

Ein schneller Blick auf Wikidata



Anlass:

World-Café

„Wikipedia, Wikimedia Commons,
Wikidata“

Berlin, 27. August 2019

@a_ka_es (Andrea Knabe-Schönemann)

Das Problem:

- Wissen verteilt auf verschiedenen Wikipedia-Ausgaben
- Inhalt ist nicht strukturiert: Faktenwissen ist in Prosa umschrieben; Daten sind nicht leicht zu durchsuchen oder zu extrahieren
- Wikimedia-Projekte können von der Verbindung zu externen Datenbanken profitieren; diese Anbindung ist nicht einfach möglich

Die Lösung:

eine freie Wissensdatenbank für die Summe des menschlichen Wissens

- alle Menschen und Maschinen können beitragen, um strukturierte Daten zu verknüpfen



WIKIDATA

- gestartet am 29. Oktober 2012
- aktuell über 59 Mio. Datenobjekte
- über 20.000 aktive Benutzerinnen und Benutzer mit +1 Bearbeitung im Monat

Stand heute, 27.08.2019, 06:00 Uhr:

Main Page **Discussion**

Read [View source](#) [View history](#) ☆

Search Wikidata



Welcome to Wikidata

the **free** knowledge base with **59,262,559** data **items** that **anyone** can edit.

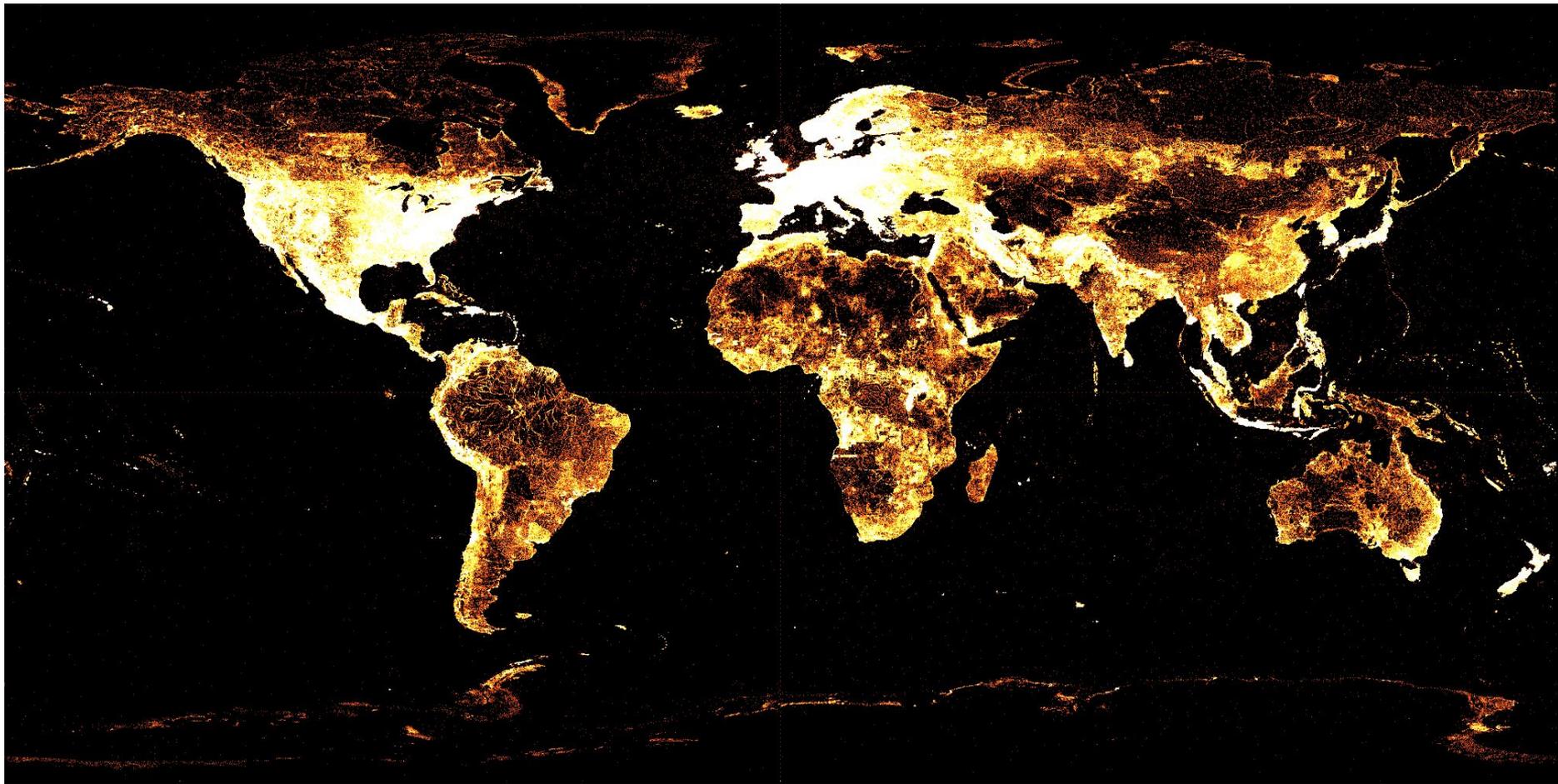
[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

