

Abstract

Die Datenlaube ist ein internationales Citizen Science-Projekt, mit dem die ehrenamtlich transkribierten Volltexte der Illustrierten *Die Gartenlaube* erschlossen werden. Das Transkriptionsprojekt für *Die Gartenlaube* (derzeit Jahrgänge 1853 bis 1899) der deutschsprachigen Wikisource profitiert auf diese Weise von strukturierten bibliografischen Metadaten in Wikidata. Die Autoren berichten von der Projektarbeit, Zielen und Projektgrundsätzen.

Die Gartenlaube und Die Datenlaube – Erschließung von Transkriptionen in Wikisource mit Wikidata

Jens Bemme, Christian Erlinger, Matthias Erfurth

Die Datenlaube ist eine digitale Citizen Science-Initiative für und mit offenen Kulturdaten. Das Projektteam erschließt und verlinkt mittels Wikidata *Die Gartenlaube*, eine Illustrierte aus Leipzig, die in Wikisource transkribiert und korrigiert wird.

Grundlage des Projekts sind die digitalen Quellentexte und reichhaltigen Illustrationen, die von der deutschsprachigen Wikisource-Gemeinschaft seit 2005 erstellt und verknüpft werden: Transkriptionen und OCR-Korrekturen gemeinfreier Werke. Wikidata fungiert dabei als Katalog für Wikisourcetexte und gescannte Illustrationen. Als strukturiertes Datenrückgrat der Wikimediafamilie ist Wikidata zugleich das Datenhub dieser digitalisierten Wissensbestände mit externen, verknüpfbaren Datenquellen als Teil des *Webs of Linked Open Data* (LOD).

Das Projekt *Die Datenlaube* startete Anfang 2019. Inzwischen besteht das internationale Team aus drei Mitgliedern: Christian Erlinger (Wien/Luzern), Matthias Erfurth und Jens Bemme (beide in Dresden). Jede und jeder kann mitmachen: Wikisource, Wikidata, Wikipedia oder Wikimedia Commons können prinzipiell ohne Anmeldung benutzt und editiert werden. Das Login zum Beispiel für Wikipedia gilt global für alle Wikimediaportale in allen Sprachvarianten.

Regelmäßig dienstags morgens trifft sich das Projektteam online eine halbe Stunde zum *DatenlaubeJam*, einer für alle Interessierten offenen Projektbesprechung ggf. mit Gästen. Für die Dokumentation dieser Treffen, der Kurse und Vorträge wird die Wikiversität/Wikiversity genutzt. Vortragsskript und die *Hands on Labs* im Rahmen des Workshops *Partizipative Transkriptionsprojekte in Museen, Archiven, Bibliotheken* wurden ebenfalls dort offen dokumentiert. Das Wikidata-Item des gesamten Workshops am 28. und 29. Oktober 2021 ist (Wikidata: (Q109299762)). Eine Projektbibliografie der *Datenlaube* wird in Wikidata unter (Q61943025) fortlaufend ergänzt (Wikidata: Aussage (Q61943025#P1343)). Das Projektblog ist

auf <https://diedatenlaube.github.io/> zu finden. Weitere Projektberichte wurden im Blog des sächsischen Landeskundeportals *Saxorum* veröffentlicht: <https://saxorum.hypotheses.org/tag/die-datenlaube>.

Projektziele

Im engeren Sinne verfolgt *Die Datenlaube* das Ziel, die Jahrgänge 1853 bis 1899 der Illustrierten *Die Gartenlaube* vollständig zu erschließen – voraussichtlich ungefähr 18.500 einzelne Zeitungsartikel. Bisher wurden etwa 16.300 dieser Artikel in Wikisource bearbeitet und durch strukturierte maschinenlesbare Metadaten in Wikidata beschrieben, also bibliografisch erschlossen. Sobald darüber hinaus weitere Jahrgänge der Illustrierten in Wikisource bearbeitet werden, wird diese Aufgabe wachsen. Ausgehend von der formalen Erschließung aller Gartenlaubeartikel reichern wir die Datenobjekte jedes Artikels durch zusätzliche Aussagen an: inhaltliche Schlagworte, ggf. Illustrationen, deren Illustratorinnen und Illustratoren, Genre, bibliografische Zusammenhänge wie Serien und Zitationen anderer Werke: Gartenlaubeartikel, wissenschaftliche Publikationen und Bücher. Damit betrachten wir *Die Datenlaube* als exploratives Pilotprojekt für andere Werke und Konvolute der deutschsprachigen Wikisource. Langfristig wäre es ideal, wenn alle Wikisourcetranskriptionen in allen Sprachversionen in Wikidata nachgewiesen und detailliert erschlossen werden. Das verbessert die Sichtbarkeit, Durchsuchbarkeit und Verknüpfbarkeit auch für Zwecke der Wissenschaft.

