



# Wikidata for Scholarly Profiles

Love Data Week 2023

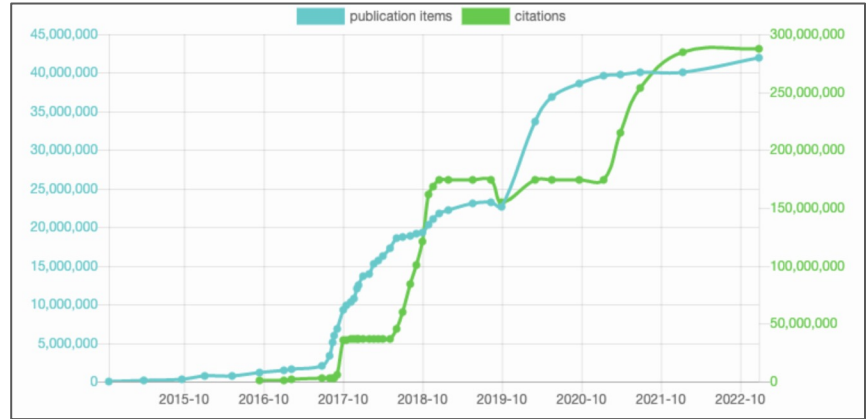
Mairelys Lemus-Rojas

Head of Open Metadata Production & Initiatives

Brown University

# What is Wikidata?

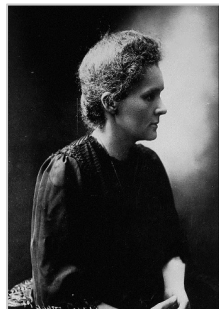
- Multilingual knowledge base
- Stores structured linked-data
- Read by humans and machines
- Openly accessible
- Launched in 2012
- Hub for Wikimedia projects
- Collaboratively edited
- Contains over 101 million items
  - Growing collection of open citation data (over 41 million publications as of December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



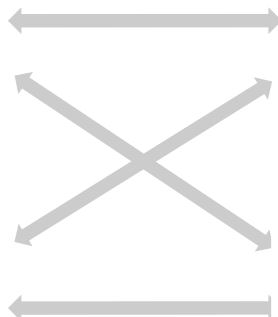
Link: <http://wikicite.org/statistics.html>



# Wiki Projects Interlinking



**Marie Salomea Skłodowska-Curie** (/ˈkjuːri/ *KURE-ee*,<sup>[4]</sup> French pronunciation: [maʁi kɥʁi], Polish pronunciation: [ˈmarja skwɔˈdɔfska kʲiˈɲi]; 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**.<sup>[5]</sup>



country of citizenship	<span>country of citizenship</span> <span><span></span></span> <span>France</span> <span><span></span></span> <span>...</span>
<span><span></span></span>	<span>start time</span> <span><span></span></span> <span>1895</span> <span><span></span></span> <span>...</span>
	<span>series ordinal</span> <span><span></span></span> <span>2</span> <span><span></span></span> <span>...</span>
	<span><span></span></span> <span>2 references</span>
	<span><span></span></span> <span>Poland</span> <span><span></span></span> <span>...</span>
	<span>start time</span> <span><span></span></span> <span>1918</span> <span><span></span></span> <span>...</span>
	<span>series ordinal</span> <span><span></span></span> <span>3</span> <span><span></span></span> <span>...</span>
	<span><span></span></span> <span>4 references</span>

