

Wikidata

A Gentle Introduction for Complete Beginners

Asaf Bartov <asaf@wikimedia.org>

Video of (almost) this talk:

<https://www.youtube.com/watch?v=eVrAx3AmUvA>

Wikimania 2023, Singapore

Hello! I'm Asaf.

- Longtime Wikipedia volunteer
- Amateur digital librarian
- Linked data enthusiast
- I have been teaching Wikidata around the world since 2016

You are lucky!

You're (only just) still in time to be considered a Wikidata hipster!

Four key problems

1. Dated data; lacking triggers; repetitive work
2. Inflexible ways of lateral queries of knowledge (Categories and their limitations)
3. Describing data with data.
4. Correlating data from disparate sources is hard work.

~~Four key problems~~

One solution

An editable central storage for **structured and linked data**, on a wiki, under a free license.

(AKA **Wikidata**)

In short:

Wikidata is 

Structured data, you say?

Claim = Item --> Property --> Value

Item is any topic (Wikipedia article, thing, person, place, concept, etc.)

Property is a specific single **kind** of data relevant to this item (e.g. height for mountains, capitals for countries, gender for humans)

Value is either a reference to another Item (the capital of Germany is Berlin) or a literal value (e.g. 8848 meters)

Items and Properties

Property

Name _____

Address _____

City _____ **Value** _____ State _____ Zip _____

Home Phone _____ Cell Phone _____

Work Phone _____ Other phone _____

Derivative work of CC-BY-SA 3.0 Unported by Doug Coldwell

An **item** is described by a series of **claims**, each asserting a single datum or fact about it.

A **property** is a specific kind of information, like a field in a form, and is used to describe what the data means.

Examples

An **item** about a **country** would have claims about the following **properties**: population, land area, official languages, borders-with, anthem, capital city, demononyms, Gross Domestic Product, etc. etc.

Examples

An item about a **person** would have claims about: name, date of birth/death, place of birth/death, citizenship, occupation, languages spoken, religion, father, mother, children, notable works, place of education, etc. etc.

Claim examples (Item --> Property --> Value)

Earth --> highest-point --> Mount Everest

Mount Everest --> elevation above sea level -->
8848 meters

Earth --> deepest point --> Challenger Deep

Challenger Deep --> elevation above sea level -->
-10,994±1 metre

Another way of looking at it

item **Earth -->**

property highest-point: value Mount Everest

property deepest point: value Challenger Deep

item **Challenger Deep -->**

property elevation above sea level: value -10,994±1 metre

item **Mount Everest -->**

property elevation above sea level: value 8848 meters

Once more with ~~feeling~~ numeric IDs

Earth (**Q2**) --> highest-point (P610) --> Mount Everest (Q513)

Mount Everest (Q513) --> elevation above sea level (P2044) -->
8848 meters

Earth (Q2) --> deepest point (P1589) --> Challenger Deep
(Q459173)

Challenger Deep (Q459173) --> elevation above sea level
(P2044) --> -10,994±1 metre

Once more without squishy humanspeak

Q2 --> P610 --> Q513

Q513 --> P2044 --> 8848 meters

Q2 --> P1589 --> Q459173

Q459173 --> P2044 --> $-10,994 \pm 1$ meters

Interlude: Why numbers?

Labels are ambiguous and sometimes redundant: What is London?

City in England, in Canada, family name, personal name, movie company, hotel?

Language-neutral - not everybody speaks English!

Robot-friendly (robots love numbers ;)

Questions so far?

Let's go explore Wikidata!

<https://wikidata.org>

Questions?

(also: remember [\[\[Wikidata:Project chat\]\]](#))

Let's teach Wikidata Something new!

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

Let's teach Wikidata
Something new!

<https://www.wikidata.org/wiki/Q131248>

Do I really have to just guess property names?

Nope. Here's a handy property browser:

<https://hay.toolforge.org/propbrowse/>
https://www.wikidata.org/wiki/Wikidata:List_of_properties

But also: Wikidata tries to auto-suggest common ones

**Contributing to Wikidata
on your commute!**

The Wikidata Game

Semi-automates micro-decisions and make incremental improvement of Wikidata super-fun.

<https://wikidata-game.toolforge.org/distributed/>

Embedding Data from Wikidata

{{#property}} and {{#statements}} are magic! :)
[[[en>User:Ijon/Wikidata](#)]]

Article Placeholder

<https://eo.wikipedia.org/w/index.php?search=Helen+DeWitt&title=Speciala%C4%B5o:Ser%C4%89i&go=Ek>

Questions?

Remember:

Wikidata is 

Extra Credit

For when you just can't get enough of Wikidata

Wikidata has an API!

<https://www.wikidata.org/w/api.php>

And people do awesome things with it:

Some examples, in brief

- [Histropedia](#)
- [WikiGenomes](#)
- [Monumental](#)
- [Scholia](#)
- [Crotos](#)
 - [Callisto](#)
- [Inventaire](#)
- [Stadt, Land, Fluss](#)

Querying Wikidata

<https://query.wikidata.org/>

Querying Wikidata

- * query Wikidata using [\[\[SPARQL\]\]](#)
- * Wikidata will tell you **everything it knows**, but **no more**
- * The more data and the more links in Wikidata, the more useful it becomes (so contribute!)
- * [Wikidata Query and SPARQL tutorial](#)

query.wikidata.org demos

Links to sample queries

- [Notable cats](#)
- [Popular surnames among fictional characters](#)
- [Overall causes of death ranking \(presented as bubble chart\)](#)
- [Painters whose fathers were also painters](#)
- Count total articles in Wikipedia language XX by gender:
<http://tinyurl.com/yaklnjism>
- [Largest cities with female mayor](#)
- [Paintings known to Wikidata, by location, on a map](#)
- [African women with articles in English but not in French](#)
 - [Same, but only good/featured articles in English \(i.e. "badges"\)](#)

Remember: “**good artists borrow, great artists steal.**” --attributed to Picasso

WD-FIST

- Free Image Search Tool for Wikidata

<https://tools.wmflabs.org/fist/wdfist/index.html>

The Wikidata Reasonator

- Pretty-format data from Wikidata
- <https://tools.wmflabs.org/reasonator>

Mix and Match

- Match Wikidata entities to external databases' entities

<https://mix-n-match.toolforge.org/>

<https://mix-n-match.toolforge.org/#/catalog/5498>

(coverage of TUBITAK Sosyal Bilimler Ansiklopedisi)

WikiProjects on Wikidata

- Do systematic work on Wikidata around a theme!
- For example:
https://www.wikidata.org/wiki/Wikidata:WikiProject_India
- (none for Singapore yet!)

Go edit Wikidata! :)

Now that you have the power,
use it for good and not for evil.

Video of (almost) this talk:

<https://www.youtube.com/watch?v=eVrAx3AmUvA>
asaf@wikimedia.org