



WIKIDATA FOR ACADEMICS

LARGE-SCALE, INTERDISCIPLINARY, MULTILINGUAL, OPEN DATA

OUTLINE

WHY SHOULD YOU BE INTERESTED IN WIKIDATA?



- A brief introduction to the largest open knowledgebase
 - Why you might draw data from it
 - Why you might deposit data into it

How to put data in



- Different ways to editing and create items
 - One at a time (direct manual editing)
 - Multiple at once (tool assisted batch editing)

How read data from



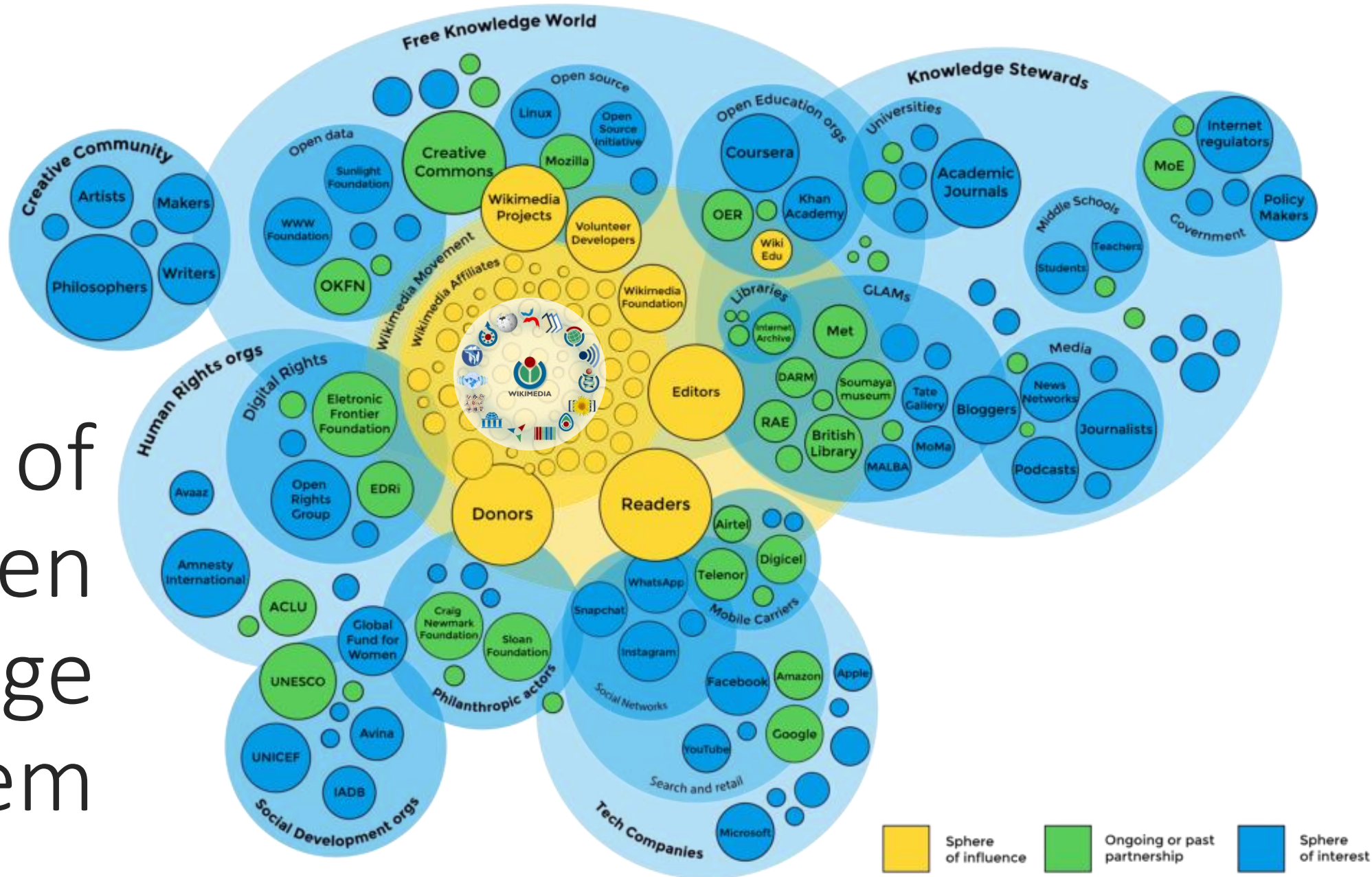
- Query information back out of Wikidata
 - SPARQL tool (and language)
 - A wide range of custom examples