## 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**

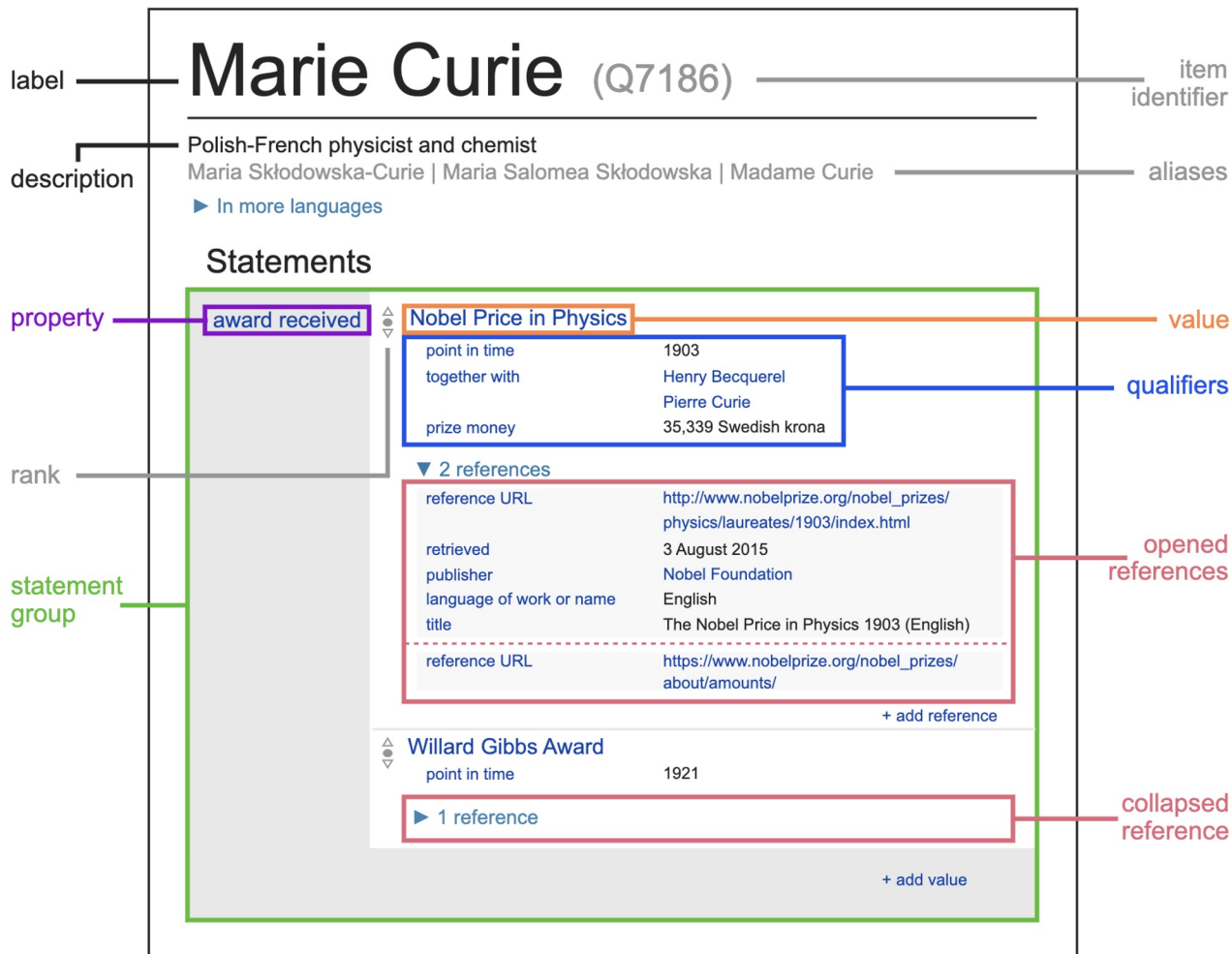
# Wikidata Basics

Data is stored in triples

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




# Wikidata Data Model



# Account Creation

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






Special page

## Create account

Language: [Deutsch](#) | [English](#) | [Esperanto](#) | [français](#) | [español](#) | [Italiano](#) | [Nederlands](#)

**Wikidata is made by people like you.**

-  **1,834,895,509** edits
-  **101,943,288** items
-  **24,738** recent contributors

**Username** [\(help me choose\)](#)

**Password**

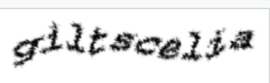
It is recommended to use a unique password that you are not using on any other website.

