

Meyermann, Alexia et al.

Working Paper

Der Verbund Forschungsdaten Bildung – Eine Forschungsdateninfrastruktur für die empirische Bildungsforschung

RatSWD Working Paper, No. 266

Provided in Cooperation with:

German Data Forum (RatSWD)

Suggested Citation: Meyermann, Alexia et al. (2017) : Der Verbund Forschungsdaten Bildung – Eine Forschungsdateninfrastruktur für die empirische Bildungsforschung, RatSWD Working Paper, No. 266, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin, <https://doi.org/10.17620/02671.27>

This Version is available at:

<https://hdl.handle.net/10419/172797>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

266

Der Verbund Forschungsdaten Bildung – Eine Forschungsdaten- infrastruktur für die empirische Bildungsforschung

Alexia Meyermann, Doris Bambey,
Malte Jansen, Reiner Mauer, Thomas Ebel,
Marcus Eisentraut, Karoline Harzenetter,
Poldi Kuhl, Claudia Neuendorf, Lisa Pegelow,
Maike Porzelt, Marc Rittberger,
Thomas Schwager, Petra Stanat, Jessica Trixa

Oktober 2017

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Working Paper Series des Rates für Sozial- und Wirtschaftsdaten (RatSWD)

Die *RatSWD Working Papers* Reihe startete Ende 2007. Seit 2009 werden in dieser Publikationsreihe nur noch konzeptionelle und historische Arbeiten, die sich mit der Gestaltung der statistischen Infrastruktur und der Forschungsinfrastruktur in den Sozial-, Verhaltens- und Wirtschaftswissenschaften beschäftigen, publiziert. Dies sind insbesondere Papiere zur Gestaltung der Amtlichen Statistik, der Ressortforschung und der akademisch getragenen Forschungsinfrastruktur sowie Beiträge, die Arbeit des RatSWD selbst betreffend. Auch Papiere, die sich auf die oben genannten Bereiche außerhalb Deutschlands und auf supranationale Aspekte beziehen, sind besonders willkommen.

RatSWD Working Papers sind nicht-exklusiv, d. h. einer Veröffentlichung an anderen Orten steht nichts im Wege. Alle Arbeiten können und sollen auch in fachlich, institutionell und örtlich spezialisierten Reihen erscheinen. Die *RatSWD Working Papers* können nicht über den Buchhandel, sondern nur online über den RatSWD bezogen werden.

Um nicht deutsch sprechenden Nutzer/innen die Arbeit mit der Reihe zu erleichtern, sind auf den englischen Internetseiten der *RatSWD Working Papers* nur die englischsprachigen Papers zu finden, auf den deutschen Seiten werden alle Nummern der Reihe chronologisch geordnet aufgelistet.

Einige ursprünglich in der *RatSWD Working Papers* Reihe erschienenen empirischen Forschungsarbeiten sind ab 2009 in der RatSWD Research Notes Reihe zu finden.

Die Inhalte der *RatSWD Working Papers* stellen ausdrücklich die Meinung der jeweiligen Autor/innen dar und nicht die des RatSWD. Das Bundesministerium für Bildung und Forschung hat die Publikationen nicht beeinflusst.

Die RatSWD Working Paper Series wird herausgegeben von:

seit 2014 Regina T. Riphahn (Vorsitzende des RatSWD)

2009–2014 Gert G. Wagner

2007–2008 Heike Solga

Der Verbund Forschungsdaten Bildung – Eine Forschungsdateninfrastruktur für die empirische Bildungsforschung

Alexia Meyermann^a, Doris Bambey^a, Malte Jansen^b, Reiner Mauer^c,
Thomas Ebel^c, Marcus Eisentraut^c, Karoline Harzenetter^c, Poldi Kuhl^d,
Claudia Neuendorf^b, Lisa Pegelow^b, Maike Porzelt^a, Marc Rittberger^a,
Thomas Schwager^a, Petra Stanat^b, Jessica Trixa^c

-
- ^a DIPF – Bildungsforschung und Bildungsinformation
^b FDZ am IQB (Institut zur Qualitätsentwicklung im Bildungswesen)
^c GESIS-Datenarchiv (GESIS – Leibniz-Institut für Sozialwissenschaften e. V.)
^d Leuphana Universität Lüneburg

doi: [10.17620/02671.27](https://doi.org/10.17620/02671.27)

1.	Der Verbund Forschungsdaten Bildung	2
2.	Bildungsforschung: Ein interdisziplinäres Forschungsgebiet mit hoher Datenvielfalt.....	3
3.	Eine Forschungsdateninfrastruktur für die Bildungsforschung	4
4.	Die Services für die Forschung	6
4.1.	Forschungsdaten finden	7
4.2.	Forschungsdaten teilen: Angebote für Datengeber/innen	8
4.3.	Forschungsdaten nutzen: Angebote für Datennutzer/innen	8
4.4.	Forschungsdatenmanagement: Angebote für Datenproduzent/innen	9
4.5.	Erfahrungen – Hürden der Datenarchivierung.....	10
5.	Zukünftige Handlungsfelder.....	12
6.	Literatur.....	14

1. Der Verbund Forschungsdaten Bildung

Für den Erkenntnisgewinn, die Qualitätssicherung und letztlich den Erfolg wissenschaftlicher Forschung ist es zentral, einen nachhaltigen und grundsätzlich offenen Zugang zu Forschungsdaten und -ergebnissen für Wissenschaftler/innen zu gewährleisten. Die verlässliche Zugänglichkeit von Forschungsdaten sorgt einerseits für Transparenz und Nachvollziehbarkeit und fördert andererseits den kumulativen Erkenntnisgewinn. Wissenschaftspolitische Akteure und Zuwendungsgeber fordern und fördern daher den Aufbau entsprechender Infrastrukturen zur Archivierung und Nachnutzung qualitäts gesicherter Forschungsdaten (vgl. z. B. Allianz der Wissenschaftsorganisationen 2010, Kommission Zukunft der Informationsinfrastrukturen 2011, Wissenschaftsrat 2012, OECD 2013, Hochschulrektorenkonferenz 2014 und 2015, Rat für Informationsinfrastrukturen 2016). Seit 2008 fordert auch das Bundesministerium für Bildung und Forschung (BMBF) durch entsprechende Förderauflagen im *Rahmenprogramm zur Förderung der empirischen Bildungsforschung* (EBF-Rahmenprogramm) gezielt die Nachhaltigkeit von Forschungsdaten der Bildungsforschung ein. Es ist Bestandteil der Förderbedingungen, dass Projektnehmer/innen die in den Projekten generierten Daten und Instrumente archivieren und der Wissenschaft zur Nachnutzung verfügbar machen.¹ Vor diesem Hintergrund widmet sich der Verbund Forschungsdaten Bildung (VerbundFDB) seit Oktober 2013 im Auftrag des BMBF dem Aufbau und der Gestaltung einer Forschungsdateninfrastruktur für die empirische Bildungsforschung. Die Aufgabe besteht darin, eine vernetzte Forschungsdateninfrastruktur für die empirische Bildungsforschung in ihrer gesamten Breite aufzubauen und im Austausch mit der Fachcommunity bedarfsgerecht zu gestalten.

Der VerbundFDB ist ein Kooperationsprojekt der Forschungsdatenzentren der drei Institute DIPF, GESIS und IQB, das sich in der Pilotphase (Förderzeitraum 2013-2016) darauf konzentrierte, die Daten aus dem EBF-Rahmenprogramm zu sichern und im Rahmen der bestehenden Möglichkeiten zur Nachnutzung verfügbar zu machen. Dabei bestand die Herausforderung zunächst darin, die große Vielzahl an Projekten (über 300) und Daten mit einer entsprechend großen methodischen Vielfalt und Heterogenität von Datentypen zu betreuen. Da dieses Vorhaben nur kooperativ zu bewältigen war, hat der VerbundFDB ein spezifisches Modell einer föderierten Infrastruktur (vgl. bspw. Rat für Informationsinfrastrukturen 2016: 43f. oder rdn o. J.) realisiert. Das Kooperationsmodell der *verteilten Archivierung* verbindet einen zentralen Datennachweis im Onlineportal www.forschungsdaten-bildung.de mit der dezentralen Archivierung und Bereitstellung der Daten bei den jeweils zuständigen Datenzentren. Derzeit befindet sich der VerbundFDB in seiner zweiten Förderphase (2016-2019). In dieser Phase wird die bisher aufgebaute Infrastruktur ausgebaut, um das zentrale Ziel zu erreichen, die bestehenden Angebote für die gesamte empirische Bildungsforschung zu öffnen.