HELP, COMMUNITY AND RESOURCES

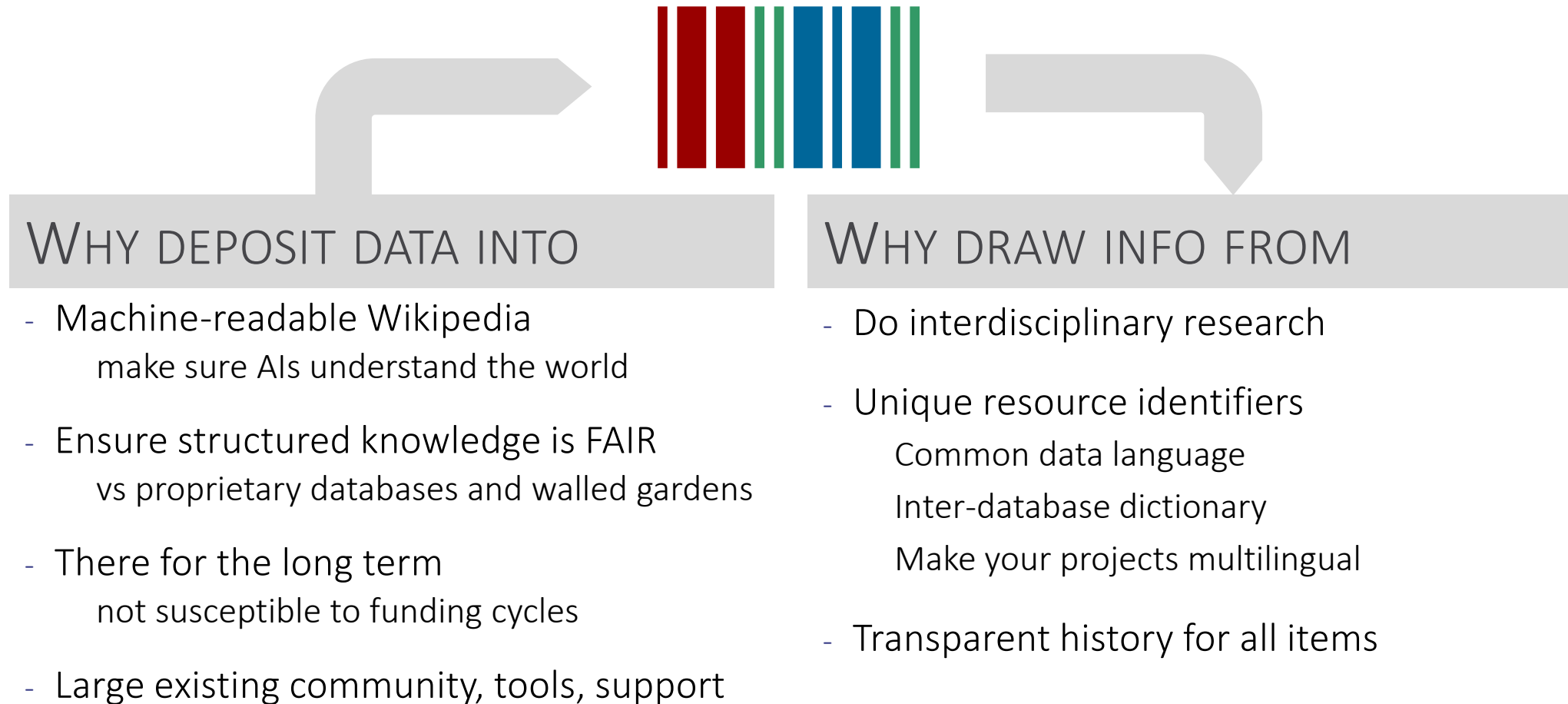


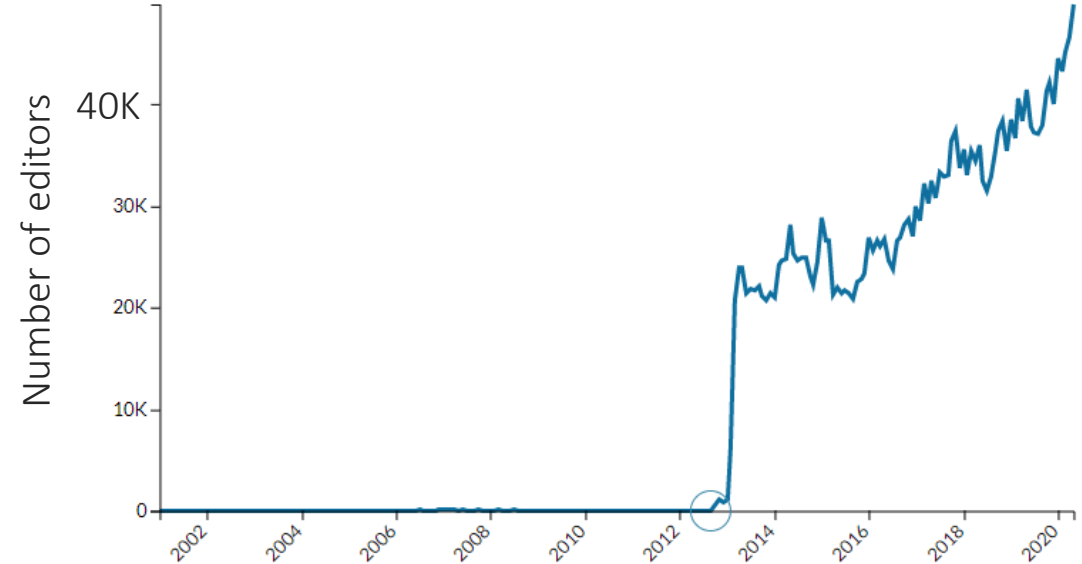
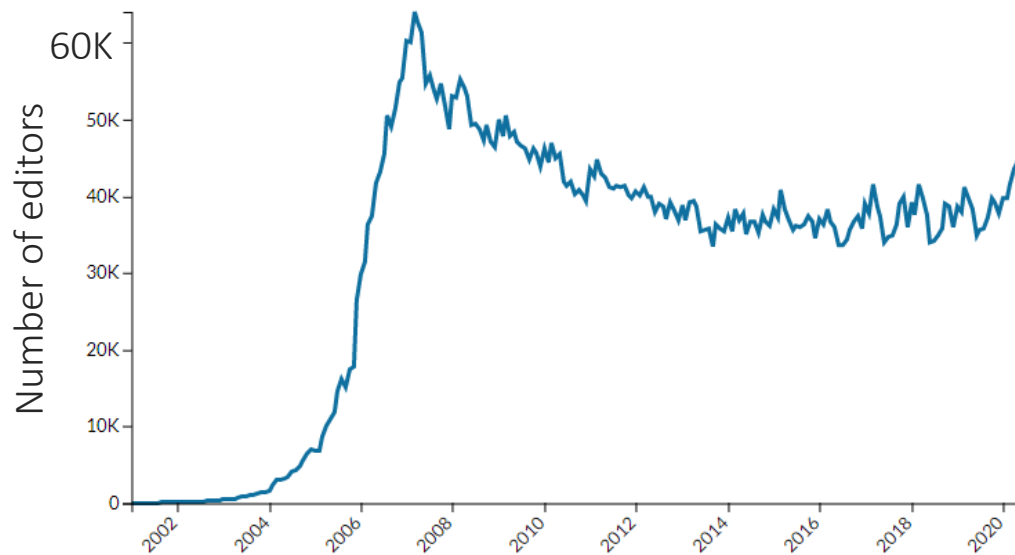
