



FDM-Portfolio **Psychologie & Bildungsforschung**

Disziplinspezifische Dienste, Praktiken & Leitlinien
für das Forschungsdatenmanagement

Janna Kienbaum | Boris Jacob
Universität Potsdam

Potsdam 2023

DIE AUTOR*INNEN

Janna Kienbaum  [0000-0002-8653-5051](https://orcid.org/0000-0002-8653-5051)

Boris Jacob  [0000-0002-8565-3312](https://orcid.org/0000-0002-8565-3312)

HERAUSGEBER

Universität Potsdam mit dem Projektschwerpunkt „FD-Strategien für Fachbereiche“.

Diese Publikation wurde im Rahmen des Verbundprojekts „FDNext“ durch die Deutsche Forschungsgemeinschaft gefördert – Projektnummer: 429 828 830



IMPRESSUM

„FDM-Portfolio Psychologie & Bildungsforschung: Disziplinspezifische Dienste, Praktiken & Leitlinien für das Forschungsdatenmanagement“ von Janna Kienbaum und Boris Jacob ist lizenziert unter der Creative Commons Attribution 4.0 International (CC BY 4.0)



ZITATIONSVORSCHLAG

Janna Kienbaum und Boris Jacob (2023). FDM-Portfolio Psychologie & Bildungsforschung: Disziplinspezifische Dienste, Praktiken & Leitlinien für das Forschungsdatenmanagement. 10.5281/zenodo.7433874

DOI

10.5281/zenodo.7433874



Inhalt

Zum Portfolio	4
1. Forschungsvorhaben planen	7
1.1 Leitlinien und Empfehlungen zum Umgang mit Forschungsdaten	7
1.1.1 Deutsche Gesellschaft für Psychologie (DGPs)	7
1.1.2 Deutsche Gesellschaft für Erziehungswissenschaft (DGfE), Gesellschaft für Empirische Bildungsforschung (GEBF) und Gesellschaft für Fachdidaktik (GFD)	7
1.1.3 DFG Fachkollegium „Sprachwissenschaften“	7
1.2 Infrastrukturdienstleister	8
1.2.1 Leibniz-Institut für Psychologie (ZPID)	8
1.2.2 Verbund Forschungsdaten Bildung (Verbund FDB)	8
1.2.3 Leibniz-Institut für Sozialwissenschaften (GESIS)	8
1.2.4 CLARIAH-DE	8
1.3 Recherche	9
1.3.1 PubPsych	9
1.3.2 Publikationsdatenbank des DIPF	9
1.3.3 Fachportal Pädagogik	9
1.4 Datenmanagementpläne (DMP)	9
1.4.1 <i>Use Case: Datenmanagementplan</i>	9
1.4.2 DMP-HuWi (Universität Potsdam)	10
1.4.3 DataWiz (DMP)	10
1.4.4 MRT-Daten	10
1.4.5 Checkliste in der empirischen Bildungsforschung & STAMP	10
1.5 Präregistrierung	11
1.5.1 <i>Use Case: Präregistrierung</i>	11
1.5.2 OSF Registries & Vorlage für die Psychologie	11
1.5.3 PreReg	11
1.5.4 Allgemeines Template für die Psychologie	11
1.5.5 Journale mit Registered Reports	12
1.5.6 Workshop-Reader	12
1.6 Elektronische Laborbücher	12
1.6.1 PsychNotebook	12
1.6.2 Redcap	12
1.6.3 Rspace	13
1.6.4 eLabFTW	13
1.7 Datenschutz & Ethik	13
1.7.1 Übersicht zur Erhebung personenbezogener Daten	13
1.7.2 QualiAnon	14
1.7.3 De-Identifizierungsmethoden	14
2. Daten erheben/Daten und Instrumente nachnutzen	15
2.1 Psychologie	15
2.1.1 PsychArchives	15
2.1.2 FDZ am Leibniz-Institut für Psychologie (ZPID)	15
2.1.3 Open Neuro	15
2.1.4 Open Science Framework (OSF)	15
2.1.5 Open Test Archive	16
2.2 Bildungswissenschaften	16
2.2.1 Das Nationale Bildungspanel (NEPS)	16
2.2.2 Institut zur Qualitätsentwicklung im Bildungswesen (IQB)	16
2.2.3 FDZ-Bildung	16
2.2.4 Datenarchiv „Kindheit und Jugend im urbanen Wandel“	16
2.2.5 Archiv für pädagogische Kasuistik (ApaeK)	17
2.2.6 Fachportal Pädagogik	17

2.2.7	Instrumentendatenbank.....	17
2.3	Sozialwissenschaftliche (Umfrage-)Daten.....	17
2.3.1	SowiDataNet datorium	17
2.3.2	Qualiservice.....	17
2.3.3	CESSDA Data Catalogue.....	17
2.3.4	Repositorium für Messinstrumente ZIS.....	17
3.	Daten verarbeiten und dokumentieren.....	18
3.1	Standardformate	18
3.1.1	<i>Use Case: Dateiformate</i>	19
3.1.2	Dateiformate bei PsychArchives.....	19
3.1.3	Dateiformate des Verbund FDB.....	19
3.1.4	Brain Imaging Data Structure (BIDS).....	19
3.1.5	Psych-DS.....	19
3.2	Aufbereitung und Dokumentation quantitativer Forschungsdaten.....	20
3.2.1	<i>Use Case: Datendokumentation 1</i>	20
3.2.2	Aufbereitung quantitativer Forschungsdaten.....	20
3.2.3	Fehlende Werte in der Aufbereitung	20
3.2.4	Open Refine.....	20
3.2.5	Syntaxen.....	21
3.2.6	Methoden- und Feldbericht.....	21
3.2.7	Methodenbericht am Bsp. der StEGStudie	21
3.2.8	Checkliste eines Methodenberichts am Bsp. Der ALLBUS-Bevölkerungsumfrage.....	21
3.2.9	Dokumentationsleitfaden am Bsp. von „mail surveys“	21
3.2.10	Dokumentationsleitfaden am Bsp. von Fragebogenübersetzungen	21
3.2.11	Checkliste zum Dokumentieren am Bsp. von Online-Umfragen	21
3.3	Aufbereitung & Dokumentation/Kontextualisierung qualitativer Forschungsdaten.....	22
3.3.1	<i>Use Case: Datendokumentation 2</i>	22
3.3.2	Studienreport	22
3.3.3	Leitfaden am Bsp. einer Gruppendiskussion.....	23
3.3.4	Methoden- und Feldbericht.....	23
3.3.5	Studien- und Datendokumentation am Bsp. narrativer Interviews	23
3.3.6	Dokumentation am Bsp. Von face-to-face-Befragungen.....	23
3.3.7	Transkription.....	23
3.3.8	Unterrichtsaufnahmen.....	23
3.3.9	Videoaufnahmen.....	23
3.4	Codebuch	24
3.4.1	DataWiz II (beta)	24
3.4.2	<i>Use Case: DataWiz II</i>	24
3.4.3	DDI-Codebuchschemata.....	24
3.4.4	Codebuch-Muster	24
3.5	Metadatenstandard	25
3.5.1	Data Documentation Initiative (DDI)	25
3.5.2	DataWiz II (beta) und PsychArchives.....	25
3.5.3	openMINDS	25
4.	Daten archivieren und publizieren	26
4.1	Psychologie	26
4.1.1	FDZ am Leibniz-Institut für Psychologie (ZPID)	26
4.1.2	PsychArchives	26
4.1.3	<i>Use Case: PsychArchives</i>	26
4.1.4	Open Neuro	27
4.1.5	Open Science Framework (OSF).....	27
4.1.6	Open Test Archive	27
4.2	Bildungswissenschaften	27
4.2.1	Institut zur Qualitätsentwicklung im Bildungswesen (IQB)	27
4.2.2	FDZ-Bildung.....	28
4.2.3	Datenarchiv „Kindheit und Jugend im urbanen Wandel“.....	28

4.2.4	Archiv für pädagogische Kasuistik (Apaek)	28
4.2.5	Fachportal Pädagogik	28
4.2.6	Instrumentendatenbank.....	28
4.3	Sozialwissenschaftliche (Umfrage-)Daten.....	28
4.3.1	SowiDataNet datorium	28
4.3.2	Qualiservice	28
4.3.3	CESSDA Data Catalogue.....	29
4.3.4	Repositorium für Mess- und Erhebungsinstrumente ZIS	29
4.4	Zugriffsklassen.....	29

Zum Portfolio¹

Forschungsdatenmanagement (FDM) gilt seit vielen Jahren als grundlegender Bestandteil der Guten Wissenschaftlichen Praxis.² Es fördert Open Science-Praktiken sowie FAIRe³ und nachhaltige Datenkulturen institutionsübergreifender Fachdisziplinen mit deren spezifischen Forschungsdaten und -methoden. Die Forderung nach einem höheren Bewusstsein für transparentes FDM und für Open Science setzte vor allem mit der mangelhaften Replizierbarkeit von Forschungsbefunden ein, die u.a. auch Psycholog*innen betraf.⁴ Das bekannte Problem des HARKings (Hypothesising After the Results are Known) sollte im Umgang mit Forschungsdaten vermieden werden, um Studien nicht nur nachvollziehen, sondern auch reproduzieren zu können.⁵

Gutes FDM und Open Science-Praktiken zu betreiben, sollte nicht nur Anspruch von Forschungsförderern, sondern größte Motivation für die Forschenden selbst sein. Für Forscher*innen muss es selbstverständlich werden, den Lebenszyklus ihrer Forschungsdaten bewusst zu gestalten. Dabei spielt die disziplinspezifische Berücksichtigung ihrer Forschungsdaten eine zentrale Rolle, verlangen doch z.B. Kulturerbe- oder Textdaten andere Metadatenstandards als Geodaten, Fragebogendaten eine unterschiedliche Datenschutz-Handhabung als Sprachkorpora und Software-Quellcodes andere Datenrepositorien als Messdaten der Elektroenzephalografie (EEG).⁶

Durch die Breite der Forschungslandschaft an den Universitäten und den dort vorherrschenden begrenzten Personalkapazitäten der Serviceeinrichtungen ist es dabei jedoch kaum möglich, eine fachspezifische oder auf besondere Organisationsformen bezogene FDM-Beratung anzubieten. Eine disziplinspezifische Unterstützung und vor allem Beratung im FDM zählt noch nicht zum Standard der FD-Serviceeinrichtungen, wird jedoch zunehmend durch domänenspezifische Initiativen⁷, nicht zuletzt von den Konsortien der Nationalen Forschungsdaten Infrastruktur (NFDI)⁸ angestrebt.

Um diese Lücke zu füllen und Forscher*innen mit zugeschnittenen FD-Diensten in ihrer Forschungspraxis vertraut zu machen, wurde das vorliegende Dokument konzipiert, das themen- und disziplinspezifische FDM-Services als Portfolios bündelt. Es bietet standortübergreifend Informationen zum Umgang mit Forschungsdaten für die Psychologie, Bildungsforschung, mitunter Sozialwissenschaften an und stellt das inhaltliche Arbeitsprodukt des Arbeitspaketes der Universität Potsdam „FD-Strategien für Fachbereiche“ vor. Dieses ist Teil des DFG-geförderten Verbundprojektes „FDNext“ (2020-2023).⁹

Inhalt & Hintergrund

Das Portfolio umfasst praxisnah ausgewählte FDM-Inhalte, wozu softwaregestützte Dienste, Leitlinien und Handreichungen gehören. Es ist ähnlich einem *Glossar* kompakt aufgebaut. Gegenüber einer

¹ Alle Internetquellen wurden zuletzt am 13.02.2023 geprüft.

² Vgl. Kodex „Leitlinien zur Sicherung guter wissenschaftliche Praxis“ der Deutschen Forschungsgemeinschaft“ (2019), Zugriff am 08.02.2023, <https://doi.org/10.5281/zenodo.3923601>.

³ Vgl. <https://www.go-fair.org/fair-principles/>.

⁴ Vgl. Goodman, Steven N.; Fanelli, Daniele; Ioannidis, John P.A. (2016): What does research reproducibility mean? In: *Sci Transl Med* 8, 341ps12341ps12. DOI: 10.1126/scitranslmed.aaf5027.

⁵ Ebd., S.5.

⁶ Vgl. die Übersicht an Informationen für disziplinspezifische Wissenschaftsbereiche von forschungsdaten.info unter <https://forschungsdaten.info/wissenschaftsbereiche/>

⁷ Vgl. beispielsweise in den Ingenieurwissenschaften <https://zenodo.org/record/3520902> und <https://www.izus.uni-stuttgart.de/fokus/fdm-projekte/dipling/>; in den Digitalen Geisteswissenschaften siehe <https://www.izus.uni-stuttgart.de/fokus/fdm-projekte/replay/>.

⁸ Vgl. <https://www.nfdi.de/>.

⁹ Das Arbeitspaket der Universität Potsdam baut dabei u.a. auf den Ergebnissen der BMBF-Projekte „FDMentor“ und „PODMAN“ (Blask, Förster, Lemaire & Minn 2018) auf und konkretisiert diese auf strategischer Ebene zur Verankerung der Prinzipien des FDM und von FAIR Data in Fachbereichen. Die Zielgruppe des Portfolios ergibt sich aus der Zusammenarbeit mit der Humanwissenschaftlichen Fakultät der Universität Potsdam, die sich aus den Strukturbereichen der Kognitions- und Bildungswissenschaften zusammensetzt. Sie eignete sich angesichts der hohen FD interessierten Akteurschaft und der Heterogenität qualitativer und quantitativer Forschungsmethoden und -daten als „Modell“ des Arbeitspaketes.

umfassenden Abdeckung aller Forschungsbereiche hat das Portfolio vor allem das Ziel, ausgewählte bzw. vorherrschende Bedarfe und Anforderungen aus den Disziplinen am Beispiel der Humanwissenschaftlichen Fakultät der Universität Potsdam zu adressieren.¹⁰ Spezifische Produkte mit besonders hohem Bedarf werden dabei anhand von Daten-typischen Use Cases bzw. Anwendungsbeispielen vorgestellt oder eingeleitet.¹¹

Grundlage für die Zusammenstellung der FD-Dienste bildete eine intern durchgeführte qualitative Bedarfsermittlung innerhalb der Studienfelder Bildungs- und Kognitionswissenschaften im Jahr 2021. Workshops und insbesondere individuelle, regelmäßige Gespräche mit den Forscher*innen (meist Doktorand*innen, aber auch Arbeitsgruppenleiter*innen) dienten der Ermittlung der Portfolio-Zusammenstellung. Da insbesondere der personenbezogene Datenschutz in den Fächern der Psychologie und Bildungsforschung einen wichtigen Stellenwert im FDM darstellt, integriert das Portfolio das Dokument „Forschungsdatenmanagement bei personenbezogenen Daten – eine Handreichung“, das zeitgleich im Rahmen von „FDNext“ erarbeitet wurde.¹²

Neben der Zusammenarbeit mit den Forscher*innen, orientiert sich die Zusammenstellung der Portfolio-Produkte an parallelen Entwicklungen disziplinspezifischer FD-Infrastrukturen. Dazu zählen insbesondere die Vorhaben der NFDI. In Bezug auf die Zielgruppe der Psycholog*innen und Bildungsforscher*innen bot beispielsweise das Service-Angebot vom „KonsortSWD“ eine umfangreiche Informationsquelle für die Auswahl von Daten und Forschungsdatenzentren. Ein Abgleich zu Vorhaben der NFDI-Sektionen „ELSA“ zu ethischen und datenschutzrechtlichen Prinzipien im FDM, aber auch „EduTrain“ zu Förderung von Datenkompetenz von FDM-Akteuren und Wissenschaftler*innen bestärkte und rahmte u.a. die Auswahl der Portfolio-Dienste. Zuzüglich verhalf eine enge Zusammenarbeit mit der disziplinspezifischen Community bei der Zusammenstellung der Portfolio-Inhalte. Dazu zählen v.a. intensive Austauschtreffen mit dem Verbund Forschungsdaten Bildung (Verbund FDB), dem Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF), dem Leibniz-Institut für Psychologie (ZPID) und mit der UAG „Textanleitungen“ der RDMO-Contentgroup.

