



# Forschungsdatenmanagement in der Psychologie

## Fachspezifisches Train-the-Trainer-Konzept

Fachspezifische Informationen

Version 1.0


**Sven Paßmann**


**Sibylle Söring**

Freie Universität Berlin

Berlin 2022

## DIE AUTORINNEN UND AUTOREN

Sven Paßmann  0000-0002-1698-3289

Sibylle Söring  0000-0002-1698-3289

## HERAUSGEBER

Freie Universität Berlin mit Projektschwerpunkt „Fachspezifisches Train-the-Trainer-Modul“ (AP 5).

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

Gefördert durch



## IMPRESSUM

„Forschungsdatenmanagement in der Psychologie: Fachspezifisches Train-the-Trainer-Konzept. Fachspezifische Informationen“ von Sven Paßmann und Sibylle Söring ist lizenziert unter der Creative Commons Attribution 4.0 International (CC BY 4.0).



## ZITATIONSVORSCHLAG

Paßmann, Sven; Söring, Sibylle. (2023). Forschungsdatenmanagement in der Psychologie: Fachspezifisches Train-the-Trainer-Konzept. DOI: 10.5281/zenodo.7560900.

## DOI

<https://doi.org/10.5281/zenodo.7560900>



## Inhaltsverzeichnis

Einleitung.....	6
Danksagung .....	8
Forschungsdaten (Einheit 3) .....	9
1.    Forschungsdaten: Definition .....	9
2.    Spezifika von Forschungsdaten der Psychologie.....	9
3.    Forschungsdatenlebenszyklus .....	10
Forschungsdatenmanagement (Einheit 4).....	12
1.    Motivation.....	12
2.    Forschungsdatenmanagement.....	12
a)    Was ist Forschungsdatenmanagement? .....	12
b)    Aufgaben des Forschungsdatenmanagements.....	14
c)    Die FAIR-Prinzipien .....	15
3.    (Psychologie-spezifische) Forschungsdaten-Policies .....	16
a)    Was ist eine Forschungsdaten-Policy? .....	16
b)    Psychologie-spezifische Policies .....	16
c)    Leitlinien von Forschungsförderern .....	17
d)    Institutionelle Policies und Leitlinien (in der Psychologie).....	18
Rechtliche Aspekte (Einheit 5) .....	20
1.    Ebenen der Regelung.....	20
2.    Datenschutz und Persönlichkeitsrecht .....	21
3.    Anonymisierung der in der Psychologie relevanten Daten.....	22
4.    Urheberrecht und verwandte Schutzrechte .....	23
5.    Vertragsrecht .....	24
Datenmanagementplan (Einheit 6) .....	25
1.    Motivation.....	25
2.    Was ist ein Datenmanagementplan?.....	25
3.    Bestandteile eines Datenmanagementplans .....	25
4.    Anforderungen der Forschungsförderer .....	26
5.    Muster und Werkzeuge für die Psychologie .....	27
Systematik im FDM (Einheit 7).....	30
1.    Motivation.....	30
2.    Verzeichnisstruktur .....	30
3.    Benennung.....	31
4.    Umbenennen mehrerer Dateien gleichzeitig .....	32
5.    Versionskontrolle .....	32
Dokumentation und Metadaten (Einheit 8) .....	34
1.    Motivation.....	34
2.    Datendokumentation.....	34
3.    Codebooks für die Psychologie .....	35
4.    Was sind Metadaten? .....	36

5. Metadatenstandards .....	36
6. Psychologie-spezifische Metadatenstandards .....	36
7. Thesaurus, Normdaten und Klassifikationen in der Psychologie .....	37
8. Elektronische Laborbücher (ELB) in der Psychologie .....	38
Zugriffssicherheit (Einheit 9) .....	40
1. Motivation.....	40
2. Verschlüsselung .....	40
3. Überprüfen der Nutzbarkeit .....	40
4. Passwortschutz und Zugriffsrechte .....	41
Speicherung und Backup (Einheit 12) .....	42
1. Motivation.....	42
2. Datenspeicherung.....	42
3. Backup-Varianten .....	43
4. Hinweise für ein sicheres Backup.....	44
Langzeitarchivierung (Einheit 13) .....	45
1. Motivation.....	45
2. Begriff „Langzeitarchivierung“ .....	45
3. Nachhaltige Dateiformate in der Psychologie.....	45
4. Anforderungen an Langzeitarchive.....	46
Institutionelle Infrastruktur (Einheit 14) .....	48
1. Motivation.....	48
2. Institutionelles Forschungsdatenmanagement.....	48
3. Forschungsdatenmanagement-Infrastrukturen .....	48
Publikation von Forschungsdaten (Einheit 15) .....	50
1. Ziele und Motivation.....	50
2. Leitfragen für die Datenauswahl .....	50
3. Publikationswege.....	51
4. Repositorien.....	52
5. Repositorien für die Psychologie .....	52
6. Auswahl eines Repositoriums.....	53
7. Lizenzen.....	54
8. Persistente Identifier .....	56
a) Digital Object Identifier (DOI).....	56
b) Open Researcher and Contributor ID (ORCID).....	57
9. Zugriffsklassen/Agreements .....	57
Weiterführende Ressourcen .....	59
<b>Einheit 3</b> .....	59
<b>Einheit 4</b> .....	59
<b>Einheit 5</b> .....	59
<b>Einheit 6</b> .....	60
<b>Einheit 7</b> .....	60
<b>Einheit 8</b> .....	61
<b>Einheit 9</b> .....	61

<b>Einheit 12</b> .....	61
<b>Einheit 13</b> .....	61
<b>Einheit 15</b> .....	61

## Abkürzungsverzeichnis

BMBF	Bundesministerium für Bildung und Forschung
CC	Creative Commons
DFG	Deutsche Forschungsgemeinschaft
DINI	Deutsche Initiative für Netzwerkinformation
DMP	Datenmanagementplan
EC	Europäische Kommission
FD	Forschungsdaten
FDM	Forschungsdatenmanagement
K	Kurze Version
L	Lange Version
M	Mittellange Version
PPTX	PowerPoint-Präsentation
TN	Teilnehmende
UAG	Unter-Arbeitsgruppe
WL	Workshopleitende

## Einleitung

Das vorliegende Konzept wurde im Rahmen des 2020-23 von der DFG geförderten Verbundvorhaben FDNNext erarbeitet und richtet sich zuvorderst an Trainer\*innen des Fachbereichs Psychologie, die einen fachspezifischen Train-the-Trainer-Workshop zum Thema Forschungsdatenmanagement (FDM) durchführen wollen. Zu diesem Zweck wurden zahlreiche, für die Psychologie relevanten Aspekte des Forschungsdatenmanagements zusammengetragen und in das Workshop-Konzept des Vorgänger-Projekts FDMentor<sup>1</sup> (finanziert durch das Bundesministerium für Bildung und Forschung, Projektnummer: 16FDM010-16FDM014) integriert.

Um den unterschiedlichen Erfahrungsstufen der Trainer\*innen in FDM und/oder Didaktik Rechnung zu tragen, wurden zwei Hauptdokumente (FDNNext\_TtTKonzept\_Fachinformationen, FDNNext\_TtTKonzept\_DidaktikLehrdrehbücher) sowie eine Methodensammlung (FDNNext\_TtTKonzept\_MethodenMaterialien) erstellt, die unabhängig voneinander, aber auch ergänzend genutzt werden können (siehe Tabelle 1). Aufgrund der wechselweisen Referenzierung empfiehlt sich die parallele Nutzung beider Dokumente. Alle Dokumente sowie die entsprechenden Begleitmaterialien (siehe unten) liegen als .zip-File<sup>2</sup> im offenen Repositorium Zenodo zum freien Download vor.

*Tabelle 1: Übersicht der Dokumente.*

Dokument „FDNNext_TtTKonzept...“	Inhalt	Zielgruppe
<b>_Fachinformationen</b>	Grundlegende Aspekte des FDM mit fachspezifischen Informationen für Psycholog*innen	unerfahren in FDM
<b>_DidaktikLehrdrehbücher</b>	Fachliche Informationen zu Didaktik; Ziele der Lerneinheiten; Übersicht der jeweils verwendeten Methoden und Übungen; Lehrdrehbücher	unerfahren in Didaktik
<b>_MethodenMaterialien</b>	Übersicht und Beschreibung der im Workshop verwendeten Methoden; Materialsammlung	
<b>_Arbeitsmaterialien_FDNNext_TtTKonzept.zip</b>	Sammlung verschiedenster Materialien zur Ausgestaltung eines Workshops (Präsentationsfolien, Lehrdreh-bücher, Sammlung fachspezif. Informationen)	

Das vorliegende Dokument „FDNNext\_TtTKonzept\_Fachinformationen“ bezieht sich dabei ausschließlich auf die fachspezifischen Aspekte des FDM (**Lerneinheiten 3-9 und 12-16**) und kann von bereits Didaktik-erfahrenen Trainer\*innen als inhaltliche Grundlage für die Workshop-Planung genutzt werden. Jede Lerneinheit referiert dabei auf einen spezifischen Aspekt des FDM und enthält Hintergrundwissen sowie weitergehende Informationen zu den einzelnen Themengebieten. Die hinter jedem Aspekt in Klammern vermerkte Nummer verweist auf das zweite Dokument

<sup>1</sup> Internetpräsenz des Projektes FDMentor, Zugriff am 27.04.2023, <https://www.forschungsdaten.org/index.php/FDMentor>.

<sup>2</sup> FDNNext-Community auf Zenodo, Zugriff am 27.04.2023, <https://zenodo.org/communities/fdnext/?page=1&size=20>.

„FDNext\_TtTKonzept\_DidaktikLehrdrehbücher“ und den darin enthaltenen Lehrdrehbüchern für die jeweilige Lerneinheit.

Thematische Einheiten, die sich ausschliesslich auf didaktische Komponenten beziehen (**Lerneinheiten 1-2, 10-11 und 17-21**), werden inhaltlich nur in dem Dokument „FDNext\_TtTKonzept\_Didaktik Lehrdrehbücher“ behandelt. Besonderer Fokus sind hier die sozialen, didaktischen und organisatorischen Komponenten eines Workshops inkl. der zu deren Verständnis notwendigen Hintergrundinformationen. Für die konkrete Ausgestaltung einzelner Themengebiete enthält das Dokument exemplarische Lehrdrehbücher für einzelne Lerneinheiten zur Orientierung und definiert mögliche, im Workshop zu erreichende Ziele.

### Fachspezifische Komponenten

Forschungsdaten (FI-3)	Forschungsdatenmanagement (FI-4)
Rechtliche Aspekte (FI-5)	Datenmanagementplan (FI-6)
Ordnung und Struktur (FI-7)	Dokumentation und Metadaten (FI-8)
Zugriffssicherheit (FI-9)	Speicherung und Backup (FI-12)
Langzeitarchivierung (FI-13)	Institutionelle Infrastruktur (FI-14)
Publikation von Forschungsdaten (FI-15)	

### Soziale Komponenten

Begrüßen & Kennenlernen (DI-1, DI-11)	Abschluss des ersten Tages (DI-10)
Feedback und Verabschiedung (DI-21)	

### Didaktische Komponenten

Didaktisches Vorgehen (DI-17)	Konzeptentwicklung (DI-18)
Didaktische Methoden (DI-19)	Praktische Übung (DI-20)

### Organisatorische Komponenten

Orientierung (DI-2)	Formaler Rahmen (DI-16)
---------------------	-------------------------

Abbildung 1: Komponenten des Train-the-Trainer-Konzepts. Die angegebenen Nummern entsprechen den Einheitsnummern im jeweiligen Dokument (FI, Fachinformation; DI, didaktische Informationen).

Die Nachnutzung der Konzept-Dokumente soll die Planung und Durchführung eines insgesamt zweitägigen (Online-)Workshops ermöglichen, dessen Inhalt auf sieben Zeitstunden pro Tag verteilt ist. Jeder Tag enthält zudem jeweils eine Mittagspause (eine Stunde) und zwei kürzere Pausen (je eine Viertelstunde). Dabei sind die einzelnen Lerneinheiten so konzipiert, dass diese sowohl eigenständig als auch in thematischen Gruppen genutzt werden können, welche z. B. „Forschungsdaten“ und „Forschungsdatenmanagement“ sein können. Aber auch „Rechtliche Aspekte“ und „Zugriffssicherheit“, oder auch „Speicherung und Backup“ sowie „Langzeitarchivierung“ bieten sich als thematische Einheiten an. Am Ende des Dokuments finden sich darüber hinaus weiterführende Lektürehinweise.

## Danksagung

Das hier vorliegende Konzept basiert auf dem „Train-the-Trainer Workshop zum Thema Forschungsdatenmanagement“-Konzept des BMBF-geförderten Vorgänger-Projektes FDMentor (Autor\*innen: Katarzyna Biernacka, Petra Buchholz, Sarah Ann Danker, Dominika Dolzycka, Claudia Engelhardt, Kerstin Helbig, Juliane Jacob, Janna Neumann, Carolin Odebrecht, Britta Petersen, Benjamin Slowig, Ute Trautwein-Bruns, Cord Wiljes, Ulrike Wuttke); zahlreiche generische Inhalte konnten übernommen werden. Den Autorinnen gilt unser Dank für die umfangreichen Vorarbeiten, auf denen aufgebaut werden konnte.



## Forschungsdaten (Einheit 3)<sup>3</sup>

### 1. Forschungsdaten: Definition

In der Vergangenheit haben sich – je nach Fachdisziplin und Projektkontext - unterschiedliche Definitionen von Forschungsdaten etabliert. Kindling/Schirmbacher (2013) betrachten digitale Forschungsdaten als „[...] alle digital vorliegenden Daten [zu verstehen sind], die während des Forschungsprozesses entstehen oder ihr Ergebnis sind“<sup>4</sup>. Laut der Deutschen Forschungsgemeinschaft (DFG) zählen zu Forschungsdaten „[...] u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Umfragedaten, Objekte aus Sammlungen oder Proben, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden. Methodische Testverfahren, wie Fragebögen, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und sollten daher ebenfalls unter den Begriff Forschungsdaten gefasst werden“<sup>5</sup>.

Die Technische Universität Berlin fasst diesen Begriff noch weiter<sup>6</sup>. Forschungsdaten sind „alle Informationen (unabhängig von Form oder Darstellung), die während eines Forschungsprozesses entstehen oder sein Ergebnis sind, einschließlich der Informationen, die zur Nachvollziehbarkeit und Reproduktion der Ergebnisse notwendig sind.“ In der Aufzählung werden u. a. „Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Objekte aus Sammlungen oder Proben, Umfragen und Interviews, aber auch Skripte, Zeitschriebe, Berechnungen, Software und Code“ als spezifische Forschungsdaten erwähnt. Die Freie Universität Berlin bestimmt in ähnlicher Weise „Informationen als Forschungsdaten, die die Dokumentation, Nachvollziehbarkeit und – abhängig vom Forschungsgebiet – Reproduzierbarkeit der Ergebnisse gewährleisten (Metadaten)“<sup>7</sup>.

In ihrer fachspezifischen Empfehlung von 2016 bezeichnet die Deutsche Gesellschaft für Psychologie (DGPs) Forschungsdaten zunächst noch als Primärdaten, welche „vollkommen unbearbeitete (d. h. untransformierte, nicht-aggregierte etc.) quantitative und qualitative Daten, zum Beispiel bei Experimenten alle manipulierten und gemessenen Variablen für jeden Experimentaldurchgang jeder Person“ sind. In den nachfolgenden Überarbeitungen<sup>8</sup> wird diese Definition erweitert: Sie umfasst neben den Primär- nunmehr auch die Roh- und Sekundärdaten.

### 2. Spezifika von Forschungsdaten der Psychologie

Aufgrund des Forschungsobjekts „Mensch“ und seines durch kognitive Prozesse geleiteten Verhaltens sind die in der Psychologie erhobenen Daten sehr häufig durch einen starken Personenbezug gekennzeichnet. Dies führt dazu, dass die Erhebung solcher Daten immer auch mit dem Recht der

<sup>3</sup> Die Vergabe der Nummern für jede Einheit basiert auf die im gesamten Workshop enthaltenen Lerneinheiten. Da in diesem Dokument nur fachspezifische Informationen enthalten sind, die didaktisch relevanten Informationen jedoch im Dokument „FDNext\_TtTKonzept\_DidaktikLehrdrehbücher“, fängt die Zählung hier mit der Ziffer 3 an.

<sup>4</sup> Kindling, Maxi und Peter Schirmbacher: „Die digitale Forschungswelt“ als Gegenstand der Forschung. Information - Wissenschaft & Praxis<sup>64</sup> (2013), S. 130. <https://doi.org/10.1515/iwp-2013-0017>.

<sup>5</sup> Deutsche Forschungsgemeinschaft. „Leitlinien zum Umgang mit Forschungsdaten.“ Zugriff am 27.04.2023, [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf).

<sup>6</sup> Forschungsdatenpolicy der Technischen Universität Berlin. Zugriff am 27.04.2023, <https://www.tu.berlin/arbeiten/wichtige-dokumente/richtlinien-leitlinien/forschungsdaten-policy>.

<sup>7</sup> Definition „Forschungsdaten“ der Freien Universität Berlin. Zugriff 27.04.2023, <https://www.fu-berlin.de/sites/forschungsdatenmanagement/ueber-forschungsdaten/forschungsdaten/index.html>.

<sup>8</sup> Gollwitzer, Mario, Andrea Abele-Brehm, Christian Fiebach, Roland Ramthun, Anne M. Scheel, Felix D. Schönbrodt, and Ulf Steinberg. 2020. „Data Management and Data Sharing in Psychological Science: Revision of the DGPs Recommendations.“ PsyArXiv. September 10. doi:10.31234/osf.io/24ncs.

Person auf den Schutz ihrer Daten kollidiert, die in der Charta der Europäischen Union verankert ist <sup>9</sup>. Potentielle Teilnehmende einer Studie müssen der Erhebung und Verarbeitung ihrer Daten deshalb durch eine informierte Einwilligung eindeutig zustimmen (siehe Einheit 5 "Rechtliche Aspekte", S. 20).

Bei einigen Datenformaten (bspw. Video- und Audioaufnahmen, Genetik, M/EEG oder MRT-Daten) lässt sich zudem ein Personenbezug auch bei anonymisierter Erhebung nicht gänzlich von einem Individuum trennen. Das Gehirn, und ebenso die Gehirnaktivität, könnte mit entsprechend weit entwickelten Programmen aufgrund der individuellen Lebensgeschichte, welche sich in spezifischen Verknüpfungen zwischen den Gehirnarealen widerspiegelt, eindeutig einer Person zuordenbar sein. Insbesondere MRT-Aufnahmen enthalten neben den Daten von Interesse auch Informationen über Knochenstruktur und Bindegewebe, welche eine eindeutige Rekonstruktion des Gesichtes einer Person ermöglichen <sup>10</sup>. Darüber hinaus besitzen menschliche Stimmen personenspezifische Merkmale<sup>11</sup> wie Stimmhöhe, Stimmfrequenz usw., bestimmt durch die individuelle Größe und Form des Kehlkopfes oder ethnische Zugehörigkeit. Videoaufnahmen zeichnen neben dem Gesicht oft auch spezifische Körpermerkmale, wie z.B. Tattoos, Frisur, Körperbewegungen (vor allem im klinischen Kontext) usw. auf.

Auch wenn Methoden existieren, die eine nachfolgende Zuordnung solcher Daten zu konkreten Personen erschweren (siehe Einheit 5 "Rechtliche Aspekte", S. 20) kann eine komplette Anonymisierung nicht immer gewährleistet werden, was bei einer geplanten Publikation der Daten beachtet werden muss (siehe Einheit 15 "Publikation von Daten", Punkt 2, S. 50). Diese, die Psychologie betreffenden, Besonderheiten der Daten müssen somit im Sinne des Forschungsdatenmanagements reflektiert und fachspezifisch umgesetzt werden.

### 3. Forschungsdatenlebenszyklus

Der Lebenszyklus von Forschungsdaten stellt die Schritte und Aspekte eines Forschungsvorhabens in Bezug auf das Forschungsdatenmanagement dar. Der Lebenszyklus nach [forschungsdaten.info](https://www.forschungsdaten.info) <sup>12</sup> basiert auf der beim UK Data Service erstellten Version<sup>13</sup> und beinhaltet die Planung, Erhebung, Aufbereitung und Analyse, Teilen und Publizieren, Archivierung sowie die Nachnutzung von Daten (siehe Abbildung 2).

---

<sup>9</sup> EU-Charta, Recht auf Privatsphäre (Art. 7) und den Schutz der Daten (Art. 8). Zugriff am 27.04.2023, <http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:12012P/TXT&amp;amp;from=DE>.

<sup>10</sup> Übersichtsarbeit über die Möglichkeiten der Gesichtsrekonstruktion von Gesichtern aus MRT-Daten. 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. [https://doi.org/10.1007/978-3-642-22348-8\\_26](https://doi.org/10.1007/978-3-642-22348-8_26).

<sup>11</sup> Zhang, Z. (2016). Mechanics of human voice production and control. *The journal of the acoustical society of america*, 140(4), 2614-2635.

<sup>12</sup> Der Datenlebenszyklus, Zugriff am 27.04.2023, <https://www.forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>.

<sup>13</sup> Data Lifecycle vom UK Data Service, Zugriff am 27.04.2023, <https://ukdataservice.ac.uk/learning-hub/research-data-management/>.

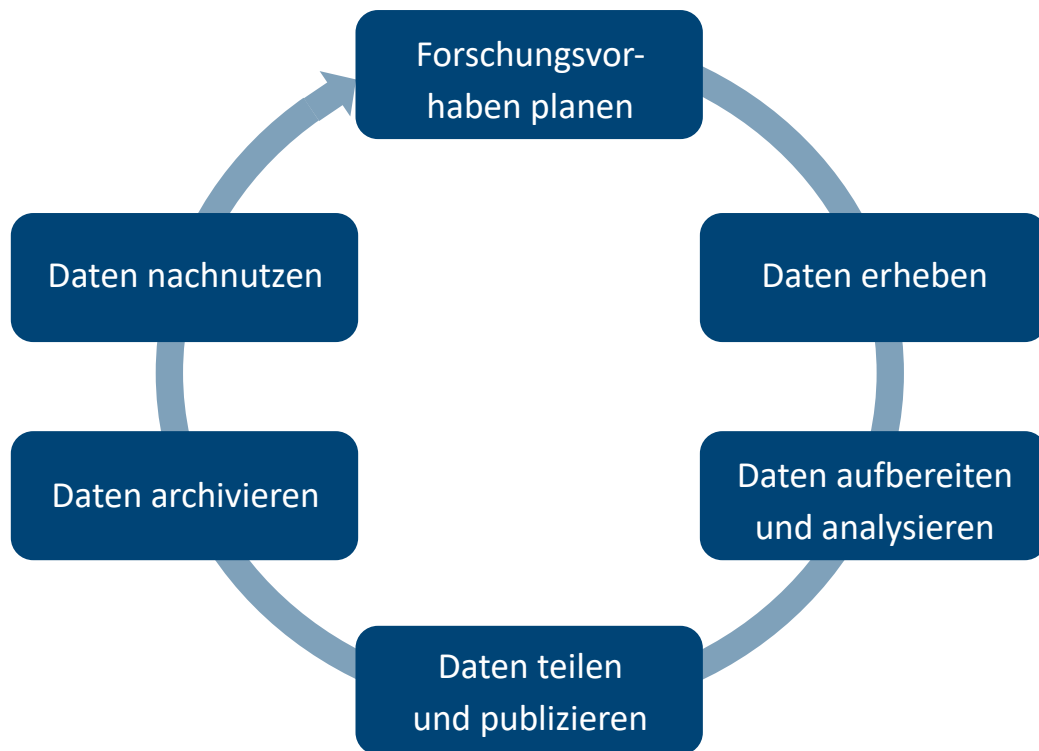


Abbildung 2: Forschungsdatenlebenszyklus nach [forschungsdaten.info](https://www.forschungsdaten.info/).

## Forschungsdatenmanagement (Einheit 4)

### 1. Motivation

Der Kodex „Leitlinien zur Sicherung guter wissenschaftliche Praxis“ der DFG<sup>14</sup> formuliert verschiedene Aspekte der wissenschaftlichen Integrität, die durch ein nachhaltiges, den FAIR-Kriterien entsprechendes Forschungsdatenmanagement (FDM) adressiert und umgesetzt werden können. Ziel ist es, Forschungsdaten langfristig und personenunabhängig zugänglich, nachnutzbar und nachprüfbar zu halten<sup>15</sup>, in dem es Aspekte der Qualitätssicherung (in Bezug auf Erhebung, Verarbeitung, Auswertung und Archivierung bis hin zur Veröffentlichung von Daten) einbezieht. FDM adressiert zudem die rechtlichen und ethischen Rahmenbedingungen eines Forschungsprozesses (Recht auf Forschungsfreiheit, Nutzungsrecht von Daten, Ethikvoten, Datenschutz). Instrumente wie der Datenmanagementplan regeln darüber hinaus die datenbezogenen Verantwortlichkeiten innerhalb eines Projekts und dokumentieren verwendete Datentypen, Software, Speicherung, Archivierung und Zugang zu den im Vorhaben generierten bzw. beforschten Daten.