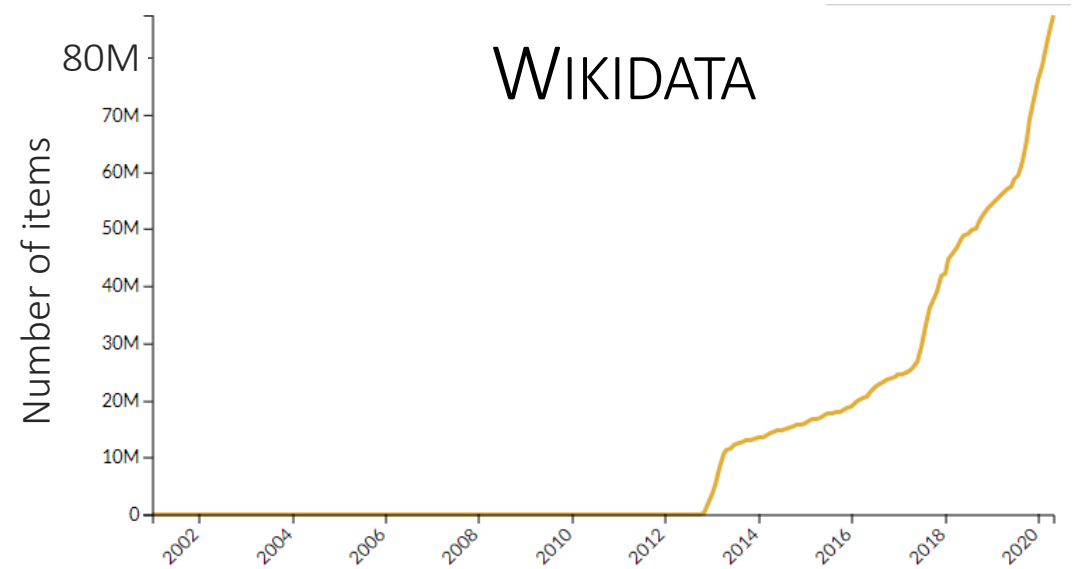
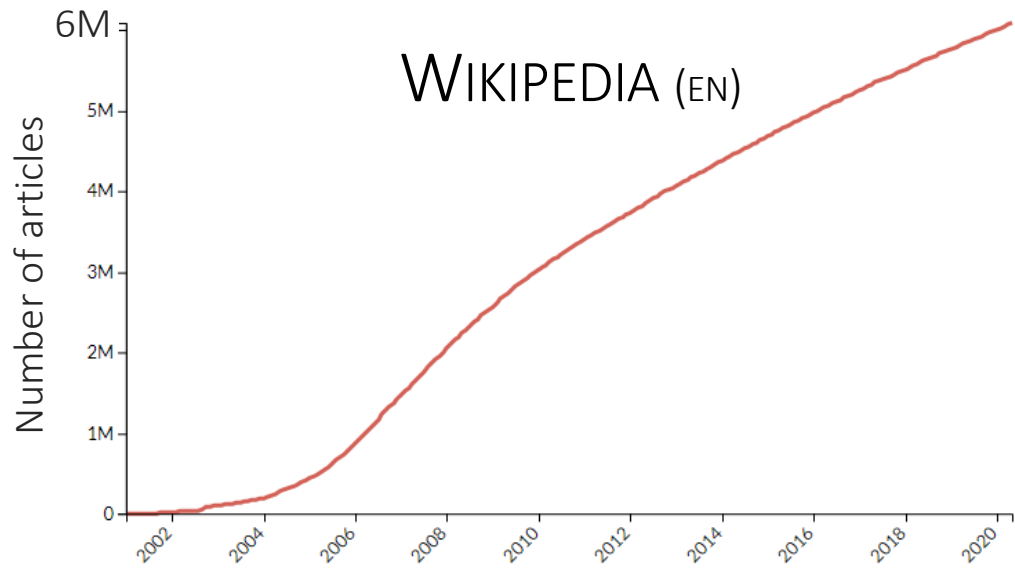
- The hidden world behind Wikidata

