

The UCLA Library and Wikidata: A Pilot

Paromita Biswas, Continuing Resources Metadata Librarian
Iman Dagher, Arabic & Islamic Studies Catalog Librarian
Erica Zhang, Metadata Librarian for Open Access

UCLA Library Research Forum
March 23, 2021

PCC-UCLA Wikidata Pilot

UCLA Library joined the PCC Wikidata Pilot in September of 2020. The Pilot is part of a broader move by the PCC to transform library data (traditional MARC based catalogs) into linked data supporting the twin goals of identity management and better information discovery across data silos.

Linked data is sharing structured data on the web that is machine readable and can be interlinked (Tim Berners-Lee: use HTTP URIs for names of things).
“Linked data is fundamentally about building a web of data.”*

Identity management moves away from creating text strings for identifying entities in favor of registered identifiers: VIAF; ISNI; ORCID (these have HTTP URIs)

* [Introduction to the Semantic Web](#)

- What is PCC?
- Recent directions of PCC from siloed data to linked data, and a big step towards that is moving from traditional cataloging to identity management (e.g., name authority disambiguation).
 - different pilots, i.e. ISNI, URIs in MARC, Wikidata

PCC-UCLA Wikidata Pilot

The Wikidata Pilot is the 3rd Pilot for the UCLA Library supporting these goals:

- ISNI Pilot (identity management)
Stephen Frank (*ISNI: 0000 0001 1598 4821*)
- PCC URIs in MARC (linked data)
Stephen Frank (*URI: <http://isni.org/isni/0000000115984821>*)
- PCC Wikidata Pilot (linked data/identity management)
Stephen Frank (*Wikidata: Q105871437*
<http://www.wikidata.org/entity/Q105871437>)

MARC and Wikidata

SRN	2587644	In Distribution
010	1	n 88219068
040		DLC #b eng #e rda #c DLC #d DLC #f CLU
024	7	0000000391577675 #2 isni #1 http://isni.org/isni/0000000115984821
024	7	14851683 #2 viaf #1 http://www.viaf.org/viaf/14851683
024	7	Q105871437 #2 wikidata #1 https://www.wikidata.org/entity/Q105871437
046		#f 1955 #2 edit
100	1	Frank, Stephen, #d 1955-
373		University of California, Los Angeles. Department of History #2 naf
374		Historians #2 lcsh
374		University and college faculty members #2 lcdgt
667		URIs added to this record for the PCC URI MARC Pilot. Please do not remove or edit the URIs
670		The World of the Russian peasant, c1990: #b CIP t.p. (Stephen Frank) data sh. (b. 8/28/55)
670		Cultures in flux, 1994: #b CIP t.p. (Stephen P. Frank)
670		UCLA history department faculty website, viewed on April 29, 2020: #b (Associate Professor of history, teaches Russian history)

Stephen Frank (Q105871437)

historian of Russia and Europe

Stephen Frank is a academic, university teacher, and faculty member.

He was born on August 28, 1955.

He studied at Brown University. His field of work includes history of Russia and history of Europe. He worked for University of California, Los Angeles.

Other properties

field of work history of Russia occurrences and people in Russia throughout history

history of Europe history of Europe, including the continent and nearby islands

employer University of California, Los Angeles public research university in Los Angeles, California, United States

on focus list of Wikimedia project WikiProject PCC Wikidata Pilot UCLA Wikidata pilot project at the UCLA Library

date of birth 1955-08-28

educated at Brown University

academic degree point in time: 1987

Q49114

University in Providence, Rhode Island



given name Stephen male giv

private university, research university, Colonial Colleges, and private not-for-profit educational institution in United States of America from Providence

academic appointment UCLA Department of History

start time: 1993

Timeline

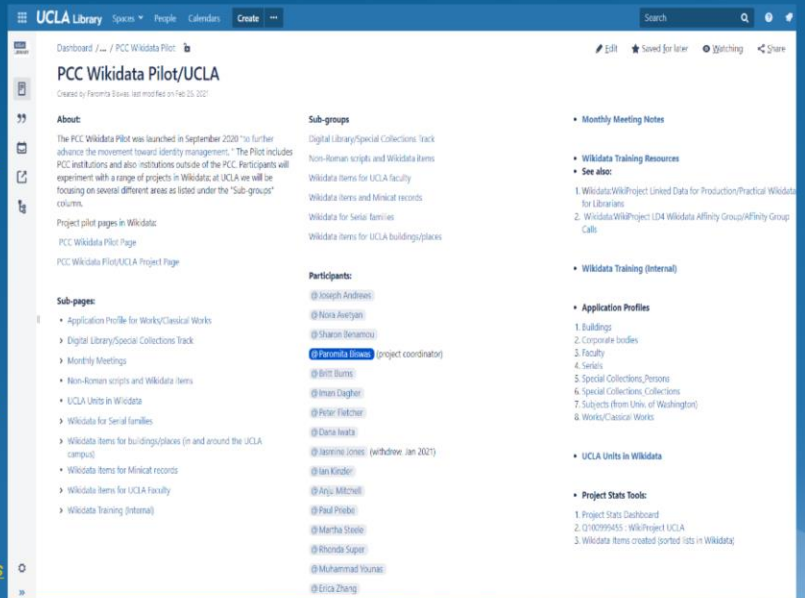
Wikidata and UCLA Pilot

Pilot ideal for virtual work:
Existing tools; low barrier
Wikidata ecosystem:
(Affinity group/Wikidata community:
26K+ contributors/93M+ items)*

Departmental meeting for the
project: 15 catalogers signed up.

6 subgroups:
Faculty; Corporate bodies; Campus
buildings; Non-Roman scripts;
eScholarship journals;
Special/Digital Collection items
(UCLA focus)

**Wikidata:Statistics*



- UCLA answered the call of participation to the pilot when it started, and also aligning/getting involved with the Wikidata Affinity Group and the general huge community
 - Perfect time to start this pilot during remote work
 - Confluence page
 - Subgroups

Wikimedia Sister Projects

- Free & publicly available
- Cross-linking
- Complement each other



We are all familiar with Wikipedia, the free large encyclopedia that we tend to consult first when we want to find easy and fast information. Wikidata is a sister project operated by the [Wikimedia Foundation](https://www.wikimedia.org/). It is considered as the linked-data repository, It is developed to provide Wikipedia and its sister projects with structured and reusable data. And that what makes libraries and GLAM institutions interested in exploring the use of Wikidata, as Paromita has mentioned? **NEXT SLIDE**

Wikidata:



- Launched in October 2012
- The largest Wikimedia project
- Free, open knowledge base (data has Creative Commons CC0 License)
- Collaborative platform
- Multilingual
- Central storage for **structured & reusable linked data**
- Integrated with the semantic web
- Read and edited by both human & machine

Wikidata was launched in October of 2012.

Wikidata is currently the largest Wikimedia project, with over 95 million content pages. It is a free and open knowledge base; anyone can create an account to add and edit content.

It is a collaborative platform to share and spread resources, with a community of thousands of volunteers.

Wikidata is an international and multilingual project.

It is also considered a central storage for **structured linked data**, **data** read and edited by both human and machine.