Das Portfolio ist Teil einer Reihe disziplinspezifischer FDM-Services, die bereits bestehen oder zukünftig für Forscher*innen zur Verfügung gestellt werden. In der Entwicklung begriffene Plattformen oder Dienste, die insbesondere qualitative Datenerhebungen ansprechen, sind beispielsweise das QualiDataNetwork¹³ sowie ein FDM-Portfolio¹⁴ vom Konsort SWD mit Fokus auf die qualitative Forschung.

Zielgruppe

Das Portfolio dient insbesondere als Einstieg für Nachwuchswissenschaftler*innen der Psychologie und Bildungswissenschaften wie Studierende und Doktorand*innen/PhD's, um sich mit den Grundlagen des FDM sowie mit Forschungsdatendiensten und -handreichungen vertraut zu machen. Es sorgt für einen erleichterten und FAIRen Umgang mit den eigenen Forschungsdaten und unterstützt die Organisation, Durchführung und Verfügbarmachung von Forschungsvorhaben.

¹⁰ Das Train-the-Trainer(TtT)-Konzept aus dem „FDMentor“-Projekt stellt dem gegenüber eine umfassende und generische Materialsammlung des FDM mit didaktischem Fokus dar, Vgl. <https://zenodo.org/record/5773203#.Y34GI3ZM2w>. Arbeitspaket 5 des FDNext-Projektes konzipiert das TtT-Konzept disziplinspezifisch für die Psychologie, Vgl. <https://www.fu-berlin.de/sites/forschungsdatenmanagement/ueber-uns/projekt-fdnext/index.html>.

¹¹ Die Ermittlung und Verwendung von Use Cases war einerseits stark von den individuellen zeitlichen Kapazitäten der Forscher*innen, andererseits von deren Schutzbedürfnissen gegenüber geistigem Eigentum abhängig. Die in das FDM-Portfolio eingebundenen Use Cases variieren daher in ihrem Umfang. Das Portfolio ist ähnlich der disziplinspezifischen FDM-Integration am Bsp. der Informatik von Katarzyna Biernacka und Sandra Schulz praxisorientiert auf die Zielgruppe der Wissenschaftler*innen ausgerichtet. Vgl. Biernacka, K. & Schulz, S. (2022). Forschungsdatenmanagement in der Informatik. Logos Verlag Berlin. <https://doi.org/10.30819/5490>.

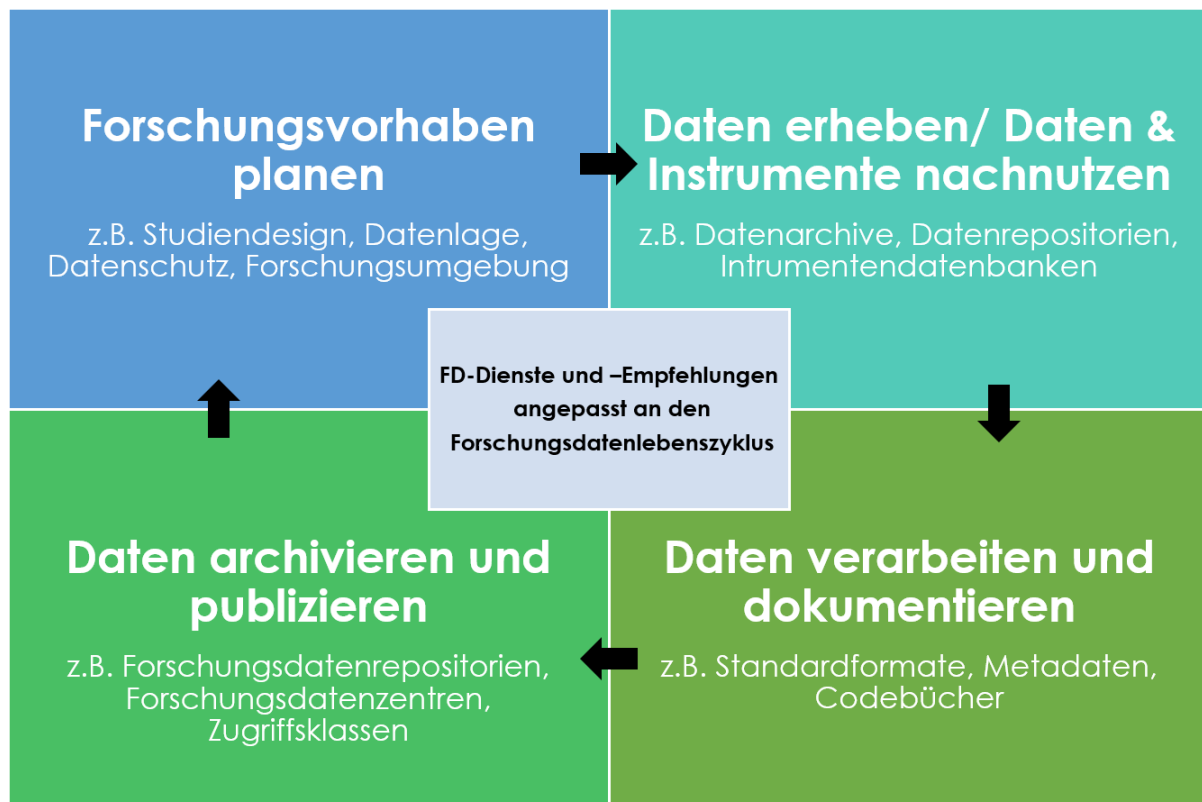
¹² Vgl. Kienbaum, Janna, Fischer, Patryk, & Paßmann, Sven. (2023). Forschungsdatenmanagement bei personenbezogenen Daten - eine Handreichung. Zenodo. <https://doi.org/10.5281/zenodo.7428524>.

¹³ Vgl. <https://www.konsortswd.de/konsortswd/das-konsortium/services/qualidatanet/>.

¹⁴ Vgl. <https://www.konsortswd.de/konsortswd/das-konsortium/services/fdm-quali-sozialforschung/>.

Das Portfolio richtet sich ferner an Arbeitsgruppen- oder Projektleiter*innen/Principle Investigators (PI's), die das Material für die Betreuung der Nachwuchswissenschaftler*innen und deren Qualifizierungsarbeiten heranziehen können.

Der Aufbau des Portfolios orientiert sich am Forschungsdatenlebenszyklus (siehe nachfolgende Abbildung). Dabei wurden die Inhalte auf vier zentrale Prozesse abstrahiert, um größere Redundanzen bei der Zuordnung der Dienste und Empfehlungen zu verhindern. Die ausgewählten Portfolio-Inhalte sind so strukturiert, dass sie dynamisch für die zentralen chronologischen Schritte einer Forschungsstudie *1. Forschungsvorhaben planen, 2. Daten erheben/Daten und Instrumente nachnutzen, 3. Daten verarbeiten und dokumentieren sowie 4. Daten archivieren und publizieren* genutzt werden können.



FD-Dienstempfehlungen und -Handreichungen angelehnt an den Forschungsdatenlebenszyklus von [forschungsdaten.info](https://www.forschungsdaten.info) und vom Verbund Forschungsdaten Bildung (Verbund FDB).¹⁵

¹⁵ Vgl. die Übersicht „Datenlebenszyklus“ unter <https://www.forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>; Vgl. Verbund Forschungsdaten Bildung: Checkliste zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung. Version 1.1. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation 2015, 9 S. - (fdb info; 2) - URN: urn:nbn:de:0111-pedocs-222986 - DOI: <https://doi.org/10.25656/01:22298>, S.1.

1. Forschungsvorhaben planen

Die ersten Schritte im FDM dienen der Planung und Organisation Ihres Forschungsvorhabens und Studiendesigns. Sie umfassen Open Science-Verfahren und FDM-Dienste, die Sie noch *vor* der eigentlichen Erhebung oder Nachnutzung der Daten berücksichtigen sollten, um späteren Lücken bei der Datenspeicherung, -archivierung, -lizenzierung und -publikation vorzubeugen. Dazu gehören insbesondere Informationen zu disziplinspezifischen Richtlinien im Umgang mit Forschungsdaten, Infrastrukturdiensteanbietern, Publikationsportalen zur Studienrecherche, Datenmanagementplänen, Umsetzungsmöglichkeiten von Prä-Registrierungen und Elektronischen Laborbüchern (ELB) u.a. zur transparenten und vernetzten Forschungsarbeit sowie zu datenschutzrechtlichen und ethischen Anforderungen.

1.1 Leitlinien und Empfehlungen zum Umgang mit Forschungsdaten

Die Deutsche Forschungsgemeinschaft (DFG) formuliert seit 2015 Leitlinien zum Umgang mit Forschungsdaten, die obligatorisch von Projektantragsstellenden zu beantworten sind.¹⁶ Diese Checkliste wurde von den Fachgesellschaften an ihre jeweiligen Disziplinen und typischen Daten angepasst.

1.1.1 Deutsche Gesellschaft für Psychologie (DGPs)

Die Deutsche Gesellschaft für Psychologie gab im Jahr 2016 eine Empfehlung zum Umgang mit Forschungsdaten für die Psychologie in Deutschland heraus. Diese wurde zum jetzigen Stand bereits zweimal überarbeitet. Die letzte Version mit dem Titel [Management und Bereitstellung von Forschungsdaten in der Psychologie: Überarbeitung der DGPs-Empfehlungen](#) stammt aus dem Jahr 2020.

1.1.2 Deutsche Gesellschaft für Erziehungswissenschaft (DGfE), Gesellschaft für Empirische Bildungsforschung (GEBF) und Gesellschaft für Fachdidaktik (GfD)

Im Auftrag der DFG formulierten die Fachgesellschaften der empirischen Bildungsforschung 2020 eine [gemeinsame Stellungnahme](#) zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten in den Erziehungs- und Bildungswissenschaften sowie Fachdidaktiken.

Möglichkeiten der Bereitstellung und Nachnutzung insbesondere *qualitativer* Daten in erziehungswissenschaftlichen Forschungsprojekten berücksichtigt die [Stellungnahme der DGfE zur Archivierung, Bereitstellung und Nachnutzung qualitativer Forschungsdaten in der Erziehungswissenschaft](#).

1.1.3 DFG Fachkollegium „Sprachwissenschaften“

Das DFG-Fachkollegium "Sprachwissenschaften" veröffentlichte 2019 [Empfehlungen zu datentechnischen Standards und Tools bei der Erhebung von Sprachkorpora](#). Als Richtlinie unterstützen sie eine optimale Archivierbarkeit und Nachnutzbarkeit der Daten von mündlichen und schriftlichen Sprachkorpora, die ggf. für kognitionslinguistische Forschungsfragen im Umgang mit Sprachkorpora relevant sind.

¹⁶ Vgl. DFG (2021): Umgang mit Forschungsdaten: Checkliste für Antragstellende zur Planung und zur Beschreibung des Umgangs mit Forschungsdaten in Forschungsvorhaben, https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf.

1.2 Infrastrukturdiensteanbieter

Forschenden stehen umfangreiche Services durch Institutionen und Einrichtungen zur Verfügung, die oft in Verbundstruktur das Qualitätsmanagement der Forschungsdaten disziplinspezifisch unterstützen. Diese Forschungsinfrastrukturen bieten je nach Fachbereich Beratungen, Empfehlungen und Dienste primär zum Management, zur Dokumentation, zur Archivierung sowie zum Veröffentlichen und Teilen der Daten an. Auch Mess- bzw. Erhebungsinstrumente zählen zum Beratungsgegenstand. Bei Disziplinen, in denen personenbezogene Daten erhoben werden, kommen insbesondere datenschutzrechtliche und ethische Informationsangebote hinzu.

Es empfiehlt sich, noch in der Planungsphase der Datenerhebung Kontakt zu den FDM-Service- und Strukturanbietern aufzunehmen, um sich über Datenschutzvorgehen oder Zugriffs- und Nachnutzungsmöglichkeiten Ihrer Forschungsdaten informieren zu lassen.

1.2.1 Leibniz-Institut für Psychologie (ZPID)

Das [Leibniz-Institut für Psychologie](#) (ZPID) versteht sich als Public-Open-Science-Institut und stellt die zentrale überregionale Adresse im deutschsprachigen Raum für die Psychologie dar. Angesiedelt an der Universität Trier bietet es Forscher*innen ein umfangreiches Angebot an Diensten, die den gesamten Forschungsdatenprozess unterstützen. Diese reichen von der Literaturrecherche und Studienplanung über die Datenerhebung und -auswertung bis hin zur Dokumentation, Archivierung und Publikation und allen Bestandteilen des Forschungsprozesses (z.B. Daten, Code, Präregistrierung, etc.).

1.2.2 Verbund Forschungsdaten Bildung (Verbund FDB)

Der [Verbund Forschungsdaten Bildung](#) setzt sich aus nationalen Forschungsdatenzentren und Partnern der Bildungsforschung sowie weiteren Kooperationspartnern wie dem Konsortium für die Sozial-, Verhaltens-, Bildungs-, und Wirtschaftswissenschaften zusammen. Er bietet ein hochwertiges Angebot an Forschungsdaten für die empirische Bildungsforschung. Dazu zählen allgemein Suchmöglichkeiten und der Zugang von Forschungsdaten, die Publikation von Forschungsdaten unter datenschutzrechtlichen und dokumentarischen Standards sowie diverse Schulungs- und Beratungsmaterialien. Die Koordinierungsstelle ist dauerhaft am Leibniz-Institut für Bildungsforschung und Bildungsinformation angesiedelt. Die Basisdienste werden von den Partnerinstitutionen GESIS, IQB und DIPF betrieben.

