

Forschungsdaten- Management-Services für Linked Open Data: Ein Vergleich

Nasarek, Robert

r.nasarek@gnm.de
Germanisches Nationalmuseum Nürnberg,
Deutschland

Rossenova, Lozana

Lozana.Rossenova@tib.eu
Leibniz Information Centre for Science and Technology

Sohmen, Lucia

Lucia.Sohmen@tib.eu
Leibniz Information Centre for Science and Technology

Duchesne, Paul

Paul.Duchesne@tib.eu
Leibniz Information Centre for Science and Technology

Einleitung

Vor fast 15 Jahren stellte Tim Berners-Lee seine Vision des "Giant Global Graph" vor: Nach dem Net, das Computer miteinander verbinde und dem Web, das Dokumente verknüpfe, setze der Graph einzelne digitale Dinge miteinander in Beziehung und stelle die dritte Stufe des Internets dar (Berners-Lee 2007). Einzelne Datensätze können referenziert werden und die Vorteile für die Forschung liegen auf der Hand: Die erleichterte Wiederverwendung vermeidet die Mehrfacherfassung von Daten und reduziert so Arbeitsaufwand und Redundanzen, was den Ressourcenverbrauch, die Mehrdeutigkeit und die Fehleranfälligkeit verringert (W3C 2011; Neumann 2021). Die Basis eines solchen Graphen ist heute als Linked Open Data (LOD) bekannt: vernetzte, eindeutig identifizierbare, offen zugängliche Daten, die mit Semantik angereichert und damit universell verständlich sind.

Seit Berners-Lee diese Vision im Jahr 2007 postulierte, haben sich die Bedingungen für die Erstellung und Verbreitung von LOD stetig verbessert. Neben Beschreibungs- und Abfragesprachen (namentlich RDF/OWL, SPARQL), Normdaten (wie z.B. die Gemeinsame Normdatei), Vokabularen (wie z.B. die Getty Vocabularies) und Standard-Ontologien (wie z.B. das Conceptual Reference Model des Internationalen Komitees für Dokumentation - CIDOC CRM) stehen uns heute eine Vielzahl von elaborierten Softwareumgebungen zur Verfügung, die uns bei der Erstellung, Bearbeitung und Veröffentlichung von LOD unterstützen (Engelhardt und Kutsch 2021, 458-460).

Zu Beginn eines Projekts stehen viele Forscher:innen jedoch vor der Herausforderung, eine geeignete Softwareumgebung auszuwählen, die alle relevanten Anforderungen der Dateneingabe, -speicherung, -modifikation, -annotation, -analyse, -präsentation, -archivierung und -publikation erfüllt. Oftmals wissen die Datenkurator:innen nicht im Voraus, welche Eigenschaften und Funktionen einer Software sie benötigen, da sich bestimmte Anforderungen im Laufe eines Forschungsprojekts ergeben oder ändern können. Die Auswahl einer geeigneten Forschungsdaten-Management-Umgebung ist eine der wichtigsten Entscheidungen eines datenfokussierten Projektes, die nur mit großem Aufwand im fortgeschrittenen Stadium geändert werden kann (Humboldt-Universität zu Berlin 2020). In diesem Zusammenhang glauben wir, dass ein Überblick über alle verfügbaren LOD-Management-Systeme und ihre relevanten Eigenschaften ein äußerst nützliches Werkzeug für alle Forscher:innen sein kann, die vor der Herausforderung stehen, den am besten geeigneten Software für ihr spezifisches Projekt auszuwählen.

