Wikidata for Scholarly Profiles

Love Data Week 2023

Mairelys Lemus-Rojas
Head of Open Metadata Production & Initiatives
Brown University

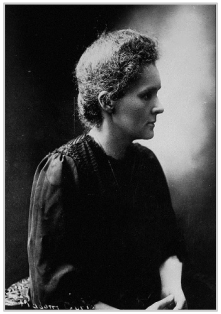
What is Wikidata?

- Multilingual knowledge base
- Stores structured linked-data
- Read by humans and machines
- Openly accessible
- Launched in 2012
- Hub for Wikimedia projects
- Collaboratively edited
- Contains over 101 million items
    - Growing collection of open citation data (over 41 million publications as of December 2022)

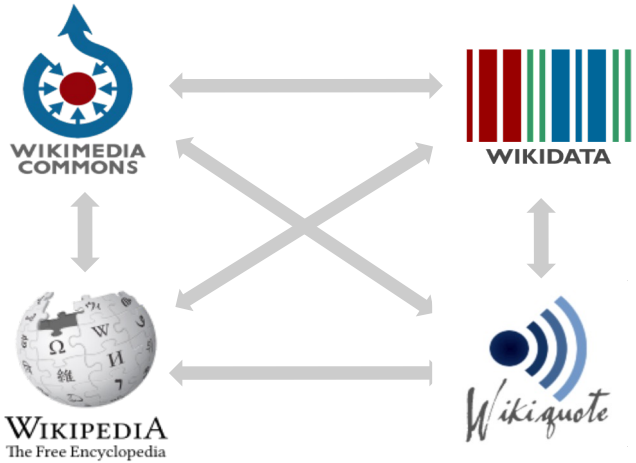
Link: http://wikicite.org/statistics.html

WIKIMEDIA

Wiki Projects Interlinking



Marie Salomea Skłodowska–Curie (/ˌkjʊəriː/ *KURE-ee*,[4] French pronunciation: [maʁi kyʁi], Polish pronunciation: [ˈmarja skwɔˈdɔfska kiˈri]; born **Maria Salomea Skłodowska**, Polish: [ˈmarja salɔˈmɛa skwɔˈdɔfska]; 7 November 1867 – 4 July 1934) was a Polish and naturalized-French physicist and chemist who conducted pioneering research on radioactivity. She was the first woman to win a Nobel Prize, the first person and the only woman to win a Nobel Prize twice, and the only person to win a Nobel Prize in two scientific fields. Her husband, Pierre Curie, was a co-winner of her first Nobel Prize, making them the first-ever married couple to win the Nobel Prize and launching the Curie family legacy of five Nobel Prizes. She was, in 1906, the first woman to become a professor at the University of Paris.[5]



country of citizenship	France •••
	start time 1895 •••
	series ordinal 2 •••
	▸ 2 references
	Poland •••
	start time 1918 •••
	series ordinal 3 •••
	▸ 4 references

Quotes [edit]

- **One never notices what has been done; one can only see what remains to be done.**
  - Letter to her brother (1894)
- I have no dress except the one I wear every day. If you are going to be kind enough to give me one, please let it be practical and dark so that I can put it on afterwards to go to the laboratory.

Wikidata for Scholarly Profiles

- Ability to contribute scholarly data to an open platform (expanding the bibliographic dataset)
  - Aligns with other projects/initiatives on campus
- Opportunities for collaboration
- Inclusion
  - Works that are not necessarily included in either Crossref or PubMed
- Increase visibility and discovery of faculty/researchers and their scholarship

Wikidata Notability

Meet at least 1 of the 3 criteria

- Have a presence in other Wikimedia project (and be linked to it)
- Be a clearly identifiable concept/object that can be supported with reliable resources
- Fulfill a structural need (validate existing statements)

Wikidata Basics

Entities refer to either **Properties**, **Items**, or **Lexemes**

Properties

- Describe data values (link **Items** to values)
- Need to be proposed and approved by the Wikidata community
- Unique identifier: **P** prefix followed by a number
  - Example: P123 is the identifier for **publisher**

