

# Forschungsdatenmanagement an der Fakultät Psychologie

Alexander Strobel, Jens Strehle, Denise Dörfel & Kontaktstelle Forschungsdaten

2022/05/17

## Contents

1 Ausgangspunkt . . . . .	2
2 FDM-Kernprozesse . . . . .	4
2.1 Datenmanagementplan . . . . .	4
2.2 Datendokumentation . . . . .	5
2.3 Datenarchivierung . . . . .	6
2.4 Datenpublikation . . . . .	7
2.5 Datenschutz . . . . .	8
3 Unterstützende Prozesse für das FDM . . . . .	9
3.1 Datenerhebung . . . . .	9
3.2 Datenorganisation . . . . .	10
3.3 Wissensmanagement Onboarding . . . . .	11
3.4 Wissensmanagement Offboarding . . . . .	11
4 Tools für das Forschungsdatenmanagement . . . . .	12
4.1 Tools der TU Dresden . . . . .	12
4.1.1 ZIH . . . . .	12
4.1.2 Kontaktstelle Forschungsdaten . . . . .	12
4.1.3 ORSEE/PaMS . . . . .	12
4.1.4 REDCap . . . . .	13
4.1.5 Gruppenlaufwerk/Cloudstore/SharePoint . . . . .	14
4.1.6 OpARA . . . . .	15
4.2 Externe Tools . . . . .	15
4.2.1 Leibniz Psychology DataWiz . . . . .	15
4.2.2 Open Science Framework (OSF) . . . . .	15
4.2.3 GitHub . . . . .	16
4.3 Individuelle Tools . . . . .	16
4.3.1 Standard Operating Procedures . . . . .	17
4.3.2 Lab Manuals . . . . .	17
5 Ressourcen und Vorlagen . . . . .	17

# 1 Ausgangspunkt

*“Die langfristige Sicherung und Bereitstellung der Forschungsdaten leistet einen Beitrag zur Nachvollziehbarkeit und Qualität der wissenschaftlichen Arbeit und eröffnet wichtige Anschlussmöglichkeiten für die weitere Forschung. (DFG-Leitlinien zum Umgang mit Forschungsdaten, S. 1)*

Ausgehend von den 2015 formulierten [Leitlinien der Deutschen Forschungsgemeinschaft \(DFG\)](#) hat die Technische Universität Dresden (TU Dresden) im Jahr 2018 [Leitlinien für den Umgang mit Forschungsdaten an der TU Dresden](#) beschlossen. Diese sind – ungeachtet der für die Psychologie konkretisierten [Empfehlungen der Deutschen Gesellschaft für Psychologie \(DGPs\)](#) für den Umgang mit Forschungsdaten aus dem Jahr 2016 – für die Fakultät Psychologie bindend. Sie lauten wie folgt (Version vom 13.06.2018, jeweils mit einem Schlagwort versehen):

## *Leitlinien*

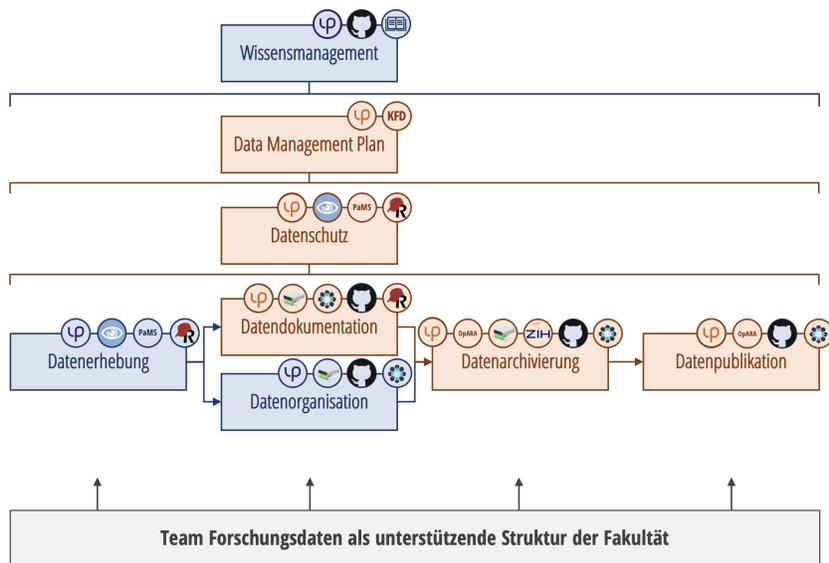
- 1. Forschungsdatenmanagement (FDM):** Verantwortungsvolle Forschung geht einher mit einem bewussten und sorgfältigen Umgang mit Forschungsdaten über deren gesamten Lebenszyklus, wie es die DFG in ihren „Leitlinien zum Umgang mit Forschungsdaten“ formuliert. Forschungsdatenmanagement umfasst die Planung, Gewinnung, Verarbeitung, Dokumentation, Publikation und langfristige Aufbewahrung der Daten.
- 2. Verantwortlichkeit:** Die Verantwortung für das Forschungsdatenmanagement liegt bei den Projektleiterinnen und Projektleitern sowie eigenverantwortlich Forschenden. Die Struktureinheiten (Bereiche, Fakultäten oder Institute) können für ihren Verantwortungsbereich verbindliche Regeln festlegen.
- 3. Data Management Plan:** Für Projekte mit Forschungsdaten soll, basierend auf allgemeinen und fachspezifischen Richtlinien (z.B. der Fördermittelgeber oder Fachgesellschaften), möglichst schon bei Projektkonzeption bzw. -antragstellung ein „Data Management Plan“ erstellt werden, der den Umgang mit Forschungsdaten dokumentiert.
- 4. Kontaktstelle Forschungsdaten:** Die TU Dresden unterstützt die Wissenschaftlerinnen und Wissenschaftler beim Forschungsdatenmanagement durch eine „Kontaktstelle Forschungsdaten“, die gemeinsam vom ZIH und der SLUB betrieben wird. Sie bietet Beratung bei der Entwicklung von Konzepten und Unterstützung durch Werkzeuge des Forschungsdatenmanagements an. Dafür ist eine frühzeitige Kontaktaufnahme vor oder zu Projektbeginn erforderlich.
- 5. Forschungsdateninfrastruktur:** Die TU Dresden verfügt über eine Grundausstattung an Forschungsdateninfrastruktur und stellt damit eine angemessene Aufbewahrung und die technische Verfügbarkeit von digitalen Forschungsdaten sicher.
- 6. Open Access:** Die TU Dresden fördert und unterstützt den freien Zugang zu Forschungsdaten (Open Access) – einerseits im Sinne guter wissenschaftlicher Praxis, andererseits wegen der großen Bedeutung für die optimale Verwertung von Forschungsergebnissen und der Qualität der Forschung. Die Entscheidung für eine Veröffentlichung und deren rechtliche Bedingungen liegt in der Eigenverantwortung der Wissenschaftlerinnen und Wissenschaftler. Der Schutz persönlicher Daten, ethische und urheberrechtliche Belange, Verpflichtungen gegenüber Dritten und Verwertungsinteressen sollen ausgewogen berücksichtigt werden.
- 7. Datendokumentation:** Für die Bereitstellung und Nachnutzung ist es notwendig, die Daten, ihren Entstehungskontext und die benutzten Werkzeuge zu dokumentieren.
- 8. Datenarchivierung:** Forschungsdaten, die Grundlage einer Publikation bilden, sollen langfristig in einem geeigneten vertrauenswürdigen Datenarchiv bzw. Repositorium archiviert und/oder geeignet veröffentlicht werden. Sie sind Teil der wissenschaftlichen Reputation der Forschenden.
- 9. Weiterbildung:** Zur nachhaltigen Etablierung und Weiterentwicklung eines qualitätsvollen Forschungsdatenmanagements sollen dessen Prinzipien und Methoden in Lehre und Weiterbildung verankert werden.

Gemäß *Leitlinie 1. Forschungsdatenmanagement* ist es im Sinne eines “bewussten und sorgfältigen Umgang[s] mit Forschungsdaten über deren gesamten Lebenszyklus” erforderlich, für jedes Forschungsprojekt Routinen für die “Planung, Gewinnung, Verarbeitung, Dokumentation, Publikation und langfristige Aufbewahrung der [erhobenen] Daten” vorzuhalten.