Die Linked Open Data Arbeitsgruppe

NFDI4Culture ist ein deutsches Konsortium von Forschungs- und Kultureinrichtungen, das auf eine gemeinsame Infrastruktur für Forschungsdaten hinarbeitet, die den Bedürfnissen der Datenersteller:innen, -kurator:innen und -nutzer:innen des 21. Jahrhunderts im breiten Spektrum des kulturellen Erbes und der Digital Humanities entspricht (NFDI4Culture 2022). Die Linked Open Data Working Group von NFDI4Culture soll als Initiative die Erstellung des Culture Knowledge Graph unterstützen, dessen Ziel es ist, Forschungsdaten zu verknüpfen, die von verschiedenen Partnerorganisationen des Konsortiums erzeugt wurden (Linked Open Data Working Group 2022a). Im Rahmen dieser unterstützenden Rolle im Konsortium organisiert die LOD-Working Group Workshops (Rossenova und Sohmen 2021a), Infoveranstaltungen (Linked Open Data Working Group 2022b) und veröffentlicht Schulungsmaterial, um Forscher:innen den Einstieg in LOD zu erleichtern (Rossenova und Sohmen 2021b). Während der monatlichen Treffen mit den Mitgliedern der NFDI4Culture-Gemeinschaft stellt die LOD-Working Group auch die besonderen Bedürfnisse und Anforderungen der Gemeinschaft heraus. Als Mitbegründer der Gruppe konnten wir den Mangel an leicht verfügbaren Ressourcen feststellen, die einen umfassenden Überblick über die verschiedenen Software-Umgebungen bieten, die Forscher:innen zur Verfügung stehen, und gleichzeitig die Anforderungen an Daten und Software berücksichtigen, die den FAIR-Prinzipien entsprechen (Wilkinson et al. 2016). Seitdem haben wir gemeinsam mit Nutzer:innen und Expert:innen aus den jeweiligen Communities einen ersten Überblick über die folgenden Forschungsdaten-Management-Software erstellt: ConedaKor, Omeka S, Metaphacts Community/ ResearchSpace, WikiBase und WissKI, die unserer Meinung nach zu den am häufigsten genutzten Linked Open Data-fähigen Umgebungen in den digitalen Geisteswissenschaften gehören. Damit einher ging die Entwicklung von Vergleichskategorien, an-

hand derer diese Software-Umgebungen bewertet werden können

Methodik

Der Vergleich wurde mit agilen, Community-orientierten Methoden entwickelt. Community-Treffen, Experteninterviews und Literaturrecherchen bestimmten die endgültige Auswahl der zu vergleichenden Softwareoptionen und die endgültigen Vergleichskategorien. Die Rekrutierung erfolgte über die Mailingliste der LOD WG und über Community-Veranstaltungen, die vom NFDI4Culture-Konsortium und der LOD WG selbst organisiert wurden. Letztendlich wurde nur Software in den Vergleich einbezogen, für die Mitglieder der jeweiligen Maintainer-Community und erfahrene Nutzer an einer Reihe von Expertenrunden teilnehmen und zur endgültigen Tabelle beitragen konnten.

Der Vergleich sollte fünf Eckpunkte berücksichtigen: Neben dem übergreifenden Bezug zu den FAIR-Prinzipien, der projekt- und technikbezogenen Nachhaltigkeit und der Benutzerfreundlichkeit, sollten sowohl Anwender:innen als auch Entwickler:innenperspektiven beachtet werden. Es wurde darauf geachtet, dass alle Kategorien auch bei allen Forschungsdaten-Management-Umgebungen angewendet werden können und es keine individuellen Vergleichspunkte gibt, die nur bei einzelnen Anwendungen sinnvoll ausgefüllt werden können.

Für jede Kategorie wurde ein kontrolliertes Vokabular mit Erklärung entwickelt, um größtmögliche Vergleichbarkeit zu gewährleisten. Da jedoch zu manchen Punkten eine detaillierte Erläuterung hilfreich ist, wurde den Expert:innen die Möglichkeit gegeben, an gesonderter Stelle in Freitextfeldern weitere Informationen zu vermerken.

Das Ergebnis ist damit eine Tabelle mit zwei Arbeitsblättern: Das erste Tabellenblatt dient einem einfachen Vergleichs und enthält in der ersten Zeile die Forschungsdaten-Management-Systeme und in der ersten Spalte die Vergleichskategorien mit Erläuterungs-Tooltip. Die Zellinhalte dürfen nur mit einem kontrollierten Vokabular (Ja/Nein, Checkbox, Auswahlliste) gefüllt werden (siehe Abbildung 1). Das zweite Arbeitsblatt hat die gleiche Tabellenstruktur, jedoch können in die Zellen nun Freitexte eingegeben werden (siehe Abbildung 2).

