Broadcasting A/V Data to the SNAC universe (and beyond!)

September 19, 2023

Emily Frazier (she/her)
formerly: Maryland Institute for Technology in the Humanities
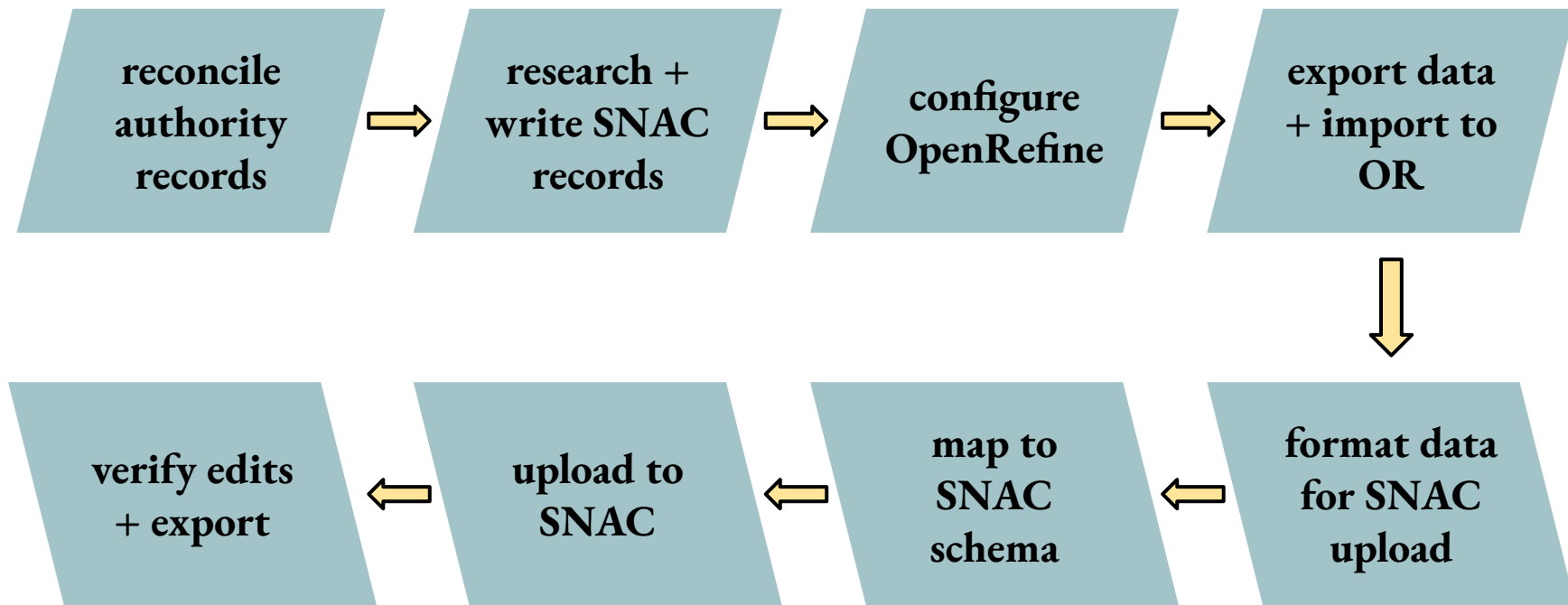currently: University of Maryland Libraries

project background

❖ follow-up project to *Unlocking the Airwaves* (virtual reunification of NAEB collection)

❖ centered on archival authorities--aka CPF records, people and organizations, or entities

❖ goal was to enhance discoverability of authorities and the fully-digitized radio programs in which they appear or helped create

❖ 2 sides:

➢ project website, with links to other authority records and streaming audio as well as data visualization

➢ linked data on the web--SNAC, Wikidata, and Wikipedia

two SNAC processes

1. getting data OUT of SNAC for use on BAVD project website and to know what data to submit back to SNAC

2. uploading data TO SNAC using the OpenRefine + SNAC extension to surface authority records of people and organizations in early educational radio to the web
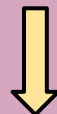
reconcile authority records → research + write SNAC records → configure OpenRefine → export data + import to OR

