

Hi. I'm Stacey Wolf, from the University of North Texas.

Welcome.

I'm here today to talk about some ways you can automate the authority control process.

Outline:

- •What is authority control?
- •Me and my library
- Different ways to automate authority control processes
- Some ideas on how to automate the authority control maintenance process
- •Tips
- •Resources and Questions

Here's what I plan to cover today:

I'll briefly explain what authority control is

I'll tell you a little bit about me and the library I work at

I'll talk about the different ways you can automate the authority control process in general

I'll show you some different methods I have come up with to automate the maintenance process

I'll give you a few tips, some links to resources and then take any questions you might have.

This presentation could easily be very technical, with regular expressions and JSON coding throughout it. Given that my audience today may not all be catalogers, I'm going to do my best to keep it much more broad. If you want more in-depth procedures, I'll provide a handout and I have a link to a much more detail version of this presentation.

What is Authority Control?

• "Authority control (or access point control) refers to the normalisation of controlled access points (headings) and the provision of alternative and related access points." (pg 45, Zumer, 2009)

- Collocates all works by a specific author
- Helps disambiguate people, corporate bodies, or titles
- •Can "identify and relate resources by subject" (pg. 45, Zumer, 2009)

So what is authority control?

- Why do I want to do it, let alone automate it?
- According to IFLA, "Authority control (or access point control) refers to the normalisation of controlled access points (headings) and the provision of alternative and related access points."
- For example, an authority record establishes a certain spelling and form of a name as preferred and lists alternate spellings, forms, etc. of that name, as well as related access points.
- The same can be true with subjects and titles.
- By using the same form of an access point (like the author or the subject or the series title), authority control helps users find everything by a certain author or on a specific subject.
- Authority records contain information that can help users disambiguate names of people, corporate bodies, and titles; this is very helpful with common names, like John Smith.
- Authority control can also help users identify other subject terms to use by providing broader or narrower terms related to that subject.



Here's an example of an authority record.

- You have the authorized access point (Apple Computer, Inc.)
- Other names of the company are Apple (Firm) or Apple Computer Company. If you were to search by one of these names, the authority record would redirect you to the authorized form of Apple Computer, Inc.
- This record also has extra information linking it to the founders of the company, Steve Jobs and Steve Wozniak. It also tells you that this company has changed its name to Apple, Inc., and provides a link to that record.
- There is a lot more descriptive information in this record, like the company's headquarters, their fields of activity (like computer software, household electronics, Macintosh, and Ipod). The sources for this information are listed at the bottom of the record.



6 years at UNT

para-professional copy cataloger

- juvenile books and authority control
- batch e-resource record editing and loading

Cataloging and Metadata Librarian

- supervise the staff person who does most of our authority work
- working towards using an outside vendor
- So that is a little bit about authority control, I'll get into it more in just a minute.
- First, a little background on me and my library and how we use authority control
- I started at UNT in 2013, just as RDA was being implemented.
- I started as a para-professional copy cataloger, specializing in juvenile books and authority control.
- My job responsibilities eventually grew to include batch e-resource record editing and loading
- In 2016, I was promoted to Cataloging and Metadata Librarian.
- My tasks remained mostly the same (just more complicated), and now I supervise the staff person who does most of the library's authority work.
- Over the last few years, I have led an enormous project to outsource our authority work to a vendor.



- The UNT library catalog runs on III's Sierra.
- We have about 3 million records
- We add about 100 thousand records each year, with about 75% of those being e-resources from vendor batch records.
- UNT libraries has decentralized cataloging, with Media, Music, Government Documents, and Special Collections doing their own cataloging.
- There are 7 catalogers in the main cataloging and metadata services department who all do their own authority work.
- One staff person is trained to do the authority work for the rest of the library, sending new names and subjects or problems to that specific department

sierr	a		FUNCTION Headings Reports	
Report Type		Limit to these Locations:		
All		AB	czm Chilton Media Library	
Headings used	d for the first time		czmrs Chilton Media Reference czmrs Chilton Media Reserves czwww MEDIA ONLINE RE SOURCE S	Local
Invalid heading	gs		d Dallas Campus Library	E 11 11
Duplicate entri	ies	Limit to these Logins:		authority
Blind reference	es	48	aacosta Axel Acosta abajracharya Anima Bajracharya	
Duplicate auth	ority records		abaron Adam Baron accoper Amy Cooper	control
Updated bibliog	graphic headings		admin administrator	
Near matches		Limit to these Programs:		- priorities
🗹 Records busy.	, not updated	-01	Record Maintenance	(Thom)
Non-unique 4X	(Xs		Rapid Update Becord Loading	(inen)
Cross-thesaur	rus matches	Limit to these Statistics:		
Egclude record	ds without a CAT DAT	E VI	System	
Entry Date Range			10 Group/Admin Login	
From:	20		50 Lib TACO Full Time	
Tg:	20		95 Willis Self-Checkout	

