

WikiCite für Bibliotheken

Jakob Voß (VZG)



26. Oktober 2018, SUB Göttingen



Hintergrund

Begriffsklärung

- Wikipedia Enzyklopädie in Form eines offenen Wikis
- Wikidata Datenbank in Form eines Wikis
- Wikimedia Organisation & Community
- Wikibase Erweiterung der Wiki-Software für Wikidata
- WikiCite ???

Wikidata

- ▶ Offene Datenbank in Form eines Wikis
- ▶ Geburtstag am 29. Oktober 2012!
- ▶ Wissensdatenbank und Normdatei für Wikidata u.v.a.m.
- ▶ Basiert auf Software Wikibase

Beispieldatensatz: Q564783 (SUB Göttingen)

Wikibase-Datenmodell

- ▶ Item-Identifizier
 - ▶ Q30000000
- ▶ Eigenschaften
 - ▶ P31 *ist ein(e) / instance of / est ...*
 - ▶ P50 *Autor / author / scriptor ...*
 - ▶ ...
- ▶ Qualifikatoren & Belege
- ▶ Constraints als Aussagen über Eigenschaften (optional!)

Bezeichnung:

ID:

Beschreibung:

The Synergistic Activity of Thyroid Transcription Factor 1 and Pax 8 Relies

(Q300000000)...

wissenschaftlicher Artikel aus dem Jahr 2002

» In weiteren Sprachen

mehrsprachig

Bearbeitungsmöglichkeiten

Aussagen

1. Aussage:

ist ein(e)



wissenschaftlicher Artikel ...

» 0 Fundstellen



bearbeiten



Fundstelle hinzufügen



Wert hinzufügen

2. Aussage:

Titel



The Synergistic Activity of Thyroid Transcription Factor 1 and Pax 8 Relies on the Promoter/Enhancer Interplay (Englisch)

» eine Fundstelle



bearbeiten



Wert hinzufügen

Fundstellen für Aussagen

3. Aussage:

Schlagwort



Transkription ...

» eine Fundstelle



bearbeiten



Wert hinzufügen

...

Wikibase

- ▶ Wikisoftware für Wissensdatenbanken, insbesondere Wikidata
- ▶ Ontologie (Eigenschaften & Constraints) als Teil des Wiki
- ▶ APIs und Programmbibliotheken
- ▶ Zunehmend auch zur Erfassung in GLAM-Einrichtungen
 - ▶ FactGrid (Geschichtswissenschaft, Erfurt), siehe Blogartikel
 - ▶ Rhizome (Digital Preservation, New York)
 - ▶ Workshop Authority Control meets Wikibase, 29. Oktober 2018, DNB

Warum WikiCite?

Was ist WikiCite?

<http://www.wikidata.org/entity/Q30035267>

ist ein(e)	 Projekt  bearbeiten
	▼ 0 Fundstelle
	 Konferenzreihe  bearbeiten
	▼ 0 Fundstelle
	 Online-Community  bearbeiten
	▼ 0 Fundstelle
+ Wert hinzufügen	

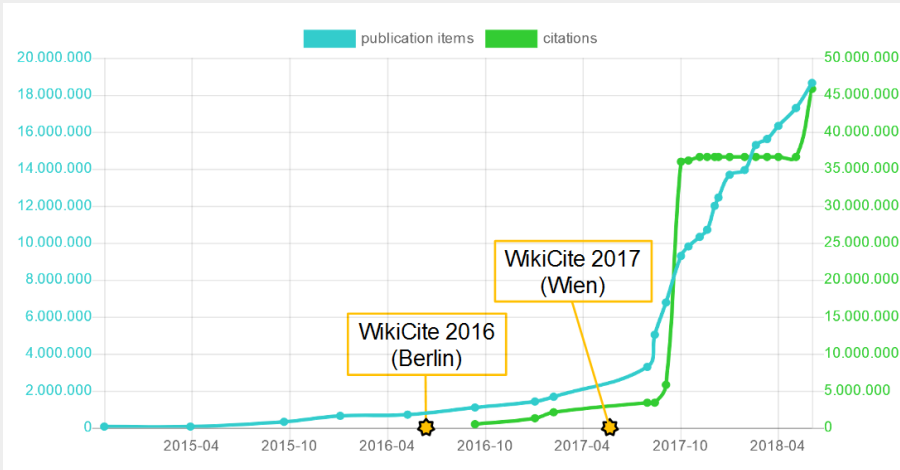
Ziele von WikiCite

- ▶ Freie bibliographische Datenbank in Wikidata
 - ▶ Bibliographische Metadaten
 - ▶ Zitationsdaten
- ▶ Selbstständige + unselbstständiger Werke + Beziehungen
 - ▶ Kollaborative Katalogisierung (Social Cataloging)
 - ▶ offene Daten (CC0)

<http://wikicite.org/>

Umfang

- ▶ **Publikationen:** 19 Millionen Wikidata-Einträgen (38%)
- ▶ **Zitationen:** 46 120 Millionen Zitationen in Wikidata



Motivation für Wikipedia

► Aussagen mit Belegen versehen, z.B.

