

# Bericht vom zweiten Netzwerktreffen des Verbund Forschungsdaten Bildung

Verbund Forschungsdaten Bildung // Juni 2018

## Einleitung

---

Am 27. April 2018 lud der Verbund Forschungsdaten Bildung Vertreter/innen von Einrichtungen der Forschungsdateninfrastruktur im Bereich Bildungsforschung zu einem zweiten Vernetzungstreffen an das IQB nach Berlin ein. Das erste Treffen dieser Art fand am 29. März 2017 statt.<sup>1</sup>

Der VerbundFDB ist ein Projekt, das durch die Einrichtungen DIPF, GESIS und IQB getragen und durch das BMBF gefördert wird (1. Förderphase 2013-2016, 2. Förderphase 2016-2019). Im VerbundFDB wurde in den vergangenen Jahren der Grundstein für eine föderierte Infrastruktur gelegt, in der verschiedene FDZ gemeinsam, untereinander abgestimmte Datendienste für Forschende anbieten. Über die Projektpartner hinaus strebt der VerbundFDB einen engen Austausch mit weiteren für die Bildungsforschung relevanten FDZ an. Gemeinsames Ziel dieses Netzwerks ist es, die Sichtbarkeit und Nachnutzbarkeit von Forschungsdaten für die multidisziplinäre und heterogene Bildungsforschung zu erhöhen. Sozial-, verhaltens- und wirtschaftswissenschaftliche Bildungsforscher/innen haben die Möglichkeit, über das Portal [www.forschungsdaten-bildung.de](http://www.forschungsdaten-bildung.de) Forschungsdaten zur Nachnutzung zu finden sowie eigene Forschungsdaten zu übermitteln. Die eingehenden Forschungsdaten werden zur Archivierung und Bereitstellung an das jeweils zuständige, fachlich spezialisierte FDZ weitergeleitet. Im Netzwerk gilt es, sich über gemeinsame Standards und Verfahren der Archivierung auszutauschen, um ein hohes Maß an Qualität, Effizienz und Nutzerfreundlichkeit der angebotenen Dienste zu erreichen.

Das Schwerpunktthema des diesjährigen Netzwerktreffens lautete „Datenqualität – Auswahl und Bewertung“. Zudem wurden Neuigkeiten aus dem Verbundprojekt des vergangenen Jahres berichtet sowie die Ergebnisse der beim letzten Treffen gegründeten und unterjährig aktiven Arbeitsgruppe Metadaten vorgestellt und im größeren Kreis diskutiert.

## Anwesende

---

(in alphabetischer Reihenfolge)

Florence Baillet, Deutsches Zentrum für Hochschul- Wissenschaftsforschung (DZHW), Hannover  
Doris Bambey, DIPF – Bildungsforschung und Bildungsinformation, Frankfurt a. M.

---

<sup>1</sup> vgl. Harzenetter 2017, verfügbar unter: [www.forschungsdaten-bildung.de/files/VFDB\\_NT1\\_Bericht01\\_201705.pdf](http://www.forschungsdaten-bildung.de/files/VFDB_NT1_Bericht01_201705.pdf) (Zugriff am 07.06.2018).

Noemi Betancort-Cabrera, QualiService der Universität Bremen, Bremen  
Katarina Blask, Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID), Trier  
Anett Friedrich, Bundesinstitut für Berufsbildung (BIBB), Bonn  
Daniel Fuß, Leibniz-Institut für Bildungsverläufe e.V. (LifBi), Bamberg  
Malte Jansen, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Berlin  
Marius Gerecht, DIPF – Bildungsforschung und Bildungsinformation, Frankfurt a.M.  
Karoline Harzenetter, GESIS - Leibniz-Institut für Sozialwissenschaften, Köln  
Ute Hoffstätter, Deutsches Zentrum für Hochschul- Wissenschaftsforschung (DZHW), Hannover  
Harald Kaluza, Deutsches Institut für Erwachsenenbildung (DIE), Bonn  
Helge Kminek, Archiv für pädagogische Kasuistik (Apaek), Frankfurt a.M.  
Reiner Mauer, GESIS - Leibniz-Institut für Sozialwissenschaften, Köln  
Alexia Meyermann, DIPF – Bildungsforschung und Bildungsinformation, Frankfurt a. M.  
Claudia Neuendorf, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Berlin  
Lisa Pegelow, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Berlin  
Maike Porzelt, DIPF – Bildungsforschung und Bildungsinformation, Frankfurt a. M.  
Holger Quellenberg, Deutsches Jugendinstitut (DJI), München  
Thomas Runge, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin  
Dirk Weisbrod, DIPF – Bildungsforschung und Bildungsinformation, Frankfurt a. M.  
Knut Wenzig, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

