

# WikiCite als kollaborativer Ansatz für offene Zitationsdaten

Vortrag auf dem Bibliothek\*tag Berlin

Philipp Zumstein (UB Mannheim)    Jakob Voß (VZG)

2018-06-13









{ } wikicite

# Einführung & Motivation

# Was ist WikiCite?

- Kontext: Wikipedia, **Wikidata**, Wikisource. . .

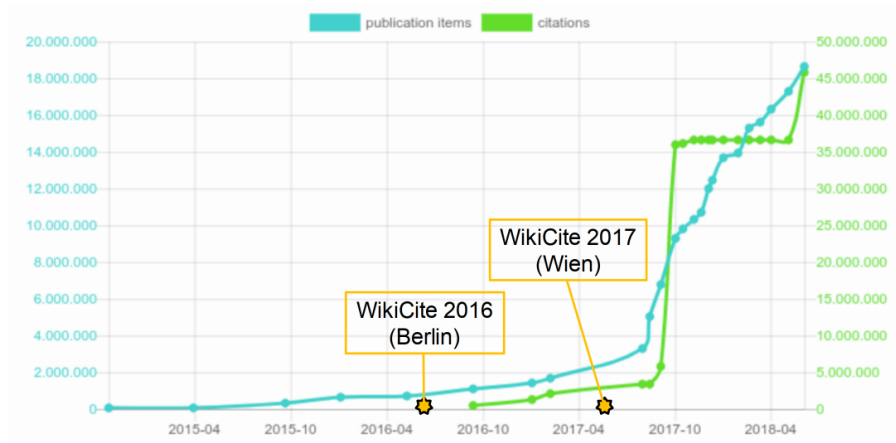
ist ein(e)	 Projekt  bearbeiten
	▼ 0 Fundstelle
	 Konferenzreihe  bearbeiten
	▼ 0 Fundstelle
	 Online-Community  bearbeiten
	▼ 0 Fundstelle
+ Wert hinzufügen	

# Ziele von WikiCite

- ▶ Freie bibliographische Datenbank in Wikidata
  - ▶ Bibliographische Metadaten
  - ▶ Zitationsdaten
- ▶ Katalogisierung selbstständiger + unselbständiger Werke + Beziehungen
  - ▶ kollaborativ (Social Cataloging)
  - ▶ offene Daten (CC0)

# Umfang

- ▶ 17/48 Millionen Wikidata-Einträgen zu Publikationen
- ▶ 36/440 Millionen Aussagen in Wikidata sind Zitationen



# Motivation für Wikipedia

- ▶ Aussagen mit Belegen versehen
- ▶ Bibliographische Metadaten zentral für alle Artikel und Sprachversionen

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

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

# Motivation für Wikidata

## ► Fakten in Wikidata mit Fundstellen belegen

höchster Punkt

 Arkenberge ...