Part of wider open knowledge ecosystem



WHY INTERACT WITH WIKIDATA





EXAMPLE ITEMS AND PROPERTIES

Concepts

E.g. locations, events, diseases, genes, languages, processes

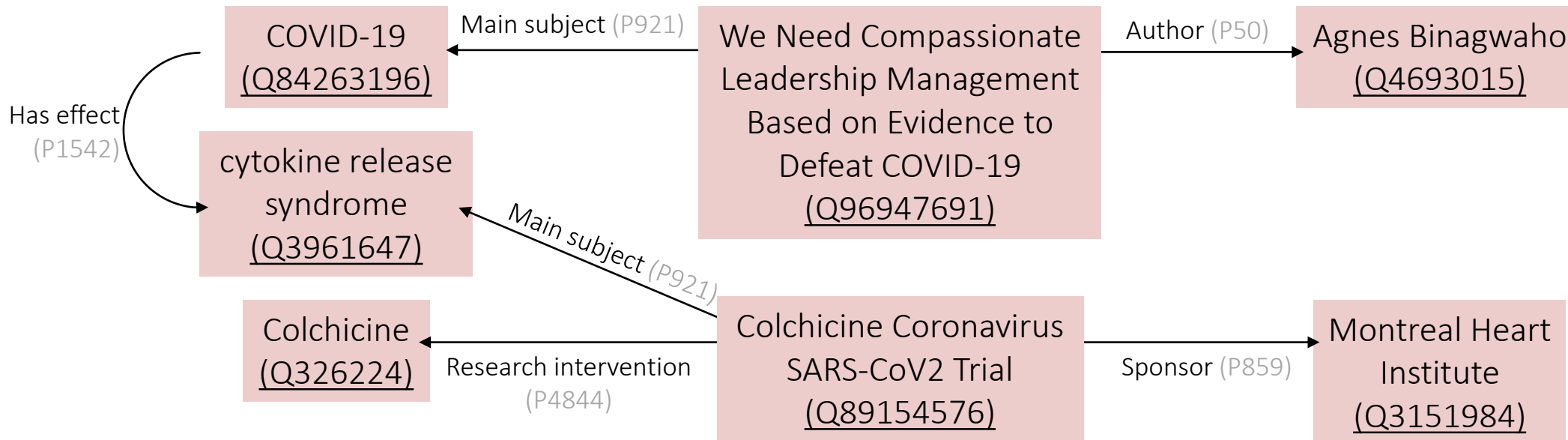
Research outputs

E.g. publications, journals, publishers, clinical trials

People & orgs

E.g. Researchers, authors, universities, institutes

... etc



EXAMPLE ITEMS AND PROPERTIES

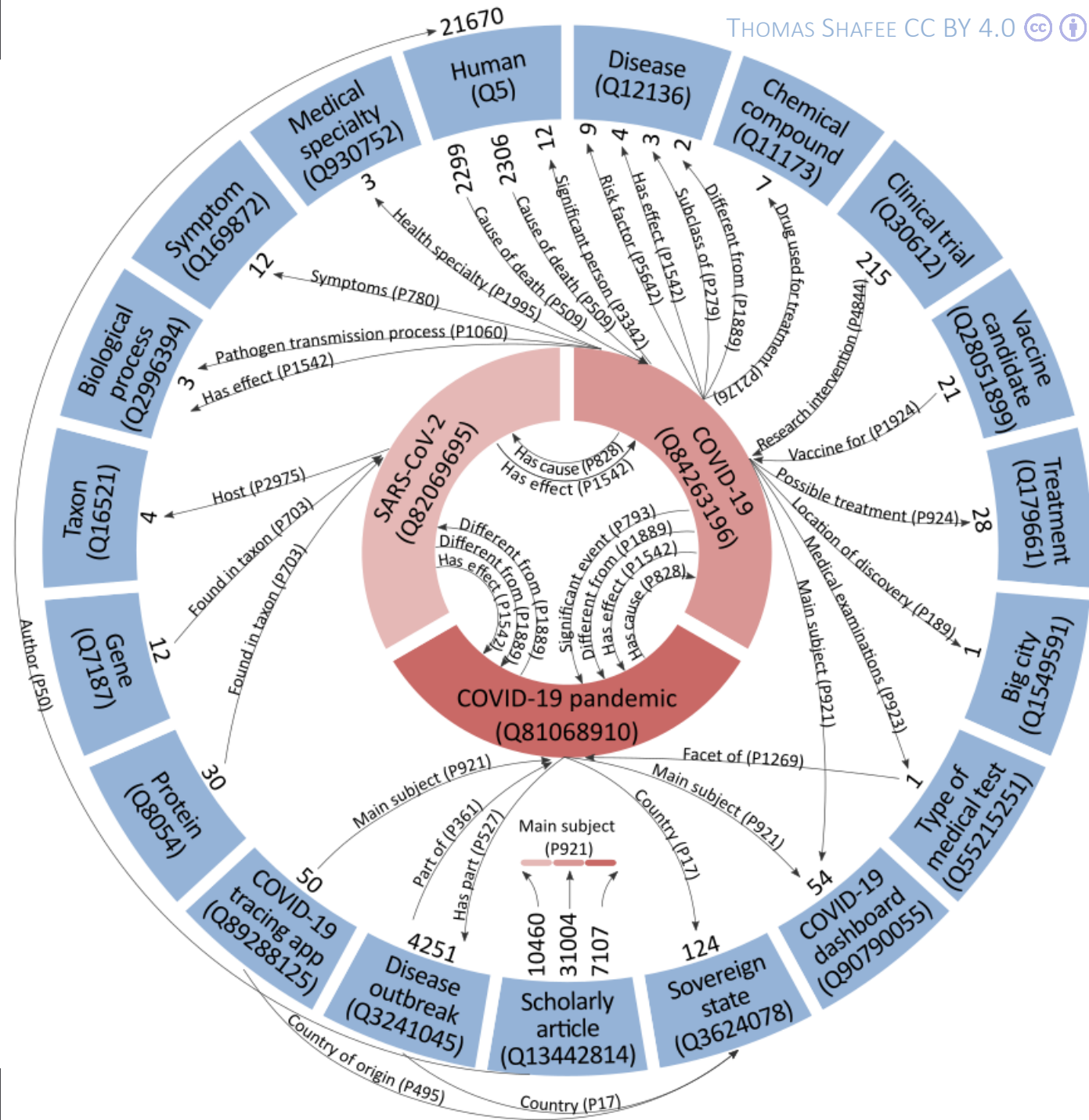
Granular disambiguation of concepts

All of Wikidata tied into a single network

90M items

>1.1B connections

Links data from wide variety of disciplines



OUTLINE

WHY SHOULD YOU BE INTERESTED IN WIKIDATA?



