

Wikidata

A Gentle Introduction for Complete Beginners

Asaf Bartov <asaf@wikimedia.org>

WikiConference India 2016, Chandigarh

You are lucky!

Two key problems

1. Dated data; lacking triggers; repetitive work
2. Inflexible ways of lateral queries of knowledge (Categories and their limitations)

~~Two 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 <3

Structured data, you say?

Statement = 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)

Statement examples

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

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

Items and Properties

An **item** is described by a series of **statements**, 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 statements 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 statements 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.

Questions so far?

Let's go explore Wikidata!

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!)

query.wikidata.org demos

Links to sample queries

- [Notable cats](#)
- [Popular surnames among fictional characters](#)
- [Overall causes of death ranking \(presented as bubble chart\)](#)
- [Whose birthday is it today?](#)
- Count total articles in wiki XX by gender: <http://tinyurl.com/j6nedlb> (women) and <http://tinyurl.com/zdvgthf> (men)
- [Largest cities with female mayor](#)

Questions?

**Contributing to Wikidata
on your commute!**

The Wikidata Game

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

<https://tools.wmflabs.org/wikidata-game/distributed>

WD-FIST

- Free Image Search Tool for Wikidata

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

Mix and Match

- Match Wikidata entities to external databases' entities

<https://tools.wmflabs.org/mix-n-match/>

[https://tools.wmflabs.org/mix-n-match/?mode=sit
estats&catalog=92](https://tools.wmflabs.org/mix-n-match/?mode=sit
estats&catalog=92) (coverage of members of the
Royal Society)

Questions?

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

Go edit Wikidata! :)

asaf@wikimedia.org