



Die Datenlaube: **Neues Wissen und Daten aus alten Texten** – Mit Wikisource, Wikidata und mit Commons

Jens Bemme: [056880673](#) (SLUB Dresden)

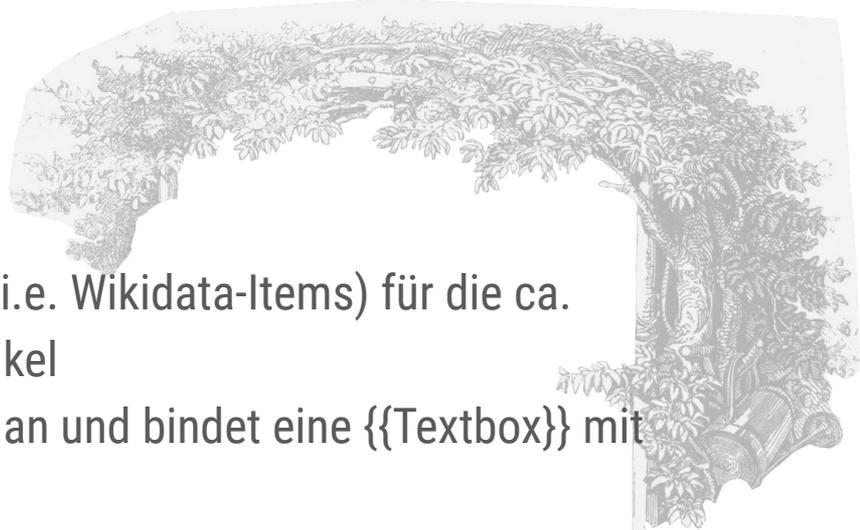
Christian Erlinger: [067173261](#) (Stadt Wien Büchereien)

Die Gartenlaube : Since 1853



- Vorläuferin der modernen Illustrierten
- erstes erfolgreiches deutsches Massenblatt
- ab 1853 in Leipzig im Verlag Ernst Keil
 - https://de.wikipedia.org/wiki/Die_Gartenlaube
 - https://de.wikisource.org/wiki/Die_Gartenlaube
 - <https://www.wikidata.org/wiki/Q655617>
- bunter Themenmix, Serien, kleine und große Illustrationen, ...
- *Die Gartenlaube* in Wikisource: Großprojekt seit 2008, Transkription und OCR-Korrektur aller Artikel bis 1899, 4-Augen-Prinzip
 - Anspruch von Wikisource: Quellensammlung für die Forschung, hohe Qualität!

[[Wikisource+Wikidata]]

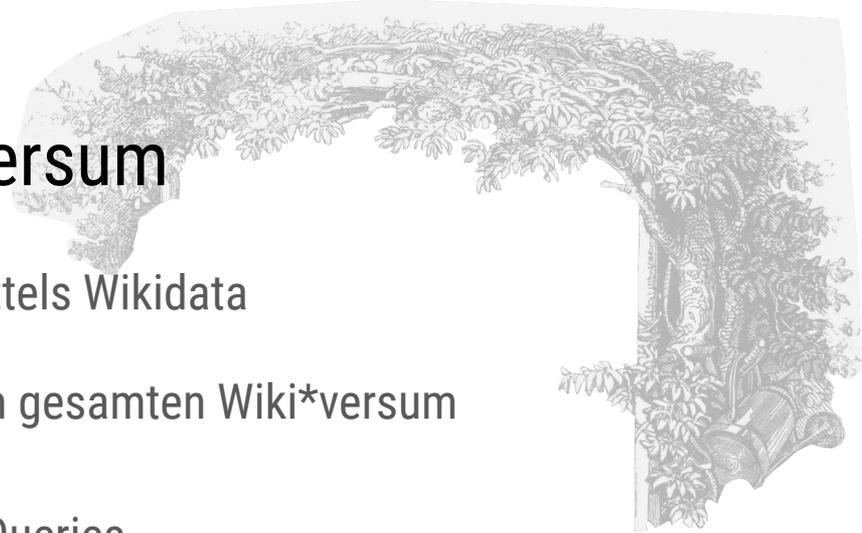


- Erstellung von bibliographischen Datensätzen (i.e. Wikidata-Items) für die ca. 13.000 bisher in Wikisource transkribierten Artikel
- Wikisource-Community legt Artikel mit Volltext an und bindet eine {{Textbox}} mit Metadaten ein
- Wikidata-Import
 - Extraktion der Daten aus der Wikisource-{{Textbox}} mittels Python-Skript
 - Import mittels QuickStatements
 - <https://github.com/DieDatenlaube>
- Wikisource-Artikel werden semantisch und graphenbasiert aus Wikidata abfragbar
- Diese Methoden sind auf andere Textsammlungen übertragbar: #DieDatenlaube i.w.S.

