

Hasche, Nick Gerrit; Wilmers, Annika; Keller, Carolin; Bayer, Sonja; Röschlein, Jens
Forschungsdatenmanagement und Forschungssynthesen. Über die Relevanz von Open Science und Forschungsdatenmanagement aus einem Metavorhaben in der Bildungsforschung

Wilmers, Annika [Hrsg.]: *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF. Methode, Auswertung, Perspektiven.* Münster ; New York : Waxmann 2024, S. 137-159. - (Digitalisierung in der Bildung; 5)



Quellenangabe/ Reference:

Hasche, Nick Gerrit; Wilmers, Annika; Keller, Carolin; Bayer, Sonja; Röschlein, Jens:
Forschungsdatenmanagement und Forschungssynthesen. Über die Relevanz von Open Science und Forschungsdatenmanagement aus einem Metavorhaben in der Bildungsforschung - In: Wilmers, Annika [Hrsg.]: *Bildung im digitalen Wandel. 20 Forschungssynthesen im Metavorhaben Digi-EBF. Methode, Auswertung, Perspektiven.* Münster ; New York : Waxmann 2024, S. 137-159 - URN: urn:nbn:de:0111-pedocs-347179 - DOI: 10.25656/01:34717; 10.31244/9783830999126.07

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-347179>

<https://doi.org/10.25656/01:34717>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Nick Gerrit Hasche, Annika Wilmers, Carolin Keller,
Sonja Bayer & Jens Röschlein

Forschungsdatenmanagement und Forschungssynthesen

Über die Relevanz von Open Science und Forschungsdatenmanagement aus einem Metavorhaben in der Bildungsforschung

Abstract

Bisher wurden im Förderschwerpunkt „Digitalisierung im Bildungsbereich“ des Rahmenprogramms empirische Bildungsforschung 48 Forschungsprojekte gefördert, die verschiedene Aspekte der Digitalisierung in der Bildung untersuchten. Das Metavorhaben Digi-EBF wurde eingerichtet, um diese Projekte zu unterstützen, zu vernetzen und aus übergeordneter Perspektive Erkenntnisse zu sammeln und zu synthetisieren. Der Beitrag beschreibt zunächst das Konzept eines systematischen Forschungsdatenmanagements und verknüpft es mit der Idee von Open Science, ehe die Implementierung dieses Konzepts im Förderschwerpunkt dargestellt wird. Die im Kontext von Digi-EBF erprobten Möglichkeiten der Bereitstellung eigener Datensätze aus Forschungssynthesen bilden einen weiteren Fokus des Kapitels. Hierfür wurde sowohl die Literaturrecherche zur Nachnutzung zur Verfügung gestellt als auch eine Überführung von Rechercheergebnissen in einen Wissensgraphen getestet. Die abschließende Diskussion fasst die Rolle des Forschungsdatenmanagements als Grundvoraussetzung für die Produktion von FAIR-Data zusammen und reflektiert Möglichkeiten und Grenzen der Nachnutzung von Forschungsdaten sowie die Herausforderungen bei der Umsetzung von Open-Science-Praktiken in der Forschungspraxis. Final bleibt die Frage, welchen Beitrag das Metavorhaben mit den vorgestellten Praktiken für den Open-Science-Diskurs leistet.

Schlüsselwörter: Metavorhaben, Forschungsdatenmanagement, Forschungssynthesen, Wissensgraphen, Open Science

Research data management and research syntheses: On the relevance of open science and research data management from the perspective of a meta-project in educational research

To date, 48 research projects investigating various aspects of digitalisation in education have been funded within the “Digitalisation in Education” priority funding line of the Framework Programme for Empirical Educational Research. The Digi-EBF meta-project was set up to support and network these projects and to collect and synthesise findings from an overarching perspective. This chapter firstly introduces the concept of systematic research data management and links it to open science, then describes the implementation of this concept in the funded projects. We also focus on data collected in Digi-EBF from the research syntheses. The literature search was made available for subsequent use and we tested the transfer of research results to a knowledge graph. The concluding discussion summarises the role of research data management as a basic requirement for the

production of FAIR data and reflects on possibilities and limitations of reusing research data as well as on the challenges of implementing open science practices in research. Finally, the question remains how the meta-project can contribute to the open science discourse given the presented practices.

Keywords: meta project, research data management, research syntheses, knowledge graph, open science

1. Einleitung

Im Förderschwerpunkt „Digitalisierung im Bildungsbereich“ des Rahmenprogramms empirische Bildungsforschung wurden in den ersten beiden Förderrichtlinien 48 Forschungsprojekte gefördert, die verschiedene Aspekte der Digitalisierung in allen Bildungsbereichen untersuchten (BMBF, 2021). Durch die interdisziplinären Ansätze ergeben sich Synergieeffekte sowie ein umfassender Blick auf das Themenfeld. Das Metavorhaben Digi-EBF wurde zusätzlich als koordinierender Akteur eingerichtet, um die einzelnen Projekte miteinander zu vernetzen und Erkenntnisse zu systematisieren. Durch die Übersicht des Metavorhabens zu allen Forschungsvorhaben konnten beispielsweise gezielt Synergien zwischen Projekten mit ähnlichen Forschungsgegenständen oder Methoden hergestellt werden. Zusätzlich unterstützte das Metaprojekt die Forschungsprojekte darin, die empirischen Forschungsdaten, die jeweils erhoben wurden, über Forschungsdatenzentren für die Sekundärforschung zu archivieren und zu teilen. Eine übergeordnete Zielsetzung war es, in Verbindung mit weiteren Kontextinformationen über den Forschungsprozess einen ganzheitlichen Überblick über das aktuelle Forschungsfeld und wie es von den unterschiedlichen Disziplinen mit ihren spezifischen Forschungsperspektiven, -fragen und -methoden untersucht wird, zu erarbeiten. Hierzu wurde ein systematisches Forschungsdatenmanagement auf Ebene der Einzelprojekte und des Metavorhabens etabliert – Letzteres als Teil des Arbeitsbereichs Forschungsdaten Bildung am DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation.

Ein weiterer Arbeitsschwerpunkt im Metavorhaben Digi-EBF, der ebenfalls Fragestellungen des Forschungsdatenmanagements berührt, war die Erstellung von 20 Forschungssynthesen in Form von Critical Reviews. Diese Reviews fassen aktuelles Forschungswissen zu Fragen der Digitalisierung in der Bildung für die Wissenschaft systematisch zusammen.¹ Damit besitzen die Forschungssynthesen sowohl einen forschungspragmatischen als auch einen wissenschaftstheoretischen Nutzen und tragen so zu einer besseren Theoriebildung über den untersuchten Zusammenhang bei (Beilmann, 2014). Ein Anliegen des Metavorhabens bestand auch darin, die Daten aus den Forschungssynthesen im Sinne der guten wissenschaftlichen Praxis (DFG, 2019) aufzubereiten und somit zur Nachnutzung zugänglich zu machen.

1 Siehe für eine Definition des Reviewtyps den Beitrag zu Reviewformat und Projektkontext von Wilmers in diesem Band.

Der vorliegende Beitrag analysiert unterschiedliche Facetten des Forschungsdatenmanagements im Kontext von Digi-EBF und zeigt auf, wie das Metavorhaben durch seine Arbeitsschwerpunkte und die Unterstützung bei einem systematischen Forschungsdatenmanagement gezielt Open Science, bzw. die Idee einer offenen Wissenschaft förderte. In Abschnitt 2 beleuchten wir die Praxis des Forschungsdatenmanagements und die Bedeutung von Forschungsinfrastrukturen, um Informationen und Daten aus Projekten nachvollziehbar aufzubereiten und niedrigschwellig mit anderen Forschenden zu teilen. Forschungsinfrastrukturen unterstützen sowohl die Recherche als auch den Zugriff auf Publikationen über Literaturdatenbanken wie das Fachportal Pädagogik. Weiterhin werden Forschungsdaten über Forschungsdatenzentren wie dem Forschungsdatenzentrum Bildung (FDZ Bildung) aufgenommen, sichtbar und verfügbar gemacht. Eine umfassende Forschungsdateninfrastruktur für die empirische Bildungsforschung bietet der Verbund Forschungsdaten Bildung. Dessen Funktion und Nutzen für die empirische Bildungsforschung wird zum Ende des zweiten Abschnitts vorgestellt, bevor im dritten Abschnitt genauer auf das Nutzungspotenzial von Forschungsdaten in Form von Forschungssynthesen eingegangen wird. Hier wird beispielhaft aufgezeigt, wie die Zusammenarbeit zwischen dem Metavorhaben und dem FDZ Bildung für die Überführung von Datensätzen aus der Literaturrecherche der Forschungssynthesen ablief.

Darüber hinaus initiierte das Metavorhaben in Zusammenarbeit mit dem Fachinformationsdienst Erziehungswissenschaft und Bildungsforschung am DIPF ein Pilotprojekt, in dem die Überführung von Daten aus den Forschungssynthesen in den Wissensgraphen der Technischen Informationsbibliothek (TIB) erprobt wurde. Wissensgraphen haben den Anspruch, Wissen in einem aus semantisch definierten Relationen gewebten Netz abzubilden. Im vierten Abschnitt dieses Beitrags wird am Beispiel des Open Research Knowledge Graph (ORKG) gezeigt, wie strukturierte Aussagen über die in den Forschungssynthesen behandelten Studien in einem Wissensgraph dargestellt werden können. Dadurch wird eine Vernetzung und Nachnutzung der Forschungsergebnisse ermöglicht. Abschließend werden Erkenntnisse aus den verschiedenen Themenbereichen – Forschungsdatenmanagement im Metavorhaben, Forschungsdatenmanagement in den Forschungssynthesen und Datenüberführung aus Forschungssynthesen in einen Wissensgraphen – zusammenfassend reflektiert und diskutiert. Dabei hinterfragen wir auch, inwiefern die hier vorgestellten Überlegungen zum Diskurs um die Etablierung von Open-Science-Paradigmen in der Bildungsforschung beitragen können.