Im Folgenden werden einführend die Anforderungen an eine Forschungsdateninfrastruktur für die empirische Bildungsforschung beschrieben, die sich aus den besonderen Charakteristika dieses Forschungsfeldes ergeben. Anschließend werden die drei zentralen Angebote des VerbundFDB vorgestellt. Diese sind

- (1) die Dokumentation von Studien der empirischen Bildungsforschung,
- (2) die Archivierung und Bereitstellung von Forschungsdaten sowie
- (3) die Information, Beratung und Schulung von Forschenden zum Forschungsdatenmanagement.

Rückblickend werden jeweils kurz die Erfahrungen aus der ersten Förderphase des VerbundFDB beschrieben, und es wird ein Ausblick auf die zukünftigen Aufgabenstellungen gegeben. Im abschließenden Resümee werden zukünftige Aufgaben und Herausforderungen diskutiert.

¹ Vgl. bspw. www.bmbf.de/foerderungen/bekanntmachung.php?B=339 (Zugegriffen: 04. Mai 2016); im Jahr 2012 wurden die Auflagen bezüglich Forschungsdaten noch darüber hinausgehend erweitert, vgl. bspw. www.bmbf.de/foerderungen/bekanntmachung.php?B=774 (Zugegriffen: 04. Mai 2016).

2. Bildungsforschung: Ein interdisziplinäres Forschungsgebiet mit hoher Datenvielfalt

Um exzellente, produktive und wettbewerbsfähige Forschung im Bereich Bildung zu ermöglichen, werden leistungsfähige und auf die speziellen Bedürfnisse der Bildungsforschung zugeschnittene Informationsinfrastrukturen benötigt. Im Folgenden gehen wir auf die besondere Lage und die speziellen Rahmenbedingungen der Datennachnutzung (d. h. des Data Sharing) in der Bildungsforschung ein.

(1) Aus *forschungspolitischer* Sicht ergeben sich aufgrund der Bedeutung des Forschungsgebiets für Bildungspolitik und -praxis Anforderungen an die Transparenz und die Nachvollziehbarkeit der Forschung. Dies zeigen auch die Diskussionen über die Themen „Bildungsmonitoring“ und „evidenzbasierte Bildungspolitik“ (etwa auf der Basis von Systematic Reviews). Die empirische Bildungsforschung trägt wesentlich dazu bei, bildungspolitische Entscheidungen wissenschaftlich zu fundieren. Für die Bildungsforschung kommt der Verfügbarkeit von Forschungsdaten zudem aus *forschungspraktischer* Sicht eine strategische Bedeutung zu. Diese ergibt sich unter anderem aus den besonders schwierigen sowie kosten- und zeitintensiven Feldzugängen (vgl. z. B. Stanat 2014). Schulen gelten zunehmend als überforscht und gerade auch videogestützte Unterrichtsbeobachtungen gelten als aufwändig in der Produktion, und der Zugang zu minderjährigen Untersuchungspersonen in institutionellen Settings ist mittlerweile stark reglementiert.

(2) Kennzeichnend für die empirische Bildungsforschung ist deren *multi- bzw. interdisziplinäres Profil*, das Forschungsfragen und methodische Ansätze aus dem Bereich der Bildungswissenschaften im engeren Sinne (Erziehungswissenschaft und Pädagogik) sowie u. a. aus den Fächern Psychologie, Soziologie, Ökonomie, den Neuro-, Geschichts- und Sprachwissenschaften umfasst. Hinzu kommen die diversen Fachdidaktiken (Mathematik, Sprachen, Musik etc.), die jeweils eigene disziplinäre Herangehensweisen einbringen. Dies geht mit einer methodischen Vielfalt einher, aus der eine große Bandbreite an Datentypen resultiert. Die generierten Daten liegen in vielfältigen Formaten vor, wie Tabellen (Umfragedaten und Daten aus Leistungstests), Bildern (Netzwerkkarten), Videodaten (Unterrichtsbeobachtungen), auditiven Aufnahmen (Experteninterviews) bis hin zu Verhaltensdaten (Reaktionszeitmessungen) oder neuropsychologischen Messdaten (EEG, MRT). In der empirischen Bildungsforschung ist ein Zuwachs im Bereich der Verwendung von audiovisuellen Erhebungsverfahren zu verzeichnen² und damit verbundenen komplexen methodischen Settings. Hinzu kommt eine allgemein für die Sozial- und Wirtschaftswissenschaften zu beobachtende Zunahme methodischer Komplexität, die mit dem vermehrten Einsatz quantitativ-qualitativer Designs (Mixed-Methods-Studien), Längsschnitt- und Mehrebenendesigns sowie experimentellen und quasi-experimentellen Designs einhergeht.

(3) Weiterhin bewegt sich die empirische Bildungsforschung in einem *besonderen rechtlichen und regulatorischen Umfeld* (vgl. bspw. Scheller 2017, Meyermann/Porzelt 2017). Häufig entstehen im Kontext von Erhebungen in diesem Feld datenschutzrechtlich und ethisch sensible Daten, z. B. Daten aus Leistungs- und Kompetenzmessungen besonders schützenswerter Populationen wie minderjährige Schüler/innen. Bestimmte Arten von Bildungsforschungsdaten sind teilweise – wie im Fall von verknüpften Schulen-Schüler-Lehrer-Daten – nur mit hohem Aufwand oder – wie im Fall von Audio- oder Videomaterial – gar nicht anonymisierbar, wenn die Möglichkeiten wissenschaftlicher Analyse nicht extrem eingeschränkt werden sollen. Häufig werden in der Bildungsforschung zudem kleinräumige Erhebungen durchgeführt, die sich auf einzelne Bundesländer, Regionen oder Städte beziehen.

² Vor allem unter dem Vorzeichen eines „visual turn“ sozialwissenschaftlicher Forschung haben videobasierte Verfahren eine starke Dynamik erfahren (vgl. Schnettler/Pöttsch 2007). Anhand einer thematischen Analyse der *FIS Bildung Literaturdatenbank* (http://www.fachportal-paedagogik.de/fis_bildung/fis_form.html) zum Forschungsbereich Unterrichtsvideografie lässt sich ebenfalls eine deutliche Zunahme an Forschungsaktivitäten im Feld feststellen (vgl. Reinhold et al. 2013).

Auch hier sind die Anonymisierungsmöglichkeiten begrenzt, wenn die Erhebungsorte und Regionalinformationen bereits im Rahmen der Projektbeschreibung veröffentlicht sind. Für Erhebungen an Schulen ist es in der Regel erforderlich, Genehmigungen von Schulaufsichtsbehörden einzuholen (vgl. Forschungsdaten Bildung, o. J.). Dies stellt administrative Zusatzaufwände für die Forschenden dar und erschwert in bestimmten Fällen aufgrund der damit einhergehenden Restriktionen eine Nachnutzung der Daten über den engeren Projektkontext hinaus (bspw. für regionale Analysen, für Ländervergleiche oder für Verknüpfungen mit anderen Datenbeständen). Beim Einsatz von Tests und insbesondere von psychologischen Messinstrumenten spielen häufig urheberrechtliche Aspekte eine Rolle, die bei Archivierung und Weitergabe zu berücksichtigen sind. Unter Umständen sind zusätzliche Zustimmungen für weitergehende Nutzungen von den Rechteinhabern einzuholen, oder die Instrumente sind nur eingeschränkt publizierbar.

Aus den genannten Charakteristika empirischer Bildungsforschung ergeben sich spezifische Anforderungen an eine Forschungsdateninfrastruktur, die den Auftrag hat, Bildungsforschungsdaten zu erhalten und verfügbar zu machen und Forschende bei der Datenproduktion und -nutzung zu unterstützen. Erforderlich sind datentypspezifische und fachwissenschaftliche Kompetenzen in den Datenzentren selbst sowie ein intensiver Austausch mit den jeweiligen Scientific Communities.

3. Eine Forschungsdateninfrastruktur für die Bildungsforschung

Auch aufgrund der Multidisziplinarität und Heterogenität der empirischen Bildungsforschung sind in Deutschland inhaltlich und methodisch spezialisierte Datenzentren verteilt an unterschiedlichen Standorten entstanden. Die stark diversifizierte Infrastrukturlandschaft und der hohe Spezialisierungsgrad der Forschungsdatenzentren haben den Vorteil, dass die einzelnen Datenzentren eine hohe fachliche Expertise und eine starke Einbindung in die jeweiligen Communities aufweisen. Dies ermöglicht eine adäquate Aufbereitung und Dokumentation der Daten auf hohem Niveau und hilft, communityspezifische Bedarfe gezielt zu berücksichtigen, wie beispielsweise besondere rechtliche Rahmenbedingungen einzelner Communities in der Bildungsforschung. Auf der anderen Seite geht die Spezialisierung und Diversifizierung der Forschungsdatenzentren mit einer gewissen Unübersichtlichkeit oder Komplexität für die Nutzer/innen einher. Die Datenbestände mit Relevanz für die Bildungsforschung finden sich an verschiedenen Orten. So weisen aus der Gruppe der vom RatSWD akkreditierten Forschungsdatenzentren (FDZ) die Zentren am LifBi, IQB und DIPF einen dezidierten Bezug zur Bildungsforschung auf. Ihnen gegenüber steht eine Vielzahl von Datenzentren mit zwar vorhandenen, jedoch unterschiedlich stark ausgeprägten Bezügen zur Bildungsforschung (z. B. GESIS, Psychdata von ZPID). Generell sind die beim RatSWD beteiligten FDZ, die über die eigene Einrichtung hinaus, überregional Dienstleistungen anbieten, in der Minderheit (FDZ-BO, FDZ Bildung am DIPF, GESIS oder Psychdata), und es überwiegen FDZ mit Fokus auf institutsintern generierte Daten und/oder Daten aus großen Projektverbänden und Umfragestudien (Large-Scale-Studies) (so die FDZ LifBI, SHARE, DZA, SOEP oder pairfam).

