

re3data – ein internationales Verzeichnis von Forschungsdateninfrastrukturen

Frank Scholze, KIT-Bibliothek, Karlsruher Institut für Technologie

60

Forschungsdaten sind wertvoll und allgegenwärtig. Die nachhaltige Verfügbarkeit von Forschungsdaten ist eine Herausforderung für alle Beteiligten in den wissenschaftlichen Communities. Dennoch bieten Langzeitarchivierung und permanenter Zugang zu Forschungsdaten grosse Chancen für die Wissenschaft. Es kann jedoch sehr schwierig sein, geeignete Infrastrukturen (meist Repositorien) zu finden, um Forschungsdaten zu speichern, zu bearbeiten und nachzunutzen.

re3data (<https://www.re3data.org/>) ist ein internationales Verzeichnis von Forschungsdaten-Repositorien, das nicht auf bestimmte akademische Disziplinen beschränkt ist. re3data fördert eine Kultur des Teilens, einen besseren Zugang und eine bessere Sichtbarkeit von Forschungsdaten und unterstützt die FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable). Es vermittelt Repositorien für die dauerhafte Speicherung und den Zugriff auf Datensätze für Forschende, Förderer, Verleger und Journalisten. Bei der Konzeption und Implementierung des Dienstes wurde deutlich, dass das Auffinden und Speichern von Forschungsdaten, das Analysieren, Aufbauen und Integrieren von Infrastruktur und Diensten auf der Grundlage der bereitgestellten Informationen im Zentrum stehen sollte.

Informationsquelle für ein breites Publikum

Als Einstieg in wissenschaftliche Ressourcen ist re3data nicht auf Forschende beschränkt. Es dient bereits als Informationsquelle für Förderorganisationen, politische Entscheidungsträger und die öffentliche Verwaltung. re3data ist z.B. ein Datenanbieter für den EU Open Science

Monitor. Rund 2100 Repositorien sind Stand Juni 2018 indiziert und decken ein breites Spektrum an Forschungsdaten ab. Diese Daten stehen ausser Forschenden auch für Datenjournalisten oder die interessierte Öffentlichkeit zur Verfügung. Da re3data nicht domänenspezifisch ist, deckt es eine grosse Anzahl von Fachgebieten aus den Ingenieur-, Natur-, Geistes- und Sozialwissenschaften sowie den Lebenswissenschaften ab.

Das Metadatenschema

re3data stellt strukturierte Metadaten für jedes Repository bereit und benötigt daher ein Metadatenschema, das von vielen anderen als De-facto-Standard übernommen wurde. Das Schema von re3data deckt verschiedene Aspekte von Repositorien ab (Allgemeine Informationen, Richtlinien, Rechtliche Aspekte, Technische Standards, Qualitätsstandards, Art des Zugriffs). Zur schnellen visuellen Orientierung spiegeln sich diese Kategorien auch im Icon-System von re3data wider, in dem komplexe Eigenschaften in einfachen grafischen Symbolen ausgedrückt werden.

Die Beschreibung von Forschungsdateninfrastrukturen ist aufgrund der fortgesetzten und vielseitigen Entwicklung des Forschungsdatenmanagements ein dynamischer Prozess. re3data begann 2012 mit einigen grundlegenden Eigenschaften und entwickelte sich zu einem umfassenden Metadatenschema. Derzeit ist das Metadatenschema in Version 3.0 mit insgesamt über 141 Eigenschaften verfügbar. Die anstehenden Anforderungen aus der Forschungsgemeinschaft wurden jedoch bereits identifiziert und weitere Versionen sollen veröffentlicht werden.

Vielfältige Suchmöglichkeiten

re3data zeigt Repository-Ergebnisse als eine lange einfache Liste, aber es bietet auch verschiedene Funktionen für Forscher, um Repositorien ohne viel Aufwand zu explorieren und zu finden. Neben der Suche nach Einträgen nach Stichworten enthält das Tool auch eine facettierte Suchfunktion und ein visuelles Suchwerkzeug. Dadurch können Benutzer die Entitäten für jede Eigenschaft mithilfe des Metadatenschemas filtern, sodass sie die Ergebnisliste ihren spezifischen Anforderungen anpassen können. Beispielsweise kann ein Journalist den gewünschten Themenbereich schnell eingrenzen oder Repositorien für ein bestimmtes Land suchen.

Mit seinem internationalen Bestand an Repositorien bietet re3data auch grundlegende statistische Analysen. Für weitere Untersuchungen innerhalb der Forschungsinfrastrukturlandschaft werden alle Daten über eine sogenannte API angeboten. Damit die Daten von re3data so einfach wie möglich nachgenutzt werden können, sind sie unter einer Creative Commons CCo-Lizenz kostenlos zugänglich.

Zum Autor

Frank Scholze



Frank Scholze ist Direktor der Bibliothek des Karlsruher Instituts für Technologie (KIT) und Mitglied des Bundesvorstandes des Deutschen Bibliotheksverbandes (dbv) sowie einer Reihe von wissenschaftlichen Ausschüssen und Beiräten; unter anderem bei DARIAH-DE (Digital Research Infrastructures for the Arts and Humanities). Frank Scholze ist Sprecher der

AG Elektronisches Publizieren der Deutschen Initiative für Netzwerkinformation (DINI) und ist an zahlreichen Projekten im Bereich Digitaler Bibliotheken und Forschungsinformation beteiligt.

Literatur

- Pampel, H., Bertelmann, R., Scholze, F., Kindling, M., Vierkant, P. (2015). Stand und Perspektive des globalen Verzeichnisses von Forschungsdaten-Repositorien re3data.org. In: Müller, P. (Eds.), *8. DFN-Forum Kommunikationstechnologien: Beiträge der Fachtagung 08.06-09.06.2015 in Lübeck*, (GI-Edition: lecture notes in informatics; 243), Bonn: Gesellschaft für Informatik, pp. 13–22. <http://gfzpublic.gfz-potsdam.de/pubman/item/escidoc:1169893:3>