

# Breaking the Bottleneck

Automating the Reconciliation of Named Entities to the  
Library of Congress Name Authority File

Jenn Wright and Matt Carruthers

University of Michigan Library

# Authority Control Overview

- What it is:
  - We send Backstage lists of headings connected to new records we acquire
    - About 5,000 records on average sent to Backstage per week
  - They run comparisons against their authority files and make as many changes as they can
    - They match about 80% of the records we send them
  - They send those without matches (the leftovers) back to us
- Questions we ask:
  - Are there any authority records for this heading?
  - If so, do we have that authority record?
    - If not, import it
  - Do any of our headings need to be changed (conformed) to match the heading established as the preferred heading in the authority record?

# How It Works: Backstage

## Backstage Authority Control Report: R07 - Unmatched Primary Headings

Report Type: Names

Created for:  
University of Michigan

Created on:  
January 29 2013

The fields in this report include a primary heading which did not match an established heading or cross reference in any authority record. The *primary heading* in a field includes different subfields depending on the heading type. For example, the primary heading in a Personal Name (x00) heading includes the \$a subfield plus any \$c, \$q, or \$d subfields present. For subject (6xx) headings, only authority records in the related national authority file (and internal BSLW authority files) are checked for matching records. For example, a Topical Subject (650) heading with a second indicator zero is only matched against the LC Subject Authority File unless special processing is requested.

The majority of non-65x headings in this report will normally be headings which are valid, but have not yet been established, i.e., an authority record has not yet been added to the national authority file for that heading. Some of the headings, however, will be incorrect forms of established headings that could not be corrected during automated authority control processing.

### Personal Name Fields (100/700)

MIU010000000000000003287189	100	1	\$a [Deschamps de Boishébert, Henri Louis, \$d 1679-1736].
MIU010000000000000002990284	100	1	\$6 880-01 \$a Adham, Sâmi.
MIU010000000000000003324416	700	1	\$a Adkins, Leslie Coleman.
MIU0100000000000000012161434	700	1	\$a Afzali, Mahnaz.
MIU0100000000000000012158555	700	1	\$a Ahmed, Abdulla Rasool.
MIU0100000000000000012185208	100	1	\$a Al Mamun, Tm Abdullah.
MIU0100000000000000009612589	700	1	\$6 880-05 \$a Aldakaeva, N. G.
MIU010000000000000002713305	700	1	\$a Alessandri, Luigi.
MIU010000000000000003288928	100	1	\$a Almada, Andreas d'.
MIU0100000000000000012185407	100	1	\$a Anderson, Latham.
MIU0100000000000000012153856	700	1	\$a Andrade, Fernando Grostein. \$4 drt \$4 cng \$4 pro
MIU0100000000000000011934615	100	1	\$a Ansaloni, Enrico, \$d 1977-
MIU0100000000000000012186656	100	1	\$a Arifin, Syarifuddin, \$d 1956-

[illegible]

From Web From Text From Other Sources Existing Connections Refresh All Connections Properties Advanced Text to Columns									
Get External Data Connections Sort & Filter									
IB									
	A	B	C	D	E	F	G		
1				Name					
2	MIU01000	700	1	Abella, V.					
3	MIU01000	700	1	Acosta, Bernarda.					
4	MIU01000	100	1	Acosta, Yesenia D.					
5	MIU01000	700	1	Acuff, Peter Z.					
6	MIU01000	100	1	Aditya, Wahyu, 1990					
7	MIU01000	700	1	Agrigoroaei, Vladimir.					
8	MIU01000	700	1	Aguirre Chávez, Felipe.					
9	MIU01000	700	1	Ahrens, Thomas J.					
10	MIU01000	700	1	Ahrens, Thomas J.					
11	MIU01000	100	1	Agagündüz, Ülkü Özel.					
12	MIU01000	100	1	Al-Werfalli, Mabroka.					
13	MIU01000	700	1	Alem, Beatriz.					
14	MIU01000	700	1	Alexander, Dwight.					
15	MIU01000	700	1	Alexander, Pamela C.					
16	MIU01000	700	1	Alim, Angel.					
17	MIU01000	700	1	Alonso, Maribel.					
18	MIU01000	700	1	Alonso, Maribel.					
19	MIU01000	700	1	Alonso, Maribel.					
20	MIU01000	700	1	Alonso, Maribel.					
21	MIU01000	700	1	Amper, Grace.					
22	MIU01000	700	1	Andersen, Ronald M.					
23	MIU01000	700	1	Anderson, Carl M.					
24	MIU01000	700	1	Andô, Sakura, actor.					
25	MIU01000	700	1	Ang, HweeSim.					
26	MIU01000	100	1	Angel, G. C.					
27	MIU01000	700	1	Antipow, Lilia.					
28	MIU01000	700	1	Antleman, Kristin.					
29	MIU01000	700	1	Aoki, Jodi, 1956-					



# How It Was: Problems

Is there an authority record?

MIU01000000000000009355132	700	1 \$a	Beklemishev, Lev.	
MIU01000000000000009187801	100	1 \$a	Benjamin, Richard.	
MIU01000000000000009356462	100	1 \$a	Beraud, Pierre.	
MIU01000000000000009459420	100	1 \$a	Bergheaud, Lise.	
MIU01000000000000005237905	100	1 \$a	Bergner, Thomas.	
MIU01000000000000009613856	700	1 \$a	Bermúdez, Daniel, 1950-	
MIU01000000000000009617467	700	1 \$a	Berquin, Armand.	
MIU01000000000000009612951	100	1 \$a	Berthier de Grandry, Frédéric de.	
MIU01000000000000009476003	700	1 \$a	Bettella, Mario.	
MIU01000000000000009617734	700	1 \$a	Bibke, John, \$former owner.	
MIU01000000000000009523988	700	1 \$a	Bickeem, Ramdasha.	
MIU01000000000000009460249	100	1 \$a	Bierlein, Ulrike	
MIU01000000000000007095104	700	1 \$a	Bijelić, Severin, 1921-1971	
MIU01000000000000005642913	700	1 \$a	Bišri, Salim al-	
MIU01000000000000009620381	700	1 \$a	Bizer, Christian.	



## How It Works: GoogleRefine

From Access

From Web

From Text

From Other Sources

Existing Connections

Refresh All

Connections

Properties

Edit Links

Sort

Filter

Reapply

Advanced

Text to Columns

Get External Data

Connections

Sort & Filter

	A	B	C	D	E	F	G
1				Name			
2	MIU01000	700	1 \$a	Abella, V.			
3	MIU01000	700	1 \$a	Acosta, Bernarda.			
4	MIU01000	100	1 \$a	Acosta, Yesenia D.			
5	MIU01000	700	1 \$a	Acuff, Peter Z.			
6	MIU01000	100	1 \$a	Aditya, Wahyu, 1980-			
7	MIU01000	700	1 \$a	Agriгораеi, Vladimir.			
8	MIU01000	700	1 \$a	Aguirre Chávez, Felipe.			
9	MIU01000	700	1 \$a	Ahrens, Thomas J.			
10	MIU01000	700	1 \$a	Ahrens, Thomas J.			
11	MIU01000	100	1 \$a	Akağündüz, Ülkü Özel.			
12	MIU01000	100	1 \$a	Al-Werfalli, Mabroka.			
13	MIU01000	700	1 \$a	Alem, Beatriz.			
14	MIU01000	700	1 \$a	Alexander, Dwight.			
15	MIU01000	700	1 \$a	Alexander, Pamela C.			
16	MIU01000	700	1 \$a	Alim, Angel.			
17	MIU01000	700	1 \$a	Alonso, Maribel.			
18	MIU01000	700	1 \$a	Alonso, Maribel.			
19	MIU01000	700	1 \$a	Alonso, Maribel.			
20	MIU01000	700	1 \$a	Alonso, Maribel.			
21	MIU01000	700	1 \$a	Amper, Grace.			
22	MIU01000	700	1 \$a	Andersen, Ronald M.			
23	MIU01000	700	1 \$a	Anderson, Carl M.			
24	MIU01000	700	1 \$6	Andô, Sakura, actor.			
25	MIU01000	700	1 \$a	Ang, HweeSim.			
26	MIU01000	100	1 \$a	Angel, G. C.			
27	MIU01000	700	1 \$a	Antipow, Lilia.			
28	MIU01000	700	1 \$a	Antleman, Kristin.			
29	MIU01000	700	1 \$a	Aoki, Jodi, 1956-			