In den letzten Jahren gewinnen zudem fachübergreifende Angebote universitärer Rechenzentren und Bibliotheken an Bedeutung.³ Universitäten sind für viele Forscher/innen aufgrund der Standortnähe die erste Anlaufstelle und für Forschungsdaten, die aus rechtlichen Gründen nicht an ein fachspezifisches FDZ weitergegeben werden können, die einzige Sicherungsoption. Gleichzeitig werden Daten hier nicht fachspezifisch kuratiert, sondern es erfolgt i. d. R. eine basale Erschließung und Sicherung.

³ Diese sind derzeit dabei, fächerübergreifende Repositorien aufzubauen, Richtlinien zum Umgang mit Forschungsdaten für ihre Einrichtungen zu entwickeln und das Thema Forschungsdatenmanagement in die universitäre Lehre aufzunehmen (vgl. Hochschulrektorenkonferenz 2015). Beispielhaft seien hier die Universitäten Heidelberg oder Bielefeld genannt, die bereits eigene Forschungsdatenrepositorien aufgebaut haben und umfangreiche Informationen zum Forschungsdatenmanagement anbieten.

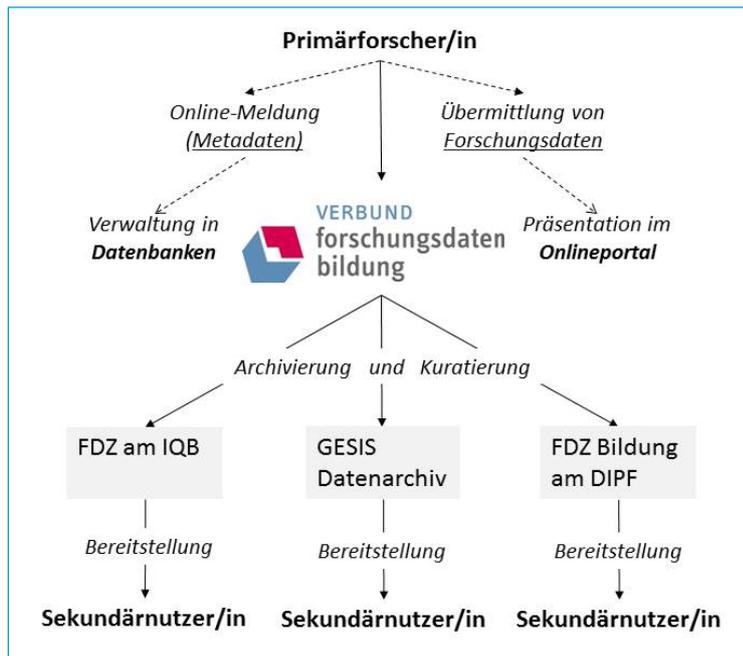
Dies bedeutet, dass eine fachspezifische Aufbereitung und Pflege von Daten ebenso wenig angeboten wird wie Beratungsdienste für Sekundärnutzer/innen der Daten. Die fachspezifische Pflege von Daten, wie sie FDZ leisten, ist für bestimmte Datentypen jedoch erforderlich, um die Nachnutzbarkeit dieser Daten zu gewährleisten oder, darüber hinausgehend, um ein hohes Nachnutzungspotential zu erzeugen und zu erhalten. Dies gilt beispielsweise für Daten, die auf komplexen methodischen Designs basieren oder besonderen rechtlichen Restriktionen unterliegen.

Für Nutzer/innen folgt aus der voranstehend geschilderten Situation, dass sie derzeit mehrfach und archivspezifisch nach Archivierungs- und Sekundärnutzungsangeboten recherchieren müssen. Es fehlt ein zentraler Zugriff auf und Überblick über Forschungsdaten der Bildungsforschung unabhängig von ihrem jeweiligen Archivierungsort. Zudem decken die vorhandenen Datenzentren derzeit mit ihren Schwerpunkten auf Large Scale Surveys und institutsinternen Daten vor allem ausgewählte Teilbereiche der Bildungsforschung ab. Andere Bereiche, wie Daten der universitären Einzelforschung oder der qualitativen Bildungsforschung, sind unterversorgt. Insbesondere die Vielfalt qualitativer Daten, die, neben Leitfadeninterviews, narrative Interviews, Ethnographien, Diskursanalysen u. v. m. bereithält, ist in der gegenwärtigen Forschungsdateninfrastruktur nicht angemessen vertreten. Auch fehlen Services zum Datenmanagement in den Projekten (begleitend im Forschungsprozess) und das Datenmanagement unterstützende Tools⁴.

Die Herausforderung für die Forschungsdateninfrastruktur besteht darin, ein gemeinsames und harmonisiertes Angebot zu schaffen, das es Bildungsforscher/innen aller Disziplinen ermöglicht, unter einer Adresse standortunabhängig auf qualitätsgesicherte Forschungsdaten zuzugreifen. Eine naheliegende Lösung hierfür bietet das Modell der *verteilten Archivierung mit einem zentralen Recherchezugang*, wie es im Rahmen des VerbundFDB pilothaft durch die Datenzentren von DIPF, GESIS und IQB in den letzten Jahren umgesetzt wurde, und das derzeit weiter ausgebaut wird. Abbildung 1 zeigt die derzeitige Architektur der verteilten Archivierung im VerbundFDB im Überblick. Das Modell verbindet die Vorteile einer dezentralen Kuratierung mit den Vorteilen eines zentralen Zugangs. Im Rahmen der verteilten Archivierung haben Forschende die Möglichkeit, ihre Datenbestände an zentraler Stelle in die Infrastruktur einzuspeisen; die Daten werden anschließend durch den VerbundFDB an die jeweils zuständigen Datenzentren weitergeleitet. Im Verbund FDB wird ein zentraler Datennachweis im Online-Portal erstellt. Durch die Datenzentren erfolgt die fachspezifische Kuratierung, Archivierung und Bereitstellung der Forschungsdaten. Das FDZ Bildung am DIPF ist für die qualitativen Daten zuständig, das GESIS Datenarchiv für Surveydaten und das FDZ am IQB für Datensätze mit Kompetenz- und Leistungsmessungen.

⁴ Derzeit wird allerdings an mehreren Stellen hieran gearbeitet. Ein Beispiel für eine deutsche Entwicklung stellt das Projekt DataWiz dar, indem ein Tool zum Management psychologischer Forschungsdaten entwickelt wird. Vgl. www.zpid.de/index.php?wahl=forschung&uwahl=informationstechnologie&uuwahl=DataWiz (Zugegriffen: 17. Feb. 2017).

Abbildung 1: Struktur der verteilten Archivierung mit zentralem Zugang



Eine solche Struktur ermöglicht es, die datentypspezifischen Kompetenzen der beteiligten Datenzentren synergetisch zu nutzen und gleichzeitig gegenüber den Forschenden einen einheitlichen und nutzerfreundlichen Auftritt zu realisieren. Durch die Zusammenführung der Angebote unterschiedlicher Datenzentren kann die Diversität und die Heterogenität der Daten und des methodischen Vorgehens, die die Bildungsforschung kennzeichnet, adäquat fachspezifisch betreut werden.⁵ Gleichzeitig führt das gemeinsame, FDZ-übergreifende Portal www.forschungsdaten-bildung.de zu einer erhöhten Sicht- und Auffindbarkeit der verteilt aufbewahrten Forschungsdaten und dient als zentrale Anlaufstelle für die Forschenden.