↓

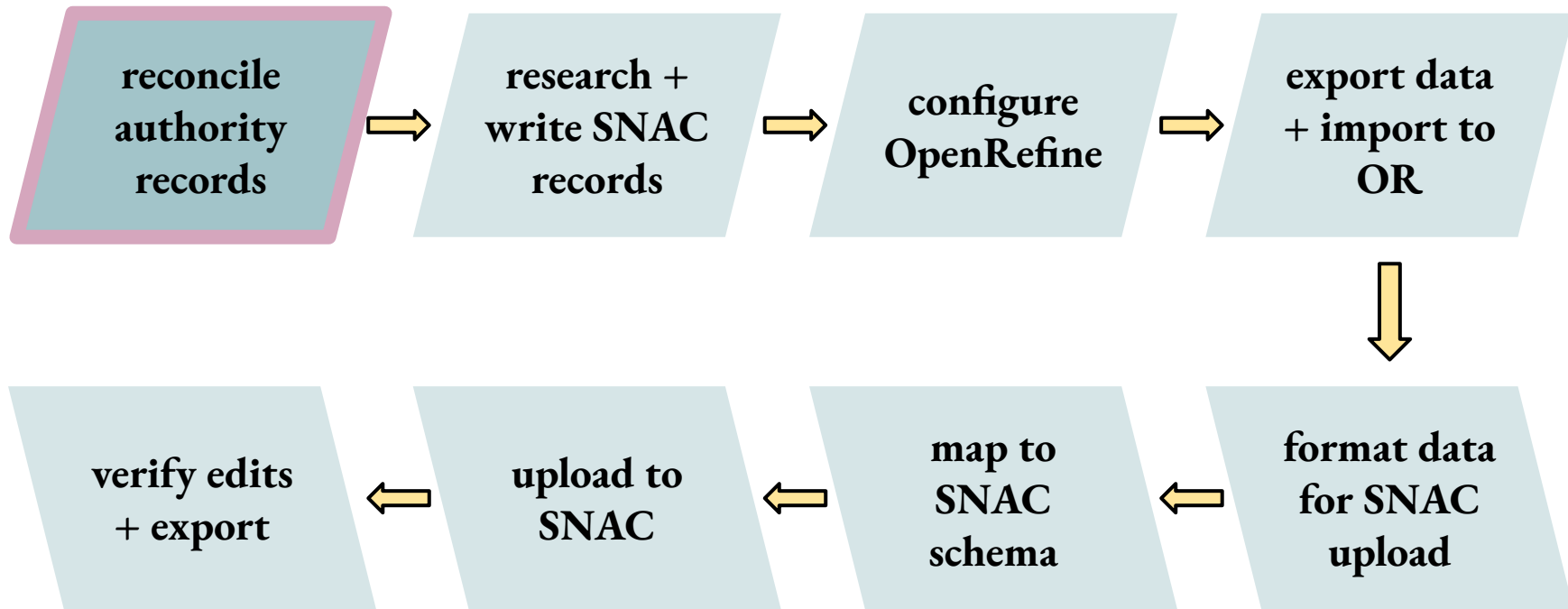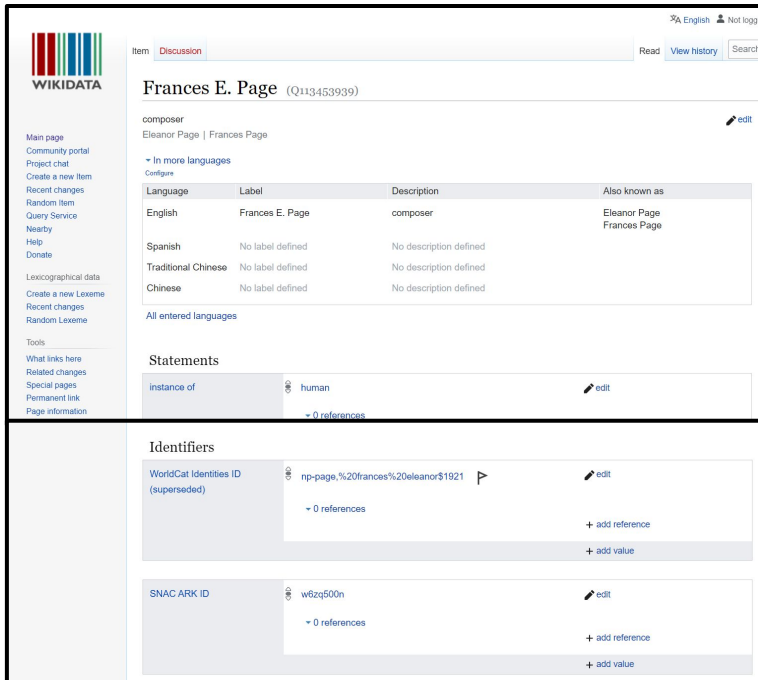verify edits + export ← upload to SNAC ← map to SNAC schema ← format data for SNAC upload

