

## FDM Basics

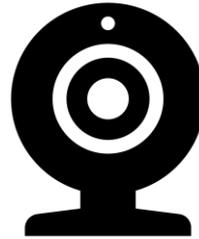
### Grundlagen des Forschungsdatenmanagements

Anja Schwarz (HS Koblenz) & Anja Nied (HS Trier),  
vertr. durch Dr. Thomas Schimmer (HS Mainz)

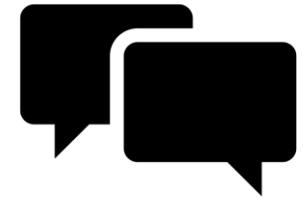
# Netiquette



Mikro  
stummschalten



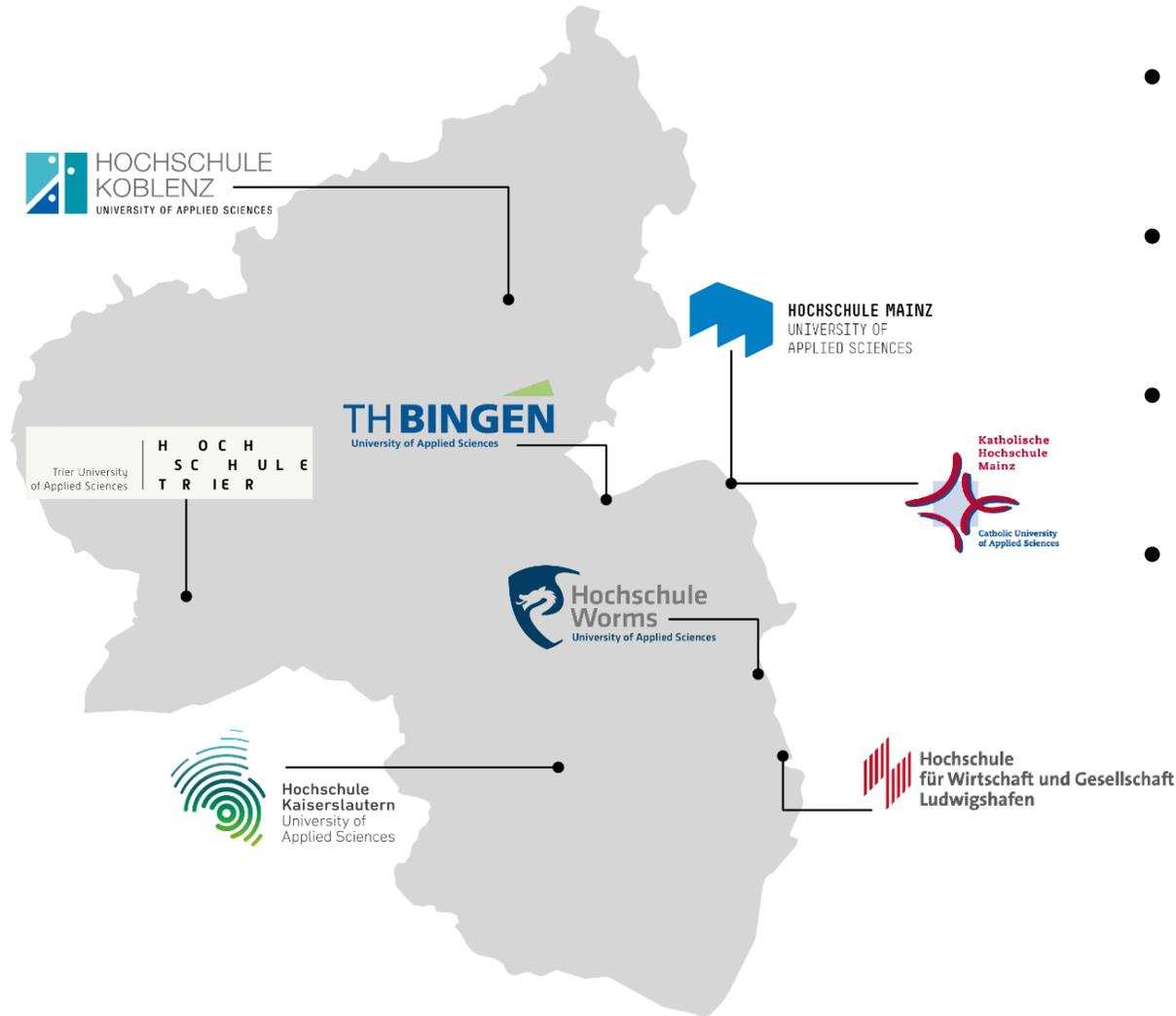
Kamera an ;-)



Fragen in den Chat

**Teilnahmebescheinigung auf Wunsch!**

# Projekt FDM@HAW.rlp



- Ziel: Nachhaltiger und qualitätsgesicherter Kompetenzaufbau zu FDM in RLP
- Verbundprojekt: 7 staatl. Hochschulen für angewandte Wissenschaften + KH Mainz
- Gefördert durch das BMBF und finanziert durch die EU
- Laufzeit 2022 - 2025