Want to help translate? [Translate the missing messages.](#)

https://www.wikidata.org/wiki/Wikidata:Main_Page

Wikidata ist...

- Ein gemeinsames Repository zur Unterstützung der Wikimedia-Projekte
- Eine freie Wissensdatenbank, die beliebig nachnutzbare Inhalte anbietet
- Ein Hub im Linked Open Data Web, der Inhalte verknüpft



Über 56 Millionen Datenobjekte

Wiki-Idee:

- “wiki” bedeutet im Hawaiianischen “schnell”
- Wikidata ist immer mehr als eine Software, die Kommunikation mit der Community gehört dazu (social software)
- Wiki heißt anfangen und machen — die Struktur entwickelt sich anhand der Erfordernisse
- Ordnung kann sich schnell ändern

Beispiel: zu Beginn 2 Werte für Sex/Gender - heute: mehr als 20

<https://www.wikidata.org/wiki/Property:P21>

Vorteile:

- die neue Wissensbasis kann schnell und effizient abgefragt werden
- die Daten sind über alle Wikimedia-Projekte hinweg konsistent
- neue Datenbank-Tools können entwickelt werden

Datenstruktur



WIKIMEDIA
DEUTSCHLAND

Wissensgegenstände werden nicht in einzelnen Tabellen einer Datenbank abgelegt, sondern miteinander in der Form

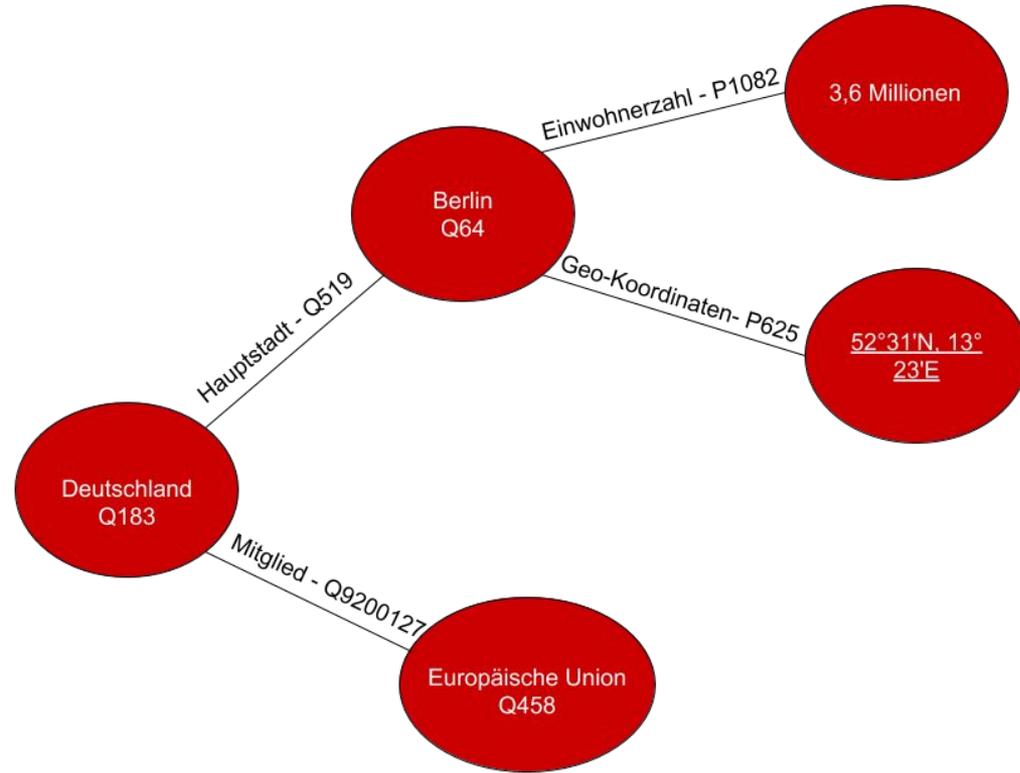
Subjekt - Prädikat - Objekt
("Tripel")