2. Forschungsdatenmanagement in Metavorhaben

Die Relevanz von strukturiertem Forschungsdatenmanagement (FDM) wird in der Bildungsforschung immer deutlicher. Prägnant zusammengefasst bezieht sich das Forschungsdatenmanagement auf alle Schritte der systematischen Organisation,

Erhebung, Aufbereitung, Speicherung und Dokumentation von Forschungsdaten im Forschungsprozess mit dem Ziel, diese im Sinne guter wissenschaftlicher Praxis langfristig zu sichern und im Idealfall weiteren Forschenden zur Verfügung zu stellen (Trixa & Ebel, 2015). FDM trägt also maßgeblich zur Qualität von wissenschaftlichen Forschungsprozessen und den daraus resultierenden Publikationen bei. Es unterstützt nicht nur die Primärforschenden bei der Strukturierung ihrer Arbeit und der Zusammenarbeit im Team, sondern ist auch essentiell für die Nachvollziehbarkeit von Forschungsergebnissen und ihrer Reproduktion (Fournier, 2021). Unter anderem dafür ist vor allem die Produktion von FAIRen Daten relevant. FAIR steht für Findable (Auffindbar), Accessible (Zugänglich), Interoperable (Miteinander verknüpfbar) und Reusable (Wiederverwendbar) (Wilkinson et al., 2016), d.h., die Daten sind mit beschreibenden Informationen (Metadaten) versehen, die es ermöglichen, sie zu finden und inhaltlich einzuordnen, sie sind für potenzielle Nutzer:innen möglichst ohne Einschränkungen zugänglich und langfristig mit gängigen Programmen und Standards verarbeitbar. Damit sind FAIRe Daten essentielle Bestandteile einer offenen Wissenschaft bzw. von Open Science.

Es gibt keine einheitliche Definition von Open Science, doch mehrere verbindende Elemente zwischen den verschiedenen Erklärungsansätzen. Demnach kann man Open Science als Oberbegriff für eine Reihe an Konzepten verstehen, die in der Wissenschaft zu einem offenen und freien Zugang von Erkenntnissen und Forschungsdaten beitragen sollen, um somit die Qualität in der Wissenschaft zu verbessern (vgl. u. a. DFG, 2018; Scherp, 2022; UNESCO, 2021; Van der Zee & Reich, 2018). Dies tun Open-Science-Praktiken, indem sie bspw. den Zugang zu Forschungsergebnissen, Daten und Methoden vereinfachen (Open Access und Open Data), oder zu einer Überprüfbarkeit von Forschung beitragen (Open Reproducible Research). Durch die Verfügbarmachung und Nachnutzung von Forschungsdaten werden außerdem Ressourcen geschont und das Informationspotenzial von Daten besser ausgenutzt.

Insgesamt zielt Open Science darauf ab, die Qualität der Wissenschaft zu verbessern, die Zusammenarbeit zu fördern und die Verbindung zwischen Forschung und Gesellschaft zu stärken (vgl. Krammer & Svecnik, 2020, S. 267). Ein gutes Forschungsdatenmanagement, das sich an den FAIR-Prinzipien orientiert, sollte daher als Standard etabliert werden; auch in der Bildungsforschung. Gerade bei der Aufbereitung von Forschungsdaten gibt es allerdings bisher keine einheitliche Vorstellung davon, wie Open- bzw. FAIR-Data aufbereitet sein müssen, damit sie auch wirklich als solche bezeichnet werden können. Hier gibt es zwar Bemühungen Standards zu schaffen (vgl. Nosek et al., 2015), doch bisher führen verschiedene Vorstellungen davon, wie Open Science in der Bildungsforschung umgesetzt werden soll, oftmals noch zu einer inkohärenten Durchführung des Datenmanagements (Fournier, 2021).

2.1 Relevanz von Forschungsdatenmanagement in der Bildungsforschung

Dieser Umstand zeigt sich auch in den Ergebnissen von Übersichtsarbeiten zur Reproduzierbarkeit von Forschungsergebnissen. So lassen sich in der Bildungsforschung viele publizierte Forschungsergebnisse nicht reproduzieren oder replizieren (vgl. Krammer & Svecnik, 2020). Eine zentrale Forderung der Autoren ist daher mehr Transparenz durch die Umsetzung von Open-Science-Praktiken zu schaffen (Krammer & Svecnik, 2020, S. 276f.). Gerade für die Veröffentlichung und Nachnutzung von Sekundärdaten spielt das FDM eine zentrale Rolle und wird vor allem in der empirischen Bildungsforschung, mit der Schule als einer der Hauptuntersuchungsorte, relevant. Datenerhebungen in der Schule sind sowohl für die Schulleitung, die Lehrkräfte und die Schüler:innen besondere Stresssituationen, die idealerweise so gering wie möglich gehalten werden sollten. Ein Umstand der ab 2020 in Folge der Covid-19-Pandemie noch verschärft wurde: Der Präsenzunterricht wurde eingeschränkt und auf einen digitalen Unterricht waren die meisten Schulen nicht vorbereitet. Durch die Herausforderungen, den Schulbetrieb aufrecht zu erhalten, waren zusätzliche Belastungen durch Datenerhebungen oftmals nicht zumutbar; gleichfalls standen die Schulen nun eben wegen der aktuellen Situation im Fokus neuer Forschungsvorhaben (vgl. Fickermann & Edelstein, 2021). Indem Forschungsdaten, die in Schulen erhoben werden können, mit anderen Forschenden geteilt werden, wird die weitere Belastung von den Bildungsinstitutionen reduziert.

Aber nicht nur in Hinblick auf die Transparenz und Sekundärdatenforschung kann ein strukturiertes FDM einen Mehrwert für die Bildungsforschung bringen. Weitere Teilbereiche des FDM sind Forschungsethik und Datenschutz. Ein systematisches FDM denkt diese Punkte von Beginn des Forschungsvorhabens an mit, was insbesondere dann relevant wird, wenn in Bildungseinrichtungen oder mit Kindern geforscht wird. Vor allem Minderjährige gelten als besonders schutzbedürftig: Eine unzureichende Anonymisierung von Forschungsdaten kann so bspw. zu erheblichen rechtlichen Konsequenzen führen. Daher werden beim FDM auch rechtliche Rahmenbedingungen berücksichtigt, notwendige Berechtigungen eingeholt (bspw. zur Erhebung von personenbezogenen Daten oder zum Umgang mit urheberrechtlich geschütztem Material) und die Daten entsprechend der Auflagen und Genehmigungen (bspw. Anonymisierung von Daten) behandelt (Putnings, 2021; Trixa & Ebel, 2015).

Ein weiterer wichtiger Aspekt, der in der Bildungsforschung oftmals kritisiert wird, ist die fehlende Praxisrelevanz und der Praxistransfer. Die Entfremdung der Wissenschaft zur Praxis oder auch der „research-practice gap“ wird in der modernen Wissensgesellschaft bereits seit etlichen Jahren besonders kritisch beobachtet (Benner, 1980). So werde die Forschung nicht selten als „realitätsferne Instanz im gemeinhin bekannten Elfenbeinturm gesehen, die sich mit Fragen beschäftigt, die keine Relevanz für die Praxis hätten, und die Wissen generiere, das in der Praxis nicht als verständlich, evident, anschlussfähig oder zumutbar [...] wahrgenommen werde

und somit die Bedürfnisse der Praxis nicht treffe“ (Diekmann, 2022, S. 2). Auch hier können und sollen FDM und Open-Science-Praktiken dazu beitragen, dass die Ergebnisse aus der Wissenschaft verständlich und nachvollziehbar aufbereitet und in die breite Gesellschaft sowie in die Praxis transferiert werden.

2.2 Anforderungen an die Förderprojekte und das Metavorhaben

Gerade der Transfer von Forschungsergebnissen war ein besonderes Anliegen des Metavorhabens. So ging bspw. mit der Förderung des BMBF im Förderschwerpunkt *Digitalisierung im Bildungsbereich* aus dem Rahmenprogramm empirische Bildungsforschung unter anderem die Verpflichtung einher, die Ergebnisse aus den Forschungsprojekten nicht nur innerhalb der Wissenschaft zu publizieren, sondern diese auch für ein breites Publikum zugänglich zu machen. Somit sollen sowohl der Praxistransfer, als auch Open Science und Transparenz gefördert werden. „Eine freie Wissenschaft lebt von der Publikation ihrer Ergebnisse. Wissenschaftlicher Erkenntnisgewinn lebt von Offenheit: der Offenheit zur Diskussion über Forschungsfragen, der Offenheit über den vielversprechendsten Weg zur Lösung dieser Fragen, der Offenheit, Forschungsergebnisse immer wieder neu zu hinterfragen, sie zu verwerfen oder zu vertiefen. Der freie Zugang zu den Ergebnissen öffentlich finanzierter Wissenschaft trägt wesentlich zu dieser Offenheit bei“ (BMBF, 2023, S. 1). Eine Offenheit bei Wissenschaftspraktiken zu erreichen ist allerdings gar nicht so einfach. Zwar befürwortet der größte Teil der Wissenschaftler:innen Open Science, das aktuelle akademische System würde diese aber oft nicht ausreichend unterstützen, so Nosek und Kolleg:innen (2015).

Open-Science-Paradigmen gehen eng einher mit gelebter offener Forschungspraxis. Das BMBF versuchte diese top-down über Auflagen in ihren Ausschreibungen zu fördern, bspw. durch die Verpflichtung, Daten über Repositorien oder Forschungsdatenzentren zu veröffentlichen, vor der Datenerhebung eine mögliche Sekundärdatennutzung in Betracht zu ziehen und bereits bei Projektantrag relevante Aspekte in FDM-Plänen zu adressieren (BMBF, 2017). Eine offene Forschungspraxis und systematisches FDM erfordern auf der anderen Seite erhöhte Ressourcen bei den Forschenden (vgl. Scherp, 2022), insbesondere zeitliche Ressourcen. Daher unterstützt das BMBF Aufwände für das FDM finanziell, indem entsprechende Kosten bei Antragstellung mit eingereicht werden können. Bei großen Förderlinien, wie dem hier beispielhaft beschriebenen Förderschwerpunkt Digitalisierung im Bildungsbereich wurde zudem eine zentrale Stelle zur Unterstützung des FDM in den Einzelprojekten durch das BMBF im Metaprojekt etabliert.

