

Ein weiteres Toolverzeichnis für die Digital Humanities?!

Aber diesmal offen und mit Wikidata

Grallert, Till

till.grallert[at]hu-berlin.de
Humboldt-Universität zu Berlin, Deutschland
ORCID-ID: 0000-0002-5739-8094

Eckenstaler, Sophie

sophie.eckenstaler.1[at]ub.hu-berlin.de
Humboldt-Universität zu Berlin, Deutschland

Tirtohusodo, Samatha

samantha.tirtohusodo.1[at]hu-berlin.de
Humboldt-Universität zu Berlin, Deutschland

Schlesinger, Claus-Michael

claus-michael.schlesinger[at]hu-berlin.de
Humboldt-Universität zu Berlin, Deutschland
ORCID-ID: 0000-0001-6718-5773

Zusammenfassung. Der eingereichte Beitrag skizziert die Landschaft der bestehenden Toolverzeichnisse in den Digital Humanities und ihrer Schwächen und stellt vor diesem Hintergrund unseren Vorschlag eines Wikidata-basierten offenen Ansatzes als einen von minimal computing, making und Open Science informierten Beitrag zu den Digital Commons vor.

Toolverzeichnisse sind Legion und ein etabliertes Genre in den Digital Humanities: von DiRT zu Bamboo und TAPoR (3.0)¹ (Grant u. a. 2020), großen EU-Projekten wie dem *Social Sciences and Humanities Open Marketplace*,² den Konsortien der deutschen Nationalen Forschungsdaten Infrastruktur (NFDI), den Fachinformationsdiensten (FID) oder individuellen Bibliotheken und Instituten. Ihnen allen ist gemeinsam, dass sie ein offensichtliches und reales Bedürfnis der Forschungscommunities nach einem Überblick über computergestützte Werkzeuge mit kuratorischen Ansätzen der Wissensorganisation bedienen. Ihnen ist außerdem gemeinsam, dass sie über keine

¹ <https://tapor.ca/>

² <https://marketplace.sshopencloud.eu/>

dauerhafte Finanzierung verfügen, dass sie primär auf die Kuratierung durch (unbezahlte) Expert_innen-Gremien setzen, diesen Prozess aber nicht dauerhaft und nachhaltig gewährleisten können, dass Datensilos mit proprietären Infrastrukturen (Datenmodelle, Backends und Frontends) geschaffen werden, und dass nur in geringem Maße APIs angeboten und dokumentiert werden. Im Ergebnis sind diese Toolverzeichnisse in dem Anspruch eines umfassenden, repräsentativen und je aktuellen Abbildes der verfügbaren Möglichkeiten computationeller Forschung und digitaler Wissenschaft als gescheitert zu verstehen (vgl. Dombrowski 2021). Vielmehr handelt es sich im besten Fall um Momentaufnahmen bzw. die Dokumentation historischer Praktiken. Der Fokus auf eine volatile, kontinuierliche Wartung erfordernde Präsentationsschicht führt dazu, dass auch die Daten als Ergebnis und wissenschaftlicher Mehrwert der Kuratierung nicht dauerhaft, mit permanenten URIs, und in einem stabilen, maschinenlesbaren Format als Basis für verlinkte Normdatensätze zur Verfügung stehen.

Für den DFG-geförderten prototypischen *Scholarly Makerspace* als Lernort zur Förderung der Werkzeugkompetenz (*tool literacy*) in den Geistes- und Kulturwissenschaften an der Universitätsbibliothek der Humboldt-Universität zu Berlin verfolgen wir daher einen von *minimal computing*, *making* und Open Science inspirierten Ansatz. *Minimal computing* erfordert dabei die Frage "was brauchen wir?" mit den Antworten auf die Frage "was steht uns je zur Verfügung?" in Einklang zu bringen (vgl. Gil und Ortega 2016; Risam und Gil 2022). Für unsere Arbeit in der Beratung benötigen wir eine Infrastruktur, die es uns erlaubt unser sich entwickelndes Wissen über computationelle Werkzeuge für die Wissenschaft nachhaltig festzuhalten und teilen zu können: was gibt es, wofür kann es eingesetzt werden, wer hat es eingesetzt und dabei welche Erfahrungen gemacht, und wo und wie kann ich den Einsatz erlernen? Den Projektzyklen der Wissenschaftsförderung unterliegend, sind unsere Möglichkeiten dabei beschränkt: wir sind darauf angewiesen auf bestehenden Datensätzen aufzusetzen und offene, kostenfreie und etablierte Software und Plattformen zu nutzen. Der wissenschaftliche Mehrwert liegt dann darin gut dokumentierte Workflows, Datenmodelle und Beispielimplementationen für den Aufbau von offenen Toolverzeichnissen zu entwickeln und den Forschungscommunities zur Verfügung zu stellen.