Gemäß *Leitlinie 2. Verantwortlichkeit* liegt die Verantwortung dafür bei den eigenverantwortlich Forschenden. Struktureinheiten können allerdings verbindliche Regeln für das Forschungsdatenmanagement festlegen. Zu diesem Zweck wurde vom Fakultätsrat der Fakultät Psychologie die Kommission Datenhaltung eingesetzt (bestehend aus Denise Dörfel, Lars Pieper, Stefan Scherbaum, Alexander Strobel und Gesine Wieder). In Zusammenarbeit mit der [Kontaktstelle Forschungsdaten](#), die Gespräche mit den Professuren der Fakultät zu deren bereits bestehenden FDM-Routinen und FDM-Erfordernissen führte, wurde ein erstes Konzept dieser Richtlinien erarbeitet und vom Kernteam Forschungsdaten (bestehend aus Alexander Strobel und Jens Strehle als Nachfolge von Lars Pieper) in Zusammenarbeit mit der Kontaktstelle Forschungsdaten und Denise Dörfel weiter ausgearbeitet.

Mit den vorliegenden Leitlinien wird ein Mittelweg beschritten: Alle Elemente eines den FDM-Leitlinien (sowie im Übrigen auch der [Satzung zur Sicherung guter wissenschaftlicher Praxis](#)) der TU Dresden entsprechenden Forschungsdatenmanagements werden adressiert bzw. konkretisiert, und es werden beispielhaft Vorschläge zu deren Umsetzung unterbreitet. Gleichwohl verbleibt es im Ermessen und der Verantwortung der eigenverantwortlich Forschenden, ihr Forschungsdatenmanagement gemäß der Leitlinien der TU Dresden und der hier unterbreiteten Vorschläge auszugestalten. Die folgende Abbildung fasst alle Elemente und einige der empfohlenen Tools des Forschungsdatenmanagements an der Fakultät zusammen. Im weiteren Verlauf werden die einzelnen Elemente einschließlich der jeweiligen Grundlagen und vorgeschlagenen Umsetzungsmöglichkeiten detaillierter erläutert. Das Dokument stellt dabei ein *Living Document* dar, d.h., alle Fakultätsmitglieder sind eingeladen, an seiner Fortentwicklung mitzuwirken. Wenden Sie sich dazu einfach per Mail an [Alexander Strobel](#).

**Übersicht über erforderliche Elemente des Forschungsdatenmanagements an der Fakultät Psychologie**  
einschließlich empfohlener Tools



**! Legende**

- Kernprozesse
- Unterstützende Prozesse
- Tools
- Tools der TU Dresden**
- KFD Kontaktstelle Forschungsdaten
- ORSEE
- PaMS
- REDCap
- ZIH
- Gruppenlaufwerk
- OpARA
- Externe Tools**
- Leibniz Psychology DataWiz
- Open Science Foundation (OSF)
- GitHub
- Individuelle Tools**
- Labmanuals, Checklisten etc.

Figure 1: Elemente des Forschungsdatenmanagements an der Fakultät Psychologie

## 2 FDM-Kernprozesse

FDM-Kernprozesse sind solche Prozesse, die von den Leitlinien der DFG und der TU Dresden gefordert und auch in der Satzung zur Sicherung guter wissenschaftlicher Praxis der TU Dresden sowie den Empfehlungen der DGPs thematisiert werden. Sie adressieren folgende Aspekte:

- [2.1 Datenmanagementplan](#)
- [2.2 Datendokumentation](#)
- [2.3 Datenarchivierung](#)
- [2.4 Datenpublikation](#)
- [2.5 Datenschutz](#)

### 2.1 Datenmanagementplan

#### Grundlage

- **DFG-FDM-Leitlinie 1.** „In einem Antrag sollten die Antragstellenden daher ausführen, welche Forschungsdaten im Verlauf eines wissenschaftlichen Forschungsvorhabens entstehen, erzeugt oder ausgewertet werden [...]. Die einschlägigen Erläuterungen müssen Informationen zu Datentypen, falls vorhanden zu disziplinspezifischen Standards und zur Wahl geeigneter Repositorien enthalten [...]. Zusätzlich werden Angaben zu ggf. betroffenen Rechten Dritter sowie erste Planungen zum zeitlichen Rahmen der Datenveröffentlichung erbeten.“
- **TUD-FDM-Leitlinie 3.** „Für Projekte mit Forschungsdaten soll, basierend auf allgemeinen und fachspezifischen Richtlinien (z.B. der Fördermittelgeber oder Fachgesellschaften), möglichst schon bei Projektkonzeption bzw. -antragstellung ein „Data Management Plan“ erstellt werden, der den Umgang mit Forschungsdaten dokumentiert.“

#### Umsetzung

Relevante Aspekte, auf die in einem Data Management Plan eingegangen werden sollte, finden sich hier:

- [DFG-Checkliste zum Umgang mit Forschungsdaten](#)

Die Kontaktstelle Forschungsdaten bietet als kostenlose Dienstleistung die Bereitstellung von Textbausteinen für einen Data Management Plan für Förderanträge an, siehe hier:

- vgl. [Checkliste/Formular des Data Management Plan \(DMP\) Service der Kontaktstelle Forschungsdaten](#)

[Leibniz Psychology](#) bietet mit dem Tool *DataWiz* „ein Assistenzsystem, das Wissenschaftlern und Wissenschaftlerinnen helfen soll, ihre Daten von Forschungsbeginn an fachgerecht aufzubereiten und zu dokumentieren. Das Tool hilft dabei, einen Data Management Plan zu erstellen, eigene Forschungsdaten zu verwalten, und es ermöglicht auch, Daten im Team zu teilen und später der Fachgemeinschaft zur Nachnutzung zur Verfügung zu stellen. DataWiz ist eine kostenfreie Open-Science-Webanwendung.“ (Text von der entsprechenden Webseite entnommen). Eigene Erfahrungen mit *DataWiz* stehen allerdings noch aus.

Obwohl Data Management Plans vorrangig für Förderanträge zu erstellen sind, wird empfohlen, für jedes Forschungsprojekt einen solchen Plan zu erstellen und für alle Projektbeteiligten einsehbar in einem Projektordner zu hinterlegen (z.B. auf einem Gruppenlaufwerk).

## 2.2 Datendokumentation

### Grundlage

- **TUD-FDM-Leitlinie 7.** „Für die Bereitstellung und Nachnutzung ist es notwendig, die Daten, ihren Entstehungskontext und die benutzten Werkzeuge zu dokumentieren.“
- **TUD-GWP-Satzung § 7 (3).** „Grundsätzlich sind alle für das Zustandekommen von Forschungsergebnissen relevanten Informationen zu dokumentieren. Dazu gehört auch, soweit dies möglich und zumutbar ist, die den Ergebnissen zugrundeliegenden Forschungsdaten.“

### Umsetzung

Für jedes Forschungsprojekt existiert ein Projektordner (z.B. auf einem Gruppenlaufwerk, in einem GitHub/GitLab Repository etc.) mit idealerweise für die betreffende Arbeitsgruppe vereinheitlichter Ordnerstruktur. Dieser Ordner enthält die im Projekt gesammelten und verarbeiteten Daten sowie alle Informationen, die zur Nutzung und ggf. Nachnutzung der Daten bzw. zur Interpretation oder auch Reproduktion von im Projekt erzielten Ergebnissen erforderlich sind. Dabei sind vor allem folgende Punkte zu beachten:

- **Nachvollziehbarkeit:** Nachvollziehbare Dokumentation ist eine Voraussetzung für die Nachnutzung von Daten.
- **Umfang:** Der Umfang der Dokumentation richtet sich nach der Zielgruppe: Dokumentationen werden nicht für fachfremdes Personal geschrieben, sondern für Kolleg:innen mit ähnlichem Ausbildungsniveau.
- **Readme-Dateien:** Im Idealfall befindet sich in jedem Ordner eine kurze Text-Datei (Readme.txt; für GitHub Repositories eine Readme.md), die den Inhalt des Ordners beschreibt (auch als „Leseanleitungen“ oder „Ordner-Steckbrief“ bezeichnet). Die Verwendung von Word-Dateien (Readme.docx) sollte eher vermieden werden.
- **Struktur:** Allgemeine Dokumentationen sollten an zentraler Stelle (z.B. oberste Hierarchie-Ebene auf dem Gruppenlaufwerk) abgelegt werden; sämtliche Dokumente sollten eindeutig und einheitlich benannt werden; bei vielen Dokumenten kann es sich lohnen, eine Übersichtsliste zu führen, die von einer verantwortlichen Person aktuell zu halten.
- **Relevante Dateien**
  - *Manuale/Protokolle:* detaillierte Beschreibung von Forschungsabläufen
  - *Skripte:* z.B. SPSS/SAS/Stata, R, Matlab, Python Code etc.; anhand derer der “Weg” von den Primärdaten zu den letztlich für statistische Analysen genutzten Daten ebenso nachvollzogen werden kann wie die statistischen Analysen selbst. Relevante Informationen sind im Header und in Kommentaren in einem Auflösungsgrad dokumentieren, der die Handhabung der Skripte durch Dritte möglich macht.
  - *Daten:* die eigentlichen Forschungsdaten, d.h. sowohl die Primärdaten (= die unmittelbarerhobenen und noch unverarbeiteten Daten) wie auch die verarbeiteten und ggf. aggregierten Daten. Idealerweise liegen die Daten in einem nicht-proprietären Format vor (also z.B. als CSV- anstatt als Excel- oder SPSS-Datei). Viele proprietäre Formate können allerdings von offenen Anwendungen wie R gelesen werden, was bei z.B. SPSS-Datenfiles u.U. hilfreicher ist als eine CSV-Datei, da die Attribute der einzelnen Variablen (wie Variable oder Value Labels) ebenso eingelesen werden können und ein Code Book (s. nächster Punkt) bei hinreichender Beschreibung der Daten ggf. sogar ersetzen können
  - *Code Books:* Erklärungen zur Bedeutung verwendeter Variablen und Abkürzungen, der Kodierung der Variablen etc. Bei Nutzung von REDCap für die Erhebung und das Management von Forschungsdaten kann ein Code Book automatisiert erstellt werden. Code Books sind ein *unerlässlicher* Bestandteil eines angemessenen Forschungsdatenmanagements!

Table 1: Beispiel für ein Code Book

Variable	Description	Range/Units/Values
id	identifier	1-138
age	age	18-38
gen	gender	1 = male, 2 = female, 3 = diverse
edu	education	0 = no University Entrance Diploma, 1 = UED
job	job	0 = training/studies, 1 = unemployed, 2 = employed
bfi2_01	Big Five Inventory 2, Item 1	-2 = stimme überhaupt nicht zu (fully disagree) ... +2 = stimme voll und ganz zu (fully agree)

- **Versionskontrolle:** Modifikationen an Dateien sollten durch (digitale) Dokumentation erfasst werden; Ziel ist die Nachvollziehbarkeit von Änderungen in Dateien. Bei Nutzung von GitHub als Projektplattform ist die Versionskontrolle bereits "eingebaut", andere Plattformen bieten diese tw. ebenfalls an.

## 2.3 Datenarchivierung

### Grundlage

- **DFG-FDM-Leitlinie 3.** „Den Regeln der Guten Wissenschaftlichen Praxis folgend sollen Forschungsdaten in der eigenen Einrichtung oder in einer fachlich einschlägigen, überregionalen Infrastruktur für mindestens 10 Jahre archiviert werden.“
- **TUD-FDM-Leitlinie 8.** „Forschungsdaten, die Grundlage einer Publikation bilden, sollen langfristig in einem geeigneten vertrauenswürdigen Datenarchiv bzw. Repositorium archiviert und/oder geeignet veröffentlicht werden. Sie sind Teil der wissenschaftlichen Reputation der Forschenden.“
- **TUD-GWP-Satzung §7 (5).** „In der Regel müssen die Primärdaten für zehn Jahre in der Einrichtung zugänglich bleiben. [...] Sind in den Primärdaten personenbeziehbare Daten enthalten [...] so sind die Merkmale, mit deren Hilfe ein Personenbezug hergestellt werden kann, getrennt zu speichern; die Merkmale sind zu löschen, sobald der Forschungszweck dies zulässt. [...] Aus Gründen der Nachvollziehbarkeit, Anschlussfähigkeit der Forschung und Nachnutzbarkeit hinterlegen Wissenschaftler und Wissenschaftlerinnen, wenn immer möglich, die der Publikation zugrundeliegenden Forschungsdaten und zentralen Materialien – den FAIR-Prinzipien („Findable, Accessible, Interoperable, Reusable“) folgend – zugänglich in anerkannten Archiven und Repositorien.“
- **TUD-FDM-Leitlinie 5.** „Die TU Dresden verfügt über eine Grundausstattung an Forschungsdateninfrastruktur und stellt damit eine angemessene Aufbewahrung und die technische Verfügbarkeit von digitalen Forschungsdaten sicher“

### Umsetzung

- **Zeitpunkt** für Archivierung von Forschungsdaten:
  - bei Erreichen vorab definierter Projekt-Meilensteine, z.B.
    - \* Archivierung der Rohdaten (z.B. Fragebogenangaben, Audiodateien von Interviews)
    - \* Archivierung vorverarbeiteter Daten (z.B. Fragebogen-Summenscores, Interviewtranskripte)
    - \* Archivierung der Ergebnisse von quantitativen oder qualitativen Datenanalysen (z.B. Korrelationsanalysen über Fragebogendaten, Ergebnisse von Inhaltsanalysen über Interviewdaten)
  - nach Projektabschluss bzw. zusammen mit der Veröffentlichung der Forschungsergebnisse
  - in beiden Fällen werden die Daten i.d.R. nicht mehr verändert

- **Vorgehen** bei der Datenarchivierung:
  - Datenselektion: Welche erhobenen Daten sind für die Nachvollziehbarkeit von in einem Projekt gewonnenen Ergebnissen bzw. für die Nachnutzung erforderlich, welche müssen (s.u. [Datenschutz](#)) oder können entfernt werden (z.B. Daten, die von Softwareanwendungen automatisch erhoben wurden, deren Berücksichtigung aber voraussichtlich bei einer etwaigen Nachnutzung der Daten keinem wissenschaftlichen Zweck dienlich sein wird, etwa der Browsertyp, in dem ein Online-Survey bearbeitet wurde)
  - Datenaufbereitung: analog zum unter 2.2 beschriebenen Vorgehen
  - bei personenbezogenen Daten zusätzlich Anonymisierung oder Pseudonymisierung der Forschungsdaten; Unterstützung bietet die [Unabhängige Treuhandstelle der TU Dresden](#) (aktuell im Aufbau ist auch die neue Datentreuhandstelle der TU Dresden: „DD Trust“)
  - Verwendung geeigneter Dateiformate: im Falle von proprietären Dateiformaten sollten die Originaldaten zusätzlich in ein offenes Dateiformat exportiert und zusammen archiviert werden (z.B. \*.sav + \*.csv, siehe auch oben 2.2.)
  - Auswahl einer geeigneten Archivlösung: z.B. [Archivierung am ZIH](#) der TU Dresden (Zwischenarchiv für eine Projektlaufzeit oder OpARA für die langfristige Archivierung)

## 2.4 Datenpublikation

### Grundlage

- **DFG-FDM-Leitlinie 2.** „Soweit einer Veröffentlichung der Forschungsdaten aus einem DFG-geförderten Projekt Rechte Dritter (insbesondere Datenschutz, Urheberrecht) nicht entgegenstehen, sollten Forschungsdaten so zeitnah wie möglich verfügbar gemacht werden.“
- **TUD-FDM-Leitlinie 6.** „Die TU Dresden fördert und unterstützt den freien Zugang zu Forschungsdaten (Open Access) – einerseits im Sinne guter wissenschaftlicher Praxis, andererseits wegen der großen Bedeutung für die optimale Verwertung von Forschungsergebnissen und der Qualität der Forschung. Die Entscheidung für eine Veröffentlichung und deren rechtliche Bedingungen liegt in der Eigenverantwortung der Wissenschaftlerinnen und Wissenschaftler. Der Schutz persönlicher Daten, ethische und urheberrechtliche Belange, Verpflichtungen gegenüber Dritten und Verwertungsinteressen sollen ausgewogen berücksichtigt werden.“
- **TUD-FDM-Leitlinie 5.** „Die TU Dresden verfügt über eine Grundausstattung an Forschungsdateninfrastruktur und stellt damit eine angemessene Aufbewahrung und die technische Verfügbarkeit von digitalen Forschungsdaten sicher“

### Umsetzung

- **Zeitpunkt** für die Veröffentlichung von Forschungsdaten:
  - Zu jeder *eingereichten* Publikation werden idealerweise die relevanten Forschungsdaten bereitgestellt (und idealerweise auch die Analyseskripte und die verwendeten Materialien zur Datenerhebung, z.B. Fragebögen, Stimuli, Experimentalskripte, sofern sie keinem Copyright unterliegen). Dies erlaubt Transparenz und Nachvollziehbarkeit schon im Reviewprozess. Für double-blind Review-Verfahren bietet OSF etwa eine Möglichkeit anonymisierte Projekte, siehe [Create a View-only Link for a Project](#)
  - Mindestens aber werden zu jeder *veröffentlichten* Publikation die relevanten Forschungsdaten und Materialien bereitgestellt, oder es wird im Manuskript erläutert, warum das nicht möglich ist. Immer mehr Zeitschriften verlangen ohnehin ein sog. *Data Availability Statement*.
- **Vorgehen** bei der Datenpublikation:

- Datenselektion und Datenaufbereitung erfolgt analog zum Vorgehen im Abschnitt „Datenarchivierung“
- Lizenzierung der veröffentlichten Forschungsdaten: Die meisten Forschungsdaten werden unter Verwendung einer [Creative Commons Lizenz](#) veröffentlicht. Dabei sollte die zu vergebende Lizenz stets genau erwogen werden: Die offenste Form der Creative Commons (CC) Lizenzierung ist die CC 0 Lizenz, bei der nachnutzende Personen vollkommen frei und ohne Namensnennung der Urheberschaft über die Daten verfügen können; eine ebenfalls sehr offene Lizenz ist CC 4.0 BY, die zumindest die Namensnennung der Urheberschaft (BY = Attribution) verlangt; eine vergleichsweise gängige CC Lizenz ist CC 4.0 BY SA NC, bei der von Nachnutzenden neben der Angabe der Urheberschaft auch verlangt wird, die Daten unter gleichen Bedingungen zu lizenzieren (SA = Share Alike) und nur für nicht-kommerzielle Zwecke (NC = Non-Commercial) zu verwenden. Bei Verwendung von Daten Dritter sollten die Daten nur nach Rücksprache mit den Dateneigentümern und unter Verwendung einer geeigneten Lizenz veröffentlicht werden (vgl. Abschnitt „Datenschutz“ weiter unten, siehe auch [Creative Commons Licence Chooser](#)).
- Im Zweifelsfall kann Rücksprache mit der [Professur für Urheberrecht und geistiges Eigentum](#) der TU Dresden gehalten werden.
- Auswahl eines geeigneten Datenrepositoriums
  - \* Verzeichnis von Forschungsdatenrepositorien: [re3data](#)
  - \* Datenrepositorium der TU Dresden: [OpARA](#), ist derzeit noch etwas umständlich in der Handhabung, zeichnet sich aber dadurch aus, dass 1. alle Daten auf Servern der TU Dresden liegen und 2. eine unabhängige Überprüfung der Daten und ihrer Dokumentation erfolgt, d.h., die Daten werden kuratiert.
  - \* Datenrepositorien mit Fokus auf oder häufiger Nutzung in der Psychologie:
    - [Open Science Foundation \(OSF\)](#): eines der in der Psychologie am häufigsten genutzten Repositorien, zeichnet sich durch vergleichsweise einfache Handhabung auf, kommt aber bei großen Datenmengen schnell an seine Grenzen, wird nicht kuratiert, man kann also auch “Datenmüll” archivieren, als Server können deutsche Server gewählt werden, der Sitz der Organisation ist allerdings in den USA
    - [GitHub](#): zunehmend häufiger auch in der Psychologie genutzt, Vorteil ist die Versionkontrolle bei GitHub, OSF-Projekte können auch mit GitHub-Repositorien verlinkt werden, sie werden dann mit jeder Aktualisierung des GitHub-Repositoriums ebenfalls aktualisiert, siehe [Connect GitHub to a Project](#)
    - [PsychArchives](#): wird von Leibniz Psychology vorgehalten, es existieren unsererseits allerdings noch keine Erfahrungen damit. Ein Vorteil dieser Plattform könnte die Abbildung des gesamten Forschungszyklus sein, wenn man das möchte: [Planung und Analyse von Studien](#), [Präregistrierung](#), teilweise [Datenerhebung](#), [Datenarchivierung](#) mit [Datenmanagementsystem](#), [Publikation](#)

## 2.5 Datenschutz

### Grundlage

- **TUD-FDM-Leitlinie 6.** “Die TU Dresden fördert und unterstützt den freien Zugang zu Forschungsdaten (Open Access) – einerseits im Sinne guter wissenschaftlicher Praxis, andererseits wegen der großen Bedeutung für die optimale Verwertung von Forschungsergebnissen und der Qualität der Forschung. Die Entscheidung für eine Veröffentlichung und deren rechtliche Bedingungen liegt in der Eigenverantwortung der Wissenschaftlerinnen und Wissenschaftler. Der Schutz persönlicher Daten, ethische und urheberrechtliche Belange, Verpflichtungen gegenüber Dritten und Verwertungsinteressen sollen ausgewogen berücksichtigt werden.”

- **TUD-GWP-Satzung GWP §7 (6).** “Sind in den Primärdaten personenbeziehbare Daten enthalten - Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person – so sind die Merkmale, mit deren Hilfe ein Personenbezug hergestellt werden kann, getrennt zu speichern; die Merkmale sind zu löschen, sobald der Forschungszweck dies zulässt. Insofern sind diese Daten aus den zu archivierenden Primärdaten zu entfernen.”
- **TUD-GWP-Satzung §8 (10).** “Soll die Veröffentlichung personenbezogener Daten enthalten – Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person – so ist dies nur zulässig, wenn die hiervon Betroffenen eingewilligt haben oder dies für die Darstellung von Forschungsergebnissen über Ereignisse der Zeitgeschichte unerlässlich ist und überwiegende schutzwürdige Interessen der Betroffenen nicht entgegenstehen.”

## Umsetzung

- **Anonymisierung bzw. Pseudonymisierung** von Daten als essenzieller Schritt des Datenschutzes
  - erfolgt, sobald es keinen wissenschaftlichen Grund mehr für die Zuordnung der Forschungsdaten zu personenbezogenen Daten gibt (also ggf. schon vor der Datenarchivierung, in jedem Fall aber vor der Datenpublikation); Nachteil: eine Löschung von Forschungsdaten einzelner Personen aufgrund einer Rücknahme der Einwilligung in die Datennutzung ist dann nicht mehr möglich, darauf muss in der die Einwilligungserklärung hingewiesen werden!
- **Datennutzungsverträge**
  - Die Überlassung von Forschungsdaten an Dritte sollte durch einen Datennutzungsvertrag abgesichert werden
  - Der [Datenschutzbeauftragte der TU-Dresden](#) unterstützt bei der Ausarbeitung von Datennutzungsverträgen
- **Technische Möglichkeiten** der sicheren Übermittlung sensibler Forschungsdaten an externe Partner:
  - Übermittlung per verschlüsselter E-Mail: der [Secure Mail-Dienst des ZIH](#) kann verwendet werden, wenn die Datenmenge nicht zu groß ist
  - Übermittlung via Gruppenlaufwerk: dies setzt die Einrichtung eines [Gruppenlaufwerkes](#) über das [Self Service Portal](#) der TU Dresden voraus; für die Kooperationspartner müssen Gast-Logins beantragt und das Gruppenlaufwerk mit entsprechenden Berechtigungen versehen werden; sensible Daten müssen mit [Boxcryptor](#) verschlüsselt werden; zum Abruf müssen die externen Partner eine [VPN-Verbindung](#) in das Netzwerk der TU Dresden aufbauen
  - Übermittlung via [Cloudstore](#)-Ordner: auch hier müssen sensible Daten mittels [Boxcryptor](#) verschlüsselt werden
- **Qualitätssicherung** kann z.B. dadurch umgesetzt werden, dass die eigenen Datenlieferungen protokolliert werden (welche Daten wurden wann an welchen Partner weitergeleitet)?

## 3 Unterstützende Prozesse für das FDM

Im Folgenden werden Prozesse umrissen, die die Umsetzung der Kernprozesse des FDM unterstützen können.

### 3.1 Datenerhebung

- **Versuchspersonenmanagement:** Ein effizientes Forschungsdatenmanagement setzt bereits in der Phase der Rekrutierung von Versuchspersonen an. Empfohlen wird dabei die Nutzung von [ORSEE/PaMS](#).
- **Datenerhebung:** Die Erhebung von Forschungsdaten folgt i.A. spezifischen Voraussetzungen bzw. Erfordernissen der einzelnen Arbeitsgruppen der Fakultät. Typischerweise werden aber in jedem Projekt Selbst- und/oder Fremdauskünfte, experimentelle oder sonstige Verhaltensdaten sowie ggf. (in

weitestem Sinne) physiologische Parameter mit Hilfe (peripher)physiologischer oder bildgebender Verfahren erhoben. Diese können z.B. via REDCap zusammengeführt werden. Der Vorteil der Nutzung von REDCap ist dabei u.a., dass REDCap direkt mit PaMS verschaltet ist bzw. verschaltet werden kann und somit personenbezogene Daten in PaMS gespeichert werden, während in REDCap nur die eigentlich projektrelevanten Daten gesammelt werden.