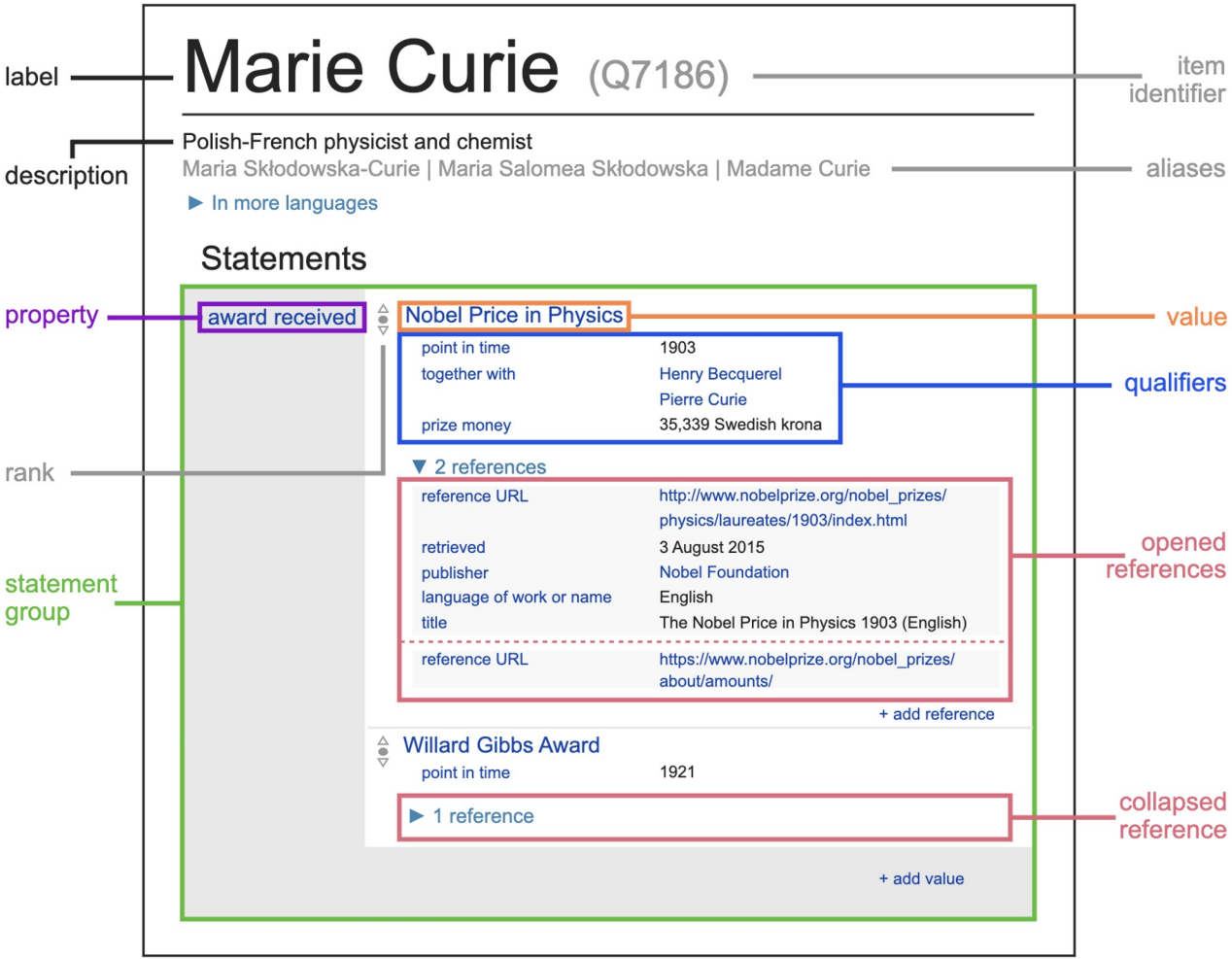
Items

- Represent human knowledge (concepts, objects, topics)
- Can be created by anyone
- Unique identifier: **Q** prefix followed by a number
  - Example: Q13442814 is the identifier for **scholarly article**