4. Die Services für die Forschung

Der VerbundFDB hat in der Pilotphase (2013-2016) ein Angebotsportfolio entwickelt, das aus verschiedenen Komponenten besteht (vgl. Abbildung 2). Über das Web-Portal www.forschungsdaten-bildung.de können sich Bildungsforscher/innen einen Überblick über Studien und die zur Sekundärnutzung verfügbaren Datenbestände verschaffen. Der VerbundFDB weist Daten an zentraler Stelle nach und gewährleistet so die standortunabhängige Sichtbarkeit und Auffindbarkeit von Daten (VerbundFDB-Arbeitsbereich Information und Dokumentation (IuD)). Das Onlineportal dient gleichzeitig als Anlaufstelle, über die Bildungsforscher/innen eigene Datenbestände in die Infrastruktur einspeisen können. Die im VerbundFDB zusammengeschlossenen Datenzentren bieten die sichere Archivierung und Bereitstellung der Daten für Dritte an. Die Nutzbarkeit bzw. Interpretierbarkeit der Daten wird geprüft und ggf. ergänzt, die Daten werden zugänglich gemacht sowie langfristig erhalten. Schließlich bietet der VerbundFDB umfassende Informations-, Beratungs- und Schulungsmaßnahmen zum Themenkomplex Datenmanagement an. Ziel dessen ist es, zur Professionalisierung des Forschungsdatenmanagements in den Projekten beizutragen und so die Produktion qualitätsgesicherter Forschungsdaten zu unterstützen. Während die Aktivitäten des VerbundFDB in der ersten Förderphase noch speziell

⁵ Die Bedeutung fachspezifischer Archivierung sieht auch die DFG: „Der Umgang mit Forschungsdaten ist in hohem Maße von den Arbeitsweisen wissenschaftlicher Disziplinen geprägt. Daher fordert die Deutsche Forschungsgemeinschaft Fächer, Fachgesellschaften und Communities dazu auf, ihren Umgang mit Forschungsdaten zu reflektieren und angemessene Regularien zur disziplinspezifischen Nutzung und ggf. offenen Bereitstellung von Forschungsdaten zu entwickeln.“(DFG o. J.)

auf die im Rahmenprogramm zur Förderung der empirischen Bildungsforschung erzeugten Daten fokussiert waren, sind die zukünftigen Services für alle Bildungsforscher/innen geöffnet.

Abbildung 2: VerbundFDB-Arbeitsbereiche und deren jeweilige Ziele und Zwecke

VerbundFDB-Arbeitsbereiche/Services	Information und Dokumentation (IuD): Daten nachweisen	Archivierung und Bereitstellung: Daten sichern und aufbereiten	Schulung, Beratung und Information zum Forschungsdatenmanagement
Ziele und Zwecke	Sichtbarkeit und Auffindbarkeit von Forschungsdaten	Nutzbarkeit und Erhalt von Forschungsdaten	Professionalisierung des Forschungsdatenmanagements
Forschende können ...	Daten suchen und finden	Daten teilen und nutzen	sich informieren und qualifizieren

4.1. Forschungsdaten finden

Ein wichtiger Baustein der Nachhaltigkeit von Forschungsdaten ist deren Sichtbarkeit und Auffindbarkeit. Um diese zu erhöhen, ist es wichtig, eine zentrale Recherchemöglichkeit anzubieten und Daten an zentraler Stelle nachzuweisen. Eine solche Möglichkeit bietet das Online-Portal www.forschungsdaten-bildung.de. Seit 2015 können Forschende empirische Studien beim VerbundFDB melden. Erfasst werden Rahmendaten zum Projekt (wie Projektteam und Laufzeit), zur einzelnen Studie (Methodik, Erhebungszeiträume u. a.) und zu den konkreten Datentypen (quantitativ, qualitativ; Erhebungsinstrumente). Die Angaben der Forschenden werden durch die dokumentarisch und fachlich ausgebildeten Mitarbeiter/innen der Datenzentren im VerbundFDB geprüft (bspw. auf Sensibilität, Konsistenz, Plausibilität), es werden (teil)automatisiert standardisierte Metadaten⁶ erstellt und die vorhandenen Metadaten mit weiteren relevanten Kontextinformationen (etwa Publikationen) verknüpft. Derzeit sind über 150 Studien im Online-Portal des VerbundFDB katalogisiert. Sowohl Projekte, die primär Daten erstellt haben, als auch solche, die sekundär Daten genutzt haben, werden über das Portal auffindbar gemacht. Über entsprechende Such- und Recherchefunktionen können interessierte Nutzer/innen sowohl über Projekte hinweg als auch innerhalb einzelner Projekte recherchieren. Nutzer/innen können sich somit im Portal über den Forschungsdatenbestand und die Verfügbarkeit bestimmter Forschungsdaten informieren. Sofern Daten aus den im Portal dokumentierten Studien zur Nachnutzung zugänglich sind, werden Interessierte an die jeweils datenhaltenden Einrichtungen des DIPF, der GESIS, des IQB oder anderer Institute weiter verwiesen. Dort erhalten sie detaillierte Informationen zu den Daten selbst und können auf die Datenbestände in der Regel nach Antragstellung zugreifen.

Eine solche zentrale Übersicht, wie sie das Portal [forschungsdaten-bildung.de](http://www.forschungsdaten-bildung.de) derzeit aufbaut, fehlte bislang für die empirische Bildungsforschung. Aufgrund der Interdisziplinarität der Bildungsforschung stellt ein Übersichtsportal eine besondere Herausforderung, aber auch eine besondere Notwendigkeit dar, um den Austausch zwischen Forschenden und um kumulative, aufeinander aufbauende Forschung zu unterstützen. Das Portal wird in den nächsten Jahren systematisch um weitere Studien der Bildungsforschung ergänzt werden. Für den weiteren Ausbau des Portals soll zukünftig – ergänzend zu den Datenmeldungen der jeweiligen Forscher/innen – ein automatischer Import von Studienmetadaten aus anderen Portalen und Datenbanken umgesetzt werden.⁷

⁶ Das Metadaten-set orientiert sich an international gängigen Standards wie DDI und ist interoperabel.

⁷ Eruiert wird derzeit, ob ein regelmäßiger Import von Metadaten aus [da|raSearchNet \(www.gesis.org/?id=10497\)](http://www.gesis.org/?id=10497) umsetzbar ist. Das Portal [da|raSearchNet](http://www.gesis.org/?id=10497) enthält u. a. Datenbestände der Bildungsforschung, die von Datenzentren dort registriert werden, um eine DOI zu erhalten.

4.2. Forschungsdaten teilen: Angebote für Datengeber/innen

Um Forschungsdaten zu teilen, können Forschende diese zentral über das Portal www.forschungsdaten-bildung.de melden und übermitteln. Anschließend werden die Forschungsdaten innerhalb des VerbundFDB an das jeweils zuständige, fachlich kompetente Datenzentrum weitergeleitet. Dort erfolgt ihre Kuratierung (Dokumentation, Aufbereitung, Aufbewahrung, Erhalt; vgl. bspw. Harvey 2010) und Bereitstellung zur Nachnutzung durch Dritte. Die Kuratierung der Daten ist erforderlich, da Sekundärnutzungsmöglichkeiten von der Interpretierbarkeit und Nutzbarkeit der Daten abhängen. Daten müssen inhaltlich verständlich sein, eine ausreichende Datenqualität besitzen sowie bestimmte technisch-formale und rechtliche Bedingungen erfüllen. Fachspezifische FDZ führen im Rahmen des sogenannten *Ingests* (der Aufnahme von Daten in den Bestand des Archivs) daher Prüfprozesse durch, die auf die Verbesserung der Interpretier- und Nutzbarkeit der Daten zielen. Dazu gehören Prüfungen der Anonymität, Sensibilität, Konsistenz, Vollständigkeit und Plausibilität. Gegebenenfalls sind weitere Anonymisierungen, Aufbereitungen oder Dokumentationen vorzunehmen, oder die Daten sind vollständig oder teilweise von der Veröffentlichung auszuschließen.

Die Bedingungen der Archivierung und Bereitstellung der übermittelten Daten werden durch die Datengeber/innen mit den jeweils datenhaltenden FDZ vereinbart.⁸ Zu klären ist, welche Daten zu welchen Zeitpunkten und in welcher Form veröffentlicht werden sollen. Die Forschungsdaten können sofort nach Aufbereitung bereitgestellt werden oder aber erst nach Ablauf einer Sperrfrist. Sperrfristen oder Embargos können beispielsweise bis zur Fertigstellung von Qualifizierungsarbeiten eingerichtet werden. Anstelle der Sperrung des gesamten Datenbestands ist es häufig ausreichend, bestimmte Forschungsfragen oder einzelne Variablen für einen gewissen Zeitraum zu sperren. Datengeber/innen haben auch die Möglichkeit zu vereinbaren, dass sie über Nutzeranfragen – ohne Nennung der Person – informiert werden. Zum Ende der ersten Förderphase (September 2016) wurden 46 Datenbestände von Studien des EBF-Rahmenprogramms an den VerbundFDB übermittelt; darunter sind 13 Studien mit qualitativen Forschungsdaten, acht Datenbestände aus Umfragestudien und 25 Datenbestände aus Studien mit Kompetenz- und Leistungsmessungen (vgl. hier und im Folgenden: Meyermann et al. 2017). Von den 46 Datenbeständen konnten bislang 23 zur Nachnutzung verfügbar gemacht werden; in 17 Fällen ist keine Bereitstellung für Dritte möglich, in der Regel aus datenschutzrechtlichen Gründen; die übrigen sechs Datenbestände befanden sich Ende der ersten Förderphase noch in Bearbeitung. In acht Fällen (von 23) wurde eine Sperrfrist vereinbart.