Nicht zuletzt durch die Replikationskrise erlangte FDM in der Psychologie<sup>16</sup> Relevanz als notwendiger Bestandteil einer validen Forschung durch das Nachvollziehen der Ergebnisse und damit deren Reproduzierbarkeit. Es erhöhte auch das Bewusstsein für die Relevanz eines entsprechend systematisierten FDM, das mittel- und langfristig die Transparenz, Nachhaltigkeit, Interoperabilität und Zugänglichkeit datenbasierter Forschung zum Ziel hat.

Aus diesem Grund

- geben sich z. B. größere Forschungsprojekte, die als Verbund mehrere wissenschaftliche Institute auf nationaler und internationaler Ebene umfassen, eine für alle gleichermaßen gültige Leitlinie im Umgang mit Forschungsdaten;
- verabschieden wissenschaftliche Institute oder Fachbereiche entsprechende Policies im Umgang mit Forschungsdaten an deren Einrichtungen;
- unterstützen Forschungsförderer oder spezifische Förderprogramme diese Bestrebungen, indem sie ein nachhaltiges FDM während der Projektlaufzeit als auch darüber hinaus fordern.

### 2. Forschungsdatenmanagement

#### a) Was ist Forschungsdatenmanagement?

Forschungsdaten zählen heute zu den zentralen Ressourcen der Wissenschaft. Dementsprechend ist ein systematischer und verantwortungsvoller Umgang mit ihnen von grundlegender Bedeutung. Im Rahmen eines Forschungsdatenmanagements werden „die eigenen Arbeitsprozesse, die die Erzeugung von und den Umgang mit Forschungsdaten betreffen, möglichst effizient und zielorientiert“<sup>17</sup>

<sup>14</sup> Deutsche Forschungsgemeinschaft. Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex. Zenodo, 2019, <http://doi.org/10.5281/zenodo.3923602>.

<sup>15</sup> Forschungsdatenmanagement, Zugriff am 27.04.2023, <https://forschungsdaten.info/themen/informieren-und-planen/was-ist-forschungsdatenmanagement/>.

<sup>16</sup> Positionspapier der Psychologie-Fachschaften Konferenz zur Replikationskrise in der Psychologie. Zugriff am 27.04.2023, <https://psyfako.org/wp-content/uploads/30-PsyFaKo-STN-Open-Science-Kulturwandel-zu-einer-offenen-Wissenschaft.pdf>.

<sup>17</sup> Meyermann, Alexia: „Datenmanagement – Eine zentrale Voraussetzung für den Erfolg der Data Sharing - Idee in den Sozialwissenschaften.“ DSZ-BO Working Paper Series Nr. 2 (2012), S. 7. Zugriff am 27.04.2023, <http://docplayer.org/83618446-Dsz-bo-working-paper-series-working-paper-nr-2.html>.

organisiert und fortlaufend gesteuert. Damit begleitet FDM den gesamten Forschungsprozess von den ersten Planungen bis hin zur Archivierung, Nachnutzung oder Löschung der Daten.

Im Zuge dessen entwickeln Forschende Methoden und Richtlinien, die sie auf ihre mit Forschungsdaten verbundenen Praktiken anwenden. Eine solche Strategie im Umgang mit Daten hilft im darauffolgenden Forschungsprozess nicht nur Daten zu verwalten, sondern steuert und vereinheitlicht den Umgang mit ihnen auch. Insbesondere vermeidet sie doppelte Arbeiten wie z. B. erneute aufwendige Einarbeitung in die Daten, bspw. bei einer Anfrage während eines Review-Prozesses. Ein standardisierter FDM-Prozess (z. B. hinsichtlich der Datendokumentation, -sicherung und Langzeitarchivierung) verringert sowohl den zukünftigen Aufwand bei eigener und fremder Nachnutzung der Daten als auch das Risiko eines Datenverlustes. Zudem stehen bestimmte Förderquellen nur Projekten mit einem dezidierten FDM-Plan zur Verfügung.

Durch das Verschriftlichen der Strategie und ihrer geplanten Methoden und Richtlinien entsteht ein erster Datenmanagementplan. Er beinhaltet technische, organisatorische, strukturelle, rechtliche und ethische Aspekte des Umgangs mit Daten für die Dauer eines Projektes. Aber auch weiterreichende Aspekte, wie beispielsweise die Archivierung und Zugänglichmachung nach Projektende oder die eigene oder fremde Nachnutzung der Daten sollten hier berücksichtigt werden.

Ausgehend vom Forschungsdatenlebenszyklus beschreibt ein sogenannter Forschungsdaten-Workflow – mehr noch als ein Datenmanagementplan – die einzelnen Verarbeitungsschritte der Forschungsdaten in Abhängigkeit von der gewählten Software und den benötigten Infrastrukturen und Services. Dabei wird eine prozessorientierte Perspektive eingenommen, die es Datenerstellenden erlaubt, auch Datenübergaben und -konvertierungen, die bspw. zwischen Datenverarbeitung und -analyse notwendig sind, konkret abzubilden. In einem Workflow werden auch Verantwortlichkeiten (Rollen und Akteure) festgehalten.

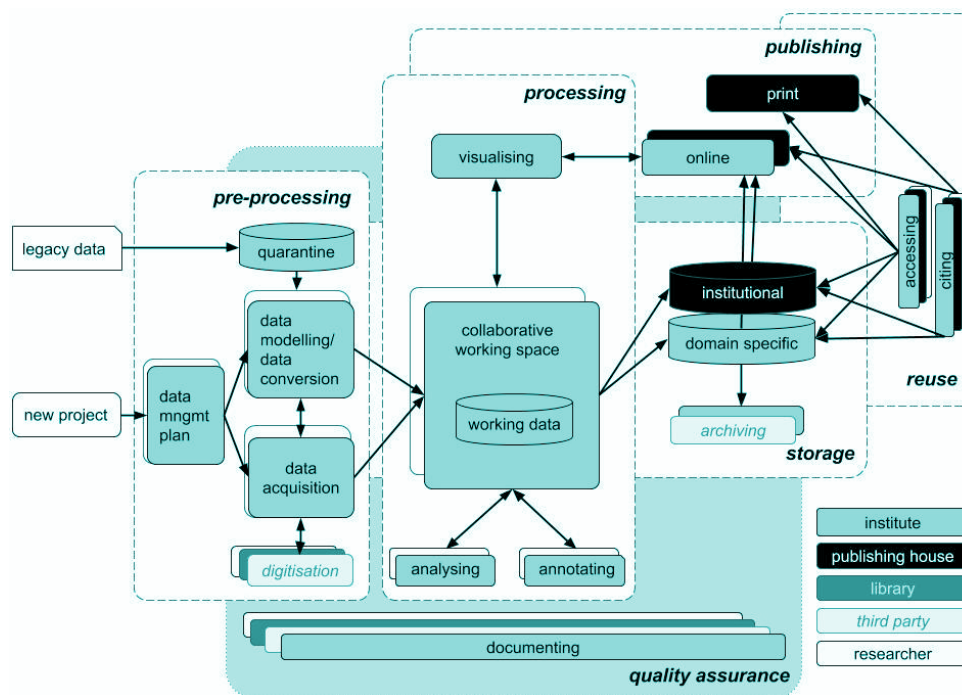


Abbildung 3: Beispielhafter Datenmanagementworkflow nach Wissik & Ďurčo 2015<sup>18</sup>

<sup>18</sup> Wissik, T., & Ďurčo, M. (2016, April). Research data workflows: from research data lifecycle models to institutional solutions. In Selected papers from the CLARIN annual conference 2015, October 14–16, 2015, Wrocław, Poland (No. 123, pp. 94-107). Linköping University Electronic Press.

Die genannten Arbeitsschritte und Instrumente im Rahmen des Forschungsdatenmanagements haben folgende Ziele:

- Nachvollziehbarkeit der Datenerhebung und –verarbeitung (z. B. keine verteilte Datenablage in unterschiedlichen Versionen auf verschiedenen Rechnern) sowie der daraus resultierenden Ergebnisse
- Gemäß den FAIR-Prinzipien<sup>1</sup> schnelleres Auffinden von Daten, bspw. durch aussagekräftige Benennung, Sicherstellung der plattformunabhängigen Datenverarbeitung durch Nutzung freier Datenformate, den Transfer der Daten in zukünftige eigene oder fremde Projekte oder die (halb-)automatische Verarbeitung durch Metadaten
- Vorbeugung von Datenverlust, u. a. aufgrund defekter Hard- oder Software oder von Ursprungsversionen von Dateien
- Weitergabe und Nachnutzung von Daten unter Verwendung von Einwilligungserklärungen
- optimierter Mitteleinsatz, bspw. Kostenersparnis durch Nachnutzung statt neuer Erhebung (Erhalt von Primär- und Sekundärdaten)
- Erfüllung von Auflagen der Drittmittelgeber
- Erleichterung der Zusammenarbeit zwischen Mitarbeitenden innerhalb eines Projektes bzw. Projektpartnern
- Wissenserhalt – Daten sind unabhängig von Personen, Projekten oder Institutionen zugänglich
- Referenzierbarkeit<sup>19</sup>

### b) Aufgaben des Forschungsdatenmanagements

Forschungsdatenmanagement umfasst sämtliche Schritte des Forschungsprozesses. Zentrale Aufgaben sind<sup>20</sup>:

- Planung des Umgangs mit Forschungsdaten zu Beginn eines Forschungsprojektes sowie ggf. Darstellung der geplanten Maßnahmen in Förderanträgen
- Festlegen einer Ordnerstruktur und Dateinamenskonventionen
- Dokumentation von Forschungsdaten und Auszeichnung mit Metadaten
- Backup und Langzeitarchivierung von Forschungsdaten
- IT-Sicherheit und Zugriffsrechte für Forschungsdaten
- Langzeitarchivierung von Forschungsdaten
- Publikation von Forschungsdaten
- Auffinden und Nachnutzen bestehender Forschungsdaten
- Berücksichtigung von Datenschutz- und Urheberrecht im Umgang mit Forschungsdaten

---

<sup>19</sup> Piwowar, Heather A., Roger S. Day und Douglas B. Fridsma. Sharing detailed research data is associated with increased citation rate. PLoS ONE 2 (2007). S. e308.  
<https://doi.org/10.1371/journal.pone.0000308>.

<sup>20</sup> Louise Corti, Veerle Van den Eynden, Libby Bishop und Matthew Woollard. Managing and Sharing Research Data: A Guide to Good Practice. Los Angeles, CA: SAGE, 2014.

### c) Die FAIR-Prinzipien

FORCE11, eine Gruppe Forschender und Mitarbeitender von Bibliotheken, Archiven, Verlagen und Forschungsförderern, hat 2016 Prinzipien zur Erschließung von Forschungsdaten verfasst<sup>21</sup>. Diese sogenannten FAIR-Prinzipien umfassen vier Ziele<sup>22</sup>:

- **Findable** (Auffindbarkeit)
- **Accessible** (Zugänglichkeit)
- **Interoperable** (Interoperabilität, d. h. mit anderen Daten kompatibel und von möglichst vielen Analyse- und Visualisierungsapplikationen nutzbar)
- **Reusable** (Wiederverwendbarkeit)

Durch die Anwendung bzw. Umsetzung der FAIR-Prinzipien werden Daten und Metadaten menschen- und maschinenlesbar, wodurch sie effektiver auffindbar und nachnutzbar sind. Darüber hinaus sollten Daten und deren Metadaten so archiviert werden, dass sie mit Hilfe von Standard-Kommunikationsprotokollen leicht und langfristig von Mensch und Maschine abgerufen, heruntergeladen oder lokal genutzt werden können. Die Daten sollten in einer Form vorliegen, in der sie in einer (teil-)automatisierten Weise mit anderen Datensätzen ausgetauscht, interpretiert und kombiniert werden können. Eine gute Beschreibung der Daten und deren Metadaten – idealerweise in standardisierter Form, siehe Abschnitt Metadatenstandard 8.3 – gewährleistet dabei die Wiederverwendung der Daten für zukünftige Forschung und den Vergleich mit anderen kompatiblen Datenquellen. Dazu zählen außerdem ordnungsgemäßes Zitieren der Daten und eine für Mensch und Maschine eindeutige Darstellung der Nachnutzungsbedingungen.

Das Hauptziel der FAIR-Prinzipien ist die Nachnutzbarkeit von Forschungsdaten<sup>23</sup>. Das bedeutet nicht, dass alle Daten offen und uneingeschränkt zugänglich sein müssen. Eine eingeschränkte Zugänglichkeit, beispielsweise aus datenschutzrechtlichen Gründen, widerspricht den FAIR-Prinzipien nicht, solange die Metadaten verfügbar und zugänglich sind. Ein sorgfältig bereits von Beginn an geplantes Forschungsdatenmanagement ist eine der Grundlagen sowohl für FAIR wie auch für Open Data (eines der 6 Prinzipien von Open Science<sup>24</sup>), denn viele der Weichen für eine möglichst weitreichende Umsetzung dieser beiden Prinzipien werden bereits sehr früh im Forschungsprozess gestellt. Darüber hinaus unterstützt das Forschungsdatenmanagement auch Aspekte, die durch FAIRness und Offenheit nicht notwendigerweise abgedeckt werden - etwa Langzeitarchivierung und gute Datenqualität<sup>25</sup>. Verschiedene Tools<sup>26</sup> bieten online Fragenkataloge zur selbstständigen Überprüfung der FAIRness an. Je höher die FAIRness und je offener die Daten sind, desto wahrscheinlicher ist die Nachnutzung.

---

<sup>21</sup> Wilkinson, Mark D., Michel Dumontier, IJsbrand J. Aalberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. „The FAIR Guiding Principles for scientific data management and stewardship.“ *Scientific Data* 3 (2016). <https://doi.org/10.1038/sdata.2016.18>.

<sup>22</sup> GO FAIR International Support and Coordination Office. „FAIR Principles.“ Zugriff am 27.04.2023, <https://www.go-fair.org/fair-principles/>.

<sup>23</sup> Boeckhout, M., Zielhuis, G.A. & Bredenoord, A.L. The FAIR guiding principles for data stewardship: fair enough?. *Eur J Hum Genet* 26, 931–936 (2018). <https://doi.org/10.1038/s41431-018-0160-0>.

<sup>24</sup> Gallagher, R. V., Falster, D. S., Maitner, B., Salguero-Gomez, R., Vandvik, V., Pearse, W., & Enquist, B. J. (2019). The open traits network: Using open science principles to accelerate trait-based science across the tree of life. *Nature Ecology & Evolution*.

<sup>25</sup> Higman, R., Bangert, D., & Jones, S. (2019). Three camps, one destination: the intersections of research data management, FAIR and Open. *Insights*, 32(1).

<sup>26</sup> Bspw. FAIRsFAIR, FAIR Aware. Zugriff am 27.04.2023, <https://www.fairsfair.eu/fair-aware> oder Australian Research Data Commons, FAIR self-assessment tool. Zugriff am 27.04.2023, <https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/>. Kollektion verschiedener Tools. Zugriff am 27.04.2023, <https://fairassist.org/#/>.



### 3. (Psychologie-spezifische) Forschungsdaten-Policies

#### a) Was ist eine Forschungsdaten-Policy?

Policies und Leitlinien beschreiben Vorgaben zum Umgang mit Forschungsdaten. Dabei gibt es verschiedene Arten:

- Fachspezifische Policies (siehe Psychologie-spezifische Policies)
- Leitlinien von Forschungsförderern
- Institutionelle Policies und Leitlinien
- Projektspezifische Policies
- Journal- und Verlags-Policies und Leitlinien

In den nachfolgenden Kapiteln werden vor allem die für die Fachspezifik relevanten Formen (Fachspezifische Policies, Leitlinien von Forschungsförderern, Institutionelle Policies und Leitlinien) näher erläutert, und mit Beispielen ergänzt.

#### b) Psychologie-spezifische Policies

Für einige Fachgebiete gibt es bereits spezifische Richtlinien für den Umgang mit Forschungsdaten (u. a. Psychologie, Genetik, Linguistik, Bildungsforschung, Sozialwissenschaften) <sup>27</sup>. Solche fachspezifischen Anforderungen sind notwendig, da sich die Art der Daten (bspw. bildgebende Formate, Interviews, Umfragen, Farbspektren usw.) und deren Quelle (z. B. Personen, chemische/physikalische Experimente usw.) von Fachbereich zu Fachbereich stark unterscheiden und damit auch der Umgang mit ihnen. Gleichzeitig dient die Formulierung dieser Standards innerhalb eines Fachbereichs dazu, Vergleichbarkeit und Interoperabilität der Daten zu spezifischen Fragestellungen herzustellen (z. B. Erhebungen zum Thema Stress).

So gab die DGPs 2016<sup>28</sup> eine in der Folge mehrfach überarbeitete Empfehlung zum allgemeinen Umgang mit Forschungsdaten<sup>29</sup> für die Psychologie in Deutschland heraus. In der Handreichung wird zwischen **Rohdaten** (Ursprungsaufzeichnungen: Fragebogen, Zeichnungen, Audio- oder Videoaufnahmen, Blickbewegungsmessungen, neuro- oder peripherphysiologische Aufzeichnungen), **Primärdaten** (unbearbeitete – untransformierte, nicht-aggregierte etc. –, in digitaler Form vorliegende quantitative oder qualitative Daten) und **Sekundärdaten** (Daten nach ersten Verarbeitungsschritten: Transformationen von Variablen, Aggregation einzelner Beobachtungen, Berechnung von Skalenwerten, Berechnung von ereigniskorrelierten Potentialen durch Mittelung, oder der Ausschluss von Werten, von denen angenommen wird, dass sie die Robustheit der Analyseergebnisse beeinträchtigen) unterschieden.

Hinsichtlich **personenbezogener Daten** adressieren die Leitlinien die informierte Einwilligung sowie die Veröffentlichung der Daten zur Nachvollziehbarkeit der Ergebnisse und Nachnutzung durch andere Forscher\*innen. In diesem Zusammenhang wird zwischen **Typ 1-Daten** (Veröffentlichung als Bestandteil einer Publikation) und **Typ 2-Daten** (Veröffentlichung nach Projektabschluss) unterschieden und Hinweise zur Wahl eines geeigneten Repositoriums für die Publikation und/oder die Archivierung der Daten gegeben.

<sup>27</sup> Übersicht über fachspezifische Leitlinien zum Umgang mit Forschungsdaten, Zugriff am 27.04.2023, [https://www.dfg.de/foerderung/grundlagen\\_rahmenbedingungen/forschungsdaten/empfehlungen/index.html](https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/forschungsdaten/empfehlungen/index.html).

<sup>28</sup> DGPs „Empfehlungen zu einem offenen und nachhaltigem Datenmanagement“. Zugriff am 27.04.2023, <https://www.dgps.de/schwerpunkte/transparente-wissenschaft/umgang-mit-forschungsdaten/>.

<sup>29</sup> Management und Bereitstellung von Forschungsdaten in der Psychologie: Überarbeitung der DGPs-Empfehlungen. Zugriff am 27.04.2023, <https://psyarxiv.com/hcxtm/>.



Die Revision von 2020 nimmt zudem Aspekte der Zugriffs- und Nutzungsrechte Dritter (Stichwort **Zugriffsklassen**), strukturelle Anreize der Bereitstellung entsprechender Dienste an den Instituten zur Umsetzung der Leitlinie sowie den Umgang mit Konfliktfällen (z. B. bei Verstößen gegen gesetzliche Vorgaben) in einem Psychologie-spezifischen Kontext in den Blick. Für die Lebenswissenschaften liefern die „Gute Klinische Praxis (GCP)“ und die „Grundsätze der Guten Laborpraxis (GLP)“ Regularien für das Forschungsdatenmanagement. Beide Grundsätze sind in Deutschland gesetzlich bindend<sup>30</sup>.

### c) Leitlinien von Forschungsförderern

Für den Fachbereich Psychologie stellen verschiedene Forschungsförderer Leitlinien für das Forschungsdatenmanagement bereit, darunter die Deutsche Forschungsgesellschaft (DFG<sup>31</sup>), das vom European Research Council (ERC) geführte Programm Horizon 2020<sup>32</sup> sowie das Bundesministerium für Bildung und Forschung (BMBF).

So sind u. a. ausführliche Angaben zum Umgang mit Forschungsdaten im Rahmen der Drittmittelbeantragung verpflichtend. Dies beinhaltet u. a.:

- die Methodik der Datenerhebung bzw. eine Begründung, warum existierende Daten nicht genutzt werden
- die Ansätze zur Beschreibung der Daten (z. B. Metadaten, Dokumentationsstandards)
- die Art der Methoden und Tools zur Verarbeitung der Daten
- Ort der Langzeitarchivierung der Daten sowie geplante Nutzungs- und Zugriffsrechte
- die für das Projekt relevanten rechtlichen Verpflichtungen und Bedingungen:
  - Auswirkungen oder Einschränkungen in Bezug auf eine spätere Veröffentlichung oder Zugänglichkeit
  - Aspekte der Nutzung und des Urheberrechts
  - Eigentumsfragen
- die Verantwortlichkeiten und Ressourcen
  - Beschreibung der Rollen und Verantwortlichkeiten innerhalb des Projekts
  - benötigte Ressourcen (Kosten, Zeit oder andere), um angemessenen Umgang mit Forschungsdaten zu realisieren
  - Verantwortlichkeiten nach Beendigung des Projektes

ERC-Einreichungen sehen die Vorlage eines Datenmanagementplans innerhalb von 6 Monaten nach Projektstart vor. Dabei kommen ähnliche Anforderungen wie seitens der DFG zum Tragen, wie etwa die Beschreibung der zu erhebenden Daten, Protokolle zur Strukturierung der Daten, die genutzten Metadatenstandards, die Benennung der persistenten Identifier, die Datenpflege (z. B. Backups, Bereinigung), Aspekte der Langzeitarchivierung sowie geplante Nutzungs- und Zugriffsrechte).

Für Anträge, die beim BMBF im Rahmen spezifischer Richtlinien bzw. Förderinitiativen eingereicht werden, können Informationen in Bezug auf den Umgang mit Forschungsdaten erforderlich sein. Dies ist abhängig vom jeweiligen Förderprogramm. So wird beispielhaft im Programm zur Förderung von

---

<sup>30</sup> BMJV. „Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung an Menschen.“ Zugriff am 27.04.2023, [https://www.bfarm.de/DE/Arzneimittel/Klinische-Pruefung/GCP-Inspektorat/\\_node.html](https://www.bfarm.de/DE/Arzneimittel/Klinische-Pruefung/GCP-Inspektorat/_node.html).

<sup>31</sup> Deutsche Forschungsgemeinschaft. „Leitlinien zum Umgang mit Forschungsdaten.“ Zugriff am 27.04.2023, <https://zenodo.org/record/6472827#.Yt40rcHP2Zw>

<sup>32</sup> Europäische Kommission. „Guidelines on FAIR Data Management in Horizon 2020. Version 3.0.“ Zugriff am 27.04.2023, [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf).

Forschungsprojekten zum Thema „Nicht-staatliche Hochschulen“ (Bundesanzeiger vom 27.04.2022<sup>33</sup>) eine Verpflichtung verlangt, Daten mit „etwaiger Relevanz zur Nutzung durch Dritte nach Abschluss des Projekts in weitergabefähiger Form auf der Basis gängiger Standards einer geeigneten Einrichtung/einem Forschungsdatenzentrum zur Verfügung zu stellen mit dem Ziel, langfristige Datensicherung, Sekundärauswertungen oder eine Nachnutzung zu ermöglichen“. Diese Vorgabe soll mittels eines eigenen Forschungsdatenmanagementplans unterstützt werden, der bei Einreichung vorgelegt werden muss. Da Förderprogramm-unabhängige Richtlinien seitens des BMBF fehlen, sollte sich im Falle des Fehlens konkreter Vorgaben bei Förderrichtlinien an den gängigen Empfehlungen der DGPs orientiert werden.

#### d) Institutionelle Policies und Leitlinien (in der Psychologie)

Die Mehrzahl der Hochschulen und außeruniversitären Forschungseinrichtungen haben bereits eine institutionelle Forschungsdaten-Policy eingeführt, um den Umgang mit Forschungsdaten zu regeln und die grundsätzlichen rechtlichen Aspekte zu klären<sup>34</sup>. Fachspezifische Policies an Forschungsinstituten sind daher in der Regel nicht nötig.

Eigenständige Forschungsinstitute im Bereich der Psychologie hingegen, wie bspw. das Leibniz-Institut für Psychologie, das ZI Mannheim<sup>35</sup> oder das IfaDo<sup>36</sup>, veröffentlichen entsprechend eigene Policies. Die Humboldt Universität Berlin verweist zudem in Bezug auf Verhaltensforschung neben den DGPs-Richtlinien auch auf die FDM-Empfehlung des Rates für Sozial- und Wirtschaftsdaten<sup>37</sup>.

In institutionellen Policies kann neben der Regelung, dass Forschungsdaten und Forschungsergebnisse im Open Access Format veröffentlicht werden sollen, auch die Kosten- und Ressourcensteuerung<sup>38</sup>, also die personellen, organisatorischen und technischen Kapazitäten der Forschungseinrichtung für das FDM festgelegt werden. Bei der Hälfte der deutschen universitären FD-Policies werden Aussagen zu den Kosten des FDM im Rahmen eines Datenmanagementplans als wichtig erachtet.

