

2022 ASIS&T Webinar Series

September 15, 2022

*“Querying Wikidata:
All the Knowledge in the World (DCMI)”*

**Daniel Garijo,
Ontology Engineering Group,
Universidad Politécnica de Madrid, Spain**

Heavily based on the K-CAP 2019 Tutorial “Linking,
Extending, Exploiting and Enhancing Tabular Data with
Wikidata” (with **Pedro Szekely**)

 @dgarijov

webinars@asist.org

asis&t

Association for Information Science and Technology

ASIRIKA

M

A brief introduction to **Wikidata**

1. What is Wikidata?
2. The Wikidata Data model
3. Organizing data in Wikidata (Wikidata schema)
4. Querying Wikidata with SPARQL
5. Summary



Goals of the Wikidata Project

WIKIDATA

Wikidata (<https://wikidata.org/>) is a free, collaborative, multilingual, secondary database, collecting structured data to provide support for Wikipedia, Wikimedia Commons, the other wikis of the [Wikimedia movement](#), and to anyone in the world.

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

Free

The data in Wikidata is published under the Creative Commons Public Domain Dedication 1.0.

Collaborative

Data is entered and maintained by Wikidata editors, who decide on the rules of content creation and management. Automated bots also enter data into Wikidata.

Multilingual

Editing, consuming, browsing, and reusing the data is fully multilingual. Data entered in any language is immediately available in all other languages.

A secondary database

Wikidata records not just statements, but also their sources, and connections to other databases.

Collecting structured data

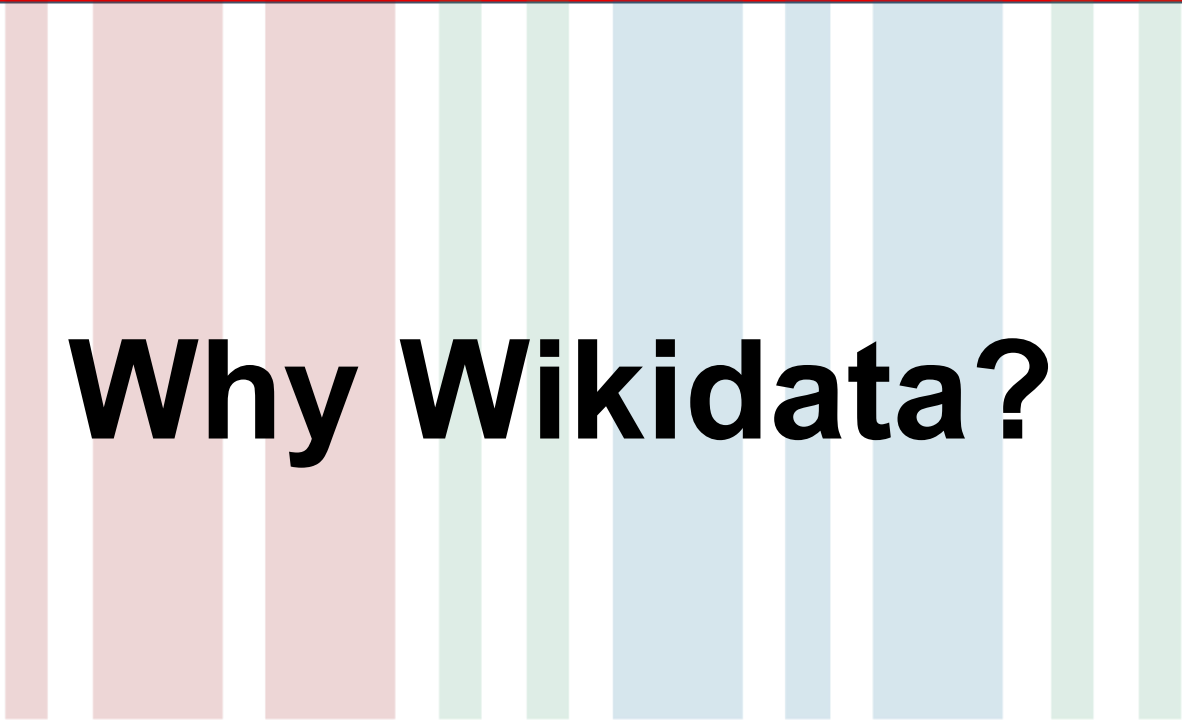
Imposing a high degree of structured organization allows for easy reuse of data by Wikimedia projects and third parties, and enables computers to process and “understand” it.

Support for Wikimedia wikis

Wikidata assists Wikipedia with more easily maintainable information boxes and links to other languages.

Anyone in the world

Anyone can use Wikidata for any number of different ways by using its application programming interface, the SPARQL endpoint and the JSON and RDF dumps.



Why Wikidata?

WIKIDATA



Marina del Rey (Q988140)

census designated place

Good data

Statements		
Own statements		From related entities
area	3.767749 square kilometre	>
population	8866 point in time : 2010	>
elevation above sea level	0 metre	>
located in time zone	Pacific Time Zone (time zone in North America)	>
	UTC-07:00 (Identifier for a time offset from UTC of -7) valid in period : daylight saving time (practice of advancing clocks so that evenings have more daylight and mornings have less)	>
	UTC-08:00 (Identifier for a time offset from UTC of -8) valid in period : standard time (synchronization of clocks within a geographical area or region)	>
coordinate location	(33°58'45" N , 118°27'10" W)	>
located in the administrative territorial entity	Los Angeles County (county in California, United States of America)	>
country	United States of America (federal republic in North America)	>

qualifiers

Media		
page banner	Marina del Rey banner.jpg	>
locator map image	LA County Incorporated Areas Marina del Rey highlighted.svg	>
Commons category	Marina del Rey, California	>
image	Marina Del Rey Looking South.jpg	>

Wikimedia Categories and Portals		
Own statements		From related entities
category of associated people	Category:People from Marina del Rey, California (Wikimedia category)	>
category for people born here	Q21000701 (Wikimedia category)	>
category for people who died here	Q21000703 (Wikimedia category)	>
topic's main category	Category:Marina del Rey, California (Wikimedia category)	>
Commons category	Marina del Rey, California	>



Links	
Wikidata page	
Wikipedia article	
Reasonator	
Identifiers	
MusicBrainz area ID	2a24c2...3f0fdd
archINFORM location ID	16300
DMOZ ID	Region...l_Rey/
FIPS 55-3 (locations in the US)	06-45806
WOEID	2445713
Who's on First ID	85928367
Freebase ID	/m/01zhsgk
Library of Congress authority ID	n81081018
VIAF ID	247815515
GND ID	4418951-5
GNIS ID	1852255
GeoNames ID	5370542

external identifiers

LOTS of data:

- > 99M items
- > 6,000 properties
- > 2,000,000 classes
- > 4,000 external id types

Kristabel Doebl-Hickok (Q19757547)

American bicycle racer

Instance of: Kristabel Doebl-Hickok is a(n) [human](#)

Statements		
Own statements		From related entities
member of sports team	Rally Cycling (United States women's cycling team) start time : 2019	>
	Cylance Pro Cycling (United States women's cycling team) start time : 2016 end time : 2018	>
	Team TIBCO (women's cycling team) start time : 2014 end time : 2015	>
sport	cycle sport (competitive physical activity using bicycles)	>
country of citizenship	United States of America (federal republic in North America)	>
place of birth	Marina del Rey (census designated place)	>
date of birth	1989-04-28	>
occupation	sport cyclist (someone who competes in the sport of cycling)	>
sex or gender	female (human gender (use with P21))	>

Media		
image	Krista Doebl-Hickok, Giro Rosa 2016.jpg ↗ media legend : Au Tour d'Italie 2016 (fr)	>



Links	
Wikidata page	
Wikipedia article	
Reasonator	

Identifiers		
ProCyclingStats cyclist ID	179437 ↗	>
Cycling Archives cyclist ID	98235 ↗	>
CQ Ranking female cyclist ID	7931 ↗	>

Kristabel Doebl-Hickok

American bicycle racer

Instance of: Kristabel Doebl-Hickok is a(n) **human**

Statements

Own statements

member of sports team	Rally Cycling (United States w start time : 2019 Cylance Pro Cycling (United start time : 2016 end time : 2018 Team TIBCO (women's cycling start time : 2014 end time : 2015
sport	cycle sport (competitive physi
country of citizenship	United States of America (
place of birth	Marina del Rey (census desig
date of birth	1989-04-28
occupation	sport cyclist (someone who c
sex or gender	female (human gender (use wit

Media

image	Krista Doebl-Hickok, Giro media legend : Au Tour d'Italie 2
--------------	--

About: Kristabel Doebl-Hickok
An Entity of Type: cyclist, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org

Kristabel Doebl-Hickok (born April 28, 1989) is an American professional racing cyclist. She rides for Cylance Pro Cycling.

Property	Value
dbp:abstract	Kristabel Doebl-Hickok (born April 28, 1989) is an American professional racing cyclist. She rides for Cylance Pro Cycling. ^(en)
dbp:birthDate	1989-04-28 ^(estimate) 1989-4-28
dbp:thumbnail	wiki-commons:Special:FilePath/Krista_Doebl-Hickok_Giro_Rosa_2016.jpg?width=300
dbp:wikiPageID	46226373 ^(estimate)
dbp:wikiPageRevisionID	729299224 ^(estimate)
dbp:caption	Kristabel Doebl-Hickok during the Giro Rosa 2016
dbp:role	Rider
dbp:description	American bicycle racer ^(en)
dbp:subject	- dbp:1989_births - dbp:Living_people - dbp:Place_of_birth_missing_fiving_people - dbp:American_female_cyclists
http://purl.org/linguistics/gold/hyphenym	dbp:Cyclist
dbp:type	- dbp:Thing - dbp:Person - dbp:Person - dbp:Agent - dbp:NaturalPerson - dbp:Q211907 - dbp:Q24229398 - dbp:Q5 - dbp:Agent - dbp:Athlete - dbp:Cyclist - dbp:Person
dbp:comment	Kristabel Doebl-Hickok (born April 28, 1989) is an American professional racing cyclist. She rides for Cylance Pro Cycling. ^(en)
dbp:label	Kristabel Doebl-Hickok ^(en)
dbp:alternateNames	- dbp:alias:Kristabel Doebl-Hickok - dbp:alias:Kristabel Doebl-Hickok - dbp:dbpedia:Kristabel Doebl-Hickok - dbp:dbpedia:Kristabel Doebl-Hickok
dbp:wasDerivedFrom	wikipedia:en:Kristabel_Doebl-Hickok?oldid=729299224
dbp:depiction	wiki-commons:Special:FilePath/Krista_Doebl-Hickok_Giro_Rosa_2016.jpg
dbp:gender	female ^(en)
dbp:isPrimaryTopicOf	wikipedia:en:Kristabel_Doebl-Hickok
dbp:isWikiPageRedirectsOf	- dbp:Kristabel_Doebl_Hickok - dbp:Kristabel_Hickok - dbp:Krista_Doebl-Hickok
dbp:isPrimaryTopicOf	wikipedia:en:Kristabel_Doebl-Hickok