- When I started at UNT in 2013, I worked with a librarian and another staff person to do all the library's authority control.
- Every day, we ran Headings Reports, which provide us with a list of names, subjects, and titles used for the first time, as well as other potential issues in the catalog.
- Back then, we looked up in OCLC every new name and subject on the reports, but only downloaded the authority record into our catalog if it had been established with cross references or a see also reference.
- We validated new names and subjects and made sure the form was correct in the newly cataloged record as well as existing records in our ILS.
- Sometimes, if a name had changed, we needed to update the authority record and the bib records.



- In addition to keeping up with the Headings Reports, we also used the Library of Congress's updated subject heading lists and OCLC's Closed dates lists to update our authority records and bib records.
- I'll explain more about them a little later.
- This all became difficult to maintain as RDA changed hundreds of thousands of authority records overnight and our library staffing changed
- A three person team working on authority control became a team of one or two (depending on staffing at the time)
- Also around this time, our library started looking at what other institutions our size do for catalog maintenance.
- With the addition of a new head of cataloging in our department, we went ahead with exploring using a vendor for authority control, so we needed to re-prioritize our current authority work.

Local authority control priorities (Now)

- •Use the "preceded by" and "followed by" lines to identify potential conflicts to investigate
- •Don't load new subject headings records
- •Don't update authority records as the national authority file changes
- General collection catalogers do their own authority work

Field: b7001 |aLee, Ellie,|d1970-|4pro| Indexed as AUTHOR: *lee ellie 1970* Preceded by "a": *lee ellie* Followed by "a": *lee elliott jason* From: b61487843 Unnatural causes: is produced by California Newsreel ; in a Center for Asian American Media ; Lat

Field: b7001 |aKrause, Dan. Indexed as AUTHOR: krause dan Preceded by "a": krause dagmar Followed by "a": krause daniel j From: b61487843 Unnatural cause produced by California Newsreel; Cantor for Asian American Media

- Now, instead of searching for every new name, we have started using the "preceded by" and "followed by" lines to see if there was a potential conflict and only investigate those in OCLC.
- If we search for a record in OCLC and find one, we can download it to our ILS if we want, even if there are no cross references. (if we spent that time to look for the record, might as well use it)
- We are not loading new subject records.
- We are not bothering to update our authority records. It took too much time to look up each listed heading in our catalog (in two different indexes for names).
- We figure the vendor will eventually provide us with subject records and updated records, so we won't use precious staff time working on that now.
- I also trained our department of 7 catalogers on how to do their own authority control work
- We still have one staff person doing the heading reports for the whole library.
- Using the new process though, it takes only an hour or two to complete, instead of several hours each day.



- The most time-effective way to handle authority work is to automate it as much as possible.
- Working in large batches can often be quicker than individually trying to search for each new name or subject heading.
- There are a variety of ways to automate authority work.
- Some tools include OCLC Connexion, MarcEdit, your ILS, Notepad++, Excel, and vendors.
- OCLC Connexion is where my library finds its bibliographic and authority records to download into our ILS.
- We use Notepad++ and Excel all the time to examine and manipulate our data.
- Excel is helpful in many ways, especially with the text to column feature and conditional formatting where you can highlight duplicates.
- Notepad++ is great to use with Regular Expressions and save data in a plain text format.



- We use Sierra for our ILS.
- This is where I can run headings reports to identify new names and subjects, invalid uses, and duplicate records.
- Some ILS have a bibliographic maintenance feature that will automatically update bib records to a match new heading over night, but only if the heading in the bib record matches a cross reference in the authority record.
- I find the bibliographic maintenance program to be very unreliable, but maybe I'm using it incorrectly.



- MarcEdit is a free software that Terry Reese provides to catalogers.
- It will authorize/validate names and subjects for you so you know you have the right form before ever loading your records.
- It will provide a report for you to fix variants, or you can set it to automatically correct them.
- It will even add the URI in \$0 for you if you are preparing for linked data in your catalog (or it's handy for batch-searching OCLC later for the authority records to add to your catalog)
- You can also have it download the authority records that match the validated headings in your file (but I haven't played with that feature very much).
- MarcEdit is a great tool for editing batches of records quickly before loading them into the ILS where making mass changes is much more clunky.
- This is where I really learned how to use Regular Expressions.