Abbildung 4 zeigt die Elemente einer institutionellen Forschungsdaten-Policy und listet inhaltliche Punkte, die in diesen Kategorien geregelt werden. Ein Beispiel dafür ist die im Jahr 2021 eingeführte Forschungsdaten-Policy der Freien Universität zu Berlin (s. begleitende Materialien).

<sup>33</sup> Richtlinie zur Förderung von Forschungsprojekten zum Thema „Nicht-staatliche Hochschulen“, Bundesanzeiger vom 27.04.2022. Zugriff am 27.04.2023, <https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2022/04/2022-04-27-Bekanntmachung-Nicht-staatl-Hochschulen.html>.

<sup>34</sup> Liste der Forschungsdaten-Policies an deutschen Hochschulen: forschungsdaten.org: „Data Policies“. Zugriff am 27.04.2023, [https://www.forschungsdaten.org/index.php/Data\\_Policies#Institutionelle\\_Policies](https://www.forschungsdaten.org/index.php/Data_Policies#Institutionelle_Policies).

<sup>35</sup> ZI Mannheim, „Handling your data“. Zugriff am 27.04.2023, <https://www.zi-mannheim.de/en/service/rechtsverstoesse-melden/handling-your-data.html>, „Privacy policy“, Zugriff am 27.04.2023, <https://www.zi-mannheim.de/en/service/privacy-policy.html>.

<sup>36</sup> Forschungsdatenpolicy des IfaDo. Zugriff am 27.04.2023, <https://www.ifado.de/datenmanagement/>.

<sup>37</sup> Humboldt Universität Berlin. Zugriff am 27.04.2023, [https://www.konsortswd.de/wp-content/uploads/RatSWD\\_Output3\\_Forschungsdatenmanagement.pdf](https://www.konsortswd.de/wp-content/uploads/RatSWD_Output3_Forschungsdatenmanagement.pdf).

<sup>38</sup> Hiemenz, Bea Maria und Monika Kuberek: Leitlinie? Grundsätze? Policy? Richtlinie? – Forschungsdaten-Policies an deutschen Universitäten. *o-bib* 5 (2018), S. 1–13. <https://doi.org/10.5282/o-bib/2018H2S1-13>.

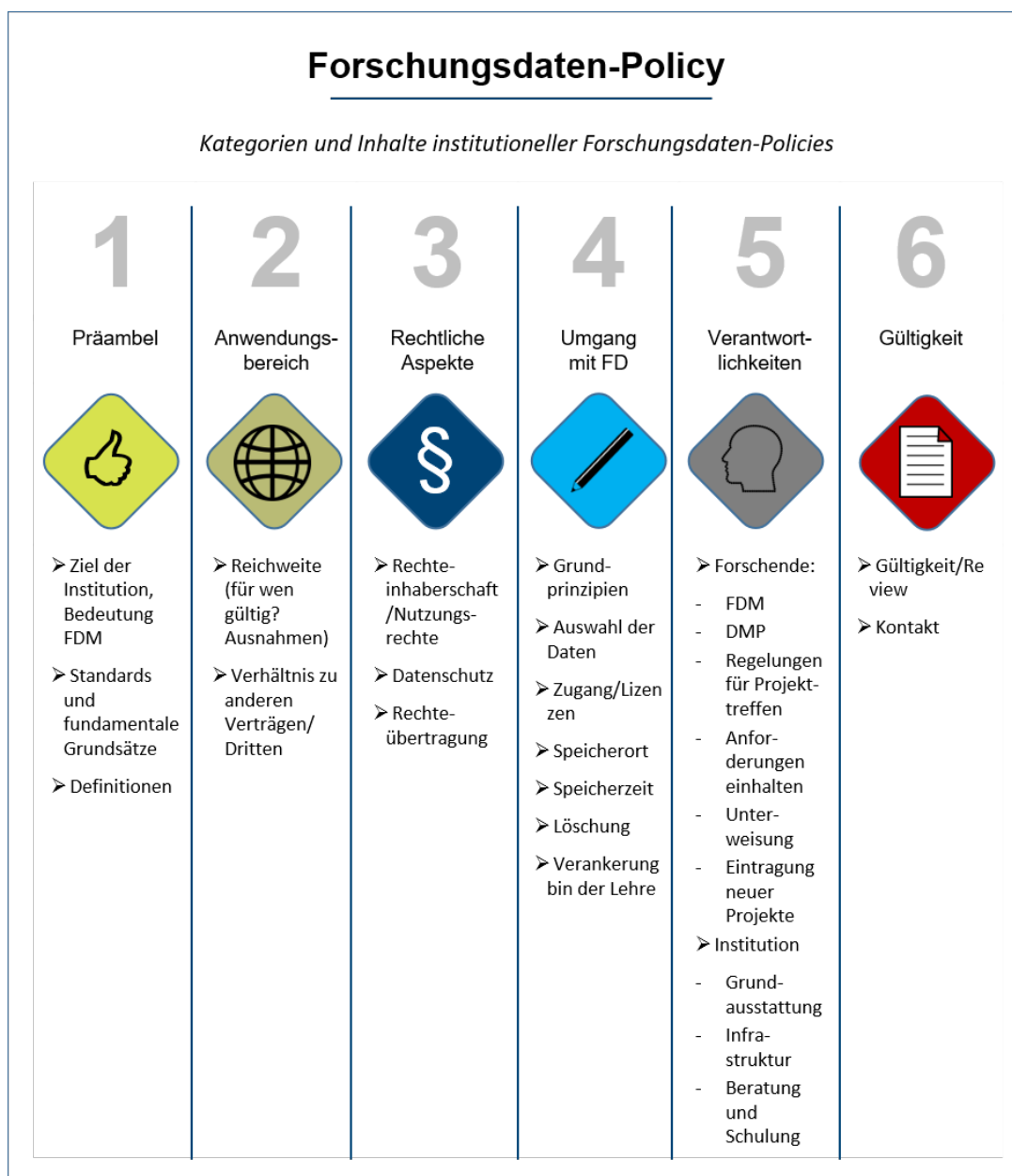


Abbildung 4: Visuelle Aufbereitung einer Forschungsdaten-Policy. Entnommen aus dem strategischen Leitfaden zur Erstellung einer Forschungsdatenpolicy<sup>39</sup>.

<sup>39</sup> Hiemenz, Bea, & Kuberek, Monika. (2018). Empfehlungen zur Erstellung institutioneller Forschungsdaten-Policies. Das Forschungsdaten-Policy-Kit als generischer Baukasten mit Leitfragen und Textbausteinen für Hochschulen in Deutschland. <https://doi.org/10.14279/depositonce-7521>.

## Rechtliche Aspekte (Einheit 5)

### 1. Ebenen der Regelung

Je nach der Beschaffenheit der Daten müssen rechtliche und ethische Rahmenbedingungen, die auf sehr verschiedenen Ebenen verortet sind, beim Forschungsdatenmanagement berücksichtigt werden (siehe Tabelle 2). So gelten, wenn es um personenbezogene Daten geht, die EU-Datenschutz-Grundverordnung (DSGVO), das Bundesdatenschutzgesetz, die Landesdatenschutzgesetze und, falls vorhanden, die Landeshochschulgesetze sowie je nach Fall weitere datenschutzrechtliche Bestimmungen, z. B. im Sozialgesetzbuch, Genetikgesetz, Arzneimittelgesetz oder Schulgesetz. Darüber hinaus kann es auch auf institutioneller Ebene Richtlinien, Policies oder auch eigene Ethikrichtlinien geben. Grundsätzlich gelten für Forschende die „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“<sup>40</sup> der Deutschen Forschungsgemeinschaft.

In manchen Fällen ist es notwendig, den Zugang zu den Daten einzuschränken, vor allem wenn ethische oder rechtliche Aspekte eine Rolle spielen, oder die Forschung Vertraulichkeitsabreden unterliegt, bspw. bei Auftragsforschung.

*Tabelle 2: Beispiele der rechtlichen Handlungsfelder, von denen die Publikation von Forschungsdaten betroffen sein kann<sup>41</sup>.*

<b>Patentrecht</b>	<b>Urheberrecht</b>	<b>Wettbewerbsrecht</b>	<b>Datenschutz</b>
Was ist zu beachten, wenn FD Patentreife erlangen (können)?	Unterliegen FD überhaupt dem Urheberrechtsgesetz?	Werden Daten im unternehmerischen Geschäftsverkehr (nicht) einvernehmlich genutzt?	Welche FD sind "schützenswert"?
<b>Wissenschaftsrecht</b>	<b>Grundrechte</b>	<b>Internationales Recht</b>	<b>EU-Recht</b>
Können Lizenz- und Veröffentlichungsvorgaben für FD per Mandatierung erfolgen?	Welche verfassungsrechtlichen Grenzen sind zu beachten?	Welche Rechtsbestimmungen bestehen außerhalb Deutschlands?	Was bringt bspw. die "European Data Economy" für FD?
<b>Verträge</b>	<b>Arbeits-/Dienstrecht</b>	<b>Förderbedingungen</b>	<b>Policies</b>
Bestehen Absprachen zum "geistigen Eigentum" an FD?	Wem "gehören" die an Hochschulen erhobenen FD?	Welche Bedingungen geben Förderer (DFG; Industrie) vor?	Welche rechtliche Bindung können Policies entfalten?

<sup>40</sup> Deutsche Forschungsgemeinschaft. *Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex*. Zenodo, 2022. Zugriff am 27.04.2023, [https://zenodo.org/record/6472827#.Yt\\_WXdrwBaQ](https://zenodo.org/record/6472827#.Yt_WXdrwBaQ)

<sup>41</sup> nach Hartmann, Thomas: „terra incognita – digitale Forschungsdaten auf der Suche nach einer rechtlichen Heimat“. „Beitrag bei der Workshoptagung Rechtliche Aspekte bei digitalen Forschungsdaten“ am 31.01.2018, Europa-Universität Viadrina, Frankfurt (Oder), Zugriff am 27.04.2023, [https://www.forschungsdaten.org/images/0/08/Hartmann\\_TerraIncognita-Forschungsdaten-RechtlicheHeimat.pdf](https://www.forschungsdaten.org/images/0/08/Hartmann_TerraIncognita-Forschungsdaten-RechtlicheHeimat.pdf).

## 2. Datenschutz und Persönlichkeitsrecht

In der europäischen Union regelt die DSGVO die Belange des Datenschutzes und des Persönlichkeitsrechts ihrer Bürger\*innen, welches grundsätzlich den Umgang mit personenbezogenen Daten unter Erlaubnisvorbehalt untersagt. Sie regelt auf europäischer Ebene das Recht auf den Schutz personenbezogener Daten und stellt den freien Verkehr personenbezogener Daten sowohl innerhalb der EU als auch gegenüber Drittstaaten sicher.

Durch eine sogenannte „**informierte Einwilligung**“ (informed consent) des Teilnehmenden ist es jedoch Forschenden gestattet, personenbezogene Daten zu erheben und zu verarbeiten. Hierbei handelt es sich um eine unmissverständlich abgegebene Einwilligung seitens des Teilnehmenden in Form einer informierten Erklärung oder einer sonstigen eindeutigen zustimmenden Handlung, die erkennen lässt, dass der\*diejenige über den genauen Zweck der Datenerhebung, -verarbeitung und -nutzung aufgeklärt ist. Diese Zustimmung kann auch mündlich erfolgen (z. B. bei Analphabeten). Bei Studien, die eine Täuschung beinhalten (zur Vermeidung des Bestätigungsfehlers, engl. *confirmation bias*<sup>42</sup>), ist nach der Aufklärung über das wahre Ziel<sup>43</sup> die Einholung einer zweiten Einwilligung notwendig, welche sich nur noch auf die Verwendung der Daten selbst beziehen muss<sup>44</sup>.

Ist die Publikation von Forschungsdaten geplant, muss diese Information in die Aufklärung aufgenommen werden. Wichtig ist dabei vor allem, alle geplanten Nutzungszwecke in der Einwilligung zu erfassen (so u. a. die Langzeitarchivierung, Datenpublikation, weitere Auswertungen zu anderen Fragestellungen usw.), da es im Nachhinein häufig nicht mehr gelingt, die Erweiterung der Einwilligung einzuholen.

Die Artikeln 4 und 9 der DSGVO gehen ausführlicher auf die Definition von **personenbezogenen Daten** ein. Im ersten Artikel werden personenbezogene Daten als „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person [...] beziehen“ definiert. Als identifizierbar werden die Betroffenen angesehen, wenn sie direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen identifiziert werden können, die Ausdruck der physischen, psychologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Personen sind.

Als eine besondere Form der personenbezogenen Daten gelten diejenigen, die eines erhöhten Schutzes bedürfen (Art. 9 DSGVO). Dazu gehören Daten, „aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder Gewerkschaftszugehörigkeit hervorgehen“, sowie genetische und biometrische Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person.

---

<sup>42</sup> Hugo Mercier: Confirmation bias – Myside bias, in: Rüdiger F. Pohl (Hrsg.): Cognitive illusions: Intriguing phenomena in thinking, judgment and memory. 2. Auflage. Routledge, London and New York 2017

<sup>43</sup> Deutsche Gesellschaft für Psychologie (DGPs). (2018). Ethisches Handeln in der psychologischen Forschung. S. 41.

<sup>44</sup> Deutsche Gesellschaft für Psychologie (DGPs). (2018). Ethisches Handeln in der psychologischen Forschung. S. 45; Rat für Sozial- und Wirtschaftsdaten (RatSWD) (Ed.) (2017): Forschungsethische Grundsätze und Prüfverfahren in den Sozial- und Wirtschaftswissenschaften, RatSWD Output, No. 9 (5), Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin, <https://doi.org/10.17620/02671.1>.

Die folgenden Seiten bieten Informationen, Hilfestellungen und Templates zur Erstellung von Einwilligungserklärungen:

- Hilfestellungen und eine Sammlung von Templates auf TransMIT-Zentrum für wissenschaftlich-psychologische Dienstleistungen (DGPs)<sup>45</sup>
- Sammlung von Hilfestellungen für Einwilligungsunterlagen auf den Seiten von KonsortSWD<sup>46</sup>
- Sammlung auf Forschungsdaten-Bildung<sup>47</sup>

### 3. Anonymisierung der in der Psychologie relevanten Daten

Insbesondere in der Psychologie ist die Nutzung personenbezogener Daten weit verbreitet. Hier hat die DGPs für die Psychologie eine datenschutzrechtliche Empfehlung zur Erstellung einer Einwilligungserklärung im Rahmen von Forschungsvorhaben erstellt<sup>48</sup>. Neben den bereits genannten Aspekten adressiert diese Empfehlung u. a. auch die Studienbeschreibung, Datenauswertung und -nachnutzung. Enthalten sind zudem ein Template für eine Einwilligungserklärung, dass die Vorgaben der DSGVO umsetzt, sowie Kontaktangaben zu den Datenschutzbeauftragten des Bundes und der Länder.

Um Daten einfacher und rechtssicher verarbeiten zu können, existieren verschiedene Möglichkeiten, den Personenbezug zu entfernen bzw. eine Anonymisierung der Daten zu erreichen. So können numerische Daten beispielsweise durch die Aggregation (Mittelung der Datenwerte) mittels herkömmlicher Analysesoftware wie R oder SPSS anonymisiert werden. Eine weitere Option der Anonymisierung ist das komplette Entfernen personenbezogener Daten oder das Ersetzen dieser durch Informationen, die nicht auf einzelne Personen rückführbar sind (bspw. PLZ statt Straße). Letzteres kann zum Beispiel mittels der Software QualiAnon<sup>49</sup> bewerkstelligt werden.

In Bezug auf MRT-Daten wird empfohlen, neben der Vergabe von Dateinamen ohne Personenbezug auch die Informationen in den Aufnahmen zu entfernen, die für die Forschungsfrage u. U. keine Relevanz haben (Knochenstruktur, extra-kraniales Gewebe)<sup>50</sup>. Dafür gibt es zwei weit verbreitete Ansätze, das sogenannte „skull stripping“ (Abisolieren des Schädels) und das „defacing“ (Entfernen von Gesichtsstrukturen). Zur Umsetzung sind verschiedene Möglichkeiten entwickelt worden, zum einen spezifische Programmroutinen (z. B. `ft_defacevolumes` für Fieldtrip<sup>51</sup>) oder speziell dafür entwickelte, open-source Software (z. B. `MRI_deface`<sup>52</sup>, `Pydeface`<sup>53</sup>). Um die (Meta-)Daten in den Aufnahmen zu entfernen, die auch zur Identifizierung spezifischer Probanden (Geschlecht, Alter, Name

<sup>45</sup> TransMIT-Zentrum, Zugriff am 27.04.2023, <https://zwpd.transmit.de/zwpd-dienstleistungen/zwpd-ethikkommission/vorlagen-antragstellung>.

<sup>46</sup> Hilfestellungen für Einwilligungserklärungen auf KonsortSWD, Zugriff am 27.04.2023, <https://www.konsortswd.de/ratswd/best-practice-forschungsethik/einwilligungsunterlagen/>.

<sup>47</sup> Hilfestellungen und Templates in Bezug auf Einwilligungserklärungen auf Forschungsdaten Bildung, Zugriff am 27.04.2023, <https://www.forschungsdaten-bildung.de/einwilligung>.

<sup>48</sup> Datenschutzrechtliche Empfehlung der DGPs. Zugriff am 27.04.2023, [https://zwpd.transmit.de/images/zwpd/dienstleistungen/ethikkommission/0.1a\\_datenschutzrechtliche\\_empfehlungen\\_einwilligungsforschungsvorhaben.pdf](https://zwpd.transmit.de/images/zwpd/dienstleistungen/ethikkommission/0.1a_datenschutzrechtliche_empfehlungen_einwilligungsforschungsvorhaben.pdf).

<sup>49</sup> Anonymisierungs-Applikation QualiAnon. Zugriff am 27.04.2023, <https://www.qualiservice.org/de/helpdesk/webinar/tools.html>.

<sup>50</sup> De-Identifizierungsmethoden: 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. [https://doi.org/10.1007/978-3-642-22348-8\\_26](https://doi.org/10.1007/978-3-642-22348-8_26).

<sup>51</sup> De-Identifizierung in Fieldtrip. Zugriff am 27.04.2023, [https://www.fieldtriptoolbox.org/faq/how\\_can\\_i\\_anonymize\\_an\\_anatomical\\_mri/](https://www.fieldtriptoolbox.org/faq/how_can_i_anonymize_an_anatomical_mri/).

<sup>52</sup> `MRI_deface`. Zugriff am 27.04.2023, <https://surfer.nmr.mgh.harvard.edu/fswiki/AutomatedDefacingTools>.

<sup>53</sup> `Pydeface`. Zugriff am 27.04.2023, <https://github.com/poldracklab/pydeface>.



usw.) benutzt werden können, gibt es ebenfalls frei verfügbare Software (u. a. die De-Identification Toolbox<sup>54</sup> oder GDCM<sup>55</sup>).

Da für M/EEG Daten gerade die individuellen Muster der Gehirnaktivität Grundlage der Forschung sind, sind Methoden zur De-Identifizierung wenig sinnvoll. Nichtsdestotrotz sollte auf größtmögliche Anonymität bei den Dateinamen und Metadaten geachtet werden.

Pseudonyme, d. h. die Ersetzung einer Information durch eine ähnliche, aber nicht mehr personenbeziehbare Information, sind ebenfalls eine Möglichkeit, um qualitative Daten ohne informierte Einwilligung weiterverarbeiten zu können. Eine Veröffentlichung ist aber ohne entsprechende Maßnahmen zur Sicherung gesetzlicher Vorgaben nicht ohne weiteres möglich (z. B. Einschränkung durch Zugriffsklassen oder Lizenzen). Weitere Schutzmaßnahmen bieten Zugriffsbeschränkungen mittels Passwortschutz und die Vergabe von Zugriffsrechten. So kann Zugang zu den Daten z. B. nur an bestimmte IPs oder nur für definierte Personenkreise gewährt werden (siehe Einheit 9 "Zugriffssicherheit", S. 40).

Die rasante Entwicklung von digitalen Medien führt heutzutage mehr als zuvor zur unbedachten Veröffentlichung von personenbezogenen Daten – zum Beispiel in Form von Bildern oder Videos. Das Recht am eigenen Bild umfasst jedoch das Recht eines jeden Einzelnen zu entscheiden, für welchen Zweck die Bilder oder andere Formen von Abbildungen der eigenen Person zur Verfügung stehen. Nur mit Einwilligung können diese verbreitet oder veröffentlicht werden.

Als eine Einwilligung gilt sowohl eine Entlohnung als auch die Beantwortung von Fragen vor laufender Kamera. Ausnahmen hiervon bilden:

- zeitgeschichtliche Abbildungen
- Bilder, auf denen die Person nicht als das Hauptmotiv dargestellt wird (z. B. bei einer Aufnahme einer Demonstration oder einer belebten Einkaufsstrasse)
- Bilder von Versammlungen o. ä.

Video- und auch Tonaufnahmen können jedoch editiert und ebenfalls anonymisiert bzw. de-personalisiert werden. Beispiele hierfür sind das Verfremden (Verpixeln, schwarzer Balken) der Personen oder Informationen sowie die Überarbeitung von Tonaufnahmen, um Stimmen zu verfremden mittels verschiedener Applikationen (bspw. Freeware wie VideoPad Video Editor<sup>56</sup> oder Avidemux<sup>57</sup> bzw. kostenpflichtige Editierprogramme wie z. B. Vegas<sup>58</sup>).

#### 4. Urheberrecht und verwandte Schutzrechte

Das Urheberrecht dient dem Schutz der schöpferischen Leistung von Kreativschaffenden und der Kreativwirtschaft. Es erteilt den Urheber\*innen das alleinige Recht an der Veröffentlichung, Verarbeitung, Vervielfältigung, Aufführung und dem Verleih ihres Werkes. Damit ein Werk als "urheberrechtlich geschützt" gilt, muss es im Zuge einer persönlichen, geistig-schöpferischen Leistung entstanden sein. Nach §2 Abs. 1 UrhG gehören dazu "Werke der Literatur, Wissenschaft und Kunst".

<sup>54</sup> De-Identification Toolbox. Zugriff am 27.04.2023, <https://www.nitrc.org/projects/de-identification/>.

<sup>55</sup> GDCM zur Entfernung von Metadaten. Zugriff am 27.04.2023, [http://gdcm.sourceforge.net/wiki/index.php/Main\\_Page](http://gdcm.sourceforge.net/wiki/index.php/Main_Page).

<sup>56</sup> VideoPad Video Editor. Zugriff am 27.04.2023, [https://www.nchsoftware.com/videopad/de/index.html?ns=true&kw=videopad%20video%20editor&gclid=CjwKCAjw6raYBhB7EiwABge5KI\\_FtxfW39dUG-AzhdxtysgkPTTgjIsveEHIGNvqRADqtX5PI2k7TxoC7h4QAvD\\_BwE](https://www.nchsoftware.com/videopad/de/index.html?ns=true&kw=videopad%20video%20editor&gclid=CjwKCAjw6raYBhB7EiwABge5KI_FtxfW39dUG-AzhdxtysgkPTTgjIsveEHIGNvqRADqtX5PI2k7TxoC7h4QAvD_BwE).

<sup>57</sup> Avidemux. Zugriff am 27.04.2023, <http://avidemux.sourceforge.net/>.

<sup>58</sup> Vegas Software. Zugriff am 27.04.2023, <https://www.vegascreativesoftware.com/de/>.

Ideen, Fakten oder Informationen werden demnach nicht vom Urheberrecht geschützt, da sie lediglich die Basis eines Werkes bilden, jedoch nicht das Werk selbst sind<sup>59</sup>.

Somit sind maschinell erzeugte und unbearbeitete Rohdaten urheberrechtlich nicht schutzfähig. Dahingehend können jedoch die für eine Veröffentlichung angefertigten Texte, Zeichnungen, individuell gestalteten Grafiken, Fotografien, Videos etc. urheberrechtlich geschützt sein – jedoch nur, wenn sie den obengenannten Kriterien entsprechen. Dies muss im Einzelfall geprüft werden<sup>60</sup>.

Die sogenannten Leistungsschutzrechte - auch "verwandte Schutzrechte" genannt - sind Schutzrechte für z. B. Fotograf\*innen, Sänger\*innen, Interpreter\*innen oder Vermittler\*innen von Inhalten. Der Leistungsschutz beinhaltet unter anderem aber auch den Schutz des Datenbankerstellers, Schutz der Lichtbilder oder den Schutz des Herstellers von Tonträgern (weitere Schutzrechte können im UrhG Teil 2 "Verwandte Schutzrechte" nachgelesen werden). Die Leistungsschutzrechte stehen demjenigen zu, der die jeweils geschützte Leistung erbracht hat.