Die Übergabe an ein FDZ hat für die Primärforscher/innen den Vorteil, dass die Datenzentren die sichere Aufbewahrung der Forschungsdaten übernehmen, Zugriffe kontrollieren und die Nutzerverwaltung und -betreuung übernehmen. Diese formalisierte, kontrollierte und rechtssichere Form der Weitergabe von Daten ersetzt so eine oftmals lediglich informell vereinbarte Weitergabe von Daten innerhalb persönlicher Netzwerke auf Vertrauensbasis. Insbesondere die Nutzer/innenbetreuung ist eine Dienstleistung, die fachspezifische FDZ von generischen Repositorien unterscheidet.

4.3. Forschungsdaten nutzen: Angebote für Datennutzer/innen

Nachdem Interessierte verfügbare Forschungsdaten über das zentrale Suchportal identifiziert haben, können sie diese bei den jeweiligen datenhaltenden Einrichtungen nachnutzen. Bei den Datenzentren von GESIS, IQB und DIPF erfolgt die Bereitstellung von Daten über verschiedene Zugriffsstufen und Zugangswege. Die *Zugriffsstufen* unterscheiden sich im Ausmaß der Zugriffsrestriktionen, bedingt

⁸ Datengebevereinbarung der GESIS: www.gesis.org/fileadmin/upload/institut/wiss_arbeitsbereiche/datenarchiv_analyse/Archivierungsvertrag_GESIS_Datenarchiv_annotiert.pdf, des IQB: www.forschungsdaten-bildung.de/files/IQB_Nutzungsvertrag_Bereitstellung_v2.docx, des DIPF: www.forschungsdaten-bildung.de/files/FDZBildung_Weitergabevertrag_V2.pdf (Zugegriffen: 11. Okt. 2017).

bspw. durch datenschutzrechtliche Einschränkungen, die für die Nutzung eines bestimmten Datenbestands gelten. Zugriffsgeschützte Datenkollektionen sind in der Regel nur für Wissenschaftler/innen und nur nach Registrierung und Beantragung zugänglich. Beim FDZ Bildung am DIPF beispielsweise muss sich ein/e Interessent/in für sensible Daten als promovierte/r Wissenschaftler/in authentifizieren, die Nutzungsanfrage sowie das konkrete Forschungsvorhaben werden geprüft. Für jedes Forschungsvorhaben, d. h. für jede Nutzung eines Datenbestands, ist ein gesonderter Antrag zu stellen. Ein/e registrierte/r Nutzer/in erhält somit keinen Zugriff auf sämtliche Daten im Bestand eines Archivs, sondern nur auf diejenigen, für die ein Nutzungsantrag gestellt und genehmigt wurde. Durch Bestätigung einer Nutzungsvereinbarung verpflichtet sich der/die Nutzer/in u. a. zur Geheimhaltung sowie dazu, Re-Identifizierungsversuche zu unterlassen. Stark anonymisierte und wenig sensible Bestände unterliegen geringeren Zugriffsrestriktionen. Sie sind beispielsweise auch für den Einsatz in der Lehre oder als sog. *Public Use File* – frei zum Download – verfügbar. Für die Nachnutzung von Forschungsdaten sind verschiedene *Zugangswege* möglich. Daten können an Nutzer/innen per CD-ROM/DVD oder zum Download herausgegeben werden. GESIS unterhält einen Gastwissenschaftlerarbeitsplatz, der an speziell geschützten Rechnern in den Räumen des Datenzentrums Zugang zu besonders sensiblen Daten ermöglicht.⁹ Bei der Bereitstellung von Daten ist es das Ziel und die Aufgabe von Forschungsdatenzentren, nutzerfreundliche Zugangswege zu realisieren, um effizientes und ressourcenschonendes Forschen zu ermöglichen. Gleichzeitig gilt es, personenbezogene Daten und sensible Inhalte besonders zu schützen.

4.4. Forschungsdatenmanagement: Angebote für Datenproduzent/innen

Die Angebote für Datengeber/innen und Datennutzer/innen unter dem Dach des VerbundFDB werden begleitet von Informations-, Beratungs- und Schulungsaktivitäten zum Thema Datenmanagement. Unter Forschungsdatenmanagement (FDM) versteht man den systematischen und planvollen Umgang mit Forschungsdaten über den gesamten Lebenszyklus der Daten hinweg (vgl. z. B. Jensen 2012, Meyermann 2012, Verbund Forschungsdaten Bildung 2015, RatSWD 2016). Um Datenmanagement in diesem Sinne zu fördern, wird das Erstellen von Datenmanagementplänen im Rahmen der Beantragung von Fördermitteln zunehmend obligatorisch. Denn Daten müssen bestimmte Voraussetzungen erfüllen, um archivier- und nachnutzbar zu sein. Dies betrifft beispielsweise Aspekte der Datendokumentation (Daten sollten so dokumentiert sein, dass sie für Dritte, nicht an der Erhebung Beteiligte, ggf. auch nach mehreren Jahren, interpretierbar sind.) und rechtliche Aspekte (Für die Weitergabe personenbezogener Daten ist in der Regel das Einverständnis der Betroffenen notwendig.). Da sich die Nachnutzbarkeit von Forschungsdaten zum großen Teil nicht durch die Arbeit von Forschungsdatenzentren *ex post* herstellen lässt, liegt sie vor allem im Aufgaben- und Verantwortungsbereich der datenerhebenden Forscher/innen. Notwendige Schritte sind bereits im Prozess der Antragstellung, Designplanung und Datenproduktion zu ergreifen. Aus diesem Grund ist es erforderlich, Forschende zum Thema Forschungsdatenmanagement umfassend und frühzeitig im Projektverlauf zu informieren und zu unterstützen.

Aufgrund der großen und steigenden Nachfrage baut der VerbundFDB seine bereits implementierten Schulungs-, Informations- und Beratungsangebote in den nächsten Jahren weiter aus. Der VerbundFDB bietet einführende Workshops zum Datenmanagement in der Bildungsforschung in verschiedenen Formaten (einstündig bis zweitägig) an (vgl. Trixa/Ebel 2015). Ziel dieser Workshops ist es, den Teilnehmenden ein grundlegendes Verständnis und praktisches Handlungswissen zur Konzeption und

⁹ Andere Datenzentren bieten darüber hinaus auch Zugänge über Remote Access oder kontrollierte Datenfernübertragung. Beim Remote Access kann ein/e Nutzer/in vom eigenen Arbeitsplatzrechner über eine gesicherte Verbindung auf Daten zugreifen, es ist ihm/ihr aber technisch nicht möglich, die Daten zu speichern. Bei der kontrollierten Datenfernverarbeitung (KDFV) schicken Nutzer/innen Auswertungsprogramme an das FDZ, FDZ-Mitarbeiter/innen führen die Auswertung der Daten durch und schicken die Ergebnisse (nach datenschutzrechtlicher Unbedenklichkeitsprüfung) an die Nutzer/innen zurück. Weiterführende Informationen finden sich bspw. in Schiller et al. 2017.

Umsetzung von FDM-Maßnahmen zu vermitteln. Um die Reichweite der Qualifizierungsangebote zu erhöhen, wird das bisherige Schulungsprogramm zukünftig durch niedrigschwelligere Angebote wie Webinare und ein begleitendes Diskussionsforum im Web ergänzt. Zusätzlich wird mit der Ausbildung von Multiplikatoren („Train-The-Trainer“) begonnen. Der VerbundFDB hat neben der Durchführung von Schulungen ein umfangreiches Informationsangebot zusammengestellt, das über die Website verfügbar ist (www.forschungsdaten-bildung.de/datenmanagement) und fortlaufend erweitert wird. Zudem bietet der VerbundFDB auf Anfrage bilaterale Beratung zu den unterschiedlichen Bereichen des Datenmanagements an.