2.3 Umsetzung der Vorgaben im Metavorhaben

Dem Metavorhaben kommen nach den Vorgaben des BMBF zwei entscheidende Aufgaben zu: Zum einen werden in dem Metavorhaben Forschungsergebnisse in Form von Critical Reviews gebündelt, analysiert und für eine breitere Öffentlichkeit, sowohl wissenschaftlich als auch aus der Praxis, aufbereitet und zugänglich gemacht. Dabei werden potenzielle Forschungslücken, unerforschte Themenbereiche und Implikationen für die Bildungspraxis identifiziert (Diekmann, 2022). Zum anderen bietet das Metavorhaben eine projektspezifische Unterstützung der Projekte bei den Anforderungen des Forschungsdatenmanagement an und orientiert sich dabei an der umfassenden Zuordnung von Trixa und Ebel: „Alle Aspekte, die im Zusammenhang mit der Verwaltung, Organisation und Verarbeitung von Daten stehen, sind [...] Teil des Datenmanagements“ (Trixa & Ebel, 2015, S. 6). Beginnend mit Informationen zum Datenmanagementplan und Informationen zu rechtlichen Herausforderungen bei der Erhebung, Forschung und ggf. Weitergabe der Daten an ein FDZ gibt die FDM-Stelle am Projektzyklus orientierte Informationen und bespricht konkrete Praxisfragen der Forschenden zum FDM. Im weiteren Verlauf werden Fragen der Dokumentation, Datensicherung und Aufbereitung besprochen. Zudem ist zu klären, welche Daten in welcher Aufbereitungs- und ggf. Anonymisierungsstufe für die Nachnutzung bereitgestellt werden können und welche Infrastruktur/welches FDZ sich für die spezifischen Daten eignet. Das Metavorhaben unterstützt die Projekte somit über den gesamten Forschungsprozess bis zur Archivierung und Veröffentlichung der Daten (Meyermann, 2012).

Der konkrete Unterstützungsbedarf bei den verschiedenen Aspekten des Forschungsdatenmanagements wird durch mehrere Projektbefragungen deutlich, die vom Metavorhaben zu Beginn, zur Mitte und zum Ende der Projektlaufzeiten durchgeführt wurden. Dabei gaben fast 60 % der befragten Mitarbeitenden aus den Projekten an, Unterstützungsbedarf beim Forschungsdatenmanagement zu haben (LearningLab Essen, 2020). Etwa der gleiche Anteil gab zum Ende ihrer jeweiligen Projektlaufzeit an, das Angebot der FDM-Stelle auch wahrgenommen zu haben (ebd.). Bereits im ersten Jahr des Metavorhabens haben 14 der damals gestarteten 29 Projekte mindestens einmal von der Beratung Gebrauch gemacht. Die Unterstützung bestand dabei in der Regel aus einer projektspezifischen Beratung. Zudem wurden mehrere Workshops zu entsprechenden Themen und offene Beratungsangebote auf regelmäßig stattfindenden Projektforen organisiert und durchgeführt.

Die meisten dokumentierten Anfragen im Zeitraum von 2019 bis 2021 bezogen sich dabei auf die Themen Datenschutz und der Formulierung von informierten Einwilligungserklärungen. Eine typische Aufgabe des Metavorhabens bestand darin, diese gegenzulesen, um u. a. die Verständlichkeit (Informiertheit), Vollständigkeit (z. B. Rechte der Teilnehmenden) und getroffene Vereinbarungen (z. B. im Umgang mit den Daten) auf potenzielle Fallstricke zu prüfen (siehe Meyermann & Porzelt, 2019, S. 16ff.). Die informierte Einwilligung und die darin getroffenen Vereinbarungen

sind zentral für den Umgang mit den (personenbezogenen) Daten und legen den Grundstein für die Möglichkeiten, die (personenbezogenen) Daten zur Nachnutzung bereitzustellen (vgl. Fournier, 2021; Putnings, 2021).

Aufgrund der Covid-19-Pandemie und der damit einhergehenden Kontaktbeschränkungen und Schulschließungen waren rechtliche Fragen für alternative Datenerhebungsmethoden, z. B. per Videokonferenzen ein wiederkehrendes Thema bei den Beratungsanfragen. In Folge der starken Nachfrage, auch über die Projekte der Förderlinie hinaus, wurden in Kooperation mit dem Verbund Forschungsdaten Bildung spezifische Informationen und Handreichungen erstellt, die bspw. Hinweise zum Datenschutz bei Datenerhebungen mittels Videokonferenzsysteme oder zur Nutzung von Social-Media-Daten in der Bildungsforschung geben (Bayer et al., 2021).

Unter anderem konnte durch diese Form der Unterstützungen sichergestellt werden, dass ein systematisches FDM in den meisten Projekten umgesetzt wurde und ein rechtssicherer Umgang mit den Forschungsdaten gewährleistet war. Ab 2022 waren die meisten Datenerhebungen in den Forschungsprojekten abgeschlossen und die Projekte der ersten und zweiten Förderlinie befanden sich bereits in ihrer Endphase. Dementsprechend gab es ab diesem Zeitpunkt hauptsächlich Beratungsanfragen zur Datenarchivierung und -veröffentlichung. Für die Veröffentlichung der Forschungsdaten wurden die Projekte an den Verbund Forschungsdaten Bildung verwiesen, welcher zu diesem Zwecke und für die Nachnutzung eine entsprechende Infrastruktur bereitstellt.

2.4 Forschungsdatenmanagement mit dem Verbund Forschungsdaten Bildung

Dateninfrastrukturen sind aus technischer Sicht (digitale) Orte, Prozesse und Technologien für die Verarbeitung von Daten wie das Archivieren, Teilen, Finden und Nutzen. Umfassender betrachtet beinhalten Dateninfrastrukturen auch Personal, Expertise und Standards für die Implementation von Prozessen und Technologien sowie den Umgang mit den Daten (National Academies of Sciences, Engineering, and Medicine, 2023). Für die vertrauenswürdige und sichere Bereitstellung von Daten aus der Forschung unterstützt die öffentliche Hand den Ausbau von Forschungsdateninfrastrukturen als Teil der nationalen Forschungsdatenstrategie, welche geteilte Forschungsdaten als entscheidende Grundlage für innovative Forschung sieht (Die Bundesregierung, 2021). Gleichzeitig gilt es Herausforderungen zu adressieren, welche mit den geteilten Daten einhergehen: Hierzu zählen unter anderem rechtliche Aspekte, die transparente Beschreibung von Daten, die langfristige technische Nutzbarkeit sowie die effiziente Recherche von Datenbeständen (s.o. FAIRe Daten; Wilkinson et al., 2016). In den letzten Jahren sind durch entsprechende Initiativen, geförderte Einzelprojekte und gezielte Förderrichtlinien zum Aufbau von Dateninfra-

strukturen verschiedene Datenservices entstanden, deren Vernetzung und Harmonisierung ein Ziel im Aufbau der Nationalen Forschungsdateninfrastruktur (NFDI) ist.

- Über *Repositorien* können Daten niedrigschwellig gespeichert, geteilt und heruntergeladen werden. Liegen keine rechtlichen oder ethischen Gründe für die Einschränkung des Datenzugangs vor, können Daten unter einer freien Nutzungs- lizenz über ein existierendes Datenrepositorium angeboten werden. Ein disziplin- übergreifendes Repositorium, welches national gefördert wird und Datenverarbei- tung entsprechend der DSGVO zusichert, ist Zenodo (<https://zenodo.org/>). Zur weiteren Definition des Begriffs „Repositorium“ und weitere Beispiele siehe die Handreichung des RatSWD (Buck et al., 2022)
- Forschungsdatenzentren bieten umfassende Services für Forschende an, die ihre Daten teilen möchten und ermöglichen in der Regel verschiedene Zugangswe- ge und -beschränkungen entsprechend der Sensibilität der Forschungsdaten. Als Qualitätsmerkmal für Forschungsdatenzentren im Bereich der Bildungsforschung gilt die Akkreditierung durch den RatSWD².
- Vernetzende Infrastrukturen und übergreifende Services haben zum Ziel, For- schende dabei zu unterstützen, geeignete Repositorien und Forschungsdaten- zentren für das Teilen und Suchen von Forschungsdaten zu finden oder Informa- tionen zu verschiedenen Bereichen des Forschungsdatenmanagements zu geben.

Für die Bildungsforschung stellt der Verbund Forschungsdaten Bildung (Ver- bundFDB) eine entsprechende Infrastruktur bereit. Im VerbundFDB sind einschlä- gige Forschungsdatenzentren und Institutionen zusammengeschlossen³, die eine zen- trale Abgabe und Bereitstellung von Forschungsdaten bieten sowie die Suche von Forschungsdaten ermöglichen – über verschiedene Forschungsdatenzentren, Insti- tutionen und Forschungsnetzwerke hinweg. Damit soll es Forschenden erleichtert werden, einen geeigneten Ort zu finden, um Forschungsdaten zu teilen, denn die Forschungsdatenzentren haben jeweils eigene thematische und methodische Schwer- punkte bei der Archivierung von Forschungsdaten. Gleichzeitig soll die zentrale Datensuche den Rechercheaufwand minimieren. Stark ausgebaut wurde in den letz- ten Jahren das Informationsangebot bei Fragen zum Umgang mit Forschungsdaten, insbesondere bei der Gestaltung von Datenmanagementplänen, bei ethisch-rechtli- chen Aspekten und bei der Dokumentation und Aufbereitung. Schwerpunkte in der Beratung betreffen die Datenerhebung bei Minderjährigen, besondere Auflagen bei Erhebungen an Schulen, Urheberrechte an Produkten von Lernenden oder Schul- material. Der VerbundFDB adressiert dabei zum einen Forschende direkt und unter- stützt andererseits beim Aufbau von Beratungsangeboten an Universitäten oder in Forschungsverbünden wie Digi-EBF.