Wikidata Basics

Data is stored in triples

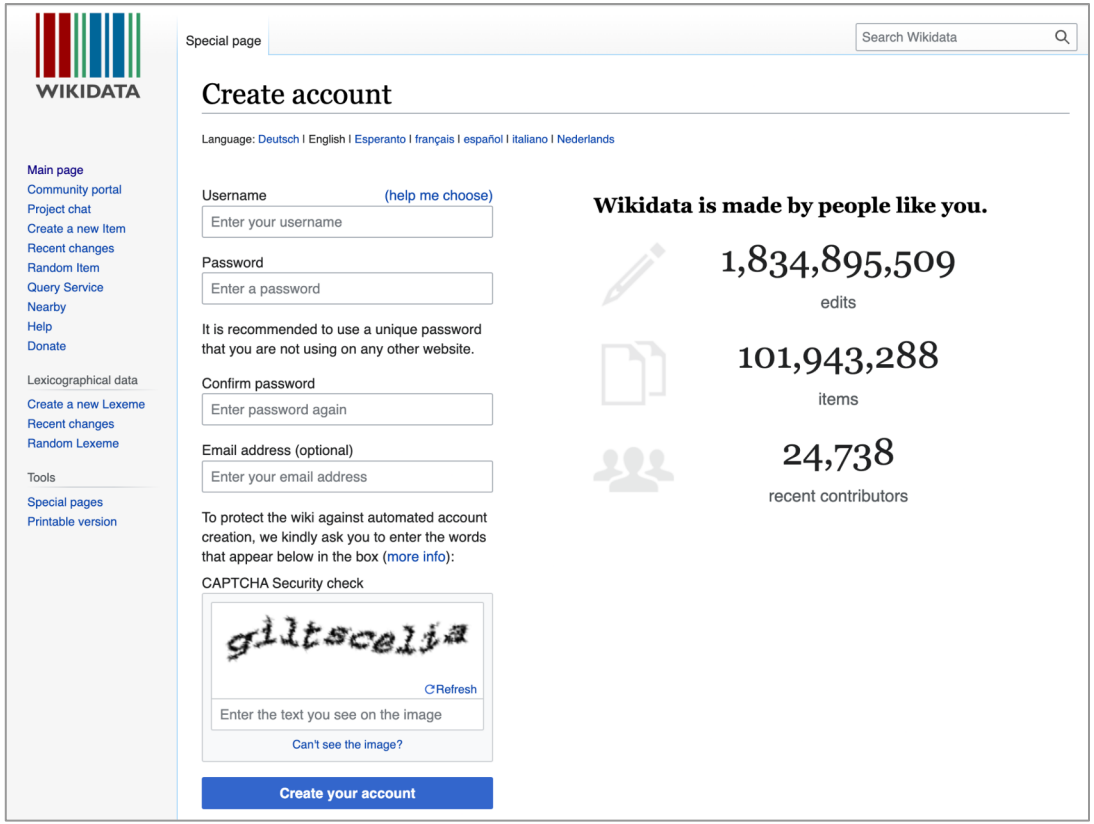
Item (Subject)	Property (Predicate)	Value (Object)
Marilyn Yalom	field of work	gender studies
Q6763649	P101	Q1662673



Wikidata Data Model

Account Creation

Link: https://w.wiki/3BeL

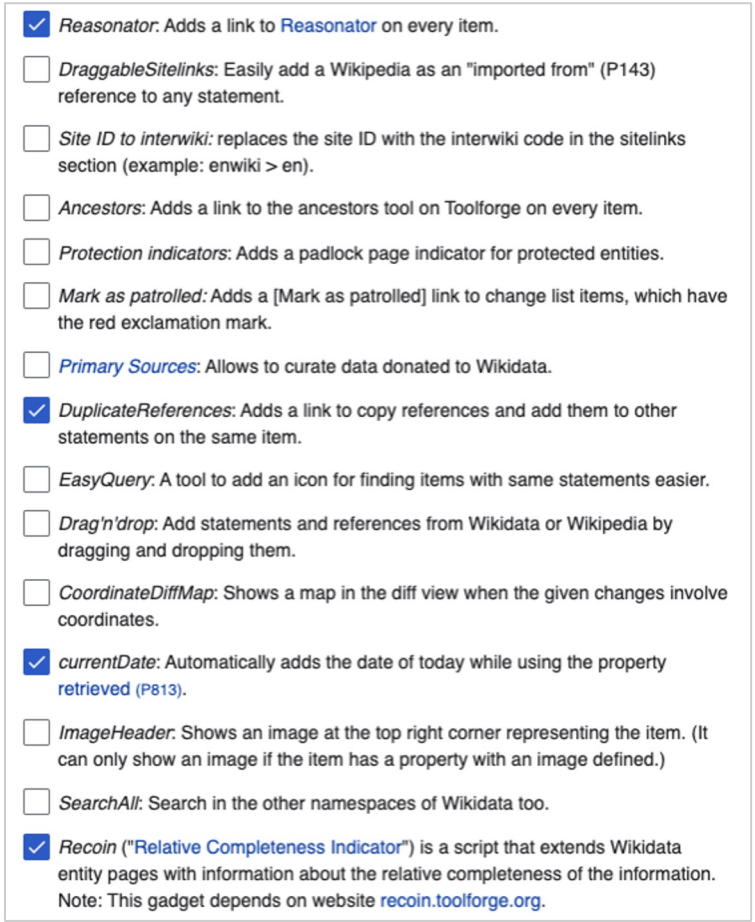


Create User Page

- Page to share information about (profile page)
  - Interest in Wikimedia projects
- Provide a channel for other users to communicate with you
- Add conflict of interest statement