batch upload workflow: overview

repeats 3x: CPFs, resources, relationships

reconcile authority records → research + write SNAC records → configure OpenRefine → export data + import to OR

↓

verify edits + export ← upload to SNAC ← map to SNAC schema ← format data for SNAC upload

```
reconcile authority records  →  research + write SNAC records  →  configure OpenRefine  →  export data + import to OR
                                                                                                      ↓
verify edits + export  ←  upload to SNAC  ←  map to SNAC schema  ←  format data for SNAC upload
```

batch upload workflow: reconciling

Frances Eleanor Page
composer for radio station KUT

method 2: SNAC reconciliation in OpenRefine

Item Discussion

Read View history Searc
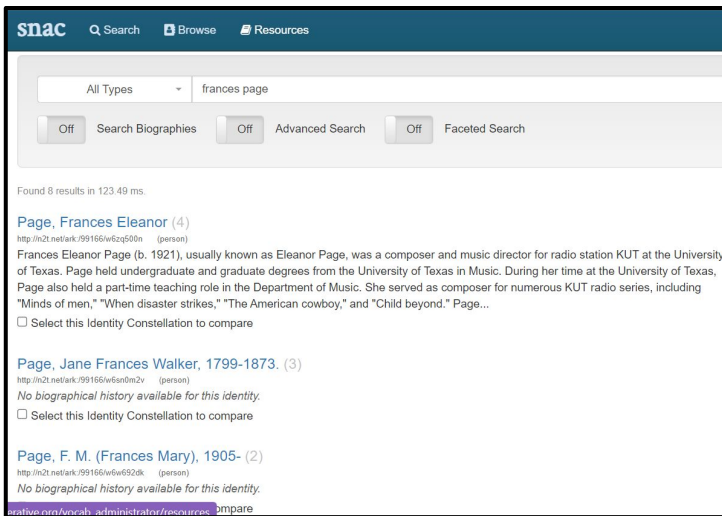
Frances E. Page (Q113453939)

composer
Eleanor Page | Frances Page

🖊 edit

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
What links here
Related changes
Special pages
Permanent link
Page information

▾ In more languages
Configure

Language	Label	Description	Also known as
English	Frances E. Page	composer	Eleanor Page Frances Page
Spanish	No label defined	No description defined	
Traditional Chinese	No label defined	No description defined	
Chinese	No label defined	No description defined	

All entered languages

Statements

instance of   👤 human   🖊 edit
                         ▾ 0 references

Identifiers

WorldCat Identities ID   👤 np-page,%20frances%20eleanor$1921 ⚑   🖊 edit
(superseded)
                                ▾ 0 references
                                                              + add reference
                                                              + add value

SNAC ARK ID   👤 w6zq500n   🖊 edit
                                ▾ 0 references
                                                              + add reference
                                                              + add value

inked CPF Authority"

Add standard service

Enter the service's URL
https://openrefine.snaccooperative.org/

Add service   Cancel

method 1: pull SNAC ID from Wikidata

snac   🔍 Search   📖 Browse   📄 Resources

All Types ▾   frances page

Off   Search Biographies   Off   Advanced Search   Off   Faceted Search

Found 8 results in 123.49 ms.

