

# Bericht vom ersten Netzwerktreffen des Verbunds Forschungsdaten Bildung

Karoline Harzenetter // Mai 2017

## Einleitung

---

Der vom BMBF geförderte Verbund Forschungsdaten Bildung (VFDB) [www.forschungsdaten-bildung.de](http://www.forschungsdaten-bildung.de) hat es sich zum Ziel gesetzt, gemeinsam mit relevanten Akteurinnen und Akteuren aus Wissenschaft, Politik und Forschungsförderung eine leistungsfähige Forschungsdateninfrastruktur für die empirische Bildungsforschung aufzubauen. Zu diesem Zweck luden die drei Kernpartner des Verbunds (GESIS, IQB und DIPF) Vertreterinnen und Vertreter der Forschungsdatenzentren (FDZ<sup>1</sup>) kooperierender Institute (Apaek, BIBB, DIE, DIW Berlin, DJI, DZHW, IAB, IIfBi, Qualiservice Bremen und ZPID) und die Leitung der Geschäftsstelle des RatSWD zum ersten Netzwerktreffen mit dem Thema *Harmonisierung und Bündelung von Metadaten in der Bildungsforschung* in die Räumlichkeiten der Humboldt-Universität zu Berlin ein.

Aufgabe des VFDB ist es, die spezifischen Bedarfe der Scientific Communities adäquat zu bedienen und die empirische Bildungsforschung dabei zu unterstützen, exzellente Forschung durchzuführen. Der Verbund erbringt dabei Services der Archivierung und Bereitstellung von Forschungsdaten und unterstützt Forschende durch Beratungs- und Schulungsangebote zum Forschungsdatenmanagement. Das Projekt befindet sich seit Oktober 2016 in der zweiten Förderphase. Insgesamt sind während der dreijährigen Laufzeit drei Netzwerktreffen geplant, die den Aufbau und die Gestaltung einer leistungsfähigen Infrastruktur zum Gegenstand haben.

Das Auftakttreffen am 29. März 2017 in Berlin diente einerseits dazu, das Interesse und die Möglichkeiten der Zusammenarbeit im Bereich Dokumentation und Archivierung von Forschungsdaten aus dem Bereich Bildung seitens der teilnehmenden Einrichtungen zu ermitteln; andererseits wurden die bisher erreichten Ziele des Projekts vorgestellt. Darüber hinaus wurden Möglichkeiten und Formen der Zusammenarbeit der FDZ im Bereich Harmonisierung und Interoperabilität studienbeschreibender Informationen (Metadaten) diskutiert und die Gründung einer Arbeitsgruppe zu diesem Thema beschlossen.

---

<sup>1</sup> Im Textverlauf wird die Abkürzung *FDZ* als Sammelbegriff stellvertretenden für die vom RatSWD akkreditierten Forschungsdatenzentren und kooperierender Institute verwendet

## Verlauf und Inhalte

---

### 10:00 - 10:30: Begrüßung, Vorstellungsrunde und Klärung der Zielstellung

Nach der Begrüßung durch die Projektleitung Malte Jansen, Doris Bambey und Reiner Mauer wurden der Hintergrund und die Ziele der Veranstaltung vorgestellt. Der seit dreieinhalb Jahren existierende Verbund strebt an, seine Aktivitäten im Bereich Archivierung und Dokumentation bildungsrelevanter Forschungsdaten zu erweitern, indem er u. a. Partner aus dem Kreise der eingeladenen FDZ für verschiedene, abzustimmende Formen der Zusammenarbeit gewinnen möchte. Ziel des ersten Treffens war es einerseits zu erörtern, inwiefern die FDZ interessiert sind, sich als Verbundpartner am Aufbau dieser föderativen Infrastruktur für die Bildungsforschung zu beteiligen und in welcher Form dies für die Beteiligten überhaupt möglich wäre. Dabei ist das Thema Metadaten bzw. im hier vorliegenden Kontext die studienbeschreibenden Informationen (studienbeschreibende Metadaten) als zentrales Schlüsselthema zu verstehen, da es einen Querschnittsbereich über alle FDZ hinweg darstellt. Als weiteres Ziel des Netzwerktreffens galt es, Ideen für mögliche weitere Kooperationen in anderen Bereichen des Forschungsdatenmanagements und der Infrastrukturentwicklung zu sammeln und die Umsetzung gemeinsamer Ziele zu diskutieren. Im Anschluss an die Eröffnung und Begrüßung durch die Projektleitenden, folgte eine knappe Selbstvorstellung aller Anwesenden.