1.2.3 Leibniz-Institut für Sozialwissenschaften (GESIS)

Das in Mannheim und Köln ansässige [Leibniz-Institut für Sozialwissenschaften](#) (GESIS) stellt die größte europäische Infrastruktureinrichtung für die empirischen Sozialwissenschaften dar. Das Institut hat mehrere Forschungsdatenzentren und bietet umfangreiche Dienst- und Beratungsangebote an, die sich ähnlich der anderen Forschungsdaten-Infrastrukturanbieter am Forschungsdatenlebenszyklus ausrichten. Schwerpunkt bildet die Erhebung und Verfügbarmachung von Forschungsdaten und die damit verknüpften Aktivitäten und Informationen. Forschungsthemen von GESIS betreffen Bereiche der Umfragemethodik, angewandten Informatik, FDM und aktueller gesellschaftlicher Fragen.

1.2.4 CLARIAH-DE

Die Forschungsinfrastruktur [CLARIAH-DE](#) richtet sich an Forscher*innen der Geistes- und Kulturwissenschaften und setzt sich aus den ehemaligen Forschungsinfrastrukturverbänden CLARIN-D und DARIAH mit europäischer Anbindung zusammen. Deren Angebote an technischen Infrastrukturen, Repositorien und Lehrangebote wurden integriert und weiterentwickelt. Schwerpunkt des Angebotes bilden die nachhaltige Bereitstellung von Forschungsdaten, technische Infrastrukturen, digitale Werkzeuge und virtuelle Forschungsumgebungen für die Geisteswissenschaften sowie Informations- und Schulungsmaterialien und Handreichungen zu u.a. Datenstandards und -verfahren. Die auf

textlichen und sprachlichen Quellen beruhenden Dienstangebote von CLARIAH-DE können für Forschungen aus der Disziplin der Kognitionslinguistik von Interesse sein.

1.3 Recherche

Bei der Recherche für ein geplantes Forschungsvorhaben und dessen Verortung im Forschungskontext, verhelfen Datenbanken, in denen eine umfassende Auswahl disziplinspezifischer Publikationen versammelt sind. Die umfassende Indexierung der Publikationen in den Datenbanken bietet sich insbesondere bei der Schlagwortauswahl bzw. -vergabe eigener Studien an, die publiziert werden sollen. Diese orientieren sich zunehmend an kontrollierten Vokabularen oder Thesauri der Fachdisziplinen.

1.3.1 PubPsych

[PubPsych](#) ist ein frei zugängliches Suchportal für internationale Psychologie-Publikationen. Es beinhaltet knapp über zwei Millionen Datensätze mit Verlinkungen zum Volltext. Als Zugriffsdienst auf die Referenzdatenbank Psyndex des Leibniz-Instituts für Psychologie (ZPID) stellt es die umfassendste Publikationsdatenbank im deutschsprachigen Raum dar.

1.3.2 Publikationsdatenbank des DIPF

Die [Publikationsdatenbank und die Publikationslisten](#) des DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation bieten Wissenschaftler*innen einen Überblick über den aktuellen Stand in der empirischen Bildungsforschung.

1.3.3 Fachportal Pädagogik

Ebenfalls vom DIPF wird das [Fachportal für Pädagogik](#) betrieben. Das Portal bietet Zugang zu wissenschaftlichen Fachinformationen für die Bildungsforschung, Erziehungswissenschaft und Fachdidaktik im deutschsprachigen Raum. Neben dem Angebot von Forschungsdaten und Forschungsinformationen bietet es ein Literaturportal, das u.a. der Recherche über aktuelle Forschungsthemen oder in Fachdatenbanken dient.

1.4 Datenmanagementpläne (DMP)

Ein Datenmanagementplan (DMP) unterstützt Forscher*innen bei der (disziplinspezifischen) Planung und Umsetzung einer Datenerhebung sowie -auswertung und orientiert sich an den Schritten des Datenlebenszyklus. Er behandelt die Organisation der Studie und ihrer Datenerhebung von der Planung über die Verarbeitung bis hin zur Archivierung, Veröffentlichung und Nachnutzbarkeit der Daten nach den FAIR Data-Prinzipien: Der DMP unterstützt also die Findability, Accessibility, Interoperability und Reusability der Daten.¹⁷ Die Erstellung eines DMPs ist zunehmend Bestandteil und Bedingung bei der Antragstellung von Fördergeldern¹⁸, ermöglicht als dynamisches Dokument jedoch auch im weiteren Verlauf der Forschungsstudie das Management der Daten im Team.

1.4.1 Use Case: Datenmanagementplan

*Ein Forscherteam aus Entwicklungspsycholog*innen und Bildungsforscher*innen plant eine kooperative Studie, die die Blickdauer und Bildverarbeitung von Kindern unterschiedlicher Kita- und Schuleinrichtungen aus Berlin und Brandenburg misst. In Teilstudien wird das Erinnerungsvermögen der Kinder unterschiedlicher Altersgruppen gemessen. Gearbeitet werden soll mit EEG und Eyetracking-Verfahren sowie qualitativen Interviews. Geplant wird demnach eine Aufteilung der Daten in zwei große Datensätze: in experimentelle und in Fragebogendaten.*

¹⁷ Vgl. <https://www.go-fair.org/fair-principles/>.

¹⁸ Vgl. <https://forschungsdaten.info/themen/informieren-und-planen/foerderrichtlinien/#c492536>.

*Das Vorhaben wird bei einem Drittmittelanbieter eingereicht, der bereits bei der Antragstellung Angaben zum Forschungsdatenmanagement verlangt, das heißt, das Team befindet sich am Beginn der Studie. Es ist nun auf der Suche nach einem Programm, das diesem interaktiv und dynamisch insbesondere bei folgenden Punkten behilflich sein soll: Bei der allgemeinen Organisation aller zentralen Schritte des Forschungsprozesses von der Methode der Datenerhebung, über Zugriffs- und Speicherregelungen der Forschungsdaten und deren Ort der Archivierung und Publikation sowie bei der rechtlichen und ethischen Regelung der Datenerhebung aufgrund der sensiblen Proband*innengruppe. Hinzu kommen Dokumentationsbedarfe bei der Festlegung der Verantwortlichkeiten und Koordination, da es sich um ein interdisziplinäres Forschungsteam von Universitäten aus Berlin und Brandenburg handelt.*

1.4.2 DMP-HuWi (Universität Potsdam)

Der Datenmanagementplan [DMP-HuWi](#) wurde im Rahmen des „FDNext“-Arbeitspaketes „FD-Strategien für Fachbereiche“ an der Universität Potsdam (UP) konzipiert. Er ist Resultat von Gesprächen und Use Case-Anwendungen mit den Forscher*innen aus der Humanwissenschaftlichen Fakultät der UP sowie den Sonderforschungsbereichen 1287 und 1294, an denen die Fakultät u.a. beteiligt ist. Der Dienst erfüllt den Bedarf nach einem Management-Tool, das an der UP zugänglich und disziplinspezifische Angaben z.B. zu Dateiformaten, Repositorien und v.a. Datenschutz- und Ethikinformationen dokumentieren soll. Diese können u.a. als Nachweis für Forschungsförderer oder für Ethik-Anträge genutzt werden. Angelehnt an den FoDaKo-Fragenkatalog des DFG-Fachkollegiums „109 Erziehungswissenschaft“¹⁹ und an Inhalte des Leibniz-Instituts für Psychologie²⁰ wurde deshalb ein DMP für Forscher*innen der Humanwissenschaftlichen Fakultät erstellt. Er basiert auf dem Softwareprogramm „Research Data Management Organizer“ (RDMO) und umfasst Schritte des FDM vor, während und nach der Datenerhebung mithilfe von Fragen und umfassenden Texterläuterungen.

1.4.3 DataWiz (DMP)

Das frei zugängliche Assistenzsystem [DataWiz](#) vom Leibniz-Institut für Psychologie (ZPID) begleitet das Datenmanagement in psychologischen Forschungsprojekten. Zielgruppe sind Wissenschaftler*innen aus der Psychologie. Das webbasierte Tool ermöglicht eine standardisierte Studiendokumentation sowie die Erstellung von einem Codebuch und weiterem Dokumentationsmaterial. Die aus der Anwendung des Tools hervorgehende Dokumentation kann sowohl für eine qualitätsgesicherte Selbstarchivierung der Forschungsdaten genutzt werden sowie auch für eine Archivierung in einem geeigneten Repository (z.B. dem disziplinspezifischen Repository PsychArchives des ZPID).

1.4.4 MRT-Daten

Das Fachgebiet "Psychologie" der [Universität Gießen](#) bietet für Neurowissenschaftler*innen eine Beratung u.a. in Hinblick auf Datenmanagement inkl. Qualitätsprüfung, -speicherung und DMP-Gestaltung an und stellt dafür Ressourcen zur Verfügung. Die Informationen beziehen sich speziell auf Messmethoden der Magnetresonanztomographie (MRT). Auch wenn sich das Angebot primär an Universitätsangehörige der Universität Gießen richtet, ist es ratsam, sich an den Fachbereich zu wenden, um einen DMP dieser Fachdisziplin anzufragen.

1.4.5 Checkliste in der empirischen Bildungsforschung & STAMP

Anstelle eines Dienstes bietet der Verbund Forschungsdaten Bildung (Verbund FDB) eine allgemeine [Checkliste](#) zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung an. Die Checkliste wird demnächst vom [STAMP](#) – dem Standardisierten

¹⁹ FoDaKo ist die Abkürzung für „Forschungsdatenmanagement in Kooperation“ der Universitäten Düsseldorf, Siegen und Wuppertal. Vgl. Katalog DFG 109 Educational Research (FoDaKo) in RDMO unter <https://github.com/rdmorganiser/rmo-catalog>.

²⁰ Vgl. die Struktur und Inhalte des DataWiz vom ZPID unter <https://datawizkb.leibniz-psychology.org/>.

Datenmanagementplan für die Bildungswissenschaft – ersetzt, der Forscher*innen speziell aus der Bildungsforschung anhand von Fallbeispielen durch acht inhaltliche Module führt (Forschungsethik, Datenschutz, Urheberrecht, Datenorganisation, Nachvollziehbarkeit, Verfügbarkeit, Langfristsicherung und Verantwortlichkeiten & Aufwendungen). Der STAMP wird sowohl in Form von pdf-Dokumenten als auch als RDMO-Fragenkatalog verfügbar sein.

1.5 Präregistrierung

Eine Präregistrierung wird noch vor der eigentlichen Datenerhebung angelegt und legt das methodische Vorgehen und Studiendesign dar, mit dem die Forschungsfrage bzw. die Forschungshypothese beantwortet werden soll. Präregistrierungen sind Bestandteil einer Offenen Wissenschaft. Präregistrierungen sind sowohl unabhängig von Journalen nutzbar, können aber auch in Form von *Registered Reports* von Journalen angeboten und geprüft werden. Die Entscheidung zur Publikation der Studie wird aufgrund der Forschungsfragen, der geplanten Methoden und Auswertung getroffen, d.h. noch bevor die Ergebnisse bekannt sind. Präregistrierungen können sowohl wissenschaftlicher Intransparenz, als auch dem *publication bias* (Verzerrung der Forschung zugunsten positiver Befunde) entgegenwirken.

1.5.1 Use Case: Präregistrierung

Eine Doktorandin der Psychologie forscht auf dem Gebiet der klinischen Psychologie und Auswirkungen der Atemfrequenz auf den menschlichen Metabolismus. Neben persönlichen Fragebögen, plant sie quantitative Untersuchungsmethoden wie z.B. EKG-Messungen einzusetzen, um schließlich statistische Ergebnisse auswerten zu können. Sie möchte einen Teil bei einem Journal ihrer Disziplin publizieren und erkundigt sich nach den Bedingungen. Das Journal bietet ihr noch vor der Datenerhebung eine Prä-Registrierung in Form eines Registered Reports an, der schließlich in mehreren Review-Schleifen vom Journal geprüft wird. So können Transparenz und Nachvollziehbarkeit ihrer Studie gewährleistet werden. Da die Forscherin zwar schon vom Verfahren der Präregistrierung in ihrer Disziplin hörte, diese bisher jedoch noch nie selbst durchgeführt hat, sucht sie nach Informationen zu diesem Thema und disziplinspezifischen Standards, nach denen sie ihre Prä-Registrierung durchführen kann.

1.5.2 OSFRegistries & Vorlage für die Psychologie

Präregistrierung können in der generischen Datenplattform Open Science Framework (OSF) über die Sektion [OSFRegistries](#) erstellt werden. OSFRegistries bietet eine Durchsuchbarkeit eingereicherter Präregistrierungen u.a. anhand von Schlagworten.

Eine Präregistrierungs-[Vorlage](#) wurde von Anna Elisabeth van't Veer und Roger Giner-Sorolla für die sozialwissenschaftliche Psychologie entworfen, welche über OSF zur Verfügung steht. Van't Veer und Giner-Sorolla geben zudem eine [Anleitung](#), wie eine Prä-Registrierung in OSF einzureichen ist.

1.5.3 PreReg

Das Leibniz-Institut für Psychologie stellt die Plattform [PreReg](#) zur Einreichung domänenspezifischer Prä-Registrierungen zur Verfügung. Die eingereichten Prä-Registrierungen werden im disziplinspezifischen Repositorium „PsychArchives“ archiviert und erhalten sowohl einen Zeitstempel als auch einen dauerhaften Identifikator bzw. *Persistent Identifier* (z.B. DOI: Digital Object Identifier), um die Zitierbarkeit und langfristige Auffindbarkeit der Prä-Registrierung zu gewährleisten.

1.5.4 Allgemeines Template für die Psychologie

Ein internationales [Template](#) für quantitative Studien der Psychologie wurde nach internationalem Standard von der „Preregistration Task Force“ entworfen. Diese setzt sich aus der American Psychological Association (APA), der British Psychological Society (BPS), der Deutschen

Gesellschaft für Psychologie (DPGs) und dem Leibniz-Institut für Psychologie (ZPID) zusammen. Das Template kann als .xlsx (Microsoft Excel), .docx (Microsoft Word), .odt (Libre Office) oder .ipynb (JupyterLab)-Datei heruntergeladen werden. Die Vorlage enthält Texterläuterungen wie z.B. zu „data accessibility statements“.