2 Kriterien zur Akkreditierung siehe unter <https://www.konsortswd.de/angebote/datenzentren/qualitaetssicherung-zertifizierung/akkreditierung/>

3 Für eine vollständige Übersicht über alle Verbundpartner siehe Forschungsdaten Bildung: <https://www.forschungsdaten-bildung.de/>

Neben der Erstberatung und Vermittlung der Forschungsprojekte aus dem Förderschwerpunkt wandte sich das Metavorhaben auch direkt an das FDZ Bildung als Teil des Verbundes FDB, um die eigenen Datenerhebungen in Form von Literaturrecherchen als Bestandteile der Forschungssynthesen zu archivieren und zur Verfügung zu stellen. Für eine solche Archivierung von Daten aus Forschungssynthesen lagen bis zu diesem Zeitpunkt indes keine Erfahrungswerte im FDZ Bildung vor. In Form eines Pilotprojektes wurden die Daten der Critical Reviews in Digi-EBF gemeinsam mit dem FDZ Bildung aufbereitet und zugänglich gemacht. Die Aufbereitungsschritte werden im Folgenden beschrieben und reflektiert.

3. Forschungssynthesen und Forschungsdatenmanagement in Digi-EBF

3.1 Der Umgang mit Forschungsdaten aus Forschungssynthesen

Der grundlegende Prozessschritt für Reviews ist die Literatursuche, auf deren Basis anschließend die Datenerhebung und Auswertung der Literatur stattfinden kann. In der Literatursuche werden mittels verschiedener Suchstrategien Titel recherchiert und zusammengetragen. Die genaue Vorgehensweise hängt dabei sowohl von der grundsätzlichen Methodik von Forschungssynthesen beziehungsweise dem ausgewählten Format einer Forschungssynthese als auch von den für ein jeweiliges Projekt festgelegten Ein- und Ausschlusskriterien ab. Zu letzterem zählt beispielsweise die Wahl bestimmter Datenbanken, Publikationsformen oder Sprachen, in denen Literatur erschienen ist. Durch dieses Vorgehen entsteht eine Datenmenge, die sich auch als Forschungsdatensatz einer Reviewstudie bezeichnen lässt. Handbücher zur Erstellung von Reviews empfehlen grundsätzlich, das Suchprotokoll in einem zusätzlichen Dokument, dem Appendix oder einem Supplement, aufzubereiten und darin die Suchstrategie so ausführlich wie möglich zu dokumentieren. Ein relevanter Hinweis ist dabei die Zugänglichkeit der Forschungsdaten dauerhaft zu gewährleisten, beispielsweise über die Aufbereitung der Daten in einem Online-Repository und einer entsprechenden Verlinkung zu dieser Datenarchivierung in der methodischen Darstellung eines Reviews (z.B. Cochrane Information Specialist Support Team, 2023; Collaboration for Environmental Evidence, 2023; Higgins et al., 2023; Institute of Medicine, 2011; Kugley et al., 2017; Rethlefsen et al., 2021).

Trotz der Empfehlungen zur öffentlich zugänglichen Datenablage ist es keine Selbstverständlichkeit, dass Forschungsdaten zusätzlich zur Publikation der Forschungssynthese selbst tatsächlich auch eingesehen werden können. Beispielsweise haben Nordström, Kalmendal und Batinovic (2023) in einem Meta-Review Systematic Reviews aus dem Themenfeld *Educational Effectiveness* untersucht, um der Frage nachzugehen, inwieweit in diesen Reviews methodische Verzerrungen auftreten beziehungsweise inwieweit es möglich wäre, die jeweiligen Forschungssynthesen zu re-

produzieren. Letzteres traf nur auf einen sehr geringen Anteil der identifizierten und ausgewählten Studien zu: Unter denjenigen 10 Studien mit einem geringen Bias-Risiko (ein Anteil von ca. 11 % der Studien insgesamt) ließ sich die jeweilige Suche beispielsweise nur in 6 Studien reproduzieren (Nordström et al., 2023). Auch die Untersuchung von Buntins et al. (2023) zu Forschungssynthesen aus dem Themenfeld der *Educational Technologies* weist mit Blick auf Replizierbarkeit auf einen hohen Anteil an Reviews mit unvollständiger methodischer Darstellung hin. Eine großangelegte Umfrage untersuchte zudem die Einstellung von Autor:innen von Systematic Reviews aus dem medizinischen Bereich zu Data Sharing und Reproduzierbarkeit von Forschungssynthesen. Die meisten der Befragten maßen beiden Punkten (Data Sharing und Reproduzierbarkeit) eine wichtige Funktion zu (84 %, bzw. 89 %); die Möglichkeiten, dies auch im Arbeitsalltag umzusetzen, wurden aufgrund verschiedener Faktoren (z. B. Verteilung von Ressourcen, Vorhandensein von unterstützenden Maßnahmen oder Wahl von anderen Prioritäten) aber als deutlich niedriger eingestuft (Nguyen et al., 2023).

In den vier Bänden der Reihe „Digitalisierung in der Bildung. Forschungsstand und -perspektiven“, erhielt die Darstellung der Literatursuche jeweils ein eigenes Kapitel. Zusätzlich wurden die gefundenen Titel und Treffermengen aus der Recherche in Excel-Tabellen festgehalten sowie die Dokumentation des Vorgehens in einem Methodenbericht veröffentlicht. Die Daten wurden über das FDZ Bildung am DIPF zur Verfügung gestellt und damit eine zitierfähige Veröffentlichung der Forschungsdaten sichergestellt (Keller & Rittberger, 2020, 2021, 2023a, 2023b). Für das Forschungsdatenzentrum handelte es sich dabei um ein Pilotprojekt, für das erst entsprechende Anpassungen im Workflow vorgenommen werden mussten, da bis zu diesem Zeitpunkt noch keine Datensätze aus Forschungssynthesen dort archiviert worden waren. Zwischenzeitlich haben auch andere Forschungsprojekte ihre Forschungsdaten aus Forschungssynthesen an das FDZ Bildung abgegeben.⁴

3.2 Datenaufbereitung und Dokumentation

Im Projekt Digi-EBF wurden die Forschungsdaten aus den insgesamt 20 Literaturrecherchen in Form von 20 Literaturlisten an das Forschungsdatenzentrum Bildung übergeben. Nach Abschluss der Literaturrecherchen zu einem Oberthema stellte das Projekt pro Band jeweils fünf Listen sowie einen Methodenbericht, der das Vorgehen dokumentarisch erfasst, zur Verfügung. Dies entspricht einer Liste pro recherchiertem Bildungssektor. Die Listen enthalten die gesammelten und aufbereiteten Rechercheergebnisse der Literatursuchen. Dabei gestaltete sich die Datenerhebung über die gesamte Publikationsphase aller Bände weitgehend gleich: Unabhängig von den an-

4 „Forschungssynthesen zur Bildungsforschung“ (ForSynBiFo): <https://www.fdz-bildung.de/studiendetails.php?id=503>, und „Genese sozialer Ungleichheiten des Bildungserwerbs“ (GesUB): <https://www.fdz-bildung.de/studiendetails.php?id=518> (beides zuletzt aufgerufen am 26.04.2024)

gewandten Suchtaktiken wurden während der Recherchen die Ergebnisse über das Literaturverwaltungsprogramm Citavi gesammelt. Bereits im Erhebungsprozess fand auch mit Blick auf die spätere Aufbereitung der Daten für ihre Nachnutzung eine Filterung doppelter Werte statt, indem die Datensammlung ausschließlich um neue Treffer angereichert wurde. Insbesondere Mehrfachanfragen in einer Datenbank generieren viele Dopplungen von Titeln. Eine dokumentarische Erfassung der totalen Treffermengen einzelner Suchanfragen in den Datenbanken erfolgte in dem zusätzlich angefertigten Methodenbericht, während die bereinigten Literaturlisten folglich nicht mehr die ursprünglichen Rechercheergebnisse widerspiegeln.

Für die Weitergabe der Forschungsdaten wurde Excel als geeignetes Format bestimmt und die Citavi-Projekte wurden entsprechend überführt. Im Anschluss an diese Übertragung erfolgte erneut eine Prüfung der Titel auf Dopplungen sowie auch auf fehlerhafte Angaben, vor allem bei der Nennung von Autor:innen, Titel, Jahreszahlen und bibliographischen Angaben zur Bestimmung der Quelle. Hierfür waren Korrekturen notwendig sowie Ergänzungen, um den Zugriff bzw. eine eindeutige Bestimmung des Titels zu gewährleisten. Infolgedessen wurden vor allem Digital Object Identifier (DOI), Online-Adressen sowie ISS- und ISB-Nummern recherchiert und manuell hinzugefügt. Vereinzelt konnten in diesem Schritt der Datenaufbereitung weitere Titel ausgeschlossen werden, die nicht den Ein- und Ausschlusskriterien des jeweiligen Reviews entsprachen, aber trotzdem im Literaturkorpus enthalten waren (z. B. unpassendes Publikationsjahr, andere Sprache als Deutsch oder Englisch). Auch wenn die erfassten Publikationen in mehreren der durchsuchten Quellen zu finden sind, wurde für jeden einzelnen Titel jeweils die Quelle angegeben, aus welcher er entnommen wurde. Damit ist es für Nutzende der Daten schneller möglich den Titel selbst zu recherchieren.⁵

Für die Dokumentation der jeweiligen Suchstrategien wurde für jeden Band ein eigener Beitrag im jeweils ersten Kapitel zum methodischen Vorgehen vorgesehen, der die Darstellung der Suchtaktiken aller in dem Band enthaltenen Critical Reviews bündelt. Der bereits erwähnte Methodenbericht schlüsselt darüber hinaus das detaillierte Vorgehen zur gesamten Suchstrategie auf. Diese Gesamtdokumentation konnte in dem jeweiligen Methodenkapitel geschlossen zitiert werden und ist über einen Digital Object Identifier (DOI) eindeutig und langfristig zugänglich. Der Methodenbericht ist in verschiedene Kapitel untergliedert: Das erste Kapitel enthält in tabellarischer Form die verwendeten deutschen und englischen Suchbegriffe pro Bildungsbereich. Die weiteren Kapitel enthalten die Suchstrategien nach Bildungssektoren geordnet, wobei die Einzeldarstellung der Suchanfragen nach den jeweiligen Quellen gegliedert ist. Abbildung 1 veranschaulicht dieses Vorgehen exemplarisch anhand zweier Datenbanken.