### 10:30 - 11:30: Der Verbund Forschungsdaten Bildung stellt sich vor

Die Aufgaben, Ziele und der aktuelle Stand des Verbund-Projekts wurden von Alexia Meyermann (DIPF) vorgestellt. Im Anschluss an die Vorstellung des Verbundes ergab sich eine Diskussion, in deren Verlauf folgende Fragen erörtert wurden:

- Ist eine gemeinsame Entwicklung von Archivierungs- und Dokumentationsstandards sinnvoll? Auf welcher Basis können gemeinsam Kriterien und Kategorien zur Bewertung des Nachnutzungspotenzials entwickelt werden?
- Welche strategische Ausrichtung hat der VFDB? Welche Rollen möchten und können die FDZ in Zukunft im Verbund wahrnehmen?
- Übernimmt der VFDB Aufgaben der Verknüpfung von Forschungsdaten innerhalb dieser Infrastruktur?
- Wie kann der VFDB die Interoperabilität von Metadaten gewährleisten und das VFDB-Suchportal weiter verbessert werden?

*Ist eine gemeinsame Entwicklung von Archivierungs- und Dokumentationsstandards sinnvoll? Auf welcher Basis können gemeinsame Kriterien und Kategorien zur Bewertung des Nachnutzungspotenzials entwickelt werden?*

FDZ sind mit immer größeren Mengen an sehr heterogenen Daten, z. B. aus kleinen universitären Projekten, konfrontiert, die Auflagen der Fördereinrichtungen zur Sicherung und Aufbewahrung erfüllen müssen. Die Art und Menge der Daten führen bei den FDZ zu ökonomischen bzw. Effizienzproblemen, so dass – aktuell schon bei einigen und zukünftig ggf. bei allen FDZ – nicht alle Forschungsdaten aufgenommen werden können, deren Archivierung seitens der Datengebenden angefragt wird. Die FDZ sehen sich meistens mit einem Kapazitätsproblem konfrontiert, das eher in Ausnahmefällen das Datenvolumen bzw. fehlende Speicherkapazitäten betrifft. Vielmehr entstehen durch zeitaufwändige und ressourcenin-

tensive Bearbeitungsprozesse der vielfältigen Projektdaten hohe Kosten. Da nicht alle Forschungsdaten von den FDZ als nachnutzbar eingeschätzt werden, wurde der Bedarf nach einheitlichen Selektionskriterien, die zur Priorisierung von Daten herangezogen und mit Datengebenden und Fördereinrichtungen kommuniziert werden können, geäußert. Dabei stellte sich jedoch die Frage, anhand welcher Kriterien das Nachnutzungspotenzial von Forschungsdaten transparent und objektiv ermittelt werden könnte, und ob es darüber hinaus Forschungsdaten gibt, die unabhängig vom Nachnutzungspotenzial archiviert werden könnten und sollten, indem z. B. miteinander abzugleichende (Re-)Finanzierungsoptionen zugrunde gelegt werden, in denen solche Archivierungsservices bepreist werden um zusätzliche Ressourcen schaffen zu können.

Einige der FDZ wenden bereits Verfahren zur Selektion an. Um dem Kapazitätsproblem zu begegnen, hat das DZHW drei Archivierungsqualitätsstandards entwickelt, die jeweils einen unterschiedlichen Ressourcenaufwand beinhalten. Ob und mit welchem Aufwand Forschungsdaten überhaupt archiviert werden, wird anhand von 80 Kriterien ermittelt, wovon zehn Minimalkriterien mindestens erfüllt sein müssen. Im Anwendungsfalle des FDZ Psychdata werden nur Forschungsdaten zur Archivierung ausgewählt, die in Verbindung mit einer Publikation in einer einschlägigen fachspezifischen Zeitschrift stehen.

Des Weiteren wurden alternative potenzielle Verfahren zur Selektion diskutiert. In Anlehnung an das Modell des DZHW könnte ein solches Verfahren die stufenweise Archivierung in Abhängigkeit der messbaren Nachfrage nach Forschungsdaten einer Studie sein. Hierbei würden zunächst unabhängig von der Nachfrage alle zur Archivierung angefragten Daten eine niedrighschwellige Archivierungsstufe durchlaufen, welche nur einen geringen Ressourcenaufwand benötigt. Dies bedeutet, dass die Forschungsdaten erst bei steigender Nachfrage weitere Stufen und Tiefen der Dokumentation durch die FDZ durchlaufen würden.

