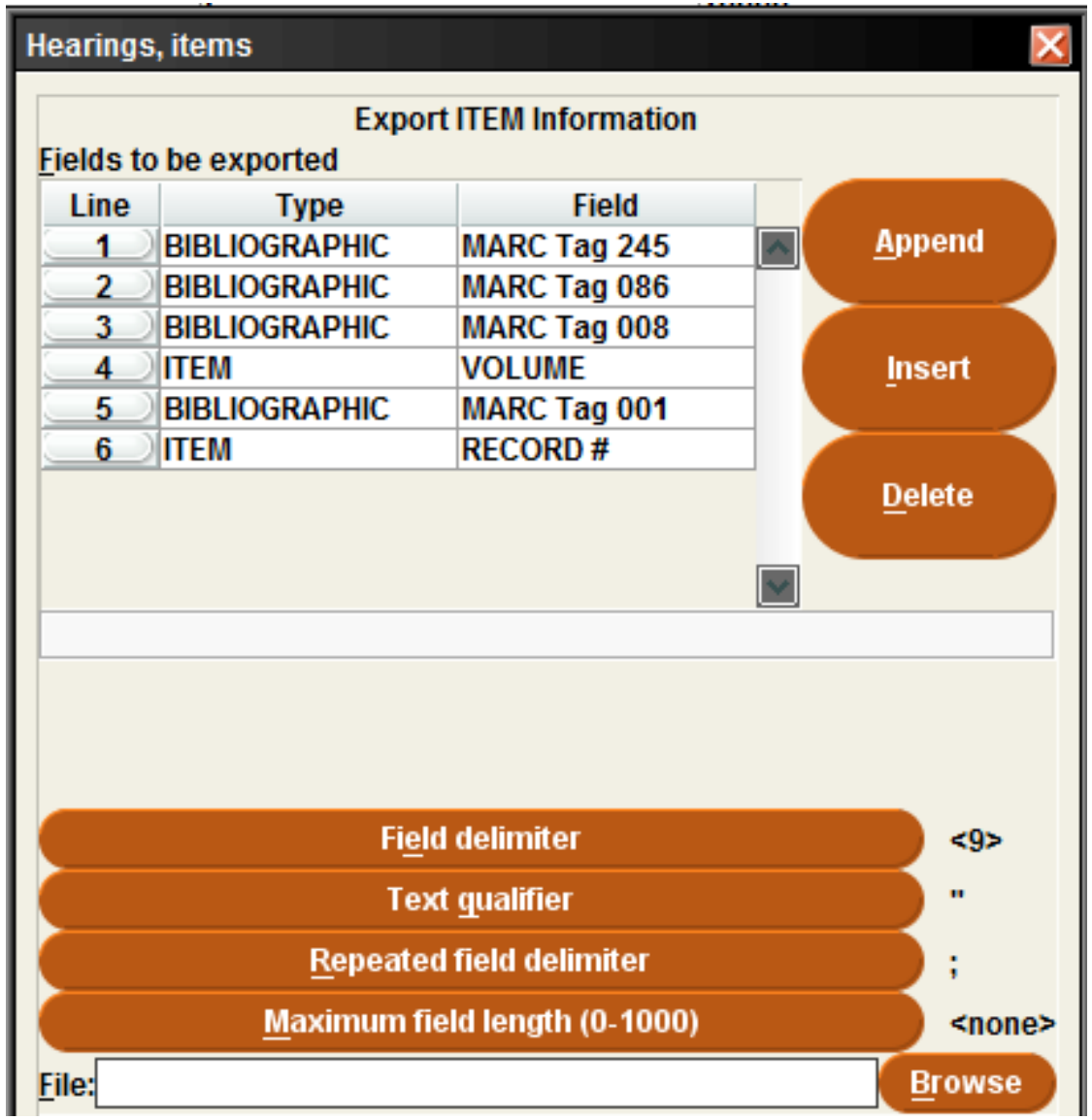
Search Strategy (Sierra Create Lists)

All hearings are assigned a SuDoc number that starts with “Y 4”

The location limits made sure we were getting only print items (we have many of these in microform and online as well)



Export



- Bib and item data
- Tab delimited

Initial results looked OK… at first…

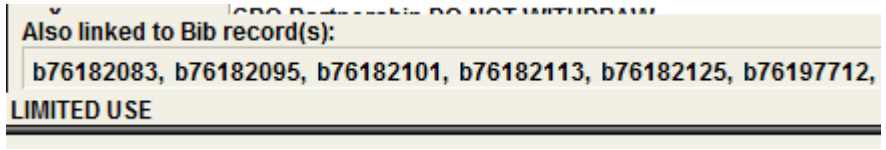
	A	B	C	D	E	F
1	245	086	008	VOLUME	001	RECORD #(ITEM)
2	The economic effects of fran	Y 4.Sm 1/2:F 84/2;"Y 4.Sm 1/2:F	711112s1971 dcu f000 0 eng cam		204933	i11738248
3	Hearing on the reauthorizati	Y 4.ED 8/1:103-55;"Y 4.ED 8/1:10	940708s1994 dcua b f000 0 eng dnam a		30725803	i20802511
4	Domestic violence : the strug	Y 4.Ap 6/2:S.HRG.102-45;"Y 4.Ap	910716s1991 dcua f000 0 eng dnam a		24083144	i20853762
5	Endangered Species Conserv	Y 4.C 73/2:92-81;"Y 4.C 73/2:92-8	730205s1972 dcu f000 0 eng nam		578668	i20912973
6	Predatory mammals and end	Y 4.M 53:92-22;"Y 4.M 53:92-22"	720822s1972 dcua 000 0 eng cam		416211	i20912997
7	Endangered species : hearin	Y 4.M 53:93-5;"Y 4.M 53:93-5"	730807s1973 dcu f000 0 eng cam a		696921	i2091300x
8	Hearings on H.R. 3160, the Co	Y 4.Ed 8/1:102-99;"Y 4.Ed 8/1:102	920904m19929999dcua b f000 0 eng dr	v.2	26535289	i2093810x
9	Hearings on H.R. 3160, the Co	Y 4.Ed 8/1:102-99;"Y 4.Ed 8/1:102	920904m19929999dcua b f000 0 eng dr	v.1	26535289	i20938111
10	Hearings on H.R. 1280, Comp	Y 4.ED 8/1:103-35;"Y 4.ED 8/1:10	940301s1994 dcua b f000 0 eng dnam a		29887223	i20938135
11	Poverty, hunger, and the we	Y 4.H 89:99-23;"Y 4.H 89:99-23"	870301s1987 dcua b f000 0 eng dcam a		15271852	i20947355
12	Redesigned Space Station Pr	Y 4.C 73/7:S.HRG.103-283;"Y 4.C	940112s1993 dcu f000 0 eng dnam a		29613705	i20947835
13	Comprehensive health care i	Y 4.L 11/4:S.HRG.103-173;"Y 4.L 1	931007s1993 dcua b f000 0 eng dnam a		28952804	i20948384
14	Coverage for mental and adc	Y 4.L 11/4:S.HRG.103-211;"Y 4.L 1	931026s1993 dcu b f000 0 eng dnam a		29195099	i20948396
15	Hearing on the Child Abuse F	Y 4.Ed 8/1:102-96;"Y 4.Ed 8/1:102	920818s1992 dcua b f000 0 eng dnam a		26427784	i21133955
16	Review of federal measurem	Y 4.P 84/10:103-7;"Y 4.P 84/10:10	940325s1994 dcua bs f000 0 eng dnam a		30021858	i21141289
17	Implementation of Indian Ga	Y 4.R 31/3:103-17/;"Y 4.R 31/3:10	931029m19939999dcua f000 0 eng dca	pt.4	29214637	i21169706
18	Indian economic developme	Y 4.R 31/3:103-18/;"Y 4.R 31/3:10	930924m19939999dcuab f000 0 eng dr	pt.1	28883671	i21169743
19	Indian economic developme	Y 4.R 31/3:103-18/;"Y 4.R 31/3:10	930924m19939999dcuab f000 0 eng dr	pt.2	28883671	i21169767
20	Implementation of Indian Ga	Y 4.R 31/3:103-17/;"Y 4.R 31/3:10	931029m19939999dcua f000 0 eng dca	pt.2	29214637	i21169779
21	Implementation of Indian Ga	Y 4.R 31/3:103-17/;"Y 4.R 31/3:10	931029m19939999dcua f000 0 eng dca	pt.3	29214637	i21169792
22	Memorandum of agreement	Y 4.P 96/11:101-50;"Y 4.P 96/11:1	910117s1990 dcu f000 0 eng dnam a		22973936	i21176176
23	Oversight hearings on eleme	Y 4.ED 8/1:ED 8/39;"Y 4.ED 8/1:EI	710610s1971 dcuad f000 0 eng nam		142075	i2124313x
24	Fast track authority and Nort	Y 4.P 96/11:102-12;"Y 4.P 96/11:1	911105s1991 dcua b f000 0 eng dnam a		24678625	i21243451
25	Field hearings on H.R. 2878	Y 4.Ed 8/1:102-109;"Y 4.Ed 8/1:10	921021s1992 dcu b f000 0 eng dnam a		26826257	i21243475