C:\Users\jennyw

13:58:43.260

s\jennywri\Doc

13:58:45.335

data path: res

13:58:45.335

data path: res

13:58:45.366

31ms>

13:58:56.700

uration

12:54:14

in 0 sec

13:59:16.700

13:59:16.366

13:59:16.939

13:59:19.947

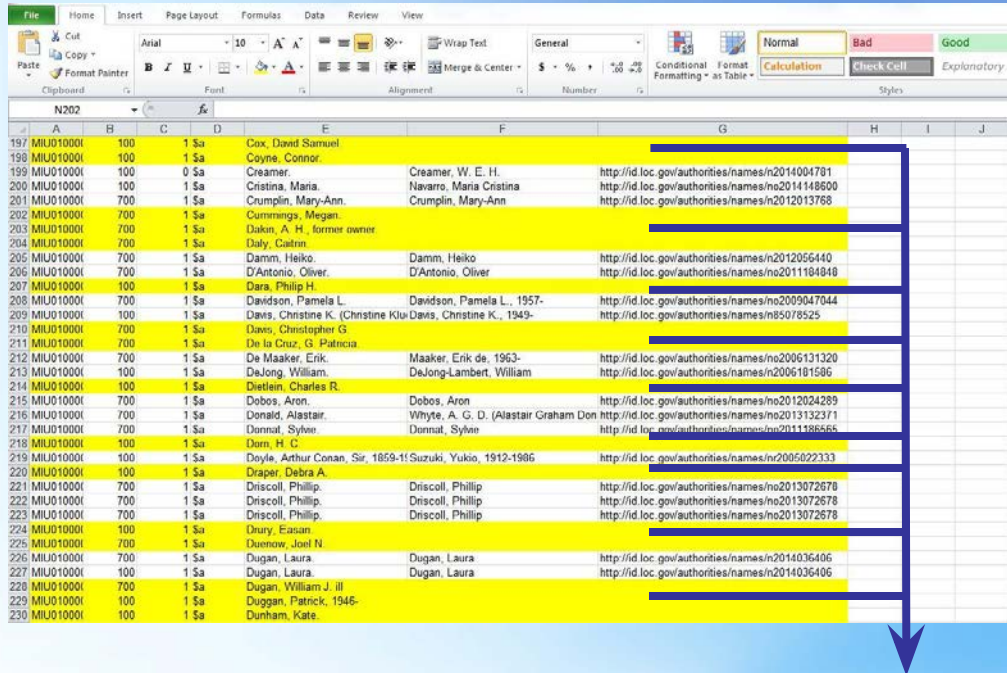
13:59:14.228

```
C:\Users\jennywn\Downloads>google-refine-2.5-r2407\google-refine.exe
```

```
13:58:43.260 [   refine_server] Initializing context: '/' from 'C:\Use  
e\jennywn\Downloads\google-refine-2.5-r2407\webapp' (31ms)  
13:58:45.335 [   refine_server] Failed to use jdatapath to detect user  
data path: resorting to environment variables (2875ms)  
13:58:45.335 [   refine_server] Failed to use jdatapath to detect user  
data path: resorting to environment variables (0ms)  
13:58:45.366 [   refine] Starting Google Refine 2.5 [r2407]...  
31ms)  
13:58:50.790 [   refine] POST /command/core/get-all-importing-confi  
guration (4ms)  
13:58:56 [   refine] GET /command/core/get-all-project-meta  
  
[   refine] GET /command/core/get-version (29ms)  
[   refine] GET /command/core/get-prefer  
  
[   refine] GET /command/core/get-proje  
project] Loaded project 235906136289  
  
in 0 sec (209ms)  
13:59:10.747 [   refine] GET /command/core/get-models (16ms)  
13:59:10.746 [   refine] GET /command/core/get-history (29ms)  
13:59:10.749 [   refine] POST /command/core/get-row (32ms)  
13:59:10.747 [   refine] GET /command/core/get-history (38ms)  
13:59:10.748 [   refine] POST /command/core/get-rows (328ms)
```

[illegible]

# How It Works: Efficiencies



	A	B	C	D	E	F	G	H	I	J
197	MIU010001	100	1	1	Cox, David Samuel					
198	MIU010001	100	1	1	Coyne, Connor					
199	MIU010001	100	0	1	Creamer,	Creamer, W. E. H.	<a href="http://id.loc.gov/authorities/names/n2014004781">http://id.loc.gov/authorities/names/n2014004781</a>			
200	MIU010001	100	1	1	Cristina, Maria	Navarro, Maria Cristina	<a href="http://id.loc.gov/authorities/names/n2014148600">http://id.loc.gov/authorities/names/n2014148600</a>			
201	MIU010001	700	1	1	Crumplin, Mary-Ann	Crumplin, Mary-Ann	<a href="http://id.loc.gov/authorities/names/n2012013768">http://id.loc.gov/authorities/names/n2012013768</a>			
202	MIU010001	700	1	1	Cummings, Megan					
203	MIU010001	700	1	1	Dakin, A. H., former owner					
204	MIU010001	700	1	1	Daly, Cadron					
205	MIU010001	700	1	1	Damm, Heiko	Damm, Heiko	<a href="http://id.loc.gov/authorities/names/n2012056440">http://id.loc.gov/authorities/names/n2012056440</a>			
206	MIU010001	700	1	1	D'Antonio, Oliver	D'Antonio, Oliver	<a href="http://id.loc.gov/authorities/names/n2011184848">http://id.loc.gov/authorities/names/n2011184848</a>			
207	MIU010001	100	1	1	Dara, Philip H.					
208	MIU010001	700	1	1	Davidson, Pamela L.	Davidson, Pamela L., 1957-	<a href="http://id.loc.gov/authorities/names/n2009047044">http://id.loc.gov/authorities/names/n2009047044</a>			
209	MIU010001	100	1	1	Davis, Christine K. (Christine Klu Davis, Christine K., 1949-		<a href="http://id.loc.gov/authorities/names/n85078525">http://id.loc.gov/authorities/names/n85078525</a>			
210	MIU010001	700	1	1	Davis, Christopher G.					
211	MIU010001	700	1	1	De la Cruz, G. Patricia					
212	MIU010001	700	1	1	De Masker, Erik	Maaker, Erik de, 1963-	<a href="http://id.loc.gov/authorities/names/n2006131320">http://id.loc.gov/authorities/names/n2006131320</a>			
213	MIU010001	100	1	1	DeJong, William	DeJong-Lambert, William	<a href="http://id.loc.gov/authorities/names/n2006181586">http://id.loc.gov/authorities/names/n2006181586</a>			
214	MIU010001	100	1	1	Dietlein, Charles R.					
215	MIU010001	700	1	1	Dobos, Aron	Dobos, Aron	<a href="http://id.loc.gov/authorities/names/n2012024289">http://id.loc.gov/authorities/names/n2012024289</a>			
216	MIU010001	700	1	1	Donald, Alastair	Whyte, A. G. D. (Alastair Graham Don	<a href="http://id.loc.gov/authorities/names/n2013132371">http://id.loc.gov/authorities/names/n2013132371</a>			
217	MIU010001	700	1	1	Donnat, Sylvie	Donnat, Sylvie	<a href="http://id.loc.gov/authorities/names/n2011188665">http://id.loc.gov/authorities/names/n2011188665</a>			
218	MIU010001	100	1	1	Dorn, H. C.					
219	MIU010001	100	1	1	Doyle, Arthur Conan, Sir, 1859-11	Suzuki, Yukio, 1912-1986	<a href="http://id.loc.gov/authorities/names/n2005022333">http://id.loc.gov/authorities/names/n2005022333</a>			
220	MIU010001	100	1	1	Draper, Debra A.					
221	MIU010001	700	1	1	Driscoll, Phillip	Driscoll, Phillip	<a href="http://id.loc.gov/authorities/names/n2013072678">http://id.loc.gov/authorities/names/n2013072678</a>			
222	MIU010001	700	1	1	Driscoll, Phillip	Driscoll, Phillip	<a href="http://id.loc.gov/authorities/names/n2013072678">http://id.loc.gov/authorities/names/n2013072678</a>			
223	MIU010001	700	1	1	Driscoll, Phillip	Driscoll, Phillip	<a href="http://id.loc.gov/authorities/names/n2013072678">http://id.loc.gov/authorities/names/n2013072678</a>			
224	MIU010001	100	1	1	Duffy, Eavan					
225	MIU010001	700	1	1	Duenow, Joel N.					
226	MIU010001	700	1	1	Dugan, Laura	Dugan, Laura	<a href="http://id.loc.gov/authorities/names/n2014036406">http://id.loc.gov/authorities/names/n2014036406</a>			
227	MIU010001	100	1	1	Dugan, Laura	Dugan, Laura	<a href="http://id.loc.gov/authorities/names/n2014036406">http://id.loc.gov/authorities/names/n2014036406</a>			
228	MIU010001	700	1	1	Dugan, William J. III					
229	MIU010001	100	1	1	Duggan, Patrick, 1946-					
230	MIU010001	100	1	1	Dunham, Kate					

No authorities?