Ein anderes mögliches Instrument zur prospektiven Ermittlung des Nachnutzungspotenzials könnte die geschätzte Nachfrage durch die beteiligten Forschenden selbst sein. Nach diesem Modell müssten die Forschenden Aussagen darüber treffen, wie wichtig bzw. speziell sie die eigene Forschungsfrage einschätzen und für wie groß sie das Potenzial zur Nachnutzung ihrer Forschungsdaten innerhalb ihrer Forschungsgemeinschaft halten. Denkbar wären auch Rezensionen zu Datensätzen durch Expertinnen und Experten, die ähnlich einem Reviewverfahren prüfen, ob inhaltliche und methodische Voraussetzungen guter wissenschaftlicher Praxis erfüllt sind. Diese Expertise könnte dann als Grundlage für die Entscheidung zur Archivierung herangezogen werden.

Die potenzielle Nachfrage der Forschungsdaten wird jedoch als kritisches Merkmal gesehen, da sie schwer vorhersagbar und messbar ist. Historischer Kontext und Forschungstrends beeinflussen diese maßgeblich. Daten werden mitunter erst im Laufe der Zeit interessant(er) und nachgefragt(er). Um Nachfrage an Forschungsdaten vergleichbar messen zu können, müssten zunächst gleiche Voraussetzungen der Auffindbarkeit, Sichtbarkeit und des Zugangs geschaffen werden. Ein Beispiel für die zeitversetzte Wirkung von Forschungstrends, wäre die Nachfrage nach Daten der Unterrichtsforschung in der DDR (Beispiel des FDZ Bildung) oder Daten zum Thema Migration (Beispiel von GESIS) - Themen, die erst nach Projektende oder Erhebung verstärkt in den Fokus der Öffentlichkeit und der Forschung gerückt sind. Das Nachnutzungspotenzial von Daten in Form der Nachfrage kann also nicht immer bereits zum Zeitpunkt der Archivierung vorausgesehen werden. Ebenso ist es als kritisch anzusehen, das Nachnutzungspotenzial einer Studie nur anhand der Nachfrage als einziges ausschlaggebendes Kriterium

heranzuziehen. Die Erfahrungen von GESIS zeigen, dass sich dort ca. 80-90% aller Sekundärnutzungen auf ca. 10% der archivierten Forschungsdaten konzentrieren. Gesteigertes Interesse der Nutzenden besteht erfahrungsgemäß an sogenannten „large scale studies“, wie es z. B. die „Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS)“ auf nationaler oder die „European Values Study (EVS)“ auf internationaler Ebene sowie Daten der PISA-Studie („Programme for International Student Assessment“) auf nationaler wie internationaler Ebene sind. Würde man sich bei der Frage der Archivierung also nur nach der potenziellen Nachfrage richten, käme es wohl zu einer starken Konzentration auf großangelegte Studien und damit zu einer Vernachlässigung kleinerer Projekte. Diese Benachteiligung kleinerer Projekte beträfe auch das BMBF-Rahmenprogramm zur Förderung der empirischen Bildungsforschung sowie die Archivierungsaufgabe des VFDB. Denn auch in der zweiten Phase des EBF-Rahmenprogramms ist mit weiteren Forschungsdaten zu rechnen, deren Archivierung unabhängig von ihrer Nachfrage und Größe erfolgen soll. Darüber hinaus sollen sich auch noch andere Datengebende, die ebenso bestimmte Auflagen zur Archivierung und Nachnutzung anderer Fördereinrichtungen erfüllen müssen, in Zukunft an die FDZ und den VFDB wenden können.

Einigkeit bestand darin, dass es sich bei den Selektionskriterien um formale und transparente und nicht um normative Eigenschaften handeln sollte und eine Basisversorgung grundsätzlich zu gewährleisten ist, um möglichst flächendeckend Forschungsdaten zu sichern und gleiche Voraussetzungen für Auffindbarkeit, Sichtbarkeit und Zugang zu schaffen.