no sport team, no identifiers



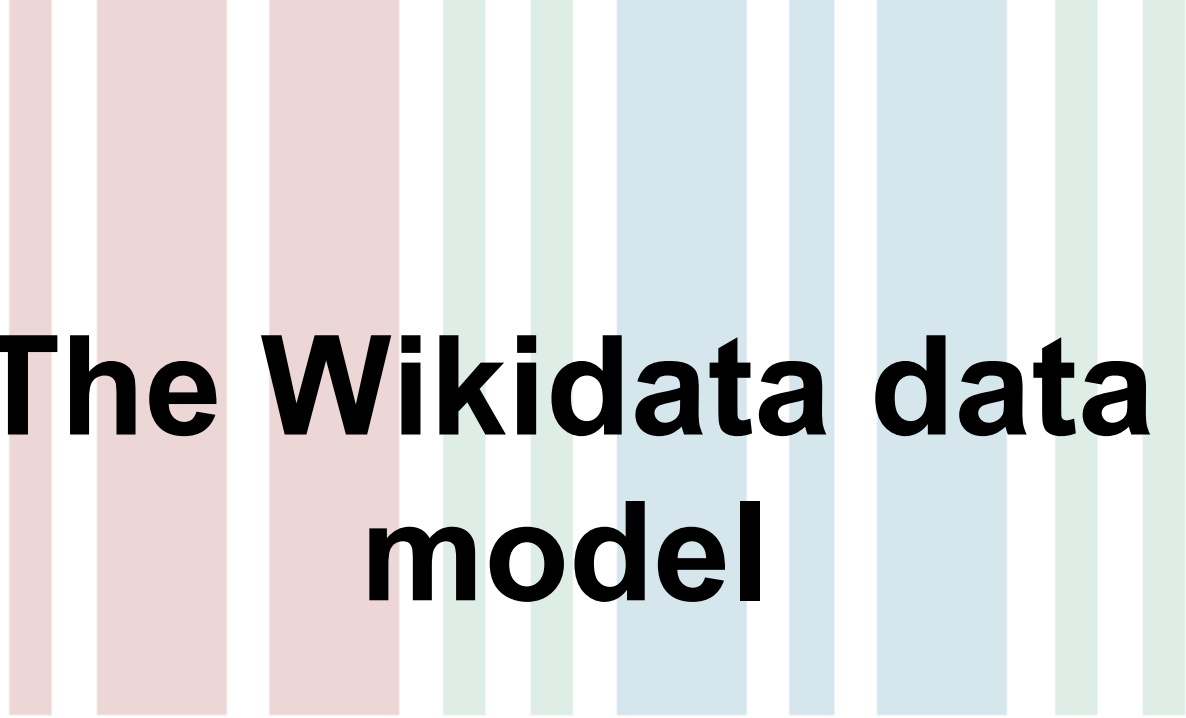
age	
article	
S	
stats	179437 ^(en)
list	98235 ^(en)
st	7931 ^(en)

Language	Label	Description	Also known as
English	Marina del Rey	census designated place	
Spanish	Marina del Rey	lugar de California	
Traditional Chinese	洛杉磯縣 (卡利福尼亞州)	No description defined	瑪麗安德爾灣 瑪麗安德爾灣 (加州)
Chinese	馬里納戴爾雷伊	No description defined	瑪麗安德爾灣
Catalan	Marina del Rey	No description defined	
Cebuano	Marina del Rey	No description defined	
Czech	Marina del Rey	obec v Los Angeles County, Kalifornie	
Danish	Marina del Rey	No description defined	
German	Marina del Rey	Census-designated place im US-Bundesstaat Kalifornien	Marina del Rey, CA
Basque	Marina del Rey (Kalifornia)	No description defined	
Persian	مارینا دل ری، کالیفرنیا	No description defined	
French	Marina Del Rey	census-designated place	
Hebrew	מרינה דל ריי	No description defined	
Croatian	Marina del Rey (Kalifornija)	No description defined	
Italian	Marina del Rey	census-designated place statunitense della California	
Japanese	マリナ・デル・レイ	No description defined	マリナデルレイ
Korean	마리나델레이	No description defined	
Latin	Marina Del Rey	No description defined	
Malagasy	Marina del Rey	No description defined	
Min Nan Chinese	Marina del Rey	No description defined	
Norwegian Bokmål	Marina del Rey	No description defined	
Newari	मरिना डेल रे, क्यालिफोर्निया	No description defined	मरिना डेल रे (क्यालिफोर्निया)
Dutch	Marina del Rey	plaats in Los Angeles County	
Polish	Marina del Rey	No description defined	
Portuguese	Marina del Rey	No description defined	
Romanian	Marina del Rey	loc desemnat pentru recensământ	
Russian	Марина-дель-Рей	No description defined	
Serbo-Croatian	Marina del Rey	No description defined	
Serbian	Марина дел Реј	No description defined	
Swedish	Marina del Rey	No description defined	
Swahili	Marina del Rey	No description defined	Mdr
Ukrainian	Марина-дель-Рей	перелісна місцевість, США, штат California	
Volapük	Marina del Rey	No description defined	
Waray	Marina del Rey, California	No description defined	
Chinese (China)	洛杉矶县 (卡利福尼亚州)	No description defined	瑪麗安德爾灣
Simplified Chinese	玛丽安德尔湾	No description defined	瑪麗安德爾灣 (加州)

Descriptions and labels available in multiple languages

same id is used for all languages!

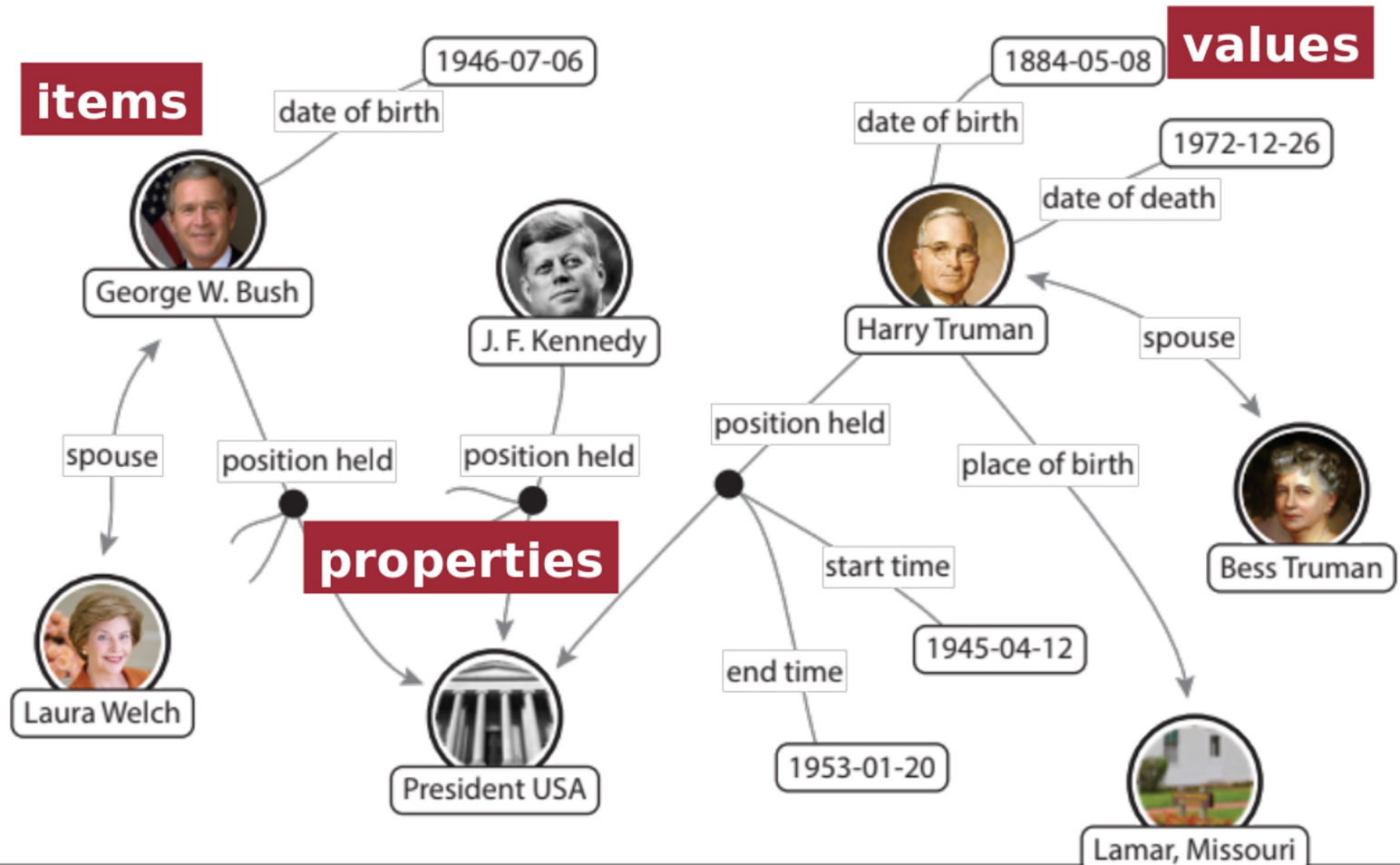
Language	Label	Description	Also known as
English	cycle sport	competitive physical activity using bicycles	bicycle sport bike sport cycling sport sport cycling sports cycling cycling sports
Spanish	ciclismo de competición	competencia deportiva en donde se usan varios tipos de bicicletas	ciclismo de competición
Traditional Chinese	自行車運動	No description defined	
Chinese	自行車運動	体育竞技	專業車手 專業競技 自由車 自行車運動
Arabic	سباق الدراجات النارية	No description defined	
Armenian	ciclismo de competición	competición deportiva onde s'usen bicicletas	deporte ciclista ciclismo deportivo
Azerbaijani	velosiped idmanı	No description defined	
Belarusian	Бенаспорт	No description defined	бенаспорт
Belarusian (Tselinka orthography)	бeнaспoрт	No description defined	
Catalan	ciclisme de competició	Esport	
Central Kurdish	بەسکەڵ سەیارە	No description defined	
Czech	sportovní cyklistika	sport používaný bicykly	cyklistport závodní cyklistika
Welsh	chwareason seiclo	No description defined	
Danish	cykelsport	No description defined	
German	Radsport	Sportart, die mit dem Fahrrad ausgeübt wird	Velosport
Swiss High German	Radsport	No description defined	
Greek	ποδηλασία	δρόμος	
Canadian English	Bicycle racing	No description defined	
British English	Bicycle racing	No description defined	
Esperanto	bicikla sporto	No description defined	
Estonian	jalgtrahisport	No description defined	
Basque	biemulakiztasun	No description defined	
Persian	فوتبال دوچرخه‌سواری	فوتبال دوچرخه‌سواری یا اسکنداره از دوچرخه	
Finnish	kilpapyöräily	pelkkyöräilyn ajamista kilpaa	cyclisme sport sport cyclisme
French	cyclisme	sport de course à vélo	cyclisme de compétition
Western Frisian	Hurdyske	No description defined	
Isish	sport nithalochtha	No description defined	
Hebrew	ספורט אופניים	No description defined	
Hungarian	kétfélekisport	No description defined	
Armenian	Չիկուկիստիկա սպորտ	No description defined	Չիկուկիստիկա
Interlingue	Ciclisme	No description defined	
Italian	ciclismo	sport in cui si utilizza la bicicletta	
Japanese	自転車競技	No description defined	
Kabyle	Ccece saɔ aɣeyya	No description defined	
Kazakh	Бенаспорттық спорт	No description defined	
Korean	자전거 경기	자전거를 타고 서로 겨루는 운동 경기이다.	
Kyrgyz	Бенаспорттуу спорт	No description defined	
Luxembourgish	Vëlossport	Sport dee mam Vëlo gemaach gëtt	
Limburgish	Fietssport	No description defined	
Lithuanian	Dviračių sportas	No description defined	
Latvian	riņķbraukšanas sports	No description defined	
Macedonian	бенаспортен спорт	No description defined	
Marathi	सरायचि	No description defined	
Mn Nan Chinese	Thh-bê kîng-cháu	No description defined	
Norwegian Bokmål	sykkling	sport	sykkelsport
Low German	Radsport	No description defined	