No problem!

Just move them on down to the end of the spreadsheet.

They'll contribute to a pool of candidates waiting for NACO authority record creation.

# Technical Details

## Behind the Scenes:

- Google Refine automatically searches the Virtual International Authority File (VIAF) for matching headings from the Library of Congress Name Authority File
- It then returns any matching headings and links to corresponding records



# Technical Details

rough NEW

Permalink

Create column Fetch at index 1 by fetching URLs based on column Name using expression  
grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22" + value +  
"%22+and+local.sources+%3D+%22lc%22&sortKeys=holdingscount&httpAccept=application/xml"  
5% complete (15 other pending processes) Cancel All

Open... Export... He

20 rows

Extensions: Freebase RDF

Show as: rows records Show: 5 10 25 50 rows

« first < previous 1 - 20 next > last

All	Name
	1. Akashkin,%20M.%20M.
	2. Alnozi,%20Rasheed.
	3. Alver,%20Fusun.
	4. An,%20Binh%20Minh.
	5. Anhalt,%20Edward.
	6. Araujo,%20Marian.
	7. Ashikawa,%20Makoto.
	8. Atkinson,%20R.%20J.%20A.
	9. Axler,%20Rachel.
	10. Babia,%20Roger.
	11. Bai,%20Juncal.
	12. Ballarín,%20Joe.
	13. Barret,%20Justin.
	14. Berg,%20William%20S.
	15. Bess,%20Ardon.
	16. Blankin,%20Petr.
	17. Blaser,%20Paul.
	18. Boatman,%20Robert.
	19. Bowles,%20Cory.
	20. Brooks,%20Geraldine,%201955-

# Technical Details

rough NEW [Permalink](#)

Create column Fetch at index 1 by fetching URLs based on column Name using expression  
grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22" + value +  
"%22+and+local.sources+%3D+%22lc%22&sortKeys=holdingscount&httpAccept=application/xml"  
5% complete (15 other pending processes) [Cancel All](#)

20 rows

Show as: **rows** records Show: 5 10 25 50 rows

Extensions: **Freebase** | RDF

« first « previous **1 - 20** next » last

All	Name
	1. Akashkin,%20M.%20M.
	2. Alnozi,%20Rasheed.
	3. Alver,%20Fusun.
	4. An,%20Binh%20Minh.
	5. Anhalt,%20Edward.
	6. Araujo,%20Marian.
	7. Ashikawa,%20Makoto.
	8. Atkinson,%20R.%20J.%20A.
	9. Axler,%20Rachel.
	10. Babia,%20Roger.
	11. Bai,%20Juncal.
	12. Ballarín,%20Joe.
	13. Barret,%20Justin.
	14. Berg,%20William%20S.
	15. Bess,%20Ardon.
	16. Blankin,%20Petr.
	17. Blaser,%20Paul.
	18. Boatman,%20Robert.
	19. Bowles,%20Cory.
	20. Brooks,%20Geraldine,%201955-

# Technical Details

Google refine Presentation FINAL Permalink Open... Export ▾ Help ▾

Facet / Filter Undo / Redo 20 18 rows Extensions: Named-entity recognition ▾ Freebase ▾ RDF ▾

Extract... Apply...

Show as: rows records Show: 5 10 25 50 rows « first « previous 1 - 18 next » last

Filter:

12. Text transform on 3 cells in column LC Record Link: `grel:replace(value, ',', '')`

13. Create column LCNAF Heading at index 3 by fetching URLs based on column LC Record Link using expression `grel:"http://viaf.org/processed/LC" + value + "?httpAccept=text/xml"`

14. Text transform on 8 cells in column LCNAF Heading: `grel:value.parseHtml().select('mx[datafield[0].htmlText()')`

15. Text transform on 0 cells in column LCNAF Heading: `value.trim()`

16. Text transform on 0 cells in column LCNAF Heading: `value.replace(/s+/, '')`

17. Text transform on 8 cells in column LC Record Link: `grel:"http://id.loc.gov/authorities/names/" + value`

18. Move column LC Record Link to position

▼ All	▼ Name	▼ LCNAF Heading	▼ LC Record Link
☆	1. Alver, Füsün.	Alver, Füsün	<a href="http://id.loc.gov/authorities/names/no2011170601">http://id.loc.gov/authorities/names/no2011170601</a>
☆	2. An, Binh Minh.	An Binh Minh, 1948-	<a href="http://id.loc.gov/authorities/names/no2011167328">http://id.loc.gov/authorities/names/no2011167328</a>
☆	3. Anhalt, Edward.	Anhalt, Edward, 1914-2000	<a href="http://id.loc.gov/authorities/names/n87911796">http://id.loc.gov/authorities/names/n87911796</a>
☆	4. Araujo, Marian.		
☆	5. Ashikawa, Makoto.		
☆	6. Atkinson, R. J. A.		
☆	7. Axler, Rachel.	Axler, Rachel	<a href="http://id.loc.gov/authorities/names/nb2011024512">http://id.loc.gov/authorities/names/nb2011024512</a>
☆	8. Babia, Roger.		
☆	9. Bai, Juncai.	Bai, Juncai, 1970-	<a href="http://id.loc.gov/authorities/names/no2009146750">http://id.loc.gov/authorities/names/no2009146750</a>
☆	10. Ballarini, Joe.		
☆	11. Barret, Justin.		
☆	12. Berg, William S.		
☆	13. Bess, Ardon.		
☆	14. Blankin, Petr.		
☆	15. Blaser, Paul.		
☆	16. Boatman, Robert.	Boatman, Robert H.	<a href="http://id.loc.gov/authorities/names/n2003010628">http://id.loc.gov/authorities/names/n2003010628</a>
☆	17. Bowles, Cory.	Bowles, Cory, 1973-	<a href="http://id.loc.gov/authorities/names/no2010126923">http://id.loc.gov/authorities/names/no2010126923</a>
☆	18. Brooks, Geraldine, 1955-	Brooks, Geraldine	<a href="http://id.loc.gov/authorities/names/n94041622">http://id.loc.gov/authorities/names/n94041622</a>

# Technical Details

Google refine Presentation FINAL Permalink

Open... Export ▾ Help

Facet / Filter Undo / Redo 0

19 rows

Extensions: Named-entity recognition ▾ Freebase ▾ RDF ▾

Extract... Apply...

Show as: rows records Show: 5 10 25 50 rows

« first < previous 1 - 19 next > last

Filter:

0. Create project

1. Reorder rows

2. Blank down 1 cells in column Name

3. Remove 1 rows

4. Text transform on 18 cells in column Name: `grel:replace(value, ',', '%20')`

5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22c%22"`

6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel:value.parseHtml().select('ns2|source')`

7. Text transform on 8 cells in column LC Record Link: `grel:substring(value, indexOf(value, 'C'))`

All	Name
☆	1. Alver, Fusun.
☆	2. An, Binh Minh.
☆	3. Blaser, Paul.
☆	4. Boatman, Robert.
☆	5. Bowles, Cory.
☆	6. Brooks, Geraldine, 1955-
☆	7. Brooks, Geraldine, 1955-
☆	8. Axler, Rachel.
☆	9. Babia, Roger.
☆	10. Bai, Juncal.
☆	11. Ballarini, Joe.
☆	12. Barret, Justin.
☆	13. Berg, William S.
☆	14. Bess, Ardon.
☆	15. Blankin, Petr.
☆	16. Anhalt, Edward.
☆	17. Araujo, Marian.
☆	18. Ashikawa, Makoto.
☆	19. Atkinson, R. J. A.

# Technical Details

Google refine Presentation FINAL Permalink

Open... Export ▾ Help

Facet / Filter Undo / Redo 0 19 rows Extensions: Named-entity recognition ▾ Freebase ▾ RDF ▾

Extract... Apply...