1.5.5 Journale mit Registered Reports

Eine Auswahl an Fachzeitschriften, in denen auf Prä-Registrierung beruhende [Registered Reports](#) für ein Review-Verfahren eingereicht werden können oder die ein solches bei Einreichung einer Studie verlangen, bietet das Center for Open Science (COS) an. Bereitgestellt unter „Ressources“ werden eine Liste der Fachzeitschriften und weitere Informationen zur Einreichung von Registered Reports.

1.5.6 Workshop-Reader

Ein [Workshop-Reader](#) zum Thema "Präregistrierung" in der Bildungsforschung wurde von Claudia Neuendorf und Christin Rüdiger bereitgestellt, die am 29. Juni 2021 den Workshop mit dem Titel „Präregistrierung von Studien in der empirischen Bildungsforschung - Wozu, Wie und Wo“ durchführten. Der Reader bietet ab S. 8 praxisbezogene Informationen wo Prä-Registrierungen abgelegt werden können (Vgl. „PreReg“ des ZPIDs), welche Kriterien sie erfüllen müssen und welche Vorlagen sich für welche Studien (z.B. Experimentelle Forschung oder Sekundärforschung, explizit qual. Forschung) anbieten.

1.6 Elektronische Laborbücher

Für die kooperative und transparente Planung, Durchführung und Auswertung von Experimenten im Labor sind elektronische Laborbücher zu empfehlen (ELB).²¹ Die Softwareprogramme ermöglichen eine strukturierte Archivierung und Protokollierung der Experimente und helfen somit nicht nur bei der Ergebnissicherung, sondern auch beim Management der Forschung, Teamarbeit oder z.B. von Laborpraktika. ELBs dienen zugleich der Dokumentation des Forschungsprozesses und der produzierten Forschungsdaten.

1.6.1 PsychNotebook

Der FDM-Dienst [PsychNotebook](#) des Leibniz-Instituts für Psychologie stellt eine Online-Plattform für die Planung, Analyse und Dokumentation psychologischer Studien dar. Die Nutzung des Dienstes ist ohne Installation über eine Registrierung per Leibniz-Account oder ORCID möglich. PsychNotebook ist in der Lehre, zum kollaborativen Arbeiten, als Supplement für Zeitschriften und für die Verknüpfung mit Forschungsdaten/Coderepositorien einsetzbar.²²

1.6.2 Redcap

Bei [Redcap](#) handelt es sich um eine Webanwendung zur Erstellung und Verwaltung von Online-Umfragen und Datenbanken, entwickelt von der privaten Vanderbilt-Universität in den USA. Ursprünglich aus dem Bereich klinischer Gesundheitsforschung konzipiert, ist das Programm mittlerweile generisch für diverse Datentypen geeignet. Die Redcap-Software eignet sich insbesondere für multizentrische Untersuchungen innerhalb der psychologischen Forschung und bietet neben der Farbmarkierung der Datensatzklassifikation, schnellen Visualisierungen und der Dokumentation von Fragebögen nach einer Visite die Möglichkeit von Datenvalidierungen

²¹ Eine Einführung zu Elektronischen Laborbüchern im Forschungsdatenmanagement bietet das ZBMed – Informationszentrum Lebenswissenschaften, Vgl. ZB MED (Hrsg.) (2019): Elektronische Laborbücher im Kontext von Forschungsdatenmanagement und guter wissenschaftlicher Praxis – ein Wegweiser für die Lebenswissenschaften, Köln. Verfügbar unter: <https://repository.publisso.de/resource/frl:6415887/data>.

²² Nähere Informationen zum Dienst „PsychNotebook“ unter https://escience.aip.de/ak-forschungsdaten/wp-content/uploads/2018/06/JupyterLab_AK_Forschungsdaten_20180612.pdf.

innerhalb der Fragebögen oder automatische Berechnungen zu Summencores. Das Programm kann von verschiedenen Orten auch mit mobilen Endgeräten verwendet werden.²³

Frei nutzbar ist die Software für Redcap-Konsortien-[Mitglieder](#), über deren Server die Software genutzt werden kann. Für eine Mitgliedschaft kann ein Endbenutzer*innen-Lizenzvertrag zwischen der jeweiligen Forschungsinstitution und der Vanderbilt-Universität geschlossen werden.

1.6.3 Rspace

Das elektronische Laborbuch des Unternehmens [Rspace](#) ist für alle Datentypen konzipiert und wird neben einer frei zugänglichen Version als zugeschnittene kostenpflichtige Softwarelösung angeboten. Es ist auf verschiedene Zielgruppen akademischer Institutionen wie z.B. Forscher*innen, IT-Personal oder Arbeitsgruppen- oder Projektleiter*innen, PI's sowie auf Personal aus der Industrie ausgerichtet. Rspace ist nach FAIR-Prinzipien entwickelt und kann mit weiteren FDM- bzw. Forschungsdiensten und -infrastrukturen vernetzt werden. In [Case Studies](#) z.B. aus dem Center for Molecular Medicine werden Anforderungen und u.a. praxisbezogene Vorteile des Dienstes vorgestellt. Möglich ist eine [Demo](#)-Anfrage.

1.6.4 eLabFTW

Das Programm [eLabFTW](#) von Deltablot ist ein elektronisches webbasiertes Open-Source-Laborbuch, das individuell und kostenlos auf dem eigenen Server installiert und genutzt werden kann oder kostenpflichtig via Daten-Hosting über Deltabot verwendbar ist. Das Programm ist generischer Natur und eignet sich für alle Datentypen. Eine live [Demo](#)-Version bietet einen praktischen Einblick in das elektronische Laborbuch.

1.7 Datenschutz & Ethik

Forschungsdaten müssen in der Regel ab dem Zeitpunkt der Erhebung für 10 Jahre archiviert werden.²⁴ Darüber hinaus verlangen Drittmittelförderer oder institutionelle Forschungsdaten-Policies nach dem Grundsatz „so offen wie möglich, so geschlossen wie nötig“ zunehmend die Zugänglichmachung oder Veröffentlichung von Forschungsdaten. Gerade bei der Erhebung personenbezogener Daten in den Kognitions- und Bildungswissenschaften ergibt sich jedoch oft eine Spannung zwischen Schutz der Probanden und Veröffentlichung der Daten. Insbesondere im schulischen Kontext handelt es sich aufgrund der Zielgruppe von Kindern und Jugendlichen um sensible Daten, die oft ethische und datenschutzrechtliche Abklärungen verlangen.

1.7.1 Übersicht zur Erhebung personenbezogener Daten

Im Rahmen des „FDNext“-Projektes wurde in Zusammenarbeit mit der Europa-Universität Viadrina Frankfurt (Oder) und der Freien Universität Berlin eine Handreichung zur Erhebung personenbezogener Forschungsdaten erstellt.²⁵ Diese fußt insbesondere auf der Handreichung zum Datenschutz des RatSWD²⁶ und bietet Nachwuchswissenschaftler*innen oder Einsteiger*innen in das Thema FDM grundlegende Datenschutzinformationen sowie forschungsethische Hinweise noch vor Durchführung der Datenerhebung. Die Übersicht widmet sich folgenden zentralen Inhalten bzw. Fragen bei der Erhebung personenbezogener Daten:

²³ Die Informationen entstammen einer Bewertungsumfrage u.a. zu diesem Dienst im Fachbereich der Psychologie.

²⁴ Vgl. Kodex „Leitlinien zur Sicherung guter wissenschaftliche Praxis“ der Deutschen Forschungsgemeinschaft“ (2019), <https://doi.org/10.5281/zenodo.3923601>.

²⁵ Vgl. Kienbaum, Janna; Fischer, Patryk; Paßmann, Sven. (2023). Forschungsdatenmanagement bei personenbezogenen Daten – eine Handreichung. 10.5281/zenodo.7428524.

²⁶ Vgl. <https://www.konsortswd.de/aktuelles/publikation/handreichung-datenschutz-2/>.

1. Welche Regelungen sollten Sie im Datenschutzrecht beachten?
2. Welche personenbezogenen Daten erheben Sie?
3. Anonymisierung: Ist die Nachnutzung Ihrer Daten zulässig?
4. Rechtsgrundlagen der Verarbeitung personenbezogener Daten
 - a. Einwilligung
 - b. Das überwiegende wissenschaftliche Interesse
5. Datenerhebungen an Schulen
6. Ethikantrag: Wann und wie muss ein Ethik-Antrag formuliert werden?
7. Veröffentlichung der Daten: Welche Zugriffsklassen stehen Ihnen zur Verfügung (siehe Abschnitt 4.4)

1.7.2 QualiAnon

Der Anonymisierungsdienst [QualiAnon](#) vom Forschungsdatennzentrum „Qualiservice“ der Universität Bremen“ ermöglicht Ihnen softwaregestützt das Markieren und Replacement von Identifikatoren qualitativer Textdaten: „QualiAnon kann Informationen markieren und unter Verwendung unterschiedlicher Abstraktionsniveaus ersetzen. Dabei können standardisierte Listen verwendet werden und die Ersetzungen können fall- und studienbezogen vorgenommen werden. QualiAnon dokumentiert die Ersetzungen und ermöglicht unterschiedlich sensible Exporte.“²⁷ Eine [Präsentation](#)²⁸ von Kati Mozygamba und Tom Nicolai sowie eine [Gebrauchsanweisung](#) von Kati Mozygamba et al. verschaffen einen Eindruck vom Dienst und verhelfen bei dessen Einrichtung.

1.7.3 De-Identifizierungsmethoden

In Bezug auf MRT-Daten wird empfohlen, neben der Vergabe von Dateinamen ohne Personenbezug auch die Datenanteile in den Aufnahmen zu entfernen, die für die Forschungsfrage keine Relevanz haben (Knochenstruktur, extra-kraniales Gewebe). Dafür gibt es zwei weit verbreitete Ansätze, das sogenannte *skull stripping* (Abisolieren des Schädels) und das *defacing* (Entfernen von Gesichtsstrukturen. Für diese De-Identifizierung sind in Analyse-Applikationen bereits Möglichkeiten entwickelt worden (z.B., `ft_defacevolumes` für Fieldtrip). Alternativ kann dafür speziell entwickelte, open-source Software genutzt werden (z.B., `MRI_deface`, `Pydeface`). Um die (Meta-)Daten in den Aufnahmen zu entfernen, die auch zur Identifizierung spezifischer Proband*innen (Geschlecht, Alter, Name usw.) benutzt werden können, gibt es ebenfalls frei verfügbare Software (z.B., die De-Identification Toolbox oder GDCM):

Vorstellung der [De-Identifizierungsmethoden](#) am Beispiel der Quickshear Defacing method: Schimke, N., Kuehler, M., & Hale, J. (2011). Preserving privacy in structural neuroimages. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6818 LNCS, 301–308.

- De-Identifizierung mit [Fieldtrip](#) (MATLAB software toolbox)
- De-Identifizierung `mri_deface` mit [MiDeFace](#) der Free Surfer Software und [Pydeface](#)
- De-Identification [Toolbox](#) (Programm in Java)
- [GDCM](#) zur Entfernung von Metadaten.²⁹

²⁷ Vgl. <https://www.qualiservice.org/de/news/qualiservice-anonymisierungstool-qualianon-nutzbar.html>.

²⁸ Vgl. Video zum Vortrag und zur Präsentation der Folien unter <https://www.youtube.com/watch?v=MbKw3LR2rVo>.

²⁹ Der Abschnitt zur De-Identifizierung ist in Absprache von dem FDNext-Arbeitspaket 5 „Konzeption von Train-the-Trainer-Workshops für Psycholog*innen“ übernommen worden, Vgl. Paßmann, Sven; Söring, Sibylle. (2023). Forschungsdatenmanagement in der Psychologie: Fachspezifisches Train-the-Trainer-Konzept. DOI: 10.5281/zenodo.7560900.

2. Daten erheben/Daten und Instrumente nachnutzen

Disziplinspezifische Forschungsdatenzentren (FDZ), Repositorien oder Datenarchive bieten nicht nur die Möglichkeit, Datensätze zu archivieren und/oder publizieren (siehe Abschnitt 4), sondern auch bereits qualitätsgeprüfte qualitative sowie quantitative Datensätze für eine Sekundärnutzung oder Reproduzierbarkeit zugänglich zu machen. Dazu zählen ebenso Messinstrumente und Erhebungsmethoden. Die Repositorien bzw. Datenarchive sind meist öffentlich und kostenlos nutzbar. Sie ermöglichen die Vergabe von Digital Object Identifier (DOI), anhand derer Daten und Forschungsmaterialien auffindbar und zitierbar werden. Sie ermöglichen die Auffindbarkeit der Forschungsstudien durch standardisierte Metadatenangaben. Die softwaregestützte Ablage bzw. Zugänglichmachung der Forschungsdaten und des Forschungsmaterials wird zudem von Beratungsservices der dahinterstehenden FD-Infrastruktur begleitet. Für die Beantwortung einer Forschungsfrage sollte vorerst ermittelt werden, ob *bereits publizierte* Datensätze für diese bestehen oder herangezogen werden können. Die Nutzung frei zugänglicher Daten fördert sowohl Austausch als auch Transparenz in der Wissenschaft.

Die folgende Auswahl zentraler Repositorien und Archive von mitunter vom RatSWD akkreditierten³⁰ Forschungsdatenzentren bieten den Zugang zu Daten bzw. Instrumenten in den Bildungswissenschaften und der Psychologie. Weitere [Datenzentren](#) können über die Suchmaske des NFDI-Konsortiums „KonsortSWD“ der Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften gefunden werden.

2.1 Psychologie

2.1.1 PsychArchives

Das Repositorium [PsychArchives](#) des Leibniz-Instituts für Psychologie (ZPID) unterstützt die Archivierung von und den Zugang zu Datensätzen und diversen weiteren digitalen Forschungsobjekten (*Digital Research Objects*, DROs). Insgesamt 20 verschiedene DROs können über PsychArchives archiviert und für die Nachnutzung zugänglich gemacht werden. Dazu zählen z.B. Artikel, Vorabdrucke, Forschungsdaten, Codes, Ergänzungen, Prä-Registrierungen, Tests und Multimedia-Objekte.

2.1.2 FDZ am Leibniz-Institut für Psychologie (ZPID)

Das [Forschungsdatenzentrum](#) des Leibniz-Instituts für Psychologie (ZPID) stellt psychologische Forschungsdaten für die Nachnutzung durch die wissenschaftliche Gemeinschaft bereit und unterstützt Forschende bei einem qualitätsgesicherten FDM während des gesamten Forschungsprozesses. Der Zugang zu den Forschungsdaten wird grundsätzlich über einen Nutzungsvertrag/eine Nutzungslizenz geregelt. Das FDZ am ZPID ist vom RatSWD akkreditiert.