5 Vgl. Keller & Heck in diesem Band, die die Sättigung der angewandten Quellen anhand der in den Reviews ausgewählten Studien untersuchen.

Web of Science

TS= teacher* AND TS=("CLASSROOM MANAGEMENT" OR "CLASSROOM ORGANIZATION" OR "CLASSROOM ORGANISATION" OR "CLASSROOM OBSERVATION" OR "CLASSROOM AWARENESS" OR "CLASSROOM COMPLEXITY" OR "INTERPERSONAL OR MONITOR*" OR "CLASSROOM COMMUNICATION" OR "TIME ON TASK" OR "DISCIPLINE REFERRAL" OR "DISCIPLINE POLICY" OR "SCHOOL DISCIPLINE" OR "BULLYING OR PRAISE OR NONVERBAL OR DISRUPTION OR DISRUPTIVE OR "GROUP CONTINGENCY" OR "GROUP CONTINGENCIES" OR "CLASSROOM COMPLEXITY" OR "TEACHING STRATEGIES" OR "behavior referrals" OR "behaviour referrals" OR "behavior management" OR "behaviour management" OR "misbehavior management" OR "misbehaviour management" OR "student behavior" OR "student behaviour" OR "student mis-behavior" OR "student mis-behaviour" OR "behavior problems" OR "behaviour problems" OR "behavior change" OR "behaviour change" OR "positive behavioral support" OR "positive behavioural support" OR "appropriate behavior" OR "appropriate behaviour") AND TS=(VIRTUAL* OR "VR" OR "MIXED REALITY" OR "AUGMENTED REALITY" OR "EXTENDED REALITY" OR IMMERSIVE* OR IMMERSION* OR "360 * VIDEO" OR SIMULATION* OR SIMULATED* OR AVATAR OR "SPHERICAL VIDEO" OR "VIRTUAL PLEX" OR "TEACHLIVE" OR "BREAKING BAD BEHAVIORS" OR "SIMINCLASS" OR "IVT-T" OR MURSION OR "SECOND LIFE")

217 Treffer 10.02.2022

TS = (Searches for topic terms in the following fields within a record: Title, Abstract, Author Keywords, Keywords Plus)

BASE

(Klassenführung* Unterrichtsführung Unterrichtsmanagement Klassenorganisation Klassenmanagement* "Classroom Management" Klassenklima Unterrichtsklima Monitor* Kommunikation Disziplin* Unterrichtsstörung* Schülerverhalten) AND (virtuell* simulation* Immersion Immersiv "360 * Video" "augmented reality" "mixed reality" "spherical video" "Virtual PLEX" "TeachLive" "Breaking Bad Behaviors" "SimInClass IVT-T Mursion "second life") AND (Lehrkräfte* Lehrer* Lehramt* Lehrkraft*)

31 Treffer 10.02.23

Abbildung 1: Recherchedokumentation am Beispiel von Suchabfragen im Bildungssektor
Lehrkräftebildung für die Datenbanken Web of Science und BASE

Um eine bestmögliche Transparenz des Vorgehens zu gewährleisten, wurden die verwendeten Suchanfragen aus den Datenbanken kopiert, die Treffermengen vermerkt und das Datum der Suche dokumentiert. Zusätzlich wurden Filterfunktionen angegeben, sofern diese in der Recherche Anwendung fanden. Auch wenn die Literaturlisten nicht den vollständigen Treffermengen entsprechen, da keine doppelten Werte aufgenommen wurden, kann der/die Nutzer:in dieser Daten oder der/die Leser:in des Reviews so nachvollziehen, welches Vorgehen den Listen zu Grunde liegt und auf welchen Literaturbestand sich die spätere Studienauswahl im Screening bezieht. Durch die offene Darstellung des Erhebungsprozesses können die Studienergebnisse sowie auch die Zahlen zu den Treffermengen besser in den Kontext des methodischen Vorgehens gesetzt werden. Darüber hinaus können die Suchen nachgestellt, evaluiert und aktualisiert werden und somit Forschenden von Reviews mit ähnlichen Fragestellungen von Nutzen sein.

3.3 Rückblick und Lernprozesse

Über die Projektlaufzeit der vier Bände hinweg wurde das methodische Vorgehen in der Literatursuche sowie deren Dokumentation modifiziert. Für die Abgabe der Daten der ersten fünf Reviews (Bd. 1) wurden zunächst Literaturlisten weitergege-

ben, die vereinzelt Titel mit unvollständigen Metadaten enthielten. Diese Listen stellen sozusagen „Rohdaten“ dar, indem sie in der ursprünglichen Form belassen wurden, die beim Datenexport aus den Datenbanken in die Citavi-Datei entstand. Wie die Datenqualität der einzelnen Titel beschaffen ist, hängt zum einen von den recherchierten Quellen, den Datenbanken, ab; zum anderen kann die Qualität auch mit dem jeweiligen Download-Format zusammenhängen. Für die Literaturrecherchen aller vier Bände wurde immer das Datenformat BibTex (.bib), EndNote format (.enw) oder MEDLINE/PubMed-style (.nbib) verwendet. Die Datenqualität bezieht sich dabei auf die Vollständigkeit und Richtigkeit der Metadaten. Beispielsweise können Angaben wie eine DOI fehlen oder auch Sonderzeichen im Titel oder bei Autorennamen nicht erkannt und verschlüsselt wiedergegeben werden. Es kann sein, dass ein Titel in einem solchen Fall nicht mehr eindeutig identifizierbar ist. Um solche Unklarheiten zu vermeiden, wurde für die folgenden drei Bände daher beschlossen, die Daten entsprechend aufzubereiten und zu ergänzen. Diese Aufbereitung der Forschungsdaten stellte sich durch zahlreiche Korrekturvorgänge und Ergänzungen der Metadaten indes als sehr zeitintensiv heraus. Um eine fristgerechte Abgabe der Daten an das Forschungsdatenzentrum zu sichern, empfiehlt es sich, diesen Arbeitsschritt mit in die Prozessfolge einzuplanen. Nachdem das Format der abzugebenden Daten einheitlich festgelegt war, stellte sich ein Workflow zur Weitergabe an das Forschungsdatenzentrum ein.

Weiterhin erforderten die unterschiedlichen Forschungsfragen der Reviews eine Erweiterung der Suche um passende Recherchestrategien. Um nach Studien außerhalb von Datenbanken zu recherchieren, kamen verschiedene Suchtaktiken zum Einsatz. Für die Dokumentation des Vorgehens bedeutet das, zusätzlich zur Datenbankrecherche auch dieses Vorgehen nachvollziehbar darzustellen, indem die Quellen, das jeweilige Datum sowie die Treffermenge vermerkt werden. Durchsuchte Webseiten oder händisch geprüfte Zeitschriftentitel sind dabei ebenfalls anzugeben. Bei einer Zitationssuche sind im Methodenbericht die Ausgangstitel genannt sowie die Titel, die durch die Vor- oder Rückwärtssuche in Google Scholar aufgenommen wurden. Für mehr Transparenz ist es von Vorteil alle Titel anzugeben, die durch die erweiterten Suchtaktiken gefunden wurden, und nicht nur diejenigen, die in der Studienauswahl der Autor:innen enthalten sind. Denn auch die unpassenden Publikationen wurden von den Autor:innen geprüft und aus bestimmten Gründen als unpassend definiert. Mit diesem Vorgehen können die Titel, die sich in den Literaturlisten befinden, aber nicht in den Datenbanken enthalten sind, trotzdem zugeordnet bzw. deren Ursprung dargestellt werden.

Zusammenfassend kann festgehalten werden, dass die Dokumentation im Laufe der Arbeiten immer umfangreicher wurde. Der Anspruch der Darstellung weitete sich dahingehend, mit jedem Band mehr Informationen weiterzugeben und bessere Varianten zu finden, damit Außenstehende einen schnellen Überblick über die Methodik erhalten und das methodische Vorgehen nachvollziehen können. Dabei zeigte sich auch, dass die vorgegebenen Eingabefelder in der Maske des FDZ Bildung mit

Blick auf Forschungssynthesen nur nach bestimmten Informationen filtern lassen, aber nicht alle Charakteristika von Forschungssynthesen abdecken. Um die Daten auch ohne eine DOI zu finden, ist eine eindeutige Verschlagwortung während der Datenaufnahme zudem unerlässlich.

4. Repräsentation und Nachnutzung von Forschungssynthesen in Wissensgraphen

Bei der Erstellung von Forschungssynthesen in der Bildungsforschung werden Studien hinsichtlich ihres Forschungsdesigns, der Methoden und der damit erzielten Ergebnisse vergleichend betrachtet. Grundlegende Aufgaben der Autor:innen bestehen darin, die zu vergleichenden Merkmale einer Studie zu identifizieren, einem (nur lose formal strukturierten) Fließtext die relevanten Informationen zu entnehmen und damit Aussagen zu einzelnen Studien zu dokumentieren. Der Weg dahin und wie diese Arbeitsprozesse transparent und FAIR nachgewiesen werden können, wurde in den Ausführungen zum Forschungsdatenmanagement dargestellt. Ein Aspekt der FAIR-Prinzipien besteht dabei in der Forderung der Maschinenlesbarkeit von (Meta-)Daten für die Recherche wie auch zur Nachnutzung (Wilkinson et al., 2016). Dies wird von den Forschungsdateninfrastrukturen aufgegriffen und verschiedenlich durch den Einsatz von Datenmodellierungen nach Linked Open Data (LOD)-Prinzipien und der Entwicklung von Wissensgraphen angegangen, da diese semantisch eindeutige Verknüpfungen von Daten ermöglichen. Bestrebungen in diese Richtung bestehen beispielsweise aktuell im Kontext der Nationalen Forschungsdateninfrastruktur (NFDI) (Rossenova et al., 2023; Sack et al., 2023; Siegers et al., 2023).