Page, Frances Eleanor (4)
http://n2t.net/ark/99166/w6zq500n   (person)
Frances Eleanor Page (b. 1921), usually known as Eleanor Page, was a composer and music director for radio station KUT at the University of Texas. Page held undergraduate and graduate degrees from the University of Texas in Music. During her time at the University of Texas, Page also held a part-time teaching role in the Department of Music. She served as composer for numerous KUT radio series, including "Minds of men," "When disaster strikes," "The American cowboy," and "Child beyond." Page...
☐ Select this Identity Constellation to compare

Page, Jane Frances Walker, 1799-1873. (3)
http://n2t.net/ark/99166/w6sn0m2v   (person)
No biographical history available for this identity.
☐ Select this Identity Constellation to compare

Page, F. M. (Frances Mary), 1905- (2)
http://n2t.net/ark/99166/w6w652dk   (person)
No biographical history available for this identity.

erative.org/vocab administrator/resources ompare

method 3: searching in SNAC
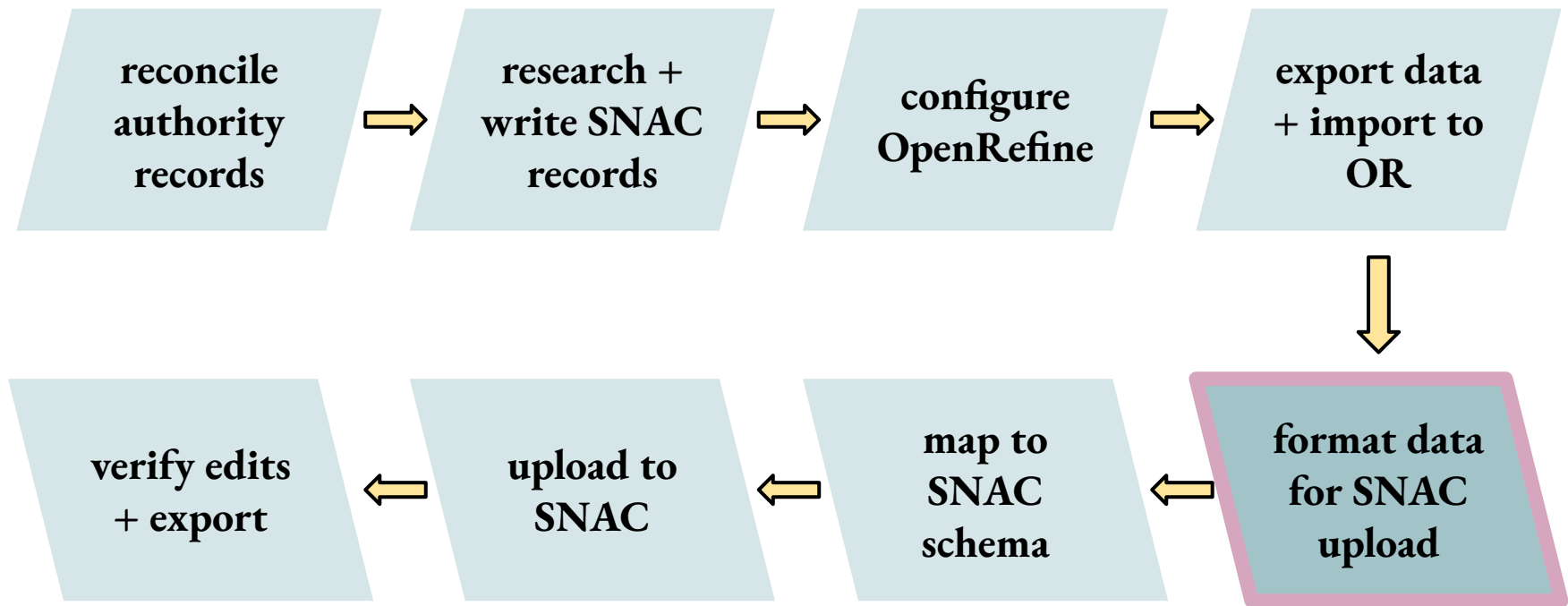
reconcile authority records → research + write SNAC records → configure OpenRefine → export data + import to OR ↓

format data for SNAC upload ← map to SNAC schema ← upload to SNAC ← verify edits + export