2.1.3 Open Neuro

Das frei zugängliche Repositorium [Open Neuro](#) der Stanford University bietet Zugang für Daten aus den kognitiven Neurowissenschaften. Schwerpunkt bilden Studien mit Daten der funktionellen Magnetresonanztomographie (fMRI-Daten). Die Ablage der Daten orientiert sich an der Brain Imaging Data Structure (BIDS). Neben fMRI-Daten umfasst das Repositorium EEG-, iEEG, PET- und MEG-Datensätze.

2.1.4 Open Science Framework (OSF)

Die Plattform [Open Science Framework](#) (OSF) ist ein disziplinübergreifendes Angebot des *Center for Open Science* (COS), kommt jedoch stark in der psychologischen Forschungscommunity zum Einsatz. OSF bietet die Möglichkeit, offene Datensätze zu finden sowie selbst erhobene Daten

³⁰ Vgl. <https://www.konsortswd.de/datenzentren/akkreditierung/>.

während oder nach dem Forschungsprozess zu dokumentieren, zu publizieren und in Kooperationen mit internationalen Forscher*innen zu treten. OSF ist flexibel nutzbar und bietet sich insbesondere für das Managen oder Arbeiten mit Forschungsdaten während einer Studie an. Eine Qualitätssicherung seitens des Betreibers sowie eine spezifische standardisierte Struktur der Datendokumentation der Daten werden nicht gegeben. Die Sicherstellung der Langzeitarchivierung der Forschungsdaten ist bei OSF an deren privatrechtliche Finanzierungsbedingungen gebunden.

2.1.5 Open Test Archive

Das Repositorium [Open Test Archive](#) des Leibniz-Instituts für Psychologie (ZPID) stellt Open-Access-Tests zur Verfügung und unterstützt Forscher*innen bei der Publikation ihrer urheberrechtlich geschützten Testinstrumente. Die Tests sind neben der Forschung in der Lehre und Praxis einsetzbar.

2.2 Bildungswissenschaften

2.2.1 Das Nationale Bildungspanel (NEPS)

Das FDZ des [Nationalen Bildungspanels](#) (NEPS) am Leibniz-Institut für Bildungsverläufe e.V. (LI fBi) berücksichtigt die Aufbereitung und Bereitstellung der Daten aus dem Nationalen Bildungspanel: Es behandelt darin u.a. empirische Analysen, Scientific Use files sowie die Erhebung längsschnittlicher Daten zu Bildungsprozessen und zur Kompetenzentwicklung.

2.2.2 Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Das an der Humboldt-Universität zu Berlin verortete FDZ am [Institut zur Qualitätsentwicklung im Bildungswesen](#) (IQB) archiviert und dokumentiert Datensätze und Begleitmaterialien insbesondere aus der quantitativ ausgerichteten empirischen Bildungsforschung. Dazu gehören die Fachbereiche der Erziehungswissenschaft, pädagogischen Psychologie, Bildungssoziologie, Bildungsökonomie und Fachdidaktik. Den Kern bilden zugängliche Datensätze der nationalen und internationalen Schulleistungsstudien in Deutschland für Re- und Sekundäranalysen. Ein Fokus wird u.a. auf Studien zur Kompetenzmessung von Schüler*innen gelegt. Das FDZ ist vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditiert.³¹

2.2.3 FDZ-Bildung

Das [FDZ-Bildung](#) vom Leibniz-Institut für Bildungsforschung und Information (DIPF) bietet Zugang zu Studien aus der qualitativen empirischen Erziehungs- und Bildungsforschung und umfasst neben Datensätzen auch Erhebungsinstrumente der quantitativen Forschung (Fragebögen, Skalen und Tests). Der Schwerpunkt der Instrumentensammlung und Datensets liegt auf der Schulqualitäts- und Unterrichtsqualitätsforschung. Das FDZ ist an der Erhebung von Daten großer Studien wie NEPS oder PISA beteiligt. Es ist vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditiert.

2.2.4 Datenarchiv „Kindheit und Jugend im urbanen Wandel“

Das Datenarchiv [Kindheit und Jugend im urbanen Wandel](#)³² des Methodenzentrums Qualitative Bildungsforschung der Universität Duisburg-Essen umfasst qualitative Studien der Kindheit und Jugendforschung (Datenmaterial seit den 1970ern). Der Fokus liegt auf Forschungsvorhaben aus den Standorten in Nordrhein-Westfalen. Zum archivierten qualitativen Datenmaterial gehören u.a.

³¹ Informationen zur Akkreditierung vom RatSWD Vgl. <https://www.konsortswd.de/datenzentren/akkreditierung/>.

³² Aufgrund technischer Störungen führen einzelne Links derzeit auf Fehlermeldungen oder leere Seiten. Hinweise oder Fragen richten Sie bitte an folgende temporäre Adresse: fachportal-pedocs@dipf-institut.de.

Audiomitschnitte von Interviews, Videoaufzeichnungen, Fotografien, Beobachtungsprotokolle, Zeitdokumente und Transkripte.

2.2.5 Archiv für pädagogische Kasuistik (Apaek)

Das [Archiv für pädagogische Kasuistik](#) (Apaek) der Goethe Universität Frankfurt am Main bietet Zugang zu Daten der qualitativen erziehungswissenschaftlichen Forschung mit einem Schwerpunkt in der Schul- und Unterrichtsforschung sowie der rekonstruktiven Bildungsforschung. Es richtet sich an Lehrende und Studierende, die sich in kasuistischer Lehre fallorientiert mit pädagogischer Praxis auseinandersetzen.

2.2.6 Fachportal Pädagogik

Das [Fachportal Pädagogik](#) des DIPF | Leibniz-Instituts für Bildungsforschung und Bildungsinformation bietet einen umfassenden, filterbaren Überblick über Studien und Daten aus verschiedenen Datenzentren der Erziehungswissenschaft, Bildungsforschung und Fachdidaktik.

2.2.7 Instrumentendatenbank

Neben Datensätzen gibt es die [Instrumentendatenbank](#) des FDZ-Bildung vom Leibniz-Institut für Bildungsforschung und Information (DIPF). Diese stellt dokumentierte quantitative Instrumente wie Skalen, Fragebögen und Tests zur Verfügung.

2.3 Sozialwissenschaftliche (Umfrage-)Daten

2.3.1 SowiDataNet|datorium

Das Repository [SowiDataNet|datorium](#) des Leibniz Instituts für Sozialwissenschaften (GESIS) bietet Forscher*innen die Archivierung von sowie den Zugang zu quantitativen Primärdaten und Materialien der sozialwissenschaftlichen Forschung an. Die Daten werden standardisiert dokumentiert und aufbereitet (z.B. Bereinigung, Variablenreports).

2.3.2 Qualiservice

Das FDZ [Qualiservice](#) der Universität Bremen kümmert sich um die Archivierung und Bereitstellung nachnutzbarer, qualitativer Forschungsmaterialien aus der Sozialforschung. Dazu zählen Datensätze multidisziplinärer und themenungebundener Studien wie Interviewstudien, Audio- und Bilddateien oder auch Social-Media-Daten. Qualiservice gehört zu den vom RatSWD akkreditierten Forschungsdatenzentren.

2.3.3 CESSDA Data Catalogue

Der [CESSDA Data Catalogue](#) vom *Consortium of European Social Science Data Archives* beinhaltet über 40.000 europaweite Datensätze der Sozialwissenschaften. Anfragen zu Datensätzen aus den im CESSDA Data Catalogues versammelten Repositorien können u.a. über Schlagworte des disziplinspezifischen Forschungsfeldes gestellt werden. Die Datenbank bietet Zugänge zu quantitativen, qualitativen oder gemischten Daten, Querschnitt- oder Längsschnittdaten, kürzlich erhobenen oder historischen Daten.

2.3.4 Repository für Messinstrumente ZIS

Das Repository für [Mess- und Erhebungsinstrumente](#) (ZIS) des Leibniz Institut für Sozialwissenschaften (GESIS) bietet englisch- und mehrsprachige Erhebungs- und Messinstrumente aus der Umfrageforschung in den Sozial- und Verhaltenswissenschaften an. Dazu gehören Items, Skalen, Fragebögen, Tests, Indexe.

3. Daten verarbeiten und dokumentieren

Der Schritt der Aufbereitung und Dokumentation der Daten im Rahmen des Forschungsvorhabens gehört zum grundlegenden Bestandteil der Methodenlehre des jeweiligen Studienfachs.³³ Aus Sicht der FAIR Data-Prinzipien und Open Science-Praktiken geht es hier nicht um die Erläuterung oder Inhalte fachspezifischer Erhebungs- oder Analysemethoden, sondern um die Beschreibung der Rahmenbedingungen bei der Auswahl, Güte und der Form der Durchführung. Im Einklang mit der „Guten Wissenschaftlichen Praxis“ steht also die *nachvollziehbare* Struktur der technischen und kontextuellen Datenverarbeitung und -analyse im Vordergrund, die eine Nachnutzung der Primär- bzw. Sekundärdaten angesichts der Forschungsfrage ermöglicht. Dazu gehört neben standardisierten, offenen Dateiformaten und Softwareprogrammen insbesondere die Dokumentation und Kontextualisierung des Datenerhebungs- und -analyseprozesses. Denn damit Forschungsdaten besser auffindbar und nachvollziehbar sind, ist eine Dokumentation der Daten als Basis wissenschaftlicher Arbeiten unverzichtbar.³⁴ Zentral wird dies naheliegenderweise bei Datenpublikationen. Hilfreich ist hier die Checkliste zum Umgang mit Forschungsdaten der DFG, die mittlerweile als verbindlich bei der Antragsstellung gilt.³⁵

Gut dokumentierte Daten werden häufiger genutzt und zitiert, was die Reputation des/der Ersteller*in erhöht. Auch für die eigene Nachnutzung und Nachvollziehbarkeit der Daten ist eine Dokumentation notwendig und gilt als „Sprachrohr“ zwischen Datenerzeuger*in und Datennutzer*in. Mit der Zeit können Details in Vergessenheit geraten, daher ist es empfehlenswert, die Dokumentation der Daten arbeitsbegleitend zu betreiben. Datendokumentationen sollten möglichst nicht lokal, sondern in *Cloud- oder Serverspeichern* Ihres Fachbereichs abgespeichert werden, um eine Nachvollziehbarkeit der Studie auch bei einem potentiellen Wechsel des Forschungspersonals zu garantieren.

Zu den allgemeinen Dokumentationsformen gehören z.B. ReadMe-Dateien³⁶, Versionierungen, Data Dictionaries³⁷, Codebücher, elektronische Laborbücher, Methodenberichte, Studienreports oder Datenpublikationen.

3.1 Standardformate

Für die Langzeitarchivierung sowie Nachnutzung von Softwareprogrammen und Rohdaten eignen sich disziplinspezifische standardisierte Dateiformate und Datenablagestrukturen.

Bei Dateiformaten können proprietäre von freien und offene von nicht offenen Dateiformaten unterschieden werden. Für die Archivierung und Publikation von Forschungsdaten sind vor allem offene Dateiformate oder (proprietäre) Formate zu empfehlen, die eine hohe Popularität innerhalb der

³³ Da in den Bildungswissenschaften und der Psychologie Umfragen eine zentrale Methodologie zur Klärung von Forschungsfragen darstellen, hat das Leibniz-Institut für Sozialwissenschaften (GESIS) Leitfäden zu diversen methodischen Aspekten in der Umfrageforschung erstellt. Sie unterteilen sich in die Themenbereiche Befragungsinstrumente, -statistiken, -operationen und Dokumentation, siehe unter <https://www.gesis.org/gesis-survey-guidelines/home>.

³⁴ Zu den wichtigsten Inhalten einer Datendokumentation, Vgl. die Übersichtsseite „Datendokumentation: Warum, was und wie?“ von forschungsdaten.info unter: <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>.

³⁵ Vgl. DFG (2021): Umgang mit Forschungsdaten: Checkliste für Antragstellende zur Planung und zur Beschreibung des Umgangs mit Forschungsdaten in Forschungsvorhaben: https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf.

³⁶ Vgl. generisches Readme-Template unter <https://zenodo.org/record/4037151/files/06-Readme-Template-Online-Workshop-Datendokumentation.docx?download=1>.

³⁷ Vgl. generisches Data Dictionary-Template unter <https://zenodo.org/record/4037151/files/05-Data-Dictionary-Template-Online-Workshop-Datendokumentation.docx?download=1>.

Methodennutzung einer Disziplin und deren Kompatibilität in alternative Dateiformate besitzen. Für statistische Erhebungen bieten sich beispielsweise die Programme SPSS, R, SAS und STATA an.³⁸

Zu standardisierten Ablagestrukturen zählen datenspezifische Ordnerhierarchien, Dateibenennungen oder auch die Angabe von Metadaten. Die Datensätze innerhalb des Arbeitsservers oder Repositoriums sind beispielsweise nach Weggang einer Wissenschaftlerin/eines Wissenschaftlers an der Universität oder für externe Forscher*innen nachvollziehbar.

3.1.1 Use Case: Dateiformate

In der psychologischen Forschung können Dateiformate der produzierten Forschungsdaten von entsprechenden Programmen der Messinstrumente abhängen. Bestimmte Softwareprogramme beispielsweise zur Messung der Herzfrequenzvariabilität (HRV) erzeugen Rohdaten, die geschlossen bzw. proprietär vorliegen und an die Nutzung des Programms gebunden sind. Die Rohdaten liegen dann bereits bereinigt oder aggregiert vor und können nicht als offenes Dateiformat (z.B. .csv) abgespeichert werden. Das Leibniz-Institut für Psychologie empfiehlt in solch einem Fall, die Rohdaten dennoch im proprietären Format in einem Repositorium oder Datenarchiv bereitzustellen, dabei jedoch alle Datenverarbeitungsschritte, die Kriterien der Bereinigung oder Aggregation sowie die benötigte Software zu dokumentieren.

3.1.2 Dateiformate bei PsychArchives

Die für psychologische Studien einzuhaltenden [Dateiformate](#) werden vom Leibniz-Institut für Psychologie im Rahmen ihres Dienstes „PsychArchives“ angegeben. Sie stellen einen allgemeingültigen Überblick nachnutzbarer Formate dar und sind zugleich Bedingung, um Datensätze im „hauseigenen“ Repositorium des ZPIDs zu archivieren oder publizieren.

3.1.3 Dateiformate des Verbund FDB

Der Verbund Forschungsdaten Bildung (Verbund FDB) empfiehlt ebenfalls allgemeingültige [Dateiformate](#) für die Sekundärnutzung und unterteilt diese in schwerpunktmäßig quantitative oder qualitative Forschungsvorhaben.