<https://www.fdmhawrlp.de/>



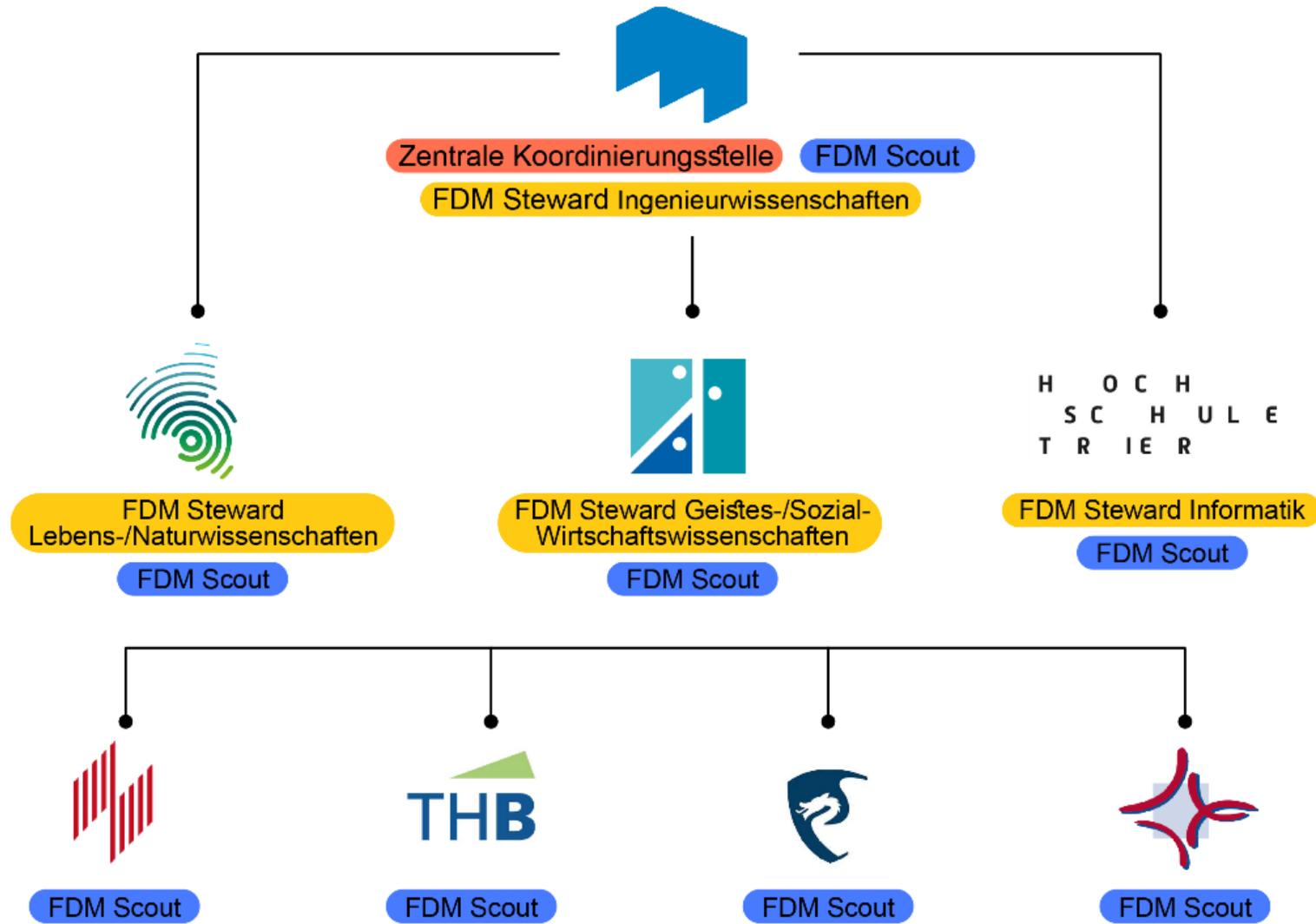
GEFÖRDERT VOM

Bundesministerium  
für Bildung  
und Forschung



Finanziert von der  
Europäischen Union  
NextGenerationEU

# Projektstruktur



# Kontakt Daten der Scouts

## Technische Hochschule Bingen

**Gery Weyd** | FDM-Scout

E: [g.weyd\(at\)th-bingen\(dot\)de](mailto:g.weyd@th-bingen.de)

## Hochschule Kaiserslautern

**Maximilian Weyland** | FDM-Scout

E: [maximilian.weyland\(at\)hs-kl\(dot\)de](mailto:maximilian.weyland@hs-kl.de)

## Hochschule Koblenz

**Anja Schwarz** | FDM-Scout

E: [schwarz1\(at\)hs-koblenz\(dot\)de](mailto:schwarz1@hs-koblenz.de)

## Hochschule für Wirtschaft und Gesellschaft Ludwigshafen

**Andrea Kuhn** | FDM-Scout

E: [Andrea.Kuhn\(at\)hwg-lu\(dot\)de](mailto:Andrea.Kuhn@hwg-lu.de)

## Hochschule Mainz

**Daniela Maier** | FDM-Scout

E: [daniela.maier\(at\)hs-mainz\(dot\)de](mailto:daniela.maier@hs-mainz.de)

## Katholische Hochschule Mainz

**Darina Schilirò** | FDM-Scout

E: [darina.schiliro\(at\)kh-mz\(dot\)de](mailto:darina.schiliro@kh-mz.de)

## Hochschule Trier

**Anja Nied** | FDM-Scout

E: [a.nied\(at\)hochschule-trier\(dot\)de](mailto:a.nied@hochschule-trier.de)

## Hochschule Worms

**Shanna Pilch** | FDM-Scout

E-Mail: [pilch\(at\)hs-worms\(dot\)de](mailto:pilch@hs-worms.de)

# Kontaktdaten der Stewards

## Ingenieurwissenschaften

**Manuela Richter** | FDM-Steward  
Hochschule Mainz  
E: [manuela.richter\(at\)hs-mainz\(dot\)de](mailto:manuela.richter(at)hs-mainz(dot)de)

## Geistes-, Sozial- und Wirtschaftswissenschaften

**Lucas Krajewski** geb. Hamel | FDM Steward  
Hochschule Koblenz  
E: [krajewski\(at\)hs-koblenz\(dot\)de](mailto:krajewski(at)hs-koblenz(dot)de)

## Informatik

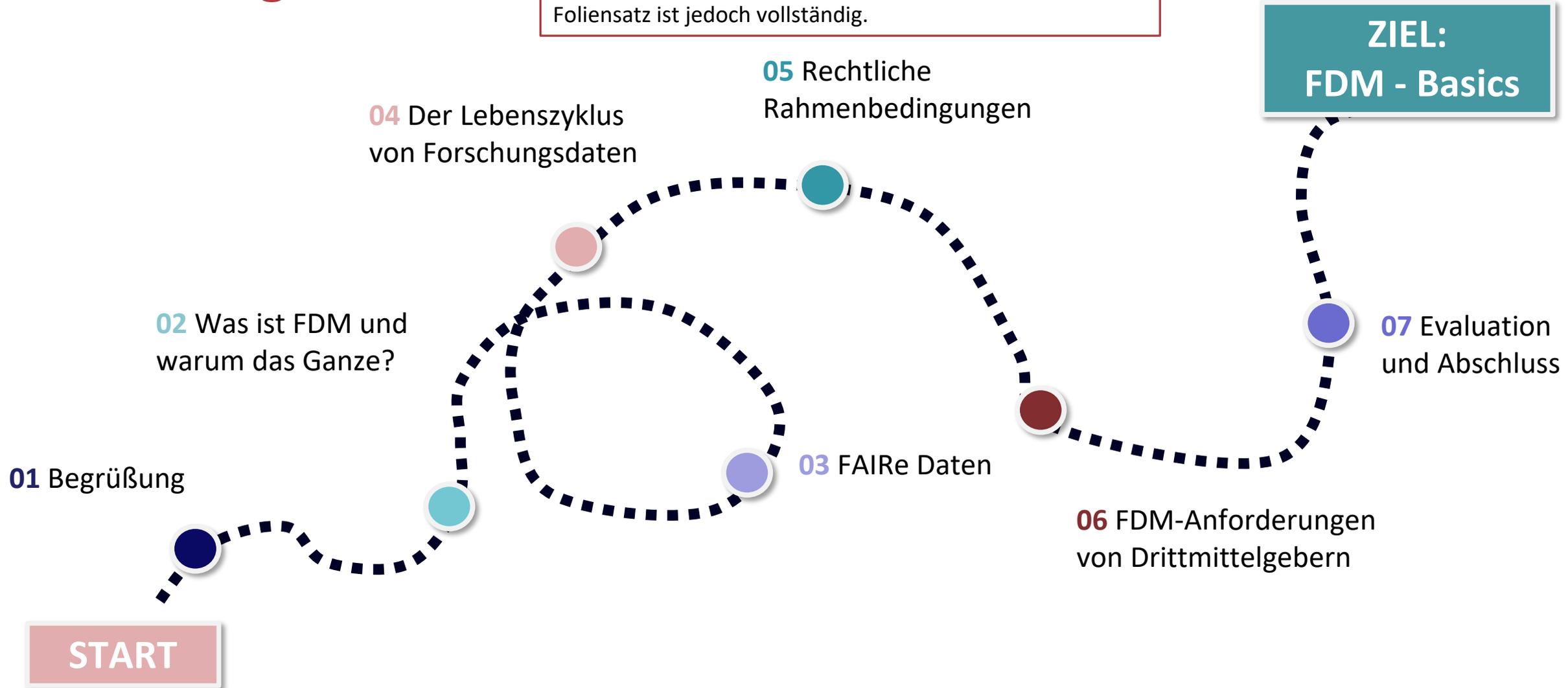
**Dr. Johannes Putzke** | FDM-Steward  
Hochschule Trier  
E: [J.Putzke\(at\)umwelt-campus\(dot\)de](mailto:J.Putzke(at)umwelt-campus(dot)de)

## Lebens- und Naturwissenschaften

**Jörg Panzer** | FDM-Steward  
Hochschule Kaiserslautern  
E: [joerg.panzer\(at\)hs-kl\(dot\)de](mailto:joerg.panzer(at)hs-kl(dot)de)

# Unser Weg heute

**Hinweis:** Dieser Foliensatz basiert auf den präsentierten Folien am 30.04.2024. Aufgrund von ausgeblendeten Folien springt die Folienzählung in diesem Dokument, der Foliensatz ist jedoch vollständig.



# Agenda

## 1. Begrüßung

2. Was ist FDM und warum das Ganze?

3. FAIRe Daten

4. Der Lebenszyklus von Forschungsdaten

5. Rechtliche Rahmenbedingungen

6. FDM-Anforderungen von Drittmittelgebern

7. Fragerunde und Abschluss

# Lernziele des Seminars

- Sie wissen, was **FDM** ist, und kennen den **Forschungsdatenlebenszyklus**.
- Sie kennen Werkzeuge, um FD gut **strukturieren und ordnen** zu können.
- Sie kennen den Nutzen und die Inhalte der **Datendokumentation**.
- Sie kennen die Voraussetzungen für ein sicheres **Backup** und die **Langzeitarchivierung**.
- Sie wissen, was ein **Repository** ist, und kennen verschiedene Publikationswege für FD.
- Sie kennen die **Creative Commons Lizenzen**.
- Sie kennen die **Rechtsbereiche**, die beim FDM von Bedeutung sein können.
- Sie kennen die **Anforderungen wichtiger Drittmittelgeber** in Bezug auf FDM.