Berlin gilt als **Weltstadt** der Kultur, Politik, Medien und Wissenschaften.^{[15][16][17][18]} Die Universitäten, Forschungseinrichtungen, Sportereignisse und Museen Berlins genießen internationalen Ruf.^[19] Die **Metropole** trägt den UNESCO-Titel *Stadt des Designs* und ist eines der meistbesuchten Zentren des Kontinents.^[20] Berlins Architektur, Festivals, Nachtleben und vielfältige Lebensbedingungen sind weltweit bekannt.^[21]

15. ↑ [Global Power City Index 2015](#) (PDF; 1,2 MB) Abgerufen am 10. April 2016.
16. ↑ [Revealed: Cities that rule the world – and those on the rise.](#) (englisch), abgerufen am 14. April 2010.
17. ↑ [Leading cities in cultural globalisations/Media](#) (englisch), GaWC, abgerufen am 14. April 2010.
18. ↑ [Berlin – Stadt der Nobelpreisträger.](#) In: *luisse-berlin.de*. Abgerufen am 3. September 2015.
19. ↑ [London crowned as SportBusiness' 2012 Ultimate Sports City](#) (Memento vom 18. August 2012 im Internet Archive). In: *sportsbusiness.com*, abgerufen am 26. November 2012.
20. ↑ [Tourismusreport 2014.](#) IHK Berlin.
21. ↑ [Berlin, the Big Canvas.](#) In: *nytimes.com*. *The New York Times*, abgerufen am 16. April 2010 (englisch).

► Bibl. Metadaten zentral für alle Artikel und Sprachversionen

Motivation für Wikidata

- ▶ Fakten in Wikidata mit Fundstellen belegen, z.B. höchster Punkt in Berlin

höchster Punkt

 Arkenberge ...

 bearbeiten

▼ eine Fundstelle

 kopieren

URL der Fundstelle	http://www.tagesspiegel.de/berlin/teufelsberg-oder-arkenberge-zum-wettstreit-um-den-hoechsten-gipfel-berlins/11413932.html
Titel	Zum Wettstreit um den höchsten Gipfel Berlins (Deutsch)

+ Fundstelle hinzufügen

+ Wert hinzufügen

Motivation für die Wissenschaft

► Auswertung und Visualisierung von Publikationsdaten

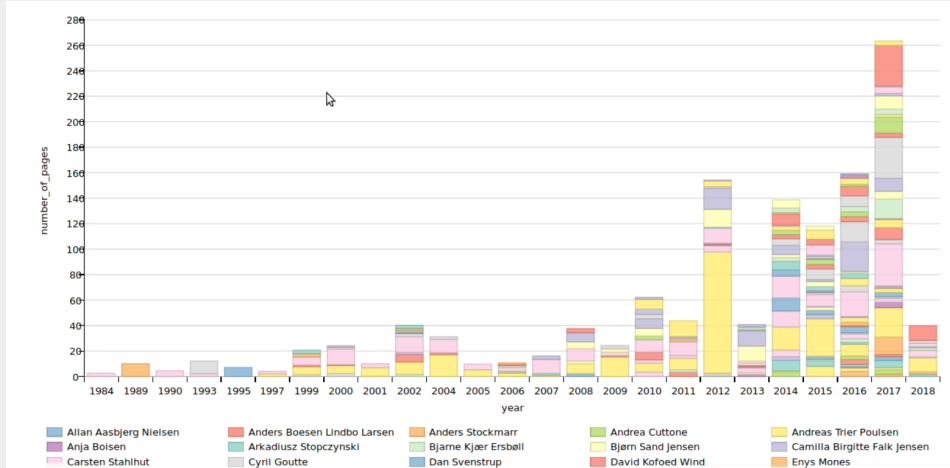
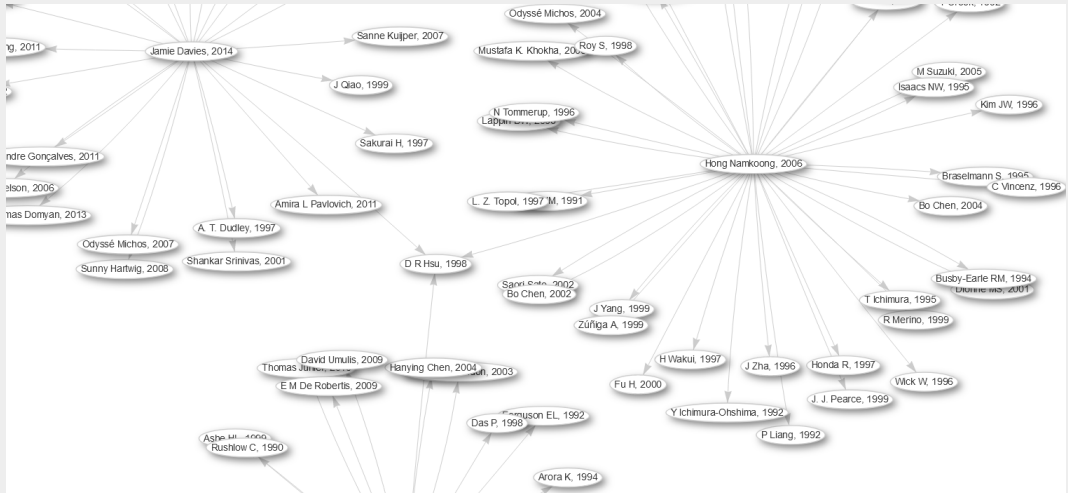