3.1.4 Brain Imaging Data Structure (BIDS)

Der Standard [Brain Imaging Data Structure](#) (BIDS) zur Ablagestruktur und Dateibenennung ist speziell auf neurophysiologische Daten (fMRI- und Verhaltensdaten) zugeschnitten. Datenbanken wie OpenNeuro oder das FDZ des Leibniz-Instituts für Psychologie importieren und exportieren Daten mit diesem Standard. Neben fMRI-Daten bezieht sich der Standard u.a. auf MRI, iEEG und Verhaltensdaten. Informationen zur Anwendung bietet eine [Anleitung](#). Aktuell wird zusätzlich an einer BIDS-Extension für [Eyetracking-Daten](#) gearbeitet.

3.1.5 Psych-DS

Der Dateistandard [Psych-DS](#) ermöglicht die für Menschen und Maschinen lesbare Nachnutzung von psychologischen sozial- sowie verhaltenswissenschaftlichen Datensätzen. Er bezieht sich aufbauend auf BIDS auf Ordnerstrukturen, Dateinamen, Dateiformate und Metadaten-Formate. Der Standard berücksichtigt dabei die getrennte Ablage von Tabellen-Datensätze (.csv) und der dazugehörigen Metadaten-Datei (.json).

³⁸ Vgl. Website zu „Formate erhalten“ von forschungsdaten.info unter <https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/formate-erhalten/>.

3.2 Aufbereitung und Dokumentation quantitativer Forschungsdaten

Leitfäden und Checklisten helfen Forscher*innen der Psychologie und der Bildungswissenschaften schon während der Planung ihres quantitativen Forschungsvorhabens oder ihrer Datenaufbereitung dabei, die Studie zu dokumentieren. Die Dokumente fragen dazu u.a. die Herleitung der Forschungsfrage, die Vorgehensweise bei der Datenerhebung und der Zielgruppenauswahl oder die Auswahl von Messinstrumenten systematisch ab.

Im Folgenden sind Dienste und insbesondere Handreichungen in Form von Literaturangaben aufgelistet, die sich auf quantitativ erhobene Forschungsdaten konzentrieren und auf die zentralen Schritte der Studien- bzw. Datendokumentation während und nach der Verarbeitung der Daten hinweisen.³⁹

3.2.1 Use Case: Datendokumentation 1

*Ein Forschungsteam aus Soziolog*innen und empirischen Bildungsforscher*innen plant eine quantitative Teilstudie zu Erfolgen aber auch auszubauenden Arbeitsbereichen in Ganztagschulen. Die Studie beinhaltet Interviews mit Lehrkräften mithilfe von Fragebögen, die statistisch ausgewertet werden sollen. Mit dem Ziel, die Forschungsergebnisse und die erhobenen Daten in anonymisierter Form (siehe Abschnitt 1.7.1) in einem Forschungsdatenrepositorium zu archivieren und für Dritte zur Sekundärnutzung bereitzustellen, beschäftigt sich das Team bereits frühzeitig mit der Dokumentation ihrer Studien. Die Dokumentation der Erhebung ist nicht zuletzt deshalb für die Wissenschaftler*innen relevant, da ihre Erhebung Teil einer Längsschnittstudie ist, die einen Vergleich der Ergebnisse mit zukünftigen periodischen Erhebungswellen vorsieht. Um Transparenz und Nachnutzbarkeit der Forschungsergebnisse, Daten und Messinstrumente zu garantieren, sucht das Forschungsteam deshalb noch vor der Datenerhebung nach Handreichungen zur Datendokumentation, die die zentralen Schritte ihrer Studie und ihrer Datenverarbeitung berücksichtigen.*

3.2.2 Aufbereitung quantitativer Forschungsdaten

Eine allgemeine Empfehlung zur [Aufbereitung quantitativer Forschungsdaten](#) bietet der Verbund Forschungsdaten Bildung, um deren Sekundärnutzung zu fördern: Vgl Trixa, Jessica; Ebel, Thomas; Harzenetter, Karoline: Hinweise zur Aufbereitung quantitativer Daten. Version 1.3. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation 2019, 11 S. - (forschungsdaten bildung informiert; 4) - URN: urn:nbn:de:0111-pedocs-219671 - DOI: 10.25656/01:21967.

3.2.3 Fehlende Werte in der Aufbereitung

Hinweise zur Beschreibung [fehlender Werte](#) während der Datenaufbereitung bietet der Verbund Forschungsdaten Bildung: Hinweise zur Codierung fehlender Werte in der Aufbereitung quantitativer Daten. Version 1.0. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation 2019, 7 S. - (fdb info; 6) - URN: urn:nbn:de:0111-pedocs-223032 - DOI: 10.25656/01:22303.

3.2.4 Open Refine

Eine Anleitung zur Open Source-Software [OpenRefine](#) wird von der Data Carpentry-Unterrichtseinheit angeboten. Der Einsatz von OpenRefine fördert die Harmonisierung und Bereinigung unstrukturierter Daten.

³⁹ Die gelisteten Handreichungen und Empfehlungen entstammen zu einem Großteil der Website vom Verbund FDB über die Aufbereitung und Dokumentation quantitativer Datenerhebungen. Siehe: <https://www.forschungsdaten-bildung.de/quantitative-daten#Quantitative-Daten-dokumentieren>.

3.2.5 Syntaxen

Eine Hilfe zur Erstellung von [Syntaxen](#) bei quantitativ-statistischen, sozialwissenschaftlichen Ergebnissen wird am Beispiel des SowiDataNet|datorium von GESIS gegeben: Vgl Ebel, Thomas, 2016: Einreichungen von Syntaxen in datorium (Replikationsserver), Stand: 2016-02-29, GESIS Datenarchiv für Sozialwissenschaften.

3.2.6 Methoden- und Feldbericht

Einen umfassenden stichpunktartigen [Leitfaden](#), der Angaben zu Projekt, Erhebungsinstrumenten, Datenerhebungsinstitut, Stichprobe, Feldarbeit, Datenaufbereitung und Datensatz der quantitativen Studie abfragt, wurde vom Verbund Forschungsdaten Bildung erstellt: Vgl. Verbund Forschungsdaten Bildung (2015): Hinweise zur Erstellung eines Methoden-/Feldberichts. Version 1.1. fdbinfno Nr. 3.

3.2.7 Methodenbericht am Bsp. der StEGStudie

Ein Beispiel eines ausgefüllten [Methodenberichts der „Studie zur Entwicklung von Ganztagschulen \(StEG\)“](#) mit Online-Umfragen als Erhebungsmethode bietet Furthmüller, Peter (2019): Methodenbericht. Systemmonitoring 2018: Stichprobe und Erhebungsverfahren: Studie zur Entwicklung von Ganztagschulen. In Zusammenarbeit mit dem StEG-Konsortium und den Mitarbeiterinnen und Mitarbeitern des StEG-Teams.

3.2.8 Checkliste eines Methodenberichts am Bsp. Der ALLBUS-Bevölkerungsumfrage

Als Orientierung der Datendokumentation dient eine Checkliste für einen Methodenbericht mit Erläuterungen (Konzept, Sampling, Datenerhebung und Datenaufbereitung) am [Beispiel der ALLBUS-Bevölkerungsumfrage](#): Watteler, Oliver (2010): Erstellung von Methodenberichten für die Archivierung von Forschungsdaten, GESIS Leibniz-Institut für Sozialwissenschaften.

3.2.9 Dokumentationsleitfaden am Bsp. von „mail surveys“

Einen Leitfaden mit Fokus auf die [Dokumentation analoger Umfragen \(mail surveys\)](#) stellt ebenfalls GESIS zur Verfügung: Stadtmüller, S., & Beuthner, C. (2020). Documentation of mail data collection (Version 1.0). (GESIS Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/gesis-sg_en_032

3.2.10 Dokumentationsleitfaden am Bsp. von Fragebogenübersetzungen

Werden Fragebögen in verschiedenen [Sprachen](#) erstellt, kann ein Leitfaden zur Dokumentation von Fragebogenübersetzungen als Orientierung dienen: Behr, D., & Zabal, A. (2020). Documenting Survey Translation. Mannheim, GESIS – Leibniz-Institute for the Social Sciences (GESIS – Survey Guidelines).

3.2.11 Checkliste zum Dokumentieren am Bsp. von Online-Umfragen

Über Fragenabschnitte zum Ankreuzen wird eine Dokumentation von speziell [Online-Umfragen](#) angeboten: Ines Schaurer, Kunz, Tanja, & Heycke, Tobias (2020). Documentation of online surveys. Mannheim, GESIS- Leibniz Institute for the Social Sciences (GESIS-Survey Guidelines).

3.3 Aufbereitung & Dokumentation/Kontextualisierung qualitativer Forschungsdaten

Datenerhebungen qualitativer Studien verlangen eine detaillierte und umfassende Kontextualisierung vom Rahmen der Datenerhebung und der darin stattfindenden Interaktionen.⁴⁰ Oft handelt es sich um Informationen, die nur von den Forscher*innen als „subjektive Eindrücke“ festgehalten werden, jedoch für die Nachvollziehbarkeit der Daten für Dritte umfassend dokumentiert werden sollten.⁴¹ Neben den Anonymisierungsprinzipien gehört zu der Datenaufbereitung also auch die Kontextualisierung der qualitativen Daten. Die Dokumentation bzw. Kontextualisierung qualitativer Datenerhebungen umfasst z.B. die Eigenschaften der Erhebungssituation und der Proband*innen, die situative Stimmung einer Gesprächsatmosphäre sowie genaue Angaben zum Anonymisierungs- oder Transkriptionsvorgang berücksichtigen.⁴² Folgende als Literaturangaben gelisteten Handreichungen oder Leitfäden dienen der Unterstützung zur Dokumentation qualitativer Datenerhebungen.⁴³

3.3.1 Use Case: Datendokumentation 2

*Ein Forschungsteam aus Bildungsforscher*innen, Soziolog*innen und Kulturwissenschaftler*innen hat qualitative ethnologische Unterrichtsbeobachtungen an Schulen zur Ermittlung interkulturellen Lernens durchgeführt. Dafür nahmen die Datenerhebenden über einen bestimmten Zeitraum mittels der Methode der „teilnehmenden Beobachtung“ am Mittelstufenunterricht von Gymnasien teil, die einen hohen Anteil von Schüler*innen mit Migrationshintergrund besitzen. Als Grundlage dienten Beobachtungsleitfäden, nach denen die Unterrichtsstunden protokolliert wurden. Zusätzlich wurden qualitative mündliche Interviews mit den Lehrkräften der beobachteten Klassen durchgeführt, die über Audiogeräte aufgenommen und anschließend transkribiert wurden. Sowohl Protokolle als auch Interviews wurden dabei von Ereignissen geprägt, die letztlich eine Auswirkung auf die erhobenen Daten hatten. Die Datenerhebung in den Unterrichtsbeobachtungen wurde beispielsweise durch einen hohen Krankenstand der Schüler*innen, durch einen spontanen Raumwechsel der Klasse sowie durch die Verlegung der Beobachtung von der ersten in die letzte Unterrichtsstunde beeinflusst. Bei der Datenerhebung durch die Lehrer*innen-Interviews mussten zum Beispiel Personenwechsel bei den Proband*innen aufgrund spontaner Absagen oder unterschiedliche Erfahrungsstufen der Lehrkraft berücksichtigt werden.*

Mit dem Ziel nicht zuletzt der Anforderung ihres Drittmittelförderers, die eine Publikation der Forschungsergebnisse und Datensätze in einem Forschungsdatenzentrum vorsieht, gerecht zu werden, beschäftigte sich das Forschungsteam bereits zu Beginn der Studie mit den Inhalten einer Dokumentation, die es für die Transparenz und Nachnutzung der Daten und Forschungsergebnisse bedarf. Es suchte nach Handreichungen oder Empfehlungen, die alle notwendigen Schritte des Forschungsvorhabens und aller zu kontextualisierenden Informationen vor, während und nach der Datenerhebung abfragen, um diese für eine potentielle spätere Sekundäranalyse nachvollziehbar zu machen.

3.3.2 Studienreport

Ein Leitfaden zur Erstellung eines umfassenden [Studienreports](#), der auch Angaben zum Aufbau des Forschungsprojekts und zu den Ergebnissen berücksichtigt, wird von Qualiservice angeboten: Heuer, Jan-Ocko; Kretzer, Susanne; Mozygemba, Kati; Huber, Elisabeth & Hollstein, Betina (2020): Kontextualisierung qualitativer Forschungsdaten für die Nachnutzung – eine

⁴⁰ Vgl. Verbund Forschungsdaten Bildung: Leitfaden zur Kontextualisierung von qualitativen Befragungsdaten. Version 1.2. Frankfurt am Main : DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation 2021, 10 S. - (fdb info; 9) - URN: urn:nbn:de:0111-pedocs-223868 - DOI: [10.25656/01:22386](https://nbn-resolving.org/urn:nbn:de:0111-pedocs-223868-10), S. 3.

⁴¹ Ebd.

⁴² Ebd.

⁴³ Die gelisteten Handreichungen und Empfehlungen entstammen zu einem Großteil der Website vom Verbund FDB über die Dokumentation bzw. Kontextualisierung qualitativer Datenerhebungen. Siehe: <https://www.forschungsdaten-bildung.de/qualitative-daten>.

Handreichung für Forschende zur Erstellung eines Studienreports. Qualiservice Working Papers 1-2020, Bremen, <http://dx.doi.org/10.26092/elib/166>.

3.3.3 Leitfaden am Bsp. einer Gruppendiskussion

Ein Leitfaden zur Kontextualisierung qualitativer Befragungen samt [Beispiel einer Gruppendiskussion](#) wird bereitgestellt vom Verbund Forschungsdaten Bildung (2019): Kontextualisierungsbogen von qualitativen Daten. Leitfaden zur Kontextualisierung für Befragungsmethoden. Version 1.0, fdbinfo Nr. 9.

3.3.4 Methoden- und Feldbericht

Eine [qualitative Version des Methoden- und Feldberichtes](#) vom Verbund Forschungsdatenbildung mit Angaben zu Projekt, Erhebungsmethode, Sampling, Feldarbeit, Datenaufbereitung und Datensatz wurde im Rahmen des FDNNext-Projektes in Zusammenarbeit mit dem DIPP | Leibniz-Institut für Bildungsforschung und Bildungsinformation entwickelt: Kienbaum, Janna, Würfl, Katja, & Favella, Gianpiero. (2022). Leitfaden für einen qualitativen Methoden- und Feldbericht. Zenodo. <https://doi.org/10.5281/zenodo.7247993>.