hearings_items_example_export

Filtering revealed a huge mess

	A	B	C	D	E	F	G	H	I	J	K
1	2	086	008	VOLUME	001	RECORI	(ITEM)				
1070	United States congressional s	Annual re	Energy action no. DOE-002 : co	Annual report / Gorgas Memo	Reorganization plan no. 2 of 1978 : mess	Internatio	Extension	Rescissior	Federal w	Review of	Cumulat
1081	United States congressional s	Federal re	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.IN 8/4:R 26/8;"Y 4.IN 8/4:R	780512c18179999dcu x f0 0eng dcas	761216s19	1976	ucbserials	2628459	i2532374x	
1087	United States congressional s	Executive	Rescissions and deferrals : me	Puerto Rico statehood act of 1	Secret service protection for certain for	Rural deve	Airport nc	Annual re	Energy act	Energy act	Defense
1091	United States congressional s	Strategic	Proceedings of the Judicial Co	Review of deferrals : commun	Congressional Budget Office, a study of	Deferrals	The feasit	Federal el	Governing	Governing	Governi
1141	Honoraria amendments to th	Senate Ele	Telephone rates and procedur	Telecommunications Equipme	The Consumer Protection Against Price	Strengthe	Saguaro N	Morristow	Authorizir	Updating	Motor V
1142	The Federal Death Penalty Ac	Approving	Transportation Employee Test	Antiterrorism and Arms Expor	Authorizing appropriations for fiscal ye	Niobrara F	Legal issu	Idaho Fore	Certificate	Coast Gua	Certifica
1152	Medicare and Medicaid Patie	Excellence	Arms Control and Disarmamer	The Developmental Disabilitie	National Academies of Practice : report	Amending	Authoriza	Authorizir	Public Bui	Amending	Relative
1161	Timeshare Consumer Protect	Waiver of	School prayer constitutional a	Voluntary school prayer amen	Emergency Wetlands Resources Act of 1	Exclusion	Merit Pay	Robbery c	Refusal of	Illinois an	Commis
1162	United States congressional s	Maritime	National Science Foundation a	National Science Foundation a	Authorizing the issuance of gold medal:	Authorizir	Authorizir	Direct por	Activities	Special int	Railroad
1165	The Federal judicial system :	Immigrati	Report on the activities of the	Administrative			m	Antitrust a	Report of	History, ju	Report on	Report c
1333	United States congressional s	Compend	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.G 74/6:B 85/			19	1977	ucbserials	2738198	i26341116	
1360	United States congressional s	Corporate	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.G 74/6:C 81/			19	1976	ucbserials	2613764	i26414442	
1377	Legislative review activities c	Report on	Annual report of the Select Co	Legislative revi			ve	United Sta	Y 1.1/8:;"X	Y 1.1/8:;"X	Y 1.1/8:;"X	Y 4.P 84/
1478	United States congressional s	Summary	Report on the activity of the C	Report on the a			on	Report of	Survey of	Legislative	Activities	Summar
1517	United States congressional s	Families a	Legislative review by the Com	Legislative revi			y	Report on	Summary	Activities	SERIAL SE	Y 1.1/8:9
1529	Activities of the Aging Comm	Legislative	Legislative review by the Com	Report on the a			on	Report of	Summary	Activities	Summary	United S
1539	United States congressional s	Summary	Report on the activities in the	Report of the C			on	Annual re	Summary	Report on	Summary	Report c
1563	United States congressional s	Treaties, c	SERIAL SET;"Y 1.1/2:"	Y 4.F 76/2:T 71/			n1	8167=Trea	ucbserials	22555009	i2704886x	
1570	United States congressional s	Report of	Report of the Committee on P	Report of the C			or	Allocation	Report of	Annual re	Centralize	Legislati
1571	United States congressional s	Annual re	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.D 36:R 29/;"			19	1977	ucbserials	1210065	i27112901	
1582	United States congressional s	Report of	Allocation of budget totals by	Summary of te			of	Allocation	Report of	Report of	The joint	Subdivis
1597	United States congressional s	Summary	Legislative review by the Com	Committee on			y	Report of	Report on	Report on	Legislative	Nationw
1612	United States congressional s	Committe	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.C 73/6:Or3;"			19	1975	ucbserials	723768	i27408930	
1613	United States congressional s	Committe	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.C 73/6:Or3;"			19	1975	ucbserials	723768	i27409065	
1614	United States congressional s	Committe	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.C 73/6:Or3;"			19	1975	ucbserials	723768	i27409089	
1615	United States congressional s	Proposing	Proposed legislation to amenc	Fifth quarterly report, progre	Proposed legislation authorizing appro	Reports o	Proposed	Amendme	Proposing	Amending	Highway	
1616	United States congressional s	Annual re	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.D 36:R 29/;"X";"Y 4.D 36:R 2	780512c18179999dcu x f0 0eng dcas	750311c19	1974	ucbserials	1210065	i27409193	
1617	United States congressional s	Temporar	Suspension of food stamp regi	Providing for consideration of	Consideration of H.R. 2634 : report (to a	Changes i	Maritime	Regional r	Proposed	Proposed	Propose
1640	United States congressional s	Committe	SERIAL SET;"Y 1.1/2:";"SERIAL S	Y 4.R 86/1:W 84o;"63-2:H.doc.7	780512c18179999dcu x f0 0eng dcas	910320s19	6754	ucbserials	23273555	i27456778	
1641	United States congressional s	Second re	Senator from Oklahoma : repo	Aircraft fuel efficiency prograi	Immigration and naturalization : report	Report on	National e	Report on	Third repc	Allocation	Fraud ar