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 19 next > last

Filter:

0. Create project

1. Reorder rows

2. Blank down 1 cells in column Name

3. Remove 1 rows

4. Text transform on 18 cells in column Name: `grel:replace(value, ' ', '%20')`

5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22c%22"`

6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel:value.parseHtml().select('ns2[source]')`

7. Text transform on 8 cells in column LC Record Link: `grel:substring(value, indexOf(value, 'LC'))`

All	Name
☆	1. Alver, Fusun.
☆	2. An, Binh Minh.
☆	3. Blaser, Paul.
☆	4. Boatman, Robert.
☆	5. Bowles, Cory.
☆	6. Brooks, Geraldine, 1955-
☆	7. Brooks, Geraldine, 1955-
☆	8. Axler, Rachel.
☆	9. Babia, Roger.
☆	10. Bai, Juncal.
☆	11. Ballarini, Joe.
☆	12. Barret, Justin.
☆	13. Berg, William S.
☆	14. Bess, Ardon.
☆	15. Blankin, Petr.
☆	16. Anhalt, Edward.
☆	17. Araujo, Marian.
☆	18. Ashikawa, Makoto.
☆	19. Atkinson, R. J. A.



# Technical Details

The screenshot displays the Google Refine web application interface. At the top, the title bar reads "Google refine Presentation FINAL Permalink". Navigation buttons include "Open...", "Export", and "Help". Below the title bar, there are tabs for "Facet / Filter", "Undo / Redo", "Extract...", and "Apply...". The left sidebar shows a list of operations: "0. Create project", "1. Reorder rows", "2. Blank down 1 cells in column Name", "3. Remove 1 rows", "4. Text transform on 18 cells in column Name: `grel:replace(value, ' ', '%20')`", "5. Create column Fetch at index 1 by fetching URLs based on column Name using expression `grel:'http://viaf.org/viaf/search/viaf?query=local_personalNames+%3D+%22'+value+'%22+and+local.sources+%3D+%22ic%2'`", "6. Create new column LC Record Link based on column Fetch by filling 18 rows with `grel:value.parseHtml().select('ns2[source]')`", "7. Text transform on 8 cells in column LC Record Link: `grel:substring(value, indexOf(value, 'LC'))`", "8. Text transform on 8 cells in column LC Record Link: `grel:substring(value, 0, indexOf(value, '<'))`", and "9. Text transform on 8 cells in column LC". The main area is partially obscured by a dialog box titled "Apply Operation History". This dialog box contains the instruction "Paste an extracted JSON history of operations to perform:" followed by a large text input area. At the bottom of the dialog are "Perform Operations" and "Cancel" buttons. In the background, on the right, there are dropdown menus for "ed-entity recognition", "Freebase", and "RDF", along with pagination controls: "« first < previous 1 - 19 next > last »".

# Technical Details

Google refine Presentation FINAL Permalink

Facet / Filter Undo / Redo Extract... Apply...

Filter:

0. Create project
1. Reorder rows
2. Blank down 1 cells in column Name
3. Remove 1 rows
4. Text transform on 18 cells in column Name: `grel.replace(value, ' ', '%20')`
5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel: 'http://viaf.org/viaf/search/viaf?query=local_personalNames+%3D+%22'+value+'%22+and+local_sources+%3D+%22ic%2'`
6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel.value.parseHtml().select('ns2[source]')`
7. Text transform on 8 cells in column LC Record Link: `grel.substring(value, indexOf(value, 'LC'))`
8. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`
9. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`

Apply Operation History

Paste an extracted JSON history of operations to perform:

```
{
  "onError": "keep-original",
  "repeat": false,
  "repeatCount": 10
},
{
  "op": "core/column-move",
  "description": "Move column LC Record Link to position 3",
  "columnName": "LC Record Link",
  "index": 3
},
{
  "op": "core/column-removal",
  "description": "Remove column Fetch",
  "columnName": "Fetch"
},
{
  "op": "core/text-transform",
  "description": "Text transform on cells in column Name using expression grel.replace(value, '%20'",
  "engineConfig": {
    "facets": [],
    "mode": "row-based"
  },
  "columnName": "Name",
  "expression": "grel.replace(value, '%20', ' ')",
  "onError": "keep-original",
  "repeat": false,
  "repeatCount": 10
}
]
```

Perform Operations Cancel

« first < previous 1 - 19 next > last »

Open... Export Help

ned-entity recognition Freebase RDF

# Technical Details

Google refine Presentation FINAL Permalink

Facet / Filter Undo / Redo Extract... Apply...

Filter:

0. Create project
1. Reorder rows
2. Blank down 1 cells in column Name
3. Remove 1 rows
4. Text transform on 18 cells in column Name: `grel.replace(value, ' ', '%20')`
5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel: 'http://viaf.org/viaf/search/viaf?query=local_personalNames+%3D+%22'+value+'%22+and+local_sources+%3D+%22ic%2'`
6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel.value.parseHtml().select('ns2[source]')`
7. Text transform on 8 cells in column LC Record Link: `grel.substring(value, indexOf(value, 'LC'))`
8. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`
9. Text transform on 8 cells in column LC

Apply Operation History

Paste an extracted JSON history of operations to perform:

```
{
  "onError": "keep-original",
  "repeat": false,
  "repeatCount": 10
},
{
  "op": "core/column-move",
  "description": "Move column LC Record Link to position 3",
  "columnName": "LC Record Link",
  "index": 3
},
{
  "op": "core/column-removal",
  "description": "Remove column Fetch",
  "columnName": "Fetch"
},
{
  "op": "core/text-transform",
  "description": "Text transform on cells in column Name using expression grel.replace(value, '%20'",
  "engineConfig": {
    "facets": [],
    "mode": "row-based"
  },
  "columnName": "Name",
  "expression": "grel.replace(value, '%20', ' ')",
  "onError": "keep-original",
  "repeat": false,
  "repeatCount": 10
}
]
```

Perform Operations Cancel

« first < previous 1 - 19 next > last »

Open... Export Help

ned-entity recognition Freebase RDF

# Technical Details

Google refine Presentation FINAL Permalink

Open... Export Help

Facet / Filter Undo / Redo

Extract... Apply...

Filter:

19 rows

Extensions: Named-entity recognition Freebase RDF

Show as: rows records Show: 5 10 25 50 rows

« first < previous 1 - 19 next > last »

▼ All ▼ Name

☆	1.	Alver, Fusun.
☆	2.	An, Binh Minh.
☆	3.	Anhalt, Edward.
☆	4.	Araujo, Marian.
☆	5.	Ashikawa, Makoto.
☆	6.	Atkinson, R. J. A.
☆	7.	Axler, Rachel.
☆	8.	Babia, Roger.
☆	9.	Bai, Juncal.
☆	10.	Ballarini, Joe.
☆	11.	Barret, Justin.
☆	12.	Berg, William S.
☆	13.	Bess, Ardon.
☆	14.	Blankin, Petr.
☆	15.	Blaser, Paul.
☆	16.	Boatman, Robert.
☆	17.	Bowles, Cory.
☆	18.	Brooks, Geraldine, 1955-
☆	19.	Brooks, Geraldine, 1955-

0. Create project

1. Reorder rows

2. Blank down 1 cells in column Name

3. Remove 1 rows

4. Text transform on 18 cells in column Name: `grel.replace(value, '', '%20')`

5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22lc%22"`

6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel.value.parseHtml().select("ns2[source]")`

7. Text transform on 8 cells in column LC Record Link: `grel.substring(value, indexOf(value, 'LC'))`

8. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`

9. Text transform on 8 cells in column LC

javascript({})

# Technical Details

Google refine Presentation FINAL Permalink

Open... Export Help

Facet / Filter Undo / Redo 2

Extract... Apply...

Filter:

19 rows

Extensions: Named-entity recognition Freebase RDF

Show as: rows records Show: 5 10 25 50 rows

« first < previous 1 - 19 next > last »

0. Create project

1. Reorder rows

2. Blank down 1 cells in column Name

3. Remove 1 rows

4. Text transform on 18 cells in column Name: `grel:replace(value, ' ', '%20')`

5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22lc%22"`

6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel:value.parseHtml().select('ns2[source]')`

7. Text transform on 8 cells in column LC Record Link: `grel:substring(value, indexOf(value, 'LC'))`

8. Text transform on 8 cells in column LC Record Link: `grel:substring(value, 0, indexOf(value, '<'))`

9. Text transform on 8 cells in column LC

javascript: {}

All	Name
1.	Alver, Füsün.
2.	An, Binh Minh.
3.	Anhalt, Edward.
4.	Araujo, Marian.
5.	Ashikawa, Makoto.
6.	Atkinson, R. J. A.
7.	Axler, Rachel.
8.	Babia, Roger.
9.	Bai, Juncal.
10.	Ballarini, Joe.
11.	Barret, Justin.
12.	Berg, William S.
13.	Bess, Ardon.
14.	Blankin, Petr.
15.	Blaser, Paul.
16.	Boatman, Robert.
17.	Bowles, Cory.
18.	Brooks, Geraldine, 1955-
19.	



