



UNIVERSITÄT  
LEIPZIG

# **Datenmanagementpläne erstellen – Teil 2**

Pia Voigt

AG Forschungsdatenmanagement

09.03.2021

# PRAXISTIPPS DATENMANAGEMENTPLÄNE

# DATENSATZBESCHREIBUNG

# DATENSATZBESCHREIBUNG

- **Datentypen/-arten:**

Text, Audio-, Videoaufnahmen, digitale Fragebögen, Tabellen etc.

→ ggf. Unterscheidung in personenbezogene und nicht personenbezogene

Forschungsdaten

→ erleichtert Übersicht, über zu anonymisierende Daten

- Unterscheidung von **Rohdaten** und **bearbeiteten Daten**

→ Rohdaten: direkt erhobene Daten, keine Bearbeitung/Analyse/Nachbearbeitung

→ für Archivierung und Veröffentlichung wichtig

# DATENSATZBESCHREIBUNG

- **nachhaltige Dateiformate**
  - unkomprimierte Formate, Softwareunabhängigkeit  
(**Open Source/freie Software verwenden, möglichst offene Formate**)
  - Beispiele: .csv, auch .xlsx (nicht .xls); .png statt .jpg
  - **R** statt SPSS, [Easytranscript](#) (Transkription), [OpenRefine](#) (Zusammenführen heterogener Datensätze, Bereinigung)
  - Hilfestellung: [Formate erhalten - forschungsdaten.info](#), [SoSciSo-Verzeichnis](#), [Verein DE-RSE](#)
- Maßnahmen zur **Qualitätssicherung**:
  - Sind Ihre Daten ausreichend und nachvollziehbar mit (fachspezifischen) **Metadaten** beschrieben? Ist die **Dokumentation** vollständig?
  - Entsprechen Ihre Daten **fachspezifischen Standards** und untermauern Ihre Forschungsergebnisse? **Begutachtung** durch KollegInnen/BetreuerIn möglich?
  - Sind Ihre Daten ausreichend **anonymisiert** und erfüllen die **Datenschutzstandards**?

# DATENSATZBESCHREIBUNG

Metadaten und Datendokumentation als Mittel der Qualitätssicherung von Forschungsdaten:

<https://www.youtube.com/watch?v=9vj2-X-ZUxU>

# LEITLINIEN ZUM UMGANG MIT FORSCHUNGSDATEN

- relevante Vorgaben (Policies, Leitlinien):
  - für UL:
    - Grundsätze für das Management von Forschungsdaten
    - Satzung zur Sicherung der guten wissenschaftlichen Praxis
  - Gute wissenschaftliche Praxis:
    - Leitlinien zur Sicherung guter wissenschaftlicher Praxis – Kodex (DFG)
  - bei externer, finanzieller Förderung:
    - Vorgaben der Förderinstitutionen (DFG, BMBF, Horizon Europe),  
UL zweitrangig
    - Vorgaben im Gesamtprojekt/Teilprojekt  
→ müssen zentral erstellt werden

# LEITLINIEN ZUM UMGANG MIT FORSCHUNGSDATEN AN DER UL

- verantwortungsbewusster Umgang mit Forschungsdaten:
  - Fachspezifische Standards
  - FAIR-Prinzipien
- Speicherung und Archivierung in institutionellen oder fachspezifischen/generischen Infrastrukturen
- Open Access als Standard → Ausnahmen möglich
- Verantwortlichkeiten:
  - DMP aufstellen, regelmäßig aktualisieren
  - Datenschutzrechtliche und ethische Bestimmungen einhalten
  - Regelungen von Drittmittelgebern, Geheimhaltungsvereinbarungen beachten



# DATENVERÖFFENTLICHUNG UND LIZENZIERUNG

## DATEN VERÖFFENTLICHEN – WARUM UND WIE?








- Sichtbarkeit und Zitierbarkeit als wissenschaftliche Publikationen
  - Nachvollziehbarkeit der Forschung durch die wissenschaftliche Community
  - Nachnutzbarkeit in anderen Zusammenhängen ermöglichen
  - Datenerhalt und Archivierung sicherstellen
  - Teils auch: Erklärung zur Datenverfügbarkeit Voraussetzung für Publikation
  - Voraussetzung für die Förderfähigkeit durch Drittmittel
- Faustregel Repositorien:  
fachspezifisch > institutionell (an der UL: OpARA) > generisch