Es existieren verschiedene Definitionen von Wissensgraphen (Ehrlinger & Wöß, 2016), eine eher allgemeinere versteht unter einem Wissensgraphen einen „graph of data intended to accumulate and convey knowledge of the real world, whose nodes represent entities of interest and whose edges represent potentially different relations between these entities“ (Hogan et al., 2021, S. 3). Zu den aktuell verbreitetsten Technologien in diesem Feld gehören das Resource Description Framework (RDF) und sog. Labeled Property Graphs (LPG). Die mit RDF realisierte Datenmodellierung bildet eine Verknüpfung von zwei Knoten durch eine Kante und ermöglicht Aussagen in der Form von Subjekt, Prädikat und Objekt (beispielhaft skizziert in Abbildung 1). Diese zeichnen sich dadurch aus, dass sie Instanzen einer oder mehrerer Klassen sein können und somit als konkrete Entitäten in einem stark formalisierten Modell, einer Ontologie, verortet werden. Durch die formale Eindeutigkeit eignen sich Ontologien besonders für den Datenaustausch zwischen Informationssystemen. Diese Eigenschaft wird in der RDF Modellierung noch weiter verstärkt, da mit ihr prinzipiell alle Klassen, Properties (Prädikate) und Instanzen über Web-Adressen eindeutig identifizierbar und adressierbar sind (ebd., passim).

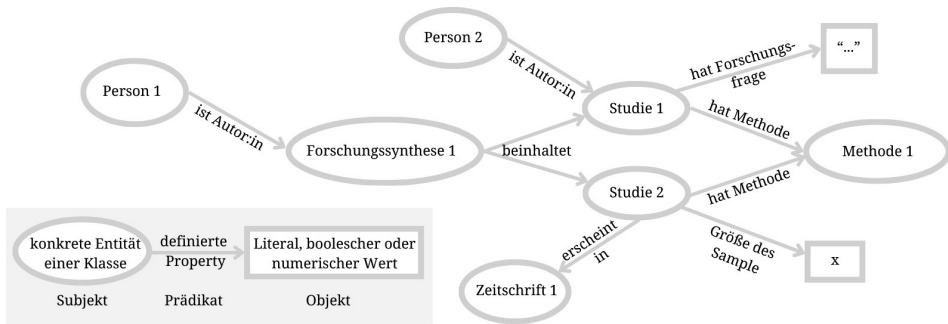


Abbildung 2: Beispielhafte schematische Darstellung eines RDF-Wissensgraphen

Nach der Einführung eines Wissensgraphen bei Google 2012 und von weiteren Unternehmen sowie der Entwicklung allgemeiner auf Wikipedia basierender LOD-Korpora (z. B. Wikidata, DBpedia⁶) entstanden zunehmend auch umfangreiche Wissensgraphen für die Wissenschaft (Manghi et al., 2022). Während die meisten wissenschaftlichen Wissensgraphen auf der Grundlage bibliografischer Metadaten weitere damit in Verbindung zu setzende Entitätstypen integrierten (z. B. Affiliationen, Projekte, Zitationen), verfolgt der Open Research Knowledge Graph (ORKG)⁷ das Ziel, inhaltliche Aussagen zu Forschungsbeiträgen über die bibliographische Beschreibung hinaus zu ermöglichen und miteinander in Beziehung zu bringen. Der ORKG wird seit 2018 an der Technischen Informationsbibliothek (TIB) in Hannover entwickelt. Die Inhalte des disziplinenübergreifenden Wissensgraphen werden mehrheitlich von Forschenden selbst generiert und auf der Open-Science-Plattform geteilt. Dafür steht ein Set an Ontologien bereit, das aber bei Bedarf auch punktuell erweitert werden kann. Ausgangspunkt ist eine wissenschaftliche Publikation, die in den vorgesehenen Formaten Comparison und Review mit anderen Publikationen, die die gleiche Forschungsfrage teilen, anhand verschiedener Merkmale (Properties) verglichen wird. Letztere sind beispielsweise verwendete Methoden, Stichprobengrößen, zeitliche und geografische Rahmen einer Untersuchung, lassen sich aber auch auf die Ergebnisse von Studien ausweiten und in Folge visualisieren. Die entstandenen Übersichtsarbeiten im ORKG können mit einer DOI versehen und damit als eigenständige Forschungsbeiträge referenziert werden. Sie versprechen dem Nutzenden einen Überblick über die Literatur zu ausgewählten Forschungsfragen zu vermitteln und damit die Literaturrecherche zu erleichtern (Auer et al., 2020).

Die Vorgehensweise der Übersichtsarbeiten im ORKG ähnelt in Teilen der Erstellung von Forschungssynthesen, wenngleich nicht mit einem vergleichbaren systematischen Anspruch. Zudem sind der Möglichkeit, strukturierte und eindeutige Aussagen in der gebotenen Knappheit zu formulieren, Grenzen gesetzt, wenn Forschungsergebnisse nicht unterkomplex dargestellt werden sollen. Daher bietet es sich

6 <https://www.wikidata.org>, <https://www.dbpedia.org/> (zuletzt aufgerufen am 26.04.2024).

7 <https://orkg.org/> (zuletzt aufgerufen am 26.04.2024).

an, die bereits vorliegenden tabellarischen Zusammenstellungen von existierenden Forschungssynthesen nachzunutzen und in den ORKG zu überführen, wofür auch eine semi-automatisierte Unterstützung angeboten wird (Oelen et al., 2020). In diesem Zusammenhang werden zur Exploration eine Auswahl der im Metavorhaben Digi-EBF erstellten Forschungssynthesen verwendet und im ORKG abgebildet.⁸ Der Umfang der einbezogenen Inhalte beschränkt sich dabei auf die in der Publikation der Forschungssynthesen enthaltenen tabellarischen Übersichten, die vor allem allgemeinere methodische Angaben, den geografischen Bezug von Untersuchungen und Dimensionen von behandelten Forschungsfragen enthalten. Die zugrundeliegende Identifizierung der relevanten Literatur und die Diskussion der aufgenommenen Studien bleiben hier außen vor. Das bedeutet zwar keine tiefe Erschließung der Critical Reviews aus dem Projekt Digi-EBF und schränkt die in Frage kommenden Properties ein, bereichert den ORKG jedoch zu den behandelten Forschungsthemen mit einer Übersicht an relevanter Literatur, die einen transparenten und systematischen Auswahlprozess durchlaufen hat.

Neben den allgemeinen Use Cases des ORKG (Brack et al., 2022) können drei Szenarien den perspektivischen Nutzen einer Aufnahme der Forschungssynthesen in den ORKG andeuten:

1. Seine Wirkungsmacht entfaltet ein Wissensgraph durch die zunehmende Aufnahme von Inhalten und die daraus entstehende Vernetzung von Entitäten. Der bisher stärker in den Inhaltsbereichen Technik und Naturwissenschaften befüllte ORKG wird damit um Studien und Themen der Bildungsforschung erweitert und kann sukzessive durch weitere Vergleichsarbeiten dieser Domäne Forschungsarbeiten miteinander vernetzen, kontextualisieren und maschinenlesbar bereitstellen.
2. Bei der Literatursuche ist die Eingrenzung von Treffermengen durch Facettierung der Ergebnisse nach bestimmten Merkmalen einer Publikation eine weit verbreitete und genutzte Funktionalität. Während üblicherweise in den bibliografischen Metadaten enthaltene formale Angaben zu den Publikationen die Grundlage für Facetten liefern, können in Wissensgraphen wie dem ORKG kontextabhängige und damit dynamische Facetten (entsprechend den verwendeten Properties) für die Recherche genutzt werden (Heidari et al., 2021). Durch die semantische Datenmodellierung erlauben Wissensgraphen allgemein die Formulierung sehr komplexer Suchfragen.
3. Über offene Schnittstellen und durch die verfügbaren Ontologien lassen sich die Inhalte des ORKG effizient in anderen, domänenspezifischen Wissensgraphen nachnutzen. Im Fachportal Pädagogik⁹ wird auf der Grundlage der vorliegenden Literaturnachweise ein Wissensgraph Bildung entwickelt. Dabei werden u. a. die

⁸ Als Beispiel: https://www.bildungsserver.de/link/forschungssynthesen_orkg. Für die Unterstützung bei der Eingabe der Angaben aus den Forschungssynthesen in den ORKG ist Loretta Schatz zu danken.

⁹ <https://www.fachportal-paedagogik.de/graph> (zuletzt aufgerufen am 26.04.2024).

Entitätstypen Personen und Zeitschriften miteinander vernetzt und mit zusätzlichen Informationen ergänzt. Zudem wird das Schlagwortvokabular semantisch anschlussfähig gemacht. Darüber hinaus sind zusätzliche Verlinkungen von Literaturnachweisen zu den dazugehörigen Forschungsdaten geplant. Somit ermöglicht eine Übernahme der Informationen aus dem ORKG, nicht vorhandene Literaturnachweise aufzunehmen oder bestehende mit zusätzlichen Informationen anzureichern. Doch auch wenn die inhaltliche Tiefe des ORKG nicht dem Anwendungsbereich vieler wissenschaftlicher Wissensgraphen entspricht, kann für die Literaturrecherche allein die Information, dass eine Studie Aufnahme in eine Forschungssynthese gefunden hat, wertvoll sein. Dies ergänzt einen Literaturnachweis um ein Relevanzmerkmal für die Forschungsfrage, die der Forschungssynthese zugrunde lag, und macht eine Verlinkung sowohl zur Forschungssynthese selbst als auch zu den anderen darin behandelten Studien möglich und erleichtert die Navigation in der Menge von Literaturnachweisen.