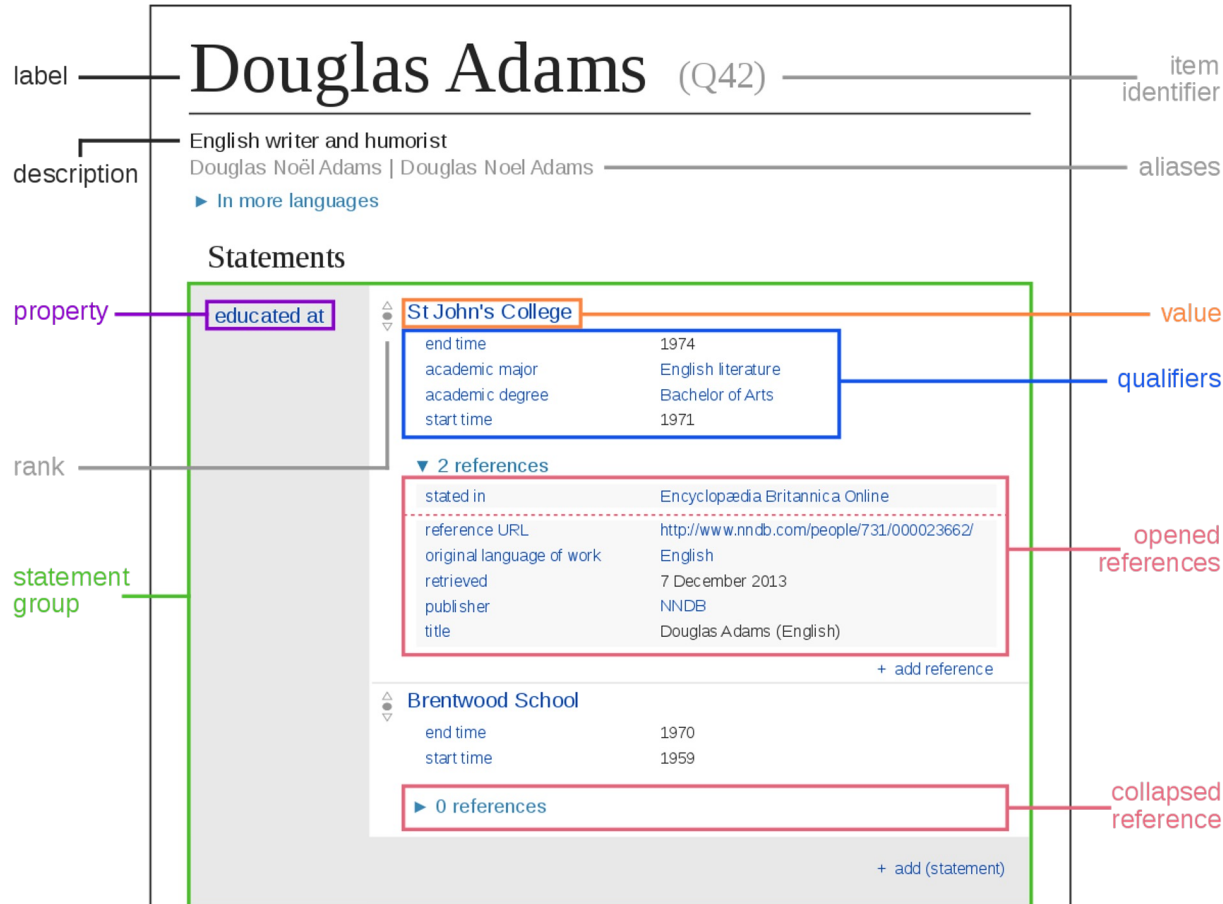


The Wikidata data model

WIKIDATA

Wikidata represents entities in a Knowledge Graph





The image shows a Wikidata profile for **Douglas Adams** (Q42). The profile is annotated with labels and lines pointing to specific parts of the interface:

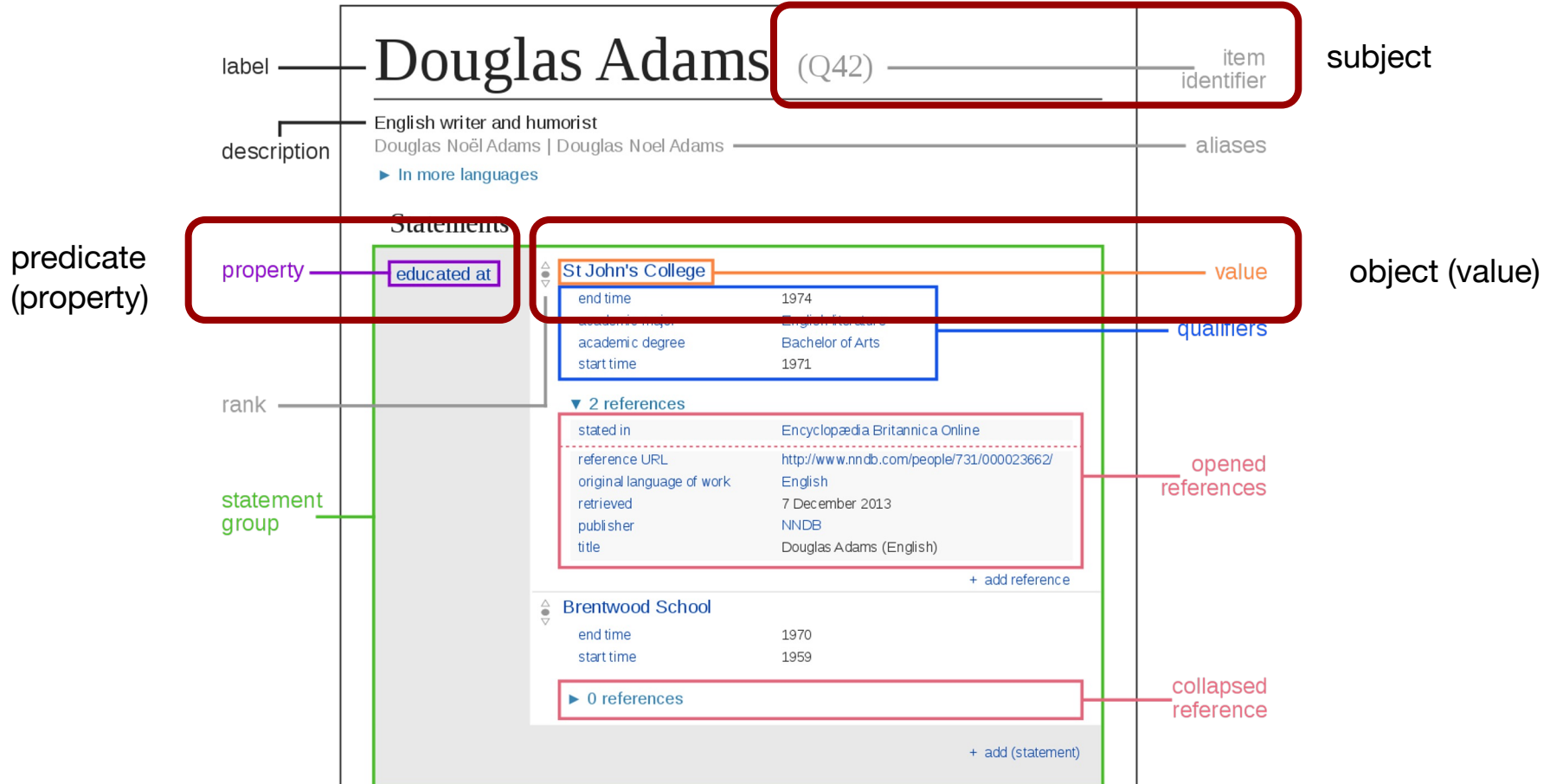
- label**: Points to the name **Douglas Adams**.
- item identifier**: Points to the identifier **(Q42)**.
- description**: Points to the text **English writer and humorist** and the aliases **Douglas Noël Adams | Douglas Noel Adams**.
- aliases**: Points to the text **Douglas Noël Adams | Douglas Noel Adams**.
- property**: Points to the **educated at** property.
- value**: Points to the value **St John's College**.
- qualifiers**: Points to the table of qualifiers for the **educated at** statement, including **end time** (1974), **academic major** (English literature), **academic degree** (Bachelor of Arts), and **start time** (1971).
- rank**: Points to the **rank** column in the statements table.
- statement group**: Points to the **educated at** statement group.
- opened references**: Points to the expanded reference for **stated in** from **Encyclopædia Britannica Online**, showing details like **reference URL**, **original language of work**, **retrieved**, **publisher**, and **title**.
- collapsed reference**: Points to the collapsed reference for **Brentwood School**, which shows **0 references**.

The image shows a Wikidata entity page for **Douglas Adams** (Q42). The page is annotated with labels and lines pointing to specific features:

- label**: Points to the main title "Douglas Adams".
- item identifier**: Points to the identifier "(Q42)".
- description**: Points to the text "English writer and humorist".
- aliases**: Points to the text "Douglas Noël Adams | Douglas Noel Adams".
- property**: Points to the "educated at" property in the "Statements" section.
- value**: Points to the value "St John's College" for the "educated at" property.
- qualifiers**: Points to the table of qualifiers for "St John's College", including "end time", "academic major", "academic degree", and "start time".
- rank**: Points to the "rank" column in the "Statements" section.
- statement group**: Points to the "Statements" section header.
- opened references**: Points to the expanded reference for "Encyclopædia Britannica Online", showing details like "reference URL", "original language of work", "retrieved", "publisher", and "title".
- collapsed reference**: Points to the collapsed reference for "Brentwood School", which shows "0 references".

Statements

property	value	qualifiers	rank	statement group												
educated at	St John's College	<table border="1"> <tr><td>end time</td><td>1974</td></tr> <tr><td>academic major</td><td>English literature</td></tr> <tr><td>academic degree</td><td>Bachelor of Arts</td></tr> <tr><td>start time</td><td>1971</td></tr> </table>	end time	1974	academic major	English literature	academic degree	Bachelor of Arts	start time	1971						
end time	1974															
academic major	English literature															
academic degree	Bachelor of Arts															
start time	1971															
	▼ 2 references															
	<table border="1"> <tr><td>stated in</td><td>Encyclopædia Britannica Online</td></tr> <tr><td>reference URL</td><td>http://www.nndb.com/people/731/000023662/</td></tr> <tr><td>original language of work</td><td>English</td></tr> <tr><td>retrieved</td><td>7 December 2013</td></tr> <tr><td>publisher</td><td>NNDB</td></tr> <tr><td>title</td><td>Douglas Adams (English)</td></tr> </table>	stated in	Encyclopædia Britannica Online	reference URL	http://www.nndb.com/people/731/000023662/	original language of work	English	retrieved	7 December 2013	publisher	NNDB	title	Douglas Adams (English)			
stated in	Encyclopædia Britannica Online															
reference URL	http://www.nndb.com/people/731/000023662/															
original language of work	English															
retrieved	7 December 2013															
publisher	NNDB															
title	Douglas Adams (English)															
	+ add reference															
	Brentwood School	<table border="1"> <tr><td>end time</td><td>1970</td></tr> <tr><td>start time</td><td>1959</td></tr> </table>	end time	1970	start time	1959										
end time	1970															
start time	1959															
	► 0 references															
	+ add (statement)															



Qualifiers (additional context about a triple)

The image shows a Wikidata page for **Douglas Adams** (Q42) with various annotations explaining the structure of statements and qualifiers.

Annotations:

- label:** Points to the main title "Douglas Adams".
- item identifier:** Points to the identifier "(Q42)".
- description:** Points to the text "English writer and humorist".
- aliases:** Points to the text "Douglas Noël Adams".
- property:** Points to the "educated at" property.
- rank:** Points to the "rank" column in the statements table.
- statement group:** Points to the entire statements section.
- Qualifier property:** Points to the "end time" property in the qualifiers table.
- Qualifier value:** Points to the "1974" value in the qualifiers table.
- qualifiers:** Points to the entire qualifiers table.
- opened references:** Points to the "opened references" section.
- collapsed reference:** Points to the "collapsed reference" section.