Daneben können aber auch gewerbliche Schutzrechte zu berücksichtigen sein, bspw. eine vertragliche Vereinbarung (durch z. B. Arbeitsverträge), die eine Übertragung des Nutzungsrechts an den Daten auf die Institution vorsieht. Des Weiteren können vertragliche Regelungen zu Nutzungsrechten auch in Förderrichtlinien von Geldgebern oder Richtlinien der beteiligten Gruppen, Abteilungen oder Institutionen zu finden sein. Auch Open-Content-Lizenzen (siehe Einheit 15 "Publikation", Punkt 7, S. 54) sind Verträge über die Einräumung von Nutzungsrechten. Die Verträge müssen dabei nicht schriftlich, sondern können auch implizit oder mündlich geschlossen werden<sup>61</sup>. Obwohl die Nutzungsrechte übertragen werden, bleiben jedoch die Urheberschaft und damit die Verwertungsrechte bei den Ersteller\*innen der Daten.

Wenn anzunehmen ist, dass die Forschung oder ihre Ergebnisse wirtschaftlich nutzbar sind, ist es empfehlenswert, sich an die Wissenstransferstellen der eigenen Einrichtung zu wenden. Hier kann geklärt werden, wie die Lizenzierung und Nachnutzung der Daten zu regeln ist. Auch über das Patentrecht kann man sich dort ausführlich informieren.

## 5. Vertragsrecht

Weitere Rechte und Pflichten für den Umgang mit Forschungsdaten können sich auch aus sonstigen vertraglichen Vereinbarungen ergeben. Dies können sowohl Verträge mit Fördermittelgebern als auch mit wissenschaftlichen oder wirtschaftlichen Kooperationspartner sein.

Sollen Daten außerhalb von Kooperationsverträgen zur gemeinsamen Bearbeitung an Kolleg\*innen weitergegeben werden, bevor sie publiziert wurden, wird dies durch Datenbereitstellungsverträge (data sharing agreements) geregelt. Hier wird festgelegt, unter welchen Bedingungen Daten durch Dritte verarbeitet werden dürfen. Dabei verfolgt ein solches Agreement zwei Ziele:

- den Schutz der\*des Datengebenden vor Datenmissbrauch
- die Verhinderung von Missverständnissen auf beiden Vertragsseiten durch Sicherstellung, dass alle Fragen der Datennutzung geklärt sind.

---

<sup>59</sup> Kreutzer, Till und Henning Lahmann. *Rechtsfragen bei Open Science: Ein Leitfaden*. Hamburg: Hamburg University Press, 2019, <https://doi.org/10.15460/HUP.195>, S. 25.

<sup>60</sup> Leibniz Universität Hannover, & Technische Informationsbibliothek. (2018). FAQs zu rechtlichen Aspekten im Umgang mit Forschungsdaten (Version 180215). Zenodo. <https://doi.org/10.5281/zenodo.1173546>.

<sup>61</sup> Ebd., S. 27.



## Datenmanagementplan (Einheit 6)

### 1. Motivation

Ein Datenmanagementplan bindet zwar Ressourcen bei der Erstellung, bietet aber gleichzeitig viele Vorteile. Er

- schafft eine verbindliche Grundlage für einen einheitlichen Umgang mit Daten im Forschungsprozess
- erleichtert das Verständnis der eigenen Daten
- erleichtert die Abstimmung zwischen Projektpartnern
- hilft mögliche Probleme frühzeitig zu erkennen und dafür Lösungen zu skizzieren
- legt Verantwortlichkeiten fest
- regelt Zugriffsrechte
- hilft Datenduplikate, Datenverlust und Sicherheitslücken zu vermeiden
- wird i.d.R. von Forschungsförderern verlangt.

Eine hilfreiche Vorüberlegung für einen Datenmanagementplan ist es, den Prozess ausgehend von der Datenarchivierung her zu beschreiben. Aus diesen Überlegungen ergibt sich die Notwendigkeit, bereits sehr früh die bspw. zu verwendenden Formate, Standards, Metadaten, Lizenzen etc. festzulegen.

### 2. Was ist ein Datenmanagementplan?

Ein Datenmanagementplan (DMP) ist ein Dokument, in dem der beabsichtigte Umgang mit Forschungsdaten beschrieben wird; er schließt Maßnahmen sowohl während des Forschungsprozesses als auch nach dessen Abschluss mit ein. Das Dokument enthält dabei alle Informationen, die die Sammlung, Aufbereitung, Speicherung, Archivierung und Veröffentlichung von Forschungsdaten beschreiben und dokumentieren, und kann – je nach Komplexität des Projektes – zwischen wenigen Absätzen und mehreren Seiten variieren. Die Diversität der Forschungsdaten sowie des Umgangs mit ihnen bestimmt die Länge eines Datenmanagementplans. Er sollte – wenn möglich – kurz, konkret und mit allen Projektbeteiligten abgestimmt sein.

### 3. Bestandteile eines Datenmanagementplans

Für die Erstellung eines Datenmanagementplans ist zentral, die Empfehlungen und Vorgaben Dritter, z. B. Fördermittelgeber oder Arbeitgeber, zu berücksichtigen. Ein DMP adressiert i.d.R. folgende Punkte:

- Projekttitle, Laufzeit und Forschungsfrage(n)
- Verantwortliche\*r für das Datenmanagement
- Nachnutzung existierender Daten
- zu erhebende Daten:
  - Beschreibung der zu erfassenden Daten, Datentypen und -formate
  - erwarteter Speicherbedarf
  - Methoden der Datenerhebung, verwendete Hard- und Software
- Datenorganisation:
  - Datenspeicherung
  - Backup
  - Ordnerstruktur

- Dateinamenskonventionen
  - Dokumentation und Metadaten
- Rechtliche Aspekte, z. B.:
  - Datenschutz
  - Urheberrecht
- Datenaustausch und -zugang:
  - im Projekt
  - mit externen Partnern und Dienstleistern
- Langzeitsicherung und Archivierung
- Datenpublikation
- Kosten des Datenmanagements
- Qualitätssicherung
- Zugriff und Nachnutzung

Idealerweise entwickelt sich ein Datenmanagementplan dynamisch: d. h. er wird im Verlauf des Projekts fortlaufend aktualisiert und erweitert und entwickelt sich somit von einer Skizze zu einer detaillierten Dokumentation des Datenmanagementprozesses (Active Data Management Plan<sup>62</sup> und trägt somit zur Nachnutzbarkeit der Daten bei (Reusability). Ein zu Anfang noch unvollständiger DMP ist daher durchaus nicht ungewöhnlich.

#### 4. Anforderungen der Forschungsförderer

In Deutschland werden Datenmanagementpläne bereits von einer Reihe von Forschungsförderern bei der Antragstellung gefordert. Forschungsförderer wie die Europäische Kommission (European Commission, EC), die Deutsche Forschungsgemeinschaft (DFG), die Volkswagen Stiftung (VW Stiftung) und das Bildungsministerium für Bildung und Forschung (BMBF) erwarten zunehmend die Bereitstellung eines Datenmanagementplans zu Förderbeginn (EC<sup>63</sup>), Angaben zum Umgang mit zu erhebenden Forschungsdaten (DFG)<sup>64</sup> sowie – je nach Förderrichtlinie – einen Verwertungsplan für Projektergebnisse oder einen detaillierten Datenmanagementplan bei Antragstellung (BMBF<sup>65</sup>, VW Stiftung<sup>66</sup>) (siehe Tabelle 3).

---

<sup>62</sup> RDA Interest Group Active Data Management Plans. Zugriff am 27.04.2023, <https://www.rda-alliance.org/node/9136/outputs>.

<sup>63</sup> European Commission, H2020 Programme. „Guidelines on FAIR Data Management in Horizon 2020.“ Version 3.0. Zugriff am 27.04.2023,

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf).

<sup>64</sup> Deutsche Forschungsgemeinschaft. „Leitlinien zum Umgang mit Forschungsdaten.“ Zugriff am 27.04.2023, [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf).

<sup>65</sup> Verbund Forschungsdaten Bildung. „Checkliste zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung.“ Version 1.1. *fdbinfo* Nr. 2 (2015). Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/files/fdbinfo\\_2.pdf](https://www.forschungsdaten-bildung.de/files/fdbinfo_2.pdf).

<sup>66</sup> VolkswagenStiftung. „OPEN SCIENCE. OPEN ACCESS – OPEN DATA – OPEN SOURCE“ Zugriff am 27.04.2023,

[https://www.volkswagenstiftung.de/sites/default/files/documents/Open\\_Science\\_Policy\\_und\\_Umsetzung\\_VolkswagenStiftung.pdf](https://www.volkswagenstiftung.de/sites/default/files/documents/Open_Science_Policy_und_Umsetzung_VolkswagenStiftung.pdf).

Tabelle 3: Vergleich der Anforderungen der Förderer bezüglich des Datenmanagementplans (Stand: 26.09.2022).

Förderer	Plan gefordert?	Abgabe bei Antrag?	Inhalt	Updates
EC Horizon Europe	Ja	Nein, erster Plan innerhalb der ersten 6 Projektmonate	Inhalte des Horizon Europe Template	Updates bei signifikanten Änderungen und zum Projektende
DFG	Angaben zum Umgang mit Forschungsdaten	Ja	Inhalte der Leitlinie zum Umgang mit Forschungsdaten	Nein
BMBF	abhängig vom Programm	Falls notwendig, ja	Inhalt abhängig vom Programm; Bildungsforschung-Checkliste	abhängig vom Programm
Volkswagenstiftung	Ja	Ja	Inhalte des Science Europe Templates	Nein

## 5. Muster und Werkzeuge für die Psychologie

Zahlreiche Instrumente können Forschende bei der Erstellung eines Datenmanagementplans unterstützen. Eine Auswahl ist in Tabelle 4 aufgelistet.

Tabelle 4: Muster und Werkzeuge zur Erstellung eines Datenmanagementplans

Checklisten	Mustervorlagen
Checkliste Verbund Forschungsdaten Bildung <sup>67</sup>	Musterplan EU Horizon 2020 (Version 3.0) <sup>68</sup>
Checkliste OSF.io <sup>69</sup>	Musterplan DFG <sup>70</sup>
Checkliste Digital Curation Centre <sup>71</sup>	Musterplan BMBF <sup>72</sup>

<sup>67</sup> Verbund Forschungsdaten Bildung (2015). „Checkliste zur Erstellung eines Datenmanagementplans in der empirischen Bildungsforschung“. Version 1.1. fdbinform Nr. 2. Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdbinform\\_2.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinform_2.pdf).

<sup>68</sup> DMP-Musterplan EU Horizon 2020 (Version 3.0). Zugriff am 27.04.2023, <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-h2020-v3>.

<sup>69</sup> Checkliste DMP auf OSF.io, Zugriff am 27.04.2023, <https://help.osf.io/article/144-creating-a-data-management-plan-dmp-document>.

<sup>70</sup> Muster-Datenmanagementplan DFG. Zugriff am 27.04.2023, <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-dfg>.

<sup>71</sup> DCC. (2013). „Checklist for a Data Management Plan. v.4.0“. Edinburgh: Digital Curation Centre. Zugriff am 27.04.2023, <http://www.dcc.ac.uk/resources/data-management-plans>.

<sup>72</sup> Muster-Datenmanagementplan BMBF. Zugriff am 27.04.2023, <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-bmbf>.

Checkliste Guidance University of Bath <sup>73</sup>	Musterplan ICPSR <sup>74</sup>
Leitfaden, entwickelt im WissGrid-Projekt <sup>75</sup>	Musterplan VW-Stiftung <sup>76</sup>
	Mustersammlung Digital Curation Centre <sup>77</sup>

DMP-Tools, allgemein		
	Ziel	Zusätzl. Informationen
DMPTool <sup>78</sup>	Erstellung DMP, insb. für amerikanische Forschungsförderer	
RDMO <sup>79</sup>	Erstellung DMP, insb. insbesondere für deutsche Forschungsförderer	bietet fachspezif. Templates; auf forschungsdaten.info ohne Installation via Shibboleth nutzbar <sup>80</sup>
DMPonline <sup>81</sup>	Erstellung DMP, insb. für britische Forschungs-förderer; DMP für EC-Förderung	DMPTool/DMPonline basieren auf dem gleichen Softwarecode (DMPRoadmap)
Data Stewardship Wizard <sup>82</sup>	Erstellung DMP	Zielgruppe: Data Stewards, die Forscher*innen bei der DMP-Erstellung unterstützen
OpenDMP/Argos <sup>83</sup>	Erstellung DMP	Tool unterstützt das RDA maDMP Template <sup>84</sup>

<sup>73</sup> Cope, J (2013). „Data Management Planning Guidance for Postgraduate Researchers“. University of Bath, Bath. Zugriff am 27.04.2023, <https://researchportal.bath.ac.uk/en/publications/data-management-planning-guidance-for-postgraduate-researchers>.

<sup>74</sup> ICPSR. „Framework for Creating a Data Management Plan“. Zugriff am 27.04.2023, <https://www.icpsr.umich.edu/web/pages/datamanagement/dmp/framework.html>.

<sup>75</sup> Ludwig, Jens; Harry Enke (Hrsg.) (2013). „Leitfaden zum Forschungsdaten-Management“. vwh-Verlag, Universitätsverlag Göttingen. ISBN: 978-3-86488-032-2. S.83ff. . am 27.04.2023, [https://www.forschungsdaten.org/images/b/b0/Leitfaden\\_Data-Management-WissGrid.pdf](https://www.forschungsdaten.org/images/b/b0/Leitfaden_Data-Management-WissGrid.pdf).

<sup>76</sup> Muster-Datenmanagementplan VW-Stiftung. Zugriff am 27.04.2023, <https://www.cms.hu-berlin.de/de/dl/dataman/muster-dmp-vwstiftung-pdf>.

<sup>77</sup> DMP-Mustersammlung des DCC. Zugriff am 27.04.2023, <https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>.

<sup>78</sup> DMPTool. „Welcome to the DMPTool.“ Zugriff am 27.04.2023, <https://dmptool.org/>.

<sup>79</sup> RDMO. „RDMO. Research Data Management Organiser.“ Zugriff am 27.04.2023, <http://rdmorganiser.github.io>.

<sup>80</sup> Research Data Management Organiser. Zugriff am 27.04.2023, <https://rdmo.forschungsdaten.info/>.

<sup>81</sup> DMPOnline. „Plan to make data work for you.“ Zugriff am 27.04.2023, <https://dmponline.dcc.ac.uk/>.

<sup>82</sup> Data Stewardship Wizard. „Data Stewardship Wizard.“ Zugriff am 27.04.2023, <https://ds-wizard.org/>.

<sup>83</sup> ARGOS. „Plan and follow your data.“ Zugriff am 27.04.2023, <https://argos.openaire.eu>.

<sup>84</sup> Allgemeiner Standard für maschinenlesbare Datenmanagementpläne der Research Data Alliance. Zugriff am 27.04.2023, <https://www.rd-alliance.org/group/dmp-common-standards-wg/outcomes/rda-dmp-common-standard-machine-actionable-data-management>.

<b>DMP-Tools, Psychologie-spezifisch</b>		
	<b>Ziel</b>	<b>Zusätzliche Informationen</b>
DataWiz <sup>85</sup>	Erstellung DMP, insb. für Psycholog*innen	
BION <sup>86</sup>	Erstellung DMP, insb. für Neurowissenschaftler*innen der Universität Gießen	
Stamp <sup>87</sup>	Erstellung DMP	Entwickelt für die Bildungsforschung, ebenso nutzbar für Psycholog*innen

---

<sup>85</sup> Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). „DataWiz.“ Zugriff am 27.04.2023, <https://datawiz.leibniz-psychology.org/DataWiz/>.

<sup>86</sup> BION. Zugriff am 27.04.2023, <https://www.uni-giessen.de/fbz/fb06/psychologie/weitere-inst/bion/bioncu-datenmanagement>.

<sup>87</sup> Standardisierter Datenmanagementplan für die Bildungsforschung (Stamp), Zugriff am 27.04.2023, <https://www.forschungsdaten-bildung.de/stamp-nutzen>.

## Systematik im FDM (Einheit 7)

### 1. Motivation

Strukturiertes Vorgehen ist erforderlich,

- damit auch nach Jahren die Art und Weise der Datenerhebung, -analyse und -veröffentlichung nachvollziehbar bleibt
- damit anderen Forschenden, aber auch ihnen selbst die Benennungskonventionen bekannt sind und die Zusammenarbeit vereinfacht wird
- damit auch andere Forschende mit den Daten arbeiten können
- um nach Daten einfacher suchen und sie schneller finden zu können
- um doppelte Arbeit zu vermeiden
- um Datenverlust durch Überschreibung oder versehentliches Löschen vorzubeugen
- um den aktuellen Stand ohne Aufwand zu identifizieren
- um Maschinenlesbarkeit zu gewährleisten

und führt insgesamt zu einem effizienteren Arbeiten, bietet Transparenz über das eigene Forschungsprojekt und führt durch das Folgen der FAIR-Prinzipien zu mehr Nachhaltigkeit durch die Nachnutzung von Daten.

### 2. Verzeichnisstruktur

Eine Verzeichnisstruktur (auch Verzeichnisbaum genannt) bezeichnet die hierarchische Anordnung der Ordner, in welchem die zu einem Projekt gehörigen Daten abgelegt werden (siehe Abbildung 5). Die Verzeichnisstruktur sollte klar ersichtlich und damit auch für andere Forschende verständlich sein, um ein leichteres Auffinden der Daten zu gewährleisten. Idealerweise folgen Verzeichnisstrukturen dem Workflow in dem jeweiligen Vorhaben und unterstützen so das meist schrittweise Erstellen, Analysieren und Publizieren der Daten. Für eine gute Übersicht sollten Verzeichnisstrukturen auf Servern identisch zu denen auf den lokalen Rechnern sein.

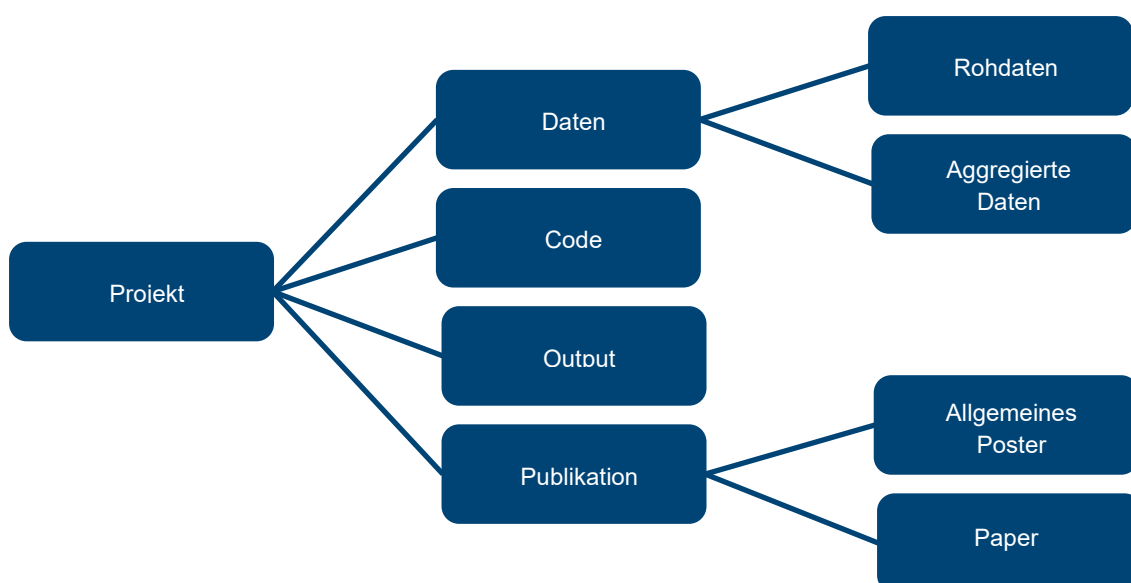


Abbildung 5: Beispiel für eine Verzeichnisstruktur mit Unterordnern.

### 3. Benennung

Der Dateiname sollte objektiv und intuitiv sowie personenunabhängig nachvollziehbar sein. Die Benennung kann dabei nach den folgenden drei Kriterien erfolgen:

- **System** – wichtig für den späteren Zugriff und Abruf der Daten ist die Berücksichtigung des Systems, unter dem die Datei gespeichert wird.
- **Kontext** – der Dateiname beinhaltet inhaltsspezifische oder deskriptive Informationen, damit unabhängig vom Speicherort klar bleibt zu welchem Kontext die Datei gehört, z. B. „Zeitplan.pdf“ oder „ZeitplanFDNext.pdf“.
- **Konsistenz** – die Namenskonvention sollte vorab gewählt werden, um sicherzustellen, dass sie systematisch befolgt werden kann und die gleichen Informationen (wie z. B. Datum und Zeit) in der selber Reihenfolge beinhaltet (z. B. JJJJMMTT).

Dateinamen sollten so lang wie nötig, aber auch so kurz wie möglich sein, um übersichtlich zu bleiben und innerhalb jedes Betriebssystems lesbar zu sein. Namensbestandteile, die bereits in den Ordnernamen abgelegt sind, müssen nicht nochmal in den Dateinamen wiederholt werden. Um eine einheitliche Namensgebung zu gewährleisten, sind folgende Namensbestandteile hilfreich:

- Inhalt
- Ersteller\*in
- Erstellungsdatum
- Bearbeitungsdatum
- Bezeichnung der Arbeitsgruppe
- Publikationsdatum
- Projektnummer
- Versionsnummer

Leerzeichen, Punkte und Sonderzeichen sollten vermieden werden, da sie unter verschiedenen Systemen unterschiedlich interpretiert werden und dies zu Fehlern führen kann. Insbesondere Leerzeichen sollten durch Unterstriche ersetzt werden. Oder man schreibt jeweils den ersten Buchstaben von Wörtern groß.

Um eine chronologische Sortierung zu ermöglichen, empfiehlt es sich, den Namen mit Datumsangabe zu beginnen, zum Beispiel JJJJMMTT\_Name oder JJJJMMTTName etc. Beispiele für einheitliche Namensgebung:

- 20200923\_VP12\_Session1\_VALlist1.txt
- 20190520\_VP09\_Session1\_EEGpre\_EO.edf
- 20210131\_VP23\_Interview1.txt

Automatisch generierte Namen (bspw. von der Digitalkamera) sollten der besseren Verständlichkeit wegen entsprechend den obengenannten Konventionen umgeändert werden. Des Weiteren ist es sowohl bei größeren Projekten, als auch bei kleinen Forschungsvorhaben lohnenswert, die gewählten Namenskonventionen schriftlich festzuhalten. Gewählte Abkürzungen sollten dabei insbesondere in einem Datenmanagementplan (siehe Einheit 6 "Datenmanagementplan", S. 25) oder einer README-Datei erläutert werden. Ohne eine solche Datei ist die Rekonstruktion dieser Konventionen nach Jahren oft nur noch schwer möglich.

## 4. Umbenennen mehrerer Dateien gleichzeitig

Das gleichzeitige Umbenennen mehrerer Dateien kann nützlich sein, u. a.:

- um die automatisch generierten Namen von der Digitalkamera oder anderer Software in einem Arbeitsschritt zu ändern;
- um Leerzeichen oder andere Sonderzeichen aus mehreren Dateinamen in einem Arbeitsschritt zu entfernen bzw. sie zu ersetzen.

Entsprechende Software existiert für die meisten Betriebssysteme, bswp.:

### Windows:

- Ant Renamer<sup>88</sup>
- Rename-IT<sup>89</sup>
- Bulk Rename Utility<sup>90</sup>

### Mac:

- Renamer 6<sup>91</sup> (for Mac)
- Name Changer<sup>92</sup>

### Linux:

- GNOME Commander<sup>93</sup>
- GPRename<sup>94</sup>

### Unix:

- Unter Unix kann das Kommando „rename“ oder „mmv“ hilfreich sein, um mit regulären Ausdrücken Dateien aufzufinden und umzubenennen.

## 5. Versionskontrolle

Versionen und deren Historie sind nützlich, um einen Überblick über die durchgeführten Schritte zu behalten und diese nachvollziehbar zu machen, und wenn nötig, auch wieder einen Arbeitsschritt zurückzugehen. Bei für die Öffentlichkeit zugänglich gemachten Versionen können diese darüber hinaus auch die Fehlerbehebung unterstützen.

Die verbreitetste Form, Versionen zu kennzeichnen, besteht in der Vergabe von ganzen Zahlen für größere Versionsänderungen und mit einem Unterstrich verbundenen Zahlen für kleinere Veränderungen (z. B. v1, v2, v1\_01, v2\_03 etc.). Es wird davon abgeraten, Bezeichnungen wie final, final2, revision, definitiv\_final zu benutzen.

Bei größeren Projekten (z. B. Verbundprojekte, SFBs) ist die Nutzung einer Versionskontroll-Software (z. B. Git oder Subversion) bei der Verwaltung von Versionen sehr hilfreich. Bei kollaborativen Dokumenten und Speicherorten wie Wikis, Google Docs oder Clouds steht ebenfalls die Funktion der Versionierung und Änderungsverfolgung zu Verfügung.

---

<sup>88</sup> Ant Renamer. Zugriff am 27.04.2023, <http://www.antp.be/software/renamer>.

<sup>89</sup> Rename-IT. Zugriff am 27.04.2023, <http://sourceforge.net/projects/renameit>.

<sup>90</sup> Bulk Rename Utility. Zugriff am 27.04.2023, <http://www.bulkrenameutility.co.uk/>.

<sup>91</sup> Renamer 6. Zugriff am 27.04.2023, <http://renamer.com/>.