# Technical Details

Google refine Presentation FINAL Permalink Open... Export Help

Facet / Filter Undo / Redo Extract... Apply... **18 rows** Extensions: Named-entity recognition Freebase RDF

Show as: rows records Show: 5 10 25 50 rows « first « previous **1 - 18** next » last

Filter:

0. Create project

1. Reorder rows

2. Blank down 1 cells in column Name

3. Remove 1 rows

4. Text transform on 18 cells in column Name: `grel.replace(value, ' ', '%20')`

5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22lc%22"`

6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel.value.parseHtml().select("ns2[source]").`

7. Text transform on 8 cells in column LC Record Link: `grel.substring(value, indexOf(value, 'LC'))`

8. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`

9. Text transform on 8 cells in column LC Record Link: `grel.trim()`

All	Name
	1. Alver, Fusun.
	2. An, Binh Minh.
	3. Anhalt, Edward.
	4. Araujo, Marian.
	5. Ashikawa, Makoto.
	6. Atkinson, R. J. A.
	7. Axler, Rachel.
	8. Babia, Roger.
	9. Bai, Juncal.
	10. Ballarini, Joe.
	11. Barret, Justin.
	12. Berg, William S.
	13. Bess, Ardon.
	14. Blankin, Petr.
	15. Blaser, Paul.
	16. Boatman, Robert.
	17. Bowles, Cory.
	18. Brooks, Geraldine, 1955-

# Technical Details

Google refine Presentation FINAL Permalink Open... Export Help

Facet / Filter Undo / Redo Extract... Apply... **18 rows** Extensions: Named-entity recognition Freebase RDF

Show as: rows records Show: 5 10 25 50 rows « first « previous **1 - 18** next » last

Filter:

- 0. Create project
- 1. Reorder rows
- 2. Blank down 1 cells in column Name
- 3. Remove 1 rows
- 4. Text transform on 18 cells in column Name: `grel.replace(value, '', '%20')`**
- 5. Create column Fetch at index 1 by fetching URLs based on column Name using expression  
`grel:"http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22"+value+"%22+and+local.sources+%3D+%22lc%22"`
- 6. Create new column LC Record Link based on column Fetch by filling 18 rows with  
`grel.value.parseHtml().select('ns2[source]')`
- 7. Text transform on 8 cells in column LC Record Link: `grel.substring(value, indexOf(value, 'LC'))`
- 8. Text transform on 8 cells in column LC Record Link: `grel.substring(value, 0, indexOf(value, '<'))`
- 9. Text transform on 8 cells in column LC

javascript:[]

All	Name
	1. Alver,%20Fusun.
	2. An,%20Binh%20Minh.
	3. Anhalt,%20Edward.
	4. Araujo,%20Marian.
	5. Ashikawa,%20Makoto.
	6. Atkinson,%20R.%20J.%20A.
	7. Axler,%20Rachel.
	8. Babia,%20Roger.
	9. Bai,%20Juncal.
	10. Ballarini,%20Joe.
	11. Barret,%20Justin.
	12. Berg,%20William%20S.
	13. Bess,%20Ardon.
	14. Blankin,%20Petr.
	15. Blaser,%20Paul.
	16. Boatman,%20Robert.
	17. Bowies,%20Cory.
	18. Brooks,%20Geraldine,%201955-

[illegible]

“http://viaf.org/viaf/search/viaf?  
query=local.personalNames+%  
3D+%22" + value + "%  
22+and+local.sources+%3D+%  
22lc%  
22&httpAccept=application/xml”

[illegible]

“[http://viaf.org/viaf/search/viaf?](http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22)  
[query=local.personalNames+%3D+%22](http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22) + value + "%  
 22+and+local.sources+%3D+%22lc%  
 22&httpAccept=application/xml”

[illegible]

“http://viaf.org/viaf/search/viaf?  
query=local.personalNames+%  
3D+%22" + value + "%  
22+and+local.sources+%3D+%  
22lc%  
22&httpAccept=application/xml”



[illegible]

“http://viaf.org/viaf/search/viaf?  
query=local.personalNames+%  
3D+%22" + value + "%  
22+and+local.sources+%3D+%  
22lc%  
22&httpAccept=application/xml”

[illegible]

“http://viaf.org/viaf/search/viaf?  
query=local.personalNames+%  
3D+%22" + value + "%  
22+and+local.sources+%3D+%  
22lc%  
22&httpAccept=application/xml”

[illegible]

“http://viaf.org/viaf/search/viaf?  
query=local.personalNames+%  
3D+%22" + value + "%  
22+and+local.sources+%3D+%  
22lc%  
22&httpAccept=application/xml”

# Technical Details

[illegible]



# Technical Details

[illegible]



## Technical Details

**Google refine** Presentation FINAL Permalink

Facet / Filter    Undo / Redo 12    Extensions: Named-entity recognition ▾    Freebase ▾    RDF ▾

Show as: rows records    Show: 5 10 25 50 rows

	Name	Fetch	LC Record Link
1.	Alver, %20Fusun.	<?xml version="1.0" encoding="UTF-8"?><searchRetrieveResponse xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.loc.gov/zing/srv/ http://www.loc.gov/standards/sru/srui-1/archive/xml-feds/srv-types.xsd" xmlns="http://www.loc.gov/zing/srv/"><version xsi:type="xsd:string">1.1</version><numberOfRecords xsi:type="xsd:nonNegativeInteger">1</numberOfRecords><resultSetTime xsi:type="xsd:positiveInteger">1</resultSetTime><records xsi:type="ns1:recordsType" xmlns:ns1="http://www.loc.gov/zing/srv/"><record xsi:type="ns1:recordType"><recordSchema xsi:type="xsd:string">http://viaf.org/VIAFCluster/</recordSchema><recordPacking xsi:type="xsd:string">xml-recordPacking=<recordData xsi:type="xsi:stringOrXMLFragment"><ns2:VIAFCluster xmlns:"http://www.loc.gov/viaf/">xmles:foaf="http://xmlns.com/foaf/1.1"/>xmles:owl="http://www.w3.org/2002/07/owl#" xmles:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmles:void="http://rdfs.org/void#" xmles:ns2="http://viaf.org/viaf/terms#"><ns2:viafID>140304808</ns2:viafID><ns2:Document about="http://viaf.org/viaf/140304808/"><ns2:inDataset resource="http://viaf.org/viaf/data/"><ns2:primaryTopic resource="http://viaf.org/viaf/140304808/"><ns2:Document><ns2:nameType>Personal</ns2:nameType><ns2:sources><ns2:source nsid="0000000357915378">ISNI 0000000357915378</ns2:source><ns2:source nsid="no2011170601">LC no2011170601</ns2:source></ns2:sources><ns2:differentiated="false" nsid="http://id-nb.info/gnd/141752971" sparse="true">DNB 141752971</ns2:source><ns2:length>16</ns2:length><ns2:mainHeadings><ns2:data><ns2:text>Alver, F&#xF0;FC,sun</ns2:text></ns2:sources><ns2:s>DNB</ns2:s><ns2:s>LC</ns2:s><ns2:s>ISNI</ns2:s><ns2:links><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170601</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, F&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>ISNI</ns2:s><ns2:sources><ns2:link>ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="viaf7">LC no2011170601</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>DNB</ns2:s><ns2:sources><ns2:link>DNB 141752971</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170601</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>DNB</ns2:s><ns2:sources><ns2:link>DNB 141752971</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170601</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>DNB</ns2:s><ns2:sources><ns2:link>DNB 141752971</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170601</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>DNB</ns2:s><ns2:sources><ns2:link>DNB 141752971</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170601</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>DNB</ns2:s><ns2:sources><ns2:link>DNB 141752971</ns2:link><ns2:link><ns2:match type="viaf7">ISNI 0000000357915378</ns2:link><ns2:link><ns2:match type="exact name">DNB 141752971</ns2:link><ns2:links><ns2:mainHeadingEl><ns2:mainHeadingEl><ns2:dataProvider>MARC21</ns2:dataProvider><ns2:id>"tag=100"><ns2:subfield code="a">Alver, Fu&#xF0;FC,sun</ns2:subfield></ns2:dataProvider></ns2:sources><ns2:s>LC</ns2:s><ns2:sources><ns2:link>LC no2011170	