batch upload workflow: OR data formatting

an example starting SNAC CPF dataset in OpenRefine, before data formatting

batch upload workflow: OR data formatting

	A	D	E	F	G	Pla...
1	**SNAC OR Schema Field Name**	Name Entry	Exist Date	Exist Date Type	Subject	Pla...
2	**Airwaves Field Name(s)**	Linked CPF Authority	Date of Birth; Date of Death; Active Date 1; Active Date 2	Date of Birth; Date of Death; Active Date 1; Active Date 2	Subject(s)	Pla...
3	**Airwaves Table(s)**	SNAC Records	SNAC Records	SNAC Records	SNAC Records	SN...
4	**BAVD Field Name(s) SAMPLE DATA**	bavdCPF	birthDate; deathDate	birthDate; deathDate	subjects	pla...
5	**BAVD Table(s)**	SNAC Records	SNAC Records	SNAC Records	SNAC Records	SN...
6	**Related Fields**		Date Type	Date		Pla...
7	**For Use With**	All	All	All	All	All
8	**OR Transformations Required**	To strip quotes, use this GREL expression: value.smartSplit(",").join("#")		Transpose cells across columns into rows, 2 columns (key: Date Type; value: Date); Then replace field names with "Birth", "Death", or "Active"	Use this GREL expression: value.smartSplit(",").join("#"), then split multi-valued cells on separator #	Use valu... mul...
9	**Ordering of actions**					
10	**Manual Tranformations Required**					
11	**Example 1**	WCAU-TV (Television Station : Philadelphia, Penn.)	1932-03-15	Active	Radio stations	Ala...
12	**Example 2**	Press, O. Leonard	1956	Birth		Ala...
13	**Example 3**	Duquesne University	1968-09	Death		Wa...
14	**Notes**	Formatting in Airwaves is aligned with SNAC preferred formatting--if entity has LCNAF entry, that formatting can be used and seems to be SNAC aligned as well	All date formats will upload exactly as submitted, no need for reformatting		SNAC now uses FAST-aligned headings, so there should be no subordination/pre-coordination of subject terms. E.g., no "--" between terms	Do in A for...
15	*Release v1.0.0-beta-18*					

data formatting and schema alignment quick reference guide

batch upload workflow: exporting IDs

reconcile authority records → research + write SNAC records → configure OpenRefine → export data + import to OR

↓

verify edits + export ← upload to SNAC ← map to SNAC schema ← format data for SNAC upload

batch upload workflow: exporting IDs

an example starting SNAC relationships dataset in OpenRefine, before data formatting

batch upload workflow: sticking points

❖ accidentally duplicating records and their contents

❖ figuring out unexpected behavior in OpenRefine

❖ development vs. production SNAC errors - FIXED!

❖ OpenRefine terminology - records mode vs. rows mode

batch upload workflow: reading the docs

SNAC data harvesting script, from workflow

[data formatting and schema alignment guide](#)

[creating SNAC constellations workflow](#)

resources

- ❖ Broadcasting A/V Data website: https://bavd.unlockingtheairwaves.org/
- ❖ SNAC OpenRefine extension: https://github.com/snac-cooperative/snac-openrefine-extension
- ❖ OpenRefine website: https://openrefine.org/
- ❖ BAVD documentation hub - see docs 9, 12, and 13:
  https://github.com/umd-mith/broadcasting/tree/main/workflows-docs
- ❖ SNAC OpenRefine extension documentation (by SNAC team):
  https://docs.google.com/document/d/1HjhKHAWXe6Xkh8F-qBnsl-psyaPd2UPBLX2OgaVbA2Q/edit?usp=sharing
- ❖ SNAC technical resources hub: https://portal.snaccooperative.org/node/489
- ❖ Library Carpentry OpenRefine lesson: https://librarycarpentry.org/lc-open-refine/
- ❖ Programming Historian OpenRefine lesson:
  https://programminghistorian.org/en/lessons/cleaning-data-with-openrefine
- ❖ SNAC-dev: https://snac-dev.iath.virginia.edu/
- ❖ Sample data:
  - ➤ CPFs: https://airtable.com/appMq6PvC5azYIsdt/shr1QKR1k0eGgPI2b/tblQI1mC0vnrEwv2T
  - ➤ resources: https://airtable.com/appMq6PvC5azYIsdt/shrhRgNNWR1UewTEJ/tblpppSJiZmbCDWZh
  - ➤ join table: https://airtable.com/appMq6PvC5azYIsdt/shrI8TproJdnjHRIh/tblQI1mC0vnrEwv2T

thank you! + contact info

Emily Frazier

Collections Data Assessment Coordinator, University of Maryland Libraries

emilyf1(at)umd.edu

Stephanie Sapienza

Project Director, Broadcasting A/V Data

Digital Humanities Archivist, Maryland Institute for Technology in the Humanities

sapienza(at)umd.edu