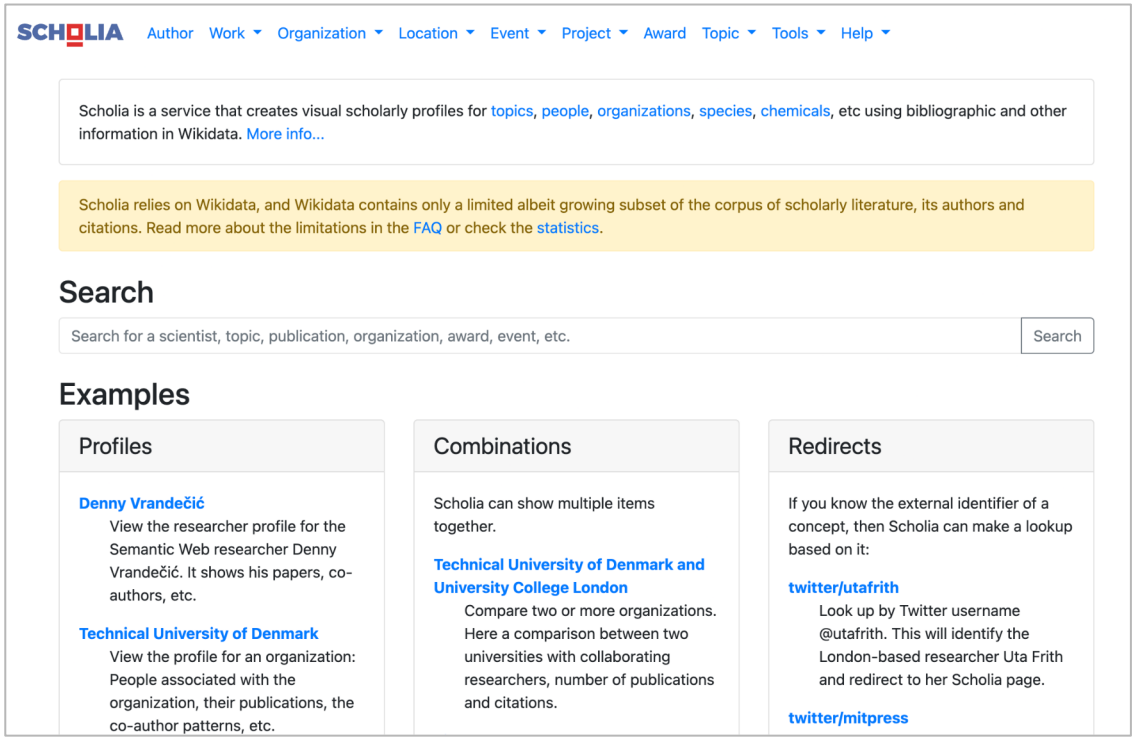
Gadgets

- Reasonator
- DuplicateReferences
- currentDate
- Recoin



Tool to Visualize Data

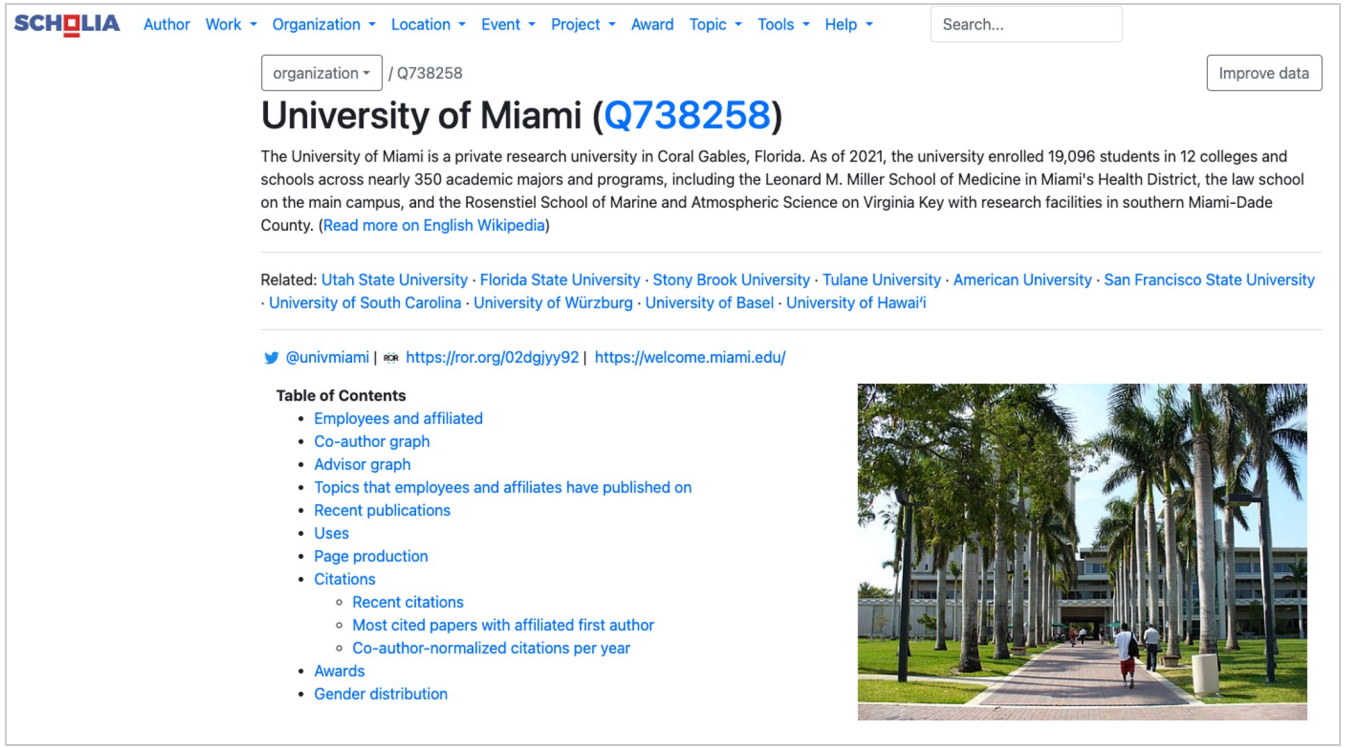
Scholia Homepage



- Web service for scholarly profiles using Wikidata

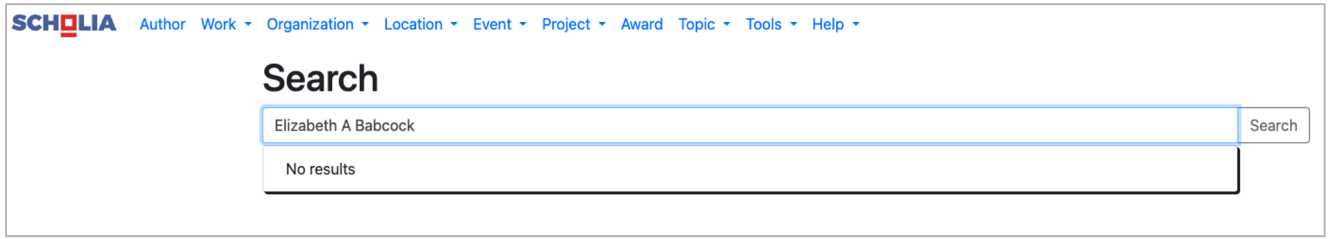
Link: https://scholia.toolforge.org/

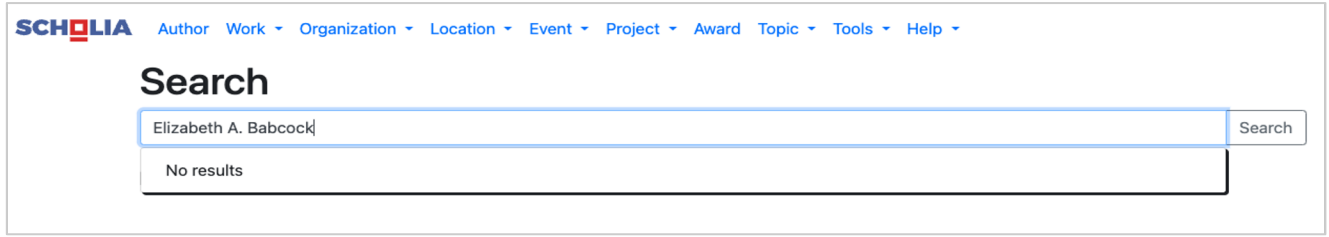
Scholia Profile