Let's look closely at the structure of the data in this knowledge base. **[NEXT SLIDE]**

The image shows a screenshot of the Wikipedia article for Noam Chomsky. At the top, there is a navigation bar with user options like 'imandag' and 'Talk: Sandbox'. Below that is a search bar and a banner for 'Celebrate International Women's Day'. The article title 'Noam Chomsky' is displayed, followed by a summary paragraph and a detailed biographical paragraph. A portrait of Chomsky is on the right. In the left sidebar, the 'Wikidata item' link is circled in red.

In Wikipedia, entities are presented in articles written in an encyclopedic style of language and textual information.

Most Wikipedia articles are presented in Wikidata; the link to access the Wikidata item of the entities is found in the left column.

We see on the slide the article on Noam Chomsky, the well-known American linguist.

NEXT SLIDE

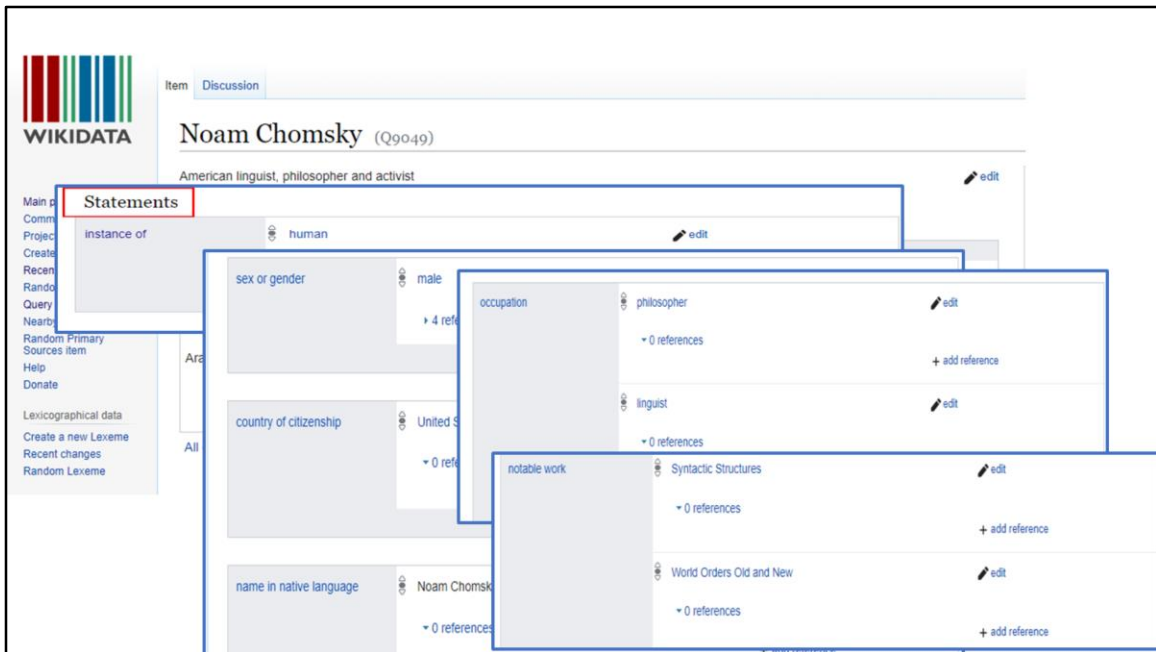
Wikidata page for Noam Chomsky (Q9049). The page shows the item name, its identifier, a description, and aliases. A table lists the item in multiple languages.

Language	Label	Description	Also known as
English	Noam Chomsky	American linguist, philosopher and activist	Avram Noam Chomsky A. Noam Chomsky Chomsky, Noam Chomsky
Arabic		فيلسوف وناشط سياسي وعالم لسانيات أمريكي نعوم تشومسكي	نعوم شومسكي ناعوم تشومسكي نوم تشومسكي نعام تشومسكي

In Wikidata, information about Chomsky is presented in a structured way. He is identified by a single item with identifier that start with Q

Each item needs to include a description; a short phrase designed to disambiguate items with the same or similar labels.

There are also **aliases**, or alternative names for items, as well as what to call Chomsky in other languages. **[NEXT SLIDE]**



But obviously there is a lot more to tell about Chomsky, such as his nationality, date and place of birth, education, employment, and notable works or publications.

These types of information can be expressed in statements.

Statements are the heart of the structured data, and what make Wikidata a source for acquiring richer descriptions and a platform for publishing, linking, and enriching library linked **data**.

[NEXT SLIDE]

Structured Data in Wikidata: Statements

Item = starts with **Q** (Q.ID) persistent & unique identifier
(people, organizations, places, articles, books, concepts, topics, chemicals, events, artworks, etc.)

Property = starts with **P**, describes the data value (instance of, location, capital of, field of work, language)

Value(s) = the actual piece of data that describes the item (can be *literal* or a *link*)

Statements = Item ?? Property ?? Value

RDF: Subject ?? Predicate ?? Object

Noam Chomsky (Q9049) ?? occupation (P106) ?? philosopher (Q4964182)

RDF:

<https://www.wikidata.org/wiki/Q9049> > < <https://www.wikidata.org/wiki/Property:P106> > < <https://www.wikidata.org/wiki/Q4964182>

In Wikidata, we can create items to represent all different things. It can basically cover all domains of knowledge: people, organizations, places, scientific articles, books, concepts, topics, chemicals, events, artworks, etc.

Each item would have a unique number that starts with Q (called QID).

In addition to items, there are also properties that start with a P. Properties are used to describe a specific aspect of data about the items.

Value is the actual piece of data that answers to the specific aspect defined in the property. We can record more than one value.

While -- Anyone can create items, Creating new properties requires a property proposal that is formally discussed online.

Items + Properties and values make statements that have the general form of triple.

This model aligns well with the RDF triple model of the semantic web.

[NEXT SLIDE]

Structured Data in Wikidata: Statements

References: Sources of information

[stated in \(P248\)](#)
referring to publications and media
[reference URL \(P854\)](#)
used for websites and online databases

Qualifiers: allow statements to be expanded on, annotated, or contextualized

The screenshot displays the Wikidata interface for editing a statement. The property is 'place of birth' and the main value is 'East Oak Lane'. Below this, there are three qualifiers: 'located in the administrative territorial entity' with values 'Philadelphia' and 'Pennsylvania', and 'country' with the value 'United States of America'. A red box highlights this entire section. Below the main statement, there is a section for '3 references'. The first reference is 'imported from Wikimedia project' with the value 'Russian Wikipedia'. The second reference is 'stated in' with the value 'Integrated Authority File' and a qualifier 'retrieved' with the value '10 December 2014'. The third reference is 'stated in' with the value 'Great Soviet Encyclopedia (1969–1978)', a 'title' qualifier with the value 'Холмский Ноам', and a 'retrieved' qualifier with the value '28 September 2015'. A green box highlights this entire reference section.

One interesting feature of Wikidata is that provenance and attribution can easily be included using the references and qualifiers, which are core to the Wikidata data model.

1) **References** are used to point to specific **sources** that back up the data provided in a statement:

2) **Qualifiers** are used in order to further describe or refine the value of a property given in a statement. Example of qualifiers: the start and ending time, specify the value (such as position held, academic major ..)

Like statements, Qualifiers and references also consist of a property and a value.