# DATEN VERÖFFENTLICHEN – LIZENZIERUNG

- Creative Commons Lizenzen, Version 4.0:
  - helfen Urheberrechte an Werke/Daten zu behalten und gleichzeitig deren Kopieren, Teilen und Nachnutzen zu erlauben
  - Lizenzen bestimmen Modalitäten der Nachnutzung, auch Public Domain (CC0) erlaubt Rechtssicherheit und Zitierbarkeit der Daten  
→ **lizenzfrei ≠ freie Lizenz!**
  - wenn Nachnutzung maximiert werden möchte, sollten kommerzielle Nutzung und freie Bearbeitung erlaubt werden  
→ **empfohlen CC BY (Namensnennung)**

➤ Choose a licence

# CC-LIZENZEN IM ÜBERBLICK

CC - Lizenzen		Bedingungen der Weiterverw.	Namensnennung	Vervielfältigung	Verbreitung	Öffentliche Zugänglichmachung	Bearbeitung	Kommerzielle Nutzung	Weitergabe
	bedingungslose Lizenz „no rights reserved“ CC Zero		+	+	+	+	+	+	Generell erlaubt
	Namensnennung CC BY	!	+	+	+	+	+	+	Generell erlaubt
	Namensnennung - Keine Bearbeitung CC BY-ND	!	+	+	+		-	+	Generell erlaubt
	Namensnennung - Nichtkommerziell CC BY-NC	!	+	+	+	+		-	Generell erlaubt
	Namensnennung - Nichtkommerziell - Keine Bearbeitung CC BY-NC-ND	!	+	+	+		-	-	Generell erlaubt
	Namensnennung - Nichtkommerziell - Weitergabe unter gleichen Bedingungen CC BY-NC-SA	!	+	+	+	+		-	Nur unter gleichen Bedingungen
	Namensnennung - Weitergabe unter gleichen Bedingungen CC BY-SA	!	+	+	+	+	+	+	Nur unter gleichen Bedingungen

**Zeichenerklärung:** ! ... Muss auf jeden Fall erfolgen + ... Ist erlaubt - ... Ist verboten

Quelle:  
Open Learn Ware Team der TU Darmstadt,  
[CC BY-SA 3.0,](http://www.e-learning.tu-darmstadt.de/werkzeuge/openlearnware/lehrrmaterial_veroeffentlichen/cc_lizenzen/index.de.jsp)  
[http://www.e-learning.tu-darmstadt.de/werkzeuge/openlearnware/lehrrmaterial\\_veroeffentlichen/cc\\_lizenzen/index.de.jsp](http://www.e-learning.tu-darmstadt.de/werkzeuge/openlearnware/lehrrmaterial_veroeffentlichen/cc_lizenzen/index.de.jsp)

# DATEN VERÖFFENTLICHEN – LIZENZIERUNG

## CC BY

### **Pro:**

- erlaubt maximale Verbreitung und größte Nutzungsfreiheit , bestmögliche Kombinierbarkeit mit anderen CC-lizenzierten Werken

### **Contra:**

- Urheber gibt Kontrolle über Großteil seiner Rechte am Werk ab

## CC BY NC

### **Pro:**

- Ausschluss kommerzieller Nutzungen

### **Contra:**

- verhindert viele Nutzungsformen, die zunächst nicht kommerziell erscheinen, es rechtlich jedoch sind  
→ bspw. Nutzung durch soziale Unternehmen oder gemeinnützige Organisationen, in wissenschaftlichen Publikationen (Fachverlage), generell zu beruflichen Zwecken oder auf Plattformen
- kommerzielle oder nicht kommerzielle Nutzung nach Nutzungskonstellation, nicht nach Typ des Nutzers bestimmt  
→ jeweils Einzelfallentscheidung → aufwendig, schreckt Nachnutzende ab

# DATENAUFBEWAHRUNG UND ARCHIVIERUNG

# DATENAUFBEWAHRUNG

- **Speicherung:**  
während der Projektlaufzeit auf geeigneten Datenträgern bzw. digital (Cloudspeicher)  
→ **Backup! (3-2-1-Regel:** mindestens 3 Kopien einer Datei, auf mindestens 2 unterschiedlichen Medien, wovon mindestens eins dezentral ist)
  - Bsp.: Kopie 1 = Laptop (lokal), Kopie 2 = Homelaufwerk (dezentral via Uni-Login), Kopie 3 = automatisches Backup aus dem Homelaufwerk (URZ)
  - institutioneller Speicher bei URZ: automatisches Backup
  - Speicherwolke: kein automatisches Backup!