*Welche strategische Ausrichtung hat der VFDB? Welche Rollen möchten und können die FDZ in Zukunft im VFDB wahrnehmen?*

» Schulung und Beratung im Forschungsdatenmanagement

Datenmanagement setzt bereits früh im Verlauf eines Forschungsprojekts an und liegt in der Verantwortung der Forschenden. Die Gewährleistung und Vorbereitung der Dokumentation durch die Primärforscherinnen und -forscher zur Sicherstellung der Nachvollziehbarkeit von Forschungsdaten für Nachnutzende verbessert Datenqualität und verringert Kosten der Kuratierung für die FDZ. Forschende müssen demnach in ihren Fähigkeiten des Datenmanagements ausgebildet werden, damit sie bereits qualitativ hochwertige Daten an die FDZ liefern können. Diese „Anschlussfähigkeit“ von Forschungsdaten ist primär Aufgabe der Forschenden selbst. Der VFDB leistet bereits einen wichtigen Beitrag zur Verbesserung der Aufbereitungs- und Dokumentationsqualität in der Bildungsforschung, indem den Forschenden Forschungsdatenmanagementkompetenz und -verantwortung für die eigenen Daten durch Schulungsangebote und Beratung vermittelt werden. Der VFDB macht es sich außerdem zur Aufgabe, Datengebende und Forschende der empirischen Bildungsforschung zu betreuen, auch wenn ihre Forschungsdaten nicht über den VFDB archiviert und dokumentiert wurden oder werden. Stand anfangs das Schulungsangebot nur für Forschende aus dem EBF-Rahmenprogramm offen, richtet sich das Angebot in der zweiten Förderphase generell an alle empirisch Forschenden im Bildungsbereich.

» Zentrales Nachweissystem für die Bildungsforschung

Der Einzelstudiennachweis/Datennachweis in der VFDB-Studiensuche stellt eine notwendige Bedingung im Bereich der dokumentarischen Erschließung und „Anschlussfähigkeit“ durch den VFDB für Forschungsdaten dar, um deren Auffindbarkeit und Sichtbarkeit zu garantieren und einen fairen zentralen Zugang zu diesen Daten zu ermöglichen. In diesem Zusammenhang können Standardisierung und Har-

monisierung der studienbeschreibenden Informationen (Metadaten) die Metadatenqualität und Darstellung im VFDB-Nachweissystem verbessern und einen wichtigen Beitrag dazu leisten. Die Vereinheitlichung der Standards und Harmonisierung wird über die Kernpartner des VFDB hinaus angestrebt.

» **Gemeinsames Metadaten-Applikationsprofil für Daten der Bildungsforschung**

Der VFDB sieht es als seine Aufgabe an, die Interoperabilität der studienbeschreibenden Informationen durch die Entwicklung eines gemeinsamen, für die Bildungsforschung geeigneten, Applikationsprofils sicherzustellen. Ein solches gemeinsames Profil zur adäquaten Beschreibung qualitativer und quantitativer Daten soll die Weiterverwendung und den Austausch von Metadaten innerhalb des VFDB ermöglichen und erleichtern.

» **Ausbau föderativer Forschungsdateninfrastruktur durch Kooperation**

Eine Erweiterung der Kooperation ist aus der Sicht des VFDB sinnvoll, wenn z. B. Datensätze in einem der FDZ des erweiterten Netzwerks inhaltlich und methodisch besser aufgehoben sind. Mögliche Formen der Zusammenarbeit bestehen beispielsweise in der Beteiligung von FDZ an der verteilten Archivierung und an der Erweiterung der bereits im VFDB implementierten Archivierungsabläufe. Grundsätzlich sind vielfältige Formen der Kooperation und Grade des Einsatzes denkbar. Der VFDB könnte die Aufgabe einer Clearingstelle erfüllen, Beratungs-, Dokumentations- und Archivierungsarbeiten an die prädestinierten kooperierenden FDZ verteilen. Mögliche verteilte Arbeiten sollen dabei unter Berücksichtigung spezifischer Expertise, Kapazitäten und bilateraler Verhandlungen sondiert und vereinbart werden. Kooperationsform und -bereitschaft könnten idealerweise in Kooperationsverträgen festgeschrieben werden. Der VFDB soll einen Verteilmechanismus für Dokumentation und Archivierung innerhalb des Verbundes entwickeln, der auch Zusammenarbeit in Form eines Staffelkonzepts beinhalten könnte. Das heißt, dass Arbeiten der Archivierung und Dokumentation einer Studie auf mehrere FDZ verteilt werden und in Phasen ablaufen.