Abb. Screenshot von Scholia zur Anzahl produzierter Seiten von Publikationen mit Autoren der DTU

Motivation für die Wissenschaft

Auswertung und Visualisierung von Zitationsdaten



Vorzeige-Projekt

Scholia: <https://tools.wmflabs.org/scholia/>

Umsetzung von WikiCite

Community

- ▶ Mailingliste, Diskussionsseiten, Konferenzen...
- ▶ Direkte Aktion
- ▶ Diskussion von Modellierungsfragen

Werkzeuge

- ▶ **sourceMD, fatameh...** Import anhand DOI/PMID
- ▶ **zotkat** Import aus Zotero (Literaturverwaltung)
- ▶ **citation-js** Export aus Wikidata
- ▶ **Mix'n'Match** Abgleich mit externen Identifikatoren
- ▶ **OpenRefine 3.0** mit erweiterter Wikidata-Anbindung
- ▶ ...

Normdaten-Verlinkung

- ▶ Wikidata ist selbst eine universelle Normdatei
- ▶ Verknüpfungen mit externen Einträgen
 - ▶ DOI, ISBN, PMID, arXiv-ID, ... (Publikationen)
 - ▶ GND, VIAF, ORCID, dblp-ID, ... (Autoren...)
 - ▶ ...
- ▶ etwa 2/3 (3150) aller Wikidata-Eigenschaften dienen der Verlinkung mit anderen Datenbanken

Qualitätssicherung

- ▶ Eigene Beobachtungslisten für Wikidata-Einträge mit Benachrichtigung bei Änderungen
- ▶ Einschränkungen von Werten von Eigenschaften
 - ▶ *ISBN* (P212/P957) folgen bestimmten Mustern
 - ▶ *Autoren* (P50) müssen Personen oder Organisation sein
- ▶ Regelverletzungen werden protokolliert (constraint reports) und können dann verbessert werden
- ▶ bottom-up statt top-down!

Abfragemöglichkeiten

- ▶ <https://query.wikidata.org/> (SPARQL-Editor)
- ▶ Beispielabfragen
 - ▶ Bücher zusammen von Elter(n) und Kinder(n) geschrieben:
<http://tinyurl.com/y7oe8uo8>
 - ▶ Autoren, deren Werke 2018 gemeinfrei wurden:
<http://tinyurl.com/yc8v6l9c>
 - ▶ Meist zitierte Wissenschaftlerinnen: <http://tinyurl.com/y8ell8nw>
- ▶ Hilfreich zur Qualitätssicherung

Perspektiven für Bibliotheken

Nachnutzung von Daten aus Wikidata

- ▶ Kataloganreicherungen bei Normdaten
- ▶ mehrsprachige Suche nach Schlagwörtern
- ▶ Artikelindex für Discovery aus Wikidata?
- ▶ Kataloganreicherungen mit offenen Zitationsdaten?

Ideen für Katalogisierungstools

- ▶ Ansätze zur Datenerfassen in Wikidata für die Katalogisierung in Bibliotheken (siehe DNB-Veranstaltung zu Wikibase & GND)
- ▶ Tools zur Batch-Bearbeitungen
- ▶ Kontroll- und Prüftools
- ▶ Oder ganz auf Wikidata/Wikibase umsteigen?

These: WikiCite ersetzt Bibliographische Datenbanken

Indiz: Wikipedia hat Nachschlagewerke ersetzt

Argument: WikiCite ist schneller

	GBV/CBS	Wikidata/Wikibase
Mit-Katalogisieren	Monate bis Jahre	Minuten
Neue Kategorie	Monate bis Jahre	Tage bis Wochen
Gesamtabzug	Wochen bis Monate	Tage bis Wochen
Tools nutzen	?	Minuten bis Stunden

Qualitätssicherung in WikiCite

- ▶ Katalogisierungsregeln im Entstehen
- ▶ Absehbar ebenso ausgefeilt wie bibliothekarische Regelwerke
- ▶ Regeln werden nicht **vorgeschrieben** sondern **nach**geprüft

These: Qualitätssicherung in WikiCite ist nachhaltiger

Bibliotheken	Wikidata
Eingangskontrolle	Beständige Änderungen
Beschränkungen	Datenanalyse & Korrekturen
Was soll rein?	Was ist drin?
Top-Down	Bottom-Up

Fragen & Diskussion