<sup>92</sup> Name Change. Zugriff am 27.04.2023, <http://mrrsoftware.com/namechanger/>.

<sup>93</sup> GNOME Commande. Zugriff am 27.04.2023, <https://gcmd.github.io/>.

<sup>94</sup> GPRename. Zugriff am 27.04.2023, <http://gprenome.sourceforge.net/>.



Beispiele für die Dateibeschriftung mit Versionierung:

- Doe\_interview\_July2010\_V1
- Lipid\_analysis\_rate\_V2
- 2017\_01\_28\_MR\_CS3\_V6\_03

## Dokumentation und Metadaten (Einheit 8)

### 1. Motivation

Damit Forschungsdaten besser auffindbar und nachvollziehbar sind, ist eine Dokumentation der Daten als Basis guter wissenschaftlicher Praxis<sup>95</sup> unverzichtbar. Hilfreich ist hier bspw. eine Checkliste zum Umgang mit Forschungsdaten der DFG<sup>96</sup> oder rudimentärer die Beantwortung der 5W1H-Fragen (siehe Dokument „FDNext\_TiTKonzept\_MethodenMaterialien\_V1“, S. 11). Eine Datendokumentation erleichtert zudem erheblich die Nachnutzung der Daten durch eigene oder externe Anschlussforschung. Sie ermöglicht auch die Reproduzierbarkeit von Befunden, die vor allem vor dem Hintergrund der Replikationskrise in der Psychologie<sup>97</sup> die Verifizierung von bisherigen Forschungsergebnissen nötig macht.

### 2. Datendokumentation

Zu einer Datendokumentation gehört unter anderen die gezielte Benennung der Dateien. Darüber hinaus umfasst sie die Ergänzung der Daten, um sogenannte Metadaten (Informationen, welche die Studie und den Erhebungsprozess beschreiben, siehe Einheit 18 "Dokumentation und Metadaten", Punkt 4, S. 36), sowie weitere Dokumente<sup>98</sup>, sofern sie zum Verständnis und Interpretation der Daten beitragen (z. B. Protokolle, eMails). Da mit der Zeit Details in Bezug auf Erhebung, Probleme bei den Messungen oder mit Programmen/Gerätschaften usw. in Vergessenheit geraten können, ist es empfehlenswert, die Dokumentation von Daten arbeitsbegleitend zu betreiben. Eine Datendokumentation dient somit der vereinfachten Handhabung und (Nach-)Nutzung der Daten, da sie Entscheidungen bei der Erhebung und Analyse rekonstruierbar und nachvollziehbar macht.

Dokumentationsformen sind bspw. ReadMe-Dateien, Tagging der Dateien, Versionierungen, Data Dictionary, Codebuch, Feld- oder Laborbuch, Artikel in einem Data Journal oder verwendete Tools.

Zu den grundlegenden Inhalten einer Dokumentation zählen:

- eine Beschreibung des Forschungsvorhabens
- die Projektziele
- Hypothesen
- Informationen zur Erhebung der Daten (Methoden, Einheiten, Zeiträume, Orte, verwendete Instrumente und Software)
- Maßnahmen zur Datenbereinigung
- Angaben zur Struktur der Daten und deren Beziehungen zueinander
- die Erläuterung von Variablen, Labels und Codes
- Angaben zur Versionshistorie

---

<sup>95</sup> Leitlinien der DFG zur Sicherung guter wissenschaftlicher Praxis. Zugriff am 27.04.2023, [https://www.dfg.de/download/pdf/foerderung/rechtliche\\_rahmenbedingungen/gute\\_wissenschaftliche\\_praxis/kodex\\_gwp.pdf](https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf).

<sup>96</sup> DFG (2021). „Checkliste zum Umgang mit Forschungsdaten“. Zugriff am 27.04.2023, [https://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/forschungsdaten/forschungsdaten\\_checkliste\\_de.pdf](https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf).

<sup>97</sup> Positionspapier der Psychologie-Fachschaften-Konferenz zur Replikationskrise in der Psychologie. Zugriff am 27.04.2023, <https://psyfako.org/wp-content/uploads/30-PsyFaKo-STN-Open-Science-Kulturwandel-zu-einer-offenen-Wissenschaft.pdf>.

<sup>98</sup> Definition Datendokumentation. Zugriff am 27.04.2023, <https://www.forschungsdaten-bildung.de/daten-dokumentieren>.

- Informationen zum Zugang und Nutzungsbedingungen

### 3. Codebooks für die Psychologie

Spätestens mit der 2016 von der DGP herausgegebenen Empfehlung sind Codebooks auch in der Psychologie ein essentieller Bestandteil der Projekt- und Datendokumentation, um die Datenanalyse zu erleichtern und eine spätere Nachvollziehbarkeit der Daten zu ermöglichen. Zahlreiche Anleitungen und Muster für fach- und anwendungsspezifische Codebooks, aber auch Tutorien sind heute u. a. auf diversen Repositorien, wie z. B. Zenodo oder OSF, zu finden.

#### Anleitung:

- Shribe-Blog<sup>99</sup>
- Cookbook for Codebook<sup>100</sup>

#### Fachspezifische Muster:

- Ludwig-Maximilian-Universität München, Codebook<sup>101</sup> der PsyCourse-Studie<sup>102</sup> (klinische Längsschnitt- und Beobachtungsstudie zur Klärung der Schizophrenie/Bipolare Störung-Dichotomie)
- Codebook auf SPSS-Basis<sup>103</sup> der ZPID-Nutzerstudie<sup>104</sup> (Bedarfserhebung zu Open-Science-Infrastrukturbedarf innerhalb der Psychologie und inwieweit dieser Bedarf durch Angebote des Leibniz-Instituts für Psychologie abgedeckt ist)
- Codebook<sup>105</sup> der PsyCorona-Studie<sup>106</sup> (Ziel: Identifikation psychologischer und kultureller Faktoren, die die Verbreitung des SARS-CoV-2 beeinflussen)

#### Weitere Muster:

- Codeplan zur Analyse von Veranstaltungsankündigungen in der Weiterbildung, Deutsches Institut für Erwachsenenbildung<sup>107</sup>

Die DFG empfiehlt darüber hinaus, in den Studien genutzte Software (Art und Versionen) sowie physische Geräte für die Datenerhebung ebenfalls in die Dokumentation aufzunehmen<sup>108</sup>. Dies hilft,

<sup>99</sup> Shribe. Zugriff am 27.04.2023, <https://shribe.de/codebuch-erstellen/>.

<sup>100</sup> Cookbook for Codebook. Zugriff am 27.04.2023, <https://osf.io/e4tqy>.

<sup>101</sup> Codebook der PsyCourse-Studie an der LMU München. Zugriff am 27.04.2023, <https://data.ub.uni-muenchen.de/251/>.

<sup>102</sup> Budde, M., Anderson-Schmidt, H., Gade, K., Reich-Erkelenz, D., Adorjan, K., Kalman, J. L., ... & Heilbronner, U. (2019). A longitudinal approach to biological psychiatric research: The PsyCourse study. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 180(2), 89-102. <https://doi.org/10.1002/ajmg.b.32639>.

<sup>103</sup> Codebook auf SPSS-Basis der ZPID-Nutzerstudie, Leibniz-Institut für Psychologie. Zugriff am 27.04.2023, <https://www.psycharchives.org/en/item/9898b1cf-25f0-4119-938b-a349aaa31b9e>.

<sup>104</sup> Ergebnisse der ZPID-Studie. Zugriff am 27.04.2023, <https://www.psycharchives.org/en/item/89a4e823-0d35-4d55-ac19-b7fda04f3935>.

<sup>105</sup> Codebook der PsyCorona-Studie. Zugriff am 27.04.2023, <https://osf.io/7kfj5>.

<sup>106</sup> PsyCorona-Studie der University Groningen. Zugriff am 27.04.2023, <https://www.rug.nl/sustainable-society/research/previous-themes/psycorona/about-the-psycorona-study?lang=en>.

<sup>107</sup> DIE Bonn. Zugriff am 27.04.2023, <https://www.die-bonn.de/docs/codeplan-analyse-veranstaltungsankuendigungen-schrader.pdf>.

<sup>108</sup> Empfehlung der DFG zur Aufnahme von Software und physischen Geräten zur Datenerhebung. Zugriff am 27.04.2023, [https://www.dfg.de/download/pdf/foerderung/grundlagen\\_dfg\\_foerderung/forschungsdaten/empfehlungen\\_forschungsdaten\\_psychologie\\_en.pdf](https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/empfehlungen_forschungsdaten_psychologie_en.pdf).

Unterschiede in der Datenerhebung und/oder Auswertung durch technisch bedingte Parameter (bspw. physische Unterschiede in der Bauweise, Unterschiede in der Programmierung) im Vergleich mit anderen Studien besser nachvollziehen zu können.

#### 4. Was sind Metadaten?

Metadaten sind strukturierte Daten, die Informationen über andere Daten beinhalten – gleichsam „Daten über Daten“. Sie werden entweder unabhängig oder gemeinsam mit den Daten, die sie beschreiben, gespeichert. Sie können ebenfalls als Primärdaten zur Beantwortung von Forschungsfragen verwendet werden. Hierbei unterscheidet man zwischen **administrativen** (geben Auskunft darüber, welche Art von Anweisungen, Richtlinien und Einschränkungen für eine Datei gelten – z. B. über Zugriffsklassen), **strukturellen** (wie verwandte Objekte miteinander verknüpft sind, bswp. die richtige Reihenfolge der Datenpunkte in einer Interventionsstudie, oder der Interviews in einer Längsschnittstudie) und **deskriptiven** Metadaten (Elemente, die Daten auffindbar und identifizierbar machen, bswp. Titel, Abstract, Author, Keywords, usw.). Sie bilden eine spezifische Untermenge der Dokumentationsangaben und dienen u. a. der Auffindbarkeit der Daten durch Suchmaschinen und in bibliothekarischen Nachweissystemen. Damit sie auch maschinenlesbar sind, zum Beispiel im Semantic Web<sup>109</sup>, werden sie häufig im XML-Format<sup>110</sup> gespeichert.

#### 5. Metadatenstandards

Um die Interoperabilität, also die Verknüpfung und gemeinsame Bearbeitung zwischen verschiedenen Metadaten zu gewährleisten, stehen sogenannte standardisierte Metadatenvokabulare zur Verfügung. Diese Standards ermöglichen eine inhaltlich und strukturell gleichförmige Beschreibung von ähnlichen Datensätzen. Zu diesem Zweck beinhalten sie eine festgelegte Auswahl an Informationen, die notwendig ist, um diese Daten auffinden und identifizieren zu können. Zu den bekanntesten bibliographischen fachübergreifenden Metadatenstandards zählen Dublin Core<sup>111</sup>, DataCite Metadata Schema<sup>112</sup> und MARC21<sup>113</sup>.

#### 6. Psychologie-spezifische Metadatenstandards

Auch wenn sich viele Fachbereiche spezifische Anforderungen zur Beschreibung und Erschließung ihrer Daten mit Metadaten erarbeitet haben, hat sich für die Psychologie noch kein entsprechender Standard entwickelt. Als Alternative etablierte sich jedoch ein für die Sozial- und Wirtschaftswissenschaften existierender Standard (Data Documentation Initiative<sup>114</sup>, DDI). Jedoch gibt es zumindest für einige Teilbereiche innerhalb der Psychologie Bestrebungen, dies zu ändern. So gibt es mittlerweile für bildgebende Formate den sogenannten BIDS-Standard<sup>115</sup>, sowie den für das Projekt

---

<sup>109</sup> Das Semantic Web ist eine Erweiterung des Webs um eine maschinenlesbare Schicht und vereinfacht den Datenaustausch zwischen Rechnern.

<sup>110</sup> XML („Extensible Markup Language“) ist eine Auszeichnungssprache zur Darstellung hierarchisch strukturierter Daten im Format einer Textdatei.

<sup>111</sup> DublinCore. Zugriff am 27.04.2023, <https://www.dublincore.org/>

<sup>112</sup> DataCite Metadata Schema. Zugriff am 27.04.2023, <https://schema.datacite.org/>

<sup>113</sup> MARC21. Zugriff am 27.04.2023,

[https://www.dnb.de/DE/Professionell/Metadatendienste/Exportformate/MARC21/marc21\\_node.html;jsessionid=838CCCAE6368D653EE43AE5BA237C815.intranet231#doc158972bodyText1](https://www.dnb.de/DE/Professionell/Metadatendienste/Exportformate/MARC21/marc21_node.html;jsessionid=838CCCAE6368D653EE43AE5BA237C815.intranet231#doc158972bodyText1).

<sup>114</sup> Data Documentation Initiative. Zugriff am 27.04.2023, <https://ddialliance.org/>.

<sup>115</sup> BIDS, Metadatenstandard für bildgebende Formate, Zugriff am 27.04.2023, <https://bids.neuroimaging.io/>.

EBRAINS entwickelten Standard openMINDS<sup>116</sup> (siehe auch Einheit 13 "Langzeitarchivierung", Punkt 4, S. 46). Für weitere konkrete Anforderungen innerhalb der Psychologie sollten ebenfalls spezifisch entwickelte Formate in Betracht gezogen werden, wie z. B. für die Statistik<sup>117</sup>.

Eine im Jahr 2021 veröffentlichte Handreichung beschreibt das Minimum an Metadaten, die in der Psychologie im Rahmen einer Veröffentlichung hilfreich wären<sup>118</sup>. Darüber hinaus bietet die DataWiz-Plattform<sup>119</sup> für die Psychologie relevante Hinweise, in welcher Art und Weise Metadaten den eigenen Daten hinzugefügt werden können. Ebenfalls hilfreich sind in diesem Zusammenhang verschiedene, von Projekten entwickelte open-source-Tools zur Generierung von Metadaten (DataCite Metadata Generator<sup>120</sup>, CFFNIT<sup>121</sup>, RADAR<sup>122</sup>).

## 7. Thesaurus, Normdaten und Klassifikationen in der Psychologie

Um Personen, Institutionen, Projekt-Eckdaten, Orte und vieles mehr ebenfalls eindeutig zu referenzieren, können sie mithilfe von sogenannten Normdateien, Thesauri und Klassifikationen erschlossen werden, welche als kontrollierte Vokabulare bezeichnet werden<sup>123</sup>.

**Normdateien** erleichtern als Identifikatoren beispielsweise die Suche nach Personen bei Namensgleichheit und können von Suchmaschinen eindeutig interpretiert werden. So ist die Psychologie in der sogenannten Gemeinsamen Normdatei (GND, siehe unten) z. B. unter der eindeutigen Nummer „4047704-6“ geführt<sup>124</sup>. Weitere Fachbereiche innerhalb der Psychologie und angrenzender Fachgebiete (u. a. Neurologie, Neuropsychologie, Medizin) sind ebenfalls durch Identifikatoren erschlossen.

Wichtige Normdateien sind:

- **GND**<sup>125</sup> – Die Gemeinsame Normdatei der Deutschen Nationalbibliothek dient neben der Katalogisierung von Literatur in Bibliotheken als etablierter Standard zur Erschließung und Referenzierung von Information.
- **ISNI**<sup>126</sup> – Der International Standard Name Identifier dient der eindeutigen Identifikation der öffentlichen Identität von Personen, die an einer Publikation beteiligt sind. Der ISNI ist ein Standard der Internationalen Organisation für Normung (ISO) und ist mit der ORCID vergleichbar.

<sup>116</sup> openMINDS, Metadatenstandard für das Projekt EBRAINS, Zugriff am 27.04.2023, <https://github.com/HumanBrainProject/openMINDS>.

<sup>117</sup> DDI-Standards für Statistik. Zugriff am 27.04.2023, <https://rdamsc.bath.ac.uk/subject/Statistics>.

<sup>118</sup> Blask, K., Jalynskij, M., & Gerhards, L. (2020). Metadata in Psychology 1.0: What researchers really need - Study description of the data referring to the expert interviews conducted in the BMBF-funded project PsyCuraDat. ZPID (Leibniz Institute for Psychology Information). <https://doi.org/10.23668/PSYCHARCHIVES.2750>.

<sup>119</sup> Metadata und Dokumentation. Zugriff am 27.04.2023, <https://datawizkb.leibniz-psychology.org/index.php/during-data-collection/what-should-i-know-about-metadata/>

<sup>120</sup> DataCite Metadata Generator. Zugriff am 27.04.2023, <https://dhvlab.gwi.uni-muenchen.de/datacite-generator/>.

<sup>121</sup> CFFNIT Metadaten. Zugriff am 27.04.2023, <https://citation-file-format.github.io/cff-initializer-javascript/#/start>.

<sup>122</sup> RADAR als Plugin für RDMO. Zugriff am 27.04.2023, <https://radar.products.fiz-karlsruhe.de/de/radarfeatures/radar-metadaten-schema>; Plugins für RDMO, Zugriff am 27.04.2023, <https://github.com/rdmorganiser/rdmo-plugins>.

<sup>123</sup> Bundesinstitut für Arzneimittel und Medizinprodukte, Zugriff am 27.04.2023, <https://www.dimdi.de/dynamic/de/klassifikationen/>.

<sup>124</sup> GND Nummer „Psychologie“. Zugriff am 27.04.2023, <https://lobid.org/gnd/4047704-6>.

<sup>125</sup> Deutsche Nationalbibliothek. „Gemeinsame Normdatei (GND)“. Zugriff am 27.04.2023, <https://www.dnb.de/>.

<sup>126</sup> International Standard Name Identifier. „About ISNI.“ Zugriff am 27.04.2023, <https://isni.org/>.

- **VIAF**<sup>127</sup> – Der Virtual International Authority File ist eine internationale Normdatei für Personen- und wird vom Online Computer Library Center (OCLC) verwaltet. Die Normdaten der GND und ISNI bilden einen Bestandteil von VIAF.
- **Open Funder Registry**<sup>128</sup> – Die Open Funder Registry dient der Identifikation von Forschungsförderern.

Des Weiteren dienen **Klassifikationen** in der Datendokumentation u. a. der Zuordnung von Objekten in (meist hierarchisch strukturierte) Klassen, welche durch bestimmte Merkmale charakterisiert sind. Für die Psychologie ergibt sich z. B. folgende Systematik des ZBW<sup>129</sup>: N.00 Nachbarwissenschaften (bswp. Geschichte, Politik und Politikwissenschaften, Kultur- und Geisteswissenschaften) > N.07 Psychologie > N.07.01 Arbeits- und Organisationspsychologie / N.07.02 Konsumentenpsychologie. Weitere solcher Klassifikationen finden auch für Fachliteratur, Tests, Methoden, Altersgruppen und Evidenzphasen Anwendung<sup>130</sup>.

Bei einem **Thesaurus** handelt es sich um eine natürlich-sprachliche, geordnete Sammlung von Begriffen und deren Beziehungen zueinander. Thesauri und kontrolliertes Vokabular werten Metadaten wesentlich auf und erhöhen die Referenzierbarkeit, Interoperabilität und Auffindbarkeit der Daten. Psychologie-spezifische Thesauri sind hier zu finden:

- Medizin und Biowissenschaften: Thesaurus Medical Subject Headings (MeSH)<sup>131</sup>
- Psychologie: deutschsprachiges Thesaurus PSYINDEX<sup>132</sup>, Thesauri der American Psychological Association<sup>133</sup>, (psychologie-spezifische Thesauri mit wirtschaftlich/ökonomischer Ausrichtung: Leibniz-Informationszentrum Wirtschaft (ZBW) Neurowissenschaften<sup>134</sup>, ZBW Psychologie<sup>135</sup>)
- Disziplinspezifische Klassifikationen und Thesauri können über das Basel Register of Thesauri, Ontologies & Classifications (BARTOC)<sup>136</sup> recherchiert werden.

## 8. Elektronische Laborbücher (ELB) in der Psychologie

Elektronische Laborbücher (ELB, oder auch ELN für Electronic Lab Notebook) dienen der Dokumentation der Konzeptionierung, Durchführung und Auswertung von wissenschaftlichen Experimenten, Beobachtungen oder Versuchen und den in diesem Zusammenhang erstellten Forschungsdaten. Sie sind die digitalen Versionen von Laborbüchern in Papierformat, die bisher überwiegend in natur- und lebenswissenschaftlichen Disziplinen einen wesentlichen Teil des wissenschaftlichen Arbeitsprozesses darstellten. Mit zunehmender Digitalisierung, v. a. bei der

<sup>127</sup> OCLC, „VIAF: Virtual International Authority File“. Zugriff am 27.04.2023, <http://viaf.org/>.

<sup>128</sup> Crossref, „Funder Registry.“ Zugriff am 27.04.2023, <https://www.crossref.org/education/funder-registry/>.

<sup>129</sup> ZBW Klassifikation Psychologie, Zugriff am 27.04.2023, <https://zbw.eu/stw/version/latest/thsys/70017/about.de.html>.

<sup>130</sup> PSYINDEX Thesaurus Psychologie. Zugriff am 27.04.2023, <https://www.psyndex.de/ueber/inhalte-aufbau/schlagwoerter-klassifikationen/#update-psyndex-terms-2019>.

<sup>131</sup> Thesaurus Medical Subject Headings. Zugriff am 27.04.2023, <https://www.nlm.nih.gov/mesh/>.

<sup>132</sup> PSYINDEX Thesaurus Psychologie. Zugriff am 27.04.2023, <https://www.psyndex.de/ueber/inhalte-aufbau/schlagwoerter-klassifikationen/#update-psyndex-terms-2019>.

<sup>133</sup> APA-Thesaurus für Psychologie. Zugriff am 27.04.2023, <https://www.apa.org/pubs/databases/training/thesaurus>.

<sup>134</sup> ZBW Leibniz-Informationszentrum Wirtschaft Thesaurus für Neurowissenschaften. Zugriff am 27.04.2023, <https://zbw.eu/stw/version/latest/descriptor/28954-1/about.de.html>.

<sup>135</sup> ZBW Leibniz-Informationszentrum Wirtschaft Thesaurus für Psychologie. Zugriff am 27.04.2023, <https://zbw.eu/stw/version/latest/descriptor/15790-5/about>.

<sup>136</sup> BARTOC.org. Zugriff am 27.04.2023, <http://www.bartoc.org>.

Erhebung von Daten, erfahren die ELB auch eine steigende Wahrnehmung und Nutzung in der Psychologie.

So zeigen sich für die Psychologie Anwendungsmöglichkeiten, wenn es z. B. um die digitale Archivierung ausgefüllter Fragebögen, Daten bildgebender Verfahren (z. B. EEG, MRT) oder die Protokolle einzelner experimenteller Messungen geht. Die Vorteile von ELBs zeigen sich vor allem in der automatisierten und digitalisierten Erhebung von Daten, deren Protokollierung und somit der Nachvollziehbarkeit hinsichtlich verschiedener Verarbeitung- und Analyseprozesse. Fragebögen, Terminvergabe, eMail-Versand usw. können durch automatisierte Prozesse gesteuert werden, was die lückenlose Erhebung von Daten erleichtert. Durch das Vorliegen der Daten in digitaler Form können diese Prozesse durch Wenn-Dann-Verknüpfungen entsprechend den Studienbedingungen gesteuert werden, wenn z. B. Cis\*-Frauen oder Trans\*-Männer durch die Angabe „Zyklusbeginn“ Terminvorschläge bekommen, um dahingehend eine Interferenz mit den Studienzielen zu vermeiden. Eine weitere Stärke zeigt sich in der automatischen Protokollierung jeder neuen Eingabe und Veränderung hinsichtlich des Autors und der Zeit (siehe Einheit 7 "Ordnung und Struktur", Punkt 5, S. 32), so dass jederzeit eine transparente Nachvollziehbarkeit gewährleistet ist.

Es existieren zahlreiche ELB-Softwareangebote - von kommerziellen Anbietern bis hin zu Open-Source-Lösungen, darunter

- PsychNotebook<sup>137</sup> (Open Source)
- REDcap<sup>138</sup> (Open Source)
- eLabFTW<sup>139</sup> (Open Source)
- Labfolder<sup>140</sup> (kommerziell)
- openBIS<sup>141</sup> (Open Source)

In der Regel haben die verschiedenen Disziplinen auch differierende Anforderungen an Funktionalitäten, die eine ELB-Software mit sich bringen sollte, sodass es keine "one-fits-all-Lösung" geben kann. In der von der ZBMed erstellten Handreichung<sup>142</sup> werden praktische Hinweise gegeben, die bei der Entscheidung und Einführung eines ELB in die jeweilige Forschungsgruppe oder das Projekt helfen können. Bei der Einführung von ELBs wird in der Regel eine technische Installation der Software in den vorhandenen IT-Infrastrukturen benötigt, für deren Betreuung auch personelle Ressourcen eingeplant werden müssen. Sollten diese Ressourcen fehlen, zeigt die TU Dresden Möglichkeiten auf, wie ELBs auch mittels herkömmlicher Software erzeugt und in den Arbeitsablauf integriert werden können<sup>143</sup>.

---

<sup>137</sup> PsychNotebook der Leibniz-Gesellschaft. Zugriff am 27.04.2023, <https://www.psychnotebook.org/>.

<sup>138</sup> REDCap, Zugriff am 27.04.2023, <https://projectredcap.org/>.