Statements Table:

rank	property	value
	educated at	St John's College

Qualifiers Table (for "educated at"):

Qualifier property	Qualifier value
end time	1974
academic major	English literature
academic degree	Bachelor of Arts
start time	1971

References:

- 2 references:**
 - opened reference:**

stated in	Encyclopædia Britannica Online
reference URL	http://www.nndb.com/people/731/000023662/
original language of work	English
retrieved	7 December 2013
publisher	NNDB
title	Douglas Adams (English)
 - collapsed reference:**

end time	1970
start time	1959

The image shows a Wikidata profile for Douglas Adams (Q42) with various annotations explaining its structure:

- label**: Points to the name "Douglas Adams (Q42)".
- item identifier**: Points to the QID "(Q42)".
- description**: Points to the text "English writer and humorist" and the aliases "Douglas Noël Adams | Douglas Noel Adams".
- aliases**: Points to the text "Douglas Noël Adams | Douglas Noel Adams".
- property**: Points to the "educated at" statement.
- value**: Points to the value "St John's College".
- qualifiers**: Points to the table of qualifiers for the "educated at" statement.

end time	1974
academic major	English literature
academic degree	Bachelor of Arts
start time	1971
- rank**: Points to the rank indicator (a dot) next to the "educated at" statement.
- statement group**: Points to the entire "Statements" section.
- opened references**: Points to the expanded reference for "Encyclopædia Britannica Online".

stated in	Encyclopædia Britannica Online
reference URL	http://www.nndb.com/people/731/000023662/
original language of work	English
retrieved	7 December 2013
publisher	NNDB
title	Douglas Adams (English)
- collapsed reference**: Points to the collapsed reference for "Brentwood School".

end time	1970
start time	1959

Ranks can be:

- Preferred
- Normal
- Deprecated

Organizing data in Wikidata (Wikidata schema)

WIKIDATA

Kristabel Doebel-Hickok (Q19757547)

American bicycle racer

P31

instance of: Kristabel Doebel-Hickok is a (human)

Statements

Own statements

From related entities

member of sports team

Rally Cycling (United States women's cycling team)
start time : 2019

Cylance Pro Cycling (United States women's cycling team)
start time : 2016
end time : 2018

Team TIBCO (women's cycling team)
start time : 2014
end time : 2015

sport

cycle sport (competitive physical activity using bicycles)

country of citizenship

United States of America (federal republic in North America)

place of birth

Marina del Rey (census designated place)

date of birth

1989-04-28

occupation

sport cyclist (someone who competes in the sport of cycling)

sex or gender

female (human gender (use with P21))

Media

image

Krista Doebel-Hickok, Giro Rosa 2016.jpg

media legend : Au Tour d'Italie 2016 [fr]



Links	
Wikidata page	
Wikipedia article	
Reasonator	

Identifiers		
ProCyclingStats cyclist ID	179437 ↗	
Cycling Archives cyclist ID	98235 ↗	
CQ Ranking female cyclist ID	7931 ↗	

human (Q5)

human being | humankind | people | homosapiens | person

common name of Homo sapiens, unique extant species of the genus Homo

subclass of: every human is also a(n) person, natural person, omnivore**instance of:** human is a(n) group of organisms known by one particular common name ⓘ

Instances

Direct instances

4

Yanka Kupala, František Plánička, Augusto Pinochet, Julian Assange, Blanche of Castile, Adolf Hitler, Piet Kraak, Louis IX of France, Patricio Aylwin, Nicolas Sarkozy, Ricardo Lagos, Eduardo Frei Ruiz-Tagle, Michelle Bachelet, Galileo Galilei, Sebastian Piñera, El Greco, Elvis Presley, Diego Velázquez, Claude Monet, Paul Morand, ... further results

All instances

5511999

Typical Properties

occupation, sex or gender, date of birth, country of citizenship, given name, place of birth, family name, date of death, VIAF ID, educated at, ISNI, place of death, Library of Congress authority ID, languages spoken, written or signed, position held, GND ID, employer, ORCID ID, sport, member of sports team, CBDB ID, name in native language, SUDOC authorities ID,

Classification

Direct superclasses

person [5567561](#), natural person [5512007](#), omnivore [5511999](#)

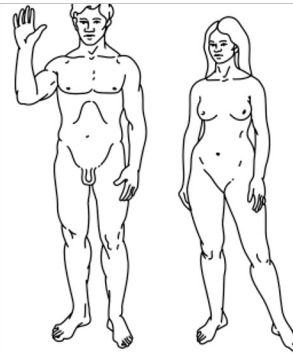
Direct subclasses

With instances [24](#)With subclasses [48](#)All [178](#)

female [491](#), woman [394](#), male [227](#), elected person [153](#), spouse [114](#), person of color [83](#), technician [61](#), child [54](#), migrant [37](#), Guantanamo Bay detainees [36](#), citizen [19](#), candidate [14](#), Caucasian [12](#), missing person [11](#), acquaintance [10](#), inhabitant [9](#), sick person [9](#), patient [8](#), widow or widower [8](#), accused [7](#), anonymous [6](#), eligible voter [6](#), disabled person [6](#), substitute [5](#), adult human [5](#), vacationer [5](#), prenatal [4](#), teenager [4](#), women in STEM fields [4](#), Israelis [3](#), live person [3](#), lover [2](#), bog body [2](#), suspect [2](#), alumnus [2](#), conscientious objection [2](#), Military dependent [2](#), Vulnerable person [2](#), Disadvantaged [1](#), special purpose artist [1](#), drug user [1](#), subject [1](#), colleague [1](#), municipality president [1](#), dead person [1](#), anatomically modern human [1](#), centenarian [1](#), plantsman [1](#)

All subclasses

1474



Classes are items
They can be instances and
subclasses at the same
time!

Links

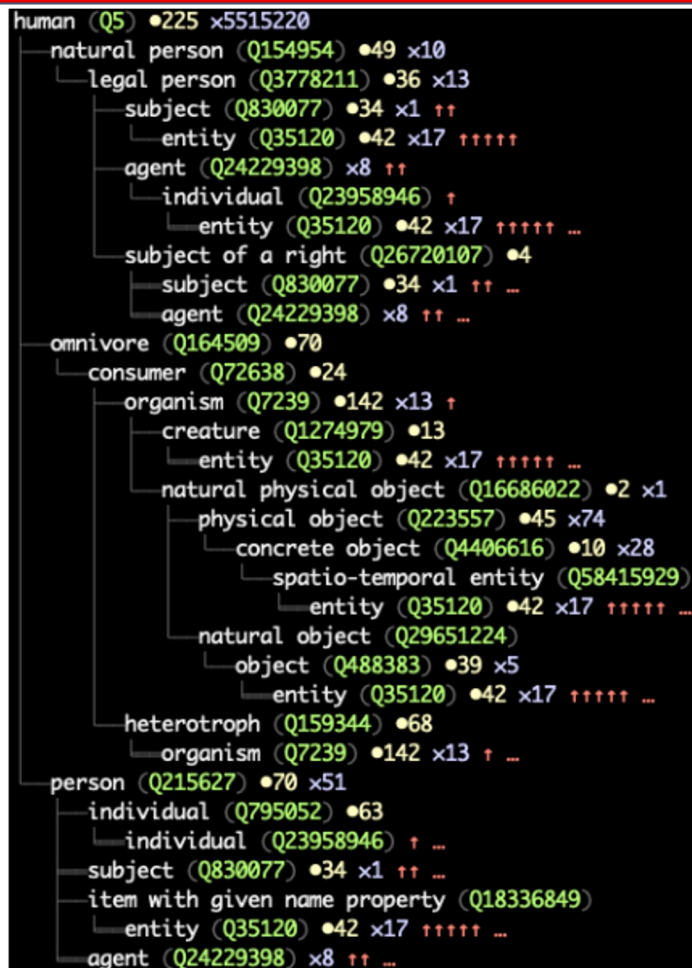
[Wikidata page](#)[Wikipedia article](#)[Reasonator](#)

Identifiers

[Klexikon article ID](#) Mensch [↗](#)[Wolfram Language entity type](#) Person [↗](#)[PSH ID](#) 2693 [↗](#)[OmegaWiki Defined Meaning](#) 2234 [↗](#)[NicoNicoPedia article ID](#) ヒト [↗](#)[JSTOR topic ID](#) humans [↗](#)[Australian Educational](#) scot/1144 [↗](#)

2,464,270 classes:
items with a **subclass**
of (P279) property

63,672 classes with
instances:
appear as values of
instance of (P31)
property



inception (P571)

date founded | date created | incorporated | foundation | created | date of foundation | date of creation | established | establishment date | foundation date | creation date | incorporation date | date incorporated | date of incorporation | date of establishment | founding date | date of founding | date formed | formation date | formed on date | founded on date | incorporated on date | established on date | created on date | commencement date | date commenced | commenced on date | date of commencement | time of foundation or creation | date of foundation or creation | written on date | time of inception | year founded | year created | year incorporated | year written | year commenced | year established | date constructed | construction date | constructed | initiated | formation | dedication date | formed in | formed at | inaugurated | launch date | introduced | introduction | start date | completed | first issue | built | foundation/ creation date | founded

date or point in time when the subject came into existence as defined

Type: Time

subproperty of: every inception is also a(n) start time

instance of: inception is a(n) Wikidata property for items about organizations, Wikidata property with datatype 'time'

Property Usage	
Entities	1713499 Venus of Petřkovice, Venus of Hohle Fels, Mehrgarh, Göbekli Tepe, Tel Megiddo, Venus of Laussel, Çatalhöyük, Knossos, Cave of Altamira, Venus of Willendorf, ... further results
Typical Properties	creator, inventory number, collection, official website, width, height, material used, location, headquarters location, image, described at URL, country, dissolved, abolished or demolished, LfDS object ID, IILF manifest, depicts, National Gallery of Art artwork ID, located in the administrative territorial entity, coordinate location, IUCN protected areas category,
Statements	1731900 (1.01 per entity)
Qualifiers	latest date 80852 , earliest date 79715 , sourcing circumstances 79150 , instance of 10430 , applies to part 8814 , stated as 4862 , start time 4458 , end time 4337 , refine date 2220 , statement is subject of 323 , has cause 151 , location 137 , object has role 123 , subject has role 103 , named as 98 , identity of subject in context 80 , of 70 ,
Uses as qualifier	1685
Uses in references	114

Statements	
Own statements	From related entities
equivalent property	http://schema.org/foundingDate  
subproperty of	start time 
subject item of this property	date of establishment (date something was established) 
Wikidata property example	Society of Jesus (male religious congregation of the Roman Catholic Church) inception: 1539  Church of Saint Michael (church in Olomouc District of Olomouc region) inception: 1300  Dracula (1897 Gothic horror novel by Irish author Bram Stoker) inception: 1890 start time: 1890-03-08 end time: 1897-05-26 

property proposal discussion

https://www.wikidata.org/wiki/Wikidata:Property_proposal/Archive/8#P571 

property constraint

 7 statements 

single value constraint (type of constraint for Wikidata properties: used to specify that this property generally contains a single value per item)
 exception to constraint: Italian National Society
 separator: applies to part
 separator: applies to jurisdiction
 separator: named as
 separator: subject has role
 separator: identity of subject in context
 separator: series ordinal
 separator: nature of statement

conflicts-with constraint (type of constraint for Wikidata properties: used to specify that an item must not have a given statement)
 property: instance of
 item of property constraint: Wikimedia disambiguation page (type of page in the Wikimedia system. Use with P21 'instance of' for disambiguation pages)
 item of property constraint: Wikimedia category (use with 'instance of' (P11) for Wikimedia category)
 item of property constraint: Wikimedia list article (page of a Wikimedia project with a list of something)
 group by: instance of

allowed qualifiers constraint (type of constraint for Wikidata properties: used to specify that only the listed qualifiers should be used. Novalue disallows any qualifier)
 property: earliest date
 property: latest date
 property: sourcing circumstances
 property: instance of
 property: applies to part
 property: statement is subject of
 property: start time
 property: end time
 property: refine date
 property: has cause
 property: time index
 property: identity of subject in context
 property: subject has role
 property: applies to jurisdiction
 property: named as
 property: statement disputed by
 property: stated as
 property: location
 property: has quality
 property: quote
 property: object has role
 property: statement supported by
 property: criterion used
 property: series ordinal
 exception to constraint: Newtontoppen (largest and highest mountain in Svalbard, Norway)
 exception to constraint: JOSM (free software editor for OpenStreetMap data)