**Confirm password**

**Email address (optional)**

To protect the wiki against automated account creation, we kindly ask you to enter the words that appear below in the box ([more info](#)):

CAPTCHA Security check



[Refresh](#)

[Can't see the image?](#)

**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**  
[Special pages](#)  
[Printable version](#)

# 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

- Reasonator*: Adds a link to [Reasonator](#) on every item.
- DraggableSitelinks*: Easily add a Wikipedia as an "imported from" (P143) reference to any statement.
- Site ID to interwiki*: replaces the site ID with the interwiki code in the sitelinks section (example: enwiki > en).
- Ancestors*: Adds a link to the ancestors tool on Toolforge on every item.
- Protection indicators*: Adds a padlock page indicator for protected entities.
- Mark as patrolled*: Adds a [Mark as patrolled] link to change list items, which have the red exclamation mark.
- Primary Sources*: Allows to curate data donated to Wikidata.
- DuplicateReferences*: Adds a link to copy references and add them to other statements on the same item.
- EasyQuery*: A tool to add an icon for finding items with same statements easier.
- Drag'n'drop*: Add statements and references from Wikidata or Wikipedia by dragging and dropping them.
- CoordinateDiffMap*: Shows a map in the diff view when the given changes involve coordinates.
- currentDate*: Automatically adds the date of today while using the property [retrieved](#) (P813).
- ImageHeader*: Shows an image at the top right corner representing the item. (It can only show an image if the item has a property with an image defined.)
- SearchAll*: Search in the other namespaces of Wikidata too.
- Recoin* ("[Relative Completeness Indicator](#)") is a script that extends Wikidata entity pages with information about the relative completeness of the information. Note: This gadget depends on website [recoin.toolforge.org](https://recoin.toolforge.org).

Tool to Visualize Data

# Scholia Homepage

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help

Scholia is a service that creates visual scholarly profiles for [topics](#), [people](#), [organizations](#), [species](#), [chemicals](#), etc using bibliographic and other information in Wikidata. [More info...](#)

Scholia relies on Wikidata, and Wikidata contains only a limited albeit growing subset of the corpus of scholarly literature, its authors and citations. Read more about the limitations in the [FAQ](#) or check the [statistics](#).

## Search

Search for a scientist, topic, publication, organization, award, event, etc.

## Examples

Profiles	Combinations	Redirects
<p><a href="#">Denny Vrandečić</a></p> <p>View the researcher profile for the Semantic Web researcher Denny Vrandečić. It shows his papers, co-authors, etc.</p> <p><a href="#">Technical University of Denmark</a></p> <p>View the profile for an organization: People associated with the organization, their publications, the co-author patterns, etc.</p>	<p>Scholia can show multiple items together.</p> <p><a href="#">Technical University of Denmark and University College London</a></p> <p>Compare two or more organizations. Here a comparison between two universities with collaborating researchers, number of publications and citations.</p>	<p>If you know the external identifier of a concept, then Scholia can make a lookup based on it:</p> <p><a href="#">twitter/utafrith</a></p> <p>Look up by Twitter username @utafrith. This will identify the London-based researcher Uta Frith and redirect to her Scholia page.</p> <p><a href="#">twitter/mitpress</a></p>

- Web service for scholarly profiles using Wikidata

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

# Scholia Profile

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help Search...

organization / Q738258 Improve data

## University of Miami (Q738258)


The University of Miami is a private research university in Coral Gables, Florida. As of 2021, the university enrolled 19,096 students in 12 colleges and schools across nearly 350 academic majors and programs, including the Leonard M. Miller School of Medicine in Miami's Health District, the law school on the main campus, and the Rosenstiel School of Marine and Atmospheric Science on Virginia Key with research facilities in southern Miami-Dade County. [\(Read more on English Wikipedia\)](#)

Related: [Utah State University](#) · [Florida State University](#) · [Stony Brook University](#) · [Tulane University](#) · [American University](#) · [San Francisco State University](#) · [University of South Carolina](#) · [University of Würzburg](#) · [University of Basel](#) · [University of Hawai'i](#)

[@univmiami](#) | <https://ror.org/02dgjyy92> | <https://welcome.miami.edu/>

### Table of Contents

- [Employees and affiliated](#)
- [Co-author graph](#)
- [Advisor graph](#)
- [Topics that employees and affiliates have published on](#)
- [Recent publications](#)
- [Uses](#)
- [Page production](#)
- [Citations](#)
  - [Recent citations](#)
  - [Most cited papers with affiliated first author](#)
  - [Co-author-normalized citations per year](#)
- [Awards](#)
- [Gender distribution](#)



- Profile for the University of Miami

# Scholia Search Results

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help

## Search

Elizabeth A Babcock Search

No results

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help

## Search

Elizabeth A. Babcock Search

No results

**SCHOLIA** Author Work Organization Location Event Project Award Topic Tools Help

## Search

Elizabeth Babcock Search

Show all results for "**Elizabeth Babcock**"

**Elizabeth Babcock** - - 1808 (Q101246113)

- 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

Source MetaData Tools Issues Talk 

The IDs and action you enter below will be added to a batch system.  
It will run in the background, you won't have to keep the browser open.