verlinkt. Auf diese Weise entsteht ein Graph bzw. ein Netzwerk.

Die Abfragesprache heißt SPARQL.

Tutorial: https://www.wikidata.org/wiki/Wikidata:SPARQL_tutorial/de
Abfrageservice: <https://query.wikidata.org/>

Beispiel Berlin:



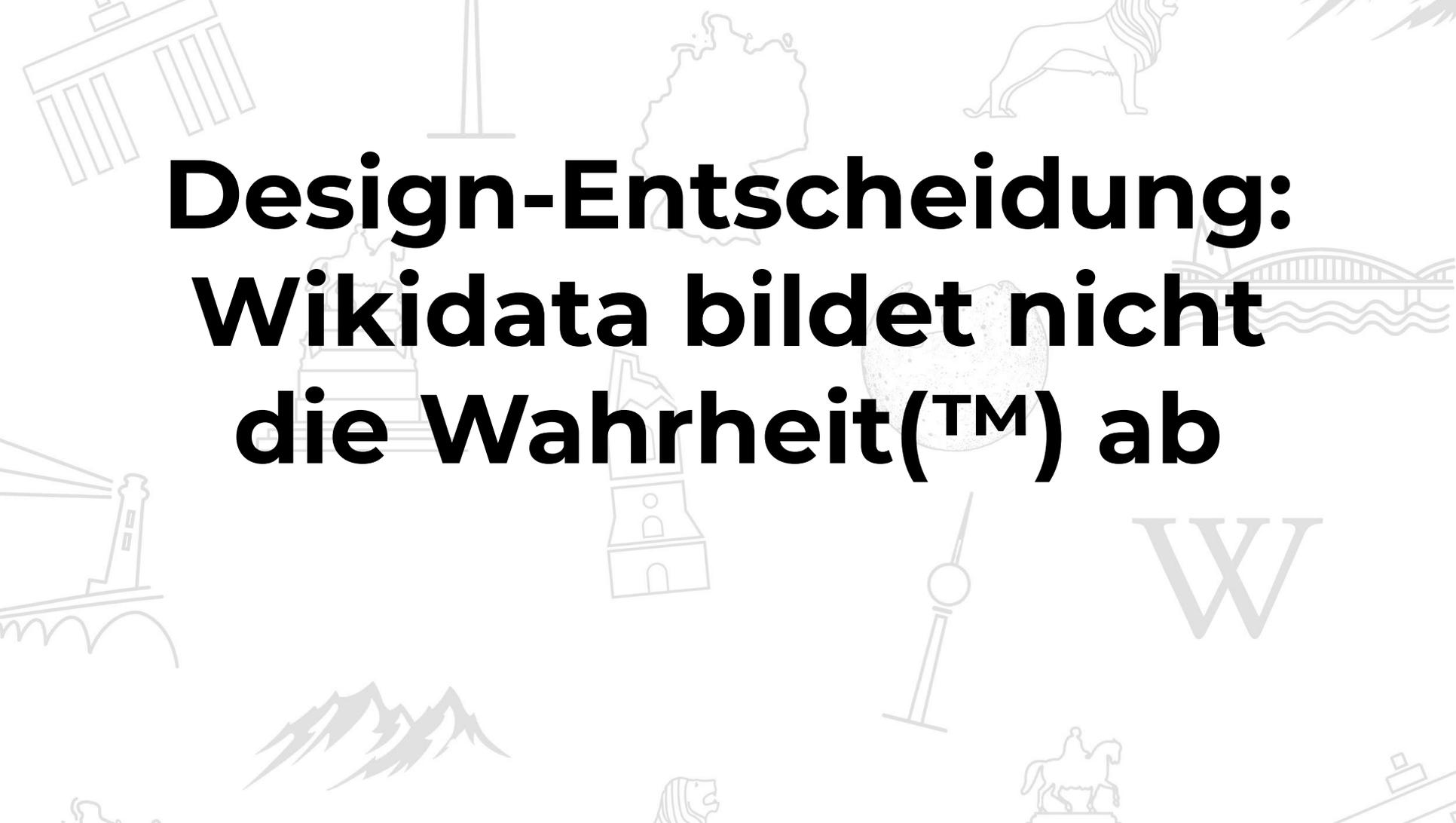


**Design-Entscheidung:
Wissen ist
übersprachlich**

W

Design-Entscheidung: Things, not Strings





**Design-Entscheidung:
Wikidata bildet nicht
die Wahrheit(™) ab**



Inhalte werden verhandelt

- https://www.wikidata.org/wiki/Wikidata:Project_chat
- https://www.wikidata.org/wiki/Wikidata:Property_proposal

- Items kann jeder anlegen
- Properties werden beantragt
- Größere Datenimporte mit der Community absprechen
- Relevant ist, was für Wikimedia-Projekte relevant ist oder was hilft, eine Struktur zu vervollständigen



Wikidata-Eintrag

[Bild "Frauen im Garten" \(Claude Monet\)](#)

- **item**
- **property - Eigenschaft**
- **value - Wert**
- **qualifier - ergänzt den Wert**
- **reference**

Was genau kann man jetzt mit Wikidata (besser) tun?

- Sprachlinks sind überall eingebaut