- A brief introduction to the largest open knowledgebase
 - Why you might draw data from it
 - Why you might deposit data into it

How to put data in



- Different ways to editing and create items
 - One at a time (direct manual editing)
 - Multiple at once (tool assisted batch editing)

How read data from



- Query information back out of Wikidata
 - SPARQL tool (and language)
 - A wide range of custom examples

HELP, COMMUNITY AND RESOURCES



- The hidden world behind Wikidata

LETS MANUALLY EDIT AN ITEM

publication date  1998  edit


▼ 1 reference  copy



reference URL	https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.246194
---------------	---


+ add reference

+ add value

PhD theses that don't list a publication date

1. Go to <https://w.wiki/ekc> and press 
2. Find the thesis's URL for the thesis
3. At the bottom of the Wikidata page, **+ add statement**
Add "publication date" based on what's listed at that URL
4. In that statement, **+ add reference**
As the "reference URL", use the thesis's URL

main subject  lipid  edit


▼ 1 reference  copy

based on heuristic	inferred from title
--------------------	---------------------

+ add reference

+ add value

PhD theses that don't list a main topic

1. Go to <https://w.wiki/ekg> and press 
2. Find the thesis's URL for the thesis
3. At the bottom of the Wikidata page, **+ add statement**
Add the main subject(s) based on the keywords or title
4. In that statement, **+ add reference**
As the "reference URL", add the thesis's URL
(or) As the reference, add "based on heuristic" "inferred from title"

LETS BULK EDIT SEVERAL ITEMS AT ONCE!

- Two main tools: QuickStatements (<https://quickstatements.toolforge.org>)
OpenRefine (<http://openrefine.org>)

- Import tabular data into Wikidata

- Example:

Create a new item
It's an instance of a human
Its label in English is...
Its sex/gender is male
Its place of birth is Rome
Its place of death is Rome

	A	B	C	D	E	F	G
1	CREATE						
2	LAST	P31	Q5	S143	Q24731821		
3	LAST	Len	"Giovanni Francesco Rossi"				
4	LAST	P21	Q6581097	S143	Q24731821	S813	+2017-10-04T00:00:00Z/11
5	LAST	P19	Q220	S143	Q24731821	S813	+2017-10-04T00:00:00Z/11
6	LAST	P20	Q220	S143	Q24731821	S813	+2017-10-04T00:00:00Z/11

Item to edit
Statements to add
Supporting reference

LETS BULK EDIT SEVERAL ITEMS AT ONCE!

	A	B	C	D	E
1	thesis	pub date is	date	source URL is	URL
119	Q62099128	P577	+2015-01-01T00:00:00Z	S854	"https://eprints.soton.ac.uk/37791
120	Q59623433	P577	+2015-01-01T00:00:00Z	S854	"http://etheses.whiterose.ac.uk/1
121	Q63457013	P577	+2015-01-01T00:00:00Z	S854	"https://chesterrep.openrepositor
122	Q78208913	P577	+2015-01-01T00:00:00Z	S854	"https://www.escholar.mancheste
123	Q90517973	P577	+2015-01-01T00:00:00Z	S854	"https://discovery.ucl.ac.uk/id/ep
124	Q98417405	P577	+2015-01-01T00:00:00Z	S854	"http://etheses.whiterose.ac.uk/1
125	Q85502175	P577	+2014-03-18T00:00:00Z	S854	"https://ore.exeter.ac.uk/repositc
126	Q28865408	P577	+2014-01-01T00:00:00Z	S854	"https://ora.ox.ac.uk/objects/uuid
127	Q28865433	P577	+2014-01-01T00:00:00Z	S854	"https://ora.ox.ac.uk/objects/uuid
128	Q28865612	P577	+2014-01-01T00:00:00Z	S854	"https://ora.ox.ac.uk/objects/uuid
129	Q28866437	P577	+2014-01-01T00:00:00Z	S854	"https://ora.ox.ac.uk/objects/uuid

PhD theses that don't list a publication date

1. Go to Docs for the list <https://tinyurl.com/QStable>
To see where this list came from, see <https://w.wiki/f25>
2. Go to <https://quickstatements.toolforge.org>
Paste across a few random rows from the table
"Import V1 commands" (if tab separated)
(or) "Import CSV commands" (if comma separated)
3. Hit "Run" to add the information to Wikidata



1	init	Searches for scharms and gluinos with the ATLAS Detector [Q28865931]	ADD	Statement	publication date [P577] : 2016-01
2	init	Searches for scharms and gluinos with the ATLAS Detector [Q28865931]	ADD	Sources to	publication date [P577] : 2016-01 reference URL [P854] : "https://ora.ox.ac.uk/objects/uuid:95c23d4c-d193-41ff-a284-2309d0a1ab99"

OUTLINE

WHY SHOULD YOU BE INTERESTED IN WIKIDATA?