- Vendors like Backstage and MARCIVE allow libraries to send a back file of all their records which are updated and validated.
- Each vendor has their own special programming to fix field tags, indicator errors, and incorrect subfield coding; both vendors offer RDA enrichment services, creating hybrid RDA records out of older AACR2 records by adding certain fields.
- The library can decide on all kinds of customizations they want made to their processing profile, usually based on local practices and policies.
- Once the records are cleaned up, they are run through as many authority files as you want to validate headings.
- The library gets the updated bibliographic records back as well as files of authority records.
- Both vendors offer a current cataloging service so that you can send newly added bibliographic records to your profile.
- Ongoing maintenance keeps your authority records up-to-date as national authority files are updated.
- This can be a wonderful option, if you have the money to spend on it.

- It can cost tens of thousands of dollars depending on the size of your catalog.
- The on-going service requires a yearly fee.
- And you will still need a staff person to run reports, create files, and reload records.
- So maybe you can afford to do a clean up project, sending your records once, but don't want to subscribe to the on-going services.
- Or you can't afford to even get your records cleaned up.
- How can you maintain your catalog's authority files and access points as efficiently as possible?



- We've used a combination of headings reports, MarcEdit and OCLC to batch load authority records for certain files of vendor records.
- I'll do a brief overview of the process since I only have a little bit of time.
- I'll have detailed instructions as a handout online



- Load records into Sierra,
- run heading reports,
- copy new names to excel

Clean up copied information so just the name remains Beceanu, Marius, eeditor. Chen, Geng, d1980-eeditor. Krause, Dan. Lee, Ellie, d1970-4pro4drt Miller, Florence.4nrt Phan, Tuocq(Tuoc Van), eeditor. Plotkin, Eugene, d1955-eeditor.

• Clean up copied information so just the name remains



- Use either the original vendor file (if these will be the only bib records in the headings report) or turn the Excel file into MARC records
- validate names in marc edit, embedding the URI in \$0 when possible
- The URI provides a link to the authority record for that specific entry, which contains the Library of Congress Control Number (or LCCN)

Found Text Found Text Found Text Found Text Found Text Jump to Record #: 0 =1001 WsaGilheary, Kevin P, \$d1963-\$eauthor.\$0 Jump to Record #: 0 =1001 WsaGilheary, Kevin P, \$d1963-\$eauthor.\$0 Jump to Record #: 0 =1011 WsaGilheary, Kevin P, \$d1963-\$eauthor.\$0 Jump to Record #: 0 =1112 WsaResearch Workshop of the Israel Scien Jump to Record #: 0 =7001 WsaChen, Geng,\$d1980-\$eeditor.\$0http://i Jump to Record #: 0 =7001 WsaPhan, Tuoc\$q(Tuoc Van),\$eeditor.\$0http Jump to Record #: 0 =7001 WsaPhan, Tuoc\$q(Tuoc Van),\$eeditor.\$0http Jump to Record #: 0 =7001 WsaPhan, Tuoc\$q(Avraham),\$eeditor.\$0http Jump to Record #: 0 =7001 WsaPhan, Tuoc\$q(Avraham),\$eeditor.\$0http Jump to Record #: 0 =7001 WsaPhan, Avy\$q(Avraham),\$eeditor.\$0http Jump to Record #: 0 =7001 WsaAtter, Shijun,\$eeditor.\$0http://id.loc Jump to Record #: 0 =7001	Copy \$0 search to Excel
--	--------------------------------

• Use a search in MarcEdit for \$0 to copy to Excel



 Isolate the Library of Congress Control number from the link in the \$0 for those names on the headings report

e	xnor	t the	reco	rd to	IIS	
	NP OI					
Criter Authority Batch Search Adju			A Str. Conten	n - Johnse Li, Jaaksen Viele Ke	teach watching of the state scie	nee Foundation on Groups, Al.,
Local Elle		Local File Merager	6845			ALL IN DR AS MAT
C - Uners' add 0045 AppDate/ Reaning-DCLC/C	orner (b) Geleukluh auh de		- Field	trings + 10 User Tack +	*ea2010163611	-0-
Search Keye Genry	Total Entered	1	ARM 1967844	ę		
		Post core and	Bec.atat	a Extend 20	Hantan Replaced 2018	1281073132.0
Use default indeg			Tune a	Upduratur	Enc.hd n	Soute e
LCCN #1	2	100	tornal Contract	fiel, status	Netter	GATELDS .
		Quint	Seres	Autom	Geo subd w	Serute 6
		Harrison	Second -	Marine m	Sabdu to m	Rules s
20		Japon.	010	ra2018163511	115 (C. 2017)	
		0.81	940	MdRoLAC #b ang #e rda #c	MeRoLAC	
		Brief	946	4x 2016-03-20 4t 2016-03-2	4 42 edtl	
			111.2	Research Workshop of the I Jonatalemi	stael Science Foundation on G	roups, Algebras and Identities #
F Patrieve all records from galine save Ne			370	to local te Jecusters 42 n	4	
Limit by Review Status			373	Universitat Banilan #a Unive	rotah ka-helt In-Yerushalasim	12 saf
					the second second process of the day beam of the second se	

- Run the batch search in OCLC,
- add any local information you might need,
- and export records either directly to your ILS or into a batch file on your desktop to load with Data Exchange.