# Technical Details

Google refine Example for Presentation Permalink

Facet / Filter Undo / Redo 31

362 rows

Show as: rows records Show: 5 10 25 50 rows

Filter

Extract... Apply

1. Create project

2. Blank down 3 cells in column Name

3. Remove 3 rows

4. Text transform on 357 cells in column Name: grel.replace(value, " ", "%20")

5. Create column Fetch at index 5 by fetching URLs based on column Name using expression: grel."http://viaf.org/viaf/search/viaf?query=local.personalNames+%3D+%22" + value + "%22+and+local.sources+%3D+%22c%2" + value + "%22"

6. Create new column LCCN based on column Fetch by filling 362 rows with: grel.value.parse(html).select(m2(source))

7. Text transform on 177 cells in column LCCN: grel.substring(value, indexOf(value, "LC"))

8. Text transform on 177 cells in column LCCN: grel.substring(value, 0, indexOf(value, "%"))

9. Text transform on 177 cells in column LCCN: value.replace("%", "/")

10. Text transform on 177 cells in column LCCN: grel.substring(value, 3)

11. Text transform on 66 cells in column LCCN: grel.replace(value, " ", "/")

12. Create column LC Heading at index 7 by fetching URLs based on column LCCN using expression: grel."http://viaf.org/processed/LC/" + value + "&httpAccept=text/xml"

All	Local ID	Tag	D1	S1	Name	LCNAF Heading	LC Record Link
1	MU010000000000000000000000728	700	1	6	Akashia, M. M.		
2	MU01000000000000000000000072199	700	1	5a	Abricc, Racheed		
3	MU0100000000000000000000000000000	100	1	5a	Aizer, Froux	Aizer, Froux	http://id.loc.gov/authfiles/names/ho2011170001
4	MU0100000000000000000000000000000	100	0	5a	As, Bish Moh	As, Bish Moh, 1945.	http://id.loc.gov/authfiles/names/ho2011107128
5	MU0100000000000000000000000000000	700	1	5a	Anhalt, Edward	Anhalt, Edward, 1914-2000	http://id.loc.gov/authfiles/names/ho7911798
6	MU0100000000000000000000000000000	700	1	5a	Arzou, Marian		
7	MU0100000000000000000000000000000	700	1	5a	Ashkara, Makdis		
8	MU0100000000000000000000000000000	700	1	5a	Atkinson, R. J. A.		
9	MU0100000000000000000000000000000	700	1	5a	Avier, Rachel	Avier, Rachel	http://id.loc.gov/authfiles/names/ho2011024312
10	MU0100000000000000000000000000000	700	1	5a	Bala, Roger		
11	MU0100000000000000000000000000000	100	1	6	Bai, Juncal	Bai, Juncal, 1975.	http://id.loc.gov/authfiles/names/ho2011067750
12	MU0100000000000000000000000000000	700	1	5a	Bakari, Joe		
13	MU0100000000000000000000000000000	700	1	5a	Barrel, Justit		
14	MU0100000000000000000000000000000	700	1	5a	Berg, William S		
15	MU0100000000000000000000000000000	700	1	5a	Bess, Ardon		
16	MU0100000000000000000000000000000	700	1	6	Blank, Petr		
17	MU0100000000000000000000000000000	700	1	5a	Blaser, Paul		
18	MU0100000000000000000000000000000	700	1	5a	Boutman, Robert	Boutman, Robert	http://id.loc.gov/authfiles/names/ho2003010628
19	MU0100000000000000000000000000000	700	1	5a	Bowles, Cory	Bowles, Cory, 1975.	http://id.loc.gov/authfiles/names/ho2011020223
20	MU0100000000000000000000000000000	700	1	5a	Brooks, Geratline	Brooks, Geratline, 1905.	http://id.loc.gov/authfiles/names/ho4041622
21	MU0100000000000000000000000000000	700	1	5a	Budono, Hendy		
22	MU0100000000000000000000000000000	100	1	5a	Bueno de Mesquita, D. M. (Daniel Meredit)	Bueno de Mesquita, D. M. (Daniel Meredit)	http://id.loc.gov/authfiles/names/ho2005031680
23	MU0100000000000000000000000000000	100	1	5a	Burke, Richard James, 1917.		
24	MU0100000000000000000000000000000	100	1	5a	Cai, Tanqun	Cai, Tanqun	http://id.loc.gov/authfiles/names/ho2011060156
25	MU0100000000000000000000000000000	700	1	5a	Capobianco, Carlisa		
26	MU0100000000000000000000000000000	700	1	5a	Carlson, Carlos Roberto, 1955.	Carlson, Carlos Roberto, 1955.	http://id.loc.gov/authfiles/names/ho2011151077
27	MU0100000000000000000000000000000	100	1	5a	Carrigan, Joseph G. Jr. 1935.		
28	MU0100000000000000000000000000000	700	1	5a	Cases, Chloé		
29	MU0100000000000000000000000000000	700	1	5a	Cawley, Tucker	Cawley, Tucker	http://id.loc.gov/authfiles/names/ho2014009752
30	MU0100000000000000000000000000000	700	1	6	Chae, Tong-ho	Chae, Tong-ho, 1945.	http://id.loc.gov/authfiles/names/ho3161568
31	MU0100000000000000000000000000000	700	1	5a	Chan, Cecely		
32	MU0100000000000000000000000000000	100	1	5a	Chapman, Dean E.		
33	MU0100000000000000000000000000000	700	1	5a	Chapman, Rubie J.	Chapman, Rubie (Rubie) John	http://id.loc.gov/authfiles/names/ho2012002236
34	MU0100000000000000000000000000000	100	1	5a	Chavez Vium, Sebastian	Chavez Vium, Sebastian	http://id.loc.gov/authfiles/names/ho2011150536

## Second Query for VIAF API:

"http://viaf.org/processed/LC|" + value + "?"  
httpAccept=text/xml"

# Technical Details

Google refine Example for Presentation Permalink

Facet / Filter Undo / Redo 31

362 rows

Show as: rows records Show: 5 to 25 56 rows

Filter Extract Apply

1. Create project

2. Blank down 3 cells in column Name

3. Remove 3 rows

4. Text transform on 357 cells in column Name: gml.replace(value, "", %20)

5. Create column Fetch at index 5 by fetching URLs based on column Name using expression: gml."http://viaf.org/viaf/search/viaf?query=local\_personalNames+%3D+%22" + value + "%22+and+local+sources+%3D+%22c%22"

6. Create new column LCCN based on column Fetch by filling 362 rows with gml.value.parse(html).select(m2.source)

7. Text transform on 177 cells in column LCCN: gml.substring(value, indexOf(value, "/"))

8. Text transform on 177 cells in column LCCN: gml.substring(value, 0, indexOf(value, "/"))

9. Text transform on 177 cells in column LCCN: value.replace(/s+/, "")

10. Text transform on 177 cells in column LCCN: gml.substring(value, 3)

11. Text transform on 66 cells in column LCCN: gml.replace(value, "", "/")

12. Create column LC Heading at index 7 by fetching URLs based on column LCCN using expression: gml."http://viaf.org/processed/LC/" + value + "&httpAccept=text/xml"