- A brief introduction to the largest open knowledgebase
 - Why you might draw data from it
 - Why you might deposit data into it

How to put data in



- Different ways to editing and create items
 - One at a time (direct manual editing)
 - Multiple at once (tool assisted batch editing)

How read data from



- Query information back out of Wikidata
 - SPARQL tool (and language)
 - A wide range of custom examples


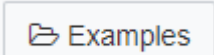
HELP, COMMUNITY AND RESOURCES



- The hidden world behind Wikidata

LETS QUERY WIKIDATA'S CONTENTS

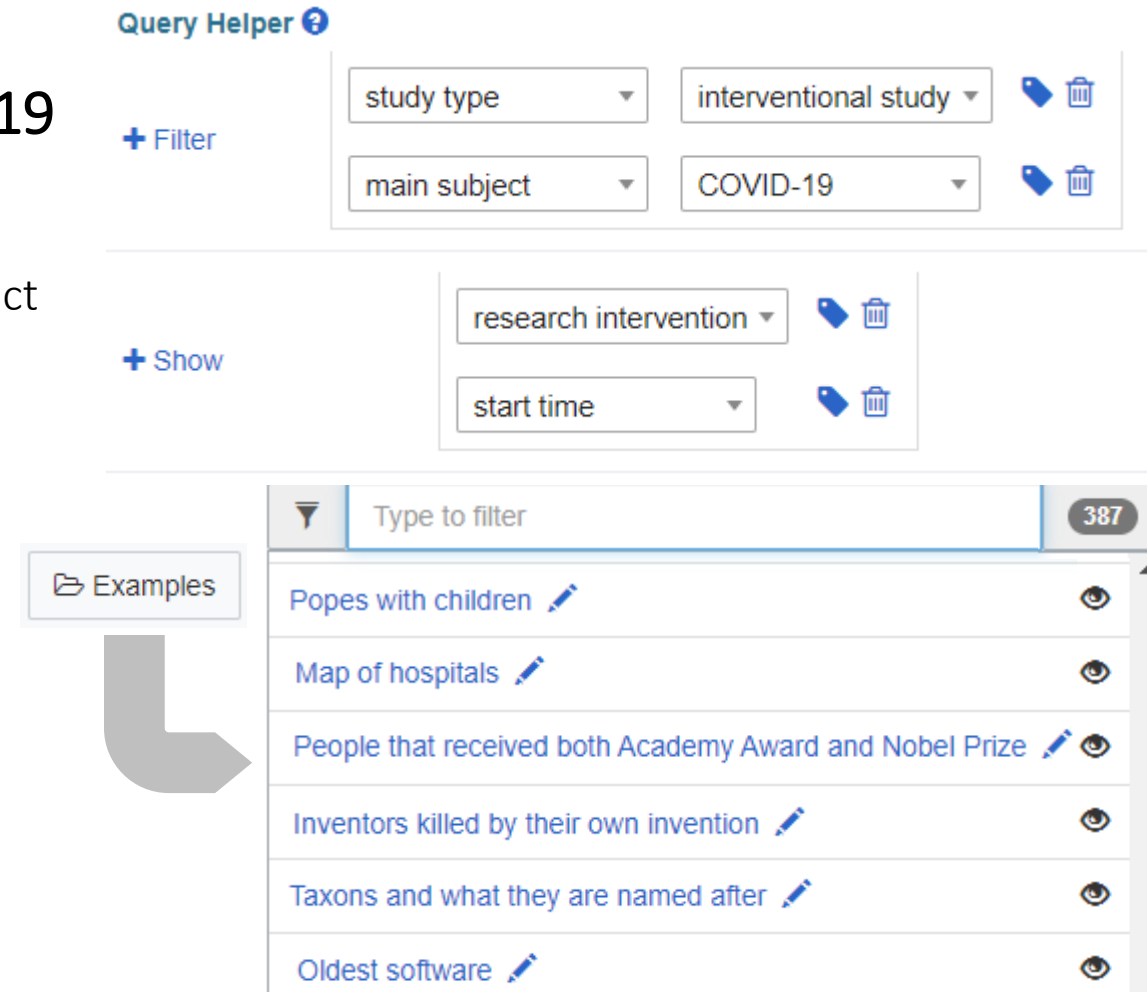
Find timeline of intervention studies on COVID-19

1. Go to query.wikidata.org
2. Filter for items that are interventional studies study type
3. Filter for items that are about COVID-19 as their main subject
4. Show their research intervention
5. Show their start time
6. Press 
7. Go to the  and find something interesting

Based on SQL language

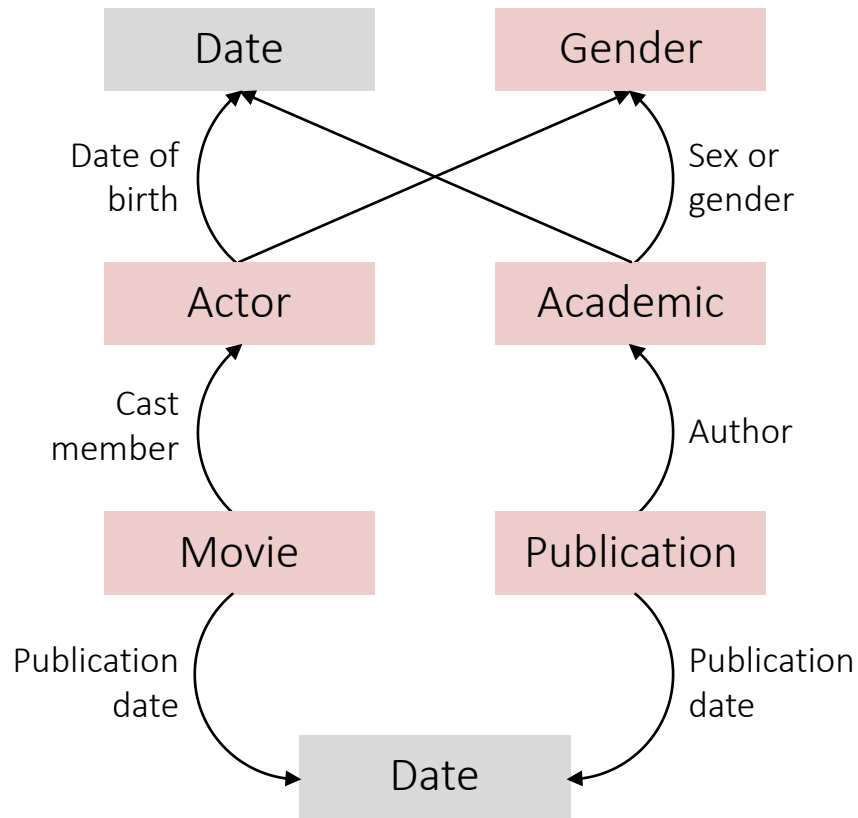
Isn't as scary as you think

The code is human-readable(ish)