[NEXT SLIDE]

Noam Chomsky (Q9049)
 American linguist, philosopher and activist
 Avram Noam Chomsky | A. Noam Chomsky | Chomsky, Noam | Chomsky

Statements

educated at

↓

property

Rank ←

University of Pennsylvania	→ value	edit
end time	1949	
academic degree	Bachelor of Arts	
start time	1945	
▼ 0 references		
+ add reference		
University of Pennsylvania		edit
end time	1951	
academic degree	Master of Arts	
start time	1949	
▼ 0 references	→ reference	
+ add reference		
Harvard University		edit
academic major	linguistics	
end time	1955	
start time	1951	
▼ 0 references		

We can also add ranks to data with statements that have multiple values. The default rank is the "normal" rank;
 it may also be marked with "preferred" or "deprecated" ranks.
 For example, the current city mayor would have the preferred rank.
 The *deprecated* rank is used for statements that are known to include errors, or that represent outdated knowledge. **[NEXT SLIDE]**

13

Wikidata: A Potential Authority Hub

The screenshot shows the Wikidata page for Noam Chomsky. The page is titled "Noam Chomsky (Q9049)" and includes a description: "American linguist, philosopher and activist". The page is divided into several sections, with a red box highlighting the "Identifiers" section. This section lists various persistent identifiers for Noam Chomsky, including VIAF ID, ISNI, BNMM authority ID, Vatican Library VcBA ID, BIBSYS ID, National Library of Brazil ID, CANTIC ID, National Library of Chile ID, and Biblioteca Nacional de España ID. Each identifier is accompanied by its value and the number of references it has.

Identifier	Value	References
VIAF ID	89803084	3 references
ISNI	0000 0001 2143 0000	1 reference
BNMM authority ID	000047047	0 references
Vatican Library VcBA ID	495/212047	1 reference
BIBSYS ID	50052767	0 references
National Library of Brazil ID	00057753	1 reference
CANTIC ID	a10065660	1 reference
National Library of Chile ID	00007482	1 reference
Biblioteca Nacional de España ID	XX849732	0 references

Wikidata can act as a central hub that connects with vast federated ecosystems of knowledge bases that can interact with Wikidata.

That can be done through the addition of persistent identifiers from all sorts of databases and authority lists, such as the ones on the slide (VIAF, ISNI, ORCID, and many others). **[NEXT SLIDE]**

Wikidata: A Potential Authority Hub

- Wikidata links to external data via hundreds of different types of identifiers
- Add Wikidata IDs in library catalogs and resources (such as in MARC records or index database)
- Wikidata is considered a low-barrier method for creating and using LOD in libraries

Wikidata links to external data via hundreds of different types of identifiers
We can add Wikidata IDs in library catalogs and resources (such as in MARC records or index database) and we have been doing this in our catalog as part the the PCC MARC URI pilot that Paramito has mentioned earlier
Wikidata is considered as a low-barrier method for creating and using LOD in libraries

[NEXT SLIDE]

Wikidata: A Multilingual Project

The screenshot shows the Wikidata 'Preferences' page for user 'Imandag'. A red box highlights the 'Babel user information' section, which includes a search bar and a list of language proficiency levels: en-5 (professional), ar-N (native Arabic), en-0 (no knowledge), en-1 (basic), en-2 (intermediate), en-3 (advanced), en-4 (near native), and en-N (native). To the right, a list of 132 language entries for 'Wikipedia' is shown, with 'en' (English) selected. The list includes various language codes and their corresponding labels in different scripts, such as 'af' (Afrikaans), 'als' (Albanian), 'am' (Amharic), 'ang' (Anglo-Frisian), 'an' (Aragonese), 'ar' (Arabic), 'arz' (Egyptian Arabic), 'ast' (Asturian), 'azb' (Azerbaijani), 'az' (Azerbaijani), 'bat_smg' (Baltic), 'ba' (Bashkir), 'be_x_old' (Old Belarusian), 'be' (Belarusian), 'bg' (Bulgarian), 'bh' (Bihari), 'bn' (Bengali), 'bo' (Tibetan), 'br' (Breton), 'bs' (Bosnian), 'bxr' (Buryat), 'ca' (Catalan), 'ce' (Chechen), 'ckb' (Kurdish), 'cs' (Czech), 'cu' (Church Slavonic), 'cy' (Welsh), 'da' (Danish), 'de' (German), 'diq' (Dialect), 'el' (Greek), 'en' (English), and 'eo' (Esperanto).

I mentioned earlier that Wikidata is an international and thus a multilingual project. Across all Wikidata entities, 410 languages are used. While English is the default interface language, the project is intended to be used by, and be useful for, users of every language possible

Every Wikidata entry for a page contains a list of links for that page in different languages.

As we see in this slide, chomsky's entry is available in 123 entries in different languages

All Wikidata editors, no matter what language they choose, are working on the same, common data set, this is accomplished by representing all entities in the system with internal ids, which are then translated on display to their labels in the user's language. Editors can edit data in more than one language. There are tools and gadgets available to make internationalization easier. **[NEXT SLIDE]**

Editing in Wikidata: Human, Bots, Gadgets, and Other Tools



- Computer program/script that can interact with Wikidata or Wikibase (software of Wikidata)
- Bots are automated tools used to make edits such as: add labels, interwiki links, descriptions, statements sources, etc. (without the necessity of human decision-making)
- Gadgets are scripts or programs written by users to make tasks easier and more efficient
 - E.g. AuthorityControl, Merge, autoEdit, labellister, ReasonatorTools, etc.
- Need to enable gadget from the user accounts
- **QuickStatements is one great tool for batch editing**

• (cur | prev) 14:30, 26 December 2023 Reinheitsgebot (talk | contribs) (11,952 bytes) (+366) ... (Created claim: Guggenheim fellows ID (P6594): andras-j-e-bodrogligeti; #quickstatements; invoked by Mix'n'match add_person_dates)

• (cur | prev) 22:22, 24 December 2023 Reinheitsgebot (talk | contribs) (11,586 bytes) (+350) ... (Created claim: CONOR.SR ID (P8851): 27585383; #quickstatements; invoked by Mix'n'match add_person_dates) (undo) (restore)

In addition to the work done by humans, Wikidata is enriched by the use of many tools, bots, scripts, and gadgets that facilitate the process, with or without the necessity of human decision-making.

Such tools can create items and add [labels](#), [descriptions](#), [statements](#), [sources](#), [interwiki links](#), and other functionalities.

QuickStatements is one great tool for batch editing

The example on the slide from the editing history of the item for "UCLA Andras Bodrogligeti", indicates the editing that was done by a bot to add identifiers to his items

[NEXT SLIDE](#)

Tools & Gadgets

Project page Discussion Read Edit View history More Search Wikidata

Wikidata:Tools

Translate this page

Other languages: Bahasa Indonesia • Bahasa Melayu • Basa Bali • British English • Canadian English • Cymraeg • Deutsch • English • Esperanto • Frysk • Ilokano • Jawa • Lëtzebuergesch • Minangkabau • Nederlands • Ripoarisch • Sesotho sa Leboa • Tiếng Việt • Türkçe • Zazaki • asturianu • bosanski • català • dansk • español • euskara • français • frysk • galego • homoserbce • hrvatski • interlingua • italiano • latviešu • lietuvių • magyar • norsk bokmål • occitan • polski • português • português do Brasil • română • shqip • slovenčina • slovenščina • srpskohrvatski / српскохрватски • suomi • svenska • siemiatka • čeština • עברית • беларуская (тарашкевіца) • македонски • русски • српски / srpski • українська • қазақша • հայերեն • עברית • اردو • العربية • سنڌي • فارسی • مصرى • پښتو • کوردی • मराठी • हिन्दी • বাংলা • ગુજરાતી • తెలుగు • தமிழ் • ཐཱ་མ་སྐད་ • བོད་སྐད་ • རྩིས་སྐད་ • 中文 • 日本語 • 粵語 • ᱵᱟᱲᱟ • 한국어

Tools

Enter search term

Search for a tool

This page is a list of tools you can use to work at Wikidata more quickly and comfortably. If you have written a script just add it and help many other users.