**Sie verstehen die Sinnhaftigkeit des strukturierten Vorgehens und erkennen den Nutzen des anfänglichen Vorbereitungsaufwands.**

# Unsere Fragen an Sie:

## Welche Vorkenntnisse bringen Sie heute mit:

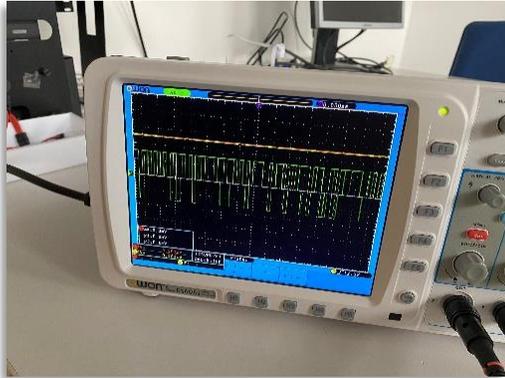
1. Ich habe den Begriff „Forschungsdatenmanagement“ (FDM) bereits gehört.
2. Ich kenne den „Forschungsdatenlebenszyklus“.
3. Ich wurde bei einem Drittmittelantrag dazu aufgefordert, Angaben zum FDM/Umgang mit Forschungsdaten zu machen.
4. Ich habe bereits einen Datenmanagementplan (DMP) erstellt.
5. Ich habe bereits Forschungsdaten veröffentlicht.



# Agenda

1. Begrüßung
- 2. Was ist FDM und warum das Ganze?**
3. FAIRe Daten
4. Der Lebenszyklus von Forschungsdaten
5. Rechtliche Rahmenbedingungen
6. FDM-Anforderungen von Drittmittelgebern
7. Fragerunde und Abschluss

# Was sind Forschungsdaten?



Messdaten/ Messreihen

Umfragen



Interview<sup>[1]</sup>

Habe nun, ach! Philosophie,  
Juristerei und Medizin,  
Und leider auch Theologie  
Durchaus studiert, mit heißem Bemühn.  
Da steh ich nun, ich armer Tor!  
Und bin so klug als wie zuvor;  
Heiße Magister, heiße Doktor gar  
Und ziehe schon an die zehen Jahr  
Herauf, herab und quer und krumm  
Meine Schüler an der Nase herum-  
Und sehe, daß wir nichts wissen können!  
Das will mir schier das Herz verbrennen.

Texte



Labordaten



Bilder/Modellierungen<sup>[2]</sup>



Video<sup>[3]</sup>

...

...und vieles mehr

# Was ist Forschungsdatenmanagement (FDM)?

planen

aufbereiten

verwalten

strukturieren



dokumentieren

archivieren

auswählen

publizieren

[1]



**FDM:** Überlegter und systematisierter Umgang mit Daten von der Planung bis zum Abschluss eines Projektes und w.m. darüber hinaus.

# Warum das Ganze? Mehrwert für Forschende



Überblick behalten



Zusammenarbeit erleichtern



Qualität wissenschaftlicher Arbeit gewährleisten



Zeit und Nerven sparen

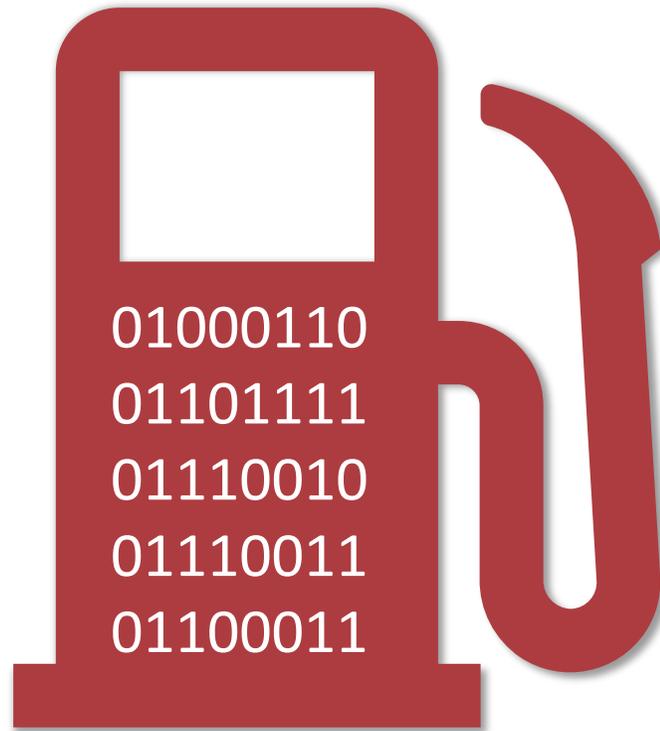


formale Anforderungen erfüllen

[1]

# Warum das Ganze? (Wissenschafts)Politische Dimension

„Daten sind das neue Öl“ [1]  
(Clive Humby, Mathematiker)



## Open Science Policy der EU

A screenshot of the European Commission website showing the Open Science policy page. The page features the European Commission logo, a search bar, and a navigation menu. The main content area is titled "Open Science" and includes a brief description: "An approach to the scientific process that focuses on spreading research results using digital and collaborative technology. Expert groups,..." Below this, there is a "PAGE CONTENTS" section with links to "The EU's open science policy", "8 ambitions of the EU's open science policy", "Future of open science under Horizon Europe", and "Tracking open research trends - Open Science Monitor". A "Documents" link is also visible. On the right side, there is a blue banner for "European Open Science Cloud - EU Node" with a navigation menu and a white box containing the text "EOSC EU Node" and "A European platform and information gateway to explore, engage, and enrich your research collaborations."

# Prinzipien von Open Science



# Agenda

1. Begrüßung
2. Was ist FDM und warum das Ganze?
- 3. FAIRe Daten**
4. Der Lebenszyklus von Forschungsdaten
5. Rechtliche Rahmenbedingungen
6. FDM-Anforderungen von Drittmittelgebern
7. Fragerunde und Abschluss