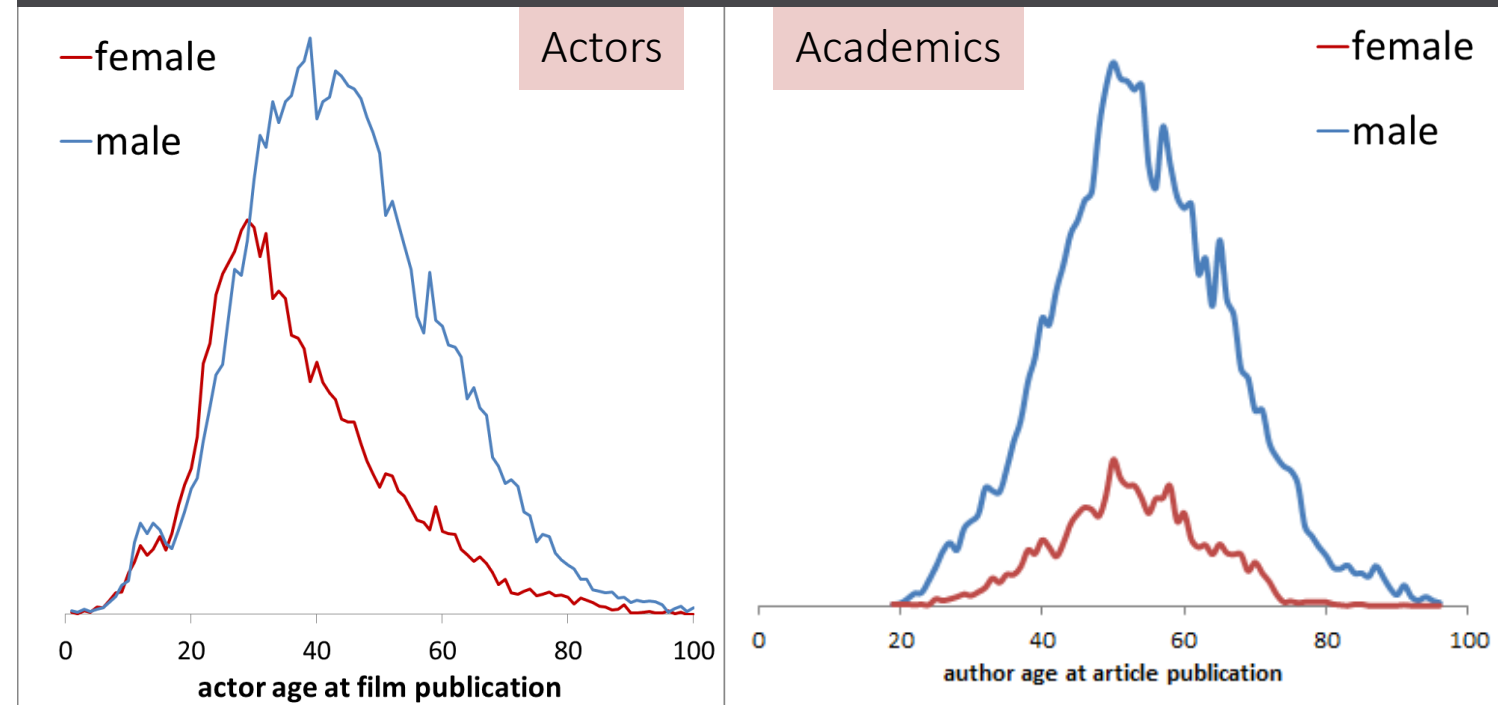


The screenshot shows the Wikidata Query Helper interface. At the top, there is a "Query Helper" section with a question mark icon. Below it, there are two filter groups. The first group, labeled "+ Filter", contains two dropdown menus: "study type" set to "interventional study" and "main subject" set to "COVID-19". The second group, labeled "+ Show", contains two dropdown menus: "research intervention" and "start time". Below the filters, there is a search bar labeled "Type to filter" with a "387" result count. A list of results is displayed, including "Popes with children", "Map of hospitals", "People that received both Academy Award and Nobel Prize", "Inventors killed by their own invention", "Taxons and what they are named after", and "Oldest software". A grey arrow points from the "Examples" button in the text above to the "Examples" button in the interface.