## 3.2 Datenorganisation

### Grundlage

- **TUD-FDM-Leitlinie 5.** „Die TU Dresden verfügt über eine Grundausstattung an Forschungsdateninfrastruktur und stellt damit eine angemessene Aufbewahrung und die technische Verfügbarkeit von digitalen Forschungsdaten sicher“

### Umsetzung

- **Datenablage**
  - Zum Einsatz kommen in der Regel ZIH-Dienste wie [Gruppenlaufwerke](#), [Cloudstore](#) oder [Share-Point](#).
  - Für bestimmte Anwendungsfälle kommt REDCap zum Einsatz .
- **Gruppenlaufwerke/Ordnerstrukturen** Hierfür gelten die Umsetzungsvorschläge zu den Punkten 2.2 bis 2.4 entsprechend. Ergänzend wird die Nutzung einer Checkliste für die Datenablage empfohlen, die wie folgt aussehen kann:

#### ***Beispiel einer Checkliste für Datenpflege auf dem Gruppenlaufwerk***

*Datentransfer* - nur relevante Daten auf Gruppenlaufwerk ablegen (Relevanz wird definiert von Projektleitung, im Zweifelsfall: Absprache im jeweiligen Team) - Datei an korrekter Stelle einfügen (im Zweifel nachfragen bei Projektleitung) - Eindeutig als zusammengehörig erkennbare Pakete aus Daten und Dokumentation schnüren

*Datenqualität* - Datenintegrität prüfen (Dateien nicht korrupt, d.h. lassen sich öffnen) - Verifizierung der verwendeten Maßeinheiten/Abkürzungen gemäß Code Book

*Dokumentation* - Ordner/Dateien gemäß vereinbarter Dateinamenskonventionen benennen - Codierungen aus Dateinamen und Spaltenköpfen im Code Book hinterlegen - Code Book auf Aktualität und Vollständigkeit prüfen - Readme.txt anlegen bzw. aufdatieren - Skripte kommentieren - Skripte nach pro Team vereinbartem Style Guide anfertigen, z.B. für R Skripte nach dem [Tidyverse Style Guide](#)) - Beschreibung der Prozessabläufe in Manualen anlegen bzw. auf Aktualität prüfen

*Datenhygiene* - 10 min pro Woche einplanen für „Datenhygiene“ auf dem Gruppenlaufwerk (z.B. gemeinsam im Team-Meeting) - Fokus-Treffen, um gemeinsam Strukturen/Standards für Readme, Gruppenlaufwerke und weitere FDM-Regeln auszuarbeiten

Jede Arbeitsgruppe sollte auf Basis dieses Beispiels eine Checkliste ausarbeiten, im Gruppenlaufwerk ablegen, regelmäßig aktualisieren und anwenden, sofern keine geeigneten alternativen Wege der Datenpflege beschritten werden (die dann allerdings gleichermaßen für alle Arbeitsgruppenmitglieder verbindlichen, dokumentierten und gut auffindbaren Regeln unterliegen müssen).

### 3.3 Wissensmanagement Onboarding

- Jede:r neue Mitarbeitende erhält diese FDM-Leitlinien zur Durchsicht (ggf. abzeichnen lassen?)
- Die FDM-Leitlinien können durch eine individuelle Onboarding-Checkliste mit Informationen zur Arbeitsweise innerhalb einer Forschungsgruppe ergänzt werden

#### *Beispiel einer Onboarding-Checkliste*

- wie werden Dateien benannt
- wo ist was auf dem Laufwerk zu finden/abzulegen
- wer in der Gruppe ist Ansprechpartner für welches Thema, welche Software, etc.
- Hinweise zum Umgang mit Datenschutz
- Hinweise zur Verwertung der Daten nach Projekteende ...

### 3.4 Wissensmanagement Offboarding

Der Offboarding-Prozess umfasst u.a. die Bereinigung und Übergabe von Daten. Dieser Prozess beginnt 3-6 Monate vor Austritt der:des Mitarbeitenden und kann unter Mitwirkung der [Kontaktstelle Forschungsdaten](#) geplant und durchgeführt werden. Die Dokumente aus dem Offboarding sind u.a. Grundlage für das Onboarding der:des nachfolgenden Mitarbeitenden. Auf diese Weise wird sichergestellt, dass von dem:der Mitarbeitenden erhobene Daten potenziell auch nachgenutzt werden können. Alle Arbeitsgruppen sind angehalten, anhand des folgenden Beispiels eine Checkliste auszuarbeiten, im Gruppenlaufwerk abzulegen und regelmäßig anzuwenden

#### *Beispiel einer Offboarding-Checkliste*

##### *Datenbereinigung*

- nicht mehr benötigte Dateien gelöscht oder archiviert
- aufzubewahrende Daten systematisch benannt und abgelegt

##### *Datendokumentation*

- worum geht es bei diesem Datensatz/Ordner?
- was ist der Ursprung dieser Daten/dieses Ordners?

##### *Datenverwendung*

- für wen sind diese Daten interessant?
- was würden Sie gern noch selbst mit den Daten anfangen?
- was könnten andere noch mit diesen Daten anfangen?
- die Daten sind auf folgende Art zu verwenden (Einsatzzweck/Methode/Software)
- die Daten haben Abhängigkeiten/Schnittstellen zu ...
- welche Probleme kennen Sie im Zusammenhang mit den Daten? wie sollte Ihr:e Nachfolger:in mit diesen Problemen umgehen?
- für welchen Zweck dürfen die Daten unter keinen Umständen eingesetzt werden?
- welchen Einstieg würden sie derjenigen Person raten, die zukünftig mit diesen Daten arbeiten wird?

##### *Soziales Netzwerk*

- wen haben Sie für Erstellung bzw. Verwendung der Daten häufig kontaktiert?
- wer hat Sie bzgl. der Verwendung der Daten häufig kontaktiert?
- wie kann Ihr Nachfolger Sie erreichen?

## 4 Tools für das Forschungsdatenmanagement

### 4.1 Tools der TU Dresden

**4.1.1 ZIH** Alle im weiteren Verlauf beschriebenen Tools der TU Dresden werden in der einen oder andern Form vom ZIH bereitgestellt. Das ZIH bietet darüber hinaus aber noch weitere, hier nicht beschriebene Archivlösungen an, die hier nicht gesondert erläutert werden. Wenden Sie sich dazu direkt an das ZIH!

**4.1.2 Kontaktstelle Forschungsdaten** Die [Kontaktstelle Forschungsdaten](#) unterstützt alle Forschenden an der TU Dresden bei Fragen rund um das Forschungsdatenmanagement. Angeboten wird Folgendes:

- **Textbausteine für Datenmanagementpläne** (DMP) in Forschungsanträgen für Einzelprojekte und Beratung zu DMPs für die Beantragung von Verbundprojekten: Über ein [Webformular](#) kann ein Textvorschlag (innerhalb von 2 Werktagen) bzw. ein Gesprächstermin angefragt werden. Der Textvorschlag für Einzelprojekte berücksichtigt die aktuellen Richtlinien der DFG. Bei Verbundprojekten werden in einem Gespräch die spezifischen Anforderungen an das Datenmanagement im geplanten Verbund besprochen, so dass die Forschenden eine Struktur und Ideen für den DMP Text entwickeln.
- **Beratung, Vorträge und Informationen zum Forschungsdatenmanagement** für einzelne Personen, für Forschungsgruppen, für Professuren oder Institute, siehe [hier](#); Um mehr Informationen rund um FDM zu erhalten, können Auskünfte per E-Mail von der Kontaktstelle eingeholt werden, auch die Website der Kontaktstelle bietet bereits viele grundlegende Informationen. Zum Erarbeiten eines FDM-Konzeptes bietet die Kontaktstelle Gesprächstermine zur individuellen Beratung an. Außerdem besteht die Möglichkeit, Vorträge durch die Kontaktstelle zu spezifischen Themen zu buchen.
- **Unterstützung bei der FDM-Umsetzung** für Arbeitsgruppen oder Struktureinheiten, siehe [hier](#); Die Kontaktstelle unterstützt dabei, Forschungsdatenmanagement in den Forschungsalltag zu integrieren. Dazu gehört die Begleitung des FDM-Vorhabens von der Idee bis hin zur gelebten Praxis, das Forschungsdaten-Offboarding für ausscheidende Mitarbeitende und/oder die technische Unterstützung bei der Nutzung der TUD-Ressourcen oder der Einführung einer neuen Technologie fürs FDM.