- So that's a quick way to import new name authority records.
- It can be adapted to subject headings too.
- But, as I mentioned before, the name authority file and the subject authority files are always changing.



So how can you keep those authority records up-to-date?

- OCLC provides a list of closed dates for names weekly
- These are name headings that have a death date or some other formatting change made to the main entry.
- Library of Congress provides a list of updated subject records monthly
- The problem is, you have no way of knowing for sure if you have the updated heading in your catalog without searching each name/subject heading one by one
- The can be very time consuming.

AL	nomati	eure	։ սբս	ale		
Regular Expre ^(.*?)\$\s+?^(?:	essions: =.*^\1\$)	Boolean Search Revie <u>w</u> File Name: [Store Record Type:	BIBLIOGRAPHIC b			
Local File: Lo Cr.Users' lod:004P-AppData' Roaming 'DCLC' Conner 'Db'Dd' Search Keys Calery: Total Entered: 0 User United The Conner Total Entered: 0	col File Marager nuk/ush auch db v Enter Discritica Add Delete Replace	Sparch Us	Existing Search	DN Retrieve Saved Que	ay Save JS0	N Load JSOK

- However, I've learned a few things since we stopped updating our authority records:
- More about regular expressions and have gotten rather comfortable writing search queries with them
- Our Create List feature in Sierra now allows for JSON queries as an input option instead of needing to enter each search command line by line
- This makes it so much easier to search for 100 or more names all at once.
- With a few simple tweaks to the data in notepad++, it's as easy as copying and pasting.
- We've also played around with and figured out how to run batch searches for authority records in OCLC
- Given all of these changes, I think I've come up with a decent system for updating these records
- Essentially, turn the lists from Library of Congress and OCLC into a JSON query in Sierra to find any records that use the updated heading
- Take just those headings and use OCLC to batch search for the new authority records
- Overlay or load new authority records

• Hope the system automatically updates the bibliographic headings overnight OR use global update to update the headings in the bib records



I'll start by explaining my process for subject headings



Here is the website you can visit to get the updated lists



 First Extract the control number from the list of changed headings and copy these numbers to notepad++



 Follow the step by step directions in the handout to format and build the JSON query in Notepad++

12014				
				in quory
"questeatin (11	
("target": ("record" ("target": ("record"	: ["type": "suthority"), "field":) ["type": "suthority"), "field":	<pre>("marcTag": "010")), "expr" ("marcTag": "010")), "expr"</pre>	a Dispita fa	nrougn
("target": ("record"	: ["type": "authority"), "field":	["marcTag": "010"]),"expr"	: ["op": ":	
("target": ("record"	: ["type": "authority"), "field":	["marcTag": "010"]), "espr"	: (*op*: *o	Ciarra
("target": ("record"	: ("type": "authority"), "field":	["marcTag": "010")],"expr"	5 (Tep1s Ts	Sierra
("target": ("record"	: ["type": "authority"), "field":	["marcTag": "010"]], "espr"	1 (TepTs Ts	
("target": ("record"	: ("type": "authority"), "field":	["marcTag": "010"]], "expr"	a chapta ta	Prooto
("target": ("record"	: ("type": "wothority"), "field":	("meroTeg": "010")), "expr"	i inspri n	Jeale
["target": ["record" ["target": ["record"	: ("type": "authority"), "field": " ("type": "authority"), "field":	<pre>("marcTeg": "010"}), "expr" ("marcTeg": "010")), "expr" ("marcTeg": "010")), "expr"</pre>	: ("ep": ":) ("ep": ":	
("target": ("record"	: ("type": "anthority"),"field":	["marcTag": "010"]], "expr"	: ("ep": "e	Linto
"target": ("record" "target": ("record"	<pre>: ("type": "asthority"), "field": : ("type": "asthority"), "field":</pre>	<pre>["marcTag": "010"]], "espr" ["marcTag": "010"]], "espr"</pre>	1 ["ep": "z	LISIS
["target": ("record"	: ["type": "authority"], "field":	["maroleg": "010"]], "espa"	1 ("ep"1 "z	
the second se	the second			
	Tangati Pi Tangati Pi Tangat	<pre>pressure = provide state = provide state</pre>	<pre>"tampet" ["record": ["type" "auchority"], "field": ["marcTag": "010"]], "app" "tampet" ["record": ["type" "auchority"], "field": ["marcTag": "010"]], "app" "tampet" ["record": ["type" "auchority"], "field": ["marcTag": "010"]], "app" ["tampet": ["record": ["type" "auchority"], "field": ["marcTag": 010"]], "app" ["tampet": ["tecord": ["type" "auchority"], "field" ["marcTag": 010"]], "app" ["tampet": ["tecord": ["type" "auchority"], "field" ["marcTag": 010"]], "app" ["tampet": ["tecord": ["type" "auchority"], "field" ["marcTag": 0</pre>	"taxpet" ("record" ("type" "axhority"), "field": ("marile": "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010")), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "010"), "sepu" ("op" ") t "taxpet" ("record" ("type" stablority"), "field": ("marile" "taxpet" ("op" ") t "taxpet" ("record" ("type" " stablority"), "field": ("marile" "taxpet" ("op" ") t "taxpet" ("record" ("type" " stablority"), "field": ("marile" "tol0"), "sepu" ("op" ") t "taxpet" ("record" ("typet" " stablority"), "field" ("marile" "tol0"), "se