see also

service entry

 9 statements 

date of official opening

location of final assembly

inverse property

dissolved, abolished or demolished

Wikimedia Categories and Portals

category for value different from Wikidata

Q22013004 (Wikimedia category)

category for value not in Wikidata

Category:inception not in Wikidata, but available on Wikipedia (Wikimedia category)

category for value same as Wikidata

Q32085204 (Wikimedia category)

property usage tracking category

Category:Pages using Wikidata property P571 (Wikimedia category)

data type

hierarchy

Type: Time

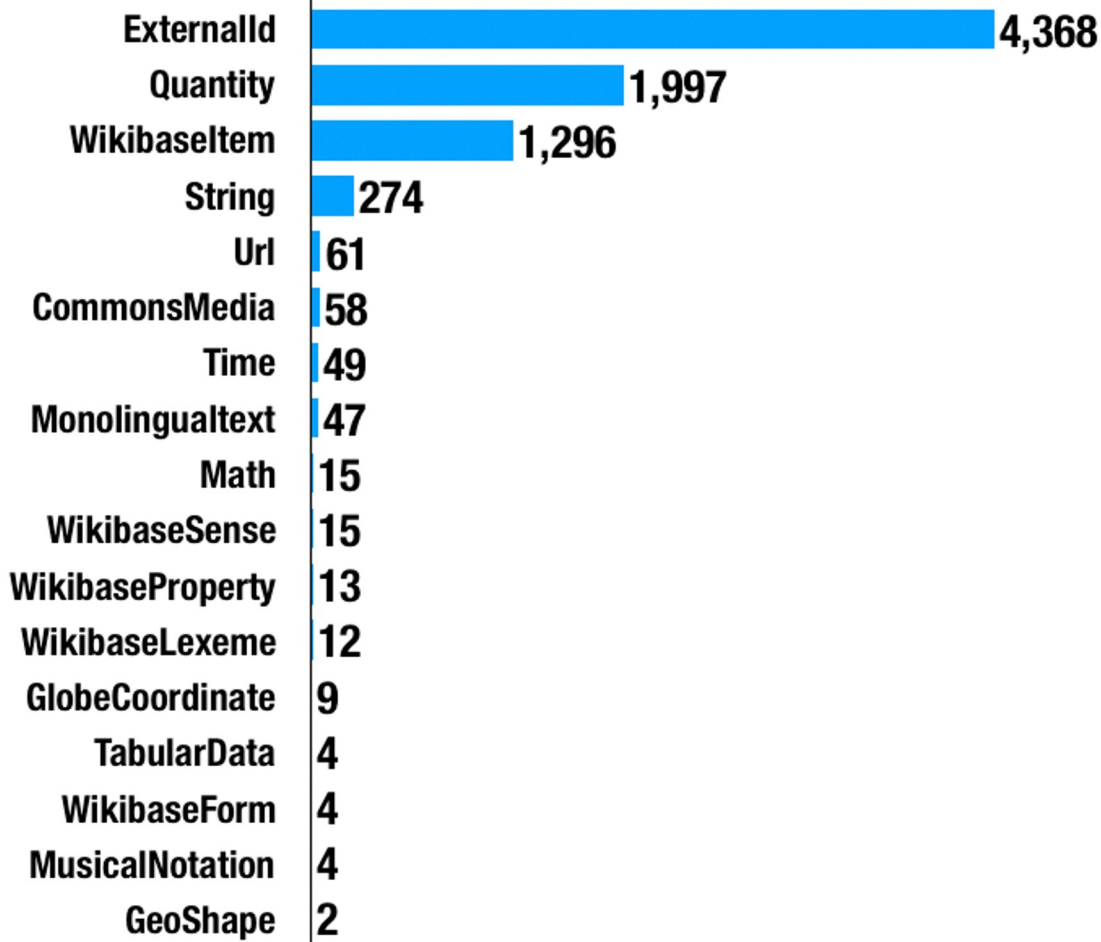
subproperty of: every inception is also a(n) start time

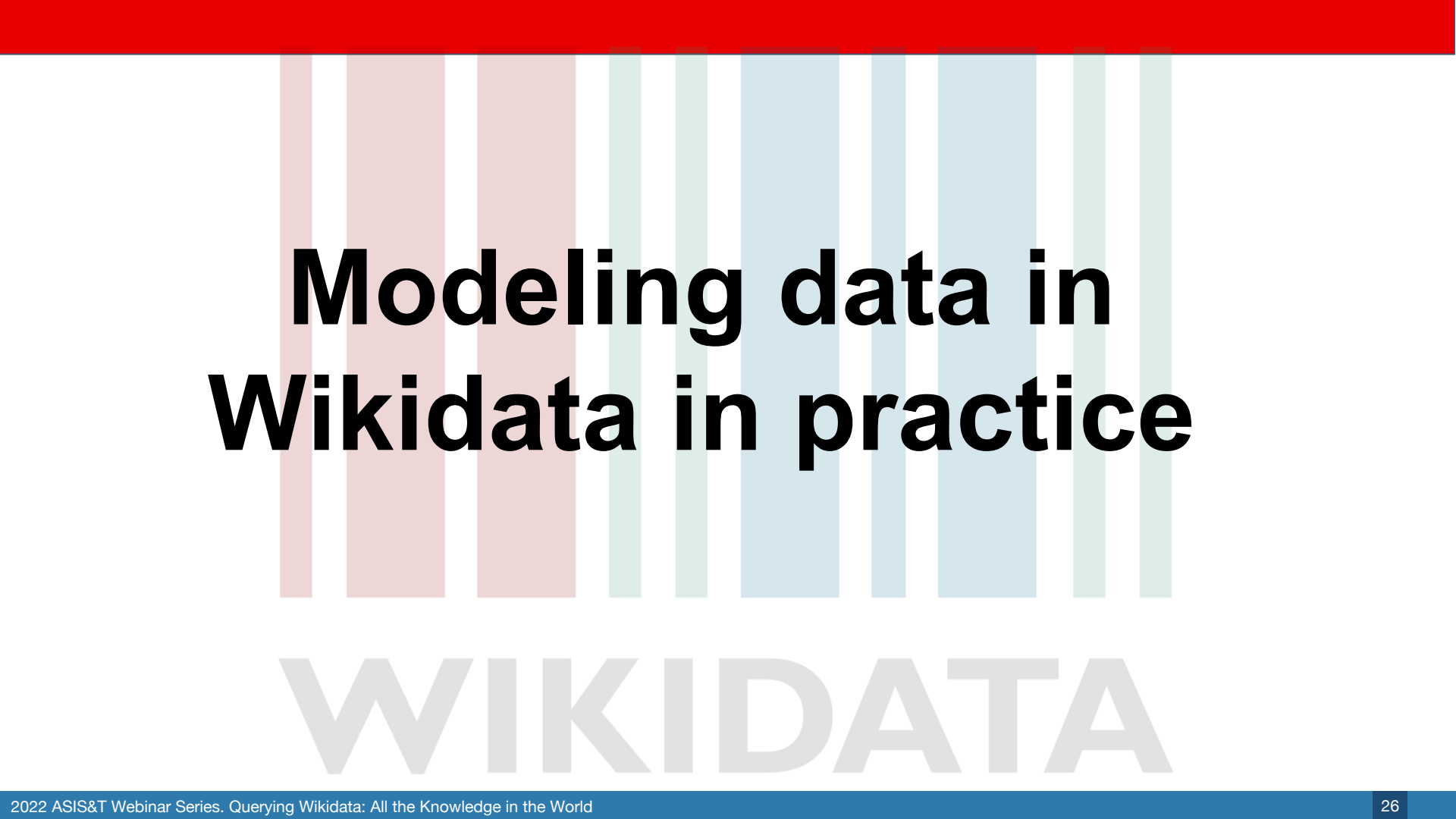
instance of: inception is a(n) Wikidata property for items about organizations, Wikidata property with datatype 'time'

Property Usage	
Entities	1713499 Venus of Petřkovice , Venus of Hohle Fels , Mehrgarh , Göbekli Tepe , Tel Megiddo , Venus of Laussel , Çatalhöyük , Knossos , Cave of Altamira , Venus of Willendorf , ... further results
Typical Properties	creator , inventory number , collection , official website , width , height , material used , location , headquarters location , image , described at URL , country , dissolved , abolished or demolished , LfDS object ID , IIIF manifest , depicts , National Gallery of Art artwork ID , located in the administrative territorial entity , coordinate location , IUCN protected areas category ,
Statements	1731900 (1.01 per entity)
Qualifiers	latest date 80852 , earliest date 79715 , sourcing circumstances 79150 , instance of 10430 , applies to part 8814 , stated as 4862 , start time 4458 , end time 4337 , refine date 2220 , statement is subject of 323 , has cause 151 , location 137 , object has role 123 , subject has role 103 , named as 98 , identity of subject in context 80 , of 70 ,
Uses as qualifier	1685
Uses in references	114

Statements	
Own statements	From related entities
equivalent property	http://schema.org/foundingDate ☑ ★
subproperty of	start time
subject item of this property	date of establishment (date something was established)
Wikidata property example	Society of Jesus (male religious congregation of the Roman Catholic Church) inception : 1539 Church of Saint Michael (church in Olomouc District of Olomouc region) inception : 1300 Dracula (1897 Gothic horror novel by Irish author Bram Stoker) inception : 1890 start time : 1890-03-08 end time : 1897-05-26

"domain"

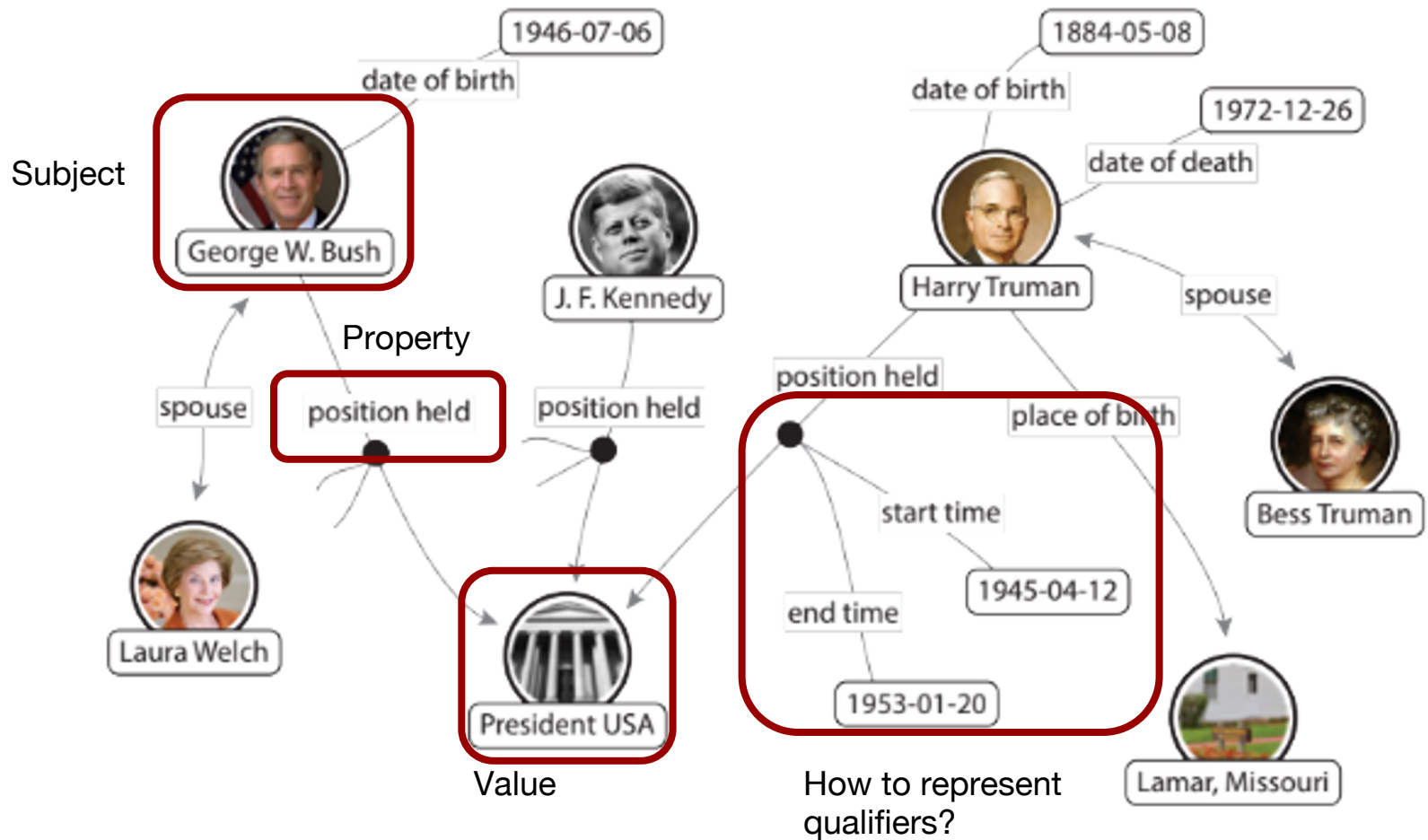




Modeling data in Wikidata in practice

WIKIDATA

We usually represent graphs with triples



How do we represent qualifiers, references and deprecated statements?