See also the list of Wikidata tagged tools# on Toolforge.

Edit items **Query data** **Enhance user interface** **Visualize data** **List properties** **Lexicographical data** **For programmers** **Non-Wikidata**

Tool of the week

Each week, a popular tool is featured in the weekly summary.

<https://www.wikidata.org/wiki/Wikidata:Tools>

This wikidata page is a list of tools you can use to work with Wikidata more quickly and comfortably.

It lists the tools by the functionality such as : editing, query, visualize data, etc. **[NEXT SLIDE]**

Wikidata Query Service

The screenshot displays the Wikidata Query Service interface. On the left, the 'Query Helper' section allows for building a query using a graphical interface. It includes filters for 'item' (with 'educated at' and 'employer' dropdowns), 'organization' (with 'part of' and 'parent organization' dropdowns), and 'ORCID ID' (with 'orcid' dropdown). A 'Limit' field is also present. On the right, a SPARQL query is shown in a text editor. The query is as follows:

```
1 #Find researchers who are employed by or have as their alma mater UCLA,
2 #and return name and their ORCID
3 SELECT DISTINCT ?item ?itemLabel ?itemDescription ?orcid ?organization
4 ?organizationLabel WHERE {
5   ?item wdt:P496 ?orcid.
6   ?item (p:P69|p:P108) [ (ps:P69|ps:P108) ?organization ] .
7   ?organization (wdt:F361|wdt:F749)* wd:Q174710 .
8   ?item ?p ?organization .
9   SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
10 }
```

At the bottom right of the interface, it indicates '2382 results in 1665 ms' and provides options for 'Code' and 'Download'.

Try It! <https://w.wiki/35p4>

Once the data is in Wikidata, you can pull the data for analysis and/or visualization through a number of ways. Wikidata's query service provides an interface to write queries using the SPARQL query language.

Wikidata also provides an API (application programming interface) if you want to use the data in your own applications, or create your own tools to work with what is in Wikidata. **[NEXT SLIDE]**

Wikidata: A Rich Ecosystem

“A rich ecosystem of human editors and bots patrols the knowledge base and its entries to enforce data quality and consistency”.

- Tools for bulk loading of data with error checks
- Discussion pages for every item in Wikidata
- Project chats: place to discuss all aspects of Wikidata; policy and proposals, individual data items, technical issues, etc.
- Document history of the item with option to undo changes
- Proposal procedures

<https://www.wikidata.org/wiki/Wikidata:Tools>

So with this rich ecosystem of human editors and bots, there is a considerable effort to enforce data quality and consistency, which is of great concern to libraries' communities.

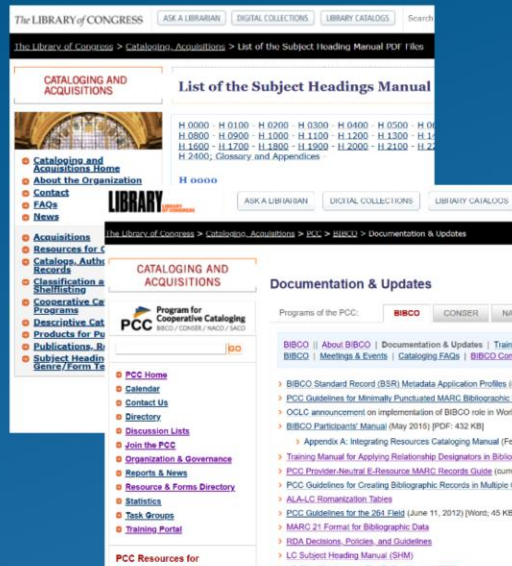
There are tools for detecting errors and the misuse of data, discussion pages for every item in Wikidata, along with a saved history of the edits applied to the item, with options to undo changes.

Also: Project chats is a place to discuss all aspects of Wikidata, such as policy and proposals, issues with individual data items, technical issues, etc.

In sum, there are lots of opportunities and potential for library use, but also some challenges that librarians/catalogers may encounter. I will switch over to Erica to tell us more about this.

[NEXT SLIDE]

From Traditional Cataloging to Wikidata



<https://www.loc.gov/aba/publications/FreeSHM/freeshm.html>
<https://www.loc.gov/aba/pcc/bibco/doc-updates.html>

So it's been almost half a year since we started this pilot and got our hands into Wikidata, and we've taken note of differences between the traditional cataloging environment where we largely work with other catalogers, and the Wikidata environment where we are essentially working with anyone who chooses to contribute.

One difference is in rules that govern work in each environment. With traditional cataloging, rules take largely a top-down approach with Library of Congress and relevant committees defining rules to address ambiguities and complexities in bibliographic data, and bring consistency to cataloging work. These screenshots are just a sample of the amount of documentation available to guide cataloging work.

[next slide]

From Traditional Cataloging to Wikidata

The screenshot shows the Library of Congress website. The top navigation bar includes 'The LIBRARY of CONGRESS', 'ASK A LIBRARIAN', 'DIGITAL COLLECTIONS', and 'LIBRARY CATALOGS'. Below this, there's a search bar and a breadcrumb trail: 'The Library of Congress > Cataloging > Acquisitions > List of the Subject Heading Manual PDF Files'. The main content area is titled 'CATALOGING AND ACQUISITIONS' and 'List of the Subject Headings Manual'. It features a table of subject headings (H 0000 to H 2500) and a sidebar with various links like 'Cataloging and Acquisitions Home', 'About the Organization', 'Contact', 'FAQs', and 'News'. Below the table, there's a section for 'Documentation & Updates' for the PCC, listing various documents and manuals.