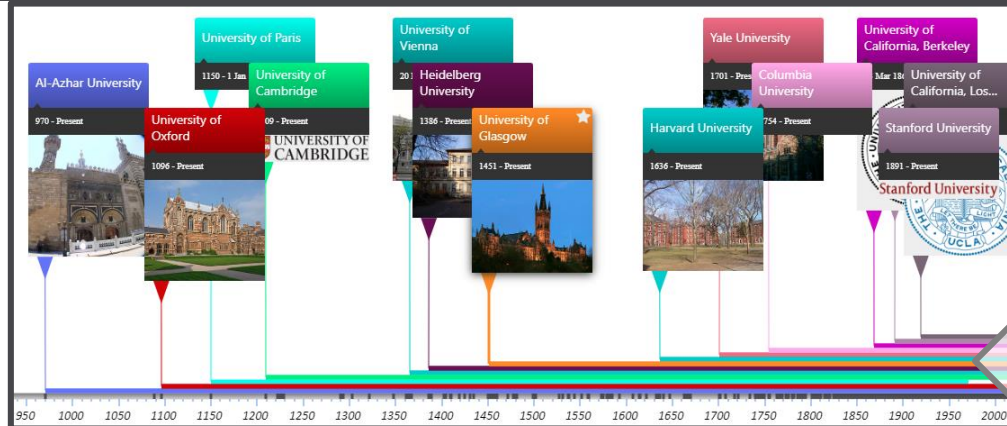
EXAMPLE QUERY RESULTS



Gendered age bias: actors vs academics



SOME INTERESTING USAGE EXAMPLES FOR INSPIRATION



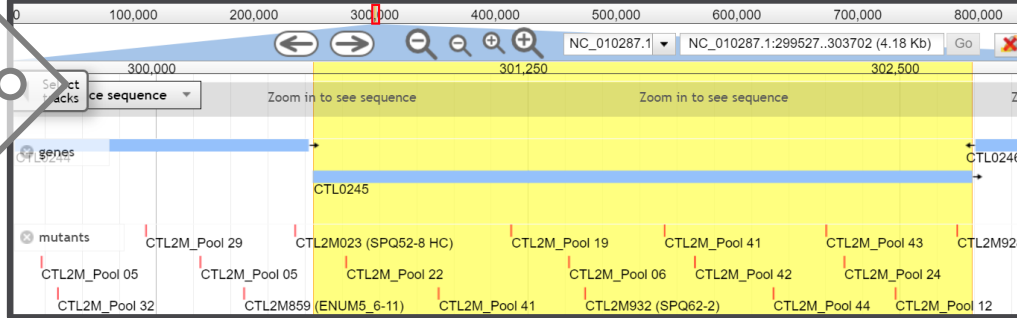
Events timelines
histropedia.com/timeline

Map of witches in Scotland
witches.is.ed.ac.uk



Scholia dashboard
scholia.toolforge.org

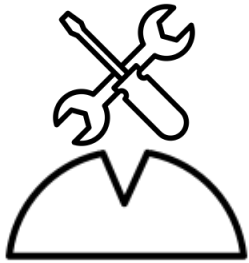
Chlamydia genome browser
chlambase.org



SO MANY TOOLS!

General tools

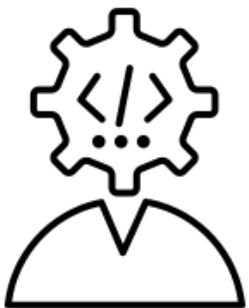
wikidata.org/wiki/Wikidata:Tools



- Mix-n-match (tools.wmflabs.org/mix-n-match)
- WD games (wikidata-game.toolforge.org/distributed)
- Author disambiguator (author-disambiguator.toolforge.org)
- Zotero importer (wikidata.org/wiki/Wikidata:Zotero)

Programmer tools

wikidata.org/wiki/Wikidata:Tools/For_programmers



- API for WikiData (wikidata.org/w/api.php)
for QuickStatements (quickstatements.toolforge.org/api.php)
- Python package (github.com/dahlia/wikidata)
- R package (github.com/TS404/WikidataR)

OUTLINE

WHY SHOULD YOU BE INTERESTED IN WIKIDATA?



- A brief introduction to the largest open knowledgebase
 - Why you might draw data from it
 - Why you might deposit data into it

How to put data in



- Different ways to editing and create items
 - One at a time (direct manual editing)
 - Multiple at once (tool assisted batch editing)

How read data from



- Query information back out of Wikidata
 - SPARQL tool (and language)
 - A wide range of custom examples

HELP, COMMUNITY AND RESOURCES



- The hidden world behind Wikidata

WHERE TO GET HELP?

GENERAL INFORMATION

- Interactive tours

wikidata.org/wiki/Wikidata:Tours

- Academic introductory publications

1. Turki, Houcemeddine, et al. "**Wikidata: A large-scale collaborative ontological medical database.**" *Journal of biomedical informatics* 99 (2019): 103292. [doi: 10.1016/j.jbi.2019.103292](https://doi.org/10.1016/j.jbi.2019.103292)
2. Allison-Cassin, et al. "**Wikidata: a platform for your library's linked open data.**" *Code4Lib Journal* 40 (2018). journal.code4lib.org/articles/13424

ONE-TO-ONE

- General help and discussion

wikidata.org/wiki/Wikidata:Project_chat

- Requesting SPARQL queries

wikidata.org/wiki/Wikidata:Request_a_query

- Special interest Wikiprojects

Examples: Chemistry	Philosophy
Source Metadata	Covid19
Open Access	Ontology

List: (wikidata.org/wiki/Wikidata:WikiProjects)

Contact

Email	Thomas.Shafee@gmail.com
Google Scholar	Thomas Shafee
ResearchGate	Thomas Shafee
LinkedIn	Thomas Shafee
Username	Search [[user:tshafee]]

Journals

<i>WikiJournal of Medicine</i>	(WikiJMed.org)
<i>WikiJournal of Science</i>	(WikiJSci.org)
<i>WikiJournal of Humanities</i>	(WikiJHum.org)
<i>PLOS (TopicPagesWiki.plos.org)</i>	

Shafee, T; Mietchen, D; Su, A. (2017).
 “Academics can help shape Wikipedia”
Science. 357 (6351): 557–558.
 doi: 10.1126/science.aao0462

Turki, Houcemeddine, et al. "Wikidata: A large-scale collaborative ontological medical database" *Journal of biomedical informatics* 99 (2019): 103292.
 doi: 10.1016/j.jbi.2019.103292

Allison-Cassin, et al. "Wikidata: a platform for your library’s linked open data" *Code4Lib Journal* 40 (2018).
 journal.code4lib.org/articles/13424

[END]