<sup>139</sup> eLabFTW ELN. Zugriff am 27.04.2023, <https://www.elabftw.net>.

<sup>140</sup> LabFolder ELN. Zugriff am 27.04.2023, <https://www.labfolder.com>.

<sup>141</sup> openBIS Laboratory Notebook. Zugriff am 27.04.2023, <https://openbis.ch>.

<sup>142</sup> ZB MED (Hrsg.). Elektronische Laborbücher im Kontext von Forschungsdatenmanagement und guter wissenschaftlicher Praxis – ein Wegweiser für die Lebenswissenschaften. Köln, 2020. <http://dx.doi.org/10.4126/FRL01-006422868>.

<sup>143</sup> Informationssammlung der TU Dresden. Zugriff am 27.04.2023, <https://tu-dresden.de/forschung-transfer/services-fuer-forschende/kontaktstelle-forschungsdaten/news/elektronisches-laborbuch>.



## Zugriffssicherheit (Einheit 9)

### 1. Motivation

Um Forschungsdaten einerseits vor Datenverlust schützen, und andererseits deren Missbrauch zu verhindern, sollten im Datenmanagement die Aspekte Sicherheit und Zugriffsrechte auch durch entsprechende Maßnahmen reflektiert werden. Insbesondere sensible Daten - personenbezogene Daten oder Daten, deren Schutz vertraglich zugesichert wurde (z. B. Firmengeheimnisse, Auftragsforschung) - erfordern spezifische Schutzmodelle. Aber auch bislang unpublizierte wissenschaftliche Ergebnisse bedürfen des Schutzes.

Ein solcher Schutz kann z. B. mittels Verschlüsselung, Backups, Speicherung auf vertrauenswürdigen Speichermedien und einer spezifischen Regelung von Zugriffsrechten sichergestellt werden. Um zu späteren Zeitpunkten entsprechende Maßnahmen nachvollziehen zu können, sollten dazu Protokolle erstellt und sicher aufbewahrt werden. Außerdem sollten nur die Personen Zugriff auf die Daten erhalten, welche diese auch zum Arbeiten benötigen. Gerade bei einer institutionenübergreifenden Zusammenarbeit ist die Regelung dieser Aspekte von besonderer Wichtigkeit.

Es ist außerdem auch sinnvoll, die für die Datensicherheit geeigneten technischen und organisatorischen Maßnahmen einzurichten, zu dokumentieren und immer wieder zu überprüfen, ob diese noch dem aktuellen Bedarf entsprechen.

### 2. Verschlüsselung

Ein physikalischer Zugriff auf einen Rechner kann einen unbefugten Zugriff auf Daten zur Folge haben. Somit erschließt sich die Notwendigkeit einer Verschlüsselung der Daten, welche jedoch nur sinnvoll ist, wenn auch sämtliche betroffenen Daten verschlüsselt werden und auch allen Beteiligten die Speicherorte bekannt sind. Dies gilt auch für deren Kopien und Backups.

Es können sowohl ausgewählte Speicherorte als auch ganze Datenträger verschlüsselt werden, z. B. durch automatische Verschlüsselungslösungen mit Programmen wie z. B. FileVault<sup>144</sup>, BitLocker<sup>145</sup> oder dm-crypt<sup>146</sup>. Dabei müssen die Parameter für alle vorgesehenen Datenspeicherorte gleich eingestellt werden. Es bietet sich ebenfalls eine Dateiverschlüsselung als zusätzliche Schutzmaßnahme bei einem Upload von Daten in eine Cloud-Umgebung an, sowie bei der Weitergabe von Daten per E-Mail. Bei letzterem verhindert sie unberechtigtes Lesen und Verändern von Nachrichtentexten und Anhängen auf dem Weg zum Empfänger (Ende-zu-Ende-Verschlüsselung).

### 3. Überprüfen der Nutzbarkeit

Durch Festplattenfehler oder Programmabstürze können Daten nicht mehr lesbar werden, weswegen eine Datensicherung nur dann hilfreich ist, wenn die Wiederherstellung der Daten gewährleistet ist. Gelegentlich verursacht auch das Kopieren von Dateien selbst Fehler. Es ist daher grundsätzlich empfehlenswert, die Datenwiederherstellung zu Beginn der Sicherung sowie in regelmäßigen Abständen auf etwaige Fehler in den Daten zu testen, um so einen Datenverlust zu verhindern.

---

<sup>144</sup> FileVault: Verschlüsselungsfunktion zum Verschlüsseln von persönlichen Daten unter Apple-Betriebssystemen

<sup>145</sup> BitLocker: Festplattenverschlüsselung von Microsoft

<sup>146</sup> dm-crypt: Verschlüsselungsfunktion unter Linux



Neben der Kontrolle der Lesbarkeit sollte auch eine Virenprüfung durchgeführt werden, um die Überschreibung unbeschädigter Backup-Dateien durch fehlerhafte oder virenbelastete Dateien auszuschließen.

#### 4. Passwortschutz und Zugriffsrechte

Bei der Arbeit mit schutzwürdigen Daten sollten sichere Passwörter vergeben und der Zugriff auf den Kreis der unmittelbar an den Daten arbeitenden Personen beschränkt werden<sup>147</sup>.

Ein sicheres Passwort hat folgende Kennzeichen<sup>148</sup>:

- es sollte aus mindestens acht Zeichen bestehen (weniger Zeichen erhöhen die Gefahr, daß das Passwort ohne großen Aufwand entschlüsselt werden kann)
- es enthält Klein- und Großbuchstaben sowie Sonderzeichen und Zahlen
- verwendete Zeichen sollten auf der Tastatur nicht nebeneinanderliegen
- das Passwort sollte möglichst nicht in Wörterbüchern vorkommen

Zu vermeiden sind einfache Passwörter wie 123456, passwort, 111111, qwertz, abc123 oder admin. Namen, Geburtstage, Begriffe aus Wörterbüchern, Filmfiguren oder Autokennzeichen sollten ebenfalls nicht als Passwort genutzt werden – auch dann nicht, wenn sie rückwärts geschrieben werden. Ebenso ergibt das Anhängen von Zahlen oder Sonderzeichen an ein Wort kein sicheres Passwort.

Über die Vergabe von Berechtigungen wird festgelegt, welche Personen bzw. Personenkreise mit welchen Rechten auf bestimmte Verzeichnisse und Dateien zugreifen dürfen. Dabei ist die Vergabe von abgestuften Lese- und Schreibrechten sowie Ausführungsrechten möglich. Es ist dabei wichtig, die Vergabe wohlgedacht vorzunehmen, um den Arbeitsfluss nicht zu behindern: sind die Zugriffsrechte zu locker vergeben, können Personen unerwünscht auf sensible Daten zugreifen; sind sie hingegen zu restriktiv vergeben, wird eine Nachnutzung erschwert oder gar unterbunden.

---

<sup>147</sup> Leitlinien der GESIS zum Umgang mit Zugriffsrechten und Sicherheit der Daten. Zugriff am 27.04.2023, [https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_methodenberichte/2012/TechnicalReport\\_2012-07.pdf](https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2012/TechnicalReport_2012-07.pdf)

<sup>148</sup> Bundesamt für Sicherheit in der Informationstechnik. „Sichere Passwörter erstellen“. Zugriff am 27.04.2023, [https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/sichere-passwoerter-erstellen\\_node.html](https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/sichere-passwoerter-erstellen_node.html).

## Speicherung und Backup (Einheit 12)

### 1. Motivation

Um dem Verlust von wertvollen Forschungsdaten entgegenzuwirken, sollten diese regelmäßig gespeichert und zusätzlich durch Backups abgesichert werden. Im Rahmen eines Forschungsdatenmanagementplans kann so für alle Mitglieder einer Forschungsgruppe verbindlich festgelegt werden, wann auf welchen Datenträgern in welchem Format die Daten abgelegt werden sollen, und in welchen Abständen Backups erfolgen sollen. Dies ist vor allem bei institutsübergreifenden Forschungsprojekten sehr sinnvoll, um ein einheitliches Vorgehen zu gewährleisten.

### 2. Datenspeicherung

Forschungsdaten können auf unterschiedlichen Medien gespeichert werden, die hinsichtlich des Schutzes vor Datenverlust und vor unbefugtem Zugriff ihre jeweiligen Vorteile und Risiken aufweisen (siehe Tabelle 5).

*Tabelle 5: Überblick über die Eigenschaften, Vorteile und Risiken der häufigsten Speichermedien und -orte.*

	Vorteile	Risiken
<b>Eigener PC</b>	<ul style="list-style-type: none"> <li>• selbst verantwortlich für Sicherheit und Backup</li> <li>• eigene Kontrolle</li> </ul>	<ul style="list-style-type: none"> <li>• evtl. fehlende Ressourcen und Knowhow zum Konfigurieren und Überprüfen der Qualität der Sicherungskopien;</li> <li>• Verlust bzw. Hard- und Softwareprobleme des PC betreffen auch das Backup</li> <li>• Individuallösungen im Rahmen einer Arbeitsgruppe sind aufwendig, kostspielig und ineffizient</li> </ul>
<b>Mobile Speichermedien</b>	<ul style="list-style-type: none"> <li>• einfach zu transportieren</li> <li>• können im verschließbaren Schrank oder Safe aufbewahrt werden</li> </ul>	<ul style="list-style-type: none"> <li>• Verlust</li> <li>• Diebstahl</li> <li>• Inhalte sind ungeschützt, wenn sie nicht zuvor verschlüsselt wurden (siehe Einheit 9, Punkt 2, S. 40)</li> <li>• stoß- und verschleißanfällig</li> </ul>

<b>Institutionelle Speicherorte</b>	<ul style="list-style-type: none"> <li>• Backup der Daten ist sichergestellt</li> <li>• professionelle Durchführung und Wartung</li> <li>• Speicherung entsprechend den Datenschutzrichtlinien der Institution</li> <li>• Datenschutz über Zugriffsrechte geregelt</li> <li>• für mobiles Arbeiten nutzbar</li> </ul>	<ul style="list-style-type: none"> <li>• Geschwindigkeit vom Netzwerk abhängig</li> <li>• Zugriff auf Backups evtl. verzögert durch Dienstweg</li> <li>• evtl. unklar, welche Sicherheitskriterien und -strategien eingesetzt werden</li> <li>• ggf. mit höheren Kosten verbunden</li> </ul>
<b>Externe Speicherorte</b>	<ul style="list-style-type: none"> <li>- einfach zu nutzen und zu verwalten</li> <li>- werden professionell gewartet</li> <li>- für mobiles Arbeiten nutzbar</li> </ul>	<ul style="list-style-type: none"> <li>- je nach Anbieter unsichere Verbindung</li> <li>- Abhängigkeit vom Internetzugang</li> <li>- Upload und Download kann lange dauern</li> <li>- Zugriff auf Backups evtl. verzögert</li> <li>- Datenschutz unklar (ob Sicherheitskriterien und -strategien den Vorgaben für sensible Daten oder der eigenen Institution entsprechen, siehe Freie Universität Berlin<sup>149</sup>)</li> </ul>

Kostenlose Cloud-Speicherdienste sind keine geeigneten Speicherorte für schützenswerte Daten, da nicht immer sichergestellt wird, wie sicher die Daten hinterlegt sind und was mit den Daten geschieht. Auch wenn es die Aufgabe der Anbieter von Cloud-Diensten ist, missbräuchliche Nutzung von Daten auszuschließen, ist es unklar, ob dies allumfassend gelingt.

Forschende können aber auch auf der physischen Ebene einiges zur Sicherung ihrer (sensiblen) Daten beitragen (siehe Einheit 9 "Zugriffssicherheit", S. 40). Passwortgeschützte externe Festplatten können zum Beispiel in einem separaten, abschließbaren Raum oder Schrank aufbewahrt werden, und Notebooks vor Diebstahl durch ein Schloss gesichert werden. Wichtig ist dabei jedoch, dass mindestens zwei Personen Zugang zu den Daten haben sollten, um auch im Krankheitsfall oder bei Abwesenheit die Verfügbarkeit der Daten zu gewährleisten.

### 3. Backup-Varianten

Backup bezeichnet die Erstellung einer Sicherungskopie der Daten auf einem anderen Speichermedium. Es sollte bereits frühzeitig eine Strategie entwickelt werden, die regelmäßige BackUps und Neukäufe von Ersatzfestplatten als Teil der Arbeitsroutinen miteinschließt, um im Bedarfsfall eine möglichst einfache Datenrekonstruktion sicherzustellen. Viele Betriebssysteme sind zu diesem Zweck bereits mit eigenen Backup-Programmen ausgestattet und erfordern keinen zusätzlichen Installationsaufwand (z. B. die Time Machine-Funktion des Mac OSX Betriebssystems).

<sup>149</sup> Freie Universität Berlin. „Richtlinie zur Auslagerung von Daten in die Cloud“. Zugriff am 27.04.2023, [https://www.fu-berlin.de/sites/it-sicherheit/downloads/Richtlinie\\_Cloud-Datenablage\\_-\\_1\\_0.pdf](https://www.fu-berlin.de/sites/it-sicherheit/downloads/Richtlinie_Cloud-Datenablage_-_1_0.pdf).

Für die Speicherung gibt es zwei verschiedene Sicherungsarten: inkrementell oder differenziell. Bei beiden Varianten wird als erstes ein Voll-Backup erstellt. Danach werden bei der

- **inkrementellen Datensicherung** nur die Dateien oder Teile von Dateien gespeichert, die sich seit der letzten inkrementellen Sicherung geändert haben oder neu hinzugekommen sind.
- **differenziellen Datensicherung** alle Daten gespeichert, die sich seit dem letzten Voll-Backup geändert haben oder neu hinzugekommen sind.

#### 4. Hinweise für ein sicheres Backup

Es wird empfohlen, die Daten mindestens einmal täglich zu sichern sowie ein wöchentliches Gesamtbackup durchzuführen. Als hilfreich hat sich die "3-2-1 Regel" herausgestellt:

- mindestens drei Kopien der Daten,
- auf mindestens zwei verschiedenen Speichermedien und
- eine davon sollte an einem anderen Ort (dezentral) hinterlegt sein.

Dezentrale Datenträger sollten von der üblicherweise genutzten Infrastruktur (z. B. Server, PC) abgekoppelt sein und die Datenwiederherstellung zu Beginn sowie in regelmäßigen Abständen auf Fehler überprüft werden.

Die meisten Institutionen bieten zudem automatisierte Lösungen an, bei der alle Daten ausschließlich auf von der EDV-Abteilung gesicherten Laufwerken abgelegt und durch Backups gesichert werden. Dies verhindert, dass Sicherungen vergessen werden und dass die Konfiguration des Backup-Systems für jeden einzelnen Datenträger einzeln zu erfolgen hat.

## Langzeitarchivierung (Einheit 13)

### 1. Motivation

Um Daten auch in Zukunft recherchier- und ggf. nachnutzbar zu halten, werden sie archiviert. Der Kodex zur Guten Wissenschaftlichen Praxis der DFG fordert hierfür eine i.d.R. 10 Jahre lange Verfügbarkeit der relevanten Forschungsdaten<sup>150</sup> ab dem Zeitpunkt der Veröffentlichung der Forschungsergebnisse. Darauf basierend haben viele wissenschaftliche Institutionen diese Forderung auch in ihren Forschungsdaten-Policies und Leitlinien aufgenommen.

### 2. Begriff „Langzeitarchivierung“

„Langzeit“ ist ein Hilfswort für die Beschreibung eines nicht näher definierten Zeitraumes, in dem technologische und soziokulturelle Veränderungen eintreten können, die den Erhalt, den Zugriff auf, die Recherche in und die Nachnutzung von digitalen Forschungsdaten beeinflussen können. Somit muss die digitale Langzeitarchivierung Maßnahmen beinhalten, die geplant, kontrolliert und umgesetzt werden müssen, um diesen Aspekten Rechnung zu tragen.

### 3. Nachhaltige Dateiformate in der Psychologie

Für eine Langzeitarchivierung ist nicht jedes Dateiformat gleich gut geeignet. Um auch eine zukünftige Arbeit mit Daten zu gewährleisten, sollten Dateiformate gewählt werden, die nicht an kostenpflichtige Software gebunden sind, um z. B. eine höhere Interoperabilität zu erreichen (siehe Tabelle 6). In diesem Zusammenhang unterscheidet man vor allem zwischen sogenannten proprietären und offenen Formaten. Proprietäre Formate sind solche, deren Nutzung eine kostenpflichtige Software erfordert (z. B. Microsoft Office, AutoCAD, SPSS, MaxQDA). Die zu archivierenden Dateien sollten dabei unverschlüsselt, nicht komprimiert, patentfrei und im offenen, dokumentierten Standard erstellt sein.

Mitunter sind proprietäre Dateiformate für die eigene Arbeit unerlässlich. Für die Langzeitarchivierung hingegen sollten etwaige Daten in andere, offene Formate konvertiert werden. Dabei ist zu prüfen, ob die Konvertierung erfolgreich war und das Format valide ist, da Software auch Fehler produzieren kann. Sowohl die Originaldatei als auch die Datei im konvertierten Format sollte gespeichert werden.

Für die Psychologie existieren neben den fachunabhängigen Empfehlungen<sup>151,152</sup> für nichtproprietäre Versionen gängiger Text-, Tabellen- und Bildformate auch Empfehlungen für die bildgebenden Verfahren (siehe Tabelle 6). Bei der Aufzeichnung struktureller und funktioneller Daten durch die Magnetresonanztomographie wurden hier schon sehr früh offene Datenformate wie z. B. DICOM<sup>153</sup> oder NifTI<sup>154</sup> genutzt. In Bezug auf die Aufzeichnung der Gehirnströme mittels Magnet-/Elektroenzephalographie sind die Aufnahmen jedoch häufig noch durch spezifische Datenformate an

<sup>150</sup> Deutsche Forschungsgemeinschaft. *Sicherung guter wissenschaftlicher Praxis. Denkschrift*. Weinheim: Wiley-VCH, 2013. S. 21–22. <https://doi.org/10.1002/9783527679188.oth1>

<sup>151</sup> IANUS - Forschungsdatenzentrum für Archäologie & Altertumswissenschaften. „IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften. Dateiformate“. Version: 1.0.1.0. Berlin: IANUS - Forschungsdatenzentrum für Archäologie & Altertumswissenschaften, 2017. <http://doi.org/10.13149/000.111000-a>.

<sup>152</sup> UK Data Service. „Recommended formats.“ Zugriff am 27.04.2023, <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>.

<sup>153</sup> DICOM Datenformatstandard. Zugriff am 27.04.2023, <https://www.dicomstandard.org/>.

<sup>154</sup> NifTI Datenformatstandard. Zugriff am 27.04.2023, <https://nifti.nimh.nih.gov/>.

die Benutzung firmeneigener (proprietärer) Software für die Analyse gebunden (z. B. bei BioSemi, BrainVision, NeuroScan). Durch die Entwicklung spezieller Softwarepakete (Matlab<sup>155</sup>-basierte open source-Anwendungen EEGLab<sup>156</sup> und Fieldtrip<sup>157</sup>) ist man jedoch in der Lage, die Verarbeitung und Analysen dieser Daten unabhängig von proprietärer Software zu gestalten<sup>158,159</sup>.

Allerdings empfiehlt es sich, die aufgezeichneten Daten zusätzlich noch in offene Datenformate umzuformatieren, um eine zukünftige Verarbeitung der Daten unabhängiger von den Entwicklungen in der Wirtschaft zu machen. Mit der Entwicklung des Datenformats BIDS<sup>160</sup> wurde Psycholog\*innen ein weiterer freier Datenstandard zur Verfügung gestellt. Dieser zeichnet sich vor allem durch die hohe Interoperabilität in Bezug auf Analyseprogramme (siehe die obengenannten Software-Pakete) aus, als auch durch die automatische Implementierung von Metadaten. Zunächst entwickelt für Magnetresonanzaufnahmen, wurde dieser mittlerweile für die Magnet/Elektroenzephalografie erweitert<sup>161,162</sup>.

*Tabelle 6: Empfehlungen nachhaltiger Formate für Langzeitarchivierungen.*

Dateiformat	Empfehlung	Vermeiden
<b>Tabellen</b>	CSV, TSV, SPSS portable, XLSX	XLS, SPSS
<b>Text</b>	TXT, HTML, RTF, PDF/A, DOCX	DOC, PDF
<b>Multimedia</b>	Container: MPEG4, MKV Codec: Theora, Dirac, FLAC	QuickTime, Flash
<b>Bilder</b>	TIFF, JPEG2000, PNG	GIF, JPG
<b>Bildgebende Verfahren</b>	European Data Format (EDF), HL7 v3.0 annotated ECG (ECG), General Data Format for biomedical signals (GDF), DICOM, BIDS (standard format for MRI, MEG and EEG data)	EEG (BrainVision), BDF (BioSemi), CNT (Neuroscan), MEG Formate spezif. Firmen (CTF, FIF, 4d, KIT, KDF, ITAB)

#### 4. Anforderungen an Langzeitarchive

Bei der Wahl eines geeigneten Speicherortes für die Langzeitarchivierung sollten folgende Punkte beachtet werden:

- **Technische Anforderungen:** Ein Dienstleister (institutionalisiert oder extern) sollte eine Strategie zur Datenkonvertierung und Migration verfolgen. Darüber hinaus sollte in

<sup>155</sup> Matlab. Zugriff am 27.04.2023, <https://www.mathworks.com/products/matlab.html>.

<sup>156</sup> Fieldtrip. Zugriff am 27.04.2023, <https://www.fieldtriptoolbox.org/>.

<sup>157</sup> EEGLab. Zugriff am 27.04.2023, <https://eeglab.org/>.

<sup>158</sup> Von Fieldtrip unterstützte Datenformate. Zugriff am 27.04.2023, <https://www.fieldtriptoolbox.org/faq/dataformat/>.

<sup>159</sup> Von EEGLab unterstützte Datenformate. Zugriff am 27.04.2023, [https://scn.ucsd.edu/eeglab/plugin\\_uploader/plugin\\_list\\_all.php](https://scn.ucsd.edu/eeglab/plugin_uploader/plugin_list_all.php).

<sup>160</sup> BIDS Spezifikationen. Zugriff am 27.04.2023, <https://zenodo.org/record/6094534#.Yt-3cd9CRaQ>.

<sup>161</sup> Niso, G., Gorgolewski, K., Bock, E. *et al.* MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Sci Data* **5**, 180110 (2018). <https://doi.org/10.1038/sdata.2018.110>.

<sup>162</sup> Pernet, C.R., Appelhoff, S., Gorgolewski, K.J. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci Data* **6**, 103 (2019). <https://doi.org/10.1038/s41597-019-0104-8>.

regelmäßigen Abständen eine Kontrolle der Lesbarkeit der Dateien sowie eine Virenprüfung durchgeführt werden. Alle Schritte sollten dokumentiert werden.

- **Siegel für vertrauenswürdige Langzeitarchive:** Für eine externe Beurteilung, ob oder in welchem Rahmen ein Langzeitarchiv vertrauenswürdig ist, werden verschiedene Siegel mit unterschiedlichen Prüfschwerpunkten, entwickelt (z. B. das nestor-Siegel <sup>163</sup> oder das CoreTrustSeal<sup>164</sup>).
- **Kosten der Services:** Es ist immer zu prüfen, ob Dienstleister Kosten für eine Datenspeicherung erheben, welche im Falle von institutionalisierten Repositorien kaum eine Rolle spielen. Diese können beispielsweise von der Menge der Daten, der Umsetzung von technischen Standards oder der Affiliation der Datenprovider abhängen.
- **Zugänglichmachung der Daten:** Vor der Wahl des Speicherortes sollte man sich die Frage stellen, ob die Daten zugänglich sein oder nur abgelegt werden sollen.
- **Langlebigkeit des Dienstleisters:** Wirtschaftliche und politische Faktoren haben Einfluss auf die Langlebigkeit der Dienstleister. Im Kontext von Forschungseinrichtungen spielt dieser Aspekt kaum eine Rolle.

---

<sup>163</sup> Nestor-Siegel. Zugriff am 27.04.2023, [https://www.langzeitarchivierung.de/Webs/nestor/DE/Zertifizierung/nestor\\_Siegel/siegel.html](https://www.langzeitarchivierung.de/Webs/nestor/DE/Zertifizierung/nestor_Siegel/siegel.html)

<sup>164</sup> CoreTrustSeal. Zugriff am 27.04.2023, <https://www.coretrustseal.org/>

## Institutionelle Infrastruktur (Einheit 14)

### 1. Motivation

Institutionelle FDM-Infrastrukturen ermöglichen den Forschungseinrichtungen ihren Forscher\*innen die Einhaltung von Vorgaben bzw. Orientierungen wie dem Kodex zur „Guten wissenschaftlichen Praxis“ oder den FAIR-Prinzipien in Bezug auf den Umgang mit Forschungsdaten unabhängig von finanziellen Ressourcen nachzukommen. Dabei unterstützen zunehmend auch zentrale Koordinationsstellen die Forscher\*innen durch Bereitstellung von Beratungen, generischen und fachspezifischen Schulungen sowie Handreichungen, Leitlinien und weiteren Informationsangeboten.