Example: is Pluto a planet?

- Until 2006, it was considered a planet
- After 2006, it is considered a trans-neptunian object

In Wikidata, this is represented as follows:

The image shows a screenshot of the Wikidata page for Pluto (Q339). The page title is "Pluto (Q339)" with the identifier "(134340) Pluto | Sol j | 134340 Pluto | (134340) Pluto nan" below it. The main text is "planet in the Solar System". There are two callout boxes: one for "end time : 2006-09-13" and another for "start time : 2006-09-13" and "statement is subject of : IAU definition of planet". The "instance of" section lists "Pluto is a(n) planet", "trans-Neptunian object", "plutino", "dwarf planet", and "plutoid". A blue box labeled "deprecated" with an orange arrow points to the "planet" statement. Below the "instance of" section is a "Statements" section.

Pluto (Q339)
(134340) Pluto | Sol j | 134340 Pluto | (134340) Pluto nan
planet in the Solar System

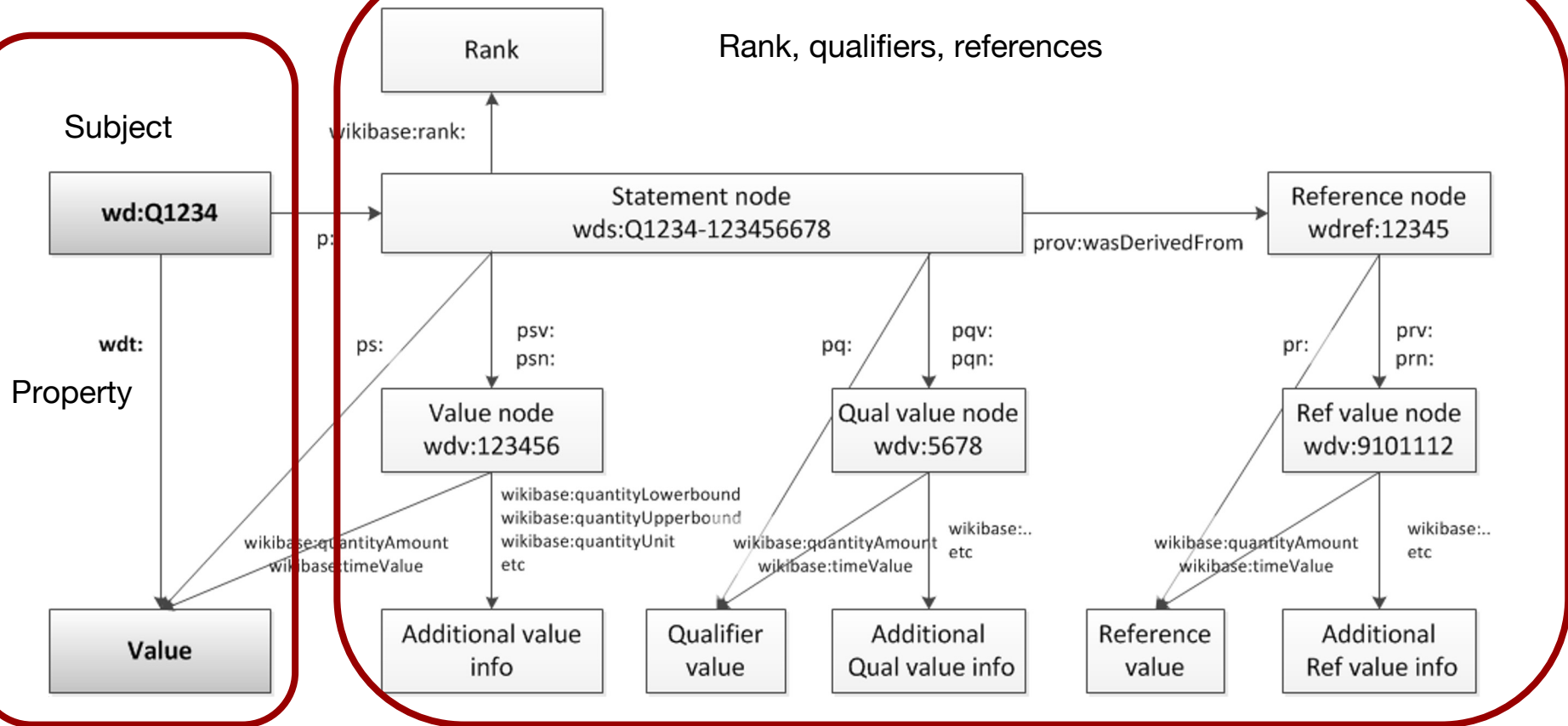
end time : 2006-09-13

start time : 2006-09-13
statement is subject of : IAU definition of planet

instance of: Pluto is a(n) **planet**, trans-Neptunian object, plutino, dwarf planet, plutoid

deprecated →

Statements

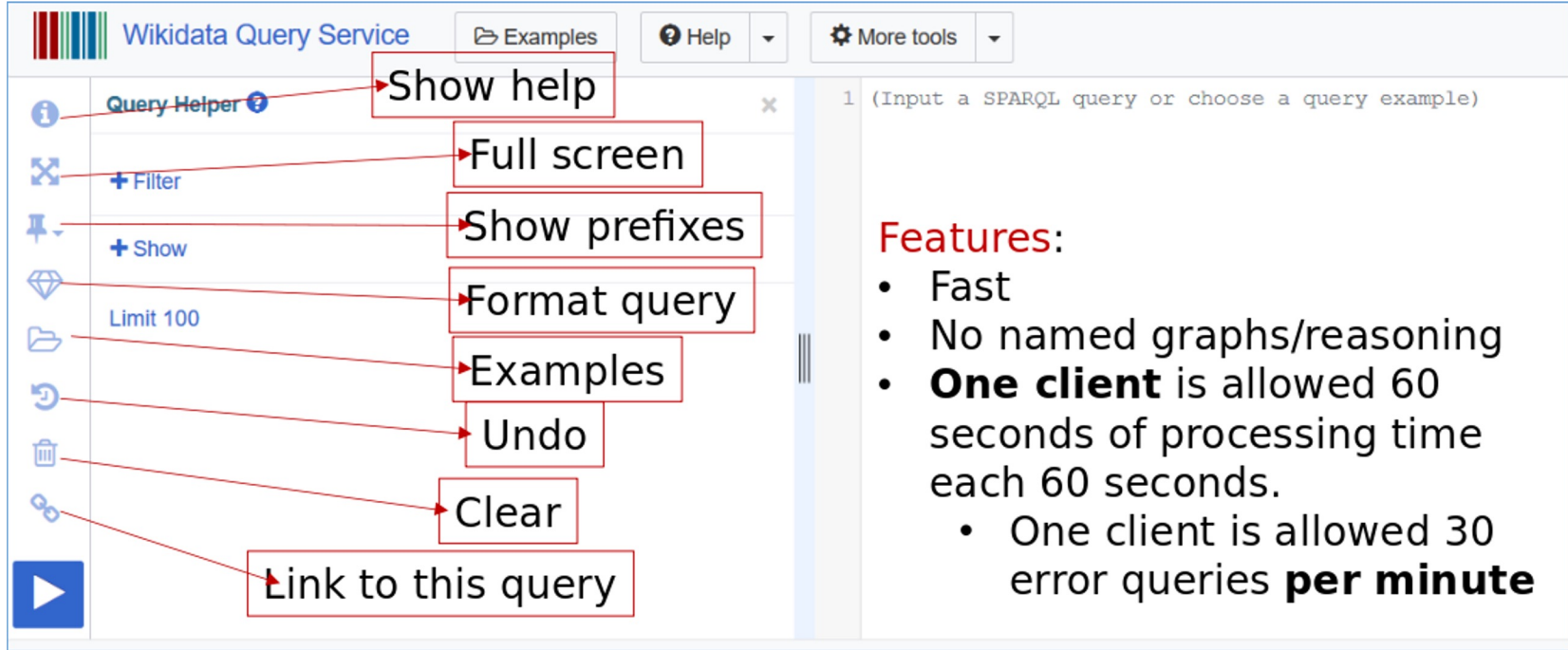




Querying Wikidata with SPARQL

WIKIDATA

<https://query.wikidata.org>



The screenshot shows the Wikidata Query Service interface. On the left is a sidebar with various icons and labels. Red arrows point from these icons to callout boxes on the right. The callouts are: 'Show help' (pointing to the 'Query Helper' icon), 'Full screen' (pointing to the 'Filter' icon), 'Show prefixes' (pointing to the 'Show' icon), 'Format query' (pointing to the 'Limit 100' icon), 'Examples' (pointing to the 'Examples' icon), 'Undo' (pointing to the 'Undo' icon), 'Clear' (pointing to the 'Clear' icon), and 'Link to this query' (pointing to the 'Play' button). The main area on the right contains a text input field with the placeholder text '1 (Input a SPARQL query or choose a query example)'.

Wikidata Query Service

Examples Help More tools

Query Helper

1 (Input a SPARQL query or choose a query example)

Features:

- Fast
- No named graphs/reasoning
- **One client** is allowed 60 seconds of processing time each 60 seconds.
 - One client is allowed 30 error queries **per minute**

Having trouble with a query? Request it: https://www.wikidata.org/wiki/Wikidata:Request_a_query

```
1 select ?s where {
2   ?s rdfs:label "Los Angeles"@en
3 }
```



Q wd:Q65

```
1 select ?p ?o where {
2   wd:Q65 ?p ?o
3 }
```

Los Angeles (Q65) county seat of Los Angeles County, California; second largest city in the United States by population

Q65 (Q504331) Dutch band

Q65 (Q27788669) Wikimedia disambiguation page



skos:altLabel	洛杉磯市
skos:altLabel	洛杉磯市
skos:altLabel	洛杉磯市
skos:altLabel	洛杉磯市
skos:altLabel	洛杉磯市
wikibase:statements	209
wikibase:sitelinks	245
wikibase:identifiers	47
wikibase:timestamp	9 November 2019
wdt:P6	Q wd:Q5386557
wdt:P17	Q wd:Q30

Available in JSON, Turtle, JSON-LD...