Der Beitrag stellt unseren Vorschlag einer gemeinsamen offenen Basisinfrastruktur für Toolverzeichnisse für die Fachcommunities vor. Dabei steht Wikidata³ als eine verteilte, community-kuratierte Normdatei und offene Softwareplattform im Zentrum unseres Vorschlages, die mehrere Schwächen bestehender Toolverzeichnisse adressiert. Wikidata erlaubt es minimale Datenmodelle iterativ zu entwickeln, Datensätze zu pflegen und diese in Wikiprojekten zu kuratierten Sammlungen zusammenzustellen. Auf der Datenebene erlaubt Wikidata die unmittelbare Nutzung sämtlicher Informationen als Linked Open Data (LOD) über SPARQL, APIs sowie das etablierte Webinterface. Wikidata ist außerdem eine der Quellen für das *Virtual International Authority File* (VIAF)⁴ und für zusammenfassende Informationen in den Ergebnislisten der dominanten Suchmaschinen, was die Sichtbarkeit der Datensätze enorm erhöht. Darüber hinaus bieten Wikidata und ihre Schwesterprojekte eine etablierte Governancestruktur für nutzergenerierte und -kuratierte Inhalte. Jede_r kann die Einträge beitragen und pflegen, die für ihre je konkrete Forschung relevant sind. Anders als bei viele Infrastrukturen der Digital Humanities ist die Vielsprachigkeit von Interfaces und Datensätzen ein grundlegendes Feature. Auf dieser Datenbasis lassen sich dann Fachcommunity-spezifische Toolverzeichnisse kuratieren und anreichern. Denkbar ist etwa eine Klassifizierung unter Anwendung der TaDiRAH-Taxonomie⁵ (Borek u. a. 2021) oder die Hinterlegung von Anwendungsbeispielen, Publikationen oder Tutorials im angereicherten Datensatz. Die Wikimedia-Software und -Plattform bietet die Möglichkeit, dies auch direkt in Wikidata zu tun. Unser Vorschlag erlaubt aber auch, Wikidata ausschließlich als Normdatei und Datenprovider für eigene Frontends einzusetzen, so wie es z.B. Scholia⁶ für die Profile von Wissenschaftler_innen tut (Nielsen, Mietchen, und Willighagen 2017). Schließlich adressiert unser Vorschlag die Nachhaltigkeit von Projektförderungen durch den kontinuierlichen Beitrag von Daten zu den *Digital Commons* (Wittel 2013) in Gestalt von Wikidata während der Projektlaufzeit und die Weiternutzung dieser Daten nach der

³ <https://wikidata.org/>

⁴ <https://viaf.org/>

⁵ <https://vocabs.dariah.eu/tadirah/>

⁶ <https://scholia.toolforge.org/>

Projektlaufzeit. Damit ist unser Vorschlag Teil einer Bewegung, Wikidata in der Wissenschaft und GLAM-Institution nicht mehr nur als Anbieter von Inhalten wahrzunehmen (vgl. Zhao 2022; Fischer und Ohlig 2019).

Unserer Kenntnis nach, gibt es mit der Research Software Encyclopaedia⁷ nur eine weitere community-kuratierte, den FAIR-Prinzipien (vgl. Barker u. a. 2022) und nachhaltigen Technologie-Stacks verpflichtete und damit unserem Vorschlag verwandte Infrastruktur. Allerdings verfolgte diese keinen LOD-Ansatz, fokussiert ausschließlich auf Forschungssoftware, die auf GitHub oder GitLab gehostet ist, und konzentriert sich nicht auf die (Digital) Humanities (Sochat u. a. 2022).