4.5. Erfahrungen – Hürden der Datenarchivierung

Trotz der durch den VerbundFDB geschaffenen Infrastruktur konnten in der Pilotphase des Projektes (2013-2016) nicht alle der im EBF-Rahmenprogramm generierten Forschungsdaten gesichert werden, und nur zu einem geringen Teil konnten Daten zur Nachnutzung verfügbar gemacht werden. Bis zum Ende der ersten Förderphase des VerbundFDB (dem 30. Sep. 2016) lagen dem VerbundFDB 151 Meldungen für Projekte vor, in 130 dieser Projekte wurden Primärdaten erhoben, 46 Datenbestände aus Einzel- und Verbundprojekten wurden an den VerbundFDB zur Sicherung übermittelt und 23 Datenbestände konnten zur Nachnutzung verfügbar gemacht werden. Verschiedene Gründe verhinderten die Dokumentation und Archivierung weiterer Datenbestände aus dem EBF-Rahmenprogramm (ausführlich hierzu siehe Meyermann et al. 2017). Die im Folgenden beschriebenen Faktoren wurden aus den dem VerbundFDB vorgelegten Stellungnahmen und Unterlagen der Forscher/innen sowie persönlichen Gesprächen mit diesen identifiziert.

Dem VerbundFDB lagen (Stand 30. Sep. 2016) aus 91 Projekten Stellungnahmen zur Sicherung und Bereitstellung von Daten nach Projektende vor, die die Projektnehmer/innen gegenüber dem DLR Projektträger im Auftrag des BMBF abgeben mussten. Aus diesen geht hervor, dass es 32 von 91 Projekten nicht möglich war, ihre Daten zur Sicherung und Bereitstellung für Dritte an ein Datenzentrum weiterzugeben. Gründe hierfür waren

- datenschutzrechtliche und ethische Aspekte (N=24),
- die Archivierung ausschließende behördliche Genehmigungen (N=14), die für Erhebungen an Schulen erforderlich waren,
- nicht abgeschlossene Qualifizierungsarbeiten oder Publikationen (N=7) und
- urheberrechtliche Einschränkungen (N=2),

wobei Mehrfachnennungen möglich waren.

Auch in den meisten uns im persönlichen Austausch geschilderten Fällen war eine Datenarchivierung bei einem FDZ aus *datenschutzrechtlichen Gründen* nicht möglich, weil die Einwilligung der Betroffenen zur Archivierung nicht vorlag. Dabei handelte es sich in vielen Fällen nicht um eine dezidierte Ablehnung der Archivierung, sondern die Archivierung war nicht Gegenstand der Einwilligung. Ähnlich verhält es sich mit den uns vorgelegten *behördlichen Genehmigungen*, die für Erhebungen an Schulen eingeholt worden waren. Der Archivierungszweck der Daten war häufig nicht Bestandteil des Genehmigungsantrags. Zudem enthielten die ausgestellten Genehmigungen häufig Vorgaben zur ausschließlichen Nutzung oder Löschung sämtlicher – nicht nur der personenbezogenen – Daten im Projekt bzw. nach Projektende. Uns bekannte Hindernisse der Datenweitergabe aufgrund von *urheberrechtlichen Aspekten* bezogen sich jeweils auf durch die Primärforscher/innen selbst entwickelte Testinstrumente oder die Verwendung von urheberrechtlich geschütztem Material in eingesetzten Testinstrumenten.

Des Weiteren scheint eine *unzureichende Ressourcenausstattung* in den Projekten die Archivierung von Forschungsdaten in einigen Fällen zu verhindern oder zumindest erheblich zu erschweren. Dies ist

insbesondere dann ein Problem, wenn die spätere Archivierung nicht bereits in der Projektplanung berücksichtigt wurde oder die Anforderung zur Archivierung der Daten erst Jahre nach Projektabschluss erfolgt. In diesen Fällen sind Projektmitarbeiter/innen häufig nicht mehr verfügbar und die Aufgabe der Datenarchivierung hat für die ehemaligen Projektleiter/innen eine vergleichsweise geringe Priorität. Für den VerbundFDB war dies aufgrund des innerhalb des EBF-Rahmenprogramms späten Projektbeginns des VerbundFDB ein spezielles Problem: Eine Vielzahl von Projekten war zu diesem Zeitpunkt schon ausgelaufen oder die Erhebungsphase bereits beendet.

Letztlich wurden in persönlichen Gesprächen von einzelnen Forschenden *Bedenken gegenüber der Archivierung* der eigenen Forschungsdaten geäußert, wie sie auch in der Literatur berichtet werden (vgl. bspw. Borgman 2010; Haas und Park 2010; Häussler 2010; RIN 2008). Ein Teil dieser Bedenken bezieht sich auf den mit der Herausgabe der Daten verbundenen Kontrollverlust über die eigenen Daten: Die Forschenden sind unsicher, ob trotz Datenweitergabe die Anonymität der Befragungspersonen zu jeder Zeit gewahrt bleibt und wie, durch wen und zu welchen Zwecken die „eigenen“ Daten nachgenutzt werden. Forschende sind hier mit widersprüchlichen Anforderungen konfrontiert: Sie sind zum einen gegenüber den Studienteilnehmer/innen zu Verschwiegenheit verpflichtet und zum anderen gegenüber den Rezipienten ihrer Forschungsergebnisse zu Transparenz. Die Angebote von Forschungsdatenzentren können dabei unterstützen, dieses Dilemma – Verschwiegenheits- vs. Transparenzverpflichtung – aufzulösen, sodass Forschende beiden Anforderungen gerecht werden können. FDZ gewährleisten auf der einen Seite Verschwiegenheit, indem sie geschützte Räume der Aufbewahrung und des Zugriffs bieten, und auf der anderen Seite Transparenz, indem Sie Auffindbarkeit und Nutzbarkeit der Forschungsdaten für Dritte ermöglichen. An FDZ stellt dies jedoch hohe Anforderungen an Qualität und Vertrauenswürdigkeit, was zu weiteren Professionalisierungsprozessen in den kommenden Jahren führen wird.

Die in der Vergangenheit bereits ergriffenen Maßnahmen, die darauf zielen, den genannten Hürden für die Datenarchivierung und Nachnutzung zu begegnen, werden in der zweiten Förderphase des VerbundFDB fortgeführt und ergänzt. Unter anderem werden die behördlichen Genehmigungspraxen der Schul- und Unterrichtsforschung, die aktuell häufig eine Weiternutzung der erhobenen Daten ausschließen, adressiert. Ziel ist es, behördliche Entscheider für das Thema der Datenarchivierung und -nachnutzung zu sensibilisieren und dazu beizutragen, vorhandene Spielräume – sofern datenschutzrechtlich möglich – zu nutzen. Des Weiteren zeigen sich insbesondere bei qualitativen Daten eine Reihe offener Fragen, die im Austausch mit den Scientific Communities diskutiert und bearbeitet werden. Die Beratungs- und Schulungsleistungen für Forscher/innen zu Fragen des Datenmanagements und insbesondere zu datenschutzrechtlichen Fragen werden ausgebaut.

In Deutschland existieren förderseitige Verpflichtungen zur Datenweitergabe, wie sie im EBF-Rahmenprogramm vorlagen, erst seit kurzem. Die Verpflichtung der EBF-Projekte bezog sich auf die retrospektive Archivierung der Daten; das heißt, Projekte wurden verpflichtet, Forschungsdaten *nach* Projektende weiterzugeben. Erschwerend hinzu kam in diesem Fall, dass die Möglichkeiten zur Archivierung über den VerbundFDB erst seit 2014 und damit vergleichsweise spät vorhanden waren. Die Erfahrungen des VerbundFDB-Projektes zeigen, dass eine derartige retrospektive Verpflichtung zu kurz greift. Die Archivier- und Nachnutzbarkeit von Forschungsdaten, lässt sich ex post nicht bzw. nur sehr aufwändig herstellen. Um archivier- und nachnutzbar zu sein, müssen Daten bestimmte Kriterien erfüllen, insbesondere rechtlicher und dokumentarischer Art. Daher sollten Forscher/innen systematisches Forschungsdatenmanagement über den gesamten Projektzeitraum hinweg betreiben (vgl. RatSWD 2016, Verbund Forschungsdaten Bildung 2015). Ein Blick in den angelsächsischen Raum, in dem Erfahrungen mit Verpflichtungen zur Datenweitergabe bereits seit längerem existieren, zeigt, dass diese mittlerweile ihre Verpflichtungen zur Datenweitergabe um solche zum Datenmanagement und zum Verfassen von Datenmanagementplänen ergänzt haben (vgl. auch Meyermann 2012). Auf Basis

der Erfahrungen des VerbundFDB ist dringend zu empfehlen, Datenmanagement in den Projekten selbst zu fördern und Beratungs- und Schulungsmaßnahmen sowie zusätzliche Ressourcen (Laufzeit, Mittel, Personal) für die Projekte bereitzustellen. Verpflichtungen zur Datenweitergabe sollten, um erfolgreich zu sein, von Maßnahmen zur Förderung des Forschungsdatenmanagements in den Projekten begleitet werden. Mittlerweile ist es sowohl bei der DFG als auch beim BMBF möglich, Ressourcen für das Forschungsdatenmanagement zu beantragen.¹⁰

Neben einer verbesserten Ressourcenausstattung und der Vermittlung von Kenntnissen des Forschungsdatenmanagements, fehlt es im Wissenschaftssystem derzeit noch immer an positiven Anreizen für die Datenarchivierung und -weitergabe. Erstrebenswert wäre eine stärkere Berücksichtigung von Forschungsdatenpublikationen neben den klassischen Publikationen bspw. im Rahmen von Begutachtungen, Evaluationen und Stellenbesetzungsverfahren, damit sich eine „Publikationskultur für Forschungsdaten“ (Büttner et al. 2011: 19) herausbilden kann.

