

Verbund Forschungsdaten Bildung – Aufbau und Gestaltung einer Forschungsdateninfrastruktur für die Bildungsforschung (10/2016 bis 09/2019)

Beteiligte Institute: DIPF, GESIS, IQB

Um die Multidisziplinarität und Größe der bundesdeutschen empirischen Bildungsforschung adressieren zu können, bedarf es einer vernetzten Forschungsdateninfrastruktur, die verteilt vorliegende Angebote unterschiedlicher Forschungsdatenzentren bündelt und Forschenden Services bedarfsorientiert und nutzerfreundlich bereitstellt. Der *Verbund Forschungsdaten Bildung* (VFDB) hat es sich daher zum Ziel gesetzt, gemeinsam mit relevanten Akteuren aus Wissenschaft, Politik und Forschungsförderung eine leistungsfähige Infrastruktur für die empirische Bildungsforschung aufzubauen. Diese soll die spezifischen Bedarfe der Scientific Communities adäquat bedienen und die empirische Bildungsforschung dabei unterstützen, exzellente Forschung durchführen zu können.

Hierzu planen DIPF, GESIS und IQB im Verbund mit anderen Forschungsdatenzentren (FDZ) in der zweiten Förderphase (2016-2019), erstens Services der Archivierung und Bereitstellung von Forschungsdaten anzubieten, zweitens das in der Pilotphase entwickelte Online-Portal mit Suchfunktion über Studien und Datenbestände der empirischen Bildungsforschung auszubauen und drittens Forscher/innen bei der Datenproduktion durch begleitende Schulungs- und Beratungsangebote zum Forschungsdatenmanagement zu unterstützen. In seiner Arbeit unterstützt der VFDB eine Vision von *Open Data*, in der die datenschutzrechtlichen und ethischen Verpflichtungen gegenüber den Betroffenen gewahrt und die *first user* Ansprüche der Datenproduzenten respektiert werden.

(1) Wie bereits in der ersten Förderphase des VFDB (10/2013-09/2016) werden die drei Institute GESIS, DIPF und IQB auch in der zweiten Förderphase *Forschungsdaten archivieren und zur Nachnutzung bereitstellen*. Um der Vielfalt der empirischen Bildungsforschung nachhaltig gerecht werden zu können, sollen die Services weiterer relevanter Forschungsdatenzentren mit dem bestehenden Angebot verknüpft werden. Dabei sind verschiedene Kooperationsformen geplant, die die verschiedenen Ausgangslagen und Möglichkeiten der jeweiligen FDZ berücksichtigen.

Ansprechpartner: DIPF Doris Bambey; Beteiligte: DIPF, GESIS, IQB

(2) Zentral für den Erhalt und die Nachnutzung von Forschungsdaten ist deren Sichtbarkeit und Auffindbarkeit. Hierzu wird der VFDB die Dokumentation von Studien und Daten im *Web-Portal forschungsdaten-bildung.de* ausbauen. Dezentral an den einzelnen FDZ gesicherte und bereitgestellte Daten werden so unabhängig von ihrem Aufbewahrungsort zentral nachgewiesen. Realisiert wird dies über einen Metadatenaustausch mit dem Portal *dalraSearchNet*. Darüber hinaus schafft der VFDB durch die Verknüpfung von Daten mit den darauf basierenden Publikationen einen zusätzlichen Mehrwert. Eine wichtige Voraussetzung für einen zentralen Datennachweis ist

die Harmonisierung der zur Beschreibung eingesetzten Metadaten und die Entwicklung von Standards und Best Practices für die Beschreibung und Erschließung von Forschungsdaten. Dieses soll in der zweiten Förderphase in Kooperation mit weiteren FDZ mit Bezug zur empirischen Bildungsforschung erarbeitet werden.

Ansprechpartner: DIPF, Thomas Schwager; Beteiligte: DIPF, GESIS, IQB

(3) Forscher/innen werden durch vielfältige *Schulungs-, Informations- und Beratungsangebote* bei der Produktion qualitativ hochwertiger und nachnutzbarer Forschungsdaten unterstützt. Das zweitägige Workshopangebot „Forschungsdatenmanagement für die empirische Bildungsforschung“ wird hierzu bedarfsorientiert weiterentwickelt und zukünftig im jährlichen Turnus durch GESIS angeboten werden. Um die Reichweite der Qualifizierungsangebote weiter zu erhöhen, wird das bisherige Schulungsprogramm zusätzlich durch Webinare, ein begleitendes Diskussionsforum im Web und die Ausbildung von Multiplikatoren („Train-The-Trainer“) ergänzt. Zudem soll das Thema Forschungsdatenmanagement zukünftig stärker in grundständiger Weise innerhalb der Ausbildung des wissenschaftlichen Nachwuchses verankert werden. Zu diesem Zweck wird mit Fachgesellschaften und Trägern renommierter Graduiertenkollegs und Summer Schools die Einbindung dieses Querschnittsthemas in deren Programmplanung umgesetzt.

Ansprechpartner: GESIS, Reiner Mauer; Beteiligte: DIPF, GESIS, IQB

Die *Projektgesamtkoordination* liegt beim DIPF, bei der Durchführung der einzelnen Aufgaben sind jeweils sämtliche Projektpartner beteiligt. Der Austausch mit weiteren FDZ soll im Rahmen jährlicher Netzwerktreffen stattfinden und gegebenenfalls themenspezifisch in Arbeitsgruppen intensiviert werden.