» *Übernimmt der VFDB Aufgaben der Verknüpfung von Forschungsdaten innerhalb dieser Infrastruktur?*

Der VFDB versteht Data Linking im Sinne der Verknüpfung von Forschungsdaten aus unterschiedlichen Datenquellen bisher nicht als seine Aufgabe, da die Verknüpfung von Forschungsdaten unterschiedlicher Studien meist als Folge sehr spezieller Anfragen und Forschungsfragen von Datennutzenden zu verstehen ist. Der VFDB bietet diese Art der Informationsverknüpfung derzeit nicht gezielt an. Der VFDB schafft jedoch Voraussetzungen für die Verknüpfung von Daten, indem Datenbestände offengelegt und gut dokumentiert werden. Er führt außerdem, durch die Abbildung verbundener Forschungsdaten und Publikationen und deren Verlinkung, Informationen sichtbar zusammen. Bestehende Relationen von Daten einer Studie bei verteilter Archivierung und Dokumentation durch die Kernpartner werden ebenfalls dargestellt, wie z. B. die Relationen qualitativer und quantitativer Daten einer Studie bei einer Kombination aus Umfrage-, Leistungs- und Videodaten innerhalb eines Forschungsprojekts. In der Zukunft nicht auszuschließen ist in Einzelfällen die Darstellung verknüpfter Daten, z. B. von Befragungsdaten und administrativen Daten, wobei dann von Fall zu Fall entschieden und festgelegt werden muss, welches Datenzentrum für die Verknüpfung und die verknüpften Daten selbst sowie für die Sicherstellung der Abbildung der Relationen zuständig ist.

*Wie kann der VFDB die Interoperabilität von Metadaten gewährleisten und das VFDB-Suchportal weiter verbessert werden?*

Das VFDB-Suchportal hat den Anspruch, als Wegweiser für das Nachweissystem zur Auffindbarkeit Studien nationaler Bildungsforschung zu dienen. Die Studienanzeige des VFDB soll in Kombination mit dem Suchinstrument einen flächendeckenden Überblick über Daten der Bildungsforschung ermöglichen und Forscherinnen und Forscher an die für sie in Frage kommenden FDZ weiterleiten. Ziel ist es, insbesondere die bei den verschiedenen FDZ bereits archivierten Bildungsforschungsdaten an zentraler Stelle nachzuweisen, durchsuchbar zu machen und durch Verlinkung zur datenhaltenden Einrichtung Daten auffindbar zu machen. Die VFDB-Suche als spezifische Suche über Daten der empirischen Bildungsforschung versteht sich hier als Ergänzung zu anderen Recherchesystemen mit anderen Fokussen und Adressatenkreisen.

Aktuell fällt es Externen schwer, die VFDB-Suche, die auf Studienebene ansetzt, von den auf der Website verlinkten und integrierten Suchsystemen von DIPF, GESIS und IQB auf Datensatzebene zu unterscheiden. Die VFDB-Websuche ist für die Suche in den Datensatzbeschreibungen der VFDB-Datenbank noch nicht konzipiert. Die bisher auf Studien fokussierte Suche wird optimiert, auch hinsichtlich einer möglichen Erweiterung der Suche auf Ebene der Datensätze. Die Durchsuchbarkeit auf Datensatzebene soll zunächst konzeptionell erarbeitet werden, um dann über deren Umsetzung entscheiden zu können. Eine weitere Lücke der Websuche wird in der Abbildung qualitativer Daten sichtbar, da die Suchmaske noch nicht allen Datenarten gerecht wird. Dies ist ebenfalls dem bisherigen Fokus der Suche auf die Studienebene geschuldet. Neben der Art der Daten und der Beschreibungsebene, schränkt auch generell die Qualität und Heterogenität der Studienbeschreibungen die Weiterentwicklung der VFDB-Suche hinsichtlich Facettierung, Filterung und Darstellung stark ein und beeinträchtigt so nicht nur die Darstellung und Auffindbarkeit in der VFDB-Websuche, sondern auch in anderen Suchsystemen.