5. Zukünftige Handlungsfelder

Voranstehend wurden die Angebote des VerbundFDB vorgestellt, die Organisationsstruktur der verteilten Archivierung im VerbundFDB erläutert und die Erfahrungen aus der ersten Förderphase berichtet. Vor dem Hintergrund der Erfahrungen der Pilotphase und der beschriebenen Ausgangslage (vgl. Kapitel 2 und 3) besteht dringender Bedarf nach einem Ausbau der bestehenden Forschungsdateninfrastruktur für die empirische Bildungsforschung. Der VerbundFDB sieht diesbezüglich drei Handlungsfelder: Erstens gilt es, das föderierte Angebot auszubauen. Ein zentraler Zugang ist notwendig, um die gegenwärtige Unübersichtlichkeit für die Nutzer/innen durch die verteilt vorliegenden Angebote zu reduzieren und die Nutzerfreundlichkeit zu erhöhen. Das Onlineportal www.forschungsdatenbildung.de kann für die Bildungsforschung und die Forschungsdateninfrastruktur die Funktionen einer zentralen Verteilerstelle und eines zentralen Datennachweises erfüllen.

Um die verteilte Archivierung und den zentralen Datennachweis auszubauen und dort auch Daten außerhalb des EBF-Rahmenprogramms und außerhalb der drei am VerbundFDB beteiligten Institute IQB, GESIS und DIPF sichtbar zu machen, ist es sinnvoll, weitere Forschungsdatenzentren sowie universitäre, fachübergreifende Repositorien einzubeziehen. In einem solchen Netzwerk gilt es einen Metadaten austausch zu realisieren sowie gemeinsam eine weitere Harmonisierung der dokumentarischen Verfahren (u. a. abgestimmte Metadaten sets, vergleichbare Qualitätsstandards) herbeizuführen, um die Vergleichbarkeit von Datendokumentationen über Institute, verschiedene Portale und Datenbanken hinweg herzustellen. Auf diese Weise kann es gelingen, die Auffindbarkeit und Sichtbarkeit von Forschungsdaten weiter zu erhöhen. Der VerbundFDB organisiert in der zweiten Förderphase jährliche Netzwerktreffen interessierter FDZ, um über gemeinsame Standards der Archivierung zu beraten und diese weiter zu entwickeln. Das erste dieser Art fand bereits im März 2017 in Berlin statt (vgl. Harzenetter 2017). Erfolgskritisch wird es sein, dass sich die föderierte Forschungsdateninfrastruktur als professioneller und vertrauenswürdiger Anbieter von Forschungsdatenservices etabliert.

Ein weiteres Handlungsfeld im Rahmen des Aufbaus einer föderierten Infrastruktur für die empirische Bildungsforschung besteht in der bedarfsgerechten Weiterentwicklung der infrastrukturellen Angebote. Der Ausbau sollte sowohl die Bandbreite der Datentypen als auch den hohen Datenoutput der empirischen Bildungsforschung abdecken. Um dies zu realisieren und die Aufnahmekapazität der Forschungsdateninfrastruktur insgesamt zu erweitern, plant der VerbundFDB soweit möglich weitere FDZ über das Modell der verteilten Archivierung (vgl. Abbildung 1) oder alternative Kooperations-

¹⁰ Vgl. bspw. DFG-Vordruck 54.01 - 07/17, www.dfg.de/formulare/54_01/54_01_de.pdf, S. 5 (Zugegriffen: 17. Jul. 2017) und BMBF-Bekanntmachung vom 22. Feb. 2017, <https://www.bmbf.de/foerderungen/bekanntmachung-1326.html> (Zugegriffen: 17. Jul. 2017)

modelle zu integrieren. Zudem sollen in der zweiten Förderphase die offenen Fragen verstärkt in den Blick genommen werden, die in Zusammenhang mit der Archivierung und Nachnutzung qualitativer Daten stehen. Um der – aufgrund der veränderten Förderauflagen – zukünftig zu erwartenden, steigenden Archivierungsnachfrage gerecht zu werden, gilt es im Übrigen, bestehende Service- und Selektionsmodelle weiterzuentwickeln und für unterschiedliche Arten von Daten je spezifische Services im Spektrum von hochspezialisiert bis niedrigschwellig anzubieten. Servicemodelle dienen der differenzierten Bearbeitung von Daten, bspw. hinsichtlich der Tiefe der Aufbereitung oder der Dauer der Aufbewahrung. Selektionsmodelle sind Verfahren und Kriterien zur Bewertung von Daten, um diese den verschiedenen Servicestufen in nachvollziehbarer und transparenter Weise zuweisen zu können (vgl. z. B. Whyte/Wilson 2010). Die Infrastruktur und die Bildungsforschung stehen vor der Herausforderung zu entscheiden, welche Daten für welchen Zeitraum, in welcher Form zu archivieren und bereitzustellen sind (bspw. über FDZ oder universitäre Repositorien, vollständig oder teilweise, umfänglich oder basal aufbereitet und dokumentiert).¹¹ Letztlich sollte das Archivierungsangebot eine Bandbreite zwischen hochspezialisierten Services und niedrigschwelligem Diensten bereithalten.

Ein drittes und wesentliches Handlungsfeld des Aufbaus einer Infrastruktur für die empirische Bildungsforschung wird darin gesehen, die Community einzubinden und Konsortien zur kooperativen Weiterentwicklung der Infrastruktur zu bilden. Die Forschungsdateninfrastruktur versteht sich als Dienstleister und die Bildungsforschung demzufolge als Adressat dieser Dienstleistungen. In den verschiedenen Rollen als Datenproduzenten, Datengeber und -nutzer sowie Experten ihres Faches sollten Bildungsforscher/innen im Prozess der Angebotsentwicklung und -weiterentwicklung beteiligt sein: frühzeitig (Bedarfsmessung), im Verlauf (Mitgestaltung) und im Nachhinein (Evaluation). Darüber hinaus sind die Scientific Communities jedoch auch als Partner und Mitgestalter der Infrastruktur gefragt. Eine aktive Beteiligung ist bei der Entwicklung archivarischer Standards selbst erforderlich sowie im Rahmen von strategischen Diskussionen und in konkreten Prozessen wie Begutachtungsverfahren von Datenmanagementplänen und Daten. Der VerbundFDB steht – neben den eigenen wissenschaftlichen Tätigkeiten durch FDZ-Angehörige – über seine Schulungs-, Informations- und Beratungsangebote zum Forschungsdatenmanagement im engen Austausch mit der Bildungsforschung. Zu ausgewählten Themen (Archivierung qualitativer Daten, Nachnutzungspotentiale von Daten) wird der VerbundFDB zukünftig Workshops in Kooperation mit Wissenschaftler/innen organisieren. Zudem bietet der VerbundFDB seine Beteiligung an Fachdiskussionen an und unterstützt bspw. bei der Erstellung von Stellungnahmen zur Archivierung von Daten.

Alle Akteure sind gefragt, damit Data Sharing zu einer leistungsstarken Bildungsforschung beitragen kann und die Prinzipien von Fair Data¹² umgesetzt werden. Das gemeinsam zu verfolgende Ziel besteht darin, ein bedarfsnahes und nutzerfreundliches Angebot zu erstellen und Forschende dabei zu unterstützen, Forschungsdaten in nachhaltiger Form zu erzeugen und nachzunutzen und damit letztlich den wissenschaftlichen Erkenntnisgewinn zu fördern.

¹¹ Im Memorandum der DFG für die Erziehungswissenschaften (Stanat 2014) wird unterschieden zwischen „a) Archivierung und auf Nachfrage nutzerfreundliche Bereitstellung direkt durch den Datenproduzenten zwecks Prüfung publizierter Ergebnisse, b) Archivierung, erweiterte Dokumentation (Codebuch) und auf Nachfrage nutzerfreundliche Bereitstellung direkt durch die Datenproduzenten zwecks weiterführender wissenschaftlicher Analysen, c) Gut dokumentierte Übergabe an ein Forschungsdatenzentrum zwecks Archivierung und allgemeiner Bereitstellung an die Scientific Community nach den Regularien des jeweiligen FDZ.“ (Stanat 2014, S. 4 f.) (vgl. auch RatSWD 2016). Die Trennlinie wird hier zwischen Aufbewahrungsorten (Datenproduzent oder FDZ) und Aufbewahrungszwecken (Replikation oder Nachnutzung) gezogen, die in der Regel mit unterschiedlichen Dokumentationsniveaus einhergehen.