**4.1.3 ORSEE/PaMS** Die Rekrutierung von Proband:innen lässt sich effizient und FDM-regelkonform über das webbasierte und kostenlose Online-Rekrutierungssystem [ORSEE](#) organisieren. Das Projekt wird auf GitHub gehostet, siehe [hier](#).

Die Hauptfunktionen sind:

- mehrere Experimentatoren/Labore/Subjektpool/Experimentklassen/Sprachunterstützung
- Auswahl von Attributabfragen
- zufällige Rekrutierung
- öffentlicher und interner Experimentkalender
- Reputationssystem
- automatisierter Versand
- PDF-Ausgabe
- Rechteverwaltung für Experimentatoren

Das ORSEE-System besteht im Wesentlichen aus zwei Schnittstellen: der Experimentatorensicht und der Probandensicht. Experimentatoren erstellen Experimente, die aus mehreren Sitzungen bestehen können, und laden Probanden ein. ([Eingeladene](#)) Probanden können sich für eine der Experimentsitzungen anmelden, um daran teilzunehmen (Subgruppen können anhand von Kriterien ausgewählt werden *Alter, Geschlecht, Händigkeit etc.*). Es besteht auch die Möglichkeit das implementierte Bezahlssystem zu verwenden (wobei hierbei die Richtlinien der Drittmittelverwaltung *Dezernat 3.1* beachtet werden).

Um allen datenschutz- und sicherheitsrelevanten Aspekte der Speicherung von personenbezogenen Daten gerecht zu werden, hat das Zentrum für Informationsdienste und Hochleistungsrechnen (ZIH) der TU-Dresden das Datenbanksystem **\*\*PaMS\*\_\*\*** (Participant Management System) neu entwickelt. Dabei werden alle personenbezogenen Daten gespeichert. Ein Übertrag aus den eingegebenen Daten der ORSEE ist automatisch möglich. Zusätzlich kann eine fortlaufende ID generiert werden, die als Identifier für REDCap (siehe folgendr Abschnitt) dient. So ist es möglich, ohne personenbezogene Daten Fragebögen via REDCap zu erheben, da dort nur der fortlaufende Identifier gespeichert wird. Das folgende Organigramm zeigt einen Überblick über den möglichen Einsatz der IT-Infrastruktur und die Funktion der darin enthaltenen Tools und Systeme.

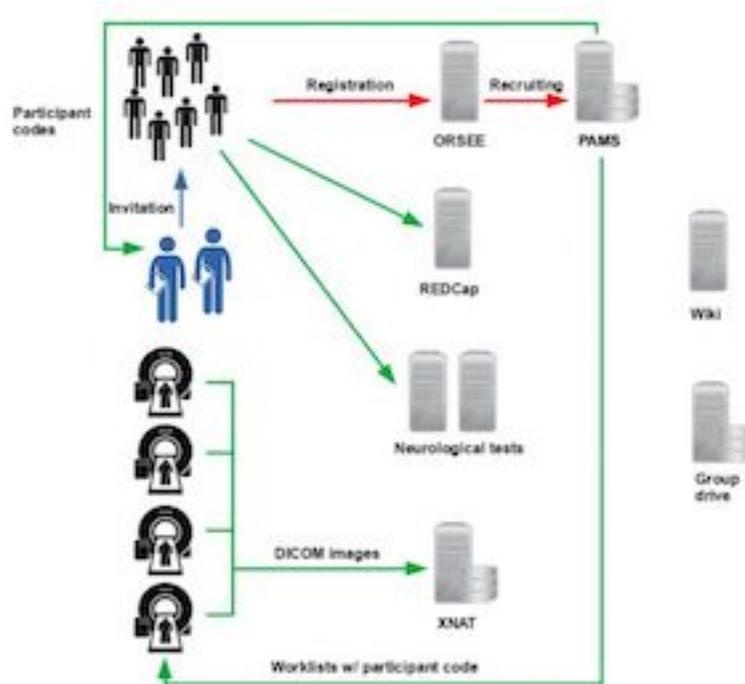


Figure 2: Organigramm der IT-Infrastruktur zum Probanden- und Forschungsdatenmanagement

**4.1.4 REDCap** REDCap (Research Electronic Data Capture) ist eine vom ZIH gehostete, browserbasierte, metadatengesteuerte EDC-Workflow-Methodik zum Entwerfen von Forschungsdatenbanken (siehe [hier](#) für weitere Informationen). Es ist in der akademischen Forschungsgemeinschaft weit verbreitet: Das REDCap-Konsortium ist ein kooperatives, internationales Netzwerk von mehr als 5800 institutionellen Partnern in über 145 Ländern mit insgesamt mehr als 2.100.00 Endbenutzern, die die [REDCap Software](#) für mehr als 450.000 laufende Forschungsprojekte einsetzen.

Obwohl REDCap institutionellen Partnern kostenlos zur Verfügung steht – abzüglich der Personalkosten für den internen IT-Support – ist REDCap ausdrücklich keine Open Source Software. Bestimmte Endbenutzer-Lizenzvereinbarungen unterscheiden sie von einer typischen Open-Source-Lizenz: Die Nutzung der Software ist nur für nichtkommerzielle Forschungszwecke gestattet.

Der REDCap-Design-Workflow ist zwar insgesamt komfortabel gestaltet, weist aber einige Einschränkungen auf. Wenn ein Projekt im Produktionsmodus bereitgestellt wird, müssen weitere Überarbeitungen des Datenbankdesigns vom Informatikern genehmigt werden. Dies gestaltet sich allerdings meist als unproblematisch. Es wird von derzeit drei Administratoren verwaltet und gewartet und aktuell u.a. in drei SFB/TRR im Rahmen der Information Infrastructure Projekte eingesetzt. Insgesamt hat die TUD-Version von REDCap über 600 Projekte angelegt. Seine Funktionalität umfasst:

- Online- oder Offline-Projektdesign: Online mit dem online Designer oder offline mit einer „Data Dictionary“-Vorlage in Microsoft Excel, die später in REDCap hochgeladen werden kann.
- Verfügbarkeit: Software ist für Partner des REDCap-Konsortiums kostenlos erhältlich.
- Sicherheit und Webbasiertheit: Dateneingabe, Online-Umfragen oder Datenbankerstellung und -verwaltung ist von überall auf der Welt über eine sichere Webverbindung mit Authentifizierung und Datenprotokollierung möglich.
- Schnelligkeit und Flexibilität: Von der Konzeption bis zur produktionsreifen Datenbank oder Umfrage in dauert es üblicherweise weniger als einen Tag.
- Standortübergreifender Zugriff: REDCap-Datenbanken/Umfragen können von Forschenden von mehreren Standorten und Institutionen genutzt werden.
- Umfangreiches Rechtemanagement: Forschungsgruppen haben vollständige Autonomie und Kontrolle, um neue Benutzer hinzuzufügen. Einzelnen Nutzerinnen können spezifische Rechte vergeben werden (so ist z.B. nur der Export von einem spezifischen Fragebogen aus einem Forschungsprojekt möglich, so dass keine Datensätze mehr anderweitig verschickt werden müssen).
- Vollständige Adaptabilität: Es besteht volle Kontrolle über die Gestaltung von Datenbanken oder Umfragen.
- Berichte/Statistiken/Projektdashboard: Alle eingegebenen Daten können anschaulich aufbereitet und bei Bedarf auch exportiert werden, was die volle Kontrolle über den Erhebungsablauf gewährleistet.
- Audit-Trails: Diese erlauben die Verfolgung von Datenmanipulationen und Benutzeraktivitäten.
- Automatisierte Exportverfahren: Diese erlauben nahtlose Datendownloads in Excel, PDF und gängige Statistikpakete (CSV, XML, SPSS, SAS, STATA, R). Dabei ist möglich, alle bekannten Identifikatoren, Freitextfelder, Datumsangaben zu entfernen.
- Integrierter Projektkalender und Planungsmodul
- Ad-hoc-Berichtstools: Diese erlauben die Erstellung von Code Books
- Viele erweiterte Funktionen wie Verzweigungslogik, Hochladen von Dateien und berechnete Felder.
- Mobile REDCap-App für iOS und Android
- API Schnittstellen für individuelle Programmierungslösungen, so z.B. für den automatisierter Upload von behavioralen, fMRT-Scanner Logfiles oder EEG-Daten
- die REDCap-Instanz wird auf Servern des ZIH der TUD gespeichert und ausgeführt
- alle Kontaktdaten der Probanden werden in der separaten Anwendung (PaMS) verwaltet. Einladungs- und Erinnerungsemails werden über eine API versendet.
- umfangreiche [Video-Tutorials](#)

**4.1.5 Gruppenlaufwerk/Cloudstore/SharePoint** [Gruppenlaufwerke](#) bieten Arbeitsgruppen die Möglichkeit, Daten zentral zu speichern und größere Datenmengen mit anderen TU-Angehörigen oder Gästen zu teilen. Dies geschieht über ein gemeinsames zentrales Verzeichnis, auf das auch von PCs und Laptops zugegriffen werden kann. Das ZIH stellt dazu die sichere Infrastruktur zur Verfügung. Die Zugriffsrechte darauf, aber auch auf Unterordner, können individuell festgelegt werden. Snapshots erlauben den Zugriff auf ältere Dateiversionen. Zusätzlich erfolgt ein tägliches Backup.