Nutzen der Datenlaube im Wiki*versum

Detaillierte Erschließung der Gartenlaube-Artikel mittels Wikidata

- vernetzt die Texte und Bilder vielfältig Weise im gesamten Wiki*versum
- strukturierte Metadaten
 - Abfrage und Visualisierung mit Wikidata-Queries
 - größere Sichtbarkeit der Wikisource-Ressourcen (SEO)
 - Verknüpfbarkeit (Linked Open Data)
- Die Gartenlaube = Trainingsplatz und Quelle für
 - [[Digital Humanities+Citizen Science]]
 - OER, ...
- Wikidata ist auf diese Weise ein Bibliothekskatalog von Wikisource



Formale Erschließung in Wikidata



- **rdfs:label** – Titel und Untertitel
- **schema:description** – (mehrsprachige Beschreibung des Wikidata-Items)
 - de: “Artikel in: Die Gartenlaube, JJJJ, Heft Nr.”
 - nl: “redactioneel artikel in Die Gartenlaube (1886), nr. 13”
- **rdfs:statement (“Kernelemente”)**
 - P31 : “dieses Item ist” -> ein Zeitschriftenartikel (verlinkt mit [Q191067](#))
 - P1476 : “Title” -> Titel des Artikels (Text-String)
 - P50 : “Autor*in” -> Verfasser*in des Artikel (verlinkt mit Personen-Item)
 - alternativ: P2093 : “Autor*innen-Name” -> (Text-String)
 - P577 : “Publikationsdatum” -> Veröffentlichungsdatum (Datum)
 - P1433 : “veröffentlicht in” -> Die Gartenlaube (verlinkt [Q655617](#))
 - P304 : “Seite(n)” -> Angabe des Seitenbereichs (Text-String)
 - P433 : “Heft” -> Angabe der Heftnummer (Text-String)
 - P407 : “Sprache” -> Sprache des Artikels “Deutsch” (verlinkt mit [Q188](#))
 - P6216 : “Copyright Status” -> public domain (verlinkt mit [Q19652](#))
- **rdfs:statement (“Erweiterte Elemente”)** - First line, Image, Audio File, Genre, ...

Inhaltliche Erschließung in Wikidata



main subject

 Franz Joseph I of Austria ...	 edit	
▼ 0 references	+ add reference	
 Berlin ...	 edit	
▼ 0 references	+ add reference	
 visit	 edit	
of	Berlin	
start time	12 August 1889 <i>Gregorian</i>	
end time	15 August 1889 <i>Gregorian</i>	
participant	Franz Joseph I of Austria	
▼ 0 references	+ add reference	
	+ add value	

Beschlagwortung des Artikels "Der Besuch des Kaiser Franz Josephs I. in Berlin. Die Gartenlaube (1889) Nr. 36, S. 605–607.

<https://www.wikidata.org/wiki/Q81253766>

Erschließung der Illustrationen



- Wikisource Community stellt Illustrationen frei und bettet sie in den Text ein.
 - Auf Wikimedia Commons bestehen somit nicht nur die vollen Seitenscans, sondern die Illustrationen als selbständige Bilddateien
- Einbindung in das Artikel-Item
 - Statement: P18 (“image”) verlinkt direkt mit WikiCommons
 - Qualifier P2086 (“media legend”) - Original Bildbeschreibung (Text-String)
 - Qualifier P180 (“abgebildet”) - verlinkte inhaltliche Beschreibung des Bildes
- Beschreibung mit “Structured Data on Commons”
 - depicts-Statement - mit Wikidata verlinkte inhaltliche Beschreibung
 - “published in” Statement - mit dem bibliographischen Wikidata-Item des Artikels verlinken (oder exakte Fundstelle angeben, ganzseitige Illustrationen oft an völlig anderer Stelle)
 - Verknüpfung mit Originalwerken
 - Frage des Datenmodells: Gartenlaube-Illustrationen sind selbst schon Reproduktionen von Werken