Ziel des VFDB ist es, die Metadaten im Sinne der Interoperabilität technisch nutzbar zu machen sowie deren Qualität zu vereinheitlichen und anzuheben. Ziel dabei ist es, einen problemlosen Fluss der Metadaten durch alle Informationssysteme zu gewährleisten. Insbesondere sollen die Metadaten in den VFDB-Suchindex mit gleicher Qualität integriert werden können, um eine nutzerfreundlichere Durchsuchbarkeit der Datenbank zu ermöglichen. Metadaten müssen hierfür gut strukturiert, standardisiert, harmonisiert und aufbereitet sein, damit sie technisch unterstützt und korrekt weiterverarbeitet werden können. Die Verwendung kontrollierter Vokabulare wie z. B. der Einsatz fachspezifischer Thesauri bei der inhaltlichen Erschließung, kann maßgeblich die Qualität von Metadaten und Suchmaschinen verbessern und so die Auffindbarkeit erhöhen. Der Großteil der von den FDZ des RatSWD betreuten Studien wird bereits bei dalra registriert und nachgewiesen (dalra-Suche, gesisDataSearch). Dadurch könnte die Infrastruktur von dalra als zentrale Quelle für einheitlich strukturierte Metadaten bereits genutzt werden. Die Identifikation der Nachweise in der dalra-Datenbank müsste jedoch ermöglicht werden. Gegebenenfalls müssten relevante Bestände durch die registrierenden FDZ für eine erleichterte Integration der Metadaten in den VFDB-Index gekennzeichnet werden. Die stärkere Abstimmung/Vereinheitlichung mit dalra wurde von den FDZ generell als sinnvoll und notwendig eingeschätzt. Unklar ist jedoch, in welcher Form dalra als Quelle von Metadaten für den VFDB-Index genutzt werden soll. Insbesondere vor dem Hintergrund, dass dalra in erster Linie eine Registrierungsagentur und keine Dokumentationstool ist, stellt sich die Frage, in welchem Ausmaß sich VFDB und die FDZ von deren Entwicklungen abhängig machen sollten.

Eine weitere Herausforderung ist es, Forschende an eine Suche über die VFDB-Website [www.forschungsdaten-bildung.de](http://www.forschungsdaten-bildung.de) heranzuführen. Ein Teil der Forschenden wird zwar als „Datenmelde“ und durch Beratungsangebote im Forschungsdatenmanagement bereits durch den VFDB unterstützt und findet so bereits den Einstieg zur Recherche über die VFDB-Suche, die Empirie zeigt jedoch, dass spezielle Suchportale von der breiten Masse der Forschenden nicht genutzt werden. Der Großteil der Forschenden nutzt nach wie vor die Google-Suche. Die Google-Suche könnte als zusätzlicher Einstieg zu den Daten des VFDB dienen, denn große Aggregatoren nutzen Metadaten einzelner Websites bereits für ihre eigenen Dienste. Ein Ziel könnte es sein, dass der Verbund im Sinne von Search Engine Optimization sich darum bemüht, Google Crawling und Indexierung der VFDB-Inhalte zu ermöglichen und zu erleichtern, um so Zugriff und Auffindbarkeit der vom VFDB nachgewiesenen Daten zu erhöhen. Andreas Daniel vom DZHW bot an, Erfahrungen mit dem Aggregator Google einzubringen. Nutzerfreundliche Gestaltung und Bedienbarkeit sollten bei der VFDB-Suche ebenfalls zukünftig im Vordergrund stehen. Auch hier könnte der VFDB von Google-Entwicklungen lernen. Usability-Tests/Nutzerbefragungen könnten zudem dabei helfen, die Nutzerperspektive bei Weiterentwicklungen der VFDB-Website stärker zu berücksichtigen und so die VFDB-Services zusätzlich zu verbessern.

#### **11:45 - 12:30 und 13:15 - 14:00: Studiendokumentationspraxis und Metadatenstandards der FDZ (Teil I und Teil II)**

In dieser und der nachfolgenden Session stellten die Vertreterinnen und Vertreter die Arbeit in ihren FDZ vor. Die Vortragenden orientierten sich hierfür an vier Leitfragen, die vorab vom VFDB vorgeschlagen wurden:

- Welche Arten von Forschungsdaten werden von Ihrem Institut erhoben, archiviert und zur Nachnutzung zur Verfügung gestellt?
- Gibt es standardisierte Dokumentationsabläufe? Verwenden Sie zum Beispiel Standards und Referenzsysteme zur Dokumentation Ihrer Studien (Metadatenschema von DDI, dalra, Dublin Core usw., kontrollierte Vokabulare, Persistente Identifier, systematische Verschlagwortung/Indexierung usw.)?
- Auf welcher Ebene beschreiben Sie Ihre Forschungsdaten (Projekt-, Studien-, Datensatz-, Variablenebene), gibt es ein Versionierungskonzept und sind Ihre Forschungsdaten eindeutig identifizierbar (anhand eines persistenten Identifikators wie z. B. DOI-System)?
- Sind Studiendokumentation und Datenaufbereitungsarbeiten technisch unterstützt? Benutzen sie z. B. bestimmte Tools, Anwendungen oder Programme zur Dokumentation Ihrer Studien?