International werden auch in anderen Sprachportalen von Wikisource Lösungen für die Anreicherung mit strukturierten Daten und Wikidata entwickelt. Diese Ansätze betreffen die bibliografische Inhaltserschließung, semantische Verfahren für die Erschließung von Details und Annotationen für digitale Editionen. An diesen Entwicklungen beteiligt sich das Projektteam der *Datenlaube*.

Im weiteren Sinne fördert das Projekt die Arbeit mit den digitalen Werkzeugen der Wikimediaportale insbesondere an Schnittstellen zwischen Gedächtnisinstitutionen (GLAM bzw. *open GLAM, Galleries, Libraries, Archives and Museums*) und nicht-institutionalisierten datenorientierten Bürgerinnen und Bürgern sowie den Bürgerwissenschaften (*Citizen Science* bzw. *Open Citizen Science*), um neue Möglichkeiten für partizipative Projekte zu eröffnen. Zu diesen Forschungs- und Entwicklungsaktivitäten gehören:

- Anleitungen für *Die Gartenlaube/Die Datenlaube* sowie für Wikisource allgemein
- Kollegiale Beratungen für Wikidata in der Wikisource-Gemeinschaft
- Kollegiale Beratungen für Wikisource und Wikidata im Bibliotheksbereich
- (semi-)automatisierte Erschließungsverfahren
- Strukturierte Erschließungen detaillierter Bilddetailpositionen für Bilder der Wikimedia Commons

- Annotationswerkzeug,
- Abfrageverfahren, insb. *Nearby-Queries* mit Wikidata (<https://saxorum.hypotheses.org/tag/1lib1nearby>), geodatenbasierte Abfragen für das multilinguale Weltwissen der nahen Umgebung
- Verknüpfungen zwischen Regiowikis, Wikidata und Wikisource, Interwiki-Verlinkungen (teils reziproke Schlagwort/Referenz-Links) insb. zwischen Wikisourcetexten und Wikipedia-Artikeln
- Vorträge, Workshops, Lehrveranstaltungen
- Digitale Wissenschaftskommunikation.

Grundsätze

Die folgenden Grundsätze lassen sich aus der bislang dreijährigen Projektarbeit ableiten:

- Jede und jeder kann selbst Metadaten erzeugen und pflegen – für historische Quellen und für eigene Projektpublikationen.
- Projekte auf der Basis von Wikimatechnologien profitieren von der wechselseitigen Nutzung, Verknüpfung, Verlinkung und von den beteiligten Gemeinschaften in Wikisource, Wikipedia, Wikidata sowie der Regiowikis.
- *Linked Open Storytelling* ist ein Weg, um digitale Metadaten für die digitale Wissenschaftskommunikation in Sozialen Medien zu nutzen, in Blogs, in Tweets oder unabhängigen digitalen Anwendungen.
- Offene Projektarbeit, offene Projektdokumentationen und offene Projektkommunikation sind Grundlagen für offene Bildungsmaterialien auf Basis der *Datenlaube*.
- Die Nische *Gartenlaube* mag klein sein: offene Bürgerwissenschaft in den Geisteswissenschaften funktioniert als *Open Citizen Science* idealerweise mit offenen Daten: Forschungsdaten, Publikationen und ihren Metadaten.
- *Die Datenlaube* ist ein nicht-institutionelles *GLAM Lab* – *since 1853*.

