



Contribution ID: 106

Type: Poster

Eine weitere Tool Registry für die DH?! Aber diesmal offen und community-zentriert auf Wikidata

Wednesday 26 February 2025 19:40 (20 minutes)

Der Zusammenhang zwischen Methoden, diese implementierenden Werkzeugen (Software) und ihrer Nützlichkeit für die Untersuchung einer Forschungsfrage und -gegenstands ist von immanen Interesse für die computationally arbeitenden Geisteswissenschaften. In der Folge hat sich das Toolverzeichnis in den Digital Humanities inzwischen als eigenes Genre etabliert: von TAPoR (3.0)¹ (Grant u. a. 2020), großen EU-Projekten wie dem Social Sciences and Humanities Open Marketplace,² oder den Konsortien der Nationalen Forschungsdateninfrastruktur (NFDI).

Alle uns bekannte Ansätze

- verfügen über keine dauerhafte Finanzierung;
- setzen primär auf die Kuratierung durch (unbezahlte) Expert_innen-Gremien oder Crowdsourcing, können diesen Prozess aber nicht dauerhaft und nachhaltig gewährleisten;
- schaffen Datensilos mit proprietären Infrastrukturen (Datenmodelle, Backends und Frontends);
- und bieten nur in geringem Maße dokumentierte APIs an.

Der Anspruch eines repräsentativen oder gar umfassenden Abbilds aktuell verfügbarer Werkzeuge für die computergestützte Forschung ist daher nicht einzulösen und muss, wo er formuliert wurde, als gescheitert gelten. (vgl. Dombrowski 2021).

Da der Informationsbedarf aber weiterhin besteht, stellt dieser Beitrag unseren Ansatz einer **offenen Basisinfrastruktur für Toolverzeichnisse** vor. Wikidata³ steht dabei als ein verteilter, community-kuratierter Wissensgraph und offene Softwareplattform im Zentrum unseres Ansatzes und adressiert die Schwächen anderer Ansätze. Wikidata erlaubt es, Datenmodelle iterativ zu entwickeln, Datensätze zu pflegen und sie in Wikiprojekten zu kuratierten Sammlungen zusammenzustellen. Auf der Datenebene erlaubt Wikidata die unmittelbare Nutzung sämtlicher Informationen als Linked Open Data über SPARQL, APIs sowie das etablierte Webinterface. Datensätze auf Wikidata sind im globalen Internet sehr sichtbar und bilden beispielsweise eine der Quellen für zusammenfassende Informationen in den Ergebnislisten von Suchmaschinen. Darüber hinaus bietet Wikidata eine etablierte Governancestruktur für nutzergenerierte und -kuratierte Inhalte. Jede_r kann die Einträge beitragen und pflegen, die für ihre je konkrete Forschung relevant sind (Stakeholderprinzip).

Die Grundlage verteilter angelegter Datensätze ist ein reduziertes Basisdatenmodell für DH-Werkzeuge, das minimale bibliografische und technische Eigenschaften definiert. Damit ist es anschlussfähig für existierende Datenmodelle, wie dem des Software Preservation Network oder von RIDE (Christopherson u. a. 2022; Sichani und Spadini 2022) und kann zu einem gemeinsamen Referenzmodell beitragen. Dieses Basisdatenmodell garantiert den Zugriff auf die verbindlich vereinbarten Basisdaten für die Nachnutzung von Einträgen in eigenen kuratierten Sammlungen mit erweiterten Datenmodellen. Für den Kontext der DH haben wir dieses Basismodell etwa um eine Klassifizierung unter Anwendung der TaDiRAH-Taxonomie⁴ (Borek u. a. 2021) und die Hinterlegung von Anwendungsbeispielen, Publikationen oder Tutorials im angereicherten Datensatz erweitert.

Als Plattform erlaubt Wikidata die Umsetzung des Ansatzes ohne weitere Software. Denkbar ist aber auch, Wikidata ausschließlich als Normdatei und Datenprovider für eigene Frontends einzusetzen, so wie es z.B. Scholia⁵ (Nielsen, Mietchen, und Willighagen 2017) tut.

Schließlich adressiert unser Ansatz 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 Projektlaufzeit. Damit ist unser Vorschlag Teil einer Bewegung, Wikidata in der Wissenschaft und GLAM-Institutionen zu verankern (vgl. Zhao 2022; Fischer und Ohlig 2019).

I want to participate in the youngRSE prize

Primary authors: GRALLERT, Till (Humboldt-Universität zu Berlin); DRESSELHAUS, Nicole Elisabeth Hitomi (Humboldt-Universität zu Berlin)

Presenters: GRALLERT, Till (Humboldt-Universität zu Berlin); DRESSELHAUS, Nicole Elisabeth Hitomi (Humboldt-Universität zu Berlin)

Session Classification: Poster and Demo Session together with Reception

Track Classification: Research Software: digital humanities