#	Category	Wikski	Wikibase	Vitro	LinkedC
1	Official website	https://www.wiki.org	https://wikibase.org	https://www.vitro-project.org	https://www.linkedc.org
2	License	GNU General Public License, Version 2	BSD 3-Clause	BSD 3-Clause License	CC-BY-SA
3	Source code availability	Is the code available online (in GIT etc.)?	Available	Available	Available
4	Community maturity	05-03: 35,634 53 forks	*7000 Comments, 70 stars, 52 forks (May 2022)	*6500 comments, 3 forks (July 2022)	
5	Statistics (time stamp #)				
6	Payment models	Available: Code base is open.	Free	Free	Free
7	Maintainer and Support	Closed: Code base is not open.	Free	Free	Free
8	Documentation				
9	Infrastructure stack	Framework: https://www.php.net	Database: https://www.mysql.com	Webserver: Tomcat	Database: Apache Tomcat 9
10	APIs	RESTful: <input checked="" type="checkbox"/>	SPARQL: <input checked="" type="checkbox"/>	Modifiable HTTP: <input checked="" type="checkbox"/>	GraphQL: <input checked="" type="checkbox"/>
11	API usability level	Well documented	Well documented	Well documented	Well documented
12	SPARQL Endpoint usability level	Availability of graphical user interfaces, examples and documentation	Availability of graphical user interfaces, examples	Availability of graphical user interfaces, examples	Availability of graphical user interfaces, examples

Abbildung 1: Arbeitsblatt "Vergleich" mit kontrolliertem Vokabular und Erläuterungen als Tooltip.

#	Category	Wikski	Wikibase	Vitro	LinkedC
1	Official website				
2	License	The source code is available at github (you have to register to see details access).	Source code available at https://github.com	Source code available at https://github.com	https://github.com
3	Reverse code availability	Wikski includes about 83 community members, 22 developers. Drupal has 13M community members (incl. 49k developers).	The underlying MediaWiki software is used by tens of thousands of websites and thousands of companies and organizations. As of 2022-05-03, the main MediaWiki GitHub page has 101,774 commits, 2.6k stars, 1.2k forks, and 962 contributors. The dedicated Wikibase GitHub page has 111 contributors.		https://github.com
4	Community maturity				
5	Statistics (time stamp #)				317 stars, 72 forks (2022)
6	Payment models	Wikski is free to use, but https://www.wiki.org is not. It is actively maintained and supported by https://www.wiki.org (Wikski) Community building within https://www.wiki.org and https://www.wiki.org Slack Channel, issue handling at https://www.wiki.org module page.	Wikimedia Deutschland		Wikimedia Deutschland
7	Maintainer and Support				Free to use
8	Documentation		Documentation for the various components is available through their respective channels on the code repository and issue tracking platforms GitHub and Wikimedia . Additionally there are resources available https://www.wikimedia.org .	There is documentation for Vitro, which is Vitro's Research Information-related functions. Most of the documentation applies to Vitro, too.	https://www.vitro-project.org
9	Infrastructure stack	Minidao depends on special implementation of the https://www.wiki.org extension, generally making use of a https://www.wiki.org database, as well as adding an additional https://www.wiki.org instance for semantic queries. The integrated table store, however, will be changed in future releases due to long-term sustainability issues with https://www.wiki.org .	Wikibase is a https://www.wikimedia.org extension, generally making use of a https://www.wikimedia.org database, as well as adding an additional https://www.wikimedia.org instance for semantic queries. The integrated table store, however, will be changed in future releases due to long-term sustainability issues with https://www.wikimedia.org .		Drupal: rights, Apache Eclipse IDE, Apache Compatible with any SF Apache Java Framework

Abbildung 2: Arbeitsblatt "Details" mit genaueren Informationen zu einzelnen Punkten.