### 2. Institutionelles Forschungsdatenmanagement

An den Bildungseinrichtungen wird das Thema des Forschungsdatenmanagements unterschiedlich gehandhabt. Es gibt Einrichtungen, die bereits seit Jahren eine Koordinationsstelle haben, bei anderen befinden sich entsprechende Angebote noch im Aufbau oder in der Planung. Häufig sind (neben den jeweiligen Fachbereichen) folgende Abteilungen bzw. Einrichtungen daran beteiligt, einen FDM-Service anzubieten:

- FDM-Kontakt- bzw. -Servicestelle (wenn vorhanden)
- Bibliothek
- IT/Rechenzentrum
- Forschungsförderung
- rechtliche Anlaufstellen (z. B. Datenschutzbeauftragte, Justizariat)

### 3. Forschungsdatenmanagement-Infrastrukturen

Hochschulen haben die Aufgabe, ihre Forschenden mit den geeigneten Werkzeugen und Dienstleistungen bestmöglich beim Forschungsdatenmanagement zu unterstützen - sowohl **organisatorisch** als auch **technisch**. Je nach Einrichtungen können folgende Services zur Verfügung stehen:

#### auf organisatorischer Ebene

- Beratung zum Umgang mit Forschungsdaten im Kontext von Projektanträgen bzw. Drittmittelbeantragungen
- ggf. fachspezifische Schwerpunkte z. B. Beratung zu Digital Humanities-Anwendungen
- Vermittlung zu internen und externen Diensten
- Unterstützung bei der Erstellung eines Datenmanagementplans
- Schulungen und/oder Workshops zum Thema Forschungsdatenmanagement
- Informationsveranstaltungen
- Informationsmaterial
- Rechtliche Beratung zum Thema Forschungsdaten (auch in Kooperation mit Experten)

#### auf technischer Ebene

- Cloud-Services
- Sync-and-Share-Lösungen
- Backup-Service
- Datenbank-Service



- Forschungsdaten-Repository
- Versionierungs-Software
- Langzeitarchivierung
- Vergabe von persistenten Identifiern
- DMP-Tool
- Elektronisches Laborbuch

## Publikation von Forschungsdaten (Einheit 15)

### 1. Ziele und Motivation

Die Publikation von Forschungsdaten nach den FAIR-Prinzipien (siehe Einheit 4 „Forschungsdatenmanagement“, Punkt 2c, S. 15) ermöglicht vor allem Anschluss- und Folgeforschung durch andere Forscher und dient der Transparenz der Forschung und der Reproduzierbarkeit der Daten, was die wissenschaftliche Integrität der datenerstellenden Forscher\*innen erhöht (eine Replikation beweist die Sorgfalt bei der Datenerhebung und -analyse der ursprünglichen Daten, wenn die Ergebnisse deckungsgleich sind). Zudem gibt es Hinweise, dass Artikel, deren Daten publiziert wurden, häufiger zitiert werden<sup>165, 166, 167, 168</sup>, was die Reputation der entsprechenden Forschenden erhöht. Durch die Kontaktaufnahme von an publizierten Daten interessierte Forscher\*innen kann sich zudem auch der Kreis möglicher zukünftiger Kooperationspartner erweitern – wenn sich über die Datennachnutzung hinaus das gemeinsame Interesse an spezifischen Fragestellungen zeigt.

Als eine starke Motivation für die Publikation von Daten kann auch die Ressourcenschonung betrachtet werden, denn sie vermeidet vor allem eine doppelte Datenerhebung zu exakt der gleichen Fragestellung – wenn die Publikation aufzeigt, dass dieser Frage schon einmal nachgegangen wurde. Zudem ist nicht jede Forschungseinrichtung in der Lage, ausreichend finanzielle und/oder technische Ressourcen für spezifische Forschung zur Verfügung zu stellen (z. B. übersteigt ein MRT-Gerät oft den finanziellen Rahmen einer Forschergruppe). So können bereitgestellte Daten auch ressourcenschwächeren Forschungskontexten zur Verfügung gestellt werden. Auch im Sinne der Studierendenausbildung in Bezug auf Methodik und Statistik ist die Publikation von Forschungsdaten sinnvoll, da eine erfolgreiche Replikation der Daten den Lernerfolg der Studierenden widerspiegelt.

Nicht zuletzt wird die Publikation von Daten zunehmend von Verlagen, Institutionen und Fördermittelgebern gefordert.

### 2. Leitfragen für die Datenauswahl

Für jede Publikation ist zu entscheiden, unter welchen Bedingungen sie veröffentlicht wird, zum Beispiel, ob sie frei zugänglich gemacht wird (Open Access) oder zugriffsgeschützt archiviert werden soll. Der Wettbewerb innerhalb der Wissenschaft kann eine eingeschränkte oder zeitverzögerte Publikation jedoch sinnvoll machen: sollen aus den erhobenen Daten noch weitere eigene Publikationen entstehen, so spielen z. B. der Zeitpunkt der Publikation der Daten und die Wahl des Publikationsmodells eine entscheidende Rolle.

---

<sup>165</sup> Pienta, Amy M., George C. Alter und Jared A. Lyle. *The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data*. Ann Arbor, Mi: University of Michigan – Deep Blue, 2010. <http://hdl.handle.net/2027.42/78307>

<sup>166</sup> Piwowar, Heather A und Todd J. Vision. Data reuse and the open data citation advantage. *PeerJ* 1 (2013), S. e175. <https://doi.org/10.7717/peerj.175>;

<sup>167</sup> Henneken, Edwin A. und Alberto Accomazzi. *Linking to data-effect on citation rates in astronomy*. arXiv: 2011, <https://arxiv.org/abs/1111.3618>;

<sup>168</sup> Piwowar, Heather A., Roger S. Day und Douglas B. Fridsma: Sharing detailed research data is associated with increased citation rate. *PLoS ONE* 2 (2007), S. e308. <https://doi.org/10.1371/journal.pone.0000308>.

Bei der Auswahl können folgende Leitfragen<sup>169</sup> helfen:

- Handelt es sich um eine abgeschlossene Datenerhebung oder um einen kumulativen Datensatz (das Projekt beinhaltet mehrere aufeinander folgende Erhebungsphasen)?
- Zu welchem Zeitpunkt im Forschungsprozess werden die Daten publiziert (z. B. sind es noch noch unbearbeitete Rohdaten, oder bereits Primär- oder Sekundärdaten)?
- Aus welcher Motivation heraus werden sie publiziert?
- Werden Rohdaten oder bearbeitete Daten veröffentlicht?
- Sollen die Daten einem Peer-Review-Prozess unterzogen werden?

Nicht immer können Nachnutzungsszenarien vollständig antizipiert werden. Nachfolgende Studien können Datensätze unter gänzlich anderen Forschungsfragen<sup>170</sup> untersuchen, als dies in der Ursprungsstudie intendiert war, oder deren Metadaten<sup>171</sup> auswerten. Grundsätzlich empfiehlt sich daher neben der Veröffentlichung auch eine umfassende Dokumentation dieser Forschungsdaten (siehe Einheit 8 „Dokumentation und Metadaten“, Punkt 2, S. 34) auch wenn zunächst nicht klar erkennbar ist, welchen weiteren Forschungsfragen nachgegangen wird.

### 3. Publikationswege

Für die Datenpublikation bieten sich verschiedene Möglichkeiten. Etabliert ist heute **(I)** die Veröffentlichung von aggregierten Daten als Supplement zu einem wissenschaftlichen Artikel (Journal-Publikation). Hinzukommen **(II)** die Datenpublikation in einem generischen, institutionsspezifischen oder fachspezifischen Repositorium als eigenständige Informationsobjekte<sup>172</sup> sowie **(III)** die Veröffentlichung der Datenbeschreibung in sog. Data Journals. Dabei handelt es sich um Fachzeitschriften, die sich auf die Veröffentlichung von Forschungsdaten spezialisiert haben.

**(I) Supplement:** Daten können als Supplemente zu Veröffentlichungen von wissenschaftlichen Forschungsergebnissen in Fachartikeln (Journals etc.) publiziert werden. Diese Daten ergänzen und belegen die im Artikel präsentierten Forschungsergebnisse. Dabei handelt es sich meist um aggregierte Daten in Form kleiner Tabellen oder Bilder. Die in dieser Form veröffentlichten Daten sind in der Regel nur in Kombination mit dem zugehörigen Artikel auffindbar und zugänglich und entsprechen in den meisten Fällen nicht den FAIR-Prinzipien.

**(II) Repositorium:** Daten können als ein eigenständiges Informationsobjekt in einem Repositorium publiziert werden. Sofern diese Daten Grundlage eines wissenschaftlichen Artikels sind, können sie in der Regel über die Metadaten mit der Publikation verknüpft werden (siehe Repositorien). In disziplinspezifischen Repositorien können diese Daten zudem besser kontextualisiert bzw. mit anderen Datensammlungen verknüpft werden, indem sie zusätzliche Recherche-, Analyse- und Visualisierungswerkzeuge bereitstellen. Disziplinübergreifende und institutionelle Repositorien sind zwar nicht fachbereichsspezifisch ausgelegt, die Daten selbst sind aber aufgrund ihrer Metadaten

<sup>169</sup> angelehnt an Martin, Elaine R. (Ed.). *New England Collaborative Data Management Curriculum. Module 6. Data Sharing & Reuse Politics*. Zugriff am 27.04.2023, [https://library.umassmed.edu/docs/necdmc\\_module6.docx](https://library.umassmed.edu/docs/necdmc_module6.docx).

<sup>170</sup> Steiner, Daniel, Heinz J. Zumbühl und Andreas Bauder. „Two Alpine Glaciers over the Past Two Centuries.“ In *Darkening Peaks: Glacier Retreat, Science, and Society*. Herausgegeben von Ben Orlove. S. 83–99. Berkeley, CA: University of California Press, 2008.

<sup>171</sup> McKiernan, Erin C., Philip E. Bourne, C. Titus Brown, Stuart Buck, Amye Kenall, Jenifer Lin, Damon McDougall et al.: Point of view: How open science helps researchers succeed. *eLife* 5 (2016), S. e16800. <https://doi.org/10.7554/eLife.16800>.

<sup>172</sup> Informationsobjekte, auch Entitäten genannt, können aus Datenbanken, Dokumenten, Nachrichten oder jeder weiteren Art von Information bestehen, Zugriff am 27.04.2023, <https://project-base.org/projektmanagement-glossar/informationsobjekt-eepek/>.

entsprechend thematisch such- und auffindbar. Zudem können institutionelle Repositorien eine gute Übersicht der an einer Institution gesammelten Daten bieten, sofern diese auf dem Repository abgelegt werden.

**(III) Data Journals:** Data Journals widmen sich der Veröffentlichung von Daten und entsprechenden Informationen zu diesen publizierten Daten, die in frei zugänglichen oder zugangsbeschränkten Repositorien publiziert werden. Bei den entsprechenden Informationen handelt es sich um ausführliche Dokumentationen von veröffentlichten Daten, deren Eigenschaften sowie Angaben zur potentiellen Nachnutzung. Die Daten im Repository und ihre Dokumentation in der Zeitschrift werden mittels eines persistenten Identifiers (siehe Identifier) miteinander verknüpft und sind somit eindeutig auffindbar. Zudem bieten einige dieser Zeitschriften ein Peer-Review-Verfahren an, bei dem der Datensatz sowie seine Dokumentation durch externe Gutachter geprüft werden, z. B. ob die jeweilige Dokumentation die damit verknüpften Daten abbildet, ob die Dokumentation die Daten ausreichend erklärt, und ob die Dateiformate standardisiert sind. Beispiele solcher Datenzeitschriften<sup>173</sup> stellen die Open-Access-Zeitschrift „Journal of Open Psychology Data“<sup>174</sup> für die Psychologie oder disziplinübergreifend „Data in Brief“<sup>175</sup> dar.

## 4. Repositorien

Repositorien sind Speicherorte, in denen Objekte archiviert, dokumentiert und publiziert werden können. Dabei handelt es sich um Speicherorte für Daten, die eine vom wissenschaftlichen Artikel zeitlich und räumlich unabhängige Veröffentlichung und Archivierung von Daten ermöglichen, i.d.R. für den Regeln der Guten wissenschaftlichen Praxis folgenden Zeitraum von mindestens 10 Jahren.

Je nach Repository können Daten, Datensätze, Experiment- und Auswertungsbeschreibungen, audiovisuelle Objekte wie Bild- und Videodateien, Modelle von Simulationen oder auch Software veröffentlicht werden. In manchen Repositorien können ganze Forschungsdatensätze in ihrer komplexen Form als Einheit abgelegt werden.

## 5. Repositorien für die Psychologie

Fachrepositorien bieten den Vorteil der höheren Sichtbarkeit in der jeweiligen Forschungscommunity sowie mehr Möglichkeiten, Daten entsprechend den Bedürfnissen und Standards des jeweiligen Fachgebiets zu dokumentieren.

Beispiele für Fachrepositorien in der Psychologie:

- PsychArchives<sup>176</sup>
- Forschungsdatenzentrum am ZPID<sup>177</sup>
- GenBank<sup>178</sup> für Genforschungsdaten
- openNeuro<sup>179</sup>

---

<sup>173</sup> Übersicht über Data Journals der HU Berlin, Zugriff am 27.04.2023, <https://www.cms.hu-berlin.de/de/dl/dataman/teilen/dokumentation/datajournal>.

<sup>174</sup> Ubiquity Press, „Journal of Open Psychology Data.“ Zugriff am 27.04.2023, <https://openpsychologydata.metajnl.com/>.

<sup>175</sup> Elsevier. „Data in Brief.“ Zugriff am 27.04.2023, <https://www.journals.elsevier.com/data-in-brief>.

<sup>176</sup> PsychArchives. Zugriff am 27.04.2023, <https://www.psycharchives.org/>.

<sup>177</sup> Forschungsdatenzentrum am ZPID, Zugriff am 27.04.2023, <https://rdc-psychology.org/de/homepage-deutsch>.

<sup>178</sup> GenBank. Zugriff am 27.04.2023, <http://www.ncbi.nlm.nih.gov/genbank>.

<sup>179</sup> OpenNeuro. Zugriff am 27.04.2023, <https://openneuro.org/>.

- brainlife<sup>180</sup>

Bei interdisziplinärer Forschung kann die Zuordnung der entstandenen Daten zu einem Fachgebiet schwierig sein. Hier können disziplinübergreifende Repositorien Abhilfe schaffen. Sie akzeptieren in der Regel sehr verschiedenartige Datenarten und bieten gute allgemeine Suchfunktionen. Es ist angeraten zu prüfen, ob eine Kuratierung der Daten oder andere Formen der inhaltlichen Qualitätskontrolle angeboten wird. Beispiele für fachübergreifende Repositorien:

- ZENODO<sup>181</sup> Digitale Daten aus allen Forschungsgebieten, Produkt des EU OpenAIRE<sup>182</sup>-Projekts (open source)
- DRYAD<sup>183</sup> Schwerpunkt: Lebenswissenschaften (open source)
- Figshare<sup>184</sup> Digitale Daten aus allen Forschungsgebieten (kommerziell)

Institutionelle Repositorien bieten eine Alternative, wenn kein geeignetes Fachrepository vorhanden ist. So können für Forschende unklare rechtliche Rahmenbedingungen in Bezug auf den Umgang mit Daten am Standort des externen Repositoriums für eine Publikation im Repository der eigenen Institution sprechen. Institutionelle Repositorien sind in der Regel für alle an der Forschungseinrichtung vertretenen Fachgebiete offen und kostenfrei nutzbar. Darüber hinaus bieten sie aber auch die Möglichkeit, lediglich die Metadaten von bereits anderweitig publizierten oder nicht publizierten Daten referenzieren zu lassen.

Beispiele für institutionelle Repositorien an Hochschulen:

- DepositOnce<sup>185</sup> der Technischen Universität Berlin
- edoc-Server<sup>186</sup> der Humboldt-Universität zu Berlin
- Refubium<sup>187</sup> der Freien Universität Berlin
- PUB<sup>188</sup> der Universität Bielefeld
- heiDATA<sup>189</sup> der Universität Heidelberg
- Edmond<sup>190</sup> der Max-Planck-Gesellschaft
- hilData<sup>191</sup> der Universität Hildesheim

## 6. Auswahl eines Repositoriums

Für die Suche nach einem geeigneten Repository dient das disziplinübergreifende Verzeichnis re3data<sup>192</sup>. Es umfasst derzeit über 3000 Einträge zu deutschen und internationalen Forschungsdaten-

---

<sup>180</sup> Brainlife. Zugriff am 27.04.2023, <https://brainlife.io/about/>.

<sup>181</sup> ZENODO. Zugriff am 27.04.2023, <https://zenodo.org>.

<sup>182</sup> OpenAIRE ist ein internationales Projekt mit Fachleuten aus Bibliotheken, nationalen e-Infrastruktur- und Datenexpert\*innen, IT- und Rechtswissenschaftler\*innen sowie offenen Stipendienorganisationen. Dieses Projekt zielt auf eine erhebliche Verbesserung der Auffindbarkeit und Wiederverwendbarkeit von Forschungspublikationen und -daten.

<sup>183</sup> DRYAD. Zugriff am 27.04.2023, <http://datadryad.org>.

<sup>184</sup> Figshare. Zugriff am 27.04.2023, <https://figshare.com>.

<sup>185</sup> DepositOnce. Zugriff am 27.04.2023, <https://depositonce.tu-berlin.de>.

<sup>186</sup> edoc-Server HU Berlin. Zugriff am 27.04.2023, <https://edoc.hu-berlin.de>.

<sup>187</sup> Refubium. Zugriff am 27.04.2023, <https://refubium.fu-berlin.de>.

<sup>188</sup> PUB. Zugriff am 27.04.2023, <https://pub.uni-bielefeld.de>.

<sup>189</sup> HeiDATA. Zugriff am 27.04.2023, <https://heidata.uni-heidelberg.de>.

<sup>190</sup> EDMOND der Max-Planck-Gesellschaft. Zugriff am 27.04.2023, <https://edmond.mpdl.mpg.de/>

<sup>191</sup> hilData, Zugriff am 27.04.2023, <https://hildata.uni-hildesheim.de/>.

<sup>192</sup> re3data.org - Registry of Research Data Repositories. Zugriff am 27.04.2023, <https://doi.org/10.17616/R3D>.

Repositorien (Stand: 27.04.2023)<sup>193</sup> und erlaubt die Suche nach Disziplinen, Art der Daten oder Land. Es ist auch möglich, sehr spezifisch nach Repositorien, bei denen beispielsweise der Daten-Upload kostenpflichtig ist oder bei denen die Datennutzung eingeschränkt ist, zu filtern. Weitere Suchmaschinen sind OpenROAR<sup>194</sup> oder ROAR<sup>195</sup>.

OpenAIRE empfiehlt folgende Priorisierung für die Wahl eines Repositoriums:<sup>196</sup>:

1. Externes, etabliertes fachspezifisches Repository
2. Institutionelles Repository
3. Kostenfreies generisches Repository
4. Suche via re3data.org

Kriterien für die Auswahl eines geeigneten Repositoriums<sup>197</sup>:

1. Zertifizierung, z. B. CoreTrustSeal (siehe Einheit 13 "Langzeitarchivierung", Punkt 4, S. 46)
2. (Automatisierte) Vergabe von persistenten Identifiern, z. B. Digital Object Identifier (siehe Einheit 15 „Publikation von Forschungsdaten“, Punkt 8a, S. 56)
3. Zugang zu Daten: offen, eingeschränkt oder unzugänglich
4. Nutzungsbedingungen für Datenautor\*innen sowie Nutzer\*innen, z. B. Entgelte, Embargofristen
5. Möglichkeit der Lizenzauswahl

## 7. Lizenzen

Die Nutzung publizierter Daten durch Dritte kann durch verschiedene Lizenzierungsmodelle geregelt werden. Etabliert haben sich z. B. die kostenfreien Lizenzen von Creative Commons (CC) für die offene Nutzung. Dabei stehen folgende Kennungen zur Verfügung:

**BY** = Namensnennung; Namensnennung des Urhebers erforderlich.

**NC** = **non-commercial** (nicht kommerziell); Nutzung unter kommerziellen Zwecken nicht erlaubt.

**ND** = **no derivatives** (keine Bearbeitung); Veränderung des Werkes nicht erlaubt

**SA** = **share alike** (Weitergabe unter gleichen Bedingungen); Nach der Bearbeitung muss das Werk unter der gleichen Lizenz weitergegeben werden.

Bei der Vergabe der Lizenz bestehen folgende Möglichkeiten:

- CC0 (Gemeinfreiheit/Public Domain)<sup>198</sup>
- CC BY (Namensnennung)<sup>199</sup>
- CC BY-ND (Namensnennung – Keine Bearbeitung)<sup>200</sup>

---

<sup>193</sup> Ebd.

<sup>194</sup> OpenROAR Suchmaschine. Zugriff am 27.04.2023, <https://v2.sherpa.ac.uk/opensoar/>.

<sup>195</sup> ROAR Suchmaschine. Zugriff am 27.04.2023, <http://roar.eprints.org/>.

<sup>196</sup> OpenAIRE. „How to select a repository?“ Zugriff am 27.04.2023, <https://www.openaire.eu/select-data-repository/>.

<sup>197</sup> re3data.org. „FAQ.“, Zugriff am 27.04.2023, <https://www.re3data.org/faq>.

<sup>198</sup> Creative Commons Public Domain Dedication (CC0). Zugriff am 27.04.2023, <https://creativecommons.org/publicdomain/zero/1.0/>.

<sup>199</sup> Creative Commons Namensnennung (CC BY). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by/4.0/>.

<sup>200</sup> Creative Commons Namensnennung - keine Bearbeitung (CC BY-ND). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by-nd/4.0/>.

- CC BY-NC (Namensnennung – Nicht-kommerziell)<sup>201</sup>
- CC BY-SA (Namensnennung – Weitergabe unter gleichen Bedingungen)<sup>202</sup>
- CC BY-NC-SA (Namensnennung – Nicht-kommerziell – Weitergabe unter gleichen Bedingungen)<sup>203</sup>
- CC BY-NC-ND (Namensnennung – Nicht-kommerziell – Keine Bearbeitung)<sup>204</sup>

Die Vergabe einer Creative Commons-Lizenz selbst hat keinen Einfluss auf das Urheberrecht, bietet jedoch eine einfache und standardisierte Lizenzierungsmöglichkeit für Forschungsdaten (siehe Einheit 5 "Rechtliche Aspekte", Punkt 4, S. 23). Dabei ist zu beachten, dass bei der Vergabe einer CC0-Lizenz der/die Urheber\*in auf alle urheberrechtlichen und verwandten Schutzrechte verzichtet, wohingegen alle weiteren CC-Lizenzen die Nutzungsrechte unterschiedlich stark einschränken.

Ein weiteres Lizenzierungsmodell bieten die Open Data Commons (ODC)<sup>205</sup>:

- Open Data Commons Open Database License (ODbL)<sup>206</sup> (entspricht der CC BY-SA Lizenz)
- Open Data Commons Attribution License (ODC-By)<sup>207</sup> (entspricht der CC BY Lizenz)
- Open Data Commons Public Domain Dedication and License (PDDL)<sup>208</sup> (entspricht der CC0 Lizenz)

Die ODC-Lizenzen eignen sich vor allem für Datenbankrechte, Patentrechte und Warenzeichen. Bei den Datenbank-Lizenzen können für die in entsprechend lizenzierten Datenbanken enthaltenen Inhalte jeweils unterschiedliche Bedingungen festgelegt werden.

Wurde im Rahmen eines Forschungsprojektes eine Software zur Analyse psychologischer Daten entwickelt, und möchte man diese Software zur freien Verfügung stellen (z. B. Fieldtrip<sup>209</sup>), benötigt man hingegen eine gesonderte Lizenz, um den besonderen Anforderungen dieses Formats gerecht zu werden (Einverständnis zur Installation, Veränderung, Ausführung, Zweck oder Ort der Nutzung, Anzahl der Benutzer etc.). Die Creative Commons-Lizenzen können dies nicht abdecken. Hierfür

---

<sup>201</sup> Creative Commons Namensnennung – Nicht-kommerziell (CC BY-NC). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by-nc/4.0/>.

<sup>202</sup> Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen (CC BY-SA). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by-sa/4.0/>.

<sup>203</sup> Creative Commons Namensnennung – Nicht-kommerziell – Weitergabe unter gleichen Bedingungen (CC BY-NC-SA). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

<sup>204</sup> Creative Commons Namensnennung – Nicht-kommerziell – keine Bearbeitung (CC BY-NC-ND). Zugriff am 27.04.2023, <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

<sup>205</sup> Open Knowledge Foundation. „Open Definition. Defining Open in Open Data, Open Content and Open Knowledge. Home.“ Zugriff am 27.04.2023, <https://opendatacommons.org/>.

<sup>206</sup> Open Knowledge Foundation. „Open Definition. Defining Open in Open Data, Open Content and Open Knowledge. Open Data Commons Open Database License (ODbL).“ Zugriff am 27.04.2023, <https://opendatacommons.org/licenses/odbl/>.

<sup>207</sup> Open Knowledge Foundation. „Open Definition. Defining Open in Open Data, Open Content and Open Knowledge. Open Data Commons Attribution License (ODC-By).“ Zugriff am 27.04.2023, <https://opendatacommons.org/licenses/by/>.