Bibliografie

Barker, Michelle, Neil P. Chue Hong, Daniel S. Katz, Anna-Lena Lamprecht, Carlos Martinez-Ortiz, Fotis Psomopoulos, Jennifer Harrow, u. a. 2022. „Introducing the FAIR Principles for Research Software“. *Scientific Data* 9 (1): 622. <https://doi.org/10.1038/s41597-022-01710-x>.

Borek, Luise, Canan Hastik, Vera Khramova, Klaus Illmayer, und Jonathan D. Geiger. 2021. „Information Organization and Access in Digital Humanities: TaDiRAH Revised, Formalized and FAIR“. In *Information Between Data and Knowledge*, 321–32. Schriften Zur Informationswissenschaft 74. Glückstadt: Werner Hülsbusch. <https://doi.org/doi.org/10.5283/epub.44951>.

Dombrowski, Quinn. 2021. „The Directory Paradox“. In *People, Practice, Power: Digital Humanities Outside the Center*, herausgegeben von Anne B. McGrail, Angel David Nieves, und Siobhan Senier. Debates in the Digital Humanities. Minneapolis: University of Minnesota Press. <https://dhdebates.gc.cuny.edu/read/people-practice-power/section/ca87ec4c-23a0-452d-8595-7cfd7e8d6f0c>.

Fischer, Barbara, und Jens Ohlig. 2019. „GND Meets Wikibase‘ - Eine Kooperation. Eine Bundesbehörde Geht Auf Expedition Im Wikiversum: Ein Neues Testfeld Für Wikibase“. *GND* (blog). 8. Mai

⁷ <https://rseng.github.io/projects/research-software-encyclopedia/>

2019.

<https://wiki.dnb.de/pages/viewpage.action?pageId=147754828>.

Gil, Alex, und Élika Ortega. 2016. „Global Outlooks in Digital Humanities: Multilingual Practices and Minimal Computing“. In *Doing digital humanities: practice, training, research*, herausgegeben von Constance Crompton, Richard J Lane, und Ray Siemens, 22–34. Abingdon: Routledge.

Grant, Kaitlyn, Quinn Dombrowski, Kamal Ranaweera, Omar Rodriguez-Arenas, Stéfan Sinclair, und Geoffrey Rockwell. 2020. „Absorbing DiRT: Tool Directories in the Digital Age“. *Digital Studies / Le Champ Numérique* 10 (1). <https://doi.org/10.16995/dscn.325>.

Nielsen, Finn Årup, Daniel Mietchen, und Egon Willighagen. 2017. „Scholia, Scientometrics and Wikidata“. In *The Semantic Web: ESWC 2017 Satellite Events*, herausgegeben von Eva Blomqvist, Katja Hose, Heiko Paulheim, Agnieszka Ławrynowicz, Fabio Ciravegna, und Olaf Hartig, 237–59. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-70407-4_36.

Risam, Roopika, und Alex Gil. 2022. „Introduction: The Questions of Minimal Computing“. Herausgegeben von Alex Gil und Roopika Risam. *Digital Humanities Quarterly* 16 (2, "Minimal Computing"). <http://digitalhumanities.org/dhq/vol/16/2/000646/000646.html>.

Sochat, Vanessa, Nicholas May, Ian Cosden, Carlos Martinez-Ortiz, und Sadie Bartholomew. 2022. „The Research Software Encyclopedia: A Community Framework to Define Research Software“. *Journal of Open Research Software*, März. <https://doi.org/10.5334/jors.359>.

Wittel, Andreas. 2013. „Counter-commodification: The economy of contribution in the digital commons“. *Culture and Organization* 19 (4): 314–31. <https://doi.org/gmqgqq>.

Zhao, Fudie. 2022. „A systematic review of Wikidata in Digital Humanities projects“. *Digital Scholarship in the Humanities*, Dezember, 1–22. <https://doi.org/10.1093/lc/fqac083>.