Das Vergleichssystem

Die mehr als 30 Vergleichskategorien beziehen sich auf wichtige Funktionen und Merkmale einer LOD-Management-Software. Diese gehen jedoch über die rein quantitativen Informationen hinaus und umfassen auch qualitative Aspekte in Bezug auf den Lebenszyklus von Forschungsdaten und Wartungsaktivitäten, die von jeder Software und der dazugehörigen Entwicklergemeinschaft bereitgestellt werden. In diesem Vortrag werden wir uns auf diese qualitativen Aspekte konzentrieren, um die Auswirkungen der verschiedenen Funktionalitäten auf die langfristige Verwaltung von FAIR-Daten aufzuzeigen. Wir werden nicht jedes Tool einzeln vergleichen, da die Konferenzteilnehmer:innen dies anhand der auf der NFDI4Culture-Website veröffentlichten Vergleichstabelle tun können. Stattdessen werden wir uns mit unserer Hauptforschungsfrage befassen: Welche Funktionalitäten und Eigenschaften sollte eine LOD-Software-Umgebung für die langfristige Verwaltung von FAIR-Forschungsdaten bieten und wie kann unsere Überblickstabelle Forscher:innen helfen, in ihrem projektspezifischen Kontext zeitnahe Entscheidungen zu treffen?

Folgende Kategorien flossen in den Vergleich ein: Generelle Informationen

- Offizielle Website
- Lizenz
- Zugänglichkeit des Quellcodes
- Community-Reife
- Metriken
- Geschäftsmodelle
- Maintainer
- Dokumentation
- Infrastruktur-Umgebung/ Stack

Schnittstellen und Endpoints

- verfügbare APIs
- Benutzerfreundlichkeit der APIs
- Zugänglichkeit des SPARQL Endpoints

Semantische Modellierung

- Reificationsmechanismus
- Reasoning
- Implementierung der Ontologie
- Umsetzung der Ontologiedefinitionen

- Wertvalidierung

- Nachnutzung von Daten

- Mapping/ Harvesting von Daten
- Generierung der Permanent-Identifiers
- Nachnutzen von Permanent-Identifiers
- Import-Exportfunktionalitäten
- Richtung des Datenzugriffs

- Datenmanagement

- Unterstützte Datentypen
- Mehrfachwerte

- Anpassbarkeit, Berechtigungen und Revisionierung

- Verwendung von Templates
- Anpassung des Interface
- Mehrsprachigkeit
- Versionierung
- Berechtigungen

- Integration, Datenvisualisierung und Skalierbarkeit

- integrierbare Medienformate
- Datenvisualisierung
- Datenanalyse
- Drittanbieterbibliotheken bzw. Erweiterungen
- Skalierbarkeit

Neben grundlegenden Aspekten wie der Codebase-Lizenzierung und der Maintainer-Community, die sich um einer bestimmten Software gebildet hat, betrachten wir die notwendige DevOps-Infrastruktur: Was ist das Frontend-Framework, was ist die Datenbanktechnologie, und welche Deployment-Pipelines sind mit diesem Tool möglich. Oft ist die notwendige Systemarchitektur die erste Hürde für Anwender:innen, wenn Web-basierte Software eine bestimmte Serverinfrastruktur und Datenbanktechnologie erfordert oder moderne Deployment-Formate wie Docker nicht unterstützt. Auch ist es wichtig zu wissen, welche Struktureinheiten ausgetauscht werden können, zum Beispiel, ob als Datenbanktyp Postgres statt MySQL oder eine Elasticsearch statt einer Solr-basierten Suche verwendet werden kann. Dies wird im Vergleichssystem oder in den für jede Software verfügbaren zusätzlichen Hinweisen erfasst.