<https://www.loc.gov/aba/publications/FreeSHM/freeshm.html>
<https://www.loc.gov/aba/pcc/bibco/doc-updates.html>

The screenshot shows the Wikidata:WikiProject Books page. The title is 'Wikidata:WikiProject Books'. Below the title, there's a list of other languages. The main content area is titled 'The aim of the WikiProject Books is:' and lists several goals: 'to define a set of properties that can be used both by book infoboxes, Commons book templates, and Wikisource', 'to map and import all relevant data from existing catalogs, Wikisource, Wikispecies, and Wikisource', and 'establish methods to interact with Wikidata'. Below this, there's a list of properties under the heading 'Bibliographic properties'. The list includes: '1 Bibliographic properties', '1.1 Wikitext properties', '1.2 Editor form properties', '1.2.1 Integrated properties', '1.2.2 Qualifiers', '1.3 Example properties', '1.4 Manuscript properties', '1.5 Physical and collection of works', '1.6 Mapping', '2 Parameters', '3 Related items', '4 Tools', '4.1 Items without claims', '4.2 Items with no available book record', '5 Navigation boxes', '6.1 Books (E&I)', '6.2 Bibliographical properties (E&I)'. Below the list, there's a section for 'Bibliographic properties'.

https://www.wikidata.org/wiki/Wikidata:WikiProject_Books
https://www.wikidata.org/wiki/Wikidata:WikiProject_Stanford_Libraries/Data_models

The screenshot shows the Wikidata:WikiProject Stanford Libraries/Data models page. The title is 'Wikidata:WikiProject Stanford Libraries/Data models'. Below the title, there's a list of other languages. The main content area is titled 'The aim of the WikiProject Stanford Libraries/Data models is:' and lists several goals: 'to define a set of properties that can be used both by book infoboxes, Commons book templates, and Wikisource', 'to map and import all relevant data from existing catalogs, Wikisource, Wikispecies, and Wikisource', and 'establish methods to interact with Wikidata'. Below this, there's a list of properties under the heading 'Bibliographic properties'. The list includes: '1 Bibliographic properties', '1.1 Wikitext properties', '1.2 Editor form properties', '1.2.1 Integrated properties', '1.2.2 Qualifiers', '1.3 Example properties', '1.4 Manuscript properties', '1.5 Physical and collection of works', '1.6 Mapping', '2 Parameters', '3 Related items', '4 Tools', '4.1 Items without claims', '4.2 Items with no available book record', '5 Navigation boxes', '6.1 Books (E&I)', '6.2 Bibliographical properties (E&I)'. Below the list, there's a section for 'Bibliographic properties'.

In contrast, Wikidata is more community driven, and as Iman mentioned, there is infrastructure by way of project chats, proposal discussions, and the like to get community input and reach consensus. Wikidata, while not absent of some core rules, generally allows contributors flexibility in how they might want to model and represent their data, and consistency usually comes from looking for examples of similar items or projects that the community has worked on. And these screenshots are an example of Wikiprojects where some group has shared how they're representing their data so that others can follow their data models. And so that has been something of an adjustment, shifting from trying to find defined rules to looking at analogous work, such as what other academic libraries are doing with respect to faculty, for instance

[next slide]

From Traditional Cataloging to Wikidata

Week 1:

A SACO proposal is created in the [SACO Proposal System](#)

-
-
-

Week 9-10:

After all the remaining proposals on the tentative list the list is posted to the Web.

A [Summary of Decisions](#) from the weekly editorial m

<https://www.loc.gov/aba/pcc/saco/SACOWorkflow.html>

Another big difference is the barrier to entry for participation. With traditional cataloging, for example, there is a rather high barrier to entry if a cataloger wants to create or update the Library of Congress Subject Headings (LCSH), or Name Authority Files. In addition to getting training on all the aforementioned documentation, there's also certain membership that a cataloger must have to participate. If you have the training and membership to participate in proposing new subject headings or modifications to subject headings, this is not an instantaneous process. These screenshots show the official process of requesting new or changes to existing Library of Congress subject headings, and as you can see, it can take up to 10 weeks, and can definitely be much longer. **[next slide]**

From Traditional Cataloging to Wikidata

Week 1:

A SACO proposal is created in the [SACO Proposal System](#)



Week 9-10:

After all the remaining proposals on the tentative list the list is posted to the Web.

A [Summary of Decisions](#) from the weekly editorial m

<https://www.loc.gov/aba/pcc/saco/SACOWorkflow.html>



WIKIDATA

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 a new Item

Please make sure that the item you want to create complies with the Wikidata guidelines. If you want to create an item about a living person, be mindful of the privacy policy. We appreciate it if you create a label and a description for all of your items. The first letter of your label should only be capitalized if it is a proper noun. To create a new lexeme (read [here first to learn how a lexeme is created](#)).

By clicking "Create", you agree to the [terms of use](#), and you irrevocably and exclusively assign the copyright to the Wikimedia Foundation.

Create a new item

Language:

Label:

Description:

Aliases, pipe-separated:

2 Create

<https://www.wikidata.org/wiki/Special:NewItem>

Compare that to creating a new item in Wikidata, which can be achieved essentially with 2 clicks, creating an item, filling in some basic information, and then hitting "Create".

One benefit to this lower barrier of entry is increased and more diverse participation in Wikidata, and can also be great in terms of being quicker to adapt to changes in language and concepts. For instance, something as controlled as the LCSH means that language can really lag behind in currency and become outdated or inappropriate. So especially for critical cataloging initiatives and ethical description, where we recognize that there are problematic terms with regard to describing people in LCSH, Wikidata may prove a better alternative for vocabularies that have long needed some change. **[next slide]**

From Traditional Cataloging to Wikidata

Difference between revisions of "Socrates (Q913)"

Browse history interactively

Revision as of 15:29, 14 September 2017 (restore)
2806:106e:a:fc0:a984:ad73:d7fa:c6e3 (talk)

(Changed [en] label, description and aliases: Socratico, classical Greek Athenian nada, Sokrátēs, Su-ko-la-ti, سُقْرَاتُ, Suqrāt, SwGhana ruehrwvgfκράτης, Sugeladi, Socrate, Sokrat, Sokrates v.; Sokrates)
(Tags: Mobile edit, Mobile web edit)
← Older edit

Revision as of 16:04, 14 September 2017 (restore) (undo) (thank)
Epidosis (talk | contribs)

(Reverted edits by 2806:106E:A:FCB0:A984:AD73:D7FA:C6E3 (talk) to last revision by NtkPrague)
Newer edit →

label / en	label / en
- Socratico	+ Socrates
aliases / en / 0	aliases / en / 0
	+ Σωκράτης
aliases / en / 1	aliases / en / 1
- SwGhana ruehrwvgfκράτης	
description / en	description / en
- classical Greek Athenian nada	+ classical Greek Athenian philosopher

<https://www.wikidata.org/w/index.php?title=Q913&type=revision&diff=558769887&oldid=558746737>

On the other hand, this low barrier to entry also means that catalogers may need to grapple with the idea of control and ownership over data. Once our data is in Wikidata, anyone can edit that data and change it, so the data in many ways is less stable than what we're used to within our own controlled environment. And when there are changes, they are hopefully for the better, but there are certainly cases where bad actors may be diminishing the quality of the data. Vandalism is something that Wikidata patrols, and here's just one example of someone doing some strange things to Socrates' Wikidata item, and thankfully, those changes were reverted back.

[next slide]

From Traditional Cataloging to Wikidata

The screenshot displays two Wikidata entries for education. The top entry is for a 'Doctor of Philosophy' degree, with the property 'educated at' linked to 'University of Pennsylvania' and 'point in time' set to '1955'. The bottom entry is for 'Princeton University', with the property 'educated at' linked to 'Princeton University', 'start time' '1992', 'end time' '1996', and 'academic degree' 'Doctor of Philosophy'. Red arrows show the mapping: one arrow points from the 'academic degree' label in the top panel to the 'academic degree' property in the bottom panel, and another arrow points from the 'educated at' property in the top panel to the 'educated at' property in the bottom panel. A URL <https://www.wikidata.org/wiki/Q9049> is visible in the bottom left.