## ARCHIVIERUNG – ABGRENZUNG ZU BACKUP

### Backup

- automatische Sicherung aller Daten, um Datenverlust vorzubeugen
- technische Ursachen, z. B. defekt
- menschliches Versagen, z. B. versehentlich gelöscht
- alle Versionen

### Archivierung

- Sicherung ausgewählter Daten, um diese langfristig aufzubewahren
- nur endgültige Versionen
- Integritätssicherung (ggf. Migration in andere Dateiformate)
- Langzeitspeicherung
- Durchsuchbarkeit (Metadaten)



# ARCHIVIERUNG

- Daten müssen für **10 Jahre nach Projektende an der Institution verbleiben**, an der sie entstanden sind  
→ Archivierung (Gute wissenschaftliche Praxis)
- **Ort:** institutionelle Archivierungsmöglichkeit oder fachspezifisches, geeignetes Repositorium/Datenarchiv
- **Datenauswahl:**
  - bearbeitete/endgültige, dem Forschungsergebnis zugrundeliegende Daten (bspw. anonymisierte Transkripte, bereinigte Analysedatensätze)
  - ggf. Rohdaten → **Einverständniserklärung!**
  - Kontextdaten (bspw. Dokumentation, Code/Scripte), wenn Forschungsdaten ohne diese nicht nachgenutzt werden könnten
  - Belegen Zustandekommen des Forschungsergebnisses → Validierung
  - einzigartige Daten (Beobachtungsdaten, einzigartige Settings, nicht wiederholbare Settings, sehr teure und aufwendige Datenerhebungen...)

# LÖSCHEN VON DATEN

- Kriterien:
    - Personenbezug = Löschung, wenn Einverständnis nicht explizit vorliegt
    - Daten ohne Wert für Nachnutzung, keine Anschlussfähigkeit für weitere Projekte
    - zu große Daten (unkomprimiert) → Kapazitäten
    - nicht aussagekräftige, unzureichend dokumentierte Daten
    - Daten, die der Validierung von Forschungsergebnissen nicht dienen
- Löschung immer begründen!

## DATENARCHIVE

- institutionelles Angebot oder Empfehlung der Fachcommunity
  - Auswahl eines Datenarchivs:
    - Zertifizierung (Core Trust Seal, dini/nestor-Siegel, ISO 16363)
    - finanzielle/personelle Ausstattung (Langlebigkeit)
    - Expertise
    - verwendete (Metadaten-)Standards und Lizenzen
    - Nutzungsbedingungen
    - Umgang mit schutzwürdigen Daten
- Auswahl archivwürdiger Daten vgl.:  
[Auswählen und Bewerten - forschungsdaten.info](https://www.forschungsdaten.info)

## DATENARCHIVE

- zertifizierte Repositorien und Datenarchive mit dem Core Trust Seal:  
[Certified Repositories](#)
- Anbindung der UL an (voraussichtlich Mitte 2021):  
[OpARA \(TU Dresden\)](#)
- akkreditierte Datenzentren für Verhaltens,-, Sozial- und Wirtschaftsdaten:  
[Datenzentren - RatSWD](#)

# RECHTLICHE ASPEKTE

## RECHTLICHE ASPEKTE

- **Urheberrecht:**  
kann für FD i.d.R. nicht angenommen werden vgl.: Schöpfungshöhe,  
Datenbanken – Leistungsschutzrecht, Interview – u.U. Urheberrecht an Fragenkatalog, wenn dann  
nur an der Form, nie an den Daten selbst
- **Nutzungsrechte:**  
abhängig von Dienstverhältnis, weisungsgebundene Forschung in Forschungsprojekten:  
FD gehören der Institution, an der sie entstanden sind → einfaches Nutzungsrecht
- **Hinweise:**
  - Kuschel, Linda (2018): Wem "gehören" Forschungsdaten?  
<https://www.forschung-und-lehre.de/wem-gehoeren-forschungsdaten-1013/>
  - [HeFDI - Hessische Forschungsdateninfrastrukturen. \(2021, March\). Rechtliche Rahmenbedingungen des Forschungsdatenmanagements \(6 Teile\). Zenodo.](#)