5. Forschungsdatenmanagement und Open Science im Metavorhaben Digi-EBF: Diskussion

In diesem Beitrag wurde ein Einblick in die Arbeit von Metavorhaben bei der Umsetzung von Open-Science-Prinzipien in einer Förderrichtlinie der empirischen Bildungsforschung gegeben. Der Fokus lag dabei vor allem auf der Rolle des Forschungsdatenmanagements, welches als Grundvoraussetzung für die Produktion von FAIR-Data und ihrer Veröffentlichung gilt. Zum einen kam dem Metavorhaben die Aufgabe zu, die Förderprojekte bei ihrem jeweiligen Forschungsdatenmanagement zu unterstützen. Dies wurde hauptsächlich durch individuelle Beratungen und der Veranstaltung von Workshops realisiert. Vor allem im Bereich der Forschungsethik und dem rechtssicheren Verfassen von Einwilligungserklärungen haben sich hier gerade zu Beginn der Projektlaufzeiten Bedarfe gezeigt. Als besondere Herausforderung kamen im Zuge der Covid-19-Pandemie Fragen zum Umgang mit alternativen Datenerhebungsmethoden hinzu. Zum Ende stellten sich hingegen insbesondere Fragen zur Veröffentlichung der Forschungsdaten, was zeigt, dass die Angebote von Forschungsdateninfrastrukturen wie dem VerbundFDB noch nicht in allen Bereichen der Bildungsforschung bekannt sind. Zum anderen war das Metavorhaben auch mit dem Forschungsdatenmanagement aus eigenen Vorhaben befasst, wie an den beiden Beispielen der Datenablage aus den Forschungssynthesen und der Überführung von Forschungsdaten in einen Wissensgraphen dargestellt werden konnte. Beide Beispiele zeigen auf, welchen Nutzen und welche Potenziale die Umsetzung von Open-Science-Praktiken für die Bildungsforschung haben können.

Für die Nachnutzung der Literaturrecherchen aus 20 Forschungssynthesen musste ein Workflow erarbeitet werden, der sowohl die Aufnahme von spezifischen Formaten wie Forschungssynthesen auf Seiten des Forschungsdaten zentrums ermöglichte

te als auch die Wiedergabe von zunehmend komplexeren Suchabfragen abdeckte. Da aufgrund der steigenden Anzahl an Forschungssynthesen in der Bildungsforschung zukünftig vermehrt mit ähnlichen Vorgängen der Datenarchivierung zu rechnen ist, können die Erfahrungen mit der Datenablage aus den Forschungssynthesen in Digi-EBF für solche Prozesse als eine (ausbaufähige) Grundlage dienen.¹⁰ Zudem veröffentlichte das DIPF zwischenzeitlich einen eigenen Open-Science-Kodex, der Transparenz und ein entsprechendes Forschungsdatenmanagement in allen Arbeitsschritten von Forschungsprozessen einfordert (DIPF, 2024). Für Forschungssynthesen bedeutet dies, dass das in Digi-EBF in Zusammenarbeit mit dem FDZ Bildung noch explorativ erprobte Verfahren der Datenablage zukünftig als Standard anzusehen und methodisch entsprechend ausgebaut werden kann.

Wissensgraphen bündeln Wissen in einer maschinenlesbaren Form. Der dazu nötige Grad an Formalisierbarkeit setzt der Umsetzung von wissenschaftlichen Texten Grenzen. Bei der Aufnahme ausgewählter Forschungssynthesen aus dem Digi-EBF Metavorhaben wurde daher bewusst nur auf bereits tabellarisch vorliegende Beschreibungen der Studien zurückgegriffen. Die durch das offene Datenformat ermöglichte Nachnutzung bietet dennoch Potenziale, die intellektuelle Arbeit bei der Erstellung der Forschungssynthesen künftig für die Recherche von wissenschaftlicher Literatur fruchtbar zu machen.

Mit den Vorgaben, die erhobenen Daten und Forschungsergebnisse aufzuarbeiten und zu veröffentlichen, fördert das BMBF den angestrebten Paradigmenwechsel zu einer offeneren Bildungsforschung, und mit der Einrichtung von Metavorhaben wird den Projekten der BMBF-Förderlinien im Rahmenprogramm empirische Bildungsforschung eine direkte Unterstützung bei der Bewältigung der Herausforderungen des Forschungsdatenmanagements zur Seite gestellt. Gleichwohl wird die Umsetzung – auch wenn der Nutzen immer häufiger anerkannt wird – von vielen Wissenschaftler:innen aus unterschiedlichen Gründen wie dem zusätzlichen Ressourcenaufwand oder einem eigenen Besitzanspruch an die Daten noch immer als Belastung empfunden (vgl. Putnigs, 2021).

In welchem Ausmaß die Unterstützung durch Metavorhaben wie Digi-EBF tatsächlich zu einer „offeneren Wissenschaft“ und zu einer besseren Praxisrelevanz beigetragen hat, kann an dieser Stelle nicht eindeutig beantwortet werden. Die Projektbefragungen zeigen allerdings, dass die Angebote des Metavorhabens durchaus angenommen wurden. Die Metaebene, auf der Digi-EBF agiert, hat darüber hinaus die Chance eröffnet, selbst in mehrere Richtungen „Testballons“ zu starten und zu erproben, inwiefern Daten aus den erstellten Forschungssynthesen geteilt und weitergenutzt werden konnten. Zur genauen Einschätzung der Wirkung dieser Maßnahmen und um das Angebot entsprechend für weitere Metavorhaben zu optimieren, bedürfte es indes ein tiefgehendes Forschungsmonitoring als in diesem Kontext gewährleistet werden konnte. In einem solchen Vorhaben ließen sich beispielsweise die

10 Eine entsprechende Zielsetzung verfolgt das Projekt ForSynData: <https://www.konsortswd.de/angebote/forschende/datenmanagementplaene/forsyndata/>

Zugriffszahlen auf die veröffentlichten Rechercheergebnisse der Forschungssynthesen bzw. die Zitation dieser auswerten, um weitere Erkenntnisse zu den Möglichkeiten der Datennachnutzung aus Forschungssynthesen zu erhalten. Auch über das Thema der Forschungssynthesen hinaus wäre es hilfreich, genauer zu beobachten, inwiefern eine Unterstützung im Forschungsdatenmanagement tatsächlich zu einer konsequenteren Umsetzung eines systematischen FDM und dadurch zu einer höheren Anzahl an Datenpublikationen bei den geförderten Projekten geführt hat. Hier wäre eine ausführliche Evaluation des Angebots denkbar.

Neben dem Ziel, mit diesem Kapitel eine Übersicht über die Potenziale des Forschungsdatenmanagements in Metavorhaben zu geben, war es uns ebenfalls ein Anliegen, einen Beitrag zum aktuellen Diskurs über eine offenere Wissenschaft in der Bildungsforschung zu leisten. Das Anliegen des BMBF, Projekte in Metavorhaben durch entsprechende Vorgaben top-down zu fördern, scheint aus unserer Perspektive zu fruchten. Wie weit die vorgestellten Bemühungen aber tatsächlich zu einem kulturellen Wandel in der Forschungspraxis beitragen, bleibt genauer zu ergründen.

Literatur

- Auer, S., Oelen, A., Haris, M., Stocker, M., D'Souza, J., Farfar, K. E., Vogt, L., Prinz, M., Wiens, V. & Jaradeh, M. Y. (2020). Improving access to scientific literature with knowledge graphs. *Bibliothek Forschung und Praxis*, 44(3), 516–529. <https://doi.org/10.1515/bfp-2020-2042>
- Bayer, S., Breuer, J., Lösch, T. & Goebel, J. W. (2021). *Nutzung von Social-Media-Daten in der Bildungsforschung*. *forschungsdaten bildung informiert*, 9, Version 1.2. <https://doi.org/10.25656/01:22121>
- Beelmann, A. (2014). Möglichkeiten und Grenzen systematischer Evidenzkumulation durch Forschungssynthesen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 17(4), 55–78. <https://doi.org/10.1007/s11618-014-0509-2>
- Benner, D. (1980). Das Theorie-Praxis-Problem in der Erziehungswissenschaft und die Frage nach Prinzipien pädagogischen Denkens und Handelns. *Zeitschrift für Pädagogik*, 26(4), 485–497. <https://doi.org/10.25656/01:14104>
- Brack, A., Hoppe, A., Stocker, M., Auer, S. & Ewerth, R. (2022). Analysing the requirements for an open research knowledge graph: Use cases, quality requirements, and construction strategies. *International Journal on Digital Libraries*, 23(1), 33–55. <https://doi.org/10.1007/s00799-021-00306-x>
- BMBF – Bundesministerium für Bildung und Forschung. (2017). *Richtlinie zur Förderung von Forschung zu „Digitalisierung im Bildungsbereich – Grundsatzfragen und Gelingensbedingungen“*. <https://www.bmbf.de/foerderungen/bekanntmachung-1420.html>
- BMBF – Bundesministerium für Bildung und Forschung. (2021). *Digitalisierung im Bildungsbereich*. <https://www.empirische-bildungsforschung-bmbf.de/de/Digitalisierung-im-Bildungsbereich-1804.html>
- BMBF – Bundesministerium für Bildung und Forschung. (2023). *Open Access in Deutschland – Gemeinsame Leitlinien von Bund und Ländern*. https://www.bmbf.de/Shared-Docs/Publikationen/de/bmbf/1/772960_Open_Access_in_Deutschland.pdf

