

Neubert, Joachim

Article — Published Version

Warum spenden wir Daten an Wikidata?

ABI Technik

Suggested Citation: Neubert, Joachim (2020) : Warum spenden wir Daten an Wikidata?, ABI Technik, ISSN 0720-6763, De Gruyter, Berlin, Vol. 40, Iss. 2, pp. 209,
<https://doi.org/10.1515/abitech-2020-2013>

This Version is available at:

<http://hdl.handle.net/11108/460>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.

ABI Technikfrage

Joachim Neubert

Warum spenden wir Daten an Wikidata?



<http://doi.org/10.1515/abitech-2020-2013>

Joachim Neubert ist wissenschaftlicher Softwareentwickler bei der ZBW – Leibniz-Informationszentrum Wirtschaft am Standort Hamburg. Seit 2008 liegen seine Schwerpunkte dort in den Bereichen Linked Open Data und Wissensorganisationssysteme, etwa beim Standard-Thesaurus Wirtschaft und der GND. Arbeiten wie die hier beschriebenen dokumentiert er u. a. im „ZBW Labs“-Blog. Er beantwortet die aktuelle ABI Technikfrage: Warum spenden wir Daten an Wikidata?

Wikidata ist eine freie Wissensdatenbank. Seit 2012 zunächst für die Verknüpfung der vielen sprachspezifischen Wikipedia-Projekte aufgebaut, ist Wikidata heute der wichtigste Knotenpunkt für freie Daten im Web. Es kann nicht nur offen abgefragt werden, sondern alle Interessierten können dort Daten ergänzen – von Hand oder unter Verwendung von effizienten Tools.

Bibliotheken können Wikidata nutzen, um etwa erweiterte Kontextinformationen oder Verknüpfungen aus über 4 000 Normdatensets zu beziehen (s. a. ABI Technikfrage 2/2018: Was ist Wikidata und wie kann es die bibliothekarische Arbeit unterstützen?). Sie können aber darüber hinaus systematisch Daten beitragen – und damit in manchen Fällen eigene Anwendungen überflüssig machen.

2017 hat die ZBW das in einem experimentellen Setting getan: Als Nebenprodukt eines Disambiguierungsprojektes war eine Liste mit gut 3 000 Zuordnungen von GND-Personen-IDs zu Autoren-Identifiern von Repec, der wichtigsten Quelle für Open-Access-Publikationen in den Wirtschaftswissenschaften, entstanden. Anstatt nun eine eigene Datenbank aufzubauen oder die Erweiterung der GND um ein Feld Repec-ID zu beantragen, haben wir eine Speicherung in Wikidata gewählt. 760 Personen wiesen dort bereits beide Identifier auf. Wir haben dem Bestand weitere IDs hinzugefügt, und wo nötig neue Personen Datensätze angelegt. Gut 4 000 waren es, als alle unsere Daten integriert waren. Seither hat die „Crowd“ der Wikidata-Beitragenden die Anzahl der jederzeit abfragbaren Autoren-Mappings auf über 6 000 erhöht.

Die „Pressemappe 20. Jahrhundert“ hat im Rahmen eines DFG-Projekts (2004–2007) Dossiers mit Presseartikeln zu Personen, Firmen und Sachthemen mit Millionen

von Seiten digitalisiert – eine einzigartige, bis ins Jahr 1949 reichende frei zugängliche Quelle zur Zeitgeschichte. Aber die Daten sind im Web nicht leicht auffindbar, und die damals erstellte Anwendung ist in die Jahre gekommen. In einem „WikiProject 20th Century Press Archives“ wird daher eine groß angelegte Data Donation an Wikidata vorbereitet und umgesetzt. Dabei werden die Metadaten der Dossiers Schritt für Schritt mit entsprechender Quellenangabe in Wikidata-Eigenschaften übernommen – vom Geburtsdatum einer Person bis zu der Information, dass sie in einer bestimmten Zeitspanne im Aufsichtsrat einer bestimmten Firma saß. Für die mehr als 5 200 Dossiers des Personenarchivs ist dieser Prozess bereits abgeschlossen, etwa 350 Wikidata-Personen wurden neu erzeugt. Die Digitalisate selbst und schlanke Referenzseiten bleiben bei der ZBW, der primäre Zugang erfolgt künftig aber über Wikidata. Die bessere Auffindbarkeit schlägt sich in deutlich höheren Nutzerzahlen nieder. Vor allem aber „leben“ die Daten in Wikidata weiter, werden ergänzt oder auch mal korrigiert, und können – durch die Multilingualität von Wikidata auch international – in vielen Zusammenhängen einfach und frei genutzt werden.

Autoreninformationen

Joachim Neubert

ZBW – Leibniz-Informationszentrum Wirtschaft
(Kiel/Hamburg), Neuer Jungfernstieg 21
20354 Hamburg

J.Neubert@zbw.eu

orcid.org/0000-0002-8086-185X