## NUTZUNGSRECHTE AN FORSCHUNGSDATEN

„Die Nutzung [der Daten] steht insbesondere der Wissenschaftlerin und dem Wissenschaftler zu, die/der sie erhebt.“

(Leitlinien zur Sicherung guter wissenschaftlicher Praxis – Kodex (DFG, 2019), Leitlinie 10: Rechtliche und ethische Rahmenbedingungen, Nutzungsrechte, S. 17.)

- Konflikt! → Aber: **Nutzungsrechte ≠ Besitz/Mitnahme** → Nutzungsrechte können durch die besitzende Institution gewährt werden, so dass die Chance besteht, den einmal begonnen wissenschaftlichen Erkenntnisprozess zu beenden.
- Lösung: **Schriftliche** Vereinbarungen mit den Betreuenden/ProjektleiterInnen treffen und Nutzungsrechte an Daten nach Beendigung einer Forschungstätigkeit oder bei vorzeitigem Ausscheiden regeln. → faire Interessenabwägung aller Seiten

## NUTZUNGSRECHTE AN FORSCHUNGSDATEN

„Die Nutzung [der Daten] steht insbesondere der Wissenschaftlerin und dem

(Leitlinien zu

➤ **Was passiert mit Ihren Daten beim (vorzeitigen) Ausscheiden aus dem Forschungsprojekt?**

nbedingungen,

→ Konflikte  
könne  
besteh  
beend

➤ **In welchen Zusammenhängen dürfen Ihre generierten Daten weiterverwendet werden?**

rechte  
e Chance  
s zu

→ Lösung  
treffen

➤ **Wer entscheidet über eine Datenveröffentlichung?**

weiterInnen

Forschungstätigkeit oder bei vorzeitigem Ausscheiden regeln.

→ faire Interessenabwägung aller Seiten!

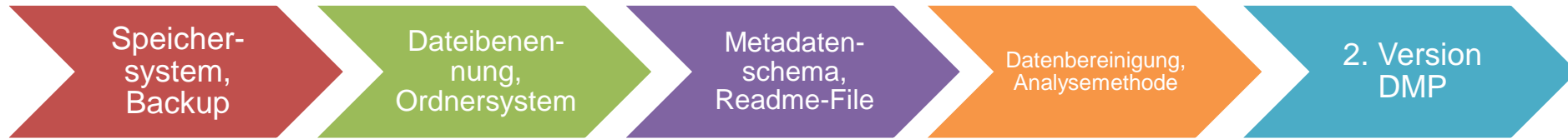


# FDM WORKFLOW

## SCHRITT 1: PROJEKTPLANUNG UND DATENERHEBUNG



## SCHRITT 2: PROJEKTDURCHFÜHRUNG – DATENDOKUMENTATION, -SPEICHERUNG UND -ANALYSE



## SCHRITT 3: PROJEKTABSCHLUSS – VERÖFFENTLICHUNG UND ARCHIVIERUNG

Anonymisie-  
rungsmaß-  
nahmen

Metadaten und  
Readme-File  
Check

Datenauswahl  
Archivierung,  
Archivierungsort

3. Version  
DMP

Rechtecheck,  
Lizenzierung,  
Veröffentlichung

**GESCHAFFT!** 😊



Quelle: Stockata (CC0): <https://stockata.de/bild/0130630-wunderkerze-funke-feuerwerk-licht.html>



UNIVERSITÄT  
LEIPZIG

# Vielen Dank!

**Kontakt: [forschungsdaten@uni-leipzig.de](mailto:forschungsdaten@uni-leipzig.de)**



Der Text dieser Präsentation steht unter der Lizenz  
Creative Commons Attribution 4.0 International (CC BY 4.0).