Run your query through Sierra's Create lists (or the equivalent in your ILS)

		1 a MUNIC Tag Oto	Append
Record 156700 194087	Descri Puppeteers Educational equalization		<u>D</u> apata.
07283 84256 32454 63360	Linoleum block-printing. Pisgah National Forest (N.C. Dido (Legendary character) Yeti		
1723x 8839x	Inventions Dido (Legendary character)	Page heading Starting record (1-14)	1
11904	Baker, Mount (Wash.) Cooking	Ending record (1.54) Number of blank lines between records	14 0
5574x 56741	Lepcha (South Asian people Lysimachia Kwakiuti language	Nymber the records in the list Display meanings of fixed-length fields in Display each yariable-length field on a new	Head of codes
1235	Callisto (Greek Injeliology)	Display labels for variable-length fields If itsting bibliographic title, print it in upper Bun in background	-
	Record 56700 94087 07283 84256 32454 53360 1723x 0838x 11904 16781 9638x 11904 5574x 5574x 5574x 5574x	Record Descri 56700 Puppeteers 94087 Educational equalization 07283 Linoleum block-printing. 84256 Pisgah National Forest (N.C. 32454 Dido (Legendary character) 5380 Yeti 1723x Inventions 0839x Dido (Legendary character) 11904 Baker, Mount (Wash.) 16781 Cooking 34229 Lepcha (South Asian people 5674x Lysimachia 5674x Kwakiuti language 71239 Callisto (Greek mythology)	Record Description 56700 Puppeteers 94087 Educational equalization 07283 Linoleum block-printing. 8256 Pisgah National Forest (N.C.) 32454 Oido (Legendary character) 63380 Yeti 1723x Inventions 6338x Oido (Legendary character) 1904 Baker, Mount (Wash.) 16781 Cooking 34229 Lepcha (South Asian people 5574x Lysimachia 56741 Kwakiuti langwage 71239 Callisto (Greek mythology)

- We have fourteen out of the two hundred and twenty records on the list.
- From here, export the control numbers in the 010 fields to a text file.
- Keep the list to delete old records later

Enter Authority Batch Search Keys				×	m (l)					
Local File			pool File Manager	1	wh Mar	Error				
C-Users'adk0040 AppDate/Rearing/00LC	Connex/Ob/Default		h.de	in the set	CI NEY	No records four	d for your	search. Plea	se change or simplify your	
Search Keys Query:	Total Entered:	14	Enter Discritics .	340	3907	search and try a	again.		cł	Dido (haract
Use default indeg: [LCCN 81]				San	enful Sea	rches (14)			re	as del cord (
In sh 85011073	_	<u>^</u>	800	Sea	rch Key	# Records	Save #	Control #	Tide	ar
In sh2010008400			Delete	Incat	85109143	1	2285	2018314	Puppetsers	
In ah 88022910		~			85841121	1	2290	2114042	Educational equalization	
In sh 95073542			Import_	in a second s	91005142		2288	3054720	Piscah National Forest (1	(C)
					85149134	1	2289	2116069	Yeti	
			Stable	Inst	85867667	1	2290	2119154	Inventions	
			Bist	lead	91000682	1	2291	11548326	Dido #c (Legendary character)	
				in si	85011073	1	2292	2060825	Bakar, Mount (Wash.)	
P Retrieve all records from griline save file				Inst	2010008400	1	2293	8535681	Cooking	
Limit by Review Status:				Ingl	85876047	1	2294	2546179	Lepcha (South Asian peo	çie)
Estavited				in st	00022910	1	2295	2243111	Lysimachia	
i segmen		•		N.9	85673542	1	2295	2151654	Kwakiuti language	

- In OCLC, enter authority batch search keys, using the default index of "LCCN (In:)".
- Process the Batch to get a "Batch Search Report" (shown here)



- Export the records.
- If your load table is set up to overlay, they will overlay the old one.
- If not, just use the list you created earlier and delete those records in the "Delete Records function" of Sierra.
- If you don't have too many in your list, you could do it individually and check for local additions that you may not want to lose.
- Or if you want to export each individual subject record from OCLC, you can use a 949 overlay text string.