# FAIR-Prinzipien

**F**indable



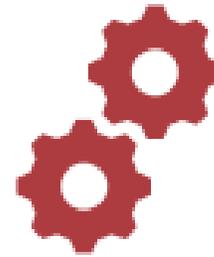
Auffindbar

**A**ccessible



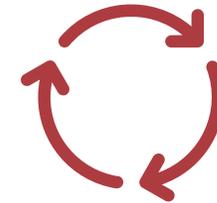
Zugänglich

**I**nteroperable



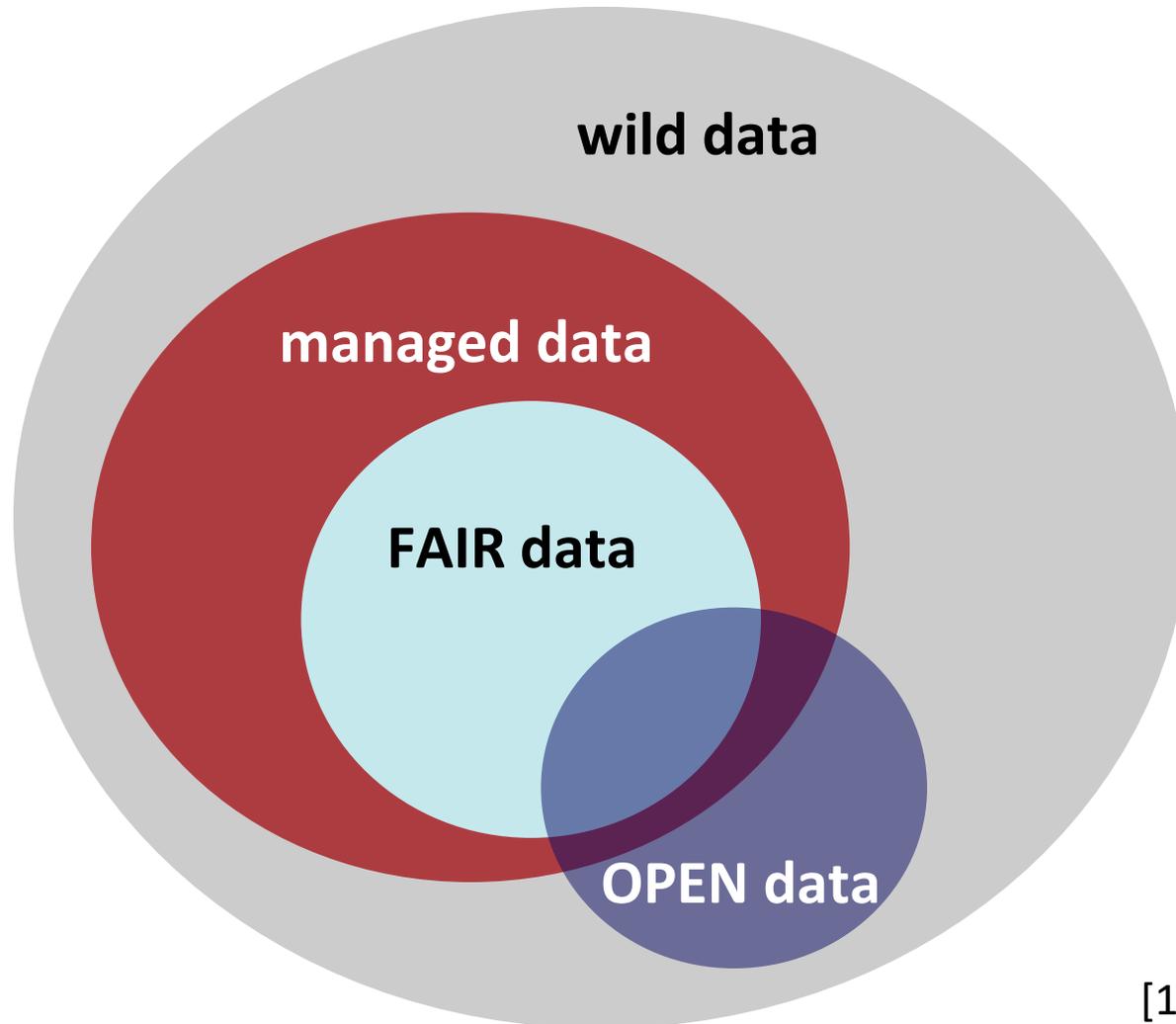
Interoperabel

**R**eusable



Wiederverwendbar

# Arten von Forschungsdaten



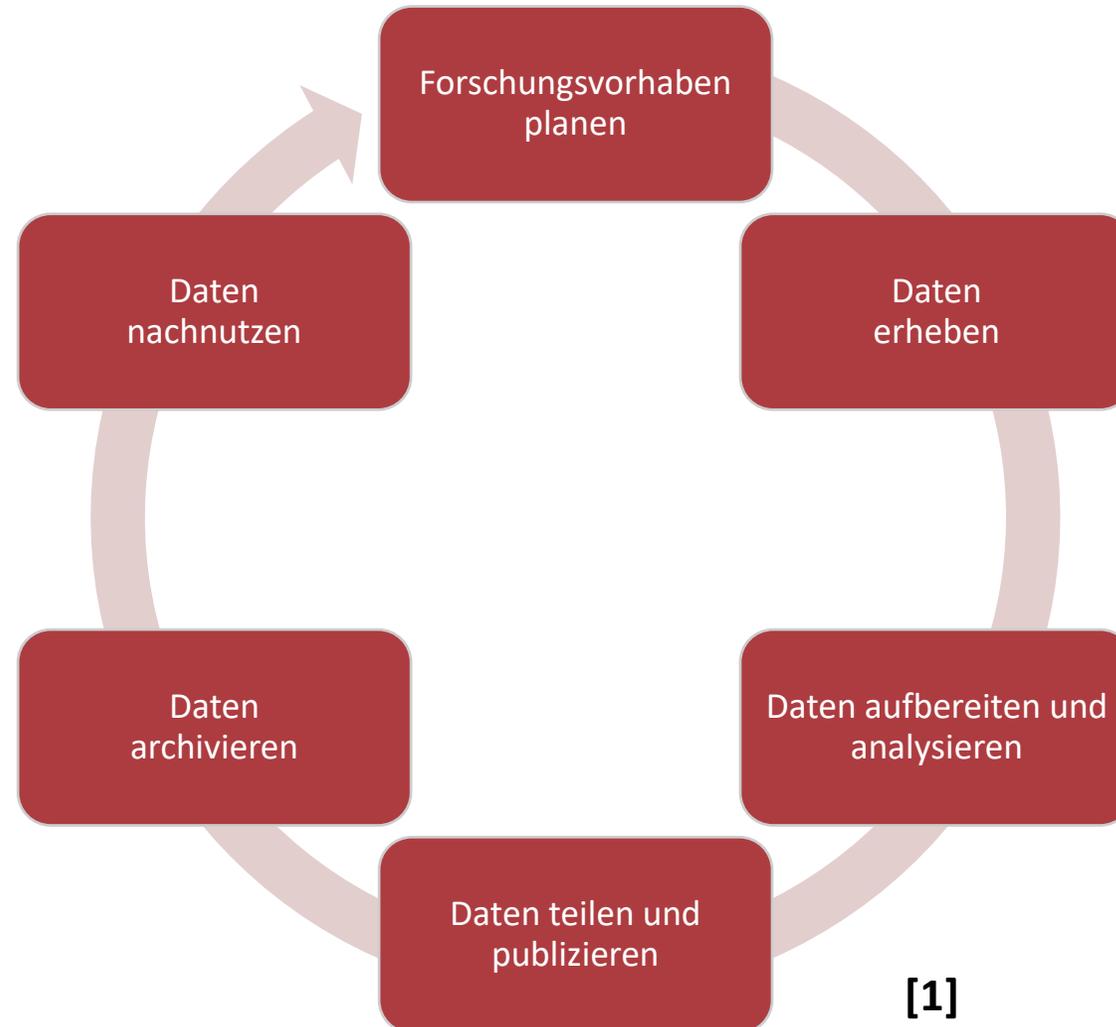
[1]

- FAIR  $\neq$  Open Data
- “Open Data” nicht immer möglich
- „As open as possible, as closed as necessary”  
(EU Open Science-Policy)
- Mehrheit der (FAIR-)Daten soll zukünftig „open“ sein

# Agenda

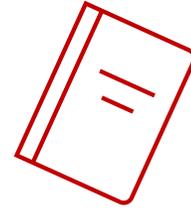
1. Begrüßung
2. Was ist FDM und warum das Ganze?
3. FAIRe Daten
- 4. Der Lebenszyklus von Forschungsdaten**
5. Rechtliche Rahmenbedingungen
6. FDM-Anforderungen von Drittmittelgebern
7. Fragerunde und Abschluss

# Forschungsdatenlebenszyklus



# 4. Der Lebenszyklus von Forschungsdaten – Inhalte

## 4.1. Planung des Forschungsvorhabens



## 4.2. Datenorganisation und -dokumentation

## 4.3. Datensicherung und -archivierung

## 4.4. Datenpublikation

# 1. Planung des Forschungsvorhabens (I)

Welche **Daten** sollen erhoben werden?

Wie organisiere ich das **Backup** meiner Daten?

Wie sollen die Daten im Projekt **geteilt** werden?

Welche **Software und Datenformate** möchte ich verwenden?

Wo sollen die Daten **gespeichert** werden?

Gibt es bereits Daten, die ich **nachnutzen** kann?

# 1. Planung des Forschungsvorhabens (II)

Welche **Vorgaben**  
seitens der  
**Drittmittelgeber**  
gibt es?

Muss/soll ein **DMP**  
erstellt werden?