## Ablauf / Tagesordnung

---

Nach einer **Begrüßung** durch die Veranstalter wurden **Neuigkeiten aus dem Verbundprojekt** aus dem letzten Jahr berichtet.

Die VerbundFDB-Koordinierungsstelle am DIPF dient über das Portal [forschungsdaten-bildung.de](http://forschungsdaten-bildung.de) als zentraler Ansprechpartner für Bildungsforscher/innen. Der VerbundFDB hat sich im letzten Jahr weiter etablieren können. So wird das Portal bspw. in den Förderausschreibungen des BMBF zum zweiten Rahmenprogramm explizit genannt. Forscher/innen haben die Möglichkeit, sich an den VerbundFDB zu wenden, um ein geeignetes Archiv zu finden oder sich zum Forschungsdatenmanagement und der Erstellung von Datenmanagementplänen zu informieren. Bislang nehmen die drei Projektpartner DIPF, GESIS und IQB Daten auf, die über das Portal übermittelt werden. Aktuell ist geplant, dass das DZHW ebenfalls Daten aufnehmen wird. Ein weiterer Ausbau darüber hinaus ist wünschenswert, da aufgrund des zweiten Rahmenprogramms zukünftig mit einer weiter steigenden Nachfrage zu rechnen ist. Mit dem VerbundFDB können in diesen Fällen die Modalitäten ausgehandelt werden, bspw. welche Aufgaben die Koordinierungsstelle, welche die dezentralen FDZ übernehmen.

In der folgenden Diskussion wurden verschiedene Themen aufgegriffen: Der VerbundFDB bietet Forschenden primär ein Archivierungsangebot für solche Daten an, die bereitgestellt und nachgenutzt werden können; in der Regel jedoch nicht für Daten, die nur gesichert werden sollen. Für letztere eignen sich die institutionellen Repositorien in den eigenen Einrichtungen. Forschende sollten bereits im Datenmanagementplan die später archivierende Einrichtung nennen und sich mit dieser oder dem VerbundFDB frühzeitig in Verbindung setzen. Von Seiten der FDZ sollte angestrebt werden, sich untereinander hinsichtlich der Archivierungsbedingungen abzustimmen.

## Metadaten

---

Im zweiten Block des Tages wurde über die **Aktivitäten der AG Metadaten** im vergangenen Jahr berichtet und die nächsten Schritte vorgestellt. Mitwirkende in der AG sind neben den VerbundFDB-Projektpartnern die Vertreter des DZHW, des LfBi und des SOEP sowie die Vertreterin von Qualiservice.

Die **AG Metadaten** hat es sich zum Ziel gesetzt, den Metadaten austausch zwischen FDZ zu organisieren und die Nutzung gemeinsamer Metadatenstandards bzw. eines gemeinsamen Metadatensets voranzutreiben. Hiermit wird das Ziel verfolgt, die Anzahl der nachgewiesenen Daten im VerbundFDB-Portal zu erhöhen. Bisher werden dort vorwiegend Studien angezeigt, die bei den Projektpartnern (DIPF, GESIS, IQB) archiviert sind und insbesondere aus dem ersten Rahmenprogramm des BMBF zur Förderung der empirischen Bildungsforschung stammen.