- Check your headings reports for duplicate authority records and blind references
- Delete the old authority records that are duplicated by the upload
- Resolve blind references (which come from headings changing, like fictional characters)
- Most of the time, subfields are only deleted from or added to the record, not changing the main entry, so there shouldn't be too many to fix.



- Closed dates are a bit more complicated because they can occur in more places.
- Names can occur in a 1XX Main entry, as a 6XX in the subjects, or as a 7XX added entry.
- That means you have to search each name three times.



• Here is the website for these lists

Buil	d the JSON q	uery in No	tepad++
n 2008005874 n 2008042334 n 2010039374 n 2019006746 n 50007231 n 50034831 n 50043013	Option 1: Record number	Option 2: Old name format	Almeida, Fernando d', 1955- Averill, Gerald P., 1919- Bennett, Rodney, 1935- Berezowski, Cezary, 1898- Berger, Klaus, 1901- Boggs, W. Arthur, 1916- Bragg, Don, 1935-
<pre>{"queries": [["target": {"record": ("010"}},"expr": {"op": "target": {"record": ("010"}},"expr": {"op": {"target": {"record": { "010"}},"expr": {"op": {"target: {"record: { "010"}},"expr": {"op:: {"target: {"record: { "010"}},"expr": {"op:: {"target: {"record: {"target: {"record: {"target: {</pre>	<pre>"type": "authority"), "field": { "marcTag": "regex", "operands": ["n.*2008005074"]}}, "or", "type": "authority"), "field": { "marcTag": "regex", "operands": ["n.*2008042334"]}}, "or", "type": "authority"), "field": { "marcTag": "regex", "operands": ["n.*2010005746"]}}, "or", "type": "authority"}, "field": { "marcTag": "regex", "operands": ["n.*201906746"]}}, "or", "type": "authority"}, "field": { "marcTag": "regex", "operands": ["n.*50007231"]}}, "or", "type": "authority"}, "field": { "marcTag": "regex", "operands": ["n.*50034831"]}}, "or", "type": "authority", "field": { "marcTag": "regex", "operands": ["n.*50034831"]}}, "or",</pre>	<pre>"queries": ["queries": ["target": ("record": ("type": "b) "op": "starts_with", "operands": "target": ("type: "b) "op": "starts_with", "operands": "op": "starts_with", "operands": "target": ("type: "b) "op": "starts_with", "operands": "target": "b) "op": "starts_with", "operands": "target": "b) "op": "starts_with", "operands": "target": ("type: "b) "op": "starts_with", "operands": "target": "b) "op": "starts_with", "operands": "target": ("type: "b) "op": "starts_with", "operands": "target": "b) "op": "starts_with", "operands": "target": "b) "op": "starts_with", "operands": "target": "b)</pre>	<pre>ib"),"field": ("marcTag": "100")),"expr": ["Almeids, Fernando d'"])),"on", ib"),"field": ("marcTag": "100")),"expr": ["Averill, Gerald P."]),"on", ib"),"field": ("marcTag": "100")),"expr": ["Bennett, Rodney"])),"on", ib"),"field": ("marcTag": "100")),"expr": ["Berger, Klaus"]),"on", ib"),"field": ("marcTag": "100")),"expr": ["Berger, Klaus"]),"on", ib"),"field": ("marcTag": "100"),"expr": ["Boggs, M. Arthun"])),"on",</pre>

- There are different levels of thoroughness you can work towards.
- Each level requires more manual work, so it's up to you to decide what's worth it in the end.
- Option 1: You can just search the authority records for the closed dates (like we did with the updated subject headings).
- This will find the authority records of any names on the list in your catalog, but then you will still need to change the corresponding bib records.
- In our catalog, we didn't load authority records for every name, so we may used the name in a bib record, but not have the authority record.
- We would never know if we only used this option.
- Option 2: You can search your bibliographic records for headings that match the old name.
- This does a pretty good job returning a name match.
- You just have to sort through many potentially false matches