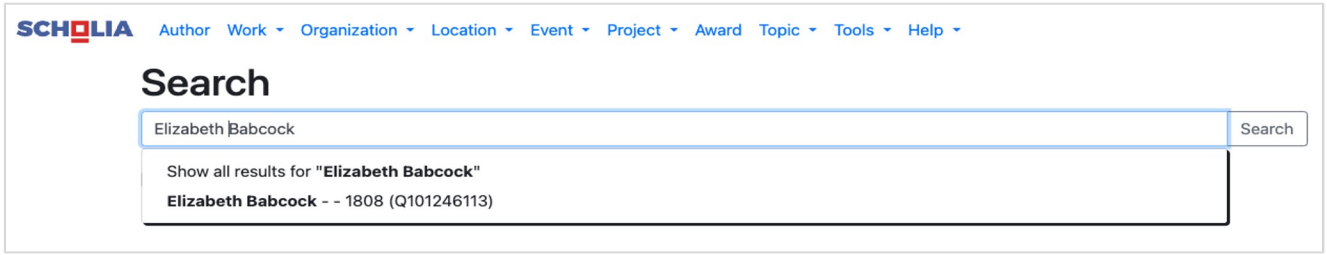


- Profile for the University of Miami

Scholia Search Results







- Search results for Elizabeth A. Babcock

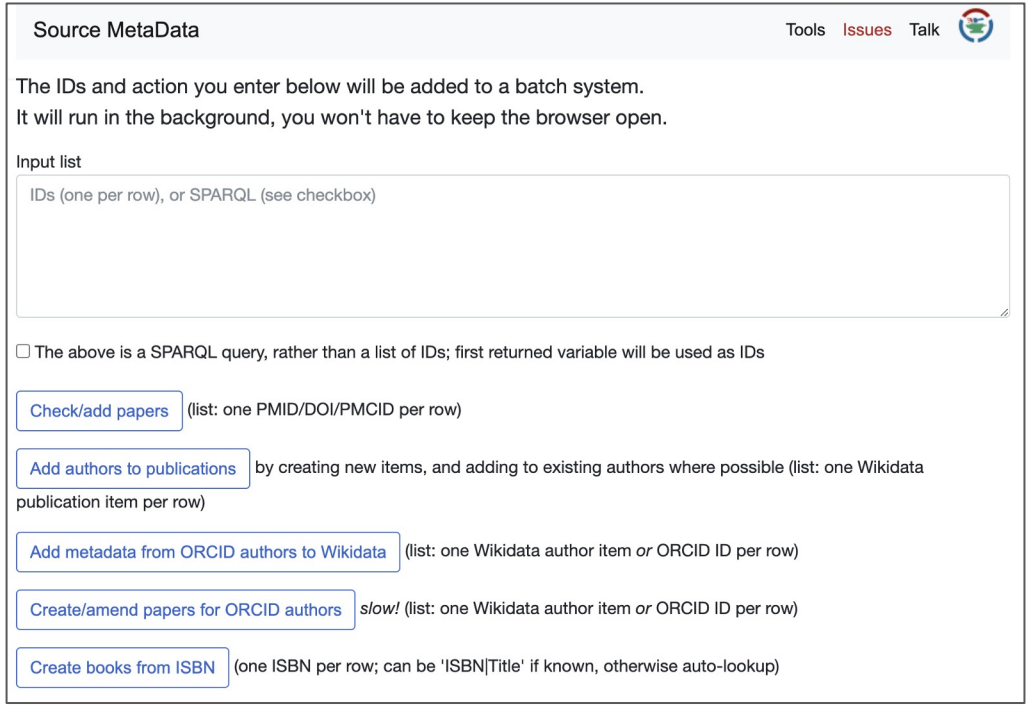
Hands-on Activity

- Add articles written by Elizabeth A. Babcock
- Use the spreadsheet to select an article to work on
- Use table on the right as reference for describing articles

Property	Value
instance of (P31)	scholarly article; working paper; preprint
title (P1476)	Title of the work
subtitle (P1680)	Subtitle of the work
author name string (P2093)	Names not established
author (P50)	Established names
language of work or name (P407)	Language of work
publication date (P577)	Date of publication
number of pages (P1104)	Number of pages
published in (P1433)	Venue where work was published in
doi (P356)	Identifier