**Welche Kosten entstehen** ggf. für  
Nachnutzung, Speicherung oder  
Archivierung? Können diese  
erstattet werden?

Welche **rechtlichen**  
**Rahmenbedingungen**  
sind relevant?

Müssen **rechtl.**  
**Vereinbarungen**  
getroffen werden?

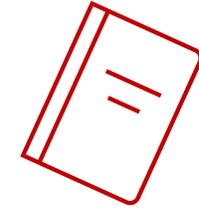
# Exkurs: Datenmanagementplan (DMP)



## Ein Datenmanagementplan (DMP) ...

- ... schafft eine verbindliche Grundlage für einen einheitlichen Umgang mit Daten im Forschungsprozess;
- ... erleichtert die Abstimmung zwischen Projektpartnern;
- ... hilft bei der frühzeitigen Erkennung von Problemen und Lösungsfindung;
- ... legt Verantwortlichkeiten fest;
- ... regelt Zugriffsrechte;
- ... beugt Datenduplikaten, Datenverlust und Sicherheitslücken vor;
- ... kann mittels Checklisten und online-Tools (z.B. RDMO) erstellt werden.

# Exkurs: Datenmanagementplan (II)



## Inhalte:

- Allgemeine **Informationen zum Projekt und den Projektzielen**
- Beschreibung **bereits vorhandener Daten**, die eventuell nachgenutzt werden können
- Beschreibung der **zu erhebenden Daten**, einschließlich **Datenmenge** und **Format**
- Angaben zur geplanten **Verwaltung und Speicherung** der Daten sowie zur **Archivierung**
- Angaben zur Erstellung der **Metadaten**
- Informationen zur **Behandlung von administrativen und rechtlichen Aspekten** (z.B. Vorgaben der mittelgebenden Einrichtungen, Zuständigkeiten, Datenschutzaspekte, Zugriffsrechte etc.)
- Angaben zu den **benötigten Ressourcen**, die zur Umsetzung des DMPs erforderlich sind

# 4. Der Lebenszyklus von Forschungsdaten

4.1. Planung des Forschungsvorhabens

**4.2. Datenorganisation und -dokumentation**

4.3. Datensicherung und -archivierung

4.4. Datenpublikation

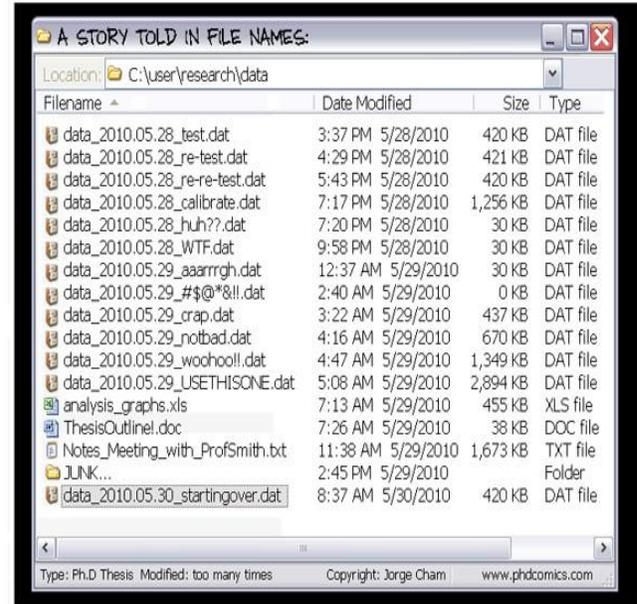


# 4.2.1. Datenorganisation

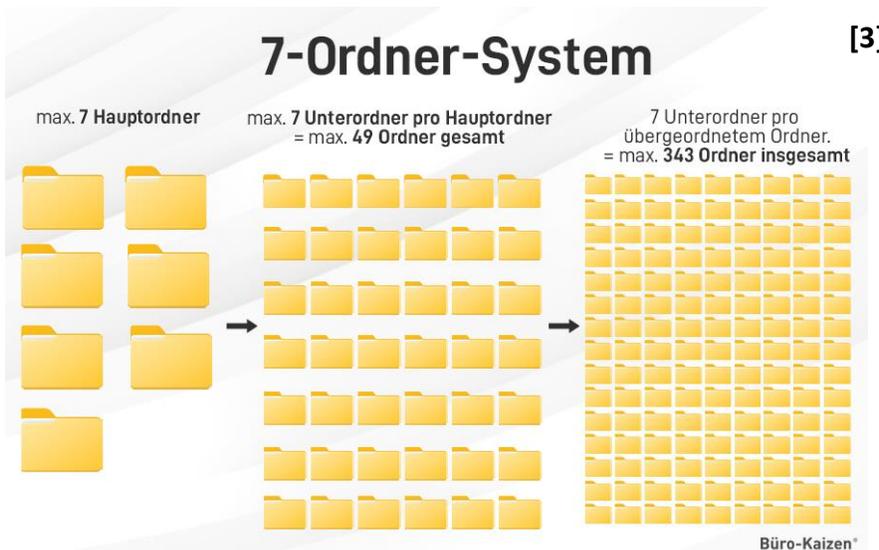
Ordnerstruktur

Dateibenennung

Versionierung

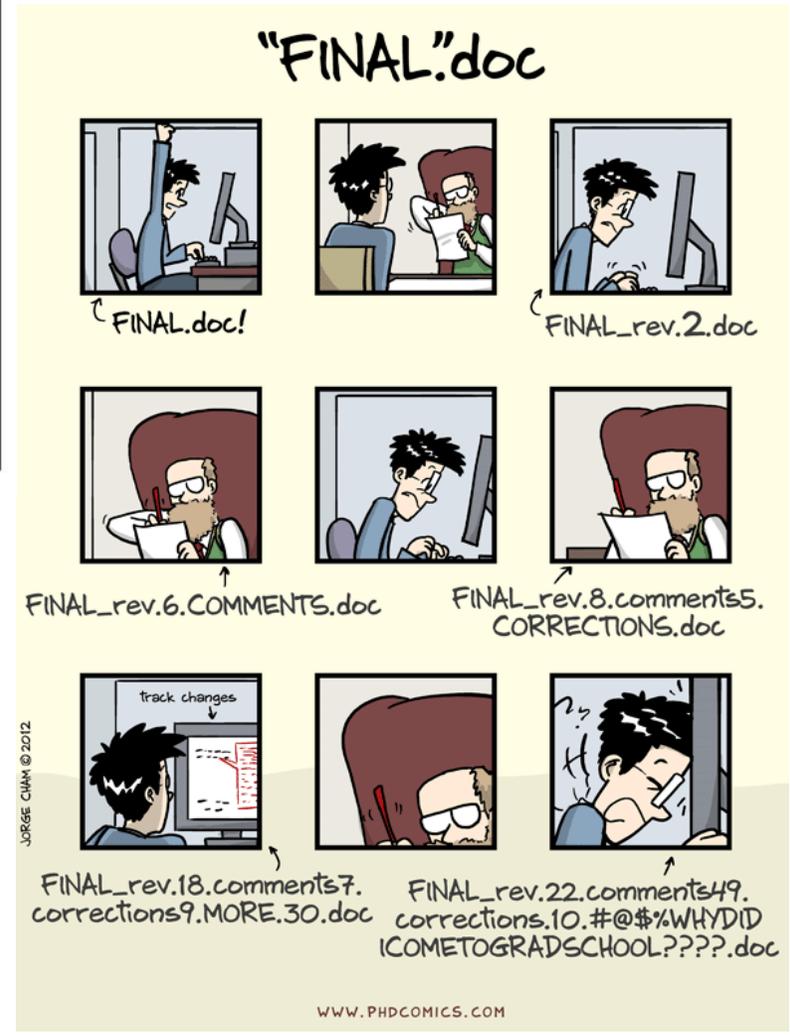


[1]



[3]

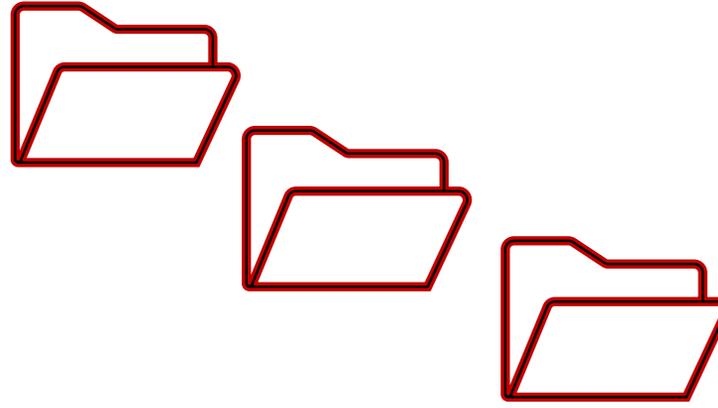
[2]



# Ordnerstruktur

Eine gute Ordnerstruktur...