Input list

IDs (one per row), or SPARQL (see checkbox)

The above is a SPARQL query, rather than a list of IDs; first returned variable will be used as IDs

[Check/add papers](#) (list: one PMID/DOI/PMCID per row)

[Add authors to publications](#) by creating new items, and adding to existing authors where possible (list: one Wikidata publication item per row)

[Add metadata from ORCID authors to Wikidata](#) (list: one Wikidata author item *or* ORCID ID per row)


[Create/amend papers for ORCID authors](#) *slow!* (list: one Wikidata author item *or* ORCID ID per row)

[Create books from ISBN](#) (one ISBN per row; can be 'ISBN|Title' if known, otherwise auto-lookup)

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

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

# Author Disambiguator Tool

Author Disambiguator [Names](#) [Authors](#) [Works](#) [Batches](#) [Lists](#) 

Wikimedia user account: Mlemusrojas ([log out](#))

Author name:   Fuzzy match  Wikibase search  Specify name strings  Limit:

Additional SPARQL filters separated by semicolons (eg. for papers on Zika virus, enter wdt:P921 wd:Q202864):   Filter potential authors as well?

## Potential publications

1 publications found

[Click here to create clusters based on exact author strings rather than rougher matches.](#)

### Misc

[Check all](#) | [Uncheck all](#)

Title	Authors (identified)	Published in	Identifier(s)	Topic	Published Date	Match?
<a href="#">Understanding Migration and Remittances: Lessons from Nigeria</a>	<a href="#">[1] Una Osili</a>	African Economic Development [curation]	DOI: <a href="#">10.1108/978-1-78743-783-820192022</a> [ORCID]		2019-01-01	

## Potential author items

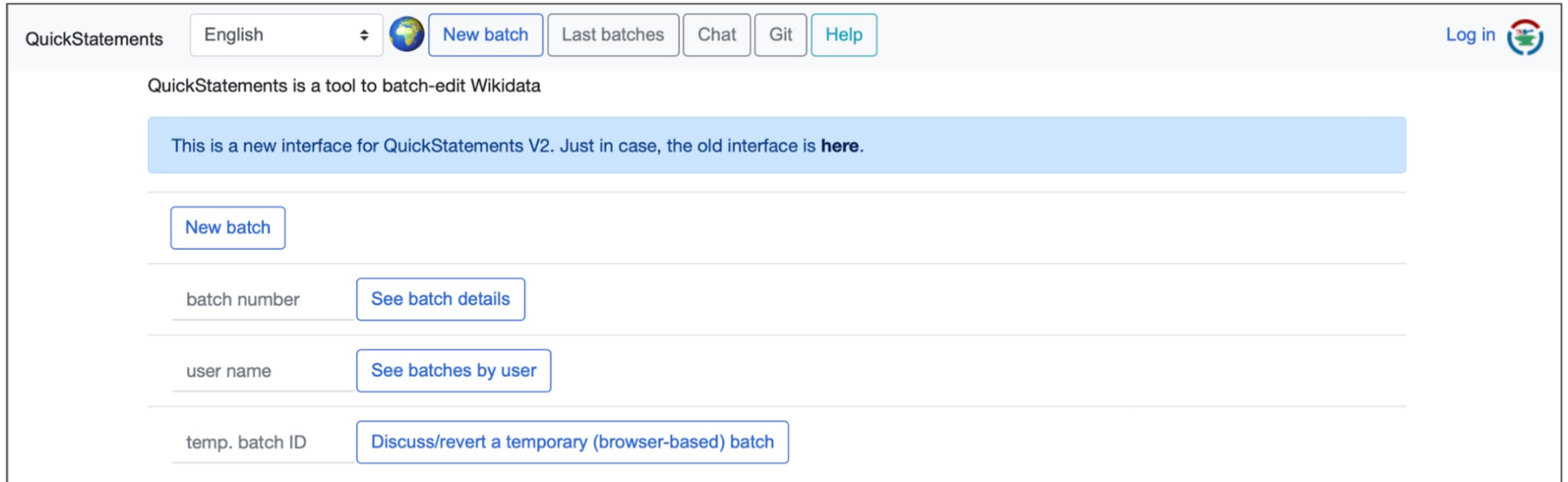
Name	Description	Authored items	Identifiers	Employer(s)
<input type="radio"/> <a href="#">Una Osili</a>	professor of economics and philanthropic studies	28	ORCID: <a href="#">0000-0001-7447-8857</a>	Indiana University – Purdue University Indianapolis [curation]