Die AG hat zwei Wege der Nachweiserweiterung erarbeitet und vorgestellt. 1) Beim *direkten Weg* erfassen Datenzentren Metadaten direkt im Datenmanagement-System des VerbundFDB (DMM). Hierzu benötigen sie eine entsprechende Authentifizierung. Soweit möglich können hierfür in Zukunft automatisierte Lösungen entwickelt werden, die eine direkte Schnittstelle zwischen DMM und dem System des jeweiligen FDZ voraussetzen. 2) Beim *indirekten Weg* erfassen Datenzentren Metadaten bei dalra im Rahmen der DOI-Registrierung. Über einen Harvester werden die bei dalra registrierten Daten (Metadaten) in das VerbundFDB-Portal übermittelt. Hierzu müssen Datenzentren bei ihrer dalra-Registrierung der jeweiligen Studie einen Flag setzen, so dass der Harvester diese als VerbundFDB-relevant erkennt. Der direkte Metadaten austausch erlaubt es, umfangreichere Metadaten zu erfassen, als der Metadaten austausch über dalra.

Um eine vergleichbare, hohe Qualität der Metadaten zu gewährleisten, hat sich die AG Metadaten auf ein Kernset an verpflichtenden und empfohlenen Metadaten geeinigt. Diese stellen Mindeststandards dar, die im Ergebnis sowohl die Qualität der Metadaten bei dalra für den Bereich Bildungsforschung als auch beim VerbundFDB-Portal erhöhen werden. Der VerbundFDB rechnet in den ersten Monaten mit einem erhöhten Aufwand an manuellen Kontrollen, um zu prüfen, ob der Harvester wie gewünscht funktioniert und welche Daten in welcher Qualität über diesen Weg in das VerbundFDB-Portal übermittelt werden sowie um eventuelle Probleme identifizieren zu können.

Diskutiert wurde anschließend im Plenum, ob auch die Möglichkeit einer Kombination des direkten und indirekten Metadaten austauschs bestünde, da es sich für Datenzentren anbietet, abhängig von den Studien unterschiedliche Wege zu verfolgen. Diese Möglichkeit besteht grundsätzlich. Der VerbundFDB muss hierfür jedoch eine Dublettenkontrolle vorsehen. Vorgeschlagen wurde ein Service für Nutzer/innen: Diese könnten bei der dalra-Registrierung darüber informiert werden, in welche weiteren Portale (wie bspw. dasjenige des VerbundFDB) die Metadaten übermittelt werden. Dies umzusetzen liegt jedoch in der Verantwortung von dalra.

Im Anschluss an die Vorstellung der Ergebnisse der AG Metadaten stellte Lisa Pegelow (IQB) ein vom IQB entwickeltes Tool, die „**Metadatenregistry**“ (MDR), vor. Die MDR ermöglicht es, FDZ-Mitarbeiter/innen oder anderen Interessierten, dort Metadatensets und kontrollierte Vokabulare einzutragen und zu verwalten sowie sich über bestehende Metadatensets Anderer zu informieren. Im besten Fall verbessert dies die kooperative Zusammenarbeit und die einheitliche Verwendung von Metadaten. Die Nutzung der MDR würde damit zu einer stärkeren Harmonisierung im Bereich Metadaten beitragen. Die MDR stellt

jedoch nur die technische Plattform dar. Darüber hinaus ist es ggf. erforderlich, Redaktionsteams bzw. redaktionelle Kontrollen einzuführen, um „Wildwuchs“ zu vermeiden. Das Tool selbst bietet hierzu die Vergabe unterschiedlicher Rollen an.

Das Tool ist Open Source und damit frei verfügbar für alle Interessierten nach einer Registrierung unter der Adresse <https://mdr.iqb.hu-berlin.de>. Die jeweiligen Sets und kontrollierten Vokabularien der MDR können mit Systemen der FDZ verknüpft werden (Eigenprogrammierung erforderlich) und müssten dann nur an einer Stelle gepflegt werden. Das Produktivsystem der MDR ist fertiggestellt, etwaige Weiterentwicklungen werden bedarfs- und ressourcenabhängig durchgeführt. Der VerbundFDB lädt die Anwesenden dazu ein, das Tool für unterschiedliche Nutzungszwecke zu erproben.

## Bewertung

---

Im dritten Block des Tages wurde das Schwerpunktthema dieses Treffens „Auswahl und Bewertung von Datensätzen“ behandelt. **Reiner Mauer** erläutert in seinem Inputvortrag, dass diese Begriffe bereits eine lange Tradition in den Archiv- und Bibliothekswissenschaften haben. Für FDZ ist die Bewertung von Datensätzen erforderlich, um eine Auswahl aus Datenbeständen treffen zu können. Die Bewertung ist Grundlage für Kuratierungsentscheidungen, wie bspw. ob Daten archiviert, in welcher Form sie bearbeitet und wie langfristig sie archiviert werden. Gründe, die solche Entscheidungen erforderlich machen, sind der gezielte Aufbau eines Bestands, technische Beschränkungen, forschungsethische oder rechtliche Einschränkungen und begrenzte Ressourcen. Solche Entscheidungen sollten stets in einem transparenten und legitimen Prozess erfolgen.