Um die Verpflichtung zu den FAIR-Prinzipien des NFDI-Konsortiums zu erfüllen, analysieren wir auch die Prozesse zur Erstellung von persistenten Identifikatoren (URIs), die verfügbaren Abfrageschnittstellen (APIs und SPARQL-Endpunkte) und die Import-/Exportoptionen der einzelnen Software. Entitäten müssen intern und extern wiederverwendbar sein, und Daten sollten leicht importiert und exportiert werden können. Wenn Daten(-strukturen) nur innerhalb einer Software funktionsfähig sind, entstehen Abhängigkeiten, die einer flexiblen Nutzung fortschreitenden Technologien im Wege stehen können. Interoperabilität bedeutet hier, dass Daten in Formaten gespeichert oder über Schnittstellen verfügbar gemacht werden, die unabhängig von der Software sind, die sie erstellt hat. Die Bereitstellung solcher Kerninformationen über die für diese Software erforderliche technologische Infrastruktur - sowie über den Stand der Wartung und der Dokumentation - kann auch die Konzeption von Forschungsprojekten insgesamt beeinflussen. Wenn sich

Forscher beispielsweise für eine bestimmte Software entscheiden, weil sie deren besondere Funktionen benötigen, kann der Grad der Komplexität des Infrastruktur-Stacks Aufschluss darüber geben, wie viel Personal sie für das Projekt einstellen müssen, für wie viele Jahre usw. Je mehr verschiedene Technologien und Frameworks an einem System beteiligt sind, desto mehr technische Personstunden sind im Allgemeinen erforderlich. Darüber hinaus kann eine unzureichende Dokumentation oder eine kleinere Maintainer-Community auch einen höheren Zeitaufwand für Debugging usw. bedeuten. Die Vergleichstabelle kann keine exakten Schätzungen liefern, soll aber als allgemeine Orientierung dienen und helfen, die Erwartungen zu steuern.

Um Forschungsdaten wirklich interoperabel und wiederverwendbar zu gestalten - wie es für die Einhaltung der FAIR-Datenanforderungen entscheidend ist - geht es beim LOD-Management nicht nur darum, die Daten im technischen Sinne referenzierbar und zugänglich zu halten, sondern auch darum, dass die Daten von Menschen und Maschinen verstanden werden können (Sack 2021). Daher muss auch der Umgang mit der Semantik und den ontologischen Grundlagen innerhalb der Software näher betrachtet werden. Die LOD-Modellierung besteht aus zwei Schritten: dem Ontologie-Engineering für das verfügbare Vokabular und dessen Syntax, das häufig auf einem Standardmodell aufbaut, und dem Entwurf des Eingabe- und Ausgabedatenmodells. Die in unserer Übersicht untersuchten Software-Umgebungen unterscheiden sich erheblich in der Handhabung der zugrundeliegenden Semantik oder der Implementierung ontologischer Strukturen und Restriktionen in ihren Eingabe-, Darstellungs- und Verteilungsmodellen. Daher werden diese Unterschiede in unserem Vergleichssystem in separaten Kategorien näher beleuchtet.

Ein gemeinsames Merkmal vieler Software ist die in einer Benutzeroberfläche gesteuerte semantische Anreicherung des Kerndatenmodells (Nasarek 2021; ResearchSpace 2021; Rossenova 2022). In solchen Fällen kann zusätzlich zur reinen Syntax - z.B. die Entität "Person" hat die Eigenschaft "Name" - eine auf einer Ontologie basierende Semantik hinzugefügt werden: die Entität "Person" kann als vom Typ "E21_Person" spezifiziert werden, wenn das CIDOC-CRM befolgt wird. In diesem CIDOC-CRM-Beispiel würde die Eigenschaft "Name" zur Klasse "E41_Appellation" gehören und die Beziehung würde mit der Aussage "E21_Person P1_ist identifiziert durch E41_Appellation" definiert werden. Die Möglichkeit, mit einem Standard wie CIDOC-CRM zu arbeiten, insbesondere wenn sie über eine grafische Benutzeroberfläche unterstützt wird, könnte für einige DH-Projekte eine Schlüsselanforderung sein. Die Hervorhebung dieses Aspekts und die Möglichkeit, schnell und einfach zu prüfen, welche Software einen solchen Arbeitsablauf unterstützen kann und welche nicht, soll Forscher:innen und Projektbeteiligten dabei helfen, fundierte Entscheidungen über die Wahl ihrer Software zu treffen.