All	Local ID	Tag	D1	S1	Name	LCNAF Heading	LC Record Link
1	MU010000000000000000000000728	700	1	6	Akashia, M. M.		
2	MU01000000000000000000000072199	700	1	5a	Abricc, Racheed		
3	MU0100000000000000000000000000000	100	1	5a	Aizer, Froux	Aizer, Froux	http://id.loc.gov/authorities/names/nb2011170001
4	MU0100000000000000000000000000000	100	1	5a	Aiz, Bish-Mish	Aiz, Bish-Mish, 1945.	http://id.loc.gov/authorities/names/nb2011107128
5	MU0100000000000000000000000000000	700	1	5a	Aitah, Edward	Aitah, Edward, 1914-2000	http://id.loc.gov/authorities/names/nb7911798
6	MU0100000000000000000000000000000	700	1	5a	Aizup, Marian		
7	MU0100000000000000000000000000000	700	1	5a	Ashkara, Makdis		
8	MU0100000000000000000000000000000	700	1	5a	Alkemon, R. J. A.		
9	MU0100000000000000000000000000000	700	1	5a	Aizer, Rachel	Aizer, Rachel	http://id.loc.gov/authorities/names/nb2011024312
10	MU0100000000000000000000000000000	700	1	5a	Bala, Roger		
11	MU0100000000000000000000000000000	100	1	6	Bai, Juncal	Bai, Juncal, 1975.	http://id.loc.gov/authorities/names/nb2011067190
12	MU0100000000000000000000000000000	700	1	5a	Balari, Joe		
13	MU0100000000000000000000000000000	700	1	5a	Barrat, Justit		
14	MU0100000000000000000000000000000	700	1	5a	Berg, William S		
15	MU0100000000000000000000000000000	700	1	5a	Bess, Ardon		
16	MU0100000000000000000000000000000	700	1	6	Blank, Petr		
17	MU0100000000000000000000000000000	700	1	5a	Blaser, Paul		
18	MU0100000000000000000000000000000	700	1	5a	Boutman, Robert	Boutman, Robert	http://id.loc.gov/authorities/names/nb2003010028
19	MU0100000000000000000000000000000	700	1	5a	Boules, Cory	Boules, Cory, 1975.	http://id.loc.gov/authorities/names/nb2010120823
20	MU0100000000000000000000000000000	100	1	5a	Brooks, Geraltine	Brooks, Geraltine, 1905.	http://id.loc.gov/authorities/names/nb04041622
21	MU0100000000000000000000000000000	700	1	5a	Budono, Handy		
22	MU0100000000000000000000000000000	100	1	5a	Bueno de Meropuz, D. M. (Daniel Meredith)	Bueno de Meropuz, D. M. (Daniel Meredith)	http://id.loc.gov/authorities/names/nb2005031680
23	MU0100000000000000000000000000000	100	1	5a	Burke, Richard James, 1917.		
24	MU0100000000000000000000000000000	100	1	5a	Cai, Tanqun	Cai, Tanqun	http://id.loc.gov/authorities/names/nb2011060156
25	MU0100000000000000000000000000000	700	1	5a	Capobianco, Carlisa		
26	MU0100000000000000000000000000000	700	1	5a	Carlone, Carlos Roberto, 1959.	Carlone, Carlos Roberto, 1959.	http://id.loc.gov/authorities/names/nb2011151077
27	MU0100000000000000000000000000000	100	1	5a	Carrigan, Joseph G. Jr. 1935.		
28	MU0100000000000000000000000000000	700	1	5a	Cases, Chloé		
29	MU0100000000000000000000000000000	700	1	5a	Cawley, Tucker	Cawley, Tucker	http://id.loc.gov/authorities/names/nb2014009752
30	MU0100000000000000000000000000000	700	1	6	Chao, Tong-ho	Chao, Tong-ho, 1945.	http://id.loc.gov/authorities/names/nb3161568
31	MU0100000000000000000000000000000	700	1	5a	Chatt, Cecely		
32	MU0100000000000000000000000000000	100	1	5a	Chapman, Dean E.		
33	MU0100000000000000000000000000000	700	1	5a	Chapman, Rubie J.	Chapman, Rubie (Rubie) John	http://id.loc.gov/authorities/names/nb2012002236
34	MU0100000000000000000000000000000	100	1	5a	Chavez Vium, Sebastian	Chavez Vium, Sebastian	http://id.loc.gov/authorities/names/nb2011150536

## Second Query for VIAF API:

"http://viaf.org/processed/LC|" + value + "?"  
httpAccept=text/xml"

# Technical Details

Google refine Example for Presentation Permalink

Facet / Filter Undo / Redo 31

362 rows

Show as: rows records Show: 5 to 25 56 rows

Filter Extract Apply

1. Create project

2. Blank down 3 cells in column Name

3. Remove 3 rows

4. Text transform on 357 cells in column Name: grel.replace(value, "", "%20")

5. Create column Fetch at index 5 by fetching URLs based on column Name using expression: `grel"http://viaf.org/viaf/search/viaf?query=local_personalNames+%3D+%22"+value+"%22+and+local+sources+%3D+%22c%22"`

6. Create new column LCCN based on column Fetch by filling 362 rows with: `grel value.parse(html).select(m2(source))`

7. Text transform on 177 cells in column LCCN: `grel.substring(value, indexOf(value, "LC"))`

8. Text transform on 177 cells in column LCCN: `grel.substring(value, 0, indexOf(value, "%"))`

9. Text transform on 177 cells in column LCCN: `value.replace("%", "/")`

10. Text transform on 177 cells in column LCCN: `grel.substring(value, 3)`

11. Text transform on 66 cells in column LCCN: `grel.replace(value, "/", "")`

12. Create column LC Heading at index 7 by fetching URLs based on column LCCN using expression: `grel"http://viaf.org/processed/LC"+value+"?httpAccept=text/xml"`

All	Local ID	Tag	D1	S1	Name	LCNAF Heading	LC Record Link
1	MU010000000000000000000000728	700	1	6	Akashia, M. M.		
2	MU01000000000000000000000072199	700	1	5a	Abricc, Racheed		
3	MU0100000000000000000000000000000	100	1	5a	Aizer, Floren	Aizer, Floren	<a href="http://id.loc.gov/authorities/names/nb2011170001">http://id.loc.gov/authorities/names/nb2011170001</a>
4	MU0100000000000000000000000000000	100	0	5a	As, Bish Moh	As, Bish Moh, 1945.	<a href="http://id.loc.gov/authorities/names/nb2011107128">http://id.loc.gov/authorities/names/nb2011107128</a>
5	MU0100000000000000000000000000000	700	1	5a	Anhalt, Edward	Anhalt, Edward, 1914-2000	<a href="http://id.loc.gov/authorities/names/nb7911798">http://id.loc.gov/authorities/names/nb7911798</a>
6	MU0100000000000000000000000000000	700	1	5a	Arzop, Marian		
7	MU0100000000000000000000000000000	700	1	5a	Ashkara, Makdis		
8	MU0100000000000000000000000000000	700	1	5a	Atkinson, R. J. A.		
9	MU0100000000000000000000000000000	700	1	5a	Avier, Rachel	Avier, Rachel	<a href="http://id.loc.gov/authorities/names/nb2011024312">http://id.loc.gov/authorities/names/nb2011024312</a>
10	MU0100000000000000000000000000000	700	1	5a	Bala, Roger		
11	MU0100000000000000000000000000000	100	1	6	Bai, Junsai	Bai, Junsai, 1975.	<a href="http://id.loc.gov/authorities/names/nb2011067190">http://id.loc.gov/authorities/names/nb2011067190</a>
12	MU0100000000000000000000000000000	700	1	5a	Bakari, Joe		
13	MU0100000000000000000000000000000	700	1	5a	Barrel, Justit		
14	MU0100000000000000000000000000000	700	1	5a	Berg, William S		
15	MU0100000000000000000000000000000	700	1	5a	Bess, Ardon		
16	MU0100000000000000000000000000000	700	1	6	Blank, Petr		
17	MU0100000000000000000000000000000	700	1	5a	Blaser, Paul		
18	MU0100000000000000000000000000000	700	1	5a	Boutman, Robert	Boutman, Robert	<a href="http://id.loc.gov/authorities/names/nb2003010028">http://id.loc.gov/authorities/names/nb2003010028</a>
19	MU0100000000000000000000000000000	700	1	5a	Bowles, Cory	Bowles, Cory, 1975.	<a href="http://id.loc.gov/authorities/names/nb2011020223">http://id.loc.gov/authorities/names/nb2011020223</a>
20	MU0100000000000000000000000000000	100	1	5a	Brooks, Geraltine	Brooks, Geraltine, 1905.	<a href="http://id.loc.gov/authorities/names/nb04041622">http://id.loc.gov/authorities/names/nb04041622</a>
21	MU0100000000000000000000000000000	700	1	5a	Budonis, Hardy		
22	MU0100000000000000000000000000000	100	1	5a	Bueno de Mesquita, D. M. (Daniel Meredith)	Bueno de Mesquita, D. M. (Daniel Meredith)	<a href="http://id.loc.gov/authorities/names/nb2005031680">http://id.loc.gov/authorities/names/nb2005031680</a>
23	MU0100000000000000000000000000000	100	1	5a	Burke, Richard James, 1917.		
24	MU0100000000000000000000000000000	100	1	5a	Cai, Tanqun	Cai, Tanqun	<a href="http://id.loc.gov/authorities/names/nb2011060156">http://id.loc.gov/authorities/names/nb2011060156</a>
25	MU0100000000000000000000000000000	700	1	5a	Capobianco, Carlisa		
26	MU0100000000000000000000000000000	700	1	5a	Carlson, Carlos Roberto, 1959.	Carlson, Carlos Roberto, 1959.	<a href="http://id.loc.gov/authorities/names/nb2011151077">http://id.loc.gov/authorities/names/nb2011151077</a>
27	MU0100000000000000000000000000000	100	1	5a	Carrigan, Joseph G. Jr. 1935.		
28	MU0100000000000000000000000000000	700	1	5a	Cases, Chlo		
29	MU0100000000000000000000000000000	700	1	5a	Cawley, Tucker		<a href="http://id.loc.gov/authorities/names/nb2014009752">http://id.loc.gov/authorities/names/nb2014009752</a>
30	MU0100000000000000000000000000000	700	1	6	Chae, Tong-ho	Chae, Tong-ho, 1945.	<a href="http://id.loc.gov/authorities/names/nb3161568">http://id.loc.gov/authorities/names/nb3161568</a>
31	MU0100000000000000000000000000000	700	1	5a	Cham, Cecely		
32	MU0100000000000000000000000000000	100	1	5a	Chapman, Dean E.		
33	MU0100000000000000000000000000000	700	1	5a	Chapman, Rubie J.	Chapman, Rubie (Rubie John)	<a href="http://id.loc.gov/authorities/names/nb2012002236">http://id.loc.gov/authorities/names/nb2012002236</a>
34	MU0100000000000000000000000000000	100	1	5a	Chavez Vium, Sebastian	Chavez Vium, Sebastian	<a href="http://id.loc.gov/authorities/names/nb2011150536">http://id.loc.gov/authorities/names/nb2011150536</a>