[Cloudstore](#) bietet Sync&Share von Dateien (Sichere Alternative zu Dropbox) mit Browser-basiertem Zugriff oder Klienten für alle wesentlichen Betriebssysteme. Dateien werden also regelmäßig zwischen dem eigenen PC und der Cloud synchronisiert. Andere Forschende können zu Ordnern eingeladen werden. Es muss beachtet werden, dass das gleichzeitige Zusammenarbeiten an einem Dokument zu Konflikten führt. Es wird seitens des ZIH **kein** Backup erstellt. Es wird empfohlen, selbständig ein Backup, zum Beispiel auf das eigene [Home-Verzeichnis](#), einzurichten.

[SharePoint](#) ist eine webbasierte Plattform zur Zusammenarbeit von Gruppen/Projekten. Der Dienst bietet eine Dateiablage und -verwaltung (inklusive paralleler Online-Bearbeitung mit MS Office), Task-Management, Notizen, Wiki, Kalender, Aufgabenplanung, Blogs und noch einiges mehr an. Es ist jedoch etwas Verwaltungsaufwand und teilweise Programmierung notwendig, um alle Funktionalitäten auszuschöpfen. SharePoint kann ein leistungsfähiger Dienst zum Verwalten von Metadaten sein (also solcher Daten, die Eigenschaften eines Datensatz möglichst umfassend beschreiben).

**4.1.6 OpARA** [OpARA](#) (Open Access Repository and Archive) ist eine Archivlösung der TU Dresden. Wie schon kurz unter 2.4. Datenpublikation ausgeführt, handelt es sich um ein kuratiertes Repository: Die zu archivierenden Daten werden zunächst durch eine von den Einreichenden zu benennende Person einem fachlichen Gutachten unterzogen. Damit soll sichergestellt werden, dass die Daten und alle weiteren einzureichenden Materialien den disziplinspezifischen Anforderungen an Nachvollziehbarkeit und Transparenz genügen. Ein weiteres technisches Review, bei dem u.a. geprüft wird, ob die Daten hinreichend durch Metadaten beschrieben sind, erfolgt durch das OpARA Team. Das Datenarchiv erhält dann einen Digital Object Identifier (DOI), der im Data Availability Statement zu Publikationen angegeben werden kann. Details zum Archivierungsprozess finden sich im [OpARA-Nutzerhandbuch](#). Wie ebenfalls oben erwähnt, ist die Archivierung auf OpARA ein wenig umständlicher als beispielsweise auf OSF, dafür ist sie im Prinzip qualitativ höherwertig, vorausgesetzt, Einreichende und fachlich Begutachtende sind mit den Prinzipien der nachvollziehbaren und transparenten Datenpublikation (und idealerweise auch mit den FAIR Principles) vertraut. Ein weiterer Vorteil von OpARA, nämlich die Unmöglichkeit, im Nachhinein Veränderungen am Archiv vorzunehmen, kann allerdings auch zugleich als Nachteil angesehen werden. Wenn Daten z.B. im Zusammenhang mit einer einzureichenden wissenschaftlichen Publikation bereitgestellt werden, ist es sinnvoll, das einzureichende Manuskript ebenfalls bereitzustellen. Etwaige Revisionen dieses Manuskripts und/oder im Reviewprozess notwendig gewordene Änderungen in Analyseskripten können dem Archiv allerdings im Nachhinein nicht mehr hinzugefügt werden. Ein Workaround könnte hier sein, dass die Daten für den Reviewprozess via OSF bereitgestellt werden und die Archivierung auf OpARA erst nach Acceptance eines Manuskriptes erfolgt und die generierte DOI im Rahmen des Proof Reading eingefügt wird.

## 4.2 Externe Tools

**4.2.1 Leibniz Psychology DataWiz** Das Tool [DataWiz](#) von [Leibniz Psychology](#) liefert eine integrierte Lösung für alle Fragen des Forschungsdatenmanagements. Wir haben bisher allerdings keine Erfahrungen mit diesem Tool sammeln können, sehen es aber – insbesondere für den Fall, dass in einzelnen Arbeitsgruppen oder Projekten noch keine etablierten Routinen für das Forschungsdatenmanagement existieren – als ausgesprochen vielversprechendes Tool an. Sollten Arbeitsgruppen dieses Tool nutzen, bitten wir darum, uns Input zu geben bzw. dieses Dokument entsprechend zu ergänzen.

**4.2.2 Open Science Framework (OSF)** Das [OSF](#) ist ein durch das Center for Open Science ins Leben gerufenes Repository, auf dem man Daten, Analyseskripte, Versuchsmaterialien und Präregistrierungen ablegen kann. Auch weitere Materialien wie etwa aus dem Projekt hervorgegangene Konferenzpräsentationen können abgelegt werden. Jedes Projekt, das von einer Person erstellt wird, kann allein oder gemeinsam mit den Co-Autor:innen verwaltet werden. Dabei kann jederzeit entschieden werden, ob das Projekt frei zugänglich ist oder nur für die Beitragenden sichtbar sein soll. Alle Änderungen am Projekt werden in der Historie festgehalten, sodass (auch für Außenstehende) einsehbar ist, wer welche Projektkomponente wann bearbeitet hat. Anders als bei GitHub (siehe unten) interagiert OSF aber nicht mit dem lokalen Projektordner auf dem Computer, OSF erkennt und markiert daher auch nicht eigenständig Änderungen in Versionen von Projektdateien. Allerdings kann, wie bereits unter 2.4. Datenpublikation angesprochen, ein OSF-Projekt mit einem GitHub Repository verlinkt werden, so dass es mit jeder Aktualisierung des GitHub-Repositorys ebenfalls aktualisiert wird (siehe [Connect GitHub to a Project](#)) Der Upload von Dateien und Daten kann auf verschiedenen Servern stattfinden; damit jedoch der deutsche Datenschutz greift, sollte immer auf den Frankfurt/Main-Server zurückgegriffen werden. Jedes Projekt kann und sollte zudem mit einer Lizenz versehen werden, die die Verwendung aller Dateien und Daten für Außenstehende regelt (siehe ebenfalls oben 2.4 Datenpublikation, Infos bezüglich geeigneter Lizenzen finden sich auf [Creative Commons](#)). Es besteht auch die Möglichkeit, ein Embargo auf die Nutzung von Daten zu verhängen.

Seinen Ursprung hatte das OSF im Bereitstellen einer Plattform für Präregistrierungen, wo Forschende ihre Hypothesen und Analysepläne in einem mit Zeitstempel versehenen Dokument festhalten, bevor sie ihre Studie durchführen. Es stehen [Vorlagen](#) für verschiedenste Studienarten zur Verfügung, um den Forschenden das Ausfüllen einer Präregistrierung zu erleichtern. Jedes Projekt ist mit einer sehr kurzen URL erreichbar, die sich sehr gut eignet, um in Journal-Artikeln direkt auf die Daten und Analyseskripte zu verweisen.