More information:

https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/A_gentle_introduction_to_the_Wikidata_Query_Service

Bear in mind that in Wikidata:

- **Everything** is an entity (classes, properties, instances)
- `rdf:type` -> P31
- `rdfs:subClassOf` -> P279
- `rdfs:domain` and `rdfs:range` are type constraint (type and value type constraints)



Functionality designed to facilitate recovering basic resource metadata

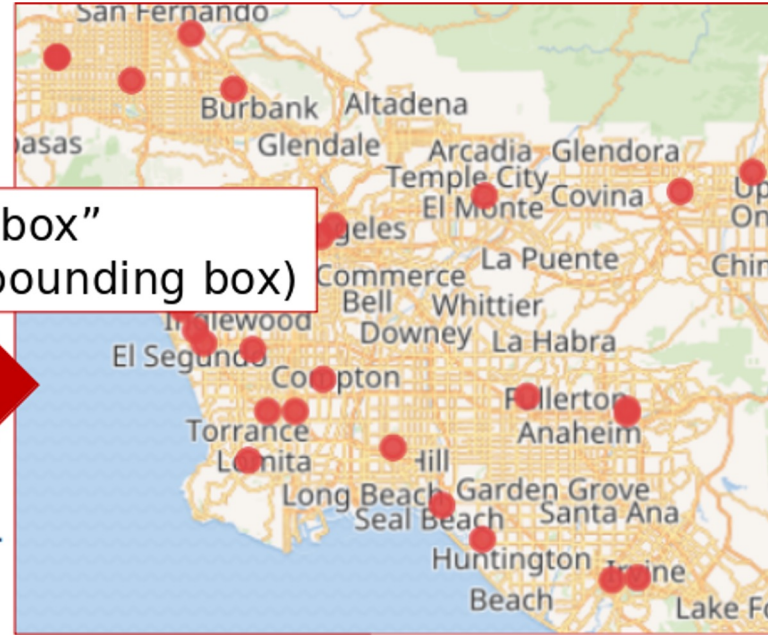
```
select ?p ?o ?oLabel ?oAltLabel ?oDescription where{
  wd:Q65 ?p ?o
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "en" .
  }
}
```

wdt:P17	 wd:Q30	United States of America	America, United States, USA, the United States, the United States of America, the US, the USA, U.S., U.S.A., US, us, US of A	federal republic in North America
wdt:P18	 commons:LA.jpg	http://commons.wikimedia.org/wiki/Special:FilePath/LA.jpg		
wdt:P30	 wd:Q49	North America	NA	continent on the Earth's northwestern quadrant
wdt:P31	 wd:Q62049	county seat	parish capitals, parish seat, parish seats, shire town	administrative center, or seat of government, for a county or civil parish
wdt:P31	 wd:Q1093829	city of the United States	city of the United States, city in the United States, city in the USA, city of the USA	administrative division at the local government level in the United States

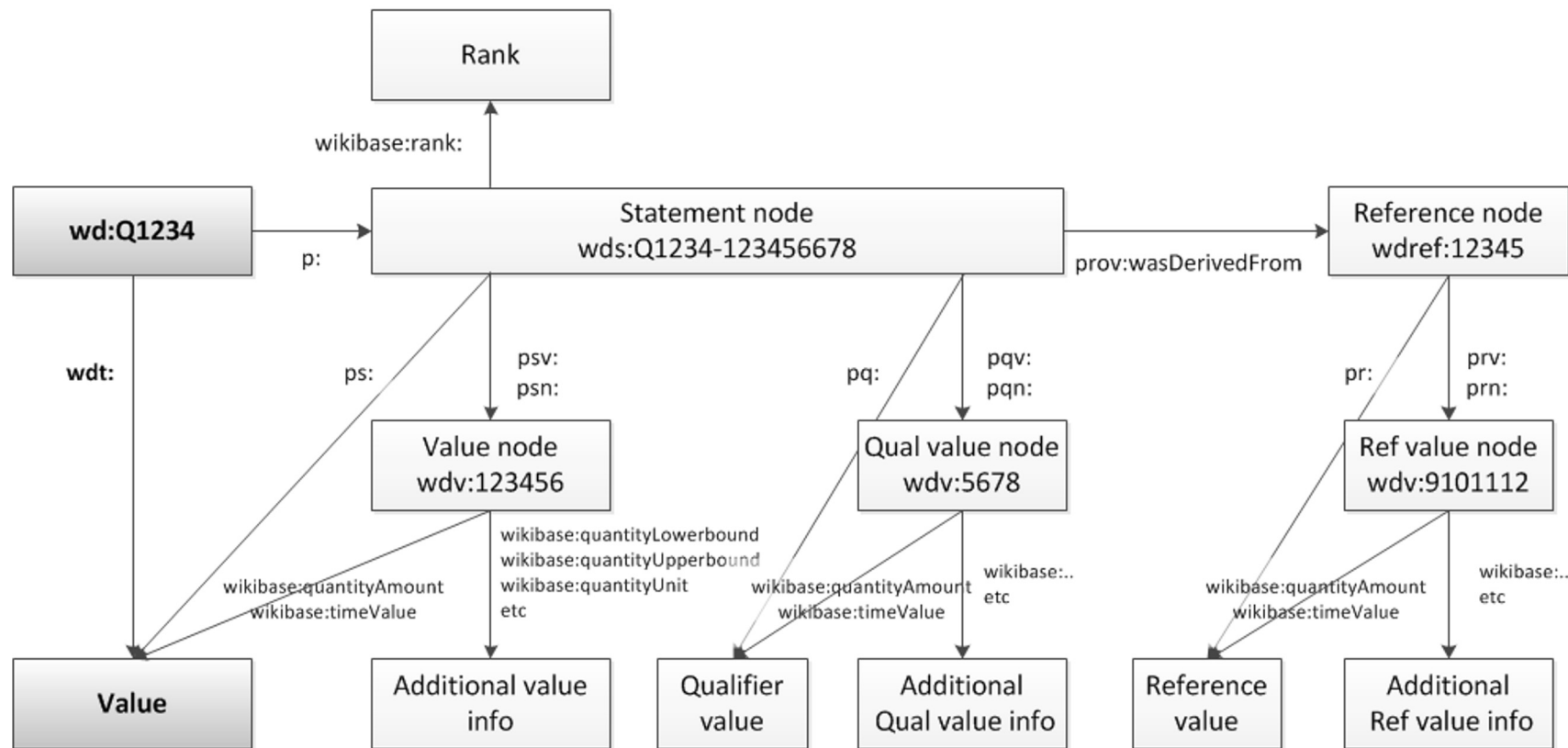
Geospatial search (may be slow)

```
#defaultView:Map
SELECT ?place ?placeLabel ?location
WHERE
{
  # Los Angeles coordinates
  wd:Q65 wdt:P625 ?LALoc .
  SERVICE wikibase:around {
    ?place wdt:P625 ?location .
    bd:serviceParam wikibase:center ?LALoc
    bd:serviceParam wikibase:radius "100" .
  } .
  # Is an airport
  FILTER EXISTS { ?place wdt:P31/wdt:P279* wd:Q1248784 } .
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" .
  }
}
```

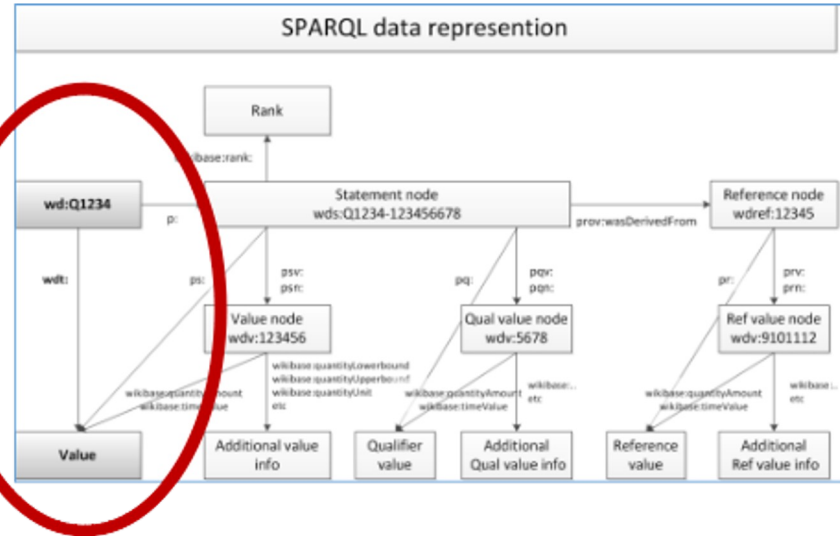
Could also be "box"
(and pass the bounding box)



More info: https://www.mediawiki.org/wiki/Wikidata_Query_Service/User_Manual



Who is the governor in LA?



Link to query: <https://w.wiki/CFf>

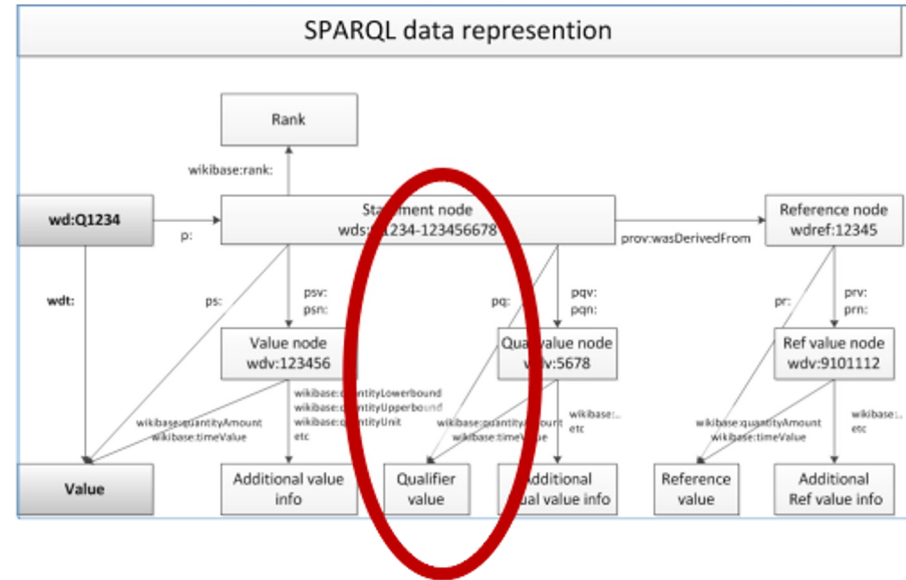
```
select ?p ?o ?oLabel ?oAltLabel ?oDescription where{
  wd:Q65 wdt:P6 ?o
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "en" .
  }
}
```



o	oLabel
wd:Q5386557	Eric Garcetti

How long has the governor been in office?

Link to query: <https://w.wiki/CFh>



```
select ?start where{
  wd:Q65 p:P6 ?statement.
  ?statement pq:P580 ?start
}
```



start
1 July 2013

What if we want to know more about how the qualifier value has been represented?

For example, what is the time precision used in the previous start date?

```
select ?p ?o where{  
  wd:Q65 p:P6 ?statement.  
  ?statement pqv:P580 ?start.  
  ?start ?p ?o  
}
```



p	o
wikibase:timeValue	1 July 2013
wikibase:timePrecision	11
wikibase:timeTimezone	0
wikibase:timeCalendarModel	 wd:Q1985727
rdf:type	wikibase:TimeValue

Link to query: <https://w.wiki/CFj>

Let's now review the population in LA

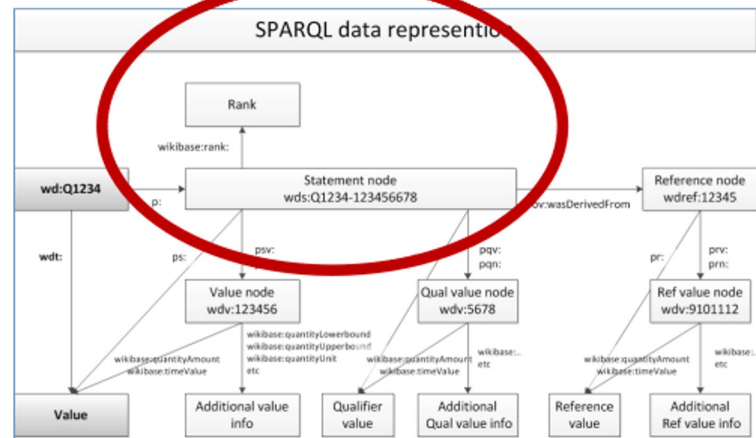
```
select ?population where{
  wd:Q65 wdt:P1082 ?population.
}
```



population
3898747

This returns a single result, which corresponds to the “truthy” statement (first rank). In Wikidata, the truthy statement for population is set to the latest available year:

```
select ?value ?year ?rank where{
  wd:Q65 p:P1082 ?ps. #1082 is population
  ?ps ps:P1082 ?value .
  ?ps wikibase:rank ?rank.
  ?ps pq:P585 ?year. #point in time
}order by ?year
```



Link to queries: <https://w.wiki/5gKZ>, <https://w.wiki/5gKa>

```
select ?value ?year ?rank where{
  wd:Q65 p:P1082 ?ps. #1082 is population
  ?ps ps:P1082 ?value .
  ?ps wikibase:rank ?rank.
  ?ps pq:P585 ?year. #point in time

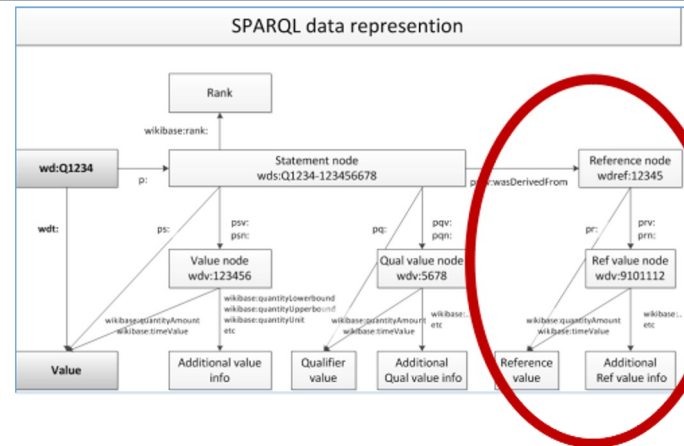
}order by ?year
```

value	year	rank
1610	1 January 1850	wikibase:NormalRank
4385	1 January 1860	wikibase:NormalRank
5728	1 January 1870	wikibase:NormalRank
11183	1 January 1880	wikibase:NormalRank
50395	1 January 1890	wikibase:NormalRank
102479	1 January 1900	wikibase:NormalRank
319198	1 January 1910	wikibase:NormalRank
3792621	1 April 2010	wikibase:NormalRank
3976322	1 January 2016	wikibase:NormalRank
3990456	1 January 2018	wikibase:NormalRank
3898747	1 January 2020	wikibase:PreferredRank
3898747	1 April 2020	wikibase:PreferredRank

What is the source for each triple regarding the population in LA?

Link to query: <https://w.wiki/CFo>

```
select ?value ?year ?rank ?refValueLabel where{
  wd:Q65 p:P1082 ?ps. #1082 is population
  ?ps ps:P1082 ?value .
  ?ps wikibase:rank ?rank.
  ?ps pq:P585 ?year. #point in time
  ?ps prov:wasDerivedFrom/pr:P854 ?refValue #reference URL
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" .
  }
}order by ?year
```



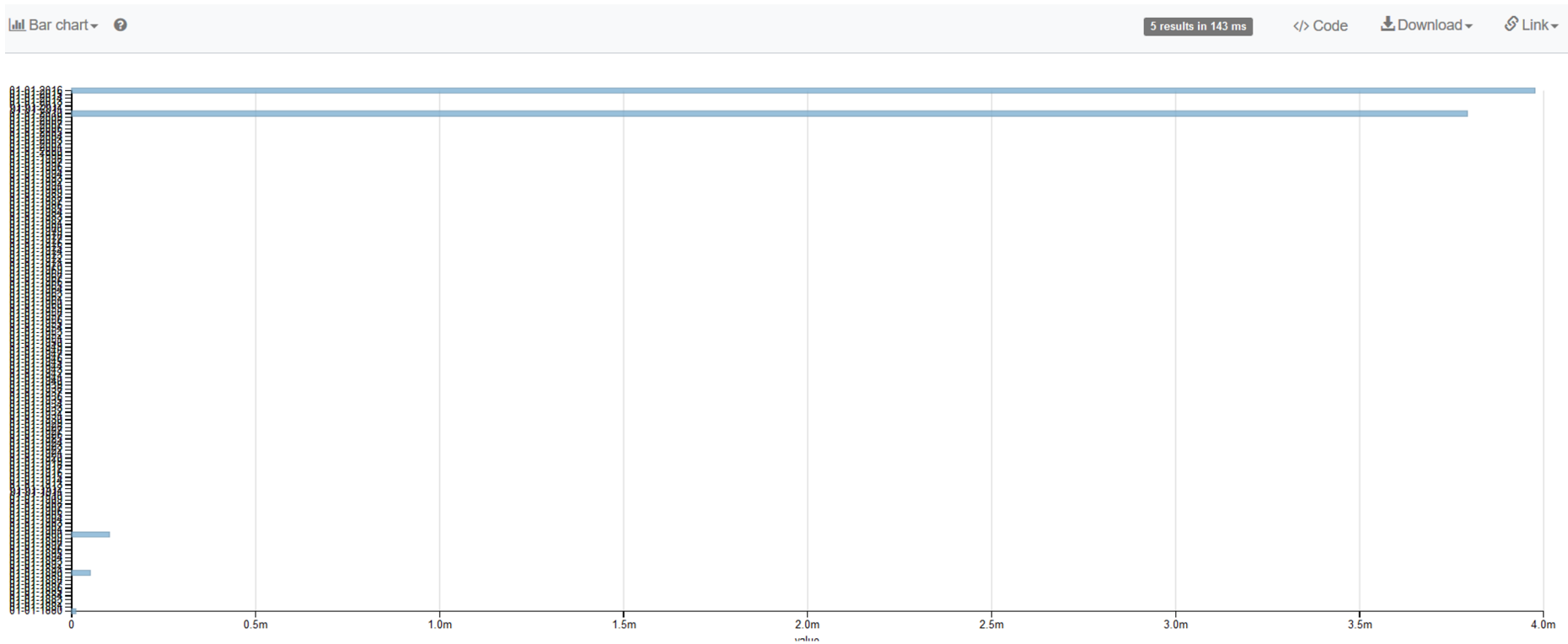
value	year	rank	refValueLabel
11183	1 January 1880	wikibase:NormalRank	http://www.la Almanac.com/population/po02.htm
50395	1 January 1890	wikibase:NormalRank	http://physics.bu.edu/~redner/projects/population/cities/la.html
102479	1 January 1900	wikibase:NormalRank	http://www.la Almanac.com/population/po02.htm
3792621	1 January 2010	wikibase:NormalRank	http://search.census.gov/search?q=cache:b2N7EEjvBBcJ:2010.census.gov/news/xls/cb118&client=2010prod&proxystylesheet=2010prod&site=2010&access=p&oe=UTF-8
3976322	1 January 2016	wikibase:PreferredRank	https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

Population of LA by year

Table



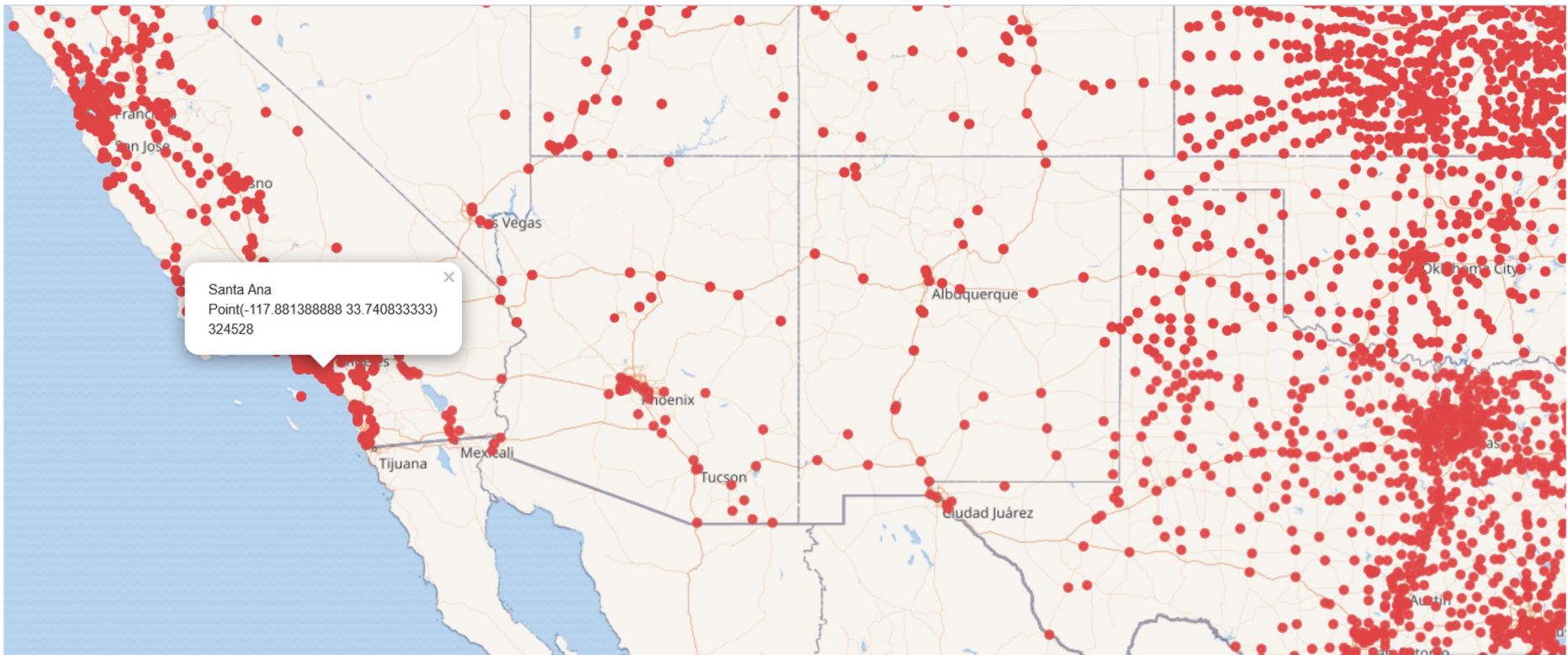
value	year
11183	1 January 1880
50395	1 January 1890
102479	1 January 1900
3792621	1 January 2010
3976322	1 January 2016



Population for all cities in the US

```
select ?cityLabel ?value ?coords where{
  ?city wdt:P31 wd:Q1093829 .
  ?city wdt:P1082 ?value .
  ?city wdt:P625 ?coords . #coordinates |
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" .
  }
}
```

Link to query: <https://w.wiki/CFy>



Entities can be classes and instances in Wikidata

```
select ?sLabel ?classLabel ?upperClassLabel where {  
  ?s wdt:P31 ?class.  
  ?s wdt:P279 ?upperClass  
  SERVICE wikibase:label {  
    bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en" .  
  }  
}LIMIT 100
```



sLabel	classLabel	upperClassLabel
Boeing 747-8	aircraft model	Boeing 747
Barossa German	dialect	German
Namibian Black German	pidgin	German
Franks	language	German
Franks	ethnic group	German
Franks	archaeological culture	German



Summing up...

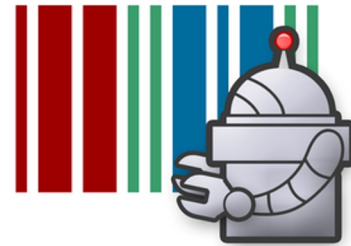
WIKIDATA

Wikidata is a large **Knowledge Graph**

- Free
- Collaborative
- Multilingual
- A secondary database
- Collecting structured data
- Support for Wikimedia wikis
- Anyone in the world
- With many applications
- A thriving community of contributors

> 99M items
> 6,000 properties
> 2,000,000 classes
> 4,000 external id types

Awesome SPARQL endpoint



THANK YOU FOR ATTENDING



Submit a webinar proposal at
<https://www.asist.org/meetings-events/webinars/>



webinars@asist.org

A copy of the recording and a follow-up survey will be emailed
within 24 hours.

asis&t

Association for Information Science and Technology