

Qualität und Nachnutzungspotenzial von LIS-Forschungsdaten in offenen Repositorien

Katharina Schobert¹, Tamara Heck², Marc Rittberger^{1,2}

¹ DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation,
Frankfurt am Main, Deutschland

² Hochschule Darmstadt, Deutschland
{k.schobert, t.heck, m.rittberger}@dipf.de

Abstract

Die Bereitstellung von Forschungsdaten kann in den meisten Disziplinen noch nicht als Standard bezeichnet werden, aber immer mehr Forschende publizieren ihre Daten in offenen Repositorien. Hierbei sind es meist die Ersteller*innen der Daten, die diese mit Metadaten versehen und beschreiben. Neben der eigentlichen Ergebnispublikation erfordert die ordentliche Dokumentation und Beschreibung der Daten nochmals Kompetenzen und Kapazitäten, damit die Daten von Dritten auch verstanden und potenziell nachgenutzt werden können. Darüber hinaus sind die Vielzahl der Datenrepositorien groß und die Portale sehr heterogen bezüglich Möglichkeiten zur Beschreibung von Datensätzen. Im Rahmen des Forschungsdatenmanagements existieren Standards zur Beschreibung von Datensätzen. Jedoch besteht hierbei die Herausforderung, dass die Standards die unterschiedlichen Methoden der Datenerhebung und -analyse ausreichend darstellen. Diese Work-in-Progress-Studie hat verschiedene offene Repositorien auf Datensätze aus der Informationswissenschaft durchsucht und diese auf ihre Qualität und ihr Nachnutzungspotenzial hin überprüft. Als Grundlage hat diese Studie Qualitätskriterien eines Forschungsdatenzentrums aus der Bildungsforschung sowie die FAIR-Prinzipien genutzt. Die Studie geht der Frage nach, welche LIS-Daten publiziert und beschrieben werden und ob die verwendeten Qualitätskriterien auf diese Datensätze gut anwendbar sind. Ausgewertet wurden 44 Datensätze aus sechs fachübergreifenden offenen Repositorien. Die Auswertungen zeigen, dass die meisten LIS-Datensätze die festgelegten Kriterien gut erfüllen, was ihre Verständlichkeit und potenzielle Nutzbarkeit unterstützt. Herausforderungen gibt es jedoch bei der Anwendung der vorhandenen Qualitätskriterien, die Anpassungen und methodenspezifische Ansätze erfordern. Zusätzlich wurden Schwierigkeiten bei der Datensatzrecherche identifiziert. Die Ergebnisse der

Studie geben Hinweise darauf, wie die Auffindbarkeit und somit die Nachnutzung von LIS-Datensätzen verbessert werden kann.

Keywords: Forschungsdaten, Repositorium, Forschungsdatenmanagement, Informationswissenschaft, Open Science

1 Einleitung

In den letzten Jahren ist die Menge an Forschungsdaten, die in öffentlichen Repositorien zur Verfügung stehen, stark gewachsen. Diese wachsende Menge an Forschungsdaten eröffnet der wissenschaftlichen Gemeinschaft enorme Potenziale (Bambey et al., 2017). Der Zugriff auf die umfangreichen Datenbestände erlaubt es Wissenschaftler*innen, publizierte Forschungsergebnisse besser nachzuvollziehen, zu überprüfen und die Daten nachzunutzen, um entweder eigene erhobene Daten anzureichern oder mit den schon vorhandenen Daten eigene Forschungsfragen zu beantworten (ebd.; Rahal et al., 2022). Letzteres kann eine enorme Zeitersparnis für Forschende bedeuten, da sie eigene Daten für ihre Untersuchung nicht erst erheben müssen (Tenopir et al., 2020).