3.3.5 Studien- und Datendokumentation am Bsp. narrativer Interviews

Eine Studien- und Datendokumentation [am Beispiel von narrativen Interviews](#) kann als Orientierung zur Beschreibung von theoretischem Rahmen, methodischem Vorgehen und Datenaufbereitung in der qualitativen Bildungsforschung dienen: Steinhardt, I. & İköz-Akıncı, D. (2020). Digitale Bildungspraktiken von Studierenden. Daten- und Methodenbericht zur qualitativen Erhebung der Studie DEPS 2018. Version 1.0.0. Kassel/Hannover.

3.3.6 Dokumentation am Bsp. Von face-to-face-Befragungen

Fragenabschnitte zum Ankreuzen dienen der [Dokumentation von Datenerhebungen via persönlich-mündlicher Befragungen \(face-to-face\)](#): Schröder, Jette, & Zabal, Anouk (2021). Documentation of Face-to-Face Surveys. Mannheim, GESIS - Leibniz Institute for the Social Sciences (GESIS - Survey Guidelines).

3.3.7 Transkription

[Hinweise](#) zur Dokumentation von Transkriptionsregeln und einen Leitfaden zur Transkription von audiovisuellen Unterrichtsaufnahmen bietet die Marburger Arbeitsgruppe für Methoden und Evaluation: Kuckartz, Udo; Rädiker, Stefan: Transkriptionsregeln, Hinweise zur automatischen Transkription sowie eine [Handreichung](#) des Apaek der Goethe-Universität Frankfurt a.M.⁴⁴

3.3.8 Unterrichtsaufnahmen

Informationen zur Dokumentation für die [gerätegestützte Aufnahme](#) von Unterricht bietet Pollmanns, Marion; geändert von Jan Belak (Juni 2010), bereitgestellt vom Apaek der Universität Frankfurt.

3.3.9 Videoaufnahmen

Informationen zur organisatorischen und technischen Durchführung von [Videoaufnahmen](#) und deren Aufbereitung bieten Breitenbach, Sebastian; Herrle, Matthias (2016): „Planung, Durchführung und Nachbereitung videogestützter Beobachtungen im Unterricht.“ In: Engartner, Tim; Herrle, Matthias; Rauin, Udo: Videoanalysen in der Unterrichtsforschung. Methodische Vorgehensweisen und Anwendungsbeispiele, Beltz Juventa, S.30-49.

⁴⁴ Weitere Empfehlungen zur Transkription bietet die Website zum Thema „Qualitative Daten“ vom Verbund Forschungsdaten Bildung unter <https://www.forschungsdaten-bildung.de/qualitative-daten#Transkription>.

3.4 Codebuch

Für die Dokumentation der gemessenen Variablen in maschinenlesbarer Form bewähren sich Codebücher. Dafür haben sich zunehmend disziplinspezifische Templates aus der Psychologie und Bildungswissenschaft bewiesen. Eine disziplinübergreifende Anleitung zur Erstellung eines Codebuches (in formr) bietet das [Cookbook for a Codebook](#) von Kai T. Horstmann samt Beispielen.

3.4.1 DataWiz II (beta)⁴⁵

Der [DataWiz II \(beta\)](#) des Leibniz-Instituts für Psychologie bietet Forscher*innen die Möglichkeit, sowohl ihre Studie über standardisierte Metadaten zu verschlagworten, Supplement Materials anzugeben als auch ein Codebuch anhand einer vorgefertigten Tabelle zu erstellen. Die Angaben und das Codebuch können als .json oder .xml-Dateien exportiert werden. Eingaben der Studienbeschreibung können dabei mit der Codebuch-Erstellung verbunden werden.

3.4.2 Use Case: DataWiz

Im Rahmen ihrer Doktorarbeit untersucht eine Forscherin Lesereaktionszeiten mithilfe der sogenannten „selbstgesteuerten Lesemethode“ (self-paced reading task)¹. In ihrer Studie analysiert sie Gedächtnisinterferenzeffekte während des Lesens: In dieser testet sie, ob Verarbeitungsschwierigkeiten durch das Lesen semantisch ähnlicher Nomen entstehen. Die Studie gibt einen Einblick darüber, welche Mechanismen in der Satzverarbeitung (insbesondere von nicht-lokalen Abhängigkeiten) involviert sind. Die Doktorandin war bereits zu Beginn ihrer Studie auf der Suche nach einem Programm, das ihr bei der standardisierten Dokumentation ihrer Daten hilft, um diese später in einem Repositorium der Psychologie für die Nachnutzung zur Verfügung zu stellen. Da sie als Forscherin zeitlich sehr eingebunden ist, suchte sie nach einem Ressourcen-sparenden Dienst, der die wichtigsten Metadaten zur Studie in Form vorstrukturierter Eingabemasken abfragt und sowohl Hilfestellungen als auch Anwendungsbeispiele nennt. Sie suchte ein Programm, das während der Studienphase Änderungen ihrer Eintragungen und ihres Datenmaterials erlaubt und Schnittstellen zu Repositoriendiensten anbietet. Die Ansprüche sah sie vom DataWiz abgedeckt, der als Dienst des Leibniz-Zentrums für Psychologie eine Studienbeschreibung und -dokumentation ermöglicht. Zudem ermöglichte ihr der Dienst den Export der Datendokumentation in Form nachnutzbarer Dateiformate. (.xml/.json).

3.4.3 DDI-Codebuchschema

Das Codebuchschema (.xml) der [Data Documentation Initiative \(DDI\)](#) dient als offen nutzbare Software zur Erstellung eines Codebuches für Datensätze der Sozial-, Verhaltens-, Wirtschafts- und Gesundheitswissenschaften.

3.4.4 Codebuch-Muster

- Als Orientierungsbeispiel kann das umfassende Codebuch des Forschungsprojektes „[PsyCorona](#)“ herangezogen werden.
- Hinweise zur Erstellung eines [Codebuchs mit SPSS](#) bietet die Website der University of California (UCLA) zu „Statistical Methods and Data Analytics“.
- Die Ludwig-Maximilian-Universität München hat das sehr ausführliche Codebuch der [PsyCourse-Studie](#) (klinische Längsschnitt- und Beobachtungsstudie zur Klärung der Schizophrenie/Bipolare Störung-Dichotomie) veröffentlicht, das ebenfalls als Muster oder Vorlage dienen kann.

⁴⁵ DataWiz II (beta) befindet sich derzeit noch in der Entwicklungsphase und steht Forscher*innen der Psychologie demnächst über das ZPID kostenfrei zur Verfügung.

3.5 Metadatenstandard

Für die bessere Auffindbarkeit, Nachvollziehbarkeit und Vergleichbarkeit sind Metadaten als strukturierte Informationen über die Forschungsdaten essentiell. Diese können inhaltlicher Natur sein und sich z.B. auf den Titel oder Zeitraum, also allgemeine Angaben zur Forschungsstudie beziehen. Sie können aber auch technisch durch Messinstrumente oder Aufnahmegeräte erzeugt werden. Metadaten sollten sowohl für Menschen als auch Maschinen lesbar sein. Zunehmend werden Metadatenstandards in den Fachdisziplinen entwickelt.

3.5.1 Data Documentation Initiative (DDI)

Die Data Documentation Initiative (DDI) bietet einen internationalen, offenen Standard zur Beschreibung und zum Management von Datensätzen der Sozial-, Verhaltens, Wirtschafts- und Gesundheitswissenschaften mit Umfragen als gängige Methodik. Die [Produkte](#), wozu das Codebuch-Schema aus Abschnitt 3.4.2 gehört, decken diverse Schritte des Forschungsdatenlebenszyklus ab.

3.5.2 DataWiz II (beta) und PsychArchives

Die als .xml- oder .json-Dateiformate exportierbaren Einträge zu Studienbeschreibung und Datendokumentation im „DataWiz II (beta)“ sowie die Angaben zur Archivierung und Veröffentlichung von Daten in „PsychArchives“ schließen ebenfalls standardisierte Metadaten zur besseren Auffindbarkeit der Daten innerhalb der Psychologie ein.

3.5.3 openMINDS

Die [Open Metadata Initiative for Neuroscience Data Structures \(Open MINDS\)](#) ist ein offener, über GitHub verfügbarer Metadaten-Standard vom Human Brain Project (HBP) und EBRAINS für neurowissenschaftliche Datenerhebungen. Die Metadaten-Modelle können im JSON-LD-Format in Graphendatenbanken verwendet werden. Vordefiniert sind Metadaten-Schema zur allgemeinen Beschreibung der Datenerhebung (openMINDS core), zur anatomischen Spezifizierung der Daten wie Gehirnatlanten (openMINDS SANDS), zu eindeutigen konsistenten Begriffen (openMINDS controlled Terms) und den technischen Analysedetails (open MINDS computation) neurowissenschaftlicher Forschung.

4. Daten archivieren und publizieren

Um Daten für mindestens 10 Jahre zu archivieren, bieten sich disziplin- bzw. themenspezifische Datenarchive an, um einerseits von der Sicherheit vor Datenverlust und Fremdzugriff, andererseits von der Auffindbarkeit der Daten durch standardisierte Metadaten innerhalb der wissenschaftlichen Community profitieren zu können. Die Datenarchive verfügen zudem angesichts einer Nachnutzung der Daten Zugriffskontrollmechanismen. Die Publikation von Daten, die über eine Archivierung hinausgeht, stellt insbesondere der Anspruch von Repositorien und Forschungsdatenzentren dar. Die Grenzen zwischen Datenarchiven und Repositorien sind teilweise schwimmend, da mitunter beide Dienste eine Archivierung und den Zugang zu Daten ermöglichen. Die mit einem persistenten Identifier versehenen Datensätze sind je nach Lizenz frei oder wie schon geschrieben, eingeschränkt nachnutzbar.

Aufgrund beider Funktionen – Finden und Archivieren bzw. Veröffentlichen von Daten – können die in Abschnitt 2 erwähnten Dienstangebote auch hier für die Bildungswissenschaft und Psychologie empfohlen werden.

4.1 Psychologie

4.1.1 FDZ am Leibniz-Institut für Psychologie (ZPID)

Das [Forschungsdatenzentrum](#) des Leibniz-Instituts für Psychologie (ZPID) stellt psychologische Forschungsdaten für die Nachnutzung durch die wissenschaftliche Gemeinschaft bereit und unterstützt Forschende bei einem qualitätsgesicherten FDM während des gesamten Forschungsprozesses. Der Zugang zu den Forschungsdaten wird grundsätzlich über einen Nutzungsvertrag/eine Nutzungslizenz geregelt. Das FDZ am ZPID ist vom RatSWD akkreditiert.

4.1.2 PsychArchives

Das Repository [PsychArchives](#) des Leibniz-Instituts für Psychologie (ZPID) unterstützt die Archivierung von und den Zugang zu Datensätzen und diversen weiteren digitalen Forschungsobjekten (*Digital Research Objects*, DROs). Insgesamt 20 verschiedene DROs können über PsychArchives archiviert und für die Nachnutzung zugänglich gemacht werden. Dazu zählen z.B. Artikel, Vorabdrucke, Forschungsdaten, Codes, Ergänzungen, Prä-Registrierungen, Tests und Multimedia-Objekte.

4.1.3 Use Case: PsychArchives

Eine Kognitionswissenschaftlerin hat das Ergebnis ihrer Studie in Form eines Artikels festgehalten. Diesen hat sie im Journal of Numerical Cognition eingereicht. Ihr Paper wurde nach einem Peer Review-Verfahren akzeptiert. Während die Forscherin bereits selbst überlegte, in welchem Datenarchiv oder Repository sie ihren Appendix, die Daten sowie das dazugehörige Material zum Nachverfolgen und zur Nachnutzung der Studie archivieren und publizieren kann, wurde ihr diese Entscheidung vom Journal abgenommen. Dieses bat sie um die Datenablage im Forschungsdatenrepository PsychArchives des Leibniz-Instituts für Psychologie.

„Zwei Fliegen wurden mit einer Klappe geschlagen“ und die Forscherin wählte diesen Dienst so mit für die Archivierung und Publikation ihres Supplementary Materials, in diesem Fall ihren Datensatz, Ergebnisse der Datenanalyse und den Appendix. Zusätzlich eröffnete dieses Prozedere die Möglichkeit, dass das Manuskript rasch online verfügbar war, nachdem dieses vom Journal akzeptiert worden ist. So konnte das Gedankengut als Autor-Akzeptiertes-Manuskript (AAM) direkt geteilt und zitiert werden noch bevor das Journal den Produktionsprozess des Artikels fertiggestellt hatte. Als besonders ansprechend hat die Forscherin hierbei empfunden, dass der Hauptartikel und die zusätzlichen Materialien zwar unterschiedliche Einträge im Archiv darstellen, aber aufeinander verweisen und somit leicht einander zuzuordnen sind.

Über den Dienst PsychArchives Submission Assistant hat die Forscherin ihre Dateien in das Repository hochgeladen. Dafür wählte sie u.a. das entsprechende [Sharing-Level](#) (siehe Abschnitt 4.4) aus, welches in Kombination mit einer entsprechenden Nutzungslizenz die Nachnutzungsbedingungen für einen

bestimmtes Forschungsobjekt regelt. Außerdem wählte sie für ihre Dateien das [Dateiformat](#), (siehe Abschnitt 3.1.1) welches für eine verbesserte Archivierung und Nachnutzbarkeit verlangt wird.

Neben der sicheren Archivierung ihrer Dateien ermöglicht PsychArchives der Forscherin die Verknüpfung ihrer verschiedenen Dateien untereinander sowie deren Zitierbarkeit und Zugänglichkeit durch die Vergabe eines Digital Object Identifiers (DOI). PsychArchives verhalf ihr als international anschlussfähiges Repositorium, ihre Daten und ihren Appendix auffindbar zu machen und mit dem Journal zu verknüpfen.

4.1.4 Open Neuro

Das frei zugängliche Repositorium [Open Neuro](#) der Stanford University bietet Zugang für Daten aus den kognitiven Neurowissenschaften. Schwerpunkt bilden Studien mit Daten der funktionellen Magnetresonanztomographie (fMRI-Daten). Die Ablage der Daten orientiert sich an der Brain Imaging Data Structure (BIDS). Neben fMRI-Daten umfasst das Repositorium EEG-, iEEG, PET- und MEG-Datensätze.