Tools to Contribute Data

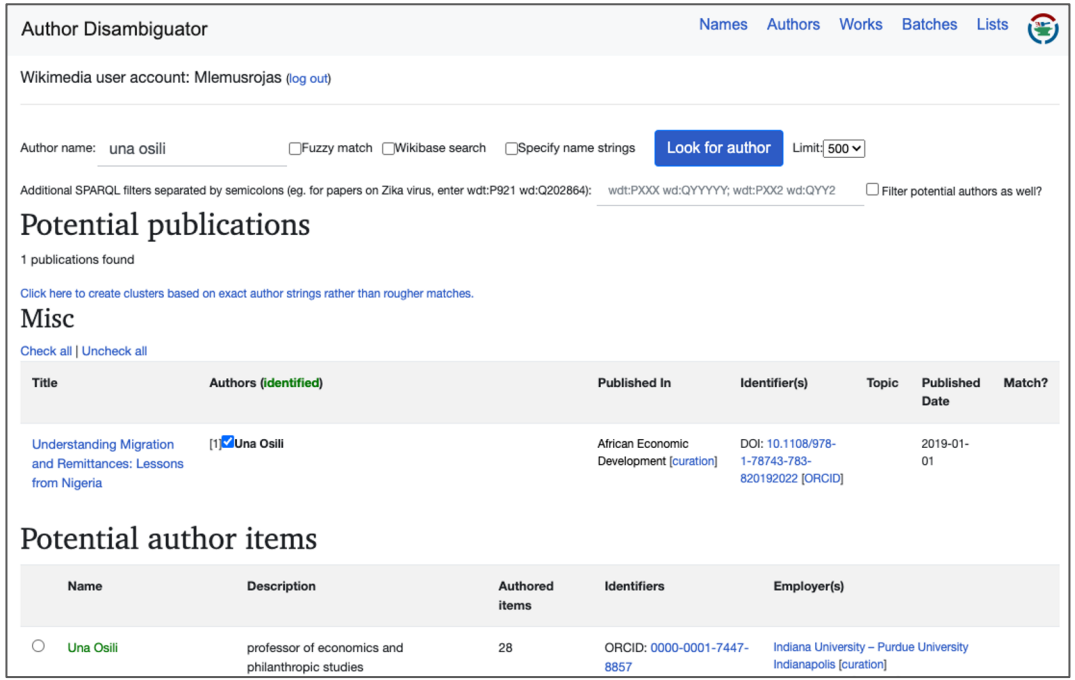
Source MetaData



Link: https://sourcemd.toolforge.org/

- Pulls bibliographic metadata based on identifiers (PMID, DOI/PMCID)