The culprit? X-links



Items linked to multiple bibliographic records (for multiple titles bound together) caused the data to be skewed

- ~~Export settings~~
- ~~OpenRefine~~
- ~~Excel experts~~
- ~~SQL queries~~ (*still trying…)
- ~~MarcEdit~~

Nothing (so far) gave me the data I needed in the format GPO requests

This is what I want

Title	SuDoc	Date(s)	vol.	OCLC#	(item)record#
The Federal Death Penalty Act of 1989 : report (to accompany S. 32)	Y 1.1/5:101-170/corr	1989		20846222	i25426898
Approving the location of the memorial to the women who served in Vietnam : report (to accompany S.J. Res. 207)	Y 1.1/5:101-171	1989		20806351	i25426898
Transportation Employee Testing Act of 1989 : report of the Senate Committee on Commerce, Science, and Transportation on S. 561	Y 1.1/5:101-172	1989		20742938	i25426898
Antiterrorism and Arms Export Amendments Act of 1989 : report (to accompany S. 347)	Y 1.1/5:101-173	1989		20685300	i25426898

This is what I got – incomplete data

Create list of **items** in Sierra (store record typ

In **Data Exchange**, create "out" file using "Cre
field]

Use **MarcEdit**'s Export Delimited Data functio

245 (title)
086$a (sudoc)
008 (to extract date(s)
945$c (volume from item)
001 (oclc #)
945$y (item record #, aka local system numb

To get the date from the 008
In **Notepad++** (copy column in to edit, copy b
Extract date1-date2 from 008 using find/repla

(\d{6}[a-z])([\du]{4})([\du]{4})(.)*
$2-$3

Extract date1 from 008 using find/replace rege

(\d{6}\S)([\du]{4})(.)*
$2

Edit delimiters & subfields out of 245

Add in columns for format, condition, storage location, access, notes

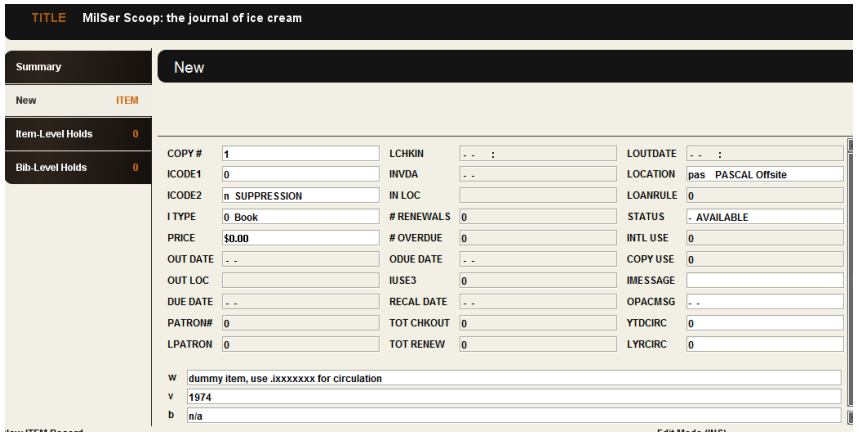
DO NOT USE

	A	B	C	D	E	F	G	H	I
1	Title	SuDoc	Publication date(s)	Vol/Issue,	Format	Condition	Storage Locati	OCLC#	Local System #
2	The economic effects of franchising / by Urban B. O	Y 4.Sm 1/2:F 84/2; Y 4.Sm 1/2:F 84/2	1971		p		Offsite Storage	204933	.i11738248
3	Hearing on the reauthorization of the community s	Y 4.ED 8/1:103-55; Y 4.ED 8/1:103-55	1994		p		Offsite Storage	30725803	.i20802511
4	Domestic violence : the struggle for survival : heari	Y 4.Ap 6/2:S.HRG.102-45; Y 4.Ap 6/2:S.HRG.102-45	1991		p		Offsite Storage	24083144	.i20853762
5	Endangered Species Conservation Act of 1972 : Hea	Y 4.C 73/2:92-81; Y 4.C 73/2:92-81	1972		p		Offsite Storage	578668	.i20912973
6	Predatory mammals and endangered species. : Hea	Y 4.M 53:92-22; Y 4.M 53:92-22	1972		p		Offsite Storage	416211	.i20912997
7	Endangered species : hearings before the Subcomn	Y 4.M 53:93-5; Y 4.M 53:93-5	1973		p		Offsite Storage	696921	.i2091300x
8	Hearings on H.R. 3160, the Comprehensive Occupat	Y 4.Ed 8/1:102-99; Y 4.Ed 8/1:102-100; Y 4.Ed 8/1:1	1992-9999	v.2	p		Offsite Storage	26535289	.i2093810x
9	Hearings on H.R. 3160, the Comprehensive Occupat	Y 4.Ed 8/1:102-99; Y 4.Ed 8/1:102-100; Y 4.Ed 8/1:1	1992-9999	v.1	p		Offsite Storage	26535289	.i20938111
10	Hearings on H.R. 1280, Comprehensive Occupationa	Y 4.ED 8/1:103-35; Y 4.ED 8/1:103-35	1994		p		Offsite Storage	29887223	.i20938135
11	Poverty, hunger, and the welfare system : hearing b	Y 4.H 89:99-23; Y 4.H 89:99-23	1987		p		Offsite Storage	15271852	.i20947355
12	Redesigned Space Station Program : hearing before	Y 4.C 73/7:S.HRG.103-283; Y 4.C 73/7:S.HRG.103-28	1993		p		Offsite Storage	29613705	.i20947835
13	Comprehensive health care reform : the need for a	Y 4.L 11/4:S.HRG.103-173; Y 4.L 11/4:S.HRG.103-17	1993		p		Offsite Storage	28952804	.i20948384
14	Coverage for mental and addictive disorders in hea	Y 4.L 11/4:S.HRG.103-211; Y 4.L 11/4:S.HRG.103-21	1993		p		Offsite Storage	29195099	.i20948396
15	Hearing on the Child Abuse Prevention, Adoption, a	Y 4.Ed 8/1:102-96; Y 4.Ed 8/1:102-96	1992		p		Offsite Storage	26427784	.i21133955
16	Review of federal measurements of race and ethni	Y 4.P 84/10:103-7; Y 4.P 84/10:103-7	1994		p		Offsite Storage	30021858	.i21141289
17	Implementation of Indian Gaming Regulatory Act :	Y 4.R 31/3:103-17/; Y 4.R 31/3:103-17/	1993-9999	pt.4	p		Offsite Storage	29214637	.i21169706
18	Indian economic development : oversight hearing b	Y 4.R 31/3:103-18/; Y 4.R 31/3:103-18/	1993-9999	pt.1	p		Offsite Storage	28883671	.i21169743
19	Indian economic development : oversight hearing b	Y 4.R 31/3:103-18/; Y 4.R 31/3:103-18/	1993-9999	pt.2	p		Offsite Storage	28883671	.i21169767
20	Implementation of Indian Gaming Regulatory Act :	Y 4.R 31/3:103-17/; Y 4.R 31/3:103-17/	1993-9999	pt.2	p		Offsite Storage	29214637	.i21169779
21	Implementation of Indian Gaming Regulatory Act :	Y 4.R 31/3:103-17/; Y 4.R 31/3:103-17/	1993-9999	pt.3	p		Offsite Storage	29214637	.i21169792
22	Memorandum of agreement between the United S	Y 4.P 96/11:101-50; Y 4.P 96/11:101-50	1990		p		Offsite Storage	22973936	.i21176176
23	Oversight hearings on elementary and secondary e	Y 4.ED 8/1:ED 8/39; Y 4.ED 8/1:ED 8/39	1971		p		Offsite Storage	142075	.i2124313x
24	Fast track authority and North American Free Trade	Y 4.P 96/11:102-12; Y 4.P 96/11:102-12	1991		p		Offsite Storage	24678625	.i21243451
25	Field hearings on H.R. 3878, the American Jobs Prot	Y 4.Ed 8/1:102-109; Y 4.Ed 8/1:102-109	1992		p		Offsite Storage	26826257	.i21243475

Maybe I was using the wrong sized container?