... ist hierarchisch gegliedert;  
... baut inhaltlich aufeinander auf.

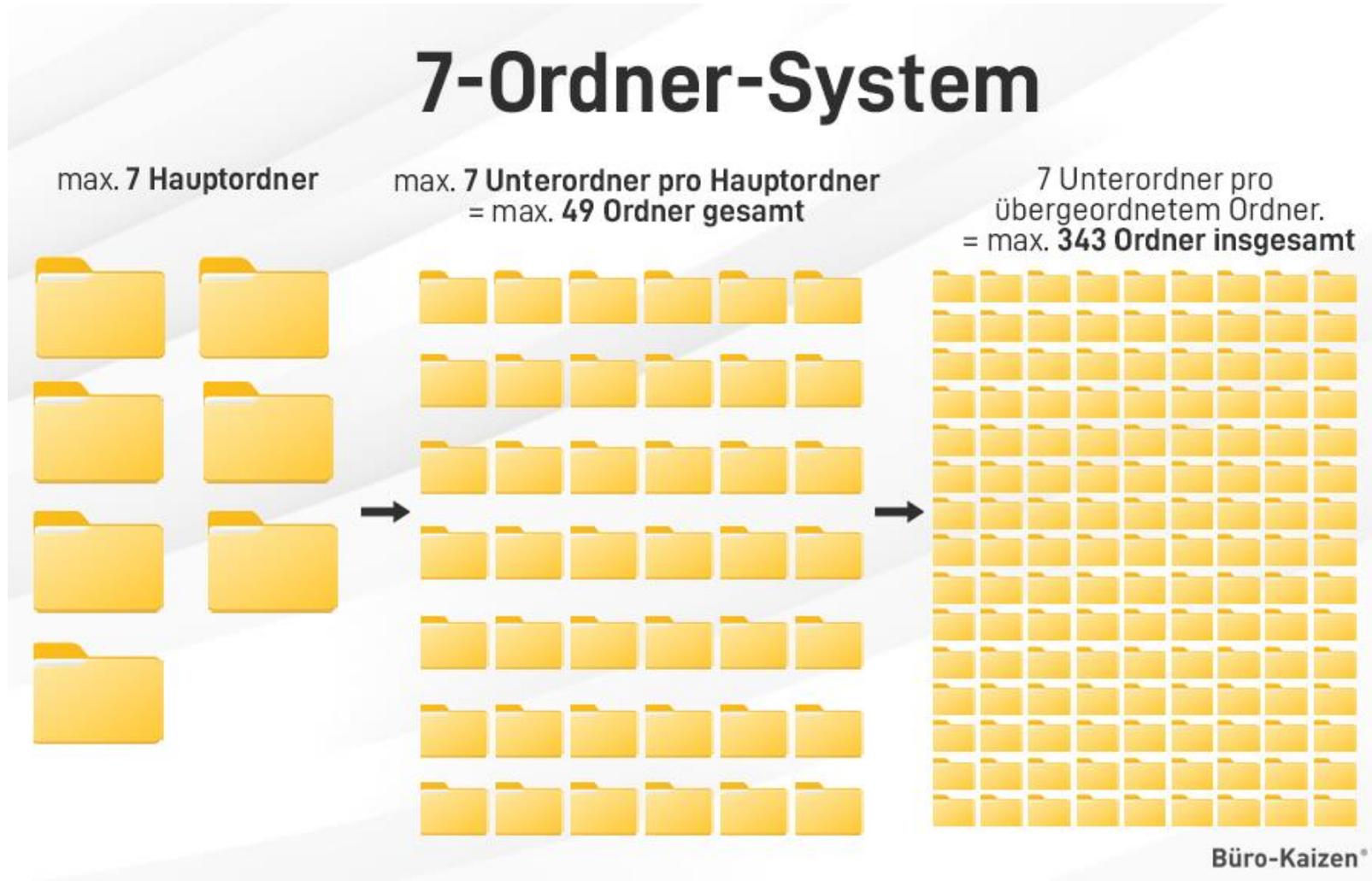


**Ordnerstrukturen auf Servern sollten möglichst identisch mit denen auf lokalen Rechnern sein.**



**Abgestimmte Benennungskonventionen etc. schriftlich festhalten (z.B. Readme-Datei oder DMP)!**

# Beispiel: 7-Ordner-3-Ebenen-System



# Beispiele Verzeichnisstruktur:

01\_Organisatorisches

02\_Literatur

03\_Software

04\_Messdaten

05\_Analyse\_Auswertung

06\_Bildmaterial

07\_Paper

01\_Verwaltung

02\_Kommunikation

03\_Forschungsdaten

04\_Auswertung

05\_Dokumente

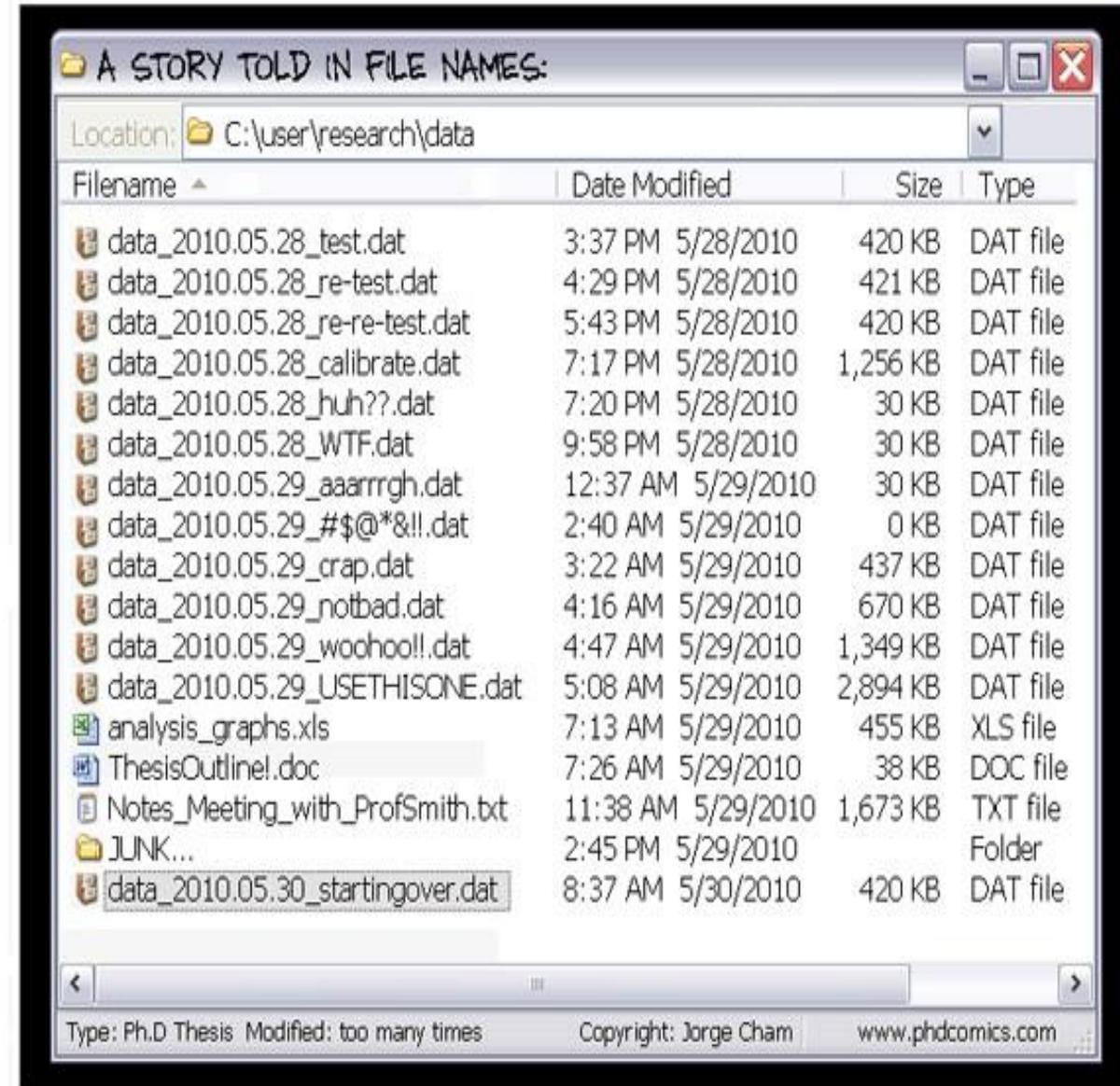
06\_Abschluss

07\_Sonstiges

# Dateibenennung

Dateien und Ordner systematisch benennen und ordnen, damit...

- ...sie jetzt und in Zukunft leicht auffindbar und zugänglich sind.
- ...längeres Suchen oder das Vergleichen verschiedener Versionen vermieden wird.
- ...Änderungen nachvollziehbar sind.
- ...die Dateien nicht versehentlich gelöscht oder überschrieben werden.



# Eine klare Dateibenennung ...

- ... ist objektiv, intuitiv und personenunabhängig nachvollziehbar;
- ... ist einheitlich & konsistent;
- ... ist eindeutig und lässt Rückschlüsse über den Inhalt zu;
- ... ist so lang wie nötig und so kurz wie möglich (zulässige Pfadlänge Windows 10: 256 Zeichen);
- ... vermeidet Leerzeichen und Sonderzeichen.

# Mögliche Namensbestandteile bei der Dateibenennung

- Inhaltsstichwort oder Kurztitel
- Name oder Kürzel des Erstellers
- Erstellungsdatum (z.B. JJMMTT oder JJJ-MM-TT)
- Bearbeitungsdatum
- Bezeichnung der Arbeitsgruppe / des Projekts
- Publikationsdatum
- Projektnummer
- Versionsnummer (z.B. V01 oder V\_01)
- Anhängsel für Software, mit der die Datei erstellt wurde (z.B. .xls/.xlsx für Excel; .txt für Textdateien etc.)

# Beispiele für einheitliche Namensgebung

Für chronologische Sortierung mit Datumsangabe beginnen:

z.B.: JJJJMMTT\_Name

- 20160512\_Klimamessung1\_original.jpg
- 20160522\_Klimamessung1\_MHU\_Ausschnitt.jpg
- 20160523\_Klimamessung1\_MHU\_Ausschnitt\_bearbeitet\_Farbe.jpg



**Automatisch generierte Namen (z.B. von Digitalkameras) vermeiden**



**Skalierbarkeit berücksichtigen: 01 oder 001**

# Versionierung im Dateinamen

- anhand natürlicher Zahlen (v1, v2, v3)
- Unterscheidung zwischen größeren und kleineren Änderungen

**Beispieldatei\_v1 -> Beispieldatei\_v2**  
**Beispieldatei\_v1\_1 -> Beispieldatei\_v1\_2**

- Verwirrende Zusatzbezeichnungen vermeiden;
- Dateinamen konstant beibehalten



# Weitere Hinweise zur Versionierung

- Jede Änderung vermerken (insbesondere in größeren Teams);
- Veraltete Versionen getrennt speichern, verwerfen oder löschen;
- Die Originaldatei sollte stets erhalten bleiben;
- Bei großen Projekten bzw. umfangreichen Änderungen: Versionskontroll-Software zur Verwaltung von Versionen nutzen (z.B. Git, Subversion);
- automatische Versionierung und Änderungsverfolgung bei kollaborativen Dokumenten und Speicherorten (z.B. Wiki, Google Docs, Cloud-Dienste)

## 4.2.2. Datendokumentation

### Metadaten

- Was sind Metadaten?
- Wozu dienen Metadaten?
- Fachspezifische Metadaten-Standards

Was muss alles dokumentiert werden?