## Second Query for VIAF API:

"http://viaf.org/processed/LC|" + **value** + "?"  
httpAccept=text/xml"



[illegible]

```
"http://viaf.org/processed/LC|" + value + "?"  
httpAccept=text/xml"
```



## Technical Details

[illegible]

## Technical Details

[illegible]

# Technical Details

Google refine Presentation FINAL Permalink Open... Export Help

Facet / Filter Undo / Redo 20 **18 rows** Extensions: Named-entity recognition Freebase RDF

Show as: rows records Show: 5 10 25 50 rows « first < previous 1 - 18 next > last »

Filter:

12. Text transform on 3 cells in column LC Record Link: `grel:replace(value, ',', '')`

13. Create column LCNAF Heading at index 3 by fetching URLs based on column LC Record Link using expression `grel:"http://viaf.org/processed/LC|" + value + "?httpAccept=text/xml"`

14. Text transform on 8 cells in column LCNAF Heading: `grel:value.parseHtml().select(mx[datafield[0]].htmlText())`

15. Text transform on 0 cells in column LCNAF Heading: `value.trim()`

16. Text transform on 0 cells in column LCNAF Heading: `value.replace(/s/, '')`

17. Text transform on 8 cells in column LC Record Link: `grel:"http://id.loc.gov/authorities/names/" + value`

18. Move column LC Record Link to position 3

19. Remove column Fetch

20. Text transform on 18 cells in column Name: `grel:replace(value, '%20', '')`

All	Name	LCNAF Heading	LC Record Link
1.	Alver, Füsün.	Alver, Füsün	<a href="http://id.loc.gov/authorities/names/no2011170601">http://id.loc.gov/authorities/names/no2011170601</a>
2.	An, Binh Minh.	An Binh Minh, 1948-	<a href="http://id.loc.gov/authorities/names/no2011167328">http://id.loc.gov/authorities/names/no2011167328</a>
3.	Anhalt, Edward.	Anhalt, Edward, 1914-2000	<a href="http://id.loc.gov/authorities/names/n87911796">http://id.loc.gov/authorities/names/n87911796</a>
4.	Araujo, Marian.		
5.	Ashikawa, Makoto.		
6.	Atkinson, R. J. A.		
7.	Axler, Rachel.	Axler, Rachel	<a href="http://id.loc.gov/authorities/names/nb2011024512">http://id.loc.gov/authorities/names/nb2011024512</a>
8.	Babia, Roger.		
9.	Bai, Juncal.	Bai, Juncal, 1970-	<a href="http://id.loc.gov/authorities/names/no2009146750">http://id.loc.gov/authorities/names/no2009146750</a>
10.	Ballarini, Joe.		
11.	Barret, Justin.		
12.	Berg, William S.		
13.	Bess, Ardon.		
14.	Blankin, Petr.		
15.	Blaser, Paul.		
16.	Boatman, Robert.	Boatman, Robert H.	<a href="http://id.loc.gov/authorities/names/n2003010628">http://id.loc.gov/authorities/names/n2003010628</a>
17.	Bowles, Cory.	Bowles, Cory, 1973-	<a href="http://id.loc.gov/authorities/names/no2010126923">http://id.loc.gov/authorities/names/no2010126923</a>
18.	Brooks, Geraldine, 1955-	Brooks, Geraldine	<a href="http://id.loc.gov/authorities/names/n94041622">http://id.loc.gov/authorities/names/n94041622</a>



# Why Use VIAF?

1. More robust dataset to search than other options
2. VIAF algorithm is better at fuzzy name searches
3. Updated weekly

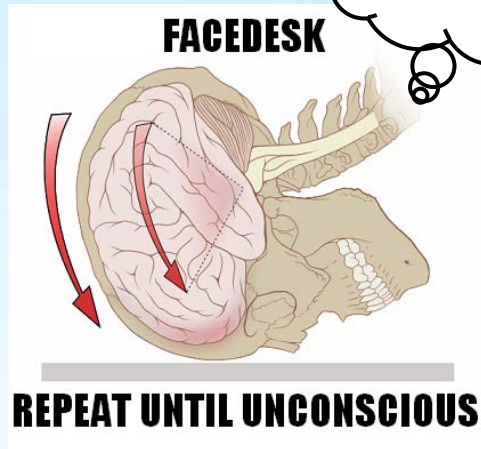
# Limitations

1. There is a time lag if you are searching hundreds or thousands of names at a time.
2. **IMPORTANT**: It is up to the user to verify the accuracy of the matches found.
3. The process does not currently reconcile geographic name headings.



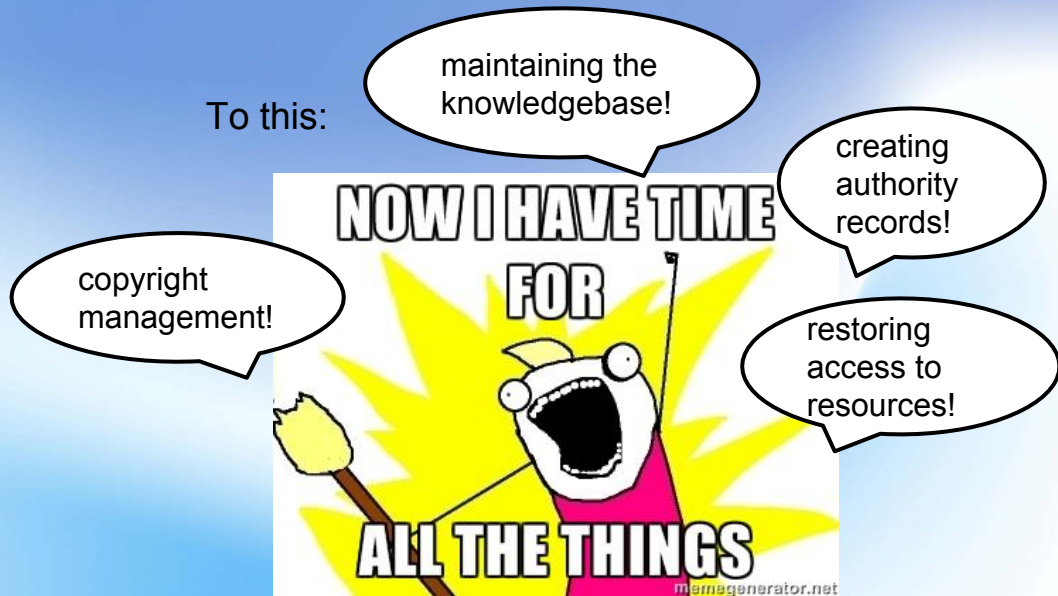
# Impact

From this:



Why must  
this take  
so long?

To this:



# Impact Details

- With this process, authority control now takes half the time it used to
  - Less time spent looking in vain for non-existent authority files, because the script tells us whether there is a file or not
- More authority control gets done
- More time for other tasks
- Less time wasted = good for morale!
- We contribute to a pool of headings that we can draw from for authority record creation workflows

# Conclusion

- This is the sort of thing we always hoped to be possible
  - Has been tremendously helpful to workflows
- We hope it helps others as much as it has us!

Get the files and instructions!

<https://github.com/mcarruthers/LCNAF-Named-Entity-Reconciliation>

Matt Carruthers  
mcarruth@umich.edu

Jenn Wright  
jennywri@umich.edu