- Buck, D., Croisier, J., Eder, C., Hoffstätter, U., Jansen, M., Meyermann, A. & Siegers, P. (2022). Handreichung: Forschungsdatenzentren gründen. RatSWD Working Paper 280/2022. Rat für Sozial- und Wirtschaftsdaten (RatSWD). <https://doi.org/10.17620/02671.70>
- Buntins, K., Bedenlier, S., Marin, V., Händel, M. & Bond, M. (2023). Methodological approaches to evidence synthesis in educational technology: A tertiary systematic mapping review. *Medienpädagogik. Zeitschrift für Theorie und Praxis in der Medienbildung*, 54, 167–191. <https://doi.org/10.21240/mpaed/54/2023.12.20.X>
- Cochrane Information Specialist Support Team. (2023). *Cochrane information specialists' handbook*. Cochrane. <https://training.cochrane.org/resource/cochrane-information-specialists-handbook>
- Collaboration for Environmental Evidence. (2023). *Guidelines and standards for evidence synthesis in environmental management* (Version 5.1). <https://environmentalevidence.org/information-for-authors/>
- DFG – Deutsche Forschungsgemeinschaft. (2018). *Förderung von Informationsinfrastrukturen für die Wissenschaft*. Ein Positionspapier der Deutschen Forschungsgemeinschaft. <https://www.dfg.de/resource/blob/173200/66e1e4296848c5f700b83898f7f48995/positionspapier-informationsinfrastrukturen-data.pdf>
- Die Bundesregierung. (2021). Datenstrategie der Bundesregierung. Eine Innovationsstrategie für gesellschaftlichen Fortschritt und nachhaltiges Wachstum, hg. v. Bundeskanzleramt. <https://www.bundesregierung.de/resource/blob/992814/1845634/1a4f7e-a800bb838562e16fdfe4ffb354/datenstrategie-der-bundesregierung-download-bpa-data.pdf?download=1>
- Diekmann, D. (2022). Praxisbericht: Mit Metavorhaben die Kommunikation und Kooperation zwischen Bildungsforschung und Bildungspraxis unterstützen. *Bildungsforschung*, 2, 1–16. <https://doi.org/10.25656/01:25497>
- DIPF – DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation (2024). *DIPF Open Science Codex*. <https://doi.org/10.25657/02:29165>
- Ehrlinger, L. & Wöß, W. (2016). Towards a definition of knowledge graphs. *SEMANTICS*, 48, 1–4. <https://ceur-ws.org/Vol-1695/paper4.pdf>
- Fickermann, D. & Edelstein, B. (Hrsg.) (2021). *Schule während der Corona-Pandemie. Neue Ergebnisse und Überblick über ein dynamisches Forschungsfeld*. Waxmann. <https://doi.org/10.31244/9783830993315>
- Fournier, J. (2021). Komplexität und Vielfalt gestalten. In M. Putnings, H. Neuroth & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement* (S. 1–4). De Gruyter Saur. <https://doi.org/10.1515/9783110657807>
- Heidari, G., Ramadan, A., Stocker, M. & Auer, S. (2021). Leveraging a federation of knowledge graphs to improve faceted search in digital libraries. In G. Berget, M. M. Hall, D. Brenn & S. Kumpulainen (Hrsg.), *Linking Theory and Practice of Digital Libraries* (S. 141–152). Springer. https://doi.org/10.1007/978-3-030-86324-1_18
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. & Welch, V. (Hrsg.). (2023). *Cochrane handbook for systematic reviews of interventions* (Version 6.4). Cochrane. <https://training.cochrane.org/handbook/current>
- Hogan, A., Blomqvist, E., Cochez, M., D'Amato, C., De Melo, G., Gutiérrez, C., Gayo, J. E. L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A. N., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J. F., Staab, S. & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), 1–37. <https://doi.org/10.1145/3447772>
- Institute of Medicine. (2011). *Finding what works in health care: Standards for systematic reviews*. National Academies Press. <https://doi.org/10.17226/13059>

- Keller, C. & Rittberger, M. (2020). *Metavorhaben „Digitalisierung im Bildungsbereich“ – Literaturrecherche für Dossier Digi EBF 1 (Digi-EBF) [Datenkollektion: Version 1.0]. Datenerhebung 2018–2019.* Forschungsdatenzentrum Bildung am DIPF. <https://doi.org/10.7477/414:1:0>
- Keller, C. & Rittberger, M. (2021). *Metavorhaben „Digitalisierung im Bildungsbereich“ – Literaturrecherche für Dossier Digi EBF 2 (Digi-EBF) [Datenkollektion: Version 1.0]. Datenerhebung 2020.* Forschungsdatenzentrum Bildung am DIPF. <http://doi.org/10.7477/414:2:0>
- Keller, C. & Rittberger, M. (2023a). *Metavorhaben „Digitalisierung im Bildungsbereich“ – Literaturrecherche für Dossier Digi EBF 3 (Digi-EBF) [Datenkollektion: Version 1.0]. Datenerhebung 2021–2022.* Forschungsdatenzentrum Bildung am DIPF. <https://doi.org/10.7477/414:3:0>
- Keller, C. & Rittberger, M. (2023b). *Metavorhaben „Digitalisierung im Bildungsbereich“ – Literaturrecherche für Dossier Digi EBF 4 (Digi-EBF) [Datenkollektion: Version 1.0]. Datenerhebung 2022–2023.* Forschungsdatenzentrum Bildung am DIPF. <https://doi.org/10.7477/414:4:0>
- Krammer, G. & Svecnik, E. (2020). Open Science als Beitrag zur Qualität in der Bildungsforschung. *Zeitschrift für Bildungsforschung*, 10(3), 263–278. <https://doi.org/10.1007/s35834-020-00286-z>
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. K., Hammerstrøm, K. T. & Sathe, N. A. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13(1), 1–73. <https://doi.org/10.4073/cmg.2016.1>
- LearningLab Essen. (2020, 2. Dezember). *Bildungsdialog.digital: Inside Bildungsforschung – Ergebnisse (Welle 2) Befragung der Förderprojekte* [Video]. YouTube. <https://www.youtube.com/watch?v=tZJh9XXzQVI&>
- Manghi, P., Mannocci, A., Osborne, F., Sacharidis, D., Salatino, A. A. & Vergoulis, T. (2022). New trends in scientific knowledge graphs and research impact assessment. *Quantitative Science Studies*, 2(4), 1296–1300. https://doi.org/10.1162/qss_e_00160
- Meyermann, A. (2012). Datenmanagement – Eine zentrale Voraussetzung für den Erfolg der Data Sharing-Idee in den Sozialwissenschaften. *DSZ-BO Working Paper 2*. www.uni-bielefeld.de/dsz-bo/pdf/2012-07-23_WP2_Datenmanagement_v2.pdf
- Meyermann, A. & Porzelt, M. (2019). Datenschutzrechtliche Anforderungen in der empirischen Bildungsforschung. Eine Handreichung. *forschungsdaten bildung informiert*, 6, Version 2.
- National Academies of Sciences, Engineering, and Medicine. (2023). *Toward a 21st Century National Data Infrastructure: Mobilizing Information for the Common Good*. The National Academies Press. <https://doi.org/10.17226/26688>
- Nguyen, P., McKenzie, J. E., Hamilton, D. G., Moher, D., Tugwell, P., Fidler, F., Haddaway, N. R., Higgins, J. P. T., Kanukula, R., Karunanathan, S., Maxwell, L., McDonald, S., Nakagawa, S., Nunan, D., Welch, V. & Page, M. J. (2023). Systematic reviewers' perspectives on replication of systematic reviews: A survey. *Cochrane Evidence Synthesis and Methods*, 1(2). <https://doi.org/10.1002/cesm.12009>
- Nordström, T., Kalmendal, A. & Batinovic, L. (2023). Risk of bias and open science practices in systematic reviews of educational effectiveness: A meta-review. *Review of Education*, 11(3). <https://doi.org/10.1002/rev3.3443>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture: Author guidelines for journals

- could help to promote transparency, openness, and reproducibility. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- Oelen, A., Jaradeh, M. Y., Stocker, M. & Auer, S. (2020). Generate FAIR literature surveys with scholarly knowledge graphs. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020 (JCDL '20)*, 97–106. <https://doi.org/10.1145/3383583.3398520>
- Putnings, M. (2021). Förderpolitische Maßgaben. In M. Putnings, H. Neuroth & J. Neumann (Hrsg.), *Praxishandbuch Forschungsdatenmanagement* (S. 53–88). De Gruyter Saur. <https://doi.org/10.1515/9783110657807>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. & PRISMA-Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10(1). <https://doi.org/10.1186/s13643-020-01542-z>
- Rossenova, L., Schubotz, M. & Shigapov, R. (2023). The case for a common, reusable knowledge graph infrastructure for NFDI. *Proceedings of the Conference on Research Data Infrastructure*, 1. <https://doi.org/10.52825/cordi.v1i.266>
- Sack, H., Schrade, T., Bruns, O., Posthumus, E., Tietz, T., Norouzi, E., Waitelonis, J., Fliegl, H., Söhn, L., Tolksdorf, J., Steller, J. J., Guzmán, A. A., Fathalla, S., Ihsan, A. Z., Hofmann, V., Sandfeld, S., Fritzen, F., Laadhar, A., Schimmler, S. & Mutschke, P. (2023). Knowledge graph based RDM solutions. *Proceedings of the Conference on Research Data Infrastructure*, 1. <https://doi.org/10.52825/cordi.v1i.371>
- Scherp, G. (2022). Open Science – Prinzipien, Entstehung, Herausforderungen. In H. Ertl & B. Rödel (Hrsg.), *Offene Zusammenhänge: Open Access in der Berufsbildungsforschung* (S. 346–355). Bundesinstitut für Berufsbildung (BIBB).
- Siegers, P., May, A. C., Saalbach, C., Nebelin, J., Kern, D., Daniel, A., Zapilko, B., Momeni, F., Wenzig, K. & Goebel, J. (2023). Linked open research data for social science. *Proceedings of the Conference on Research Data Infrastructure*, 1. <https://doi.org/10.52825/cordi.v1i.300>
- Trixa, J. & Ebel, T. (2015). *Einführung in das Forschungsdatenmanagement in der empirischen Bildungsforschung. Zusammenfassung des Workshops im Rahmen des CESSDA-Trainingsprogramms* (Version 1.1.). <https://doi.org/10.25656/01:21969>
- UNESCO. (2021). *UNESCO Recommendation on Open Science*. UNESCO. <https://doi.org/10.54677/MNMH8546>
- Van der Zee, T. & Reich, J. (2018). Open education science. *AERA Open*, 4(3). <https://doi.org/10.1177/2332858418787466>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Da Silva Santos, L. O. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T. W., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1). <https://doi.org/10.1038/sdata.2016.18>