**Maike Porzelt** stellte das Bewertungsschema vor, das im Rahmen der Eingangsprüfung durch den VerbundFDB verwendet wird. Im Rahmen der Eingangsprüfung erfolgt eine Bewertung der Archivierungsfähigkeit der Daten. Die Kriterien setzen sich zusammen aus technischen, dokumentarischen und rechtlichen Aspekten. Die Bewertung erfolgt durch das archivierende FDZ je Datenbestand, der archiviert werden soll. Wenngleich die Bewertung somit dezentral erfolgt, stimmen sich die beteiligten FDZ fortlaufend ab.

Die *Archivierungsfähigkeit* ist von der *Archivierungswürdigkeit* bzw. dem *Nachnutzungspotential* eines Datenbestands zu unterscheiden. Im Plenum wurde anschließend auch über die Bewertung von *Nachnutzungspotentialen* diskutiert. Hierbei erscheinen sowohl die Einschätzungen der Datengeber/innen als auch die Einschätzungen externer Gutachter/innen instruktiv. Es wurde darauf hingewiesen, dass auch Daten mit geringer Archivierungsfähigkeit ein besonders hohes *Nachnutzungspotential* aufweisen können.

Für die FDZ werden Prüfungen und Bewertungen der Datenbestände als wichtig angesehen, da die Kuratierungsleistungen der FDZ und die daraus resultierende erhöhte Datenqualität und -auffindbarkeit sich von den Angeboten fachübergreifender Repositorien unterscheiden.

Auf die Bestimmung des *Nachnutzungspotentials* ging **Claudia Neuendorf (IQB)** in ihrem Vortrag ausführlich ein. Durch das IQB wurde ein Bewertungsschema entwickelt, anhand dessen sich das *Nachnutzungspotential* von Datenbeständen für interne Kuratierungsentscheidungen bestimmen lässt. Das Schema besteht aus vornehmlich objektiven Kriterien (z. B. Größe der Stichprobe, geographischer und thematischer Scope), denen jeweils Bewertungspunkte zugewiesen sind. Je größer die Summe an Be-

wertungspunkten ist, die ein Datenbestand erhält, desto höher ist dessen Nachnutzungspotential. Das Schema beinhaltet – durch die unterschiedliche Anzahl Punkte, die den einzelnen Kriterien zugewiesen ist – eine Gewichtung. Das Schema soll zukünftig mit weiteren Forscher/innen diskutiert sowie auf seine Anwendbarkeit hin empirisch überprüft werden.

In der anschließenden Diskussion wurde angemerkt, dass mit einem derartigen Schema insbesondere die großen Large Scale Surveys, die thematisch breit angelegt sind, identifiziert würden. Relevant für FDZ wäre es aber auch, eine Differenzierung gerade bei kleinen Studien erreichen zu können. Daneben wird die vorgenommene Gewichtung kritisch diskutiert, die subjektive Faktoren enthält, die sich in der Praxis stark unterscheiden können. Verschiedene Vorschläge zur Erweiterung und Optimierung des Schemas wurden gemacht.

Im letzten Vortrag dieses Blocks stellte **Ute Hofstätter (DZHW)** die beim DZHW entworfenen Dokumentationsstandards vor, die sich auf Datensätze, Metadaten und Daten- und Methodenberichte beziehen. Die Standards beinhalten Mindestanforderungen an Merkmale, die Datenbestände, die beim DZHW archiviert werden, erfüllen sollten. Hierzu gehören beispielsweise, wie umfangreich methodische Informationen zu dem Datensatz vorliegen oder welche Arten von Bereinigungen und Aufbereitungen durchgeführt wurden.