Finally, going back to the relative lack of strict rules in Wikidata, contributors are free to model their data in ways that make sense to them, and whereas in traditional cataloging, the cataloging community more or less works from the same ontologies and data models, such as the RDA or Resource Description and Access standard, for the same set of data, there can be multiple ways to represent that in Wikidata, and it's been an interesting learning experience coming into contact with the variety of data models for the same concepts. Here is a screenshot showing how, for data regarding someone's education details, one contributor has chosen to use "educated at" as a property in the statement, whereas another has put it as a qualifier for "academic degree", and vice versa. **[next slide]**

From Traditional Cataloging to Wikidata

The image displays two Wikidata property value boxes. The top box is for the property 'academic degree' with the value 'Doctor of Philosophy'. It lists 'educated at' as 'University of Pennsylvania' and 'point in time' as '1955'. The bottom box is for the property 'educated at' with the value 'Princeton University'. It lists 'start time' as '1992', 'end time' as '1996', and 'academic degree' as 'Doctor of Philosophy'. A red arrow points from the 'point in time' property in the top box to the 'start time' property in the bottom box.

We can also see differences in the representation of dates, one being a more specific representation than the other. This can make it harder to query Wikidata than our own siloed data since a query-er will need to account for these modelling differences. Add to that, while in our own database, we may have a clear idea of the scope of our data, that can be difficult to gauge in Wikidata; how complete is my results set?

So these are some of the characteristics of Wikidata work that we've run into through the course of this pilot so far. But we're also just really excited about the possibilities of linking data of interest to the larger Wikidata knowledgebase, and thinking about what we can now do with our data that we couldn't do before? And we've found a few cool ideas that we hope can also spark your ideas of the possibilities with Wikidata.

[next slide]

Opportunities in Wikidata

Gregory A Bryant (Q59489478)

researcher and professor of communication

Gregory Bryant

 edit

▸ Reconcil: Most relevant properties which are absent

▾ In more languages

Configure

Language	Label	Description	Also known as
English	Gregory A Bryant	researcher and professor of communication	Gregory Bryant
Moroccan Arabic	No label defined	No description defined	
Russian	No label defined	No description defined	
Armenian	No label defined	No description defined	
Asturian	Gregory A Bryant	No description defined	
Spanish	Gregory A Bryant	No description defined	
Dutch	Gregory A Bryant	onderzoeker	

<https://www.wikidata.org/wiki/Q59489478>

So as Paromita mentioned, one of our projects is adding faculty to Wikidata. Here is a faculty member's item in Wikidata, Gregory A. Bryant, and just right off the bat, after this item was created, a bot came in and added some Dutch, speaking to the multilingual aspect of Wikidata. That was not data that I needed to put in, but it has enhanced the data and a researcher can now take advantage of these types of additions. **[next slide]**

Opportunities in Wikidata

Gregory A Bryant (Q59489478)

researcher and professor of communication

Gregory Bryant

▸ Reconcil: Most relevant properties which are absent

▾ In more languages

Configure

Language	Label	Description
English	Gregory A Bryant	researcher and professor of communication
Moroccan Arabic	No label defined	No description defined
Russian	No label defined	No description defined
Armenian	No label defined	No description defined
Asturian	Gregory A Bryant	No description defined
Spanish	Gregory A Bryant	No description defined
Dutch	Gregory A Bryant	onderzoeker

Identifiers

Loop ID ORCID 12406

▸ 1 reference

ORCID ID ORCID 0000-0002-7240-4026

▾ 0 references

Scopus author ID ORCID 10438889600

▾ 0 references

<https://www.wikidata.org/wiki/Q59489478>

Same with these additional identifiers. I didn't add those identifiers, someone else did, and so now this item links out to even more information in these different databases.

So that's just the Wikidata item itself. **[next slide]**

Opportunities in Wikidata

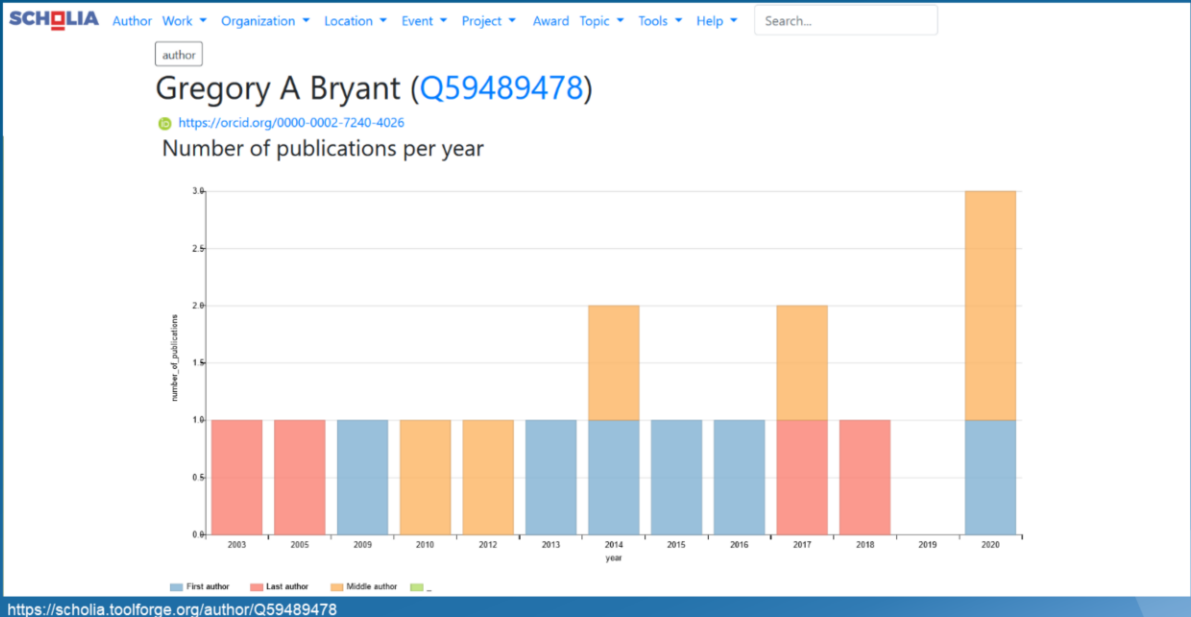
The screenshot shows the Scholia interface for the Wikidata entity Gregory A Bryant (Q59489478). The page includes a search bar, a list of publications, and a table of publication details. The table has columns for Date, Work, Type, Pages, Venue, and Authors.