One more idea— “dummy” items?

What about attaching a suppressed item record to each individual x-linked bibliographic record?



This would have been an ideal solution… **if** we’d been doing it all along

- Over 2,000 x-linked items for our Hearings alone, many of which are attached to bibs with multiple items (not all of which are x-linked)
- No way to batch insert & too labor-intensive to add manually

Plan B is not Plan A harder

photos: Eric Abramson; Fourmile Canyon Fire, 2010

Persistence pays off

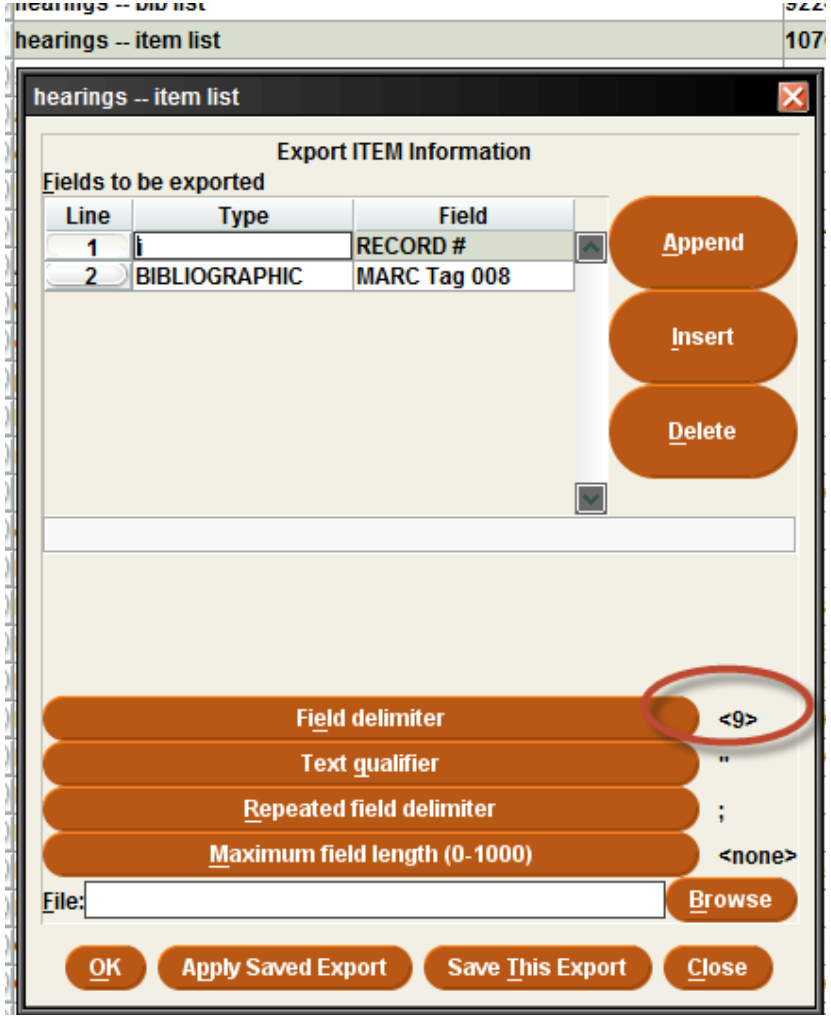
Iterative Process Outline

- Export from ITEM list (each export is done separately)
- .i number, 245
  - Concatenate* 245s
- .i number, SuDoc
  - Concatenate SuDocs, add to main sheet
- .i number, 008 [I'll explain later why I used the entire 008, not date1, date2]
  - Extract date(s) in Notepad++ (see next slide), concatenate, add to main sheet
- .i number, OCLC number
  - Concatenate, add to main sheet
- .i number, volume field
  - Add to main sheet

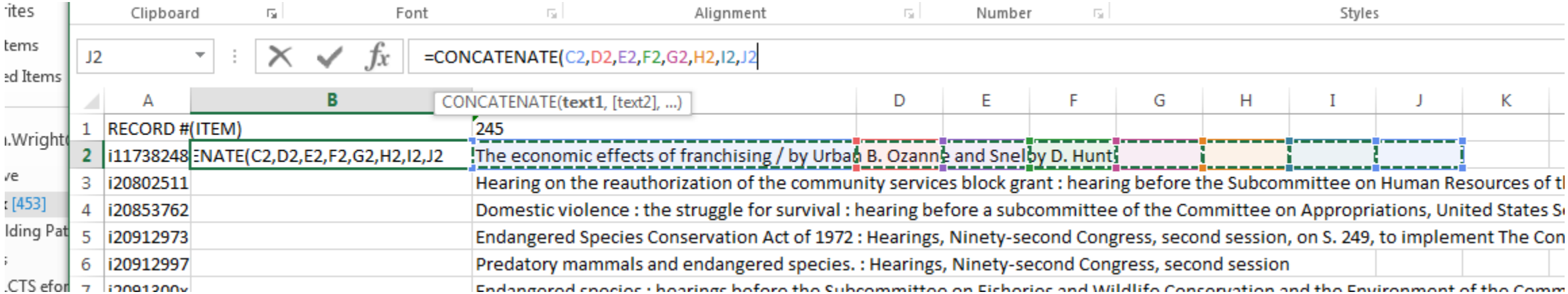
A few notes and caveats

- Export item list from Sierra as tab-delimited text, then open text file in Excel
- Use filter in Excel to identify how many columns are populated with data
- Make sure your data is stored as “General” for concatenation function to work (some concatenation methods require “Text” format)
- Concatenation (formula)
  - I separated the repeated values with “;” because it seemed clear (and the semicolon doesn’t appear in any of the data)
  - After concatenating, make sure to delete trailing semicolons
- Copy values (not the formula)
- Export bib record list also and add in separate sheet for x-reference (.b number, OCLC #, 245, SuDocs)

Export as tab delimited



Concatenate in Excel



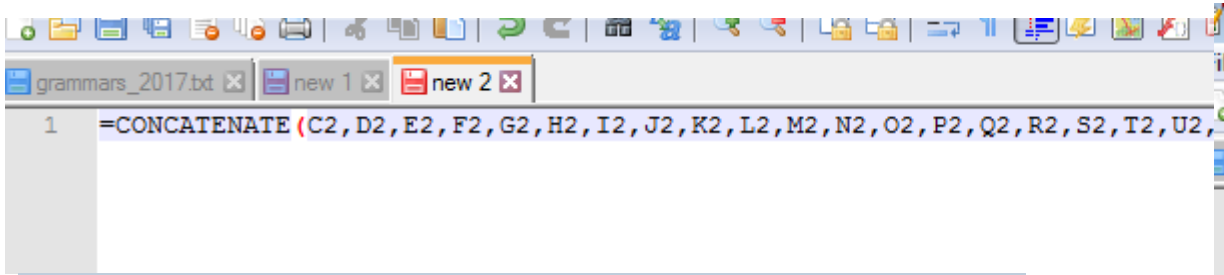
- Insert a new column, making sure the format is “General”