Author Disambiguator Tool

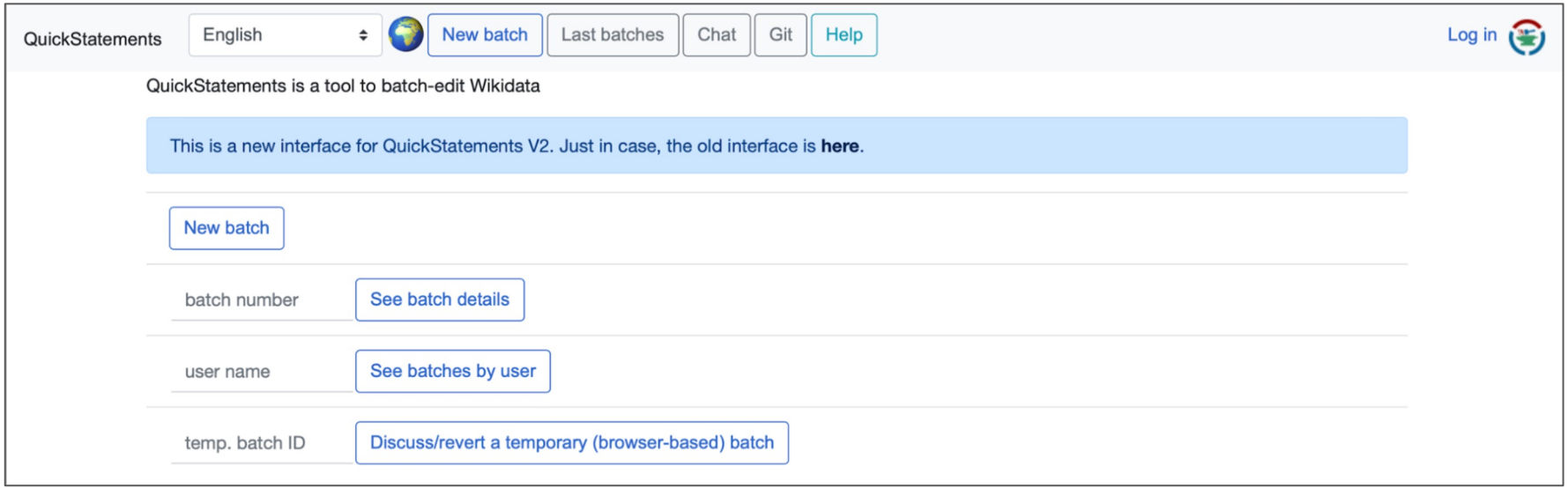


Link: https://author-disambiguator.toolforge.org/

- Provides list of possible matches
- Prepares data for batch editing (QuickStatements)

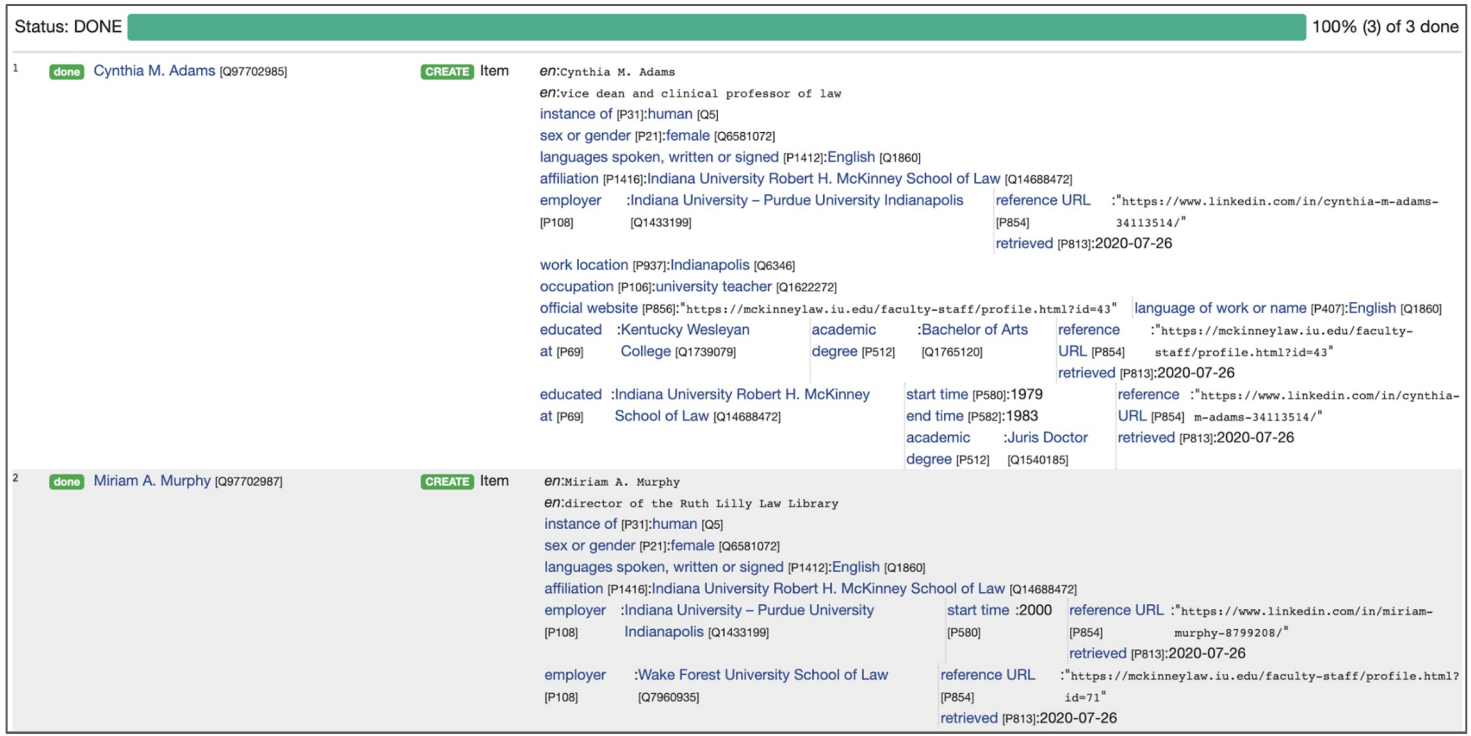
QuickStatements

- Executes automatic edits in Wikidata



Link: https://quickstatements.toolforge.org/#/

QuickStatements Example



Thank You!

Q19279214