#### **14:00 - 15:00: Metadaten und Studiendokumentation im VFDB-Netzwerk**

Karoline Harzenetter (GESIS) zeigte die Bedeutung von Metadaten, ihre prozesshafte Entwicklung und die Formen der Nachnutzung zur Verbesserung der Auffindbarkeit und Identifikation von Forschungsdaten auf. Danach stellte Thomas Schwager (DIPF) das von den drei Partnern entwickelte VFDB-Metadatenset und die kurative Arbeit des VFDB zur Qualitätssteigerung und -sicherung der Studiendokumentation vor.

Im Anschluss an die zwei Präsentationen entstand eine Diskussion, in der insbesondere der Metadaten-austausch zwischen dem VFDB, dalra und den einzelnen FDZ thematisiert wurde. Darüber hinaus wurde das Ausschöpfungspotenzial bereits angewandter Standards im Partnerkreis diskutiert.

**15:15 - 16:00: Abschlussdiskussion**

In der Abschlussdiskussion ging es unter Berücksichtigung der Bedürfnisse der Netzwerkpartner um die Formulierung möglicher gemeinsamer Ziele sowie ihren Bedarf an weiterem Austausch und der Zusammenarbeit innerhalb des Netzwerks. Zunächst wurden aber eine indirekte und eine direkte Variante des Austauschs studienbeschreibender, kuratierter und autorisierter Metadaten zwischen VFDB und FDZ diskutiert: bei der indirekten Variante würden die Metadaten in die VFDB Datenbank über die Registrierungsagentur dalra aufgenommen werden; bei der direkten Variante würden die Metadaten direkt durch die FDZ selbst dem VFDB zugänglich gemacht. Für das Management studienbeschreibender Metadaten und eine technisch unterstützte Aufnahme in den Index der VFDB-Suche könnte zumindest in der indirekten Variante auf bereits bestehende technische Entwicklungen von dalra zurückgegriffen werden, um die Weiterverwendung und den Austausch der Metadaten innerhalb des VFDB zu organisieren.

*Modelle der Metadatenaufnahme in das VFDB-Nachweissystem*

Der Aufwand, Forschungsdaten in den Nachweis aufzunehmen, sollte bei beiden Varianten gering sein, da an mehreren Stellen kaum Kapazitäten für die (v. a. rückwirkende) Pflege von Metadaten vorhanden sind. Es gilt demnach zu klären, welches Maß an Anreicherung und Pflege der Metadaten bei den unterschiedlichen Modellen überhaupt von wem umsetzbar ist und umzusetzen wäre. Es kam auch die Frage auf, wer wie bestimmt, welche Daten der Bildungsforschung zuzurechnen sind und in den Scope des VFDB fallen.

Wesentlicher Pluspunkt der Variante, den VFDB als zentrale Stelle zu nutzen und die Metadaten direkt von den FDZ in die VFDB-Datenbank einpflegen zu lassen, sind zum einen die Nutzerfreundlichkeit, da in einem Index mit höherer Dokumentationstiefe gesucht werden könnte, und zum anderen die Bündelung von Metadaten bildungsrelevanter Forschungsdaten, da die FDZ auf ein gemeinsames/zentrales Nachweissystem zugreifen könnten. VFDB und FDZ sind bei dieser Variante völlig unabhängig von Entwicklungen bei dalra. Mit jedem FDZ könnte eine individuelle Schnittstelle definiert werden, sodass die Daten aus dem FDZ ausgelesen würden. Außerdem spricht gegen die indirekte Lösung über dalra, dass nicht alle für die Bildungsforschung relevanten Datenbestände bisher dort registriert werden. Darüber hinaus gibt es in dalra derzeit Lücken bei der Repräsentanz qualitativer Daten und Instrumente. Standardisierung der Dokumentation wäre insofern nötig, da jede datenmeldende Stelle aktuell ihre registrierten Bestände in unterschiedlicher Tiefe bei der Registrierung abbildet.

Die Umsetzbarkeit der direkten Lösung wird jedoch als schwierig angesehen, da Aufwand- und Aktualisierungsprobleme entstehen. Die Umsetzung FDZ-spezifischer Schnittstellen kostet die FDZ und den VFDB Ressourcen, und händische Prozesse sind kaum praktikabel. Vorgeschlagen wurde deshalb, dass dalra trotz ihrer bisher vorrangigen Funktion als Registrierungs- und nicht als Dokumentationstool mit ihrem Kernset an Metadaten den momentan vorrangig zu prüfenden Weg für den Austausch von Metadaten zwischen FDZ und VFDB darstellt, da hier eine bereits technisch skalierende Infrastruktur vorliegt (z. B. offene OAI-PMH-Schnittstellen). Bei dieser indirekten Variante wird es als Aufgabe des VFDB und der Partner gesehen, die Qualität der Metadaten und die Weiterentwicklung des dalra-Metadatenschemas und Mappings zu nutzen und zu beeinflussen und dafür zu sorgen, dass die für die Bildungsforschung relevanten registrierten Bestände bei dalra von den FDZ zur Identifikation durch den VFDB gekennzeichnet werden können.