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

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

# QuickStatements

- Executes automatic edits in Wikidata



The screenshot shows the QuickStatements website interface. At the top left, there is a navigation bar with the text "QuickStatements" and a language dropdown menu set to "English". To the right of the language menu are several buttons: "New batch" (highlighted in blue), "Last batches", "Chat", "Git", and "Help". In the top right corner, there is a "Log in" link and a small globe icon.

Below the navigation bar, the main content area starts with the text "QuickStatements is a tool to batch-edit Wikidata". A light blue banner contains the message: "This is a new interface for QuickStatements V2. Just in case, the old interface is **here**." Below this banner, there is a "New batch" button. Further down, there are three rows of links:

- Next to the text "batch number" is a button labeled "See batch details".
- Next to the text "user name" is a button labeled "See batches by user".
- Next to the text "temp. batch ID" is a button labeled "Discuss/revert a temporary (browser-based) batch".

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

# QuickStatements Example

Status:	DONE	100% (3) of 3 done
1	<b>done</b> Cynthia M. Adams [Q97702985]	<b>CREATE</b> Item
<p><i>en</i>:Cynthia M. Adams <i>en</i>:vice dean and clinical professor of law instance of [P31]:human [Q5] sex or gender [P21]:female [Q6581072] languages spoken, written or signed [P1412]:English [Q1860] affiliation [P1416]:Indiana University Robert H. McKinney School of Law [Q14688472] employer [P108]:Indiana University – Purdue University Indianapolis [Q1433199] reference URL [P854]:<a href="https://www.linkedin.com/in/cynthia-m-adams-34113514/">https://www.linkedin.com/in/cynthia-m-adams-34113514/</a> retrieved [P813]:2020-07-26 work location [P937]:Indianapolis [Q6346] occupation [P106]:university teacher [Q1622272] official website [P856]:<a href="https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=43">https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=43</a> language of work or name [P407]:English [Q1860] educated at [P69]:Kentucky Wesleyan College [Q1739079] academic degree [P512]:Bachelor of Arts [Q1765120] reference URL [P854]:<a href="https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=43">https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=43</a> retrieved [P813]:2020-07-26 educated at [P69]:Indiana University Robert H. McKinney School of Law [Q14688472] start time [P580]:1979 reference URL [P854]:<a href="https://www.linkedin.com/in/cynthia-m-adams-34113514/">https://www.linkedin.com/in/cynthia-m-adams-34113514/</a> end time [P582]:1983 academic degree [P512]:Juris Doctor [Q1540185] retrieved [P813]:2020-07-26</p>		
2	<b>done</b> Miriam A. Murphy [Q97702987]	<b>CREATE</b> Item
<p><i>en</i>:Miriam A. Murphy <i>en</i>:director of the Ruth Lilly Law Library instance of [P31]:human [Q5] sex or gender [P21]:female [Q6581072] languages spoken, written or signed [P1412]:English [Q1860] affiliation [P1416]:Indiana University Robert H. McKinney School of Law [Q14688472] employer [P108]:Indiana University – Purdue University Indianapolis [Q1433199] start time [P580]:2000 reference URL [P854]:<a href="https://www.linkedin.com/in/miriam-murphy-8799208/">https://www.linkedin.com/in/miriam-murphy-8799208/</a> retrieved [P813]:2020-07-26 employer [P108]:Wake Forest University School of Law [Q7960935] reference URL [P854]:<a href="https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=71">https://mckinneylaw.iu.edu/faculty-staff/profile.html?id=71</a> retrieved [P813]:2020-07-26</p>		

Thank You!

[Q19279214](#)