Um die Auffindbarkeit und Wiederverwendbarkeit von Daten weiter zu unterstützen, sollte eine LOD-Management-Software über die Standardfunktionalitäten der Dateneingabe, wie z. B. die Validierung des Datentyps oder der Kardinalität hinausgehen, indem er Verifizierungs- oder Disambiguierungsfunktionen bietet (Sack, 2021). Die Software benötigt auch Routinen für das Map-

ping, Harvesting und/oder die Integration von Daten aus anderen bestehenden LOD-Quellen, um das Potenzial eines Giant Global Graph im Sinne von Berners-Lee zu erfüllen (Anders et al. 2022; Matthiak et al. 2022). Denn was nützen Dienste wie die der Gemeinsamen Normdatei, wenn sie nicht reibungslos mit existierenden Strukturen und Datenbeständen zusammengeführt werden können?

Natürlich sollten auch die eigenen Daten möglichst integrationsfähig sein, doch der bei Linked Open Data zentrale Grundsatz der Offenheit gestaltet sich meist komplizierter. Datenkurator:innen - insbesondere im Kulturbereich, wo das Urheberrecht noch immer stark durchgesetzt wird - sehen sich beispielsweise aus Gründen des Datenschutzes allzuoft gezwungen die Nutzer- und Zugriffsrechte einzuschränken (Rodríguez-Doncel et al. 2016). Die Granularität (System, Webseite, Entität, Feld) oder Dynamik (Benutzerrolle, Entitätstyp, Feldwert) der Zugriffsberechtigungen bestimmen letztlich, wer welche Daten unter welchen Umständen sehen kann. Dieser Punkt kann zu einer vollständigen Abschottung der Datenbank nach außen führen, wenn die Berechtigungen nicht dynamisch in der Form "Zeige allen nicht angemeldeten Benutzer:innen (Rolle) nur Datensätze von Personen (Entitätstyp), die seit mehr als 75 Jahren verstorben sind (Feldwert)" festgelegt werden können. Daher war die Einbeziehung von Details zu Berechtigungssystemen in der LOD-Management-Software eine weiterer wichtiger Punkt für unser Vergleichssystem.

Hinzu kommen Funktionalitäten wie Mehrsprachigkeit - besonders wichtig, wenn Forschung über Fachgebiete und Ländergrenzen hinweg sinnvoll in Systeme des bidirektionalen Datenaustauschs integriert werden soll (Labra Gayo, Kontokostas u. Auer 2015); Möglichkeiten der Versionskontrolle bei Datenbankänderungen in verschiedenen Granularitätsgraden - bedeutend, wenn die Datenprovenienz eine entscheidende Rolle für die Vertrauenswürdigkeit der Daten und die wissenschaftliche Integrität spielt (Papakonstantinou et al. 2022); aber auch das Vorhandensein erweiterter Formen der Datendarstellung, das heißt welche Suchmasken, Formulare und Datenvisualisierungen als Zugang zur Datenbank dienen oder inwieweit Frontend-Schnittstellen angepasst werden können, um die Benutzerfreundlichkeit zu erhöhen (Garcia et al. 2016).

Nicht zuletzt spielen auch Leistung und Skalierbarkeit eine wichtige Rolle bei der Entscheidungsfindung, wenn das Forschungsprojekt Hunderttausende oder gar Millionen von Datensätzen betrifft (Duchesne 2022). Im Rahmen des Vergleichs der Software wird versucht, dies anhand von Benchmark-Beispielen zu beurteilen und die Fähigkeit eines jeden Projekts zu bewerten, Add-ons, Erweiterungen und, falls erforderlich, Techniken zur Leistungsoptimierung zu entwickeln, die an die Gemeinschaft der Software-Maintainer zurückgegeben werden können.