Beim Hochladen von Dateien in ein OSF-Projekt gelten die gleichen Grundsätze zur Datenqualität wie beim Ablegen von Dateien auf einem Gruppenlaufwerk oder auf OpARA. Ein Beispiel, wie eine sinnvolle OSF-Ordnerstruktur prinzipiell aussehen könnte, findet sich hier: <https://osf.io/n7263>. Ein Beispiel eines (mit GitHub verlinkten) OSF-Projektes aus unserer Fakultät findet sich hier: <https://osf.io/9gwe7/>

**4.2.3 GitHub** Für die Zusammenarbeit an Forschungsprojekten ist eine gemeinsame Arbeitsumgebung hilfreich. Der Sharepoint, Google Docs oder Microsoft Teams stellen hierfür ebenso eine Möglichkeit dar wie Gruppenlaufwerke. Wenn sich in diesen Arbeitsumgebungen etwas ändert, ist es möglicherweise nicht einfach, diese Änderungen zu verfolgen. Wenn eine Änderung in eine Sackgasse führt, wäre es wünschenswert, zu einem früheren Punkt in der “Zeitachse” des Projektes zurückzukehren, aber kann sich u.U. schwierig gestalten, diesen wiederherzustellen, weil nicht immer klar ist, ob alle seither gemachten Änderungen verworfen werden müssen oder nur bestimmte. Hier ist Versionskontrolle enorm hilfreich.

**GitHub** (oder auch die an der TU Dresden vorgehaltene und eher datenschutzkonforme, dafür aber wieder mit anderen Problemen behaftete Variante **GitLab**) ermöglicht eine solche Versionskontrolle, d.h., es können jederzeit alle Änderungen an Projektdateien verfolgt und überprüft werden, was wann von wem geändert wurde.

Außerdem ist es möglich, neben der hauptsächlichen “Zeitachse” eines Projektes (*main* genannt), auch alternative “Zeitachsen” zu etablieren, etwa, wenn Mitarbeitende in einem Projekt Änderungen vornehmen, die sie zunächst einmal zur Diskussion stellen wollen (z.B. Änderungen im Analysecode oder in Manuskripten, letzteres analog zum *Überarbeiten*-Modus bei Word). Hierzu kann eine sogenannten *branch* erstellt werden, in der Änderungen ohne Veränderungen am *main* Projekt implementiert, dann diskutiert und schlussendlich mit dem *main* Projekt verbunden (*merged*) werden können.

Ein weiteres hilfreiches Feature von via GitHub organisierten Projekten ist, dass Daten, Analysecode, Materialien etc. sehr leicht mit anderen Projektbeteiligten bzw. anderen Arbeitsgruppen geteilt werden können, indem das Projekt einfach lokal heruntergeladen (*cloned*) werden kann. Das Teilen von Projekten kann freilich auch via OSF geschehen (indem ein OSF-Projekt als *.zip Datei heruntergeladen wird*), aber eine Mitarbeit am Projekt oder eine Änderung von Projektdateien ist ungleich komplizierter und unterliegt auch keiner so strengen Versionskontrolle wie unter *GitHub*. Es ist bei *GitHub* möglich, Projekte entweder *private\** oder *public* zu setzen. Im ersteren Falle haben nur direkt vom Projekt-Administrator eingeladene Personen Zugriff auf das Projekt (hierfür ist ein GitHub Account erforderlich), im letzteren Falle prinzipiell alle, die über einen Computer mit Internet-Zugang verfügen (hierfür ist dann auch kein GitHub Account erforderlich).

Wie oben schon angesprochen, können GitHub/GitLab Projekte auch **mit OSF verbunden werden**, so dass alle Änderungen am via Git organisierten Projekt direkt auf OSF gespiegelt werden. OSF wäre somit eine Variante, ein Projekt mit anderen Gruppen zu teilen, die Daten und/oder Code nachnutzen wollen, während GitHub/GitLab eine Variante wäre, mit anderen direkt an einem Projekt zusammenzuarbeiten.

Ein Beispiel für ein via GitHub organisiertes Forschungsprojekt, das mit OSF verbunden ist, findet sich hier:

- [GitHub Projekt](#)
- [verlinktes OSF Projekt](#)

## 4.3 Individuelle Tools

Jede Arbeitsgruppe arbeitet darüber hinaus mit individuellen Tools zur Unterstützung des Forschungsdatenmanagements. Diese gestalten sich höchst unterschiedlich, sollten aber in jedem Falle in einer strukturierten Form vorgehalten werden. Zu vermeiden ist unbedingt, dass Wissen über Arbeitsabläufe des Forschungsdatenmanagements (aber auch darüber hinaus) nur in den Köpfen von Arbeitsgruppenmitgliedern vorliegt. Alle für eine Arbeitsgruppe relevanten Arbeitsabläufe – ob nun neu etablierte oder bereits länger genutzte und daher *vermeintlich* allseits bekannte – sollten daher in jedem Falle schriftlich dokumentiert werden. Einige Hinweise, wie das geschehen kann, werden im Folgenden gegeben.

**4.3.1 Standard Operating Procedures** Standard Operating Procedures (SOP), alternativ Technisch-Organisatorische Maßnahmen (TOM) genannt, beschreiben bestimmte wiederkehrende und standardisiert durchzuführende Arbeitsabläufe einer Arbeitsgruppe. Beispiele wären Beschreibungen von Konventionen zu Ordnerstrukturen und Dateibenennungen auf einem Gruppenlaufwerk, Vereinbarungen, Anleitungen und/oder Checklisten für wiederkehrende Prozesse etc. Ein Beispiel für eine SOP der Arbeitsgruppe Differentielle und Persönlichkeitspsychologie ist die zur [Behandlung von Ausreißern in statistischen Analysen](#).

**4.3.2 Lab Manuals** Ein Lab Manual bündelt eine Reihe von (oder alle) SOPs bzw. TOMs einer Arbeitsgruppe. Auf diese Weise können – via Sharepoint, Gruppenlaufwerk, GitHub o.ä. – alle Arbeitsabläufe einer Arbeitsgruppe an einem Ort bereitgestellt (und kontinuierlich überarbeitet) werden. Das ist insbesondere für das Onboarding neuer Mitarbeitender oder für (von einigen oder allen Team-Mitgliedern) seltener genutzte Arbeitsabläufe ausgesprochen hilfreich und unterstützt damit das Wissensmanagement. Ein Beispiel für ein (via GitHub organisiertes) Lab Manual findet sich hier:

- [LabManual der Arbeitsgruppe Differentielle und Persönlichkeitspsychologie](#)

In diesem Beispiel sind in einem Living Document die wesentlichen Arbeitsabläufe in Bezug auf Forschung, Lehre und (in einem separaten, privaten Repositorium) Administration in der Arbeitsgruppe Differentielle und Persönlichkeitspsychologie zusammengefasst. Auf diese Weise ist es beispielsweise möglich, neue Team-Mitglieder oder auch Studierende, die in einer Arbeitsgruppe Abschlussarbeiten schreiben, gezielt auf spezielle, für ihre Arbeit relevante Arbeitsabläufe hinzuweisen, ohne dafür (zunächst) konkrete Unterweisungstermine zu vereinbaren.

## 5 Ressourcen und Vorlagen

Die folgenden Ressourcen stellen in Teilen nur eine Auflistung von bereitzustellenden Ressourcen dar. Die Auflistung wird laufend aktualisiert und ergänzt.

- Landing Page mit Neuigkeiten zur Fakultät
- Gesammelte Links zu Dokument-Vorlagen aus dieser FDM-Leitlinie
  - Checkliste Onboarding
  - Checkliste Offboarding
  - Checkliste für Datenpflege auf dem Gruppenlaufwerk
  - ggf. Link zu REDCap-Templates
  - ggf. Link zu OSF-Template
- IT-Dienste des ZIH
  - Buchung von IT-Services des ZIH über das [Self Service Portal](#)
  - Zentrale Datenspeicher der TU Dresden: [Gruppenlaufwerke](#)
  - Verschlüsselung von Laufwerken mit [Boxcryptor](#)
  - [Cloudstore](#) der TU Dresden
  - Versendung verschlüsselter Mails mit [Secure Mail](#)
- Datenschutz
  - Übersicht [Creative Commons Lizenzen](#) bzw. [Creative Commons Licence Chooser](#)
  - [Professur für Urheberrecht und geistiges Eigentum](#) der TU Dresden
  - [Datenschutzbeauftragte:r](#) der TU Dresden
- Links mit administrativen Informationen, z.B. Anleitungen zu Verwaltungsabläufen der Fakultät oder Leitfaden für die Anfertigung von [Bachelorarbeiten](#) bzw. [Masterarbeiten](#)
  - ggf. Liste mit Weiterbildungsangeboten (Zentrum für Weiterbildung, SLUB, Graduiertenakademie) und nützlichen Links (z.B. zu Video-Tutorials)
  - ggf. Glossar mit Begriffen aus dem Kontext FDM und Fakultät Psychologie