 bearbeiten

▼ eine Fundstelle

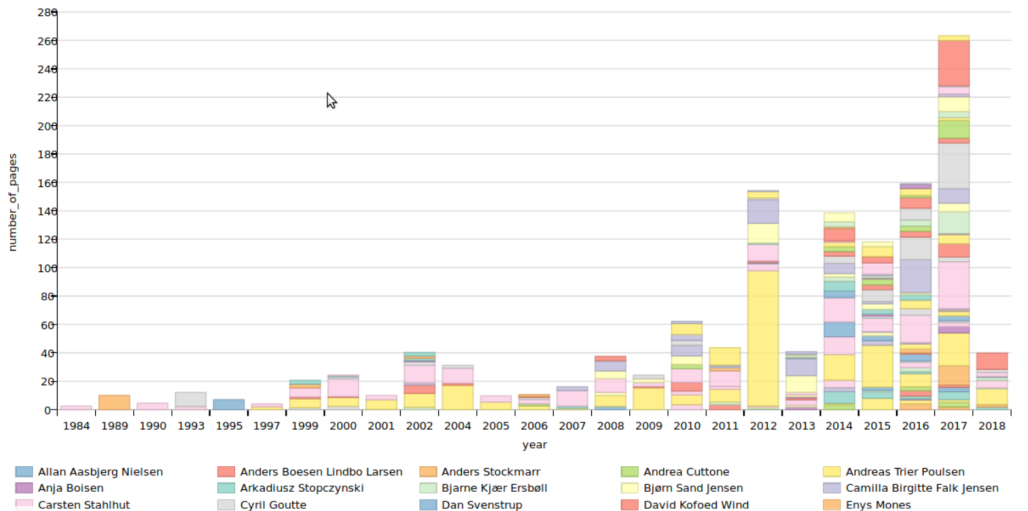
 kopieren

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

+ Fundstelle hinzufügen

+ Wert hinzufügen

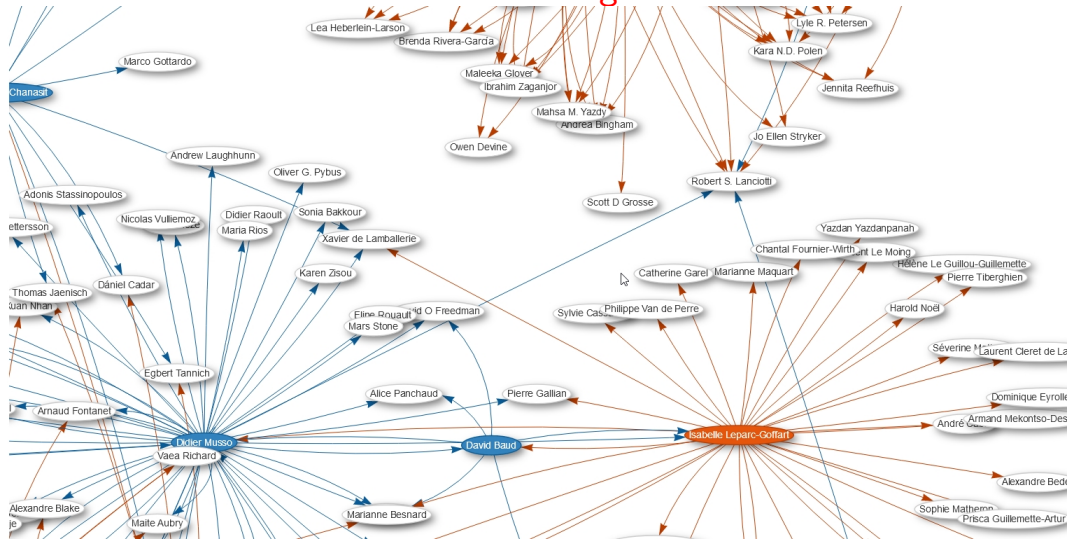
# Motivation für bibliometrische Auswertungen



Auswertung und Visualisierung von **Publikationsdaten**



# Motivation für bibliometrische Auswertungen



Auswertung und Visualisierung von **Zitationsdaten**

# Umsetzung von WikiCite

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

**ID:** (Q30000000)\*\*\*

**Beschreibung:** 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

...

# Wikidata-Datenmodell

- ▶ Item-Identifizier
  - ▶ <http://www.wikidata.org/entity/Q30000000>
- ▶ Eigenschaften
  - ▶ **P31** *ist ein(e) / instance of / est ...*
  - ▶ **P50** *Autor / author / scriptor ...*
- ▶ Datenmodell ist dynamisch
  - ▶ Neue Eigenschaften vorschlagen, diskutieren und erstellen

## Tools zur Bearbeitung

- ▶ Direkte Bearbeitung auf <https://wikidata.org/>
- ▶ **sourceMD/fatameh** Import von Metadaten anhand DOI/Pubmed-ID
- ▶ **zotkat** Import aus Zotero (Literaturverwaltung)
- ▶ **Mix'n'Match** Abgleich von Normdaten mit externen Identifikatoren
- ▶ **OpenRefine 3.0** mit erweiterter Wikidata-Anbindung
- ▶ ...

# Qualitätssicherung

- ▶ Eigene Beobachtungslist für Wikidata-Einträge
  - ▶ Benachrichtigung bei Änderungen
- ▶ Einschränkungen von Werten von Eigenschaften
  - ▶ ISBN (P.../P...) folgen bestimmten Muster
  - ▶ Autoren (P50) müssen Personen oder Organisation sein
- ▶ Regelverletzungen werden protokolliert (constraint reports) und können dann verbessert werden

# Normdaten-Verlinkung

- ▶ Wikidata ist selbst eine universelle Normdatei
- ▶ Verknüpfungen mit externen Einträgen
  - ▶ DOI, ISBN, PMID, arXiv-ID, ... (Publikationen)
  - ▶ GND, VIAF, ORCID, dblp-ID, ... (Autoren, Organisationen...)
  - ▶ ... (Orte, Substanzen, Ereignisse...)
- ▶ 2824/4741 Wikidata-Eigenschaften dienen der Verlinkung mit anderen Datenbanken

# 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>>



# Perspektiven für Bibliotheken

# Bibliotheksdaten in Wikidata

- ▶ Wie können Bibliotheksdaten für Wikidata genutzt werden?
  - ▶ Verknüpfungen über externe Verweise?
  - ▶ Exportformate, APIs bereitstellen?
- ▶ sich bei Datenmodellierung einbringen?

# Katalogisierungstools

- ▶ Kann man Ansätze zum Datenerfassen in Wikidata auch bei der Katalogisierung in Bibliotheken adaptieren?
- ▶ Sind nicht auch stärker Tools für Bearbeitungen im Batch gefragt?
- ▶ Was ist mit Kontroll- und Prüftools?

# Umfang der Katalogisierung

- ▶ Aufsatzkatalogisierung
  - ▶ Thema für Bibliotheken? FIDs?
  - ▶ WikiCite zeigt auf, dass es möglich ist
- ▶ Katalogisierung von Zitationen
  - ▶ Können Bibliotheken zu offenen Zitationsdaten beitragen?
  - ▶ Zitationsdaten in Verbundkatalog speichern?
  - ▶ Linked Open Citation Database (LOC-DB)

Fragen / Diskussion / Antworten ?

Weiterführende Links