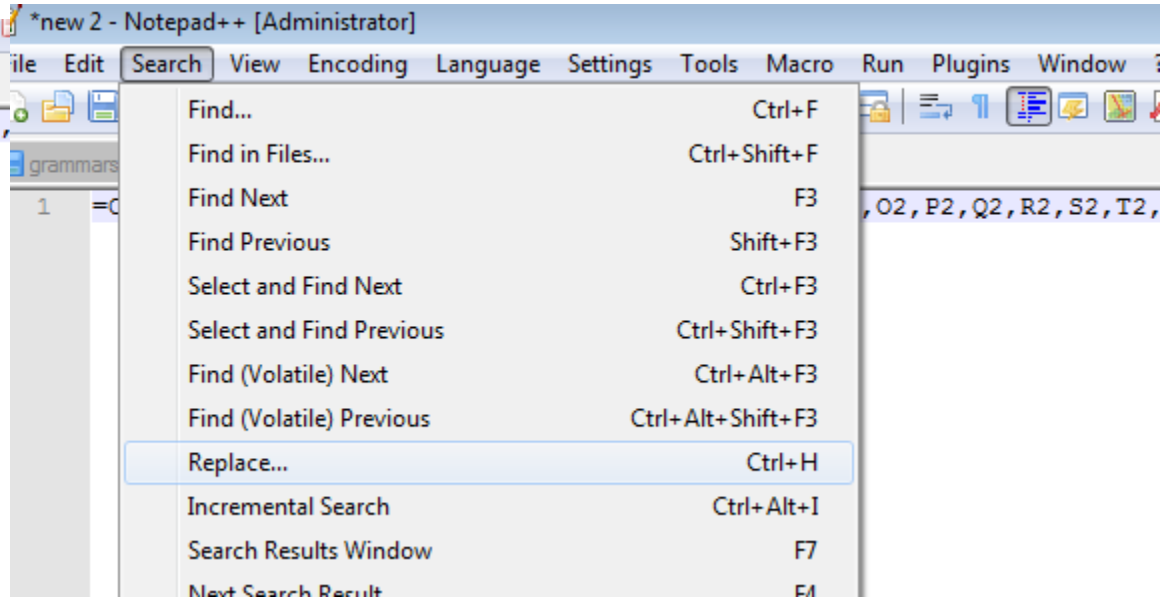
  (a function typed into a cell formatted as text will be read as literal text)

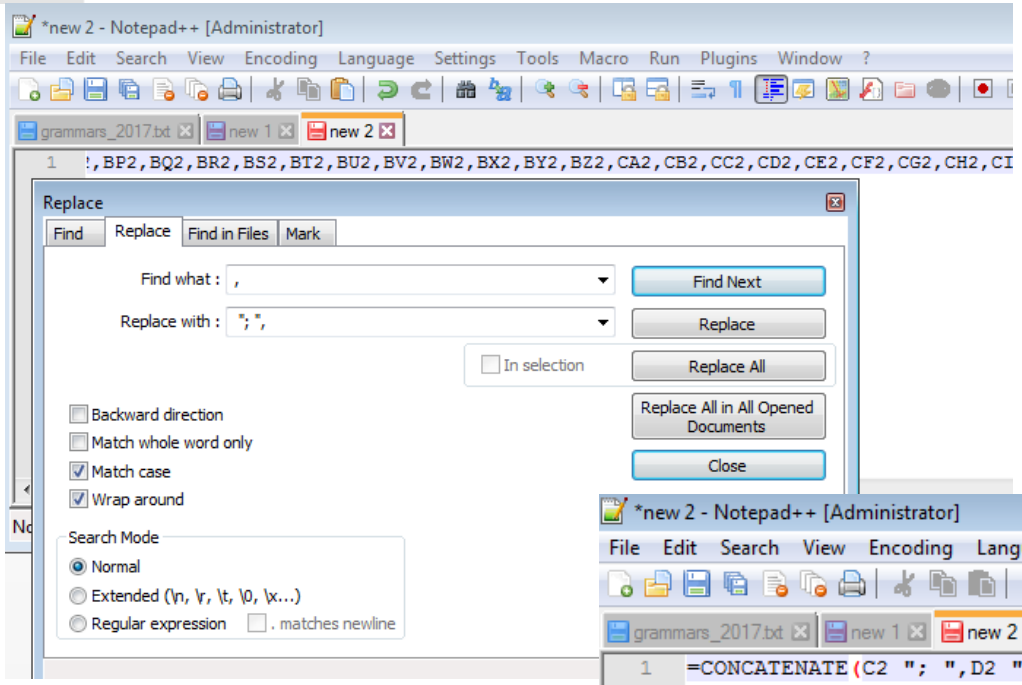
- Using the “CONCATENATE” function while holding down “CTRL,” select each cell individually

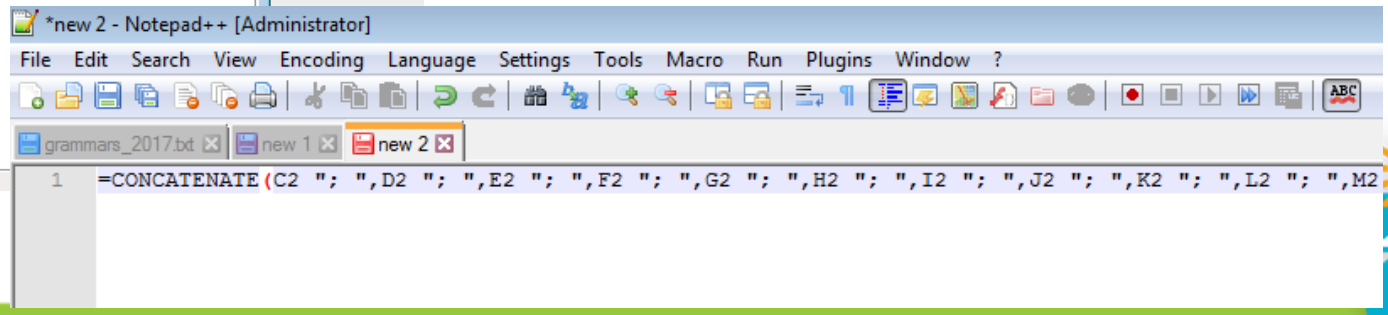
  (selecting a range of cells does not work)