In der nachfolgenden Diskussion wurde angemerkt, dass Vorgaben zur Art und Weise und Umfang der Dokumentation und Aufbereitung von Daten häufig von den Datengeber/innen selbst nachgefragt werden. Leitfäden für Datengeber werden daher als höchst sinnvoll eingeschätzt. Der VerbundFDB verfügt derzeit über eine Publikation zum Thema „Aufbereitung quantitativer Daten“, die bereits Hinweise hierzu enthält. Es wurde beschlossen, dass sich die mit quantitativen Daten arbeitenden Einrichtungen DZHW, IQB, GESIS austauschen und die jeweiligen Anforderungen untereinander abstimmen. Ggf. soll das bestehende Papier des VerbundFDB angepasst werden. Auch wurde beschlossen, dass sich die mit qualitativen Daten arbeitenden Einrichtungen (Qualiservice, FDZ Bildung) untereinander zu diesem Punkt abstimmen.

Des Weiteren wurde in der Diskussion sichtbar, dass manche FDZ die Vorgaben des DZHW für einen Mindeststandard als vergleichsweise hoch ansehen; die Praxis variiere stark. Einig sind sich die Anwesenden über die Bottom Line, dass Dritte den Datenbestand in einem angemessenen Zeitraum nachvollziehen können. Dies könne auch bei Daten gegeben sein, die weniger umfangreich dokumentiert sind. Demnach ist Nachvollziehbarkeit nicht gleichzusetzen mit Dokumentationsniveau, wenngleich ein hohes Dokumentationsniveau die Nachvollziehbarkeit erleichtert. Geringe Mindeststandards haben den Vorteil eines geringen Aufwands für die Forschenden; hohe Standards können hingegen zu einer hohen Reputation der FDZ und der dort archivierten Daten beitragen.

## Abschlussdiskussion

---

In der Abschlussdiskussion wurde in die Runde gefragt, ob und welche Bedarfe unter den Netzwerkpartnern bestehen, die im VerbundFDB bearbeitet und diskutiert werden sollten. Das ZPID ist an einem Austausch zu Kompetenzanforderungen von FDZ-Mitarbeiter/innen interessiert. Dieses Thema wird derzeit auch im FDI-Ausschuss des RatSWD (Arbeitsgruppe Kompetenzbildung) bearbeitet.

Die Bedeutung des Themas **Nachnutzungspotentiale** wurde nochmal betont. Zukünftig werden sich Forschende auf die Archivierungsverpflichtung einstellen, und es wird schwieriger für FDZ, Daten aufgrund von formalen Voraussetzungen abzulehnen. Gleichzeitig steigen die Aufnahmekapazitäten der FDZ, wenn die Dokumentationsstandards/-niveaus der zu archivierenden Daten steigen. FDZ werden zukünftig nicht alle Daten aufnehmen können.

Ausführlicher wurde über **Zertifizierungen** gesprochen, den Nutzen und den damit verbundenen Aufwand. Der VerbundFDB wird in den nächsten Wochen nochmal per E-Mail das Interesse an diesem Thema abfragen und ggf. zu einem gemeinsamen Termin zum Informationsaustausch einladen.

Für das nächste NT-Treffen wurden als Schwerpunktthemen das Thema **Nutzungsbedingungen und Zugangswege bei FDZ** sowie das Thema **Geschäftsmodelle** vorgeschlagen. Bezüglich Nutzungsbedingungen ist zu fragen, inwiefern sich die Zugangsbedingungen unterscheiden und ob eine Harmonisierung im Hinblick auf eine erhöhte Nutzerfreundlichkeit zu erreichen ist. Empfohlen wird eine enge Abstimmung mit RatSWD-Aktivitäten.

Abschließend wurde von den VerbundFDB-Vertreter/innen betont, dass der VerbundFDB eine enge Zusammenarbeit mit den FDZ begrüßt. Durch den Austausch im Netzwerk auch auf operativer Ebene wird das Netzwerk als sinnvolle Ergänzung zu den Arbeiten auf Ebene des RatSWD-FDI-Ausschusses angesehen. Dies stieß auf Zustimmung bei den anwesenden FDZ-Vertreter/innen.

Angekündigt wurde ein **nächstes Netzwerktreffen im Jahr 2019**, das wieder zeitnah zur RatSWD-FDI-Ausschusssitzung im Frühjahr stattfinden wird.