Alles, was Dritte zum Verständnis des Datensatzes benötigen!



# Metadaten...

- ... beschreiben oder definieren andere Daten näher (z.B. Datensätze od. Code).
- ... enthalten technische, rechtliche, administrative, fachliche und deskriptive Informationen (Datenvolumen, Datenformate, Lizenzen, Autor:in, Titel, Thema, Kurzbeschreibung, Schlagworte etc.).
- ... werden unabhängig von oder zusammen mit den Daten, die sie beschreiben, gespeichert oder verknüpft.

rich metadata



poor metadata



[1]

# Wozu dienen Metadaten?

## Datendokumentation ...

... macht Daten und Medien auffindbar und nachvollziehbar (z.B. in Datenbanken/Suchmaschinen);

... erleichtert die maschinelle und automatisierte Verarbeitung von Daten;

... erleichtert die Nachnutzung (auch die eigene) und ermöglicht die Reproduzierbarkeit;



**Gut dokumentierte Daten werden häufiger zitiert und genutzt!\***

# Metadaten-Standards

- ... gewährleisten die Interoperabilität von Metadaten.
- Für fachspezifische Metadaten existiert eine Vielzahl von Standards, z.B. **Dublin Core, DataCite** usw.
- Repositorien haben meist eigene Standards

Eine **Übersicht über existierende Standards:**

[Metadata Standards Catalog](#) der RDA und die Seite der [RDA Metadata Standards Directory Working Group](#), [FairSharing.org](#) oder [DDC \(Digital Curation Centre\)](#). [Tools für die Erstellung von Metadaten](#) der [Research Data Alliance](#)

# Formate für die Datendokumentation:

## Die Dokumentation kann z.B. erfolgen ...

- ... in einer begleitenden ReadMe-Datei;
- ... in einer Metadatenbank;
- ... in einem projektinternen Wiki;
- ... in einem [\(elektronischen\) Laborbuch](#);
- ... in einem [Datenmanagementplan \(DMP\)](#);
- ... innerhalb der Ordnerstruktur und Dateibenennung (z.B. Autor, Titel, Thema, Schlagworte);

# 4. Forschungsdaten im Lebenszyklus

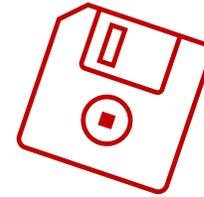
4.1. Planung des Forschungsvorhabens

4.2. Datenorganisation und -dokumentation

**4.3. Datensicherung und -archivierung**

4.4. Datenpublikation

4.5. Rechtliche Rahmenbedingungen



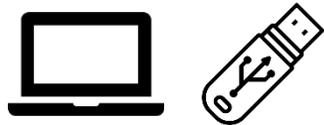
# Datensicherung

- Regelmäßige **Backups** (Nutzung von Hochschulservern inkl. Backup-Systeme; festen Zeitpunkt festlegen)
- Sicherungskopien **getrennt von Originalen** aufbewahren
- Softwareanwendungen in die Sicherungsmaßnahmen einbeziehen
- mehrere **Speichermedien** verwenden
- **Datenschutz** berücksichtigen (Pseudonymisierung und Anonymisierung vertraulicher Daten) => ggf. Datenschutzbeauftragten hinzuziehen
- **Zugriffsrechte** klären bzw. Zugriff sichern (z.B. durch Passwortschutz)

# Die 3-2-1-Regel



3 Kopien der Datei



2 unterschiedliche Medien



1 externe Speicherung

## 4.2. Backup, Publikation, Archivierung: Ein Vergleich

	Backup	Publikation	Archivierung
Ziel	Aktive Nutzung	Teilen der Daten	Sichern finaler Daten
Zeitraum	kurzzeitig (z.B. Analysezeitraum, Projektlaufzeit)	Offen	langfristig bzw. dauerhaft (10 Jahre DFG-Vorgabe)
Metadaten	Hauptsächlich Zugriffssteuerungsliste (ACL) und Dateieigenschaften, Versionierung	FAIR und nach Vorgaben des Repositoriums	datenspezifisch, standardisiert