Add separating punctuation



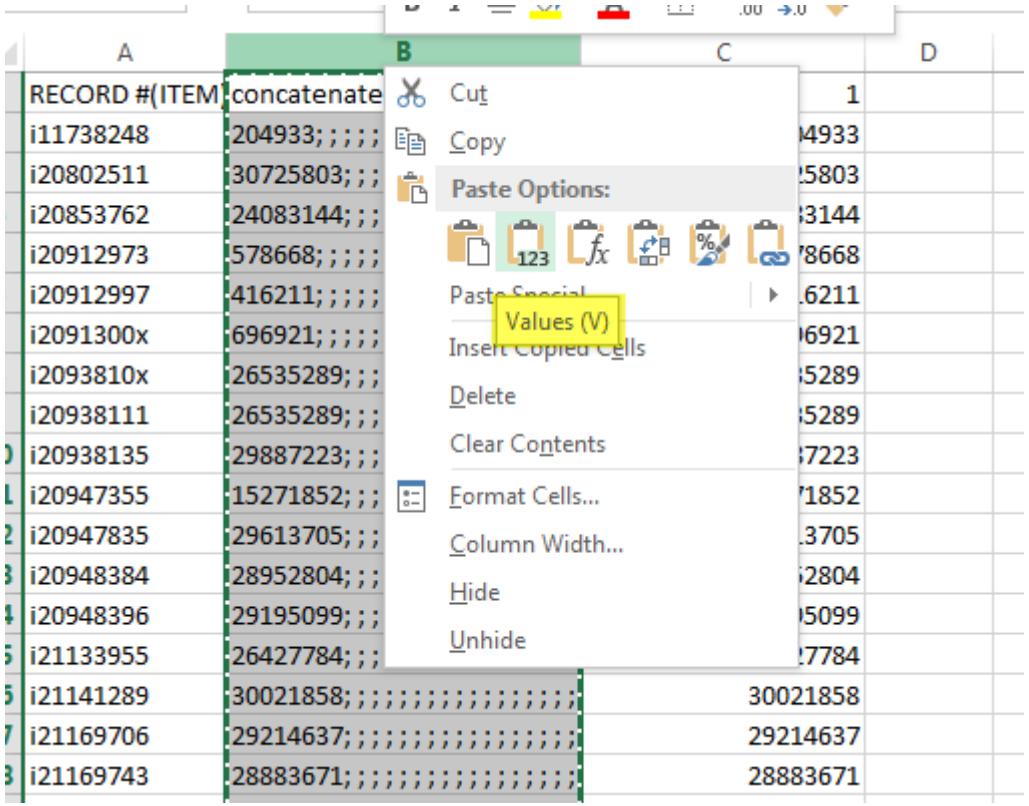




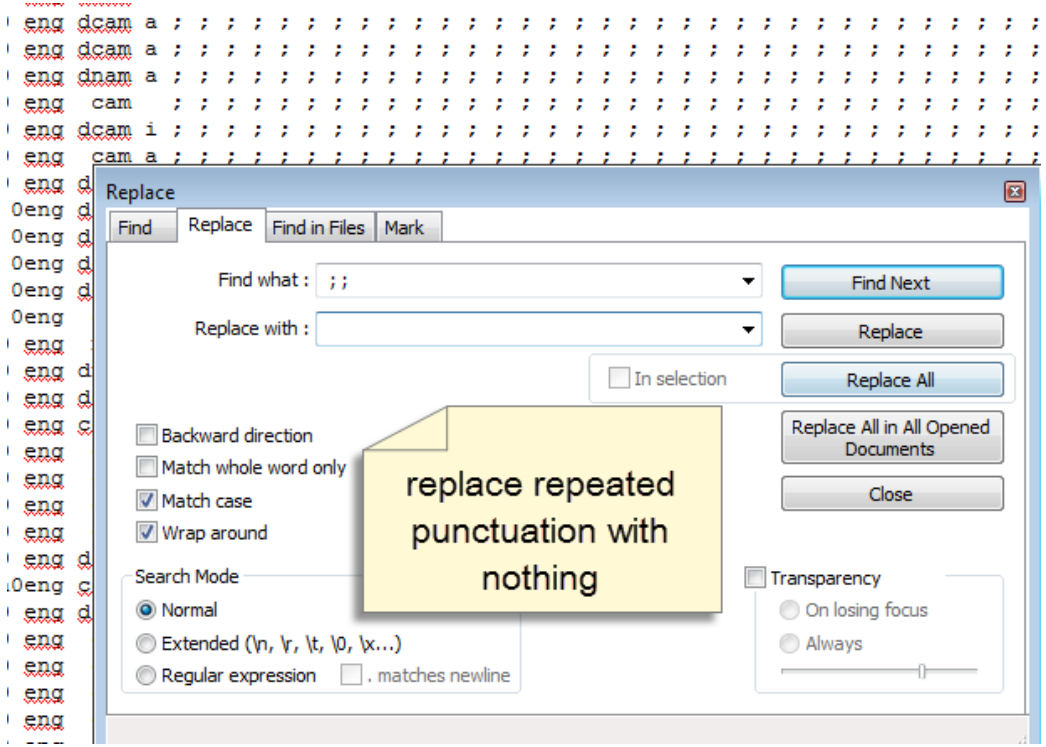


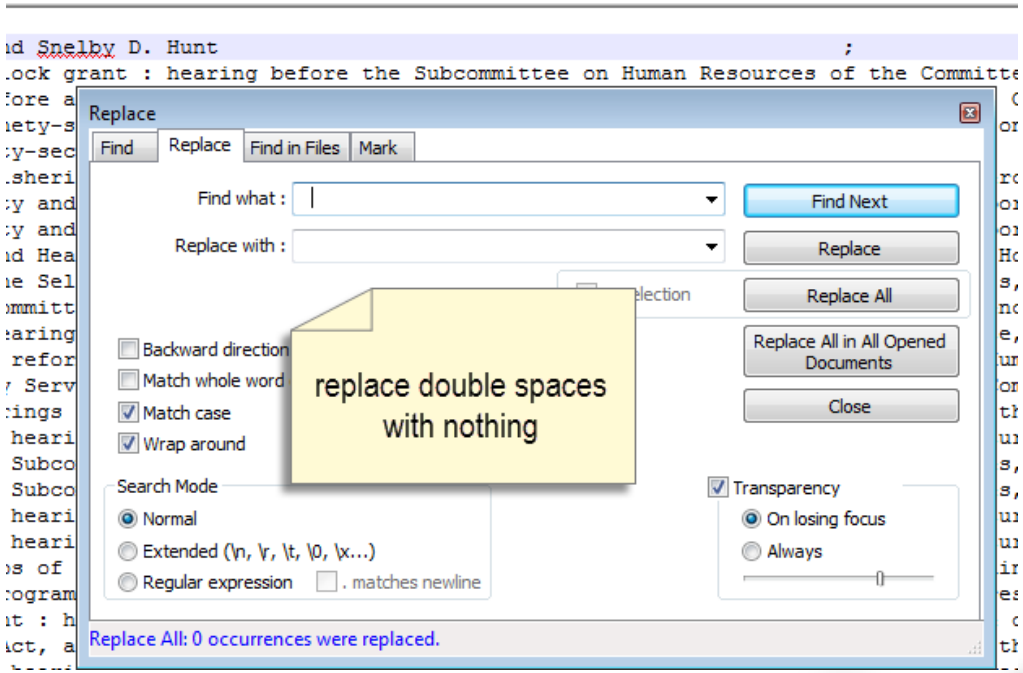
I used Notepad++ but there are ways you could do this within Excel as well

Copy/Paste concatenated data as “values”