Store Record " gps: WU	лнояту в	
Range	Spart a1000007 Stop a0315980 RUI	n allerv
Classic	JSON	quory
Enhanced	th	rough
	"queries":	rouan
JSON	<pre>("target": ("record": ("type": "authority"), "field": ("marciag": "olo")), "expr": ["op": "s ("target": ("record": ("type": "authority"), "field": ("marciag": "olo")), "expr": ("op": "s</pre>	
	("target": ("record": ["type": "acthority"),"field": ["marcTag": "010"]),"expr": ["op": "s	1
	("terget") ("record") ("type": "authority"),"field": ("marolag") "010")),"espr") ("op": "s	Jerra
	("target": ["record": ["type": "athority"),"field": ["marcTag": "010"]),"expr": ("op": "d	Jona
	["temper": ("record": ("type": "authority"), "inid": ["maring": (UO")), "type": ("op": "	
	Thereast is recorded in the second of the se	
	"target": ["record": ["ture": "acthority"]."field": ["sardTar": "00"]]."expr": ["top": "s	reate
	("target": ("record": ("type": "authority"), "field": ("marcTeg": "010")), "eppr": ("p": "	loaic
	("target": ("second": ("type": "sothority"),"field": ("saroTag": "010")),"exps": ("op": "s	
	("target": ("record": ("type": "anthority"), "field": ("marcleg": "010")), "expr": ("op": "z	· · · · ·
	<pre>("tempet": ("teopet": "subboxity")."field": ("saysTag": "010"))."espt": ("op": ")</pre>	lete
	("target": ("record": ("type": "anthority"),"field": ("mancTag": "010")),"espr": ("cp": "d	
	("target": ("record") ["type") "ashnozivy"), "Leid") ["marchag": "010"];, "espu"; ["op": "p	
	[] using the second of type 1: "second type, the late of the constraint should be the second seco	
	("tanget": ("pecord": ("type: "authority"), "field": ("marchar": "010"), "espr": ("op": "s	
	interested descents interests descented standards interested within terms into the	
	Search (Ine Internet Search) (Reserve Saved Outry) (Save JSON) (Load JSON) (Dear)	

• Copy and paste the JSON query into a create lists line

Search Results List AUTHORITY I Fields to be listed Type Line 1 a MARC Tag 010 Option 1: 1 a210443x Broecker, Wallace S., 1931 •12 of 86 record numbers 2 a2240944 Berger, Klaus, 1901 3 42869950 Gough, Kathleen, 1925 4 a401845x Devonshire, R. L., Mrs., b. 1864 5 a4461769 • Option 2: Lunt, James D., 1917-6 a4755170 Loucks, William N. (Wil lepele), 1899 7 a5509476 Dupont, Paul, d. 1957 8 a6727323 Aretino, Pietro, 1492-1556, Works, 1992 •100+ results out of 86 old 9 a7213918 Traquair, John Stewart, Earl of, d. 1659. 10 a7220650 Traquair, John Stewart, Earl of, d. 1659. 11 a7282631 Easton, Michael, 1954names searched in 258 12 49020809 Houston, David, 1937-Display meanings of fixed-length fields instead of code total searches; 18 were Cisplay each yariable-length field on a new line Display labels for variable-length fields more closely investigated; ✓ If listing bibliographic title, print it in uppercase Bun in background 9 more names found QK Apply Saved List Save This List Gose

- Searching by just the record number, I found 12 instances where we already had an authority record for the name in either our subject or name index.
- When I searched by the old names (option 2), I found several more that we did not have authority records for.
- I found at least 18 names that seemed promising for further investigation.
- A few of them were indeed other people, but there were about 8 names that needed to be updated in the name index alone that didn't have authority records.



- Look through the results and compare them to the original list of names
- Copy the control number from the OCLC list of matched names and use them to run a batch search in OCLC

Enter Authority Batch Search Keys							
Local Elle:	Los	cel File Menager	Batch Se	arch R	eport		Local File: C/User
C-1/Jaen/adk/00/E/AppCate/Reaming/OCL	C'Comex'Db'DefautA	whath do					
Search Kays Gorry: Tota	d Entered: 11	Enter Discritics	Session ID Report Start Tim Session Duration	м 3 п (nxes06. proc 1019-06-081 10100:04.56	f. ocic. org-338 (15: 38:30-05: 84056	85-jvleqibs-vebbbp 00
Use default indeg: [LCCN (m)	-	58	Bibliographic export Authority export	ports reported s reported	0		
In:n 85259745	ĵ,	Delete	Successful Sea	rches (11)			
nii 99053324 nii 99053324 nii 97115353 nii 90 6028217		Brent.	Search Key Ircn 50034831 Ircn 50043013 Ircn 82122243 Ircn 83005345 Ircn 85100269 Ircn 85143026	# Records 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Save # 1 2 3 4 5 6	Control # 70152 78183 818587 901260 1310554 1492831	Title Gough, Kathleen, 1d 1925-1990 Lunt, James, 4d 1917-2001 Broecker, Wallace S., 1d 1931-20 Houston, David, 1d 1935-1993 Loucka, William N. 4q (William N Loucka, William N. 4q (William N Toquair, John Stewart, 1c Earl of
Retrieve all records from griline save fi Limit by Review Status:	lie		In:n 85269745 In:n 86872869 In:n 96053324	1	7 8	1578753 1733024 4099489	Berger, Klaus, #d 1901-2000 Devonshire, R. L., #c Mrs., #d 18 Antico, Plates, #d 1492-1555, #1
			RT 8 99000004			4003403	AGE(10, P10/0, 10 1402/1000, 11