¹² Vgl. www.force11.org/group/fairgroup/fairprinciples (Zugegriffen: 17. Jul. 2017).

6. Literatur

- Allianz der Wissenschaftsorganisationen. 2010. Grundsätze zum Umgang mit Forschungsdaten. www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Grundsaeetze_Forschungsdaten_2010.pdf (Zugegriffen: 09. Mai 2017).
- Borgman, Christine L. 2010. Research Data: who will share what, with whom, when, and why? *RatSWD Working Paper Series* 161. www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_161.pdf (Zugegriffen: 08. Mai 2017).
- Büttner, Stephan, Hans-Christoph Hobohm und Lars Müller. 2011. Handbuch Forschungsdatenmanagement. Bad Honnef: BOCK + HERCHEN Verlag.
- DFG. o. J.. www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf (Zugegriffen: 04. Mai 2016).
- Forschungsdaten Bildung. o. J. Übersicht über die länderspezifischen Besonderheiten für Befragungen an Schulen. www.forschungsdaten-bildung.de/genehmigungen (Zugegriffen: 18. Apr. 2016).
- Haas, Martine R., und Sangchan Park. 2010. To Share or Not to Share? Professional Norms, Reference Groups, and Information Withholding Among Life Scientists. *Organization Science* 21 (4): 873–891. <http://orgsci.journal.informs.org/cgi/content/abstract/21/4/873> (Zugegriffen: 08. Mai 2017).
- Häussler, Carolin. 2010. Information-Sharing in Academia and the Industry. A Comparative Study. *RatSWD Working Paper Series* 154. www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_154.pdf (Zugegriffen: 08. Mai 2017).
- Harvey, Ross. 2010. Digital Curation, A How-To-Do-It Manual. New York, London: Neal-Schuman Publishers.
- Harzenetter, Karoline. 2017. Bericht vom ersten Netzwerktreffen des Verbunds Forschungsdaten Bildung. www.forschungsdaten-bildung.de/files/VFDB_NT1_Bericht01_201705.pdf (Zugegriffen: 26. Juni 2017).
- Hochschulrektorenkonferenz (HRK). 2014. Management von Forschungsdaten – eine zentrale strategische Herausforderung für Hochschulleitungen. Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main. www.hrk.de/fileadmin/migrated/content_uploads/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf (Zugegriffen: 31. Mai 2017).
- Hochschulrektorenkonferenz (HRK). 2015. Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien. Empfehlung der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel. www.hrk.de/uploads/tx_szconvention/Empfehlung_Forschungsdatenmanagement_final_Stand_11.11.2015.pdf (Zugegriffen: 04. Mai 2016).
- Jensen, Uwe. 2012. Leitlinien zum Management von Forschungsdaten, Sozialwissenschaftliche Umfragedaten. *gesis Technical Reports* 7. www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf (Zugegriffen: 09. Mai 2017).
- Kommission Zukunft der Informationsinfrastruktur. 2011. Gesamtkonzept für die Informationsinfrastruktur in Deutschland. Empfehlungen der Kommission Zukunft der Informationsinfrastruktur im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder. www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Infrastruktur/KII_Gesamtkonzept.pdf (Zugegriffen: 05. Mai 2017).

- Meyermann, Alexia. 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 (Zugegriffen: 05. Mai 2017).
- Meyermann, Alexia und Maike Porzelt. 2017. Datenschutzrechtliche Anforderungen in der empirischen Bildungsforschung - eine Handreichung. *forschungsdaten bildung informiert 6*. Verfügbar unter: www.forschungsdaten-bildung.de/forschungsdaten-bildung-informiert (Zugegriffen: 11. Okt. 2017).
- Meyermann, Alexia, Doris Bambey, Thomas Ebel, Marcus Eisentraut, Malte Jansen, Poldi Kuhl, Reiner Mauer, Claudia Neuendorf, Lisa Pegelow, Maike Porzelt, Marc Rittberger, Thomas Schwager, Petra Stanat, Jessica Trixa. 2017. Schlussbericht zum Verbundprojekt „Sicherung und Nachnutzung der Forschungsdaten des Rahmenprogramms zur Förderung der Empirischen Bildungsforschung“, Förderkennzeichen 01JS1302A-C. <https://www.tib.eu/suchen/id/TIBKAT:897124898/> (Zugegriffen: 11. Okt. 2017).
- OECD. 2013. New Data for Understanding the Human Condition: International Perspectives. OECD Global Science Forum Report on Data and Research Infrastructure for the Social Sciences. www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.pdf (Zugegriffen: 16. Juni 2017).
- RatSWD. 2016. Forschungsdatenmanagement in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. Orientierungshilfen für die Beantragung und Begutachtung datengenerierender und datennutzender Forschungsprojekte. Output 3. www.ratswd.de/dl/RatSWD_Output3_Forschungsdatenmanagement.pdf. (Zugegriffen: 11. Okt. 2017).
- RIN (Report commissioned by the Research Information Network). 2008. To Share or Not To Share: Publication and Quality Assurance of Research Data Outputs. Main Report. https://eprints.soton.ac.uk/266742/1/Published_report_-_main_-_final.pdf (Zugegriffen: 05. Mai 2017).
- rdn (research data netherlands). o. J. A federated data infrastructure for the Netherlands: the front-office – back-office model. http://www.researchdata.nl/fileadmin/content/RDNL_algemeen/Documenten/RDNL_FOBOModel-UK-web.pdf (Zugegriffen: 09. Mai 2017).
- Rat für Informationsinfrastrukturen (RfII). 2016. Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland. www.rfii.de/?wpdmdl=1998 (Zugegriffen: 05. Mai 2017).
- Reinhold, Anke, Marc Rittberger und Sven Keil. 2013. A thematic analysis to facilitate data curation relevance decisions - The case of research data in video studies. In *Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten*, Hrsg. Hans-Christoph Hobohm, 198-209. Glückstadt: Hülsbusch.
- Scheller, Jürgen. 2017. Rechtliche Rahmenbedingungen der Verwendung von Videos in der Schul- und Unterrichtsforschung. Diskrepanzen zwischen Datenschutzrecht, Förder- und Genehmigungsaufgaben. *forschungsdaten bildung informiert 5*. www.forschungsdaten-bildung.de/forschungsdaten-bildung-informiert (Zugegriffen: 11. Okt. 2017).
- Schiller, David H., Johanna Eberle, Daniel Fuß, Jan Goebel, Jörg Heining, Tatjana Mika, Dana Müller, Frank Röder, Michael Stegmann und Karsten Stephan. 2017. Standards des sicheren Datenzugangs in den Sozial- und Wirtschaftswissenschaften: Überblick über verschiedene Remote-Access-Verfahren. *RatSWD Working Paper Series 261*. www.ratswd.de/dl/RatSWD_WP_261.pdf (Zugegriffen: 17. Jul. 2017).
- Schnettler, Bernt und Frederik S. Pötzsch. 2007. Visuelles Wissen. In *Handbuch Wissenssoziologie und Wissensforschung*, Hrsg. R. Schützeichel, 472–484. Konstanz: UVK.
- Stanat, Petra. 2014. Bereitstellung und Nutzung quantitativer Forschungsdaten in der Bildungsforschung: Memorandum des Fachkollegiums „Erziehungswissenschaft“ der DFG. http://dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/informationen_fachwissenschaften/geisteswissenschaften/forschungsdaten_memorandum_fk_109.pdf (Zugegriffen: 08. Mai 2017).

Trixa, Jessica, und Thomas Ebel. 2015. „Einführung in das Forschungsdatenmanagement in der empirischen Bildungsforschung“. Zusammenfassung des Workshops im Rahmen des CESSDA-Trainingprogramms. *forschungsdaten bildung informiert 2*. www.forschungsdaten-bildung.de/forschungsdaten-bildung-informiert (Zugegriffen: 08. Mai 2017).

Verbund Forschungsdaten Bildung. 2015. Checkliste zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung. Version 1.1. *fdbinfo 2*. www.forschungsdaten-bildung.de/files/fdbinfo_2.pdf (Zugegriffen: 09. Mai 2017).

Whyte, Angus, und Andrew Wilson. 2010. How to Appraise and Select for Curation. A Digital Curation Centre und Australian National Data Service 'working level' guide. <http://www.dcc.ac.uk/sites/default/files/documents/How%20to%20Appraise%20and%20Select%20Research%20Data.pdf> (Zugegriffen: 02. Jul. 2014).

Wissenschaftsrat. 2012. Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020. www.wissenschaftsrat.de/download/archiv/2359-12.pdf (Zugegriffen: 05. Mai 2017).