4.1.5 Open Science Framework (OSF)

Die Plattform [Open Science Framework](#) (OSF) ist ein disziplinübergreifendes Angebot des *Center for Open Science* (COS), kommt jedoch stark in der psychologischen Forschungscommunity zum Einsatz. OSF bietet die Möglichkeit offene Datensätze zu finden sowie während oder nach dem Forschungsprozess die erhobenen Daten zu dokumentieren, zu publizieren und in Kooperationen mit internationalen Forscher*innen zu treten. OSF ist flexibel nutzbar und bietet sich insbesondere für das Managen oder Arbeiten mit Forschungsdaten während einer Studie an. Eine Qualitätssicherung seitens des Betreibers sowie eine spezifische standardisierte Struktur der Datendokumentation der Daten werden nicht gegeben. Die Sicherstellung der Langzeitarchivierung der Forschungsdaten ist bei OSF an deren privatrechtliche Finanzierungsbedingungen gebunden.

4.1.6 Open Test Archive

Das Repositorium [Open Test Archive](#) des Leibniz-Instituts für Psychologie (ZPID) stellt Open-Access-Tests zur Verfügung und unterstützt Forscher*innen bei der Publikation ihrer urheberrechtlich geschützten Testinstrumente. Die Tests sind neben der Forschung in der Lehre und Praxis einsetzbar.

4.2 Bildungswissenschaften⁴⁶

4.2.1 Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Das an der Humboldt-Universität zu Berlin verortete FDZ am [Institut zur Qualitätsentwicklung im Bildungswesen](#) (IQB) archiviert und dokumentiert Datensätze und Begleitmaterialien insbesondere aus der quantitativ ausgerichteten empirischen Bildungsforschung. Dazu gehören die Fachbereiche der Erziehungswissenschaft, pädagogischen Psychologie, Bildungssoziologie, Bildungsökonomie und Fachdidaktik. Den Kern bilden zugängliche Datensätze der nationalen und internationalen Schulleistungsstudien in Deutschland für Re- und Sekundäranalysen. Ein Fokus wird u.a. auf Studien zur Kompetenzmessung von Schülerinn*innen gelegt. Das FDZ ist vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditiert.⁴⁷

⁴⁶ Weitere Adressen neben den hier gelisteten zentralen Forschungsdatenzentren für die empirische Bildungsforschung finden Sie im Fachportal Bildung unter https://www.fachportal-paedagogik.de/forschungsdaten/Infrastruktur.php#inst_10 sowie beim Verbund Forschungsdaten Bildung (Verbund FDB) unter <https://www.forschungsdaten-bildung.de/datenzentren>.

⁴⁷ Informationen zur Akkreditierung vom RatSWD Vgl. <https://www.konsortswd.de/datenzentren/akkreditierung/>.

4.2.2 FDZ-Bildung

Das [FDZ-Bildung](#) vom Leibniz-Institut für Bildungsforschung und Information (DIPF) sowie vom Verbund Forschungsdaten Bildung (Verbund FDB) bietet Zugang zu Studien aus der qualitativen empirischen Erziehungs- und Bildungsforschung und umfasst neben Datensätzen auch Erhebungsinstrumente der quantitativen Forschung (Fragebögen, Skalen und Tests). Der Schwerpunkt der Instrumentensammlung und Datensets liegt auf der Schulqualitäts- und Unterrichtsqualitätsforschung. Das FDZ ist an der Erhebung von Daten großer Studien wie NEPS oder PISA beteiligt. Es ist vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditiert.

4.2.3 Datenarchiv „Kindheit und Jugend im urbanen Wandel“

Das Datenarchiv „[Kindheit und Jugend im urbanen Wandel](#)“⁴⁸ des Methodenzentrums Qualitative Bildungsforschung der Universität Duisburg-Essen umfasst qualitative Studien der Kindheit und Jugendforschung (Datenmaterial seit den 1970ern). Der Fokus liegt auf Forschungsvorhaben aus den Standorten in Nordrhein-Westfalen. Zum archivierten qualitativen Datenmaterial gehören u.a. Audiomitschnitte von Interviews, Videoaufzeichnungen, Fotografien, Beobachtungsprotokolle, Zeitdokumente und Transkripte.

4.2.4 Archiv für pädagogische Kasuistik (ApaeK)

Das [Archiv für pädagogische Kasuistik](#) (ApaeK) der Goethe Universität Frankfurt am Main bietet Zugang zu Daten der qualitativen erziehungswissenschaftlichen Forschung mit einem Schwerpunkt in der Schul- und Unterrichtsforschung sowie der rekonstruktiven Bildungsforschung. Es richtet sich an Lehrende und Studierende, die sich in kasuistischer Lehre fallorientiert mit pädagogischer Praxis auseinandersetzen.

4.2.5 Fachportal Pädagogik

Das [Fachportal Pädagogik](#) des Leibniz-Instituts für Bildungsforschung und Bildungsinformation bietet einen umfassenden, filterbaren Überblick über Studien und Daten aus verschiedenen Datenzentren der Erziehungswissenschaft, Bildungsforschung und Fachdidaktik.

4.2.6 Instrumentendatenbank

Neben Datensätzen gibt es die [Instrumentendatenbank](#) des FDZ-Bildung vom Leibniz-Institut für Bildungsforschung und Information (DIPF). Diese stellt dokumentierte quantitative Instrumente wie Skalen, Fragebögen und Tests zur Verfügung.

4.3 Sozialwissenschaftliche (Umfrage-)Daten

4.3.1 SowiDataNet|datorium

Das Repositorium [SowiDataNet|datorium](#) des Leibniz Instituts für Sozialwissenschaften (GESIS) bietet Forscher*innen die Archivierung von sowie den Zugang zu quantitativen Primärdaten und Materialien der sozialwissenschaftlichen Forschung an. Die Daten werden standardisiert dokumentiert und aufbereitet (z.B. Bereinigung, Variablenreports).

4.3.2 Qualiservice

Das FDZ [Qualiservice](#) der Universität Bremen kümmert sich um die Archivierung und Bereitstellung nachnutzbarer qualitativer Forschungsmaterialien aus der Sozialforschung. Dazu zählen Datensätze multidisziplinärer und themenungebundener Studien wie Interviewstudien, Audio- und Bilddateien oder auch Social-Media-Daten. Qualiservice gehört zu den vom RatSWD akkreditierten Forschungsdatenzentren.

⁴⁸ Aufgrund technischer Störungen führen einzelne Links derzeit auf Fehlermeldungen oder leere Seiten. Hinweise oder Fragen richten Sie bitte an folgende temporäre Adresse: fachportal-pedocs@dipf-institut.de.

4.3.3 CESSDA Data Catalogue

Der [CESSDA Data Catalogue](#) stellt eine Datenbank vom *Consortium of European Social Science Data Archives* dar. Sie bietet über 40.000 europaweite Datensätzen der Sozialwissenschaften, die mehrsprachig, meist englisch, dokumentiert sind. Anfragen zu Datensätzen aus den im CESSDA Data Catalogues versammelten Repositorien können u.a. über Schlagworte des disziplinspezifischen Forschungsfeldes angestellt werden. Die Datenbank bietet Zugänge zu quantitativen, qualitativen oder gemischten Daten, Querschnitt- oder Längsschnittdaten, kürzlich erhobenen oder historischen Daten.

4.3.4 Repositorium für Mess- und Erhebungsinstrumente ZIS

Das [Repositorium für Mess- und Erhebungsinstrumente](#) (ZIS) des Leibniz Institut für Sozialwissenschaften (GESIS) bietet englisch- und mehrsprachige Erhebungs- und Messinstrumente aus der Umfrageforschung in den Sozial- und Verhaltenswissenschaften an. Dazu gehören Items, Skalen, Fragebögen, Tests, Indexe.

4.4 Zugriffsklassen⁴⁹

Wenn Roh- oder Sekundärdaten gemäß den Anforderungen der GWP für mindestens 10 Jahre archiviert und ggf. *geteilt* werden sollen⁵⁰, ergeben sich generell je nach Datentyp verschiedene Möglichkeiten, die sich auf die Weitergabe der Daten beziehen: „(a) die Übergabe der Daten an ein Repositorium oder FDZ, (b) die Übergabe der Daten an ein öffentliches Archiv, (c) begründeter Verzicht auf die Übergabe von Daten und eigene Archivierung.“⁵¹

Die Publikation personenbezogener Daten als eine Form der Datenverarbeitung bedarf einer Rechtfertigung.⁵² Sofern Ihre Forschungsdaten Personenbezug aufweisen, ist die Publikation der Daten z.B. mit einer Einwilligung (siehe Abschnitt 4.1) zu rechtfertigen (Art. 6 Abs. 1 Lit. a DSGVO, § 27 Abs. 4 BDSG)⁵³.

Nicht immer bietet sich jedoch eine Anonymisierung der Daten für deren Nachnutzung. Eine Anonymisierung der pb-Daten erweist sich z.B. bei Videoaufnahmen, die Emotionen und Gesichtsausdrücke der Probanden als Inhalt des Forschungsgegenstandes behandeln, als nicht sinnvoll: der Untersuchungsgegenstand wäre maskiert und somit unkenntlich an.⁵⁴

Grundsätzlich gilt, dass die „Einverständniserklärungen [...] Aussagen zur Verwendung der Daten enthalten“ müssen.⁵⁵ „Die Forschenden, die um Einwilligung ersuchen, sind an die dort genannten

⁴⁹ Der Abschnitt 4.4 „Zugriffsklassen“ ist dem Dokument „Forschungsdatenmanagement bei personenbezogenen Daten – eine Handreichung“ entnommen, Vgl. Kienbaum, Janna; Fischer, Patryk; Paßmann, Sven. (2023).

Forschungsdatenmanagement bei personenbezogenen Daten – eine Handreichung. 10.5281/zenodo.7428524.

⁵⁰ Vgl. die Übersicht vom Verbund Forschungsdaten Bildung (Verbund FDB) „Data-Sharing: Gründe zum Teilen“, <https://www.forschungsdaten-bildung.de/data-sharing>.

⁵¹ Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung, S. 17, https://www.forschungsdaten-bildung.de/files/stellungnahme_zum_fdm_dgfe-gebf-gfd.pdf.

⁵² Bei anonymisierten Daten erlischt der Personenbezug und die Forschung mit den Daten ist rechtlich unproblematisch, Vgl. Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung, S. 7, https://www.forschungsdaten-bildung.de/files/stellungnahme_zum_fdm_dgfe-gebf-gfd.pdf

⁵³ Vgl. Datenschutz-Handbuch des RatSWD, ab S. 30, https://www.konsortswd.de/wp-content/uploads/RatSWD_Output8.6_HandreichungDatenschutz_2.pdf.

⁵⁴ Vgl. Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung, S. 10 u. 11, https://www.forschungsdaten-bildung.de/files/stellungnahme_zum_fdm_dgfe-gebf-gfd.pdf.

⁵⁵ Meyermann, Alexia; Porzelt, Maike (2019): Datenschutzrechtliche Anforderungen in der empirischen Bildungsforschung. Eine Handreichung. Version 2.1. Frankfurt am Main: DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation 2019, 39 S. -(forschungsdaten bildung informiert; 6), S. 19, <https://doi.org/10.25656/01:21990>.

Zwecke gebunden. Die Zwecke sollten so eng wie möglich, aber so breit wie nötig angegeben werden.⁵⁶

Angesichts datenschutzrechtlicher oder forschungsethischer Anforderungen sollten Sie jedoch Zugriffsregelungen und/oder spezifische Lizenzen⁵⁷ der Nachnutzung angeben. Diese reichen von der freien Verfügbarmachung bis hin zu streng kontrollierten Bedingungen wie Nutzungsverträgen mit den Nutzenden. Forschungsdatenzentren (FDZ) oder Repositorien bieten dafür den technischen und organisatorischen Rahmen.

In der Bildungsforschung wird beispielsweise zwischen drei groben Zugriffsklassen unterschieden:⁵⁸

- „*Freier Zugang* („Public Use File“): Daten können entweder direkt oder nach Registrierung und Zustimmung zu Nutzungsbedingungen eingesehen oder heruntergeladen werden.
- *Zugang auf Antrag* („Scientific Use File“): Daten können erst nach Stellung eines Antrags und auf Basis einer Datennutzungsvereinbarung genutzt werden. Der Datenzugang ist auf wissenschaftliche Nutzung beschränkt. Scientific Use Files sind typischerweise weniger stark anonymisiert als Public Use Files.
- *Gesicherter Zugang*: Um Zugang zu Datensätzen mit niedrigerem Anonymisierungsniveau anbieten zu können, bieten einige FDZ geschützte Zugänge an. Dazu zählt die Möglichkeit, Auswertungen an einem Gastarbeitsplatz durchzuführen oder – bei quantitativen Daten – über ein Fernrechner-System (z.B. RemoteNEPS, JoSuA) zu arbeiten. Hierbei erfolgt keine physische Übergabe des Datensatzes an die Nutzerinnen und Nutzer.“

Im Falle von Videodaten mit starkem Personenbezug, greift die dritte Zugangsstufe „Gesicherter Zugang.“

Die Empfehlungen der Deutschen Gesellschaft für Psychologie (DGP)⁵⁹ sind ähnlich formuliert:

- „Zugriffsklasse 0 („open data“): Weder der Zugriff noch die Nutzung der Daten sind in irgendeiner Form restringiert. (...)“
- Zugriffsklasse 1 („open data / conditional access“): Der Zugriff bzw. die Nachnutzung ist an bestimmte, von den Datenbereitstellenden festgelegte Bedingungen geknüpft, mit denen sich Datennachnutzende explizit einverstanden erklären müssen. (...)“
- Zugriffsklasse 2 („restricted access“): Der Zugriff bzw. die Nachnutzung ist über die in Zugriffsklasse 1 geltenden Bestimmungen hinaus an weitere, vorher von den Datenbereitstellenden festgelegte Bedingungen geknüpft, mit denen sich Nachnutzende explizit einverstanden erklären müssen (individualisierte Verträge zwischen Datenbereitstellenden und Nachnutzenden). (...)“
- Zugriffsklasse 3 („secure data“): Ein Zugriff auf die Daten ist nur unter datenschutzrechtlich gebotenen Restriktionen (z. B. online über einen entsprechend abgesicherten Kanal oder vor Ort in einem Datenzentrum) möglich. (...)“

⁵⁶ Meyermann, Alexia; Porzelt, Maike (2019), Vgl. auch S. 19 u. 20 zur engen und weiten Zweckbindung.

⁵⁷ Zur Bestimmung der Lizenz siehe: <http://ufal.github.io/public-license-selector/>.

⁵⁸ Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung, S. 14 u. 15, https://www.forschungsdaten-bildung.de/files/stellungnahme_zum_fdm_dgfe-gebf-gfd.pdf.

⁵⁹ Vgl. Management und Bereitstellung von Forschungsdaten in der Psychologie: Überarbeitung der DGPs-Empfehlungen (2020), S. 13-16, <https://psyarxiv.com/hcxtm/>.