Informationen zu einzelnen Arbeitspaketen der zweiten Förderphase des VFDB

Für eine nachhaltige Infrastruktur für die Archivierung und Bereitstellung von Forschungsdaten werden im Verbund in der zweiten Förderphase folgende konzeptionelle Fragen im Rahmen einzelner Arbeitspakete bearbeitet:

Welche Forschungsdaten sollen zukünftig akquiriert werden? (AP Akquisitionskonzept)

Während der ersten Förderphase war der VFDB ausschließlich für Projekte des BMBF-Rahmenprogramms zur Förderung der empirischen Bildungsforschung zuständig. Im Zuge der zweiten Förderphase erfolgt eine Öffnung auch für andere Projekte. Dies erfordert ein Akquisitionskonzept zur systematischen Erfassung und Auswahl produzierter Datenbestände der empirischen Bildungsforschung. Hierzu sind Recherchen abgeschlossener und laufender Studien ebenso erforderlich wie der Kontakt zu Förderorganisationen sowie zu den Universitäten und Forschungsinstituten, an denen Bildungsforschungsdaten produziert und aufbewahrt werden. Im Ergebnis sollte ein Akquisitionskonzept entwickelt werden, aus dem hervorgeht, welche Daten durch den VFDB aktiv akquiriert und vorrangig bearbeitet werden sollten.

Ansprechpartner: GESIS, Reiner Mauer

Wie (intensiv) sollten und können diese Daten jeweils bearbeitet und aufbewahrt (kuratiert) werden? (AP Archivierungs- und Servicepolicy)

Im Rahmen einer Archivierungs- und Servicepolicy sollen Standards zur differenzierten Bearbeitung von Forschungsdaten definiert werden. Ausgangspunkt ist die Tatsache, dass vollumfassende Archivierungs- und Bereitstellungsservices aus Kapazitäts- bzw. Kostengründen nicht für alle Forschungsdaten gleichermaßen angeboten werden können und auch nicht in gleichem Maße angeboten werden müssen. Je nach technischer Beschaffenheit, Sensibilität oder der Dokumentationsqualität variieren sowohl die auf Seiten der FDZ erforderlichen Bearbeitungsaufwände als auch die Nachnutzungspotenziale der Daten. Daher soll ein Stufenmodell ausgearbeitet werden, das die Heterogenität der Ausgangslagen der Forschungsdaten berücksichtigt und entsprechend differenzierte Services der Dokumentation und Aufbereitung erlaubt. Die Servicestufen unterscheiden sich darin, wie feingranular Datenbestände dokumentiert (z. B. auf Ebene der einzelnen Transkripte, Aufzeichnungseinheiten oder Variablen) und wie intensiv sie aufbereitet werden (bspw. Erstellen zusätzlicher Kontextmaterialien, Durchführung zusätzlicher Fehlerprüfungen, weitergehende Anonymisierungsmaßnahmen). Das Vorliegen einer Archivierungs- und Servicepolicy soll die Bearbeitung von Daten durch den VFDB kriteriengestützt, transparent und nachvollziehbar begründen.

Ansprechpartnerin: DIPF, Maïke Porzelt

Wie können unterschiedliche Services finanziert werden und welche Kosten entstehen für die Bearbeitung der Daten? (AP Geschäftsmodelle)

Zur Beantwortung dieser Fragen werden u. a. alternative Gebührenmodelle für die konkreten Serviceleistungen der FDZ im VFDB erarbeitet. In einem ersten Schritt gilt es hierzu, Zeit- und Ressourcenaufwände zu erfassen, die für die Archivierung und Bereitstellung unterschiedlicher Datenbestände bei den jeweiligen FDZ entstehen sowie die jeweiligen Kostentreiber (z. B. Datentyp, Materialart, Umfang) zu identifizieren. In einem zweiten Schritt werden verschiedene Szenarien der Refinanzierung einzelner Services und deren Allokationswirkungen vergleichend diskutiert. Ein tragfähiges Geschäftsmodell ist notwendig, um den Ressourceneinsatz zur Archivierung und Bereitstellung von Forschungsdaten effektiv planen und steuern zu können. Der VFDB befasst sich hier mit einer neuen Herausforderung und einem Desiderat für die bundesdeutsche Forschungsdateninfrastruktur.

Ansprechpartnerin: DIPF, Alexia Meyermann

Welche Selektionskriterien und -verfahren können für eine differenzierte Bearbeitung von Forschungsdaten angewandt werden? (AP Selektion von Forschungsdaten)

Im Austausch mit der Forschungscommunity soll in Ergänzung zu dem vom VFDB entwickelten Akquisitionskonzept und der entwickelten Archivierungs- und Servicepolicy (vgl. oben) der Frage der Selektion von Forschungsdaten auf Basis ihres Nachnutzungspotentials nachgegangen werden. Angesichts großer und steigender Mengen an Forschungsdaten bei gleichzeitig begrenzten Ressourcen für die Archivierung, d. h. die Aufbereitung, Dokumentation und Sicherung der nachhaltigen Zugänglichkeit, geht es darum, zu entscheiden, welche Daten für welchen Zeitraum, in welcher Form zu archivieren und bereitzustellen sind (bspw. über FDZ oder universitäre Repositorien, vollständig oder teilweise). Hierzu sind Kriterien und Verfahren zur Beurteilung der Nachnutzungspotenziale von Forschungsdaten zu entwickeln, die derartige Entscheidungen begründen und transparent und nachvollziehbar gestalten. Das Thema wird im Rahmen eines Workshops unter anderem mit Wissenschaftlern und Wissenschaftlerinnen beispielsweise aus LERN, DGfE und GEBF erarbeitet werden.

Ansprechpartner: IQB