Ausblick

Trotz eines reichhaltigen und vielfältigen Angebot an Werkzeugen für die Verwaltung, die Veröffentlichung und den Austausch von LOD, ist die Übernahme von LOD in vielen Wissensbereichen noch eher die Ausnahme als die Regel. Innerhalb des NFDI4Culture-Konsortiums haben

wir viele Projekte und Partnerorganisationen identifiziert, die bereit sind, den Sprung zur Speicherung und Zugänglichkeit ihrer Daten als LOD zu wagen. Die Entscheidung für ein Tool hängt wie dargelegt von den Bedarfen des Projektkontextes ab, d. h. vom Umfang und Menge der Daten, der gewünschten Schemakonformität (Notwendigkeit der Einhaltung spezifischer Standards, die von ihrem Wissensgebiet festgelegt wurden), dem Grad der Benutzerfreundlichkeit (Benutzer mit wenig oder gar keiner Erfahrung in der Datenwissenschaft und Anforderungen an zugängliche grafische Benutzeroberflächen) und auch dem Grad der Offenheit (Notwendigkeit, Teile des Datensatzes zu sperren und nur Teile davon zu öffnen). Mit der Veröffentlichung unserer Vergleichstabelle und dem dazugehörigen Dokumentations-Toolkit hoffen wir, die Akzeptanz verschiedener LOD-Software-Umgebungen innerhalb der NFDI4Culture-Community und darüber hinaus zu erhöhen. In diesem Vortrag werden wir das Vergleichssystem und die wichtigsten Erfolgsgeschichten von Community-Projekten vorstellen, die sich auf der Grundlage einer fundierten Bewertung der Fähigkeiten des Werkzeugs im Vergleich zu den Anforderungen des Projekts für die Nutzung einer bestimmten Software entschieden haben.

Bibliographie

Anders, Ivonne, Tobias Arera-Rütenik, Susanne Arndt, et al. 2022. "Ontology Harmonization and Mapping - Working Group Charter (NFDI section-metadata) (1.0)." Zenodo. <https://doi.org/10.5281/zenodo.6726519>

Berners-Lee, Tim. 2007. "Giant Global Graph." Internet Archive Wayback Machine <https://web.archive.org/web/20160713021037/http://dig.csail.mit.edu/breadcrumbs/node/215> (zugegriffen: 14. Juli 2022).

Duchesne, Paul. 2022. "Wikidata and Wikibase as complementary research services for cultural heritage data." TIB blog. <https://blogs.tib.eu/wp/tib/2022/03/17/wikidata-and-wikibase-as-complementary-research-services-for-cultural-heritage-data/> (zugegriffen: 02. August 2022).

Engelhardt, Claudia u. Harald Kusch. 2021. "Kollaboratives Arbeiten mit Daten." In Putnings, Markus, Heike Neuroth u. Janna Neumann, Praxishandbuch Forschungsdatenmanagement, 451-475. Berlin, Boston: De Gruyter, Saur.

García, Roberto, Rosa Gil, Juan Manuel Gimeno, Eirik Bakke, u. David R. Karger. 2016. "BESDUI: A benchmark for end-user structured data user interfaces." In The Semantic Web - ISWC 2016. ISWC 2016. Lecture Notes in Computer Science, vol 9982. Springer, Cham. https://doi.org/10.1007/978-3-319-46547-0_8

Humboldt-Universität zu Berlin. 2020. "Digitales Netzwerk Sammlungen - ein Projekt der Berlin University Alliance." <https://www.ub.hu-berlin.de/de/ueber-uns/projekte/digitales-netzwerk-sammlungen/projekt-digitales-netzwerk-sammlungen> (zugegriffen: 02. August 2022).

Labra Gayo, Jose Emilio, Dimitris Kontokostas u. Sören Auer. 2015. "Multilingual linked data patterns." Semantic Web 6 (2015): 319-337.