### Weitere Vorschläge

Es kam außerdem der Vorschlag auf, die Vorstellung darüber, wer Datengebende für den VFDB sein könnten, zu erweitern, denn auch die FDZ könnten selbst als Datengebende angesehen werden, für die der Verbund Serviceleistungen erbringt. Eine starke Berücksichtigung der Bedürfnisse von Forschenden und deren Einbindungen in die Arbeiten des VFDB, beispielsweise bei der Weiterentwicklung der Websuche oder bei der Selektion von Forschungsdaten, wird zudem als sinnvoll erachtet. Befragungen im Vorfeld und die Einbindung in Prozesse wären denkbar, um nicht am Bedarf dieses Nutzerkreises vorbei zu gehen, da dieser unter Umständen stark von denen der Datenzentren abweicht. So wäre es zum Beispiel sinnvoll, Entscheidungen zu Standardisierungen im erweiterten Partnerkreis zu treffen und erst nach Berücksichtigung des Feedbacks der Forschenden (Forschende auf anderer Fachebene) umzusetzen.

### Fazit

---

Operatives Ergebnis des Treffens war die Gründung einer Arbeitsgruppe unter Beteiligung von Andreas Daniel (DZHW), Dr. Jan Goebel (DIW Berlin), Dr. Daniel Fuß (LifBi), Thomas Schwager (DIPF), Karoline Harzenetter und Reiner Mauer (GESIS) und Lisa Pegelow (IQB) zur Thematik Metadatenharmonisierung. Der Verbund wird die hierfür notwendigen Vorleistungen in die AG einbringen. Ziel dieser Arbeitsgruppe ist es, (1) die mit den beiden Modellen der Nachweiserweiterung einhergehenden Konsequenzen wie z. B. zusätzlicher Aufwand in der Metadatenkuratierung oder Doppelerfassung von Forschungsdaten detailliert zu beschreiben, (2) die Möglichkeit, die dalra-Infrastruktur für den Metadatenaustausch zu nutzen und zu evaluieren sowie (3) ein gemeinsames für die Bildungsforschung geeignetes Applikationsprofil zu entwickeln, das vom gesamten VFDB-Partnerkreis genutzt werden kann. Geplant ist ein Treffen der Gruppe im Herbst 2017.

Als mögliche zukünftige Themen der nächsten Netzwerktreffen wurden die VFDB-Suche und ihre Usability sowie mögliche Kriterien für Auswahl und Bewertung von Forschungsdaten erwähnt.

### Anwesende

---

#### **Anwesende Vertreterinnen und Vertreter der Datenzentren (in alphabetischer Reihenfolge):**

Noemi Betancort-Cabrera, QualiService der Universität Bremen, Bremen

Andreas Daniel, Deutsches Zentrum für Hochschul- Wissenschaftsforschung (DZHW), Hannover

Anett Friedrich, Bundesinstitut für Berufsbildung (BIBB), Bonn

Dr. Daniel Fuß, Leibniz-Institut für Bildungsverläufe e.V. (LifBi), Bamberg

Dr. Marius Gerecht, DIPF, Frankfurt a.M.

Dr. Jan Goebel, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

Harald Kaluza, Deutsches Institut für Erwachsenenbildung (DIE), Bonn

Helge Kminek, Archiv für pädagogische Kasuistik (Apaek), Frankfurt a.M.

Holger Quellenberg, Deutsches Jugendinstitut (DJI), München

Alexandra Schmucker, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg

PD Dr. Erich Weichselgartner, Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID), Trier

**Anwesende Vertreterin der RatSWD Geschäftsstelle:**

Claudia Oellers, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin

**Anwesende des Verbunds Forschungsdaten Bildung:**

Dr. Doris Bambey, Alexia Meyermann, Thomas Schwager, Deutsche Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt a.M.

Karoline Harzenetter, Reiner Mauer, Maike Metz-Peeters, Leibniz-Institut für Sozialwissenschaften (GESIS), Köln

Dr. Malte Jansen, Claudia Neuendorf, Lisa Pegelow, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Berlin