Date	Work	Type	Pages	Venue	Authors
2020-10-07	Recognizing affiliation in colughter and cospeech	scholarly article		Royal Society Open Science	Riccardo Fusaroli , Christine S Wang , Gregory A Bryant
2020-09-16	Body synchrony in triadic interaction	scholarly article		Royal Society Open Science	Rick Dale , Gregory A Bryant
2020-08-26	Origins of music in credible signaling	scholarly article		Behavioral and Brain Sciences	Max M Krasnow , Edward Hagen , Samuel A Mehr , Gregory A Bryant
2018-05-02	General trust impedes perception of self-reported primary psychopathy in thin slices of social interaction.	scholarly article		PLOS ONE	Joseph H. Manson , Gregory A Bryant
2017-06-01	Three-month-old human infants use vocal cues of body size.	scholarly article		Proceedings of the Royal Society B	David Pietraszewski , Karen Wynn , Gregory A Bryant

<https://scholia.toolforge.org/author/Q59489478>

Now I want to introduce Scholia, which is kind of a visualization interface that runs on top of Wikidata, and can visualize a lot of data relevant to academic scholarship. So if I just search Gregory A. Bryant again here, suddenly I can see connections to a whole host of information that was not in his item, but are linked to his item. Information such as scholarly articles that he has written and metadata about those articles. **[next slide]**

Opportunities in Wikidata



I can also see his publishing trend over the years, and also whether he was first or corresponding author. **[next slide]**

Opportunities in Wikidata

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

author

Gregory A Bryant (Q59489478)

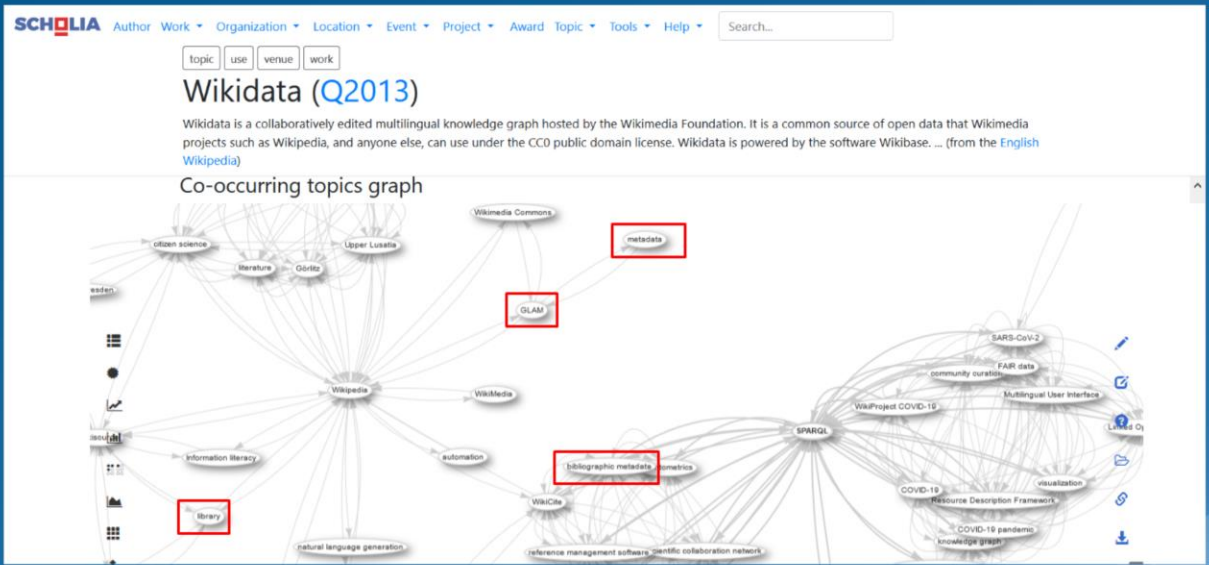
<https://orcid.org/0000-0002-7240-4026>

Co-author graph

<https://scholia.toolforge.org/author/Q59489478>

Not only that, but I can also see who he has co-authored with, according to Wikidata. All of this information has been contributed by many parties, and we can all collectively benefit from each of us putting different pieces of data into Wikidata. I think this kind of analysis of data would have been much more difficult if we were to try and collect all of this information ourselves. **[next slide]**

Opportunities in Wikidata



<https://scholia.toolforge.org/topic/Q2013>

Scholia can also provide visualizations regarding topics, not just people. I just did a search on Wikidata itself to see what I can learn about it as a topic, and here's a visualization of co-occurring topics with Wikidata. And you can see that the main nodes are Wikipedia and SPARQL, very related topics to Wikidata, but you can also see GLAM and bibliographic metadata, which speaks to how involved the GLAM community is with Wikidata. **[next slide]**

Opportunities in Wikidata

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

topic use venue work

Wikidata (Q2013)

Wikidata is a collaboratively edited multilingual knowledge graph hosted by the Wikimedia Foundation. It is a common source of open data that Wikimedia projects such as Wikipedia, and anyone else, can use under the CC0 public domain license. Wikidata is powered by the software Wikibase. ... (from the [English Wikipedia](#))

Co-occurring topics graph

<https://scholia.toolforge.org/topic/Q2013>

COVID has also shown up, which leads me to believe that Wikidata is, in a very timely way, playing a part in our understanding and knowledgebase of COVID as well.

[next slide]

Opportunities in Wikidata

Wikidata:WikiProject sum of all paintings

Shortcuts: [WD:WPSOAP](#), [WD:SOAP](#)

[Translate this page.](#) This page contains changes. Please contact a translation admin to mark them for translation.

Other languages: [Bahasa Indonesia](#) • [Deutsch](#) • **[English](#)** • [Nederlands](#) • [Türkçe](#) • [català](#) • [dansk](#) • [español](#) • [euskara](#) • [français](#) • [hrvatski](#) • [italiano](#) • [polski](#) • [suomi](#) • [svenska](#) • [русский](#) • [українська](#) • [հայերեն](#) • [اردو](#) • [العربية](#) • [日本語](#)

The **WikiProject sum of all paintings** is a WikiProject to have a Wikidata item for every notable painting.

Contents [\[hide\]](#)

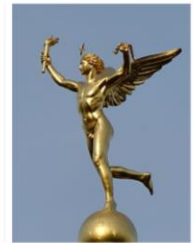
- 1 Background
- 2 Scope
 - 2.1 Tips
- 3 Current overview
 - 3.1 Creator overview
 - 3.2 Collection overview
 - 3.3 Genre overview
 - 3.4 Main subject overview
- 4 Work
 - 4.1 Help bring these lists down to zero
 - 4.1.1 Items without claims
 - 4.2 Add collections that need to be worked on by country

https://www.wikidata.org/wiki/Wikidata:WikiProject_sum_of_all_paintings

Le Génie de la Liberté (Q15726563)

sculpture by Auguste Dumont

Location properties	
Instance of	Statue
Maps	
Location	
Other properties	
genre	public art in public space mobile work of art that has as its primary subject the unclothed human body allegorical sculpture type of sculpture
creator	Augustin-Alexandre Dumont French sculptor
depicts	26 items Show items
made from material	bronze metal alloy