- [Query-Abfrage über die gesamte Datenbank](#)
- [Datenvisualisierung \(z.B. Timeline im Query-Service\)](#)
- [Histropedia](#) - Beispiel: Zeitleiste aller Bilder von Monet
- [Reasonator - Beispiel "Frauen im Garten" \(Monet\)](#)
- [Infoboxen für Wikipedia - Beispiel "Women in the Garden" \(Monet\)](#)
- [Templates für Wikipedia - Beispiel: Databox](#)

Beispiel Infoboxen für WP

Wikipedia-Seiten für Berlin - Einwohnerzahl in Sprachversionen

2017: 3 613 495 [de](#)

2018: 3 748 148 [en](#)

2017: 3 613 495 [es](#)

2017: 3 613 495 [ar](#)

2018: 3 748 148 [fr](#)

2017: 3 613 495 [hu](#)

2014: 3 437 916 [nl](#)

2017: 3 613 495 [ca](#)

wikidata: 3.613.495 (höchster Rang: Dez. 2017, DESTATIS)

Keine Idee, wie anfangen?

- [crotos - eine Suchmaschine für Kunstwerke](#)
- [WikiProjekt "Alle Gemälde"](#)
- [Strukturierte Daten für Wikimedia Commons](#)
- [ISA - Tool zum Beitragen von strukturierten Daten](#)
- [Wikidata - The Game](#)
- [Nearby](#)

Oder einen tollen Satz Offener Daten mit einem Tool (z. B. OpenRefine) aufbereiten und im Stapel uploaden...

Sprechen Sie mich an. :)