- Linked Open Data Working Group. 2022a.** "Overarching technical, ethical and legal activities." <https://nfdi4culture.de/what-we-do/task-areas/task-area-5.html> (zugegriffen: 02. August 2022).
- Linked Open Data Working Group. 2022b.** "First open meeting of the 4Culture Linked Open Data Working Group." <https://nfdi4culture.de/news-events/events/first-open-meeting-of-the-4culture-linked-open-data-working-group.html> (zugegriffen: 02. August 2022).
- Mathiak, Brigitte, Heinrich Widmann, Luca Ghiringhelli, Holger Israel, Fidan Limani, Christin Henzen u. Gerhard Heyer. 2022.** "Working Group Charter - Search and Harvesting." Zenodo. <https://doi.org/10.5281/zenodo.6770763>
- Nasarek, Robert. 2021.** "Ontology Engineering." <https://wiss-ki.eu/documentation/data-modeling/ontology-engineering> (zugegriffen: 02. August 2022).
- Neumann, Janna. 2021.** "Datentransfer und -nachnutzung." In: Putnings, Markus, Heike Neuroth u. Janna Neumann, Praxishandbuch Forschungsdatenmanagement, 399-403. Berlin, Boston: De Gruyter, Saur.
- NFDI4Culture. 2022.** "NFDI4Culture - Consortium for Research Data on Material and Immaterial Cultural Heritage." <https://nfdi4culture.de/index.html> (zugegriffen: 02. August 2022).
- Papakonstantinou, Vassilis, Giorgos Flouris, Iri Fundulaki, Kostas Stefanidis u. Yannis Roussakis. 2016.** "Versioning for Linked Data: Archiving Systems and Benchmarks." BLINK@ISWC. <http://users.ics.forth.gr/~fgeo/files/BLINK16Versioning.pdf> (zugegriffen: 02. August 2022).
- ResearchSpace. 2021.** "Semantic Tools: Become author of meaningful structured data." <https://researchspace.org/semantic-tools/> (zugegriffen: 02. August 2022).
- Rodríguez-Doncel, Víctor, Cristiana Santos, Pompeu Casanovas u. Asunción Gómez-Pérez. 2016.** "Legal aspects of linked data - The European framework." Computer Law & Security Review (32) 6: 99-813, <https://doi.org/10.1016/j.clsr.2016.07.005>.
- Rossenova, Lozana. 2022.** "Examining Wikidata and Wikibase in the context of research data management applications." TIB Blog. <https://blogs.tib.eu/wp/tib/2022/03/16/examining-wikidata-and-wikibase-in-the-context-of-research-data-management-applications/> (zugegriffen: 02. August 2022).
- Rossenova, Lozana u. Lucia Sohmen. 2021a.** "BYOD workshop (Bring your own dataset) - How to use Wikibase, OpenRefine and Linked Data." <https://nfdi4culture.de/news-events/events/byod-workshop-bring-your-own-dataset-how-to-use-wikibase-openrefine-and-linked-data.html> (zugegriffen: 02. August 2022).
- Rossenova, Lozana u. Lucia Sohmen. 2021b.** "OpenRefine to Wikibase: Data Upload Pipeline." Wikiversity. https://en.wikiversity.org/wiki/OpenRefine_to_Wikibase:_Data_Upload_Pipeline (zugegriffen: 02. August 2022).
- Sack, Harald. 2021.** "The NFDI4Culture Knowledge Graph and Wikibase." WikidataCon, Oct 29-31th, 2021. https://www.youtube.com/watch?v=JieuR-Jz14Sk%26ab_channel=wikimediaDE (zugegriffen: 02. August 2022).
- W3C. 2011.** "Benefits - Library Linked Data." <https://www.w3.org/2005/Incubator/1ld/wiki/Benefits> (zugegriffen: 02. August 2022).
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. 2016.** "The FAIR Guiding Principles for scientific data management and stewardship." Sci Data 3, 160018. <https://doi.org/10.1038/sdata.2016.18> (zugegriffen: 02. August 2022).