# Besonderheiten der Langzeitarchivierung

## Speicherung „für immer“

- Gut überlegen, was wirklich für die Nachwelt zur Verfügung stehen sollte
- Verwendung von Formaten möglichst auf Basis offener, dokumentierter und nicht proprietärer Standards
- Vermeidung von Kompression
- Migration der Daten in aktuelle Datenformate
- Nutzung professioneller Archive

# LZA – Nachhaltige Dateiformate

Datentyp	empfohlene Formate	weniger geeignete/ ungeeignete Formate
Audio	*.wav   *.flac	*.mp3
Computer-aided Design (CAD)	*.dwg   *.dxf   *.x3d   *.x3db   *.x3dv	-
Datenbanken	*.sql   *.xml	*.accdb   *.mdb
Rastergrafiken & Bilder	*.tif (unkomprimiert)   *.jp2   *.jpg2   *.png	*.gif   *.jpeg   *.jpg   *.psd
statistische Daten	*.por	*.sav (IBM®SPSS)
Tabellen	*.csv   *.tsv   *.tab	*.xls   *.xlsx   *.xlx
Texte	*.odf   *.rtf   *.txt   PDF/A	*.docx   *.doc   PDF
Vektorgrafiken	*.svg   *.svgz	*.cdr
Video	*.mp4   *.mkv   *.mj2   *.avi (unkomprimiert)	*.mov   *.wmv

# 4. Forschungsdaten im Lebenszyklus

4.1. Planung des Forschungsvorhabens

4.2. Datenorganisation und -dokumentation

4.3. Datensicherung und Langzeitarchivierung

**4.4. Datenpublikation**



## 4.4. Datenpublikation - Inhalte

4.4.1. **Publikationswege** für Forschungsdaten

4.4.2. **Persistierende Identifikatoren** (DOI; ORCID)

4.4.3. **Vergabe von Nutzungsrechten** am Beispiel von Creative Commons-Lizenzen

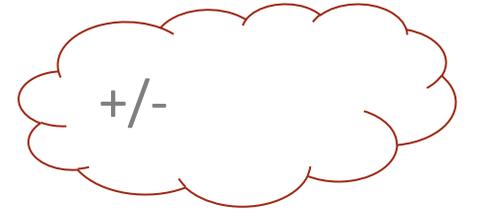
## 4.4.1. Publikationswege für Forschungsdaten



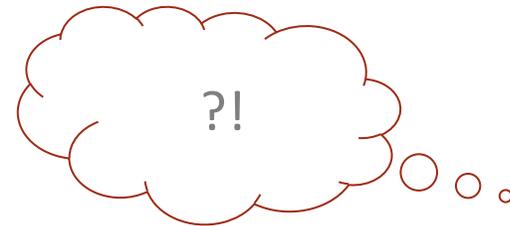
Publikation von Forschungsdaten **als eigenständiges Informationsobjekt** in einem Repository



Publikation der Datenbeschreibung in einem **Data Journal**



Publikation von Forschungsdaten **als Supplement** zur Veröffentlichung der Ergebnisse in Fachartikeln bei einem Verlag

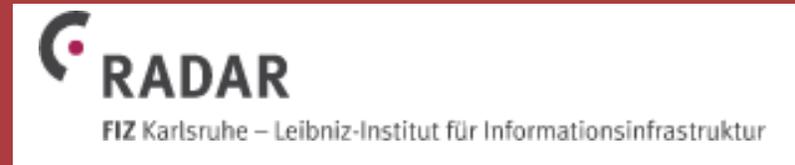


# 1010 Publikation in einem Repository

**Repositorien = Speicherorte für digitale Objekte**

## Arten von Repositorien:

- Institutionell (z.B. einige Universitäten)
- Fachspezifisch (z.B. SowiDataNet, gesis, DARIAH-DE etc.)
- Disziplinübergreifend bzw. generisch (ZENODO, RADAR, GitHub, Figshare)



# Wie finde ich ein Repository?

- Weltweit: [www.re3data.org](http://www.re3data.org)
- In Deutschland: [risources.dfg.de](http://risources.dfg.de),  
[www.nfdi.de](http://www.nfdi.de) (über Konsortium),  
[fairsharing.org](http://fairsharing.org)
- Open Access Repositorien:  
<https://v2.sherpa.ac.uk/opensoar/>

## Auswahlkriterien:

- Archivierung mind. 10 Jahre
- Automatische Vergabe von persistierenden Identifikatoren (z.B. DOI)
- Zertifizierung (z.B. Core Trust Seal)
- Klare Nutzungsbedingungen (Entgelte, Embargofristen etc.)
- Metadatenstandards

Link:

[Checkliste. So erkennen Sie ein FAIRes Datenrepository](#)



# Publikation in Data Journals

- Datenbeschreibung und Dokumentation der in einem Repository veröffentlichten Daten
- Qualitätssicherung durch Peer-Review-Verfahren
- Open-Access (z.B. Data in Brief) und Closed-Access-Zeitschriften
- Meist gebührenpflichtig

**Link:**

[Eine Liste mit Data Journals](#)

# Publikation als Supplement

- Unterstützung und Verdeutlichung der Forschungsergebnisse
- Meist aggregierte Daten (Bilder, Tabellen etc.)

## Link:

[Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B \(2020\) The citation advantage of linking publications to research data. PLoS ONE 15\(4\): e0230416. <https://doi.org/10.1371/journal.pone.0230416>](https://doi.org/10.1371/journal.pone.0230416)

## Abstract

Introduction

Materials and methods

Results

Discussion

Conclusion

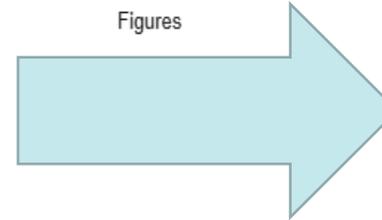
Data and code availability

Acknowledgments

References

Reader Comments

Figures



**Citation:** Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. <https://doi.org/10.1371/journal.pone.0230416>

**Editor:** Jelte M. Wicherts, Tilburg University, NETHERLANDS

**Received:** July 5, 2019; **Accepted:** February 28, 2020; **Published:** April 22, 2020

**Copyright:** © 2020 Colavizza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** Code and data can be found at: <https://doi.org/10.5281/zenodo.3470062>.

**Funding:** This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1 and by Macmillan Education Ltd, part of Springer Nature, through grant RG92108 "Effect of data sharing policies on articles' citation counts" granted to BM. Springer Nature provided support in the form of salaries for author IH, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

**Competing interests:** One of the authors (IH) is at the time of publication in the journal, employed by PLOS, publisher of PLOS ONE. IH was employed by Springer Nature, publisher of the BMC journals, at the time of planning and conducting the research and writing of the original manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents, products in development or marketed products associated with this research to declare. All other authors have declared that no other competing interests exist.

**Publisher's Note:** The article involves the independent analysis of data from

[1]

## 4.4.2. Persistierende Identifikatoren

### Digital Object Identifier – DOIs...

- ...dienen der dauerhaften und eindeutigen Referenzierbarkeit von Forschungsdaten;
- ... machen Daten zitierfähig;
- ... werden von Registrierungsagenturen (z.B. DataCite) über Verlage oder Bibliotheken vergeben

⇒ Für Autor:innen kostenfrei!

### Open Research and Contributors ID - ORCID

- eindeutige und dauerhafte 16-stellige Identifizierungsnummer für Autor:innen / Forschende
- Publikationen können einer Person eindeutig zugeordnet werden – auch bei Namensgleichheit

Registrierung: <https://orcid.org/register>

## 4.4.3. Vergabe von Nutzungsrechten – CC-Lizenzen



Namensnennung



Namensnennung, Weitergabe unter gleichen Bedingungen

Empfohlen im Sinne von  
Open Science



Namensnennung, keine Bearbeitungen