Wirkung : Recherche



- **Verbesserung der Auffindbarkeit (Wege aus der “Krise der Auffindbarkeit” [Open Knowledge Maps])**
- **WikiSource - klassische MediaWiki Suche**
 - Volltext
 - Kategorien
 - Manuell kuratierte Autoren-/Orte-/Themenseiten und Inhaltsverzeichnisse und Artikelserien
- **Linked Data aus Wikidata**
 - Vernetzung der Autoren und Themen über alle in Wikidata vorhandenen Eigenschaften
 - Vernetzung mit der Vielfalt an externen Identifikatoren
- **“Federation”**
 - Semantische Abfrage der Metadaten in Wikidata
 - + semantische Abfrage der Illustrationen in Wikimedia Commons
 - + Suche nach Volltextelementen via MediaWiki API in Wikisource

Wirkung : Benutzung und neue Forschungsfragen

- Deutschlands merkwürdige Bäume (Serie)
 - Erschließung, Query-Visualisierung mit Karte der Standorte im Deutschen Reich
 - Saxonica (Dr. Solvejg Nitzke, 2019): <https://saxorum.hypotheses.org/2396>
- OER (HistoDigitale/ Uni Leipzig)
 - <https://de.wikisource.org/wiki/Wikisource:OER>
- Aufbau eines respektablen Korpus aus Volltexten, extrahierten Illustrationen und Linked Open Data Metadaten als Basis für Forschung der DH und als Prototyp zur Realisierung ähnlicher Projekte



Wirkung : Karte aller merkwürdigen Bäume



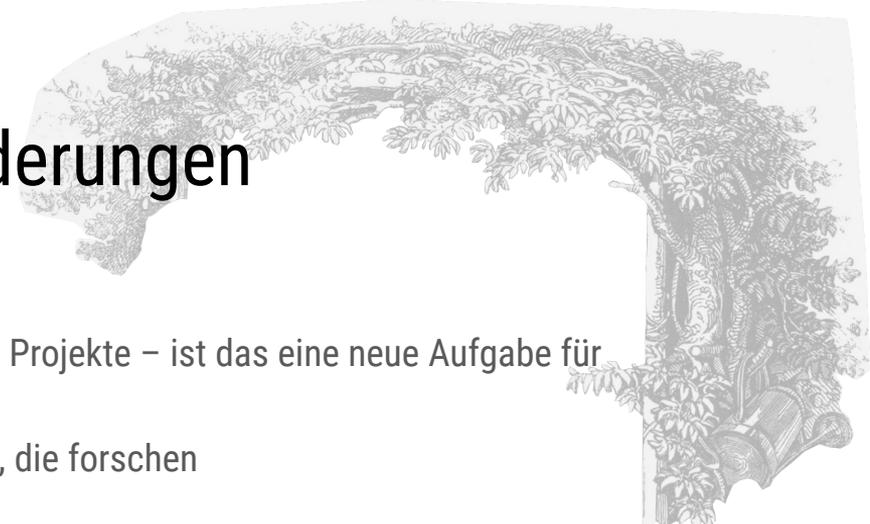
Wirkung : Vernetzung



- Material für zeitgenössische Themen
 - Historische Texte nutzen für aktuelle Ereignisse
 - Infektionskrankheiten im 19. Jahrhundert: w.wiki/Kim
 - Historische Texte zu (anderen) historischen Texten (z.B. Gartenlaube-Biographien + BLKÖ)
- Social Media
 - #DieDatenlaube-Graph - Abfragen und neue Daten vorstellen. Forschung verknüpfen
 - Wikidata-Metadaten liefern OpenGraph-Infos (-> Bibliothekskataloge aufgepasst!)
- International
 - DACH-EU (Bemme, Erlinger et al.) – Wikimedia-DH-Community
 - ¿ *Wikisource Conference 2020* ? im Oktober in Warschau

[https://meta.wikimedia.org/wiki/Grants:Conference/ ... /Wikisource_Conference_2020](https://meta.wikimedia.org/wiki/Grants:Conference/.../Wikisource_Conference_2020)

Ausblick / Aufgaben / Herausforderungen



- Offene Kulturdaten für alle
 - Erschließung und Begleitung digitaler Citizen Science Projekte – ist das eine neue Aufgabe für (Landes-)Bibliotheken?
 - Open Citizen Science: open GLAM <+++> Bürger*innen, die forschen
- Wissen – Können – Machen (Bibliotheksaufgabe)
 - Vermittlung bibliothekarischen Metzhodenwissens für *Citizen Scientists* und Crowdsourcing
 - Begleitung und Aufbau engagierter Communities
- Bautrupp *Digital Humanities* (Forderungen, Wünsche, Möglichkeiten)
 - Anleitungen und Wissenstransfer, um Datengrundlagen für digitale Forschung zu erweitern
 - Wikidata als zentralen bibliographischen Speicher im Wiki*versum etablieren. Anstelle von Copy-Cataloging direkte Bearbeitung der Metadaten aus Wikisource

Referenzen

Zum Weiterklicken

- <https://diedatenlaube.github.io/> #DieDatenlaube Blog
- https://de.wikisource.org/wiki/Wikisource:Wikidata#Die_Gartenlaube Abfragen, Links, Diskussionen

Zum Weiterlesen

- Bemme, Jens 2019. Hilfe für die Datenlaube: mit [[Wikisource+Wikidata]] die freie Quellensammlung verbessern Wikimedia Deutschland Blog. <https://blog.wikimedia.de/2019/10/16/> [Stand 2020-05-15].
- Bemme, Jens & Erlinger, Christian 2019. Die Datenlaube der Gartenlaube diedatenlaube.github.io. https://diedatenlaube.github.io/die_datenlaube_der_gartenlaube.html [Stand 2020-05-15].
- Erlinger, Christian 2020a. Beschreiben wir Bilder punktgenau! diedatenlaube.github.io. https://diedatenlaube.github.io/beschreiben_wir_bilder_punktgenau.html [Stand 2020-05-14].
- Erlinger, Christian 2020b. „...Fortsetzung folgt.“ - Metadaten für Artikel der Gartenlaube mit geteilter Erscheinungsweise diedatenlaube.github.io. https://diedatenlaube.github.io/fortsetzung_folgt_seitenzahlen.html [Stand 2020-05-15].
- Erlinger, Christian 2020c. Linked Data für Bildrepositorien. Structured Data on [Wiki] Commons. Tagung Bildarchive 2020. ISGV Dresden. <https://bildarchive.isgv.de/panels/bilder-und-digitalisierung/christian-erlinger/> [Stand 2020-05-15].
- Gill Satdeep 2019. How can Structured Data on Commons, Wikidata, and Wikisource walk hand in hand? A pilot project with Punjabi Qisse - Projects - Wikimedia Space. <https://discuss-space.wmflabs.org/> [Stand 2020-05-14].
- Munke, Martin & Bemme, Jens 2019. Bürgerwissenschaften in wissenschaftlichen Bibliotheken: o-bib. Das offene Bibliotheksjournal / Herausgeber VDB 6, 4, 178–203. [doi:10.5282/o-bib/2019H4S178-203](https://doi.org/10.5282/o-bib/2019H4S178-203)
- Munke, Martin & Bemme, Jens 2019. Offene Daten und die Zukunft der Bürgerforschung in Wissenschaftlichen Bibliotheken. <https://osf.io/qhrc4> [Stand 2020-05-15].



Queries - neue Wege der Auffindbarkeit

Queries: de.wikisource.org/wiki/Wikisource:Wikidata#Abfragen (Auswahl)

- Abfrage aller Artikel nach Jahrgang und Heft: w.wiki/RXg
- Artikel ohne Schlagwort: w.wiki/43s
- Karte der Geburtsorte der Autoren der Gartenlaube: w.wiki/RXi
- Artikel relevant für die "Sächsische Bibliografie", optional mit K10plus-Editions ID: w.wiki/Ry7
- Frauenporträts in der Gartenlaube: w.wiki/JyU
- Bildgalerie der Gartenlaube Autorinnen: w.wiki/Jye
- Zeitstrahl der Reiseberichte Alfred Brehms: w.wiki/eqH
- Datenlaube-Artikel zu Infektionskrankheiten: w.wiki/Kim