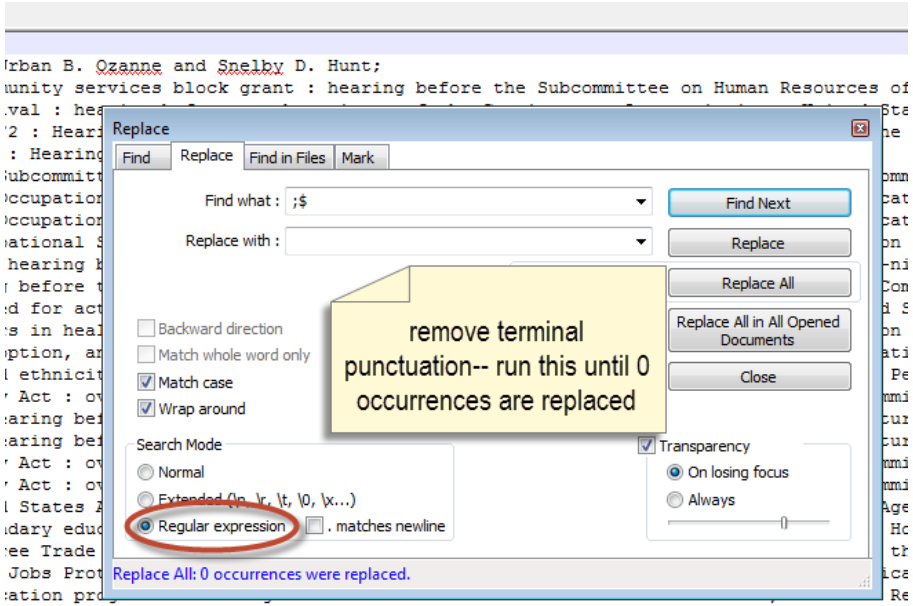


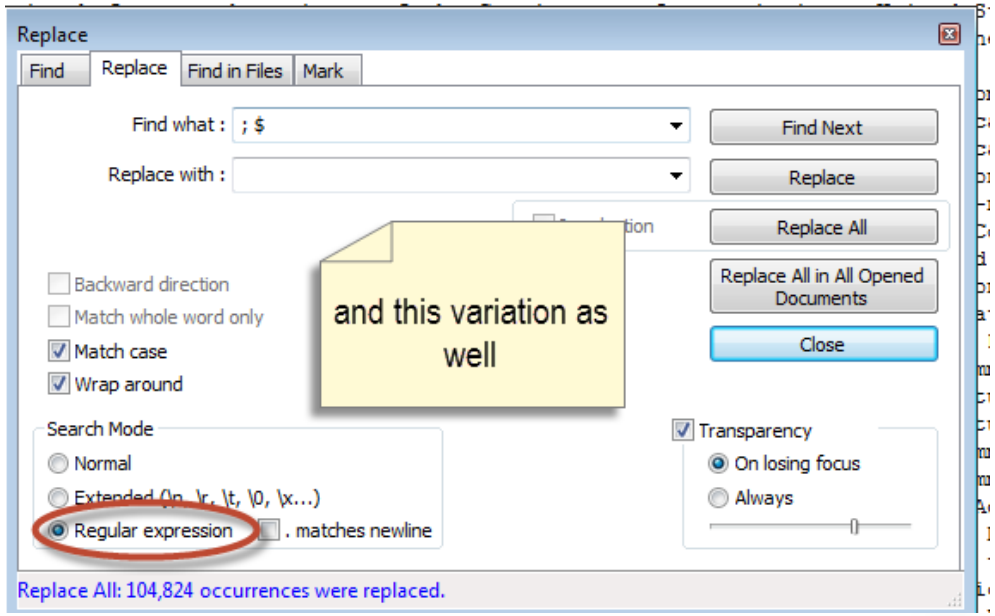
More clean-up of concatenated fields





And if you're really picky





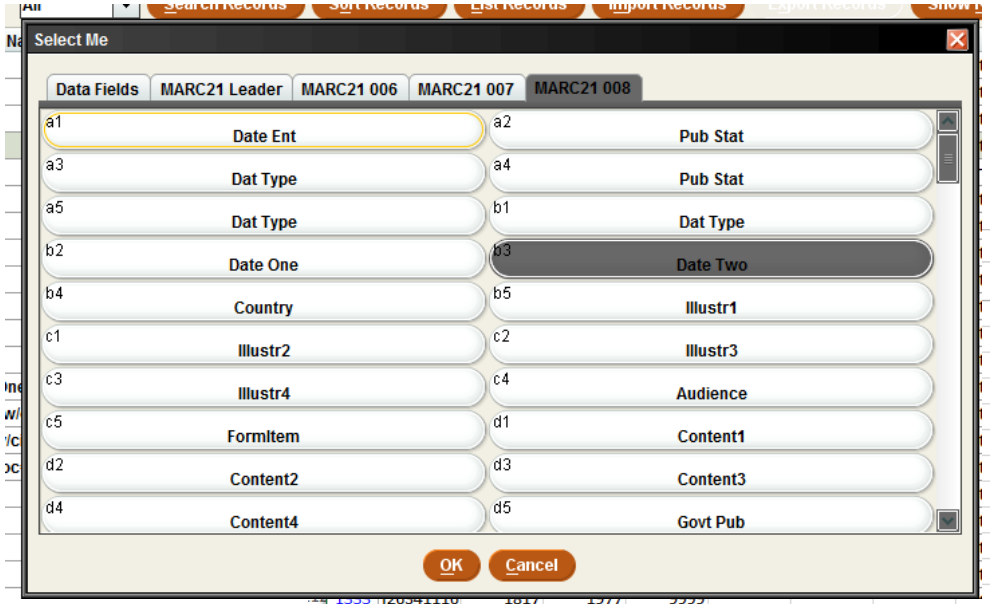
After clean-up, paste entire column back into spreadsheet.

Copy and paste the concatenation formula to repeat this process for each field (e.g., SuDoc, dates [more on this coming soon], OCLC number, volume [from item record], .b record number), pasting the new fields into the master spreadsheet

This is what it ends up looking like

	A	B	C	D	E	F	G	H	I	
L	TITLE(s)	SUDOC(s)	DATE(s)	VOLUME/ISS	FORMAT	CONDIT	STORA(	ACCESS	OCLC NUMBER(s)	LOCAL SYSTEM
38	Harmful non-indigenous species in the U.S. : hearings before the Committee on Governmental Affairs, United Sta	Y 4.G 74/9:S.HRG.103-602	1994		p				30931127	i25411780
39	UNESCO convention on cultural property : hearings before the Subcommittee on Trade of the Committee on Way	Y 4.W 36:95-28	1977		p				3397915	i25411925
40	Arctic National Wildlife Refuge : hearings before the Subcommittee on Fisheries and Wildlife Conservation and t	Y 4.M 53:102-26	1991-9999	[pt.1]	p				24882193	i25411937
41	Honoraria amendments to the Ethics in Government Act : report of the Committee on Governmental Affairs, Unit	Y 1.1/5:102-29; Y 1.1/5:102-37; Y 1.	1991 ; 199	1991	p				23948837; 237346	i25426394
42	The Federal Death Penalty Act of 1989 : report (to accompany S. 32); Approving the location of the memorial to th	Y 1.1/5:101-170/corr; Y 1.1/5:101-1	1989 ; 198	1989	p				20846222; 208063	i25426898
43	China's economy looks toward the year 2000 : selected papers / submitted to the Joint Economic Committee, Con	Y 4.Ec 7:C 44/10/	1986	v.1	p				13848964	i25428238
44	Drugs and violence : the criminal justice system in crisis : hearings before the Permanent Subcommittee on Inves	Y 4.G 74/9:S.HRG.101-290/	1989-9999	pt.3	p				20800798	i25428287
45	Miscellaneous. : Hearings before the Committee on Agriculture, House of Representatives, Ninety-first Congress	Y 4.Ag 8/1:M 68/2/970	1971		p				130234	i25433854
46	Arctic National Wildlife Refuge (ANWR) : oversight hearing before the Subcommittee on Water and Power Resou	Y 4.In 8/14:100-52/	1988-9999	pt.6	p				18898646	i25438232
47	Arctic National Wildlife Refuge (ANWR) : oversight hearing before the Subcommittee on Water and Power Resou	Y 4.In 8/14:100-52/	1988-9999	pt.3	p				18898646	i25438244
48	Arctic National Wildlife Refuge (ANWR) : oversight hearing before the Subcommittee on Water and Power Resou	Y 4.In 8/14:100-52/	1988-9999	pt.2	p				18898646	i2543827x
49	Arctic National Wildlife Refuge (ANWR) : oversight hearing before the Subcommittee on Water and Power Resou	Y 4.In 8/14:100-52/	1988-9999	pt.4	p				18898646	i25438293
50	Arctic National Wildlife Refuge development : joint hearing before the Subcommittee on Fisheries and Wildlife (	Y 4.M 53:100-45	1988		p				17986987	i25438335
51	Satellite technology and world food security : report / prepared for the Select Committee on Hunger, U.S. House	Y 4.H 89:Sa 8	1989		p				20690015	i25438712
52	Medicare and Medicaid Patient and Program Protection Act of 1987 : report (to accompany S. 661) (including cost	Y 1.1/5:100-109; Y 1.1/5:100-110; Y	1987 ; 198	1987	p				16389019; 163986	i25441875