Um die Archivierung und den Zugang zu qualitativen Forschungsdaten zu erleichtern, hat der RatSWD¹ Empfehlungen zur Etablierung von Standards für die Archivierung und den Zugang zu Forschungsdaten, zur Entwicklung von Metadatenstandards und zur Schaffung von Repositorien für Forschungsdaten herausgegeben (Bambey/Corti et al., 2018). Damit Forschungsdaten später effektiv genutzt werden, ist es von großer Bedeutung, dass sie zuerst gefunden werden können (Fecher/Puschmann, 2015; Wallis et al., 2013). Hierbei spielen öffentliche Repositorien wie Universitäts-Repositorien und Forschungsdatenzentren (FDZ) eine entscheidende Rolle. Plattformen wie re3data agieren darüber hinaus als Metasuche und erlauben eine Recherche nach verfügbaren Datenquellen (Pampel et al., 2013; Watteler, 2022). Durch diese Infrastrukturen soll die gezielte Suche nach Forschungsdaten erleichtert und eine effiziente Nachnutzung ermöglicht werden (Feldsien-Sudhaus/Rajski, 2016; Pampel et al., 2013).

1 <https://www.konsortswd.de/ratswd>

- Heck, T.; Schneider, G. (2021): Quality of open research data in education. Vorab-Onlinepublikation. <https://doi.org/10.5281/zenodo.4672653>
- Kindling, M. (2013): Qualitätssicherung im Umgang mit digitalen Forschungsdaten / Quality assurance of digital research data / La garantie de la qualité des données numériques de recherche. In: *Information – Wissenschaft & Praxis*, 64 (2–3), 137 bis 147. <https://doi.org/10.1515/iwp-2013-0020>
- Lösch, T.; Bayer, S.; Heck, T.; Blask, K.; Perry, A.; Toepfer, R.; Vlaeminck, S. (2022): Forschungsdaten nachnutzen in der Krise: Wie wurden vorhandene Daten während COVID-19 genutzt? [Re-using research data in crisis: How was open data used during COVID-19?]. In: *9. Conference of Society of Empirical Educational Research (GEBF)*.
- Neuroth, H.; Oevel, G. (2021): Aktuelle Entwicklung und Herausforderungen im Forschungsdatenmanagement in Deutschland. In: M. Putnings, H. Neuroth, J. Neumann (Hrsg.): *Praxishandbuch Forschungsdatenmanagement* (S. 537–556). Berlin: De Gruyter Saur. <https://doi.org/10.1515/9783110657807-029>
- Neuroth, H.; Rothfritz, L.; Petras, V.; Kindling, M. (2019): Digitales Datenmanagement als neue Aufgabe für wissenschaftliche Bibliotheken. Vorab-Onlinepublikation. <https://doi.org/10.18452/20680>
- Pampel, H.; Vierkant, P.; Scholze, F.; Bertelmann, R.; Kindling, M.; Klump, J.; Goebelbecker, H.-J.; Gundlach, J.; Schirmbacher, P.; Dierolf, U. (2013): Making research data repositories visible: the re3data.org Registry. In: *PLOS ONE*, 8 (11), e78080. <https://doi.org/10.1371/journal.pone.0078080>
- Piwowar, H. A.; Vision, T. J. (2013): Data reuse and the open data citation advantage *PeerJ* 1:e175. <https://doi.org/10.7717/peerj.175>
- Rahal, R.-M.; Hamann, H.; Brohmer, H.; Pethig, F. (2022): Sharing the Recipe: Reproducibility and Replicability in Research Across Disciplines. In: M. Schubotz, S. Behrens, T. Heck, R.-M. Rahal, I. Steinhardt, D. Scholl (Hrsg.): *Bootstrapping the Open Science culture: The fellowship approach*. Pensoft Publishers. <https://doi.org/10.3897/rio.8.e89980>
- Rösch, H. (2021): Forschungsethik und Forschungsdaten. In: M. Putnings, H. Neuroth, J. Neumann (Hrsg.): *Praxishandbuch Forschungsdatenmanagement* (S. 115 bis 140). Berlin: De Gruyter Saur. <https://doi.org/10.1515/9783110657807-006>
- Scherp, G.; Siegfried, D.; Biesenbender, K.; Breuer, C. (2020): Die Bedeutung von Open Science in den Wirtschaftswissenschaften. Ergebnisbericht einer Online-Befragung unter Forschenden der Wirtschaftswissenschaften an deutschen Hochschulen 2019,. Kiel/Hamburg. <http://hdl.handle.net/11108/448>
- Schobert, K.; Heck, T.; Rittberger, M. (2023): Qualität und Nachnutzungspotential von LIS-Forschungsdaten in offenen Repositorien: Datensatz. <https://doi.org/10.5281/zenodo.8360112>