<sup>208</sup> Open Knowledge Foundation. „Open Definition. Defining Open in Open Data, Open Content and Open Knowledge. Open Data Commons Public Domain Dedication and License (PDDL).“ Zugriff am 27.04.2023, <https://opendatacommons.org/licenses/pddl/>.

<sup>209</sup> Lizenzen der Toolbox Fieldtrip, Zugriff am 27.04.2023, <https://www.gnu.org/licenses/gpl-3.0.html>.



empfiehlt sich eine der gängigen Software-Lizenzen wie z. B. die MIT-Lizenz<sup>210</sup>, GNU General Public License (GPL)<sup>211</sup>, GNU Lesser General Public License (LGPL)<sup>212</sup> oder die Apache-Lizenz<sup>213</sup>.

Es wird empfohlen, die Nutzung der Forschungsdaten so umfänglich wie möglich zu ermöglichen. Zu den Lizenzen, die die Idee von Open Data, und somit die Idee von Open Science (siehe Einheit 4 „Forschungsdatenmanagement“, Punkt 2c, S. 15) fördern, zählen CC0, CC BY, CC BY-SA, PDDL, ODC-By und ODbL<sup>214</sup>. Wird jedoch keine Lizenz vergeben, ist eine Nachnutzung ohne das Einverständnis der Urheber\*innen nicht möglich, d. h. eine fehlende Lizenz bedeutet nicht automatisch lizenzfrei im Sinne von z. B. CC0 (keine Lizenz ≠ nicht lizenzfrei).

## 8. Persistente Identifier

Um digitale Publikationen langfristig auffindbar sowie eindeutig referenzier- und zitierbar zu machen, sowie zur Verbesserung der Mensch-/Maschinenlesbarkeit, werden persistente Identifier (PIDs) eingesetzt. Jedes Objekt (z. B. eine Veröffentlichung) wird hierbei eineindeutig gekennzeichnet. Diese Kennzeichnung wird dann überall dort mit angegeben, wo auf dieses Objekt Bezug genommen wird, d. h. das Objekt ist mit diesem Identifier verlinkt, welcher somit ein Teil der Metadaten wird.

Etablierte ID-Systeme sind der „Digital Object Identifier“ (**DOI**) für die Kennzeichnung von Publikationen und Daten (DOI-Dienstleister: CrossRef<sup>215</sup> für Zeitschriftenartikel, DataCite<sup>216</sup> für Datensätze) und die „Open Researcher and Contributor ID“ (**ORCID**) für eine eindeutige Identifikation wissenschaftlicher Autor\*innen. Die Vergabe von persistenten Identifier kann u. U. mit Kosten verbunden sein, die in dem Datenmanagement mitbedacht werden sollten.

### a) Digital Object Identifier (DOI)

Sehr weit verbreitet ist der zitierfähige Digital Object Identifier (DOI). Die International DOI Foundation (IDF) garantiert einheitliche Standards und Workflows für den Einsatz von DOIs<sup>217</sup>, reguliert durch einen seit Mai 2012 gültigen ISO-Standard<sup>218</sup>.

Ein DOI

- ist eine einzigartige Folge von alphanumerischen Zeichen
- Zugelassene Zeichen sind: a – z, A – Z, 0 – 9, . (Punkt), - (Bindestrich), \_ (Unterstrich), : (Doppelpunkt) sowie / (Schrägstrich).

<sup>210</sup> Open-Source Initiative. „The MIT License.“ Zugriff am 27.04.2023, <https://opensource.org/licenses/mit-license.php>.

<sup>211</sup> GNU General Public License Version 3.0. Zugriff am 27.04.2023, <https://www.gnu.org/licenses/gpl-3.0.en.html>.

<sup>212</sup> GNU Lesser General Public License Version 3.0. Zugriff am 27.04.2023, <https://www.gnu.org/licenses/lgpl-3.0.html.en>.

<sup>213</sup> The Apache Software Foundation. „Apache License 2.0“. Zugriff am 27.04.2023, <https://www.apache.org/licenses/>.

<sup>214</sup> Open Knowledge Foundation. „Open Definition. Defining Open in Open Data, Open Content and Open Knowledge. Conformant Licenses.“ Zugriff am 27.04.2023, <https://opendefinition.org/licenses/>.

<sup>215</sup> CrossRef, Zugriff am 27.04.2023, <https://www.crossref.org/>.

<sup>216</sup> DataCite. Zugriff am 27.04.2023, <https://datacite.org>.

<sup>217</sup> Paskin, Norman. „Digital object identifier (DOI®) system.“ In *Encyclopedia of library and information sciences* herausgegeben von Marcia J. Bates und Mary Niles Maack. 3. Auflage, S. 1586–1592. Boca Raton, FL: CRC Press, 2010. <https://doi.org/10.1081/E-ELIS3-120044418>.

<sup>218</sup> International Organization for Standardization. ISO 26324:2012 Information and documentation — Digital object identifier system. 2012. Zugriff am 27.04.2023, <https://www.iso.org/standard/43506.html>.



- besteht aus zwei Teilen: einem Präfix, der die vergebende Organisation kennzeichnet, und einem Suffix, der das Objekt identifiziert

Beispiele für die Gestaltung des Suffixes:

- Ursprungs-DOI: 10.1234/abc123
- DOI einer neuen Version: 10.1234/abc123.1
- DOI eines Teils: 10.1234/abc123/2

Ein DOI ermöglicht es u. a. die Daten mit zugehörigen Publikationen eindeutig zu verknüpfen. Dadurch bleiben die Daten dauerhaft zitierbar. DOIs werden häufig durch Repositorien bzw. Institutionen wie Bibliotheken vergeben. DataCite verteilt z. B. sogenannte Nummernkreise an Institutionen (Registrar), die diese dann individuell vergeben (Registrant).

Im Gegensatz zu DOIs sind Handles<sup>219</sup> als globales Referenzsystem für große Datenmengen weniger persistent und nicht durch einen Standard abgesichert. Sie eignen sich daher vor allem für die Referenzierung von Daten vor dem Publizieren.

### b) Open Researcher and Contributor ID (ORCID)

Um zu gewährleisten, dass wissenschaftliche Arbeiten einem\*r Autor\*in unabhängig von Institutions- und Namenswechseln eindeutig zugeordnet werden, kann die sogenannte ORCID<sup>220</sup> verwendet werden. ORCID vereinfacht zudem die Referenzierung von Publikationslisten und Angaben zur Autor\*innenschaft, z. B. durch die Vermeidung einer falschen Referenzierung bei Namensgleichheit in derselben Fachdisziplin.

Die ORCID:

- ist ein (alpha-)numerischer 16-stelliger Code.
- Verschafft eine eindeutige wissenschaftliche Identität (auch über Namenswechsel, Tippfehler oder Namensgleichheit hinweg).
- wird von Zeitschriften, Forschungsförderern und Institutionen genutzt.
- wird von Forschenden selbst gepflegt.
- wird von einer gemeinnützigen Initiative betrieben.
- wächst kontinuierlich (Stand März 2023: 17.040.840 ORCIDs, davon aktiv im letzten Jahr: 9.401.122).
- hat Verbindung zu Web of Science, Scopus, Zenodo, DataCite, u. a.

## 9. Zugriffsklassen/Agreements

Um den Zugriff auf Forschungsdaten für eine Ansicht oder Nachnutzung zu regeln, können auch sogenannte Zugriffsklassen bestimmt werden, die vor allem in Repositorien Anwendung finden. In der aktuellen Fassung ihrer Forschungsdaten-Policy<sup>221</sup> definiert die DGPs 4 Klassen (Klasse 0-3), um diesen Zugriff zu regeln. Auf der niedrigsten Stufe (**Zugangsklasse 0**) besitzen veröffentlichte Daten keinerlei Zugangsbeschränkungen und werden unter CO- oder ODC-Lizenzen Nutzenden zur Verfügung gestellt. Die Abfrage von Informationen über die Identität des Nutzenden, deren\*dessen Zugehörigkeit zu einer Forschungseinrichtung und/oder die angedachte Nutzung der Daten ist möglich.

<sup>219</sup> Handle.net Registry, Zugriff am 27.04.2023, <https://www.handle.net/>.

<sup>220</sup> ORCID, Zugriff am 27.04.2023, <https://orcid.org/>.

<sup>221</sup> Ebd.

Die Daten der nächsthöheren **Zugangsklasse 1** werden ebenfalls grundsätzlich frei zur Verfügung gestellt; ihre Nutzung ist jedoch nur unter der expliziten Zustimmung der Nutzungsbedingungen des\*r Urhebers\*in (data sharer) möglich. Jegliche Zuwiderhandlung kann hier somit rechtlich verfolgt werden. Zusätzlich erfragen manche Repositorien hier ebenfalls Informationen über die Identität, deren\*dessen Zugehörigkeit und/oder die angedachte Nutzung.

Daten mit einer grundsätzlichen Zugangsbeschränkung haben mindestens die **Zugangsklasse 2**. Die Nutzungsbestimmungen werden oftmals mittels individualisierter Nutzungsverträge zwischen Urheber\*in und Nutzenden geregelt. Auch hier, wie auch in der nächst höheren Stufe, gilt: Zuwiderhandlungen der explizit zugestimmten Nutzungsbedingungen können rechtliche Folgen nach sich ziehen.

Die **Zugangsklasse 3** bildet die restriktivste Klasse; sie erlaubt den Datenzugriff nur unter der expliziten Zustimmung zu den Bedingungen des Datenschutzgesetzes (z. B. online mittels einer geeigneten gesicherten Verbindung zwischen Repositorium und Datennachnutzenden oder lokal). Hierbei gilt: Art und Umfang der Nutzung unterliegen dem Ermessen des Teilenden und werden durch die datenhaltende Einrichtung ermöglicht.

Mitunter definieren Repositorien Zugriffsklassen, die u. a. den von der DGPs empfohlenen Richtlinien entsprechen, z. B.:

- PsyArchives<sup>222</sup>: 4 Klassen (Klassen: 0/0+, 1, **Klassen 2 und 3 sind in Vorbereitung**)
- PsychData<sup>223</sup>: 3 Klassen (Klassen: 0,1,2)

Andere Repositorien definieren u. U. gar keine Klassen, regeln aber den Datenzugang z. B. per Vertrag oder Affiliation.

---

<sup>222</sup> Zugriffsklassen PsychArchives. Zugriff am 27.04.2023, [https://www.psycharchives.org/en/about#sharing\\_levels/](https://www.psycharchives.org/en/about#sharing_levels/).

<sup>223</sup> Zugriffsklassen PsychData. Zugriff am 27.04.2023, <https://rdc-psychology.org/de/datenzugang/>.

## Weiterführende Ressourcen

### Einheit 3

- Aust, P., Helbig, K., Schenk, U., Zielke, D., Rosenbaum, A. und Schulze, J. *Was sind Forschungsdaten?* Video. Berlin: Humboldt-Universität zu Berlin - Medien-Repositorium, 2016. <https://doi.org/10.18450/dataman/90>.
- Dumontier, M.: *Are we FAIR yet? Webinar*. RDA, 2018. Zugriff am 27.04.2023, <https://rd-alliance.org/webinar-are-we-fair-yet>.
- PARTHENOS. "Manage, improve and open up your research and data." Zugriff am 27.04.2023, <https://training.parthenos-project.eu/sample-page/manage-improve-and-open-up-your-research-and-data/>.

### Einheit 4

- Nosek, Brian A., Alter, G., Banks, George C., Borsboom, D., Bowman, Sara D., Breckler, Steven J., Buck, S., et al. *Promoting an open research culture*. Science 348 (2015). <https://doi.org/10.1126/science.aab2374>.
- Fachspezifische Policy der American Psychological Association: Berenson, K. R. (2018). Managing your research data and documentation. American Psychological Association.
- Fachspezifische Policy des Schweizer Nationalfond, Zugriff am 27.04.2023, <https://www.scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management>.

### Einheit 5

- Brettschneider, P., Biernacka, K., Böker, E., Danker, Sarah A., Jacob, J., Perry, A., Wiljes, C. und Wuttke, U. *Urheberrecht und Lizenzierung bei Forschungsdaten*. Zenodo, 2021. <https://doi.org/10.5281/zenodo.5243232>.
- Forschungsdaten.info. Rechte und Pflichten 2021. Zugriff am 27.04.2023, <https://www.forschungsdaten.info/themen/rechte-und-pflichten/>.
- Forschungsdaten.info. Rechte und Forschungsdaten - ein Überblick 2021. Zugriff am 27.04.2023, <https://www.forschungsdaten.info/themen/rechte-und-pflichten/recht-und-forschungsdaten-ein-ueberblick/>.
- Guibault, L. und Wiebe, A. (Ed.). *Safe to be open: Study on the protection of research data and recommendations for access and usage*. Göttingen: Universitätsverlag Göttingen, 2013. <https://doi.org/10.17875/gup2013-160>.
- Hartmann, T. *Zur urheberrechtlichen Schutzzfähigkeit von Forschungsdaten*. In „Law as a Service (LaaS) - Recht im Internet- und Cloud-Zeitalter“ herausgegeben von Jürgen Taeger. Band 1, S. 505–515. Edewecht: OIWR, Oldenburger Verlag für Wirtschaft, Informatik und Recht, 2013. <http://hdl.handle.net/11858/00-001M-0000-0014-1208-E>.
- Helbig, K. *Quiz FDM-Rechtsexpertise*. Zenodo, 2019. <https://doi.org/10.5281/zenodo.3479148>.
- Lauber-Rönsberg, A. *1.4 Rechtliche Aspekte des Forschungsdatenmanagements*. In Praxishandbuch Forschungsdatenmanagement. Markus Putnings, Heike Neuroth und Janna Neumann (Hrsg.), 89-114. Berlin, Boston: De Gruyter Saur, 2021. <https://doi.org/10.1515/9783110657807-005>.
- Lauber-Rönsberg, A., Krahn, P. und Baumann, P. *Kurzfassung: Gutachten zu den rechtlichen Rahmenbedingungen des Forschungsdatenmanagements im Rahmen des DataJus-Projekts*. Dresden: Technische Universität, 2018. Zugriff am 27.04.2023, [https://tu-dresden.de/gsw/phil/irget/jfbimd13/ressourcen/dateien/dateien/DataJus/DataJus\\_Zusammenfassung\\_Gutachten\\_12-07-18.pdf](https://tu-dresden.de/gsw/phil/irget/jfbimd13/ressourcen/dateien/dateien/DataJus/DataJus_Zusammenfassung_Gutachten_12-07-18.pdf).

- Lauber-Rönsberg, A., Krahn, P. und Baumann, P. *Forschungsdatenmanagement und Recht. Datenschutz-, Urheber- und Vertragsrecht*. Arbeitshefte der Arbeitsgemeinschaft für juristisches Bibliotheks- und Dokumentationswesen 28. Düns/Feldkirch: W. Neugebauer, 2021. ISBN 978–3- 85376–328-5.
- Leibniz Universität Hannover und Technische Informationsbibliothek. *FAQs zu rechtlichen Aspekten im Umgang mit Forschungsdaten*. Zenodo, 2018, <http://doi.org/10.5281/zenodo.1173546>.
- Meyermann, A. und Porzelt, M. *Datenschutzrechtliche Anforderungen in der empirischen Bildungsforschung - eine Handreichung. Version 2.0*. Forschungsdaten Bildung informiert Nr. 6 (2019). Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdb-informiert-nr-6\\_v2.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-6_v2.pdf).
- RatSWD - Rat für Sozial- und Wirtschaftsdaten. *Handreichung Datenschutz. 2. Vollständig überarbeitete Auflage*. RatSWD Output 8 (6). Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD), 2020. <https://doi.org/10.17620/02671.50>.
- Scheller, J. *Rechtliche Rahmenbedingungen der Verwendung von Videos in der Schul- und Unterrichtsforschung. Diskrepanzen zwischen Datenschutzrecht, Förder- und Genehmigungsaufgaben. Version 1.0*. Forschungsdaten Bildung informiert Nr. 5 (2017). Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdb-informiert-nr-5.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-5.pdf).
- Technische Universität Dresden. „Forschungsprojekt DataJus.“ Zugriff am 27.04.2023, <https://tu-dresden.de/gsw/jura/igetem/jfbimd13/forschung/forschungsprojekt-datajus>.
- Verbund Forschungsdaten Bildung. „Informierte Einwilligung.“ Zugriff am 27.04.2023, <https://www.forschungsdaten-bildung.de/einwilligung>.

## Einheit 6

- CESSDA ERIC: „Adapt your Data Management Plan.“ Zugriff am 27.04.2023, [https://www.cessda.eu/content/download/4302/48656/file/TTT\\_DO\\_DMPExpertGuide\\_v1.3.pdf](https://www.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.3.pdf).
- Helbig, K. *Hinweise und Checkliste zur Erstellung eines Datenmanagementplans*. Zenodo, 2015. Zugriff am 27.04.2023, <https://doi.org/10.5281/zenodo.33485>
- Helbig, K., Paul-Stüve, T. und Rex, J. *DMP-Toolguide 1.0*. Zenodo, 2021. Zugriff am 27.04.2023, <https://zenodo.org/record/4632308#.YuEKMt9CRaR>
- Helbig, K. und Aust, P. *Datenmanagementpläne für EU, DFG und BMBF. Präsentation*. Zenodo, 2015. Zugriff am 27.04.2023, <https://doi.org/10.5281/zenodo.33482>.
- Neuroth, H., Engelhardt, C., Klar, J., Ludwig, J. und Enke, H. *Aktives Forschungsdatenmanagement*. ABI Technik 38 (2018), S. 55–64. Zugriff am 27.04.2023, <https://doi.org/10.1515/abitech-2018-0008>.

## Einheit 7

- CESSDA Training Working Group: „CESSDA Data Management Expert Guide. Bergen: CESSDA ERIC, 2017-2018. Zugriff am 27.04.2023, <https://www.cessda.eu/DMGuide>.
- Corti, L., van den Eynden, V., Bishop, L. und Woollard, M. *Managing and Sharing Research Data: A Guide to Good Practice*. Los Angeles, CA: SAGE, 2014.
- Haenel, V. und Plenz, J. *Git: verteilte Versionsverwaltung für Code und Dokumente*. 2. Aufl. München: Open Source Press, 2014.
- Pilato, C.M., Collins-Sussman, B. and Fitzpatrick, Brian W. *Versionskontrolle mit Subversion*. 3. Aufl. Köln: O'Reilly, 2009.
- Software Carpentry. *Data - Episode 1 - Data Management*. Video. Youtube, 2011. Zugriff am 27.04.2023, <https://www.youtube.com/watch?v=3MEJ38BO6Mo>.

- Software Carpentry. "Data Management: Data Management." Zugriff am 27.04.2023, <https://v4.software-carpentry.org/data/mgmt.html>.
- Technology, Information and Learning Support (TILS). *TILS Document Naming Convention*. Queensland University of Technology, 2009. Zugriff am 27.04.2023, [https://www.data.cam.ac.uk/files/gdl\\_tilsdocnaming\\_v1\\_20090612.pdf](https://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf).

## Einheit 8

- CESSDA Training Working Group. „CESSDA Data Management Expert Guide." Bergen: CESSDA ERIC, 2017-2018. Zugriff am 27.04.2023, <https://www.cessda.eu/DMGuide>.
- forschungsdaten.info. „Datendokumentation." Zugriff am 27.04.2023, <https://www.forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>.
- Social Science Research Council. „*Principles of Documenting Data*." In: Managing Qualitative Social Science Data. An interactive online course. Zugriff am 27.04.2023, <https://managing-qualitative-data.org/modules/2/a/>.
- University of Helsinki. „Data Support." Zugriff am 27.04.2023, <https://www.helsinki.fi/en/research/services-researchers/data-support>.
- ZB MED (Hrsg.). Elektronische Laborbücher im Kontext von Forschungsdatenmanagement und guter wissenschaftlicher Praxis – ein Wegweiser für die Lebenswissenschaften. Köln, 2019. <https://doi.org/10.4126/FRL01-006415715>.

## Einheit 9

- Bundesamt für Sicherheit in der Informationstechnik. „Sicherer Umgang mit Passwörtern Schritt-für-Schritt erklärt." Zugriff am 27.04.2023, [https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/Umgang-mit-Passwoertern/umgang-mit-passwoertern\\_node.html](https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/Umgang-mit-Passwoertern/umgang-mit-passwoertern_node.html).
- Leibniz Universität Hannover. „Datenverschlüsselung." Zugriff am 27.04.2023, <https://www.luis.uni-hannover.de/de/services/it-sicherheit/praevention/datenverschlueselung/>

## Einheit 12

- Hanson, K., Surkis, A. und Yacobucci, K. *Data Sharing and Management Snafu in 3 Short Acts*. Video. Youtube, 2012. Zugriff am 27.04.2023, [https://youtu.be/66oNv\\_DJuPc](https://youtu.be/66oNv_DJuPc).

## Einheit 13

- Digital Curation Centre. *Five steps to decide what to keep: a checklist for appraising research data v.1*. Edinburgh: Digital Curation Centre, 2014. Zugriff am 27.04.2023, <https://www.dcc.ac.uk/guidance/how-guides/five-steps-decide-what-data-keep>.
- Ein zukünftiges plattform- und disziplinunabhängiges Datenformat, welches für wide- oder longformat-Daten entwickelt wird, Schnittstellen zu anderen Datenformaten (Excel, Texteditor usw.) und Software (Python, R, Matlab usw.) bereithält, sowie die Vergabe von Metadaten und persistentem Identifier ermöglicht. Zugriff am 27.04.2023, <https://www.konsortswd.de/konsortswd/das-konsortium/services/open-data-format/>

## Einheit 15

- Data Archiving and Networked Services. „The PID Forum." Zugriff am 27.04.2023, <https://pidforum.org/>.

- Ebel, T. und Meyermann, A. *Hinweise zur Anonymisierung von quantitativen Daten. Version 1.0.* Forschungsdaten Bildung informiert Nr. 3 (2015). Frankfurt a.M.: Forschungsdatenzentrum Bildung am DIPF, 2015. Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdb-informiert-nr-3.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-3.pdf).
- Forschungsdaten.info. "Persistente Identifikatoren." Zugriff am 27.04.2023, <https://www.forschungsdaten.info/themen/veroeffentlichen-und-archivieren/persistente-identifikatoren/>.
- International DOI Foundation. „DOI.“ Zugriff am 27.04.2023, <https://www.doi.org>.
- Kümmer, S., Lücke, S., Schulz, J., Spenger, M., & Weber, T. (2019, November 15). *DataCite Best Practice Guide (Version Version 1.0)*. Zenodo. <http://doi.org/10.5281/zenodo.3559800>.
- McKiernan, Erin C., Bourne, Philip E., Brown, C. Titus, Buck, S., Kenall, A., Lin, J., McDougall, D., et al. *Point of view: How open science helps researchers succeed.* eLife 5(2016). <https://doi.org/10.7554/eLife.16800>.
- Meyermann, A. und Porzelt, M. „Hinweise zur Anonymisierung von qualitativen Daten. Version 1.0.“ *Forschungsdaten Bildung informiert* Nr. 1 (2014). Frankfurt a.M.: Forschungsdatenzentrum Bildung am DIPF, 2014. Zugriff am 27.04.2023, [https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdb-informiert-nr-1.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-1.pdf).
- Pienta, Amy M., Alter, George C. und Lyle, Jared A. *The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data.* Ann Arbor, Mi: University of Michigan–Deep Blue, 2010. <http://hdl.handle.net/2027.42/78307>.
- Piwowar, Heather A und Vision, Todd J. *Data reuse and the open data citation advantage.* PeerJ 1 (2013), S. e175. <https://doi.org/10.7717/peerj.175>;
- Henneken, Edwin A. und Accomazzi, A. *Linking to data-effect on citation rates in astronomy.* arXiv (2011), <https://arxiv.org/abs/1111.3618>.
- Piwowar, Heather A., Day, Roger S. und Fridsma, Douglas B. *Sharing detailed research data is associated with increased citation rate.* PLoS ONE 2 (2007). <https://doi.org/10.1371/journal.pone.0000308>.
- Poisot, Timothée E., Mounce, R. und Gravel, D. *Moving toward a sustainable ecological science: don't let data go to waste!* Ideas in Ecology and Evolution 6 (2013). <https://doi.org/10.4033/iee.2013.6b.14.f>.
- Project FREYA. Zugriff am 27.04.2023, <https://www.project-freya.eu/en/about/mission>.
- Verbund FDB. „Informierte Einwilligung.“ Zugriff am 27.04.2023, <https://www.forschungsdaten-bildung.de/einwilligung>.
- Vierkant, P., Beucke, D., Deinzer, G., Hartmann, S., Herwig, S., Höhner, K., Müller, U. et al. *Autorenidentifikation anhand der Open Researcher and ContributorID (ORCID) - Positionspapier.* Berlin: Humboldt-Universität zu Berlin – edocServer, 2018. <https://doi.org/10.18452/19528>.
- Wilkinson, Mark D., Dumontier, M., Aalbersberg, IJsbrand J., Appleton, G., Axton, M., Baak, A., Blomberg, N., et al. *The FAIR Guiding Principles for scientific datamanagement and stewardship.* Scientific Data 3 (2016). <https://doi.org/10.1038/sdata.2016.18>.