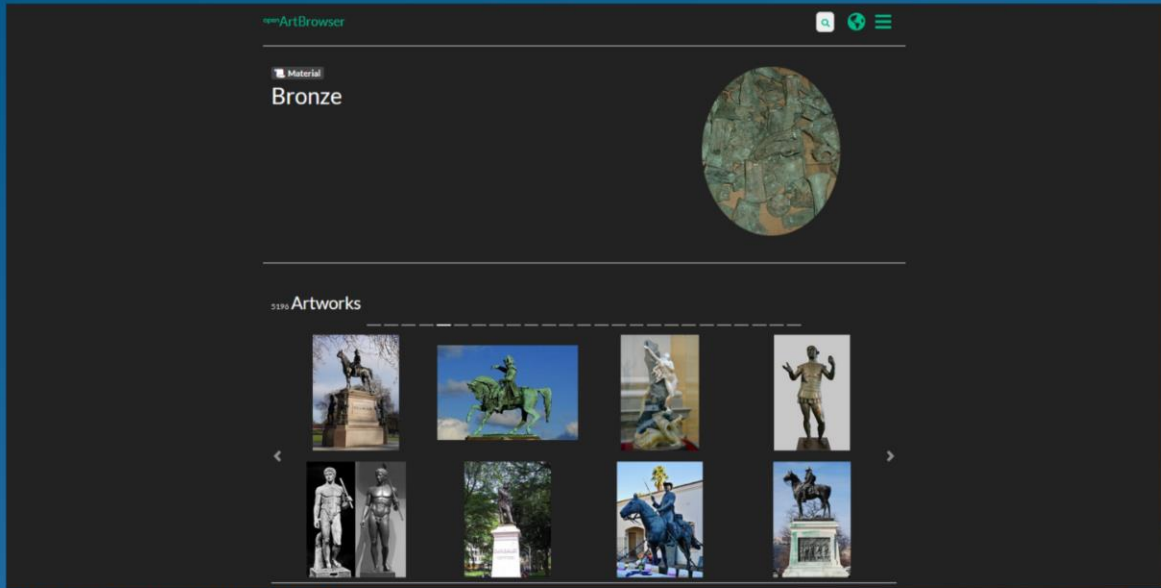
External sources
[Google Knowledge Graph](#) [/g/123pws85n](#)

External identifiers

<https://reasonator.toolforge.org/?q=15726563>

One last example in the realm of the arts and with potential for digital humanities is a project called the Sum of All Paintings, and its goal is just to have a Wikidata item for every notable painting in the world through crowd-sourcing efforts and volunteer contributions. And by adding details about paintings in a structured way, such as listing materials it's made from, like bronze, you can then build another visualization tool on top of this dataset, such as [\[next slide\]](#)

Opportunities in Wikidata



the OpenArtBrowser, and you can ask questions such as what artwork in the world is made of bronze, or even **[next slide]**

Opportunities in Wikidata

House cat

Domesticated feline

More ▾



232 Artworks



<https://openartbrowser.org/en/motif/Q146>

what artwork around the world has cats in it. **[next slide]**

Where Are We Now

327

Items Created

524

Items Edited

8.88K

Total Edits

10

Editors

1.78K[®]

References
Added

Last statistics update

UCLA PCC Wikidata Pilot Program

We will be working on several Wikidata projects that support discoverability of UCLA held/subscribed resources and the work of identity management for UCLA faculty. Our work will include creating and enhancing items for UCLA's Special and Digital Collections; adding non-Roman scripts to Wikidata items; creating/editing items on UCLA faculty and units; supporting cataloging workflow including creating Wikidata items that enhance brief bibs in the library catalog and MARC authority

Details

Facilitators: [Parobis](#)

Institution: University of California, Los Angeles

Location: Los Angeles

Website: [https://wikidata.wmflabs.org/wiki/Wikidata:UCLA_PCC_Pilot_Program](#)

Visited March 22, 2021:

https://outreachdashboard.wmflabs.org/courses/University_of_California_Los_Angeles/UCLA_PCC_Wikidata_Pilot_Program

142

UCLA FACULTY

107

UCLA BUILDINGS

51

UCLA ACADEMIC UNITS

42

ITEMS ENHANCED WITH
NON-ROMAN LANGUAGES

38

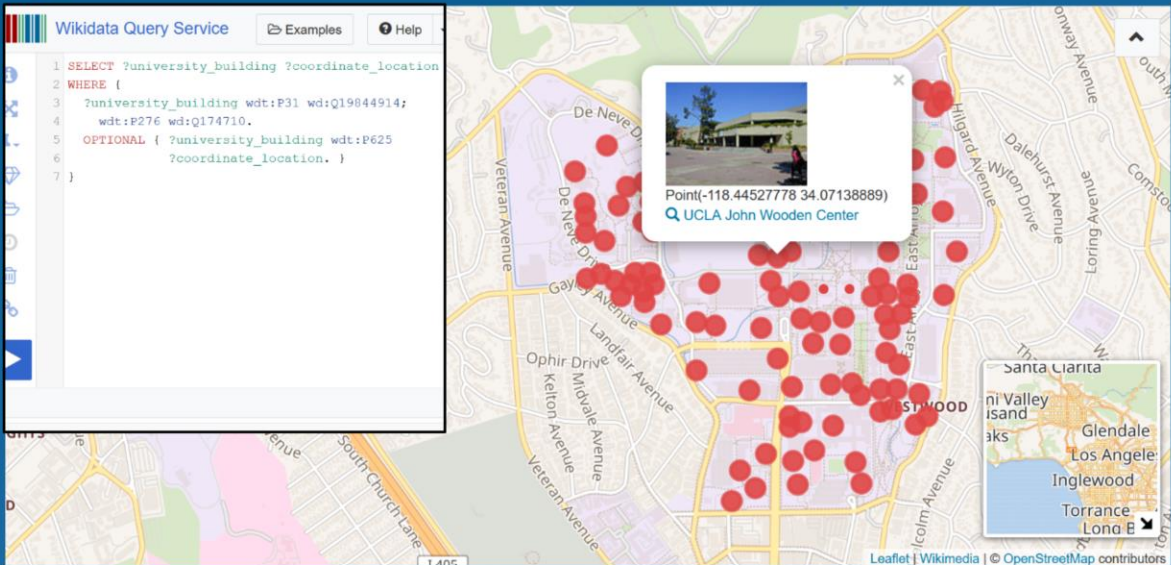
SPECIAL COLLECTIONS

18

UCLA OA JOURNALS

So to quickly sum up the Pilot team's progress so far, here are our current stats for the pilot. To date, we've created upwards of 300 items and edited even more, and you can see the work we've done with items for faculty, for academic units and buildings, for enhancing items with non-Roman languages, including Persian, Arabic, and Armenian, and for special collections, and UCLA open access journals. **[next slide]**

Opportunities in Wikidata



And while the art browser uses data that we aren't working on, here's a quick visualization of data we've entered into Wikidata through this pilot, specifically our team working on UCLA buildings. Our project team added building items with coordinate data, I ran a query using Wikidata's query service, and without needing any kind of special software, Wikidata allowed us to quickly visualize their work on this map. **next slide**

Join the Wikidata Fun!

“ While creating and editing items for UCLA buildings, I've **learned a lot about their history and their occupants**, which I've been able to share as structured data in Wikidata's low-barrier environment. **This includes highlighting the contributions of local and BIPOC leaders and architects as well as bringing attention to UCLA's sustainability efforts through its green building construction and renovation projects.** ”

- Joseph Andrews, RAMS

Find us at: **Confluence**

Reach us at:

[pccucla-wikidata-pilot\[at\]googlegroups.com](mailto:pccucla-wikidata-pilot[at]googlegroups.com)

Thank you!

As we move to questions, we just wanted to leave you with a testimonial from one of our team members, Joseph Andrews, and how his work with UCLA buildings has helped him quote “learn a lot about their history and their occupants”...and “this includes highlighting the contributions of local and BIPOC leaders and architects as well as bringing attention to UCLA’s sustainability efforts through its green building construction and renovation projects.”