- Save the control numbers as a text file.
- In OCLC, enter authority batch search keys, using the default index of "LCCN (In:)".
- Process the Batch to get a "Batch Search Report" (shown here)

Export updated authority records

Change the headings in the bibliographic record or wait for bibliographic maintenance to make the changes.



- Export into a batch file or individually.
- Here I overlaid each record one at a time so I could look to see what we had.
- Most seem to have a cross reference that should allow bibliographic maintenance to make the changes for me.
- I think in reality, it did two of the seven names I loaded, so I still needed to go back and correct several bib records.



• Check your headings reports for duplicate authority records and invalid headings

Tips

- •Google is your friend to find Regular Expressions or Excel how-to's
- Build a search in the Classic view, then switch over to the JSON view to find how to write the code. Use
 "INSERT VALUE HERE" as your value A and it'll be that much easier to copy your own values
- Now that I've explained my processes, I have a few parting tips.
- Google is your friend.
- If I don't know how to do something in Excel or Regular Expressions, I google it and can usually find a decent formula or two that can meet my needs.
- In Create Lists, build a search in the classic view, then switch over to the JSON view to find how to write the code. Use "INSERT VALUE HERE" as your value A and it'll be that much easier to copy your own values

Tips

·Combine several lists search all at once

- •Avoid copying your Regular Expressions from Word
- •Reuse! Reuse! Reuse!
- Troubleshoot Sierra error codes in Notepad++ to find the line and column easier.
- •Experience will make this easier.
- Wait until you have several lists to combine and then set up one search all at once, instead of setting it up weekly
- Try to avoid copying your Regular Expressions from Word because the automatic-formatting might cause problems.
- For example, a quotation mark gets changed to a slanted quotation mark automatically, which the computer interprets differently.
- Reuse! Reuse! reuse!
- We figure out how to set up one JSON query and then reuse it whenever practical
- You will run into times where the code doesn't want to work.
- Sierra will prevent you from continuing your search with a dialog box.
- This box shows where the problem is but it's hard to work with it in Sierra.
- Use Notepad++ to find the line and column and see what might be wrong.
- Usually you don't have enough brackets or too many, or you might be missing a comma.
- Experience will make this easier.

Resources

- Getting Started with Regular Expressions in MarcEdit
 - <u>https://www.youtube.com/watch?v=7YXvS4xBEfw</u>
- •Regular Expressions in Create Lists
 - https://wiliug.files.wordpress.com/2015/02/matches-handout-2015.pdf
- •Regular Expressions Cheat Sheet
 - <u>https://www.rexegg.com/regex-quickstart.html</u>
- •Batch Processing in the Client: Overview- OCLC
 - https://help.oclc.org/@api/deki/files/5147/connexion-client-batchprocessing.pdf
- Here are some resources that I use when doing this type of work.
- The first is a YouTube video tutorial by Terry Reese, the creator of MarcEdit.
- He really breaks down regular expressions for anyone unfamiliar with them and shows how to do some basic expressions.
- This video is over an hour long.
- The second is a great document on how to use regular expressions in Sierra Create lists.
- You can select "R matches" under "condition" to use these regular expressions in Sierra's create lists.
- The document also covers Millenium.
- I love this document's explanations and examples on how to search for different situations.
- It's really helped me find things I never thought I'd be able to in order to do clean up for our automated authority control vendor prep.
- The third document is a concise listing of different regular expressions, what they mean, and an example.

- Helpful for writing new ones I don't use very often.
- The last document is a link to OCLC's overview of their batch processing in the Connexion client.
- They have a lot of documentation on their website on how to do different things.



- •Zumer, M. (2009). *National bibliographies in the digital age*. Munchen: Saur.
- <u>https://www.loc.gov/aba/cataloging/subject/weeklylists/</u>
- <u>https://www.oclc.org/authority-records.html</u>



My slides and a handout of the procedure should eventually be posted to the conference website as well as to the UNT Institutional Repository. I created the handout as a text file because it's so much easier to copy and paste from without messing up your regular expression. However, there are very limited options for formatting within a text file so it can look a little overwhelming

- Here is the much more technical version of the presentation that has all of the regular expressions and every step explained throughout.
- And if you have any questions I can't answer here or come across a problem when you are trying this back at your library, email me. I'd love to help.
- Thank you for your time!