*Wer kennt unsere „Gartenlaube“ nicht? In Europa kennt man sie und in anderen Weltteilen: aber in Leipzig heißt die „Gartenlaube“ nicht bloß die Zeitschrift, sondern auch das von Ernst Keil gebaute stattliche Haus, in dem sie zur Ausgabe gelangt. Im Jahre 1862 fand hier der Richtschmaus statt und im folgenden Jahre entfaltete sich darin bereits die volle Thätigkeit der Journalistik und des Buchhandels. (...)*¹

Literaturverzeichnis

Bemme, Jens: Hilfe für die Datenlaube: mit [[Wikisource+Wikidata]] die freie Quellensammlung verbessern, Berlin 2019, URL:

¹ URL: https://de.wikisource.org/wiki/Ein_Donnerstag_an_der_Expedition_der_Gartenlaube.

<https://blog.wikimedia.de/2019/10/16/hilfe-fuer-die-datenlaube-mit-wikisourcewikidata-die-freie-quellensammlung-verbessern/>.

Bemme, Jens: #Nearby. Landeskunde und Citizen Science mit Pandemie im Frühjahr 2020, Informationspraxis, Bd. 6 Nr. 2 (2020), DOI: <https://doi.org/10.11588/ip.2020.2.73402>.

Bemme, Jens; Erlinger, Christian: Zwischenbericht des „SFB 1853“ – zwei Jahre #DieDatenlaube, Dresden 11. Februar 2021, URL: <https://saxorum.hypothes.es.org/5692>.

Bemme, Jens; Erlinger, Christian; Erfurth, Matthias: DieDatenlaube, Wikiversity.org 2021, URL: <https://de.wikiversity.org/wiki/DieDatenlaube>.

Bemme, Jens; Erfurth, Matthias: Partizipative Transkriptionsprojekte (DieDatenlaube), Wikiversity.org, 29. Oktober 2021, URL: [https://de.wikiversity.org/wiki/Kurs:Partizipative_Transkriptionsprojekte_\(DieDatenlaube\)](https://de.wikiversity.org/wiki/Kurs:Partizipative_Transkriptionsprojekte_(DieDatenlaube)).

Bemme, Jens: Die Datenlaube: Open Citizen Science mit Linked Open Storytelling beim Erschließen der ‚Gartenlaube‘, 9. November 2021, URL: <https://publicdh.hypothes.es.org/176>.

Die Datenlaube, Blog: <https://diedatenlaube.github.io/>.
Wikimedia Deutschland: Wikisource Broschüre, Berlin 2019, URL: <https://commons.wikimedia.org/wiki/File:Wikisource-Brosch%C3%BCre.pdf>.

Wikisource.de: Die Gartenlaube, Leipzig 1853–1899, URL: https://de.wikisource.org/wiki/Die_Gartenlaube.

Wikisource.de: Ein Donnerstag an der Expedition der „Gartenlaube“, Die Gartenlaube, Leipzig 1895, Heft 32, S. 548, URL: https://de.wikisource.org/wiki/Ein_Donnerstag_an_der_Expedition_der_Gartenlaube.

Illustrationen

https://commons.wikimedia.org/wiki/File:Die_Datenlaube.xcf
Projektlogo: Die Datenlaube.

https://upload.wikimedia.org/wikipedia/commons/archive/2/20/20211010181758%21Die_Gartenlaube_%281895%29_b_548_2.jpg

Otto Gerlach: Buchhändlerkarren vor dem Thor der Gartenlaube, in: Ein Donnerstag an der Expedition der „Gartenlaube“, Die Gartenlaube, 1895, Heft 32, S. 548, URL: https://de.wikisource.org/wiki/Ein_Donnerstag_an_der_Expedition_der_Gartenlaube.

Autoreninformation

Jens Bemme (SLUB Dresden) wirkt im Bereich Saxonica/ Sächsische Landeskunde/ Citizen Science.

Christian Erlinger arbeitet als Projektmitarbeiter an der Zentral- und Hochschulbibliothek Luzern.

Matthias Erfurth (IPF Dresden) engagiert sich im Stadtwiki Dresden.