Finally, the dates



which date(s) go with which volumes?

		H	I	J	K	L	M	N	O	P	Q	R
		1978	1978	1978	19uu	19uu	1978	1978	1978	1978	1978	1
		1977	1968	1977	1977	1977	1977	1977	1977	1977	1977	1
		1977	1977	1977	1977	1977	1977	1977	1977	1977	1977	1
		1991	1991	1991	1991	1991	1991	1991	19uu	19uu	198u	19uu
		1989	1989	1989	1989	1989	1989	1989	1989	1989	1989	1
		1987	1987	1987	1987	1987	1987	1987	1987	1987	1987	1
		1983	1984	1984	1984	1984	1984	1984	1984	1984	19uu	1
		1979	1979	19uu	1979	1979	1979	1979	1979	1979	1979	1
		1977	1977	1977	1977	1977	1977	1977	1977	1977	1977	1
1360	i26414442	1817	1976	9999								
1377	i26465966	1980	19uu	1977	19uu	19uu	19uu	1817	1995	1995	197u	199u	1995	1992	9999			
1478	i26645518	1817	1976	19uu	19uu	1977	19uu	19uu	19uu	19uu	19uu	19uu	19uu	19uu	9999	9999	1995	1
1517	i26723281	1817	1985	19uu	1980	19uu	19uu	19uu	1975	19uu	9999		199u	1995	9999	1995	9999	2
1538	i26857534	1817	1975	1978	19uu	19uu	1978	1977	1977	19uu	19uu	19uu	1980	19uu	1980	19uu	19uu	19uu
1562	i2704886x	1817	1910	9999	1938													
1570	i27111179	1817	19uu	19uu	19uu	19uu	19uu	19uu	19uu	1978	1978	19uu	19uu	19uu	1976	19uu	1978	1
1571	i27112901	1817	1951	9999	9999													
1582	i27204868	1817	19uu	19uu	1977	1977	19uu	1977	19uu	19uu	196u	19uu	1977	1977	19uu	19uu	19uu	19uu
1597	i27250969	1817	19uu	19uu	1976	19uu	1976	19uu	19uu	19uu	19uu	1976	1976	1976	19uu	19uu	19uu	9
1612	i27408930	1817	1973	9999														
1613	i27409065	1817	1973	9999														
1614	i27409089	1817	1973	9999														
1615	i27409119	1817	1975	1975	1975	1975	1975	1975	1975	1975	1975	1969	19uu	1975	1975	1975	1975	1
1616	i27409193	1817	1951	9999	9999													
1617	i27409533	1817	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1975	1

Regex to the rescue (again) to extract dates from 008

In **Notepad++ (or other text editor)**

Extract date1-date2 from 008 using find/replace regex:

(\d{6}[a-z])([\du]{4})([\du]{4})(.{30})

replace with

$2-$3

Extract date1 from 008 using find/replace regex:

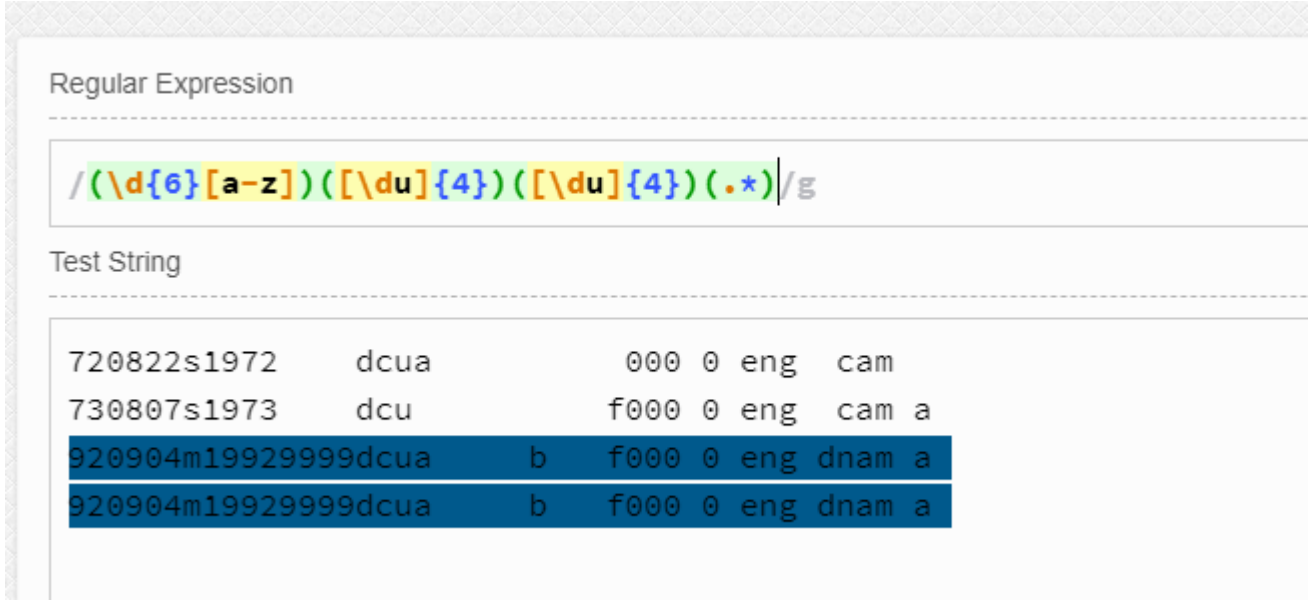
(\d{6}\S)([\du]{4})(\s{4})(.{30})

replace with

$2

For testing regular expressions

regexpal.com



Conclusions

- Consider potential ramifications of decisions, not just how they affect display
- Invoke multiple perspectives
- Explore all available tools
- Document both failures and successes

Questions? Comments? Suggestions?

Laura.Wright@colorado.edu