Wikidata

Structured data repository that anyone can edit

Filip Maljković
PSSOH, October 26, 2019

Two main problems with Wikipedia

1. Outdated data; no triggers for change; boring and repetitive work
   - There's a large number of Wikipedias; some will have updated info, some won't
2. Inflexible ways for fetching specific data
   - Limits of categories
   - Example: all authors whose parents were born in Paris

Two main problems with Wikipedia

One solution for both problems:

Wikidata - central multilingual repository for linked and structured data, on a wiki, that anyone can edit, under a free license

Structured data

Statement = Item: property → value

- Item (Q) - any topic (person, object, place, concept)
- Property (P) - specific type of data relevant for an item (e.g. *gender* for people, *capital* for countries, *length* for rivers)
- Value - either a "literal value" (e.g. length of Danube is 2860 km), or a reference to another item (e.g. capital of Serbia is <u>Belgrade</u>)

- Item is described by a set of statements (each statement provides a fact or a data point about an item)

Structured data

- Earth (**Q2**): highest point (**P610**) → Mount Everest (**Q513**)
- Mount Everest (**Q513**): elevation above sea level (**P2044**) → 8848 m
- Earth (**Q2**): lowest point (**P1589**) → Challenger Deep (**Q459173**)
- Challenger Deep (**Q459173**): elevation above sea level (**P2044**) → -10,994±40 m

Structured data

Or:

- Q2: P610 → Q513
- Q513: P2044 → 8848 m
- Q2: P1589 → Q459173
- Q459173: P2044 → -10,994±1 m

Earth (Q2)

the third planet in the Solar System

Gaia | ⊕ | the world | Blue Planet | Terra Mater | Terra | Planet Earth | Tellus | The Earth | Sol III | Sol 3 | Sol d | ♁

▾ In more languages Configure

Language	Label	Description	Also known as
English	Earth	the third planet in the Solar System	Gaia ⊕ the world Blue Planet Terra Mater Terra Planet Earth Tellus The Earth Sol III Sol 3 Sol d ♁
Croatian	Zemlja	treći unutarnji planet Sunčeva sustava	Svijet Zemlja Majka Zemlja Geja Terra
Serbian	Земља	трећа планета у Сунчевом систему	
српски	No label defined	No description defined	

All entered languages

Statements

instance of		inner planet of the Solar System
		0 references

part of		Solar System
		0 references

Wikipedia (238 entries)   [Collapse]

ab	Адгьыл
ace	Bumoë
ady	Чӏыгу
af	Aarde ☀
als	Erde
am	መሬት
ang	Eorðe
an	Tierra
arc	ܐܪܥܐ
ar	الأرض ☀
arz	الأرض
ast	Tierra
as	পৃথিবী
av	Ракь (планета)
ay	Aka pacha
azb	یئر
az	Yer
bar	Eadn
bat_smg	Žemė
ba	Ер
bcl	Kinaban
be_x_old	Зямля
be	Планета Зямля ☀
bg	Земя ☀
bh	धरती
bjn	Bumi
bn	পৃথিবী
bo	མ�ི�ལ་
br	Douar (planedenn)
bs	Zemlja (planeta)
bxr	Дэлхэй
ca	Terra
cbk_zam	Tierra
cdo	Dê-giù

- Localization
- Description
- Aliases
- Interwiki

highest point

Chimborazo

applies to part — centre

country — Ecuador

1 reference

reference URL — https://www.theweathernetwork.com/us/news/articles/climate-and-environment/ecuadors-mt-chimborazo-is-officially-highest-spot-on-earth/66219

Mount Everest

elevation above sea level — 8,848±20 metre

applies to part — Earth's surface

country — Nepal

0 references

deepest point

Challenger Deep

elevation above sea level — -10,994±40 metre

1 reference

reference URL — http://abstractsearch.agu.org/meetings/2011/FM/OS13B-1517.html

retrieved — 8 January 2017

publication date — 2011

Galathea Depth

end time — 1952

elevation above sea level — -10,540±1 metre

1 reference

imported from — German Wikipedia

Vityas Deep 1

start time — 1957

elevation above sea level — -10,950±50 metre

hak	Thi-khiù
ha	Duniya
haw	Honua
he	כדור הארץ
hif	Dunia
hi	पृथ्वी
hr	Zemlja
hsb	Zemja
ht	Latè
hu	Föld
hy	Երկիր
ia	Terra
id	Bumi 🏅
ie	Terra
ig	Àlà
ilo	Daga (planeta)
io	Tero
is	Jörðin 🏅
it	Terra 🏅
iu	ᓄᓇ
jam	Oert
ja	地球 🏅
jbo	terdi
jv	Bumi
kaa	Jer
kab	Tagnit
ka	დედამიწა
kg	Ntoto
kk	Жер 🏅
kl	Nunarsuaq
km	ផែនដី
kn	ಭೂಮಿ
koi	Мушар
ko	지구
krc	Джер
ksh	Ääd (Planeet)
ku	Erd

- Qualifiers
- Temporal data
- Data types
- Error ranges
- References

Data retrieval

- SPARQL
- Wikidata will give you all they *know*, but not more than that
- The more data there is, the more useful Wikidata is
- https://query.wikidata.org

Female mayors

```
 9
10    ?city wdt:P31/wdt:P279* ?c .  # find instances of subclasses of city
11    ?city p:P6 ?statement .             # with a P6 (head of goverment) statement
12    ?statement ps:P6 ?mayor .           # ... that has the value ?mayor
13    ?mayor wdt:P21 ?sex .       # ... where the ?mayor has P21 (sex or gender) female
14    FILTER NOT EXISTS { ?statement pq:P582 ?x }  # ... but the statement has no P582 (end date) qualifier
15
16    # Now select the population value of the ?city
17    # (wdt: properties use only statements of "preferred" rank if any, usually meaning "current population")
18    ?city wdt:P1082 ?population .
19    # Optionally, find English labels for city and mayor:
20    SERVICE wikibase:label {
21        bd:serviceParam wikibase:language "sr", "en" .
22    }
23 }
24 ORDER BY DESC(?population)
25 LIMIT 10
```

Press [CTRL-SPACE] to activate auto completion.   Data updated `a few seconds` ago

▶ Run   Clear   10 Results in 995 ms   ◉ Display ▾   ⬇ Download ▾   🔗 Link ▾

city	cityLabel	mayor	mayorLabel
🔍 wd:Q1490	Токио	🔍 wd:Q261703	Yuriko Koike
🔍 wd:Q1156	Мумбај	🔍 wd:Q18218029	Snehal Ambekar
🔍 wd:Q38283	Јокохама	🔍 wd:Q529363	Fumiko Hayashi
🔍 wd:Q1533	Каракас	🔍 wd:Q21573102	Helen Fernández
🔍 wd:Q2807	Мадрид	🔍 wd:Q19592761	Manuela Carmena
🔍 wd:Q11462	Сурабаја	🔍 wd:Q12522317	Tri Rismaharini
🔍 wd:Q220	Рим	🔍 wd:Q23766020	Вирџинија Рађи
🔍 wd:Q90	Париз	🔍 wd:Q2851133	Anne Hidalgo
🔍 wd:Q16555	Хјустон	🔍 wd:Q213847	Annise Parker
🔍 wd:Q3915	Антананариво	🔍 wd:Q14618311	Lalao Ravalomanana

Billionaires

▶ Run	Clear	67 Results in 466 ms			⊙ Display ▾	⬇ Download ▾	🔗 Link ▾

item	itemLabel	billions
🔍 wd:Q5284	Бил Гејтс	77.7
🔍 wd:Q29279	Амансио Ортега	73.5
🔍 wd:Q47213	Ворен Бафет	60.8
🔍 wd:Q170419	Карлос Слим	49.9
🔍 wd:Q1066551	Charles G. Koch	42.9
🔍 wd:Q4934	Лари Пејџ	34.8
🔍 wd:Q92764	Сергеј Брин	34
🔍 wd:Q92764	Сергеј Брин	33.4
🔍 wd:Q92764	Сергеј Брин	32.6
🔍 wd:Q2105413	Wang Jianlin	28.7
🔍 wd:Q12908	Џорџ Сорос	24.9
🔍 wd:Q181162	Steve Ballmer	24.8
🔍 wd:Q557305	Phil Knight	23.4
🔍 wd:Q92764	Сергеј Брин	18.7
🔍 wd:Q293679	Al-Waleed bin Talal	18.3
🔍 wd:Q1394965	Aliko Dangote	18
🔍 wd:Q1225503	Dilip Shanghvi	17.5
🔍 wd:Q717038	Masayoshi Son	17
🔍 wd:Q92764	Сергеј Брин	16.6

Eponyms

eponym	eponymLabel	count	sample	sampleLabel
Q wd:Q41688	Хендрик Антон Лоренц	97	Q wd:Q172137	Лоренцова сила
Q wd:Q6722	Карл Фридрих Гаус	80	Q wd:Q2658	Gaussian elimination
Q wd:Q41585	Давид Хилберт	61	Q wd:Q190056	Хилбертов простор
Q wd:Q173746	Пал Ердеш	49	Q wd:Q243972	Ердешев број
Q wd:Q935	Исак Њутн	44	Q wd:Q12438	њутн
Q wd:Q7604	Леонард Ојлер	43	Q wd:Q82435	e
Q wd:Q318362	Johan de Witt	42	Q wd:Q290492	HNLMS Johan de Witt
Q wd:Q42299	Bernhard Riman	36	Q wd:Q152860	4167 Riemann
Q wd:Q8814	Огистен Луј Коши	32	Q wd:Q45286	Cauchy–Binet formula
Q wd:Q9041	Џејмс Ват	29	Q wd:Q25236	ват
Q wd:Q23696	Simon Stevin	27	Q wd:Q1344596	Stevinus
Q wd:Q176659	Андреј Марков	25	Q wd:Q2733	Lempel–Ziv–Markov chain Algorithm
Q wd:Q80222	Жозеф Луј Лагранж	25	Q wd:Q190463	Лагранжова тачка
Q wd:Q153224	Андреј Колмогоров	24	Q wd:Q575766	Kolmogorov–Smirnov test
Q wd:Q77141	Александар Гротендик	24	Q wd:Q1062242	Grothendieck topology
Q wd:Q762	Леонардо да Винчи	23	Q wd:Q19101	Leonardo da Vinci–Fiumicino Airport
Q wd:Q17455	Џон фон Нојман	22	Q wd:Q77887	Von Neumann universe
Q wd:Q7251	Алан Тјуринг	21	Q wd:Q163310	Тјурингова машина
Q wd:Q44481	Пјер Симон Лаплас	20	Q wd:Q44666	Laplace's demon
Q wd:Q81082	Анри Поенкаре	19	Q wd:Q146944	2021 Поенкаре
Q wd:Q619	Никола Коперник	18	Q wd:Q1278	коперницијум

Cats

commons:James bowen (author).jpeg
Bob
wd:Q27744042

commons:Ayasofya Gli cat.jpg
Gli
wd:Q28792126

Mrs. Chippy
commons:Mrschippy2.png
wd:Q851190

Hodge
commons:Hodgecat flickr.jpg
wd:Q1622688

Scarlett
commons:ScarlettTheCat.jpg
wd:Q4421603

Mr. Nuts
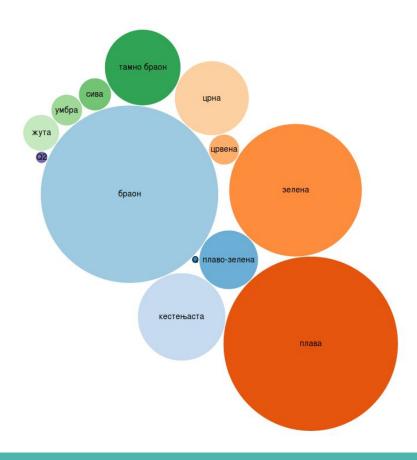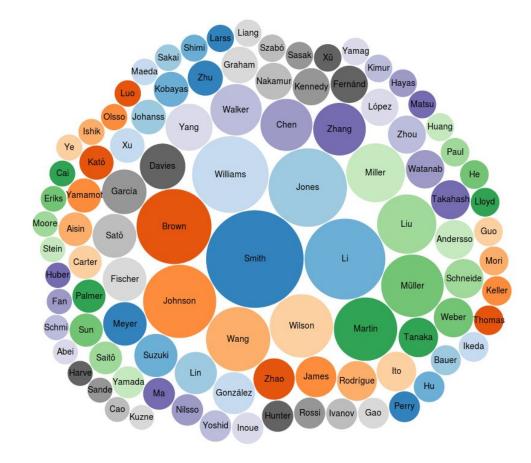commons:Mr Nuts Executive Portrait.jpg
wd:Q6928344

Tsim Tung Brother Cream
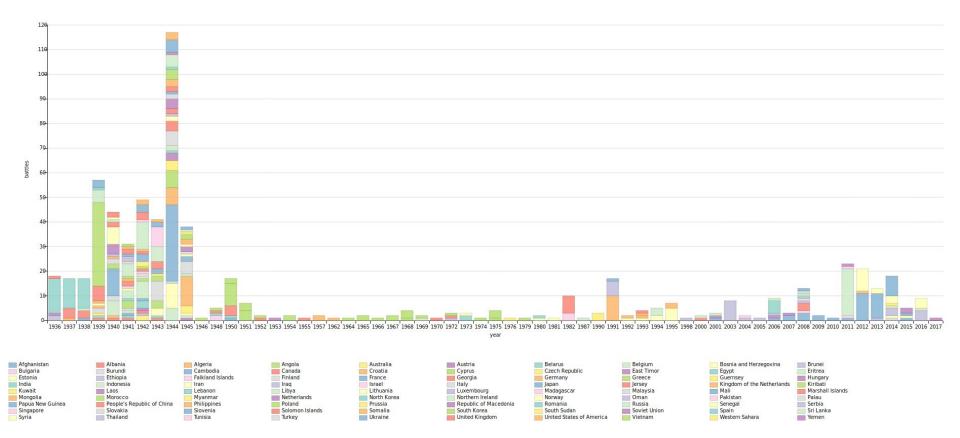commons:Brother cream in stall.jpg
wd:Q10813994

Eye colors
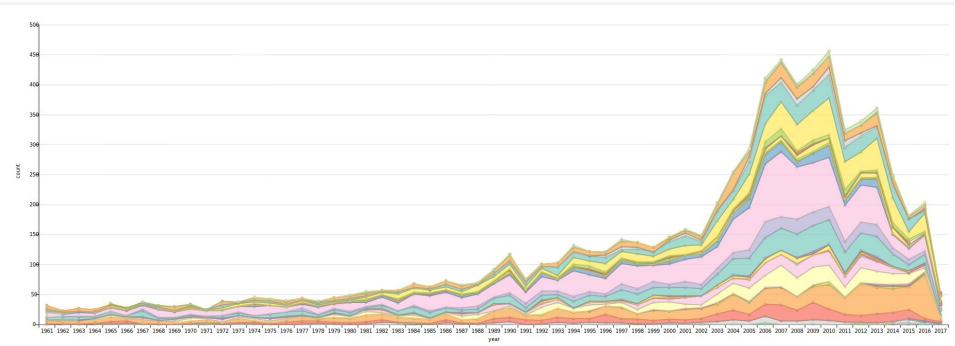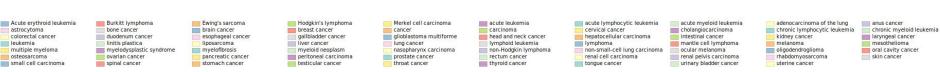
Most frequent last names

Wars by country by year

Incidence of cancer by year

Airports of Belgium

Links

- http://tools.wmflabs.org/wikidata-game/ - Gamification
- http://tools.wmflabs.org/fist/wdfist/ - Image search
- http://tools.wmflabs.org/mix-n-match/ - Linking to external databases
- https://tools.wmflabs.org/scholia/ - Citations, papers, authors
- https://tools.wmflabs.org/reasonator/ - Narrative view
- http://tools.wmflabs.org/magnustools/#wikidata - More...
- https://www.wikidata.org/wiki/Wikidata:Tools/External_tools - Misc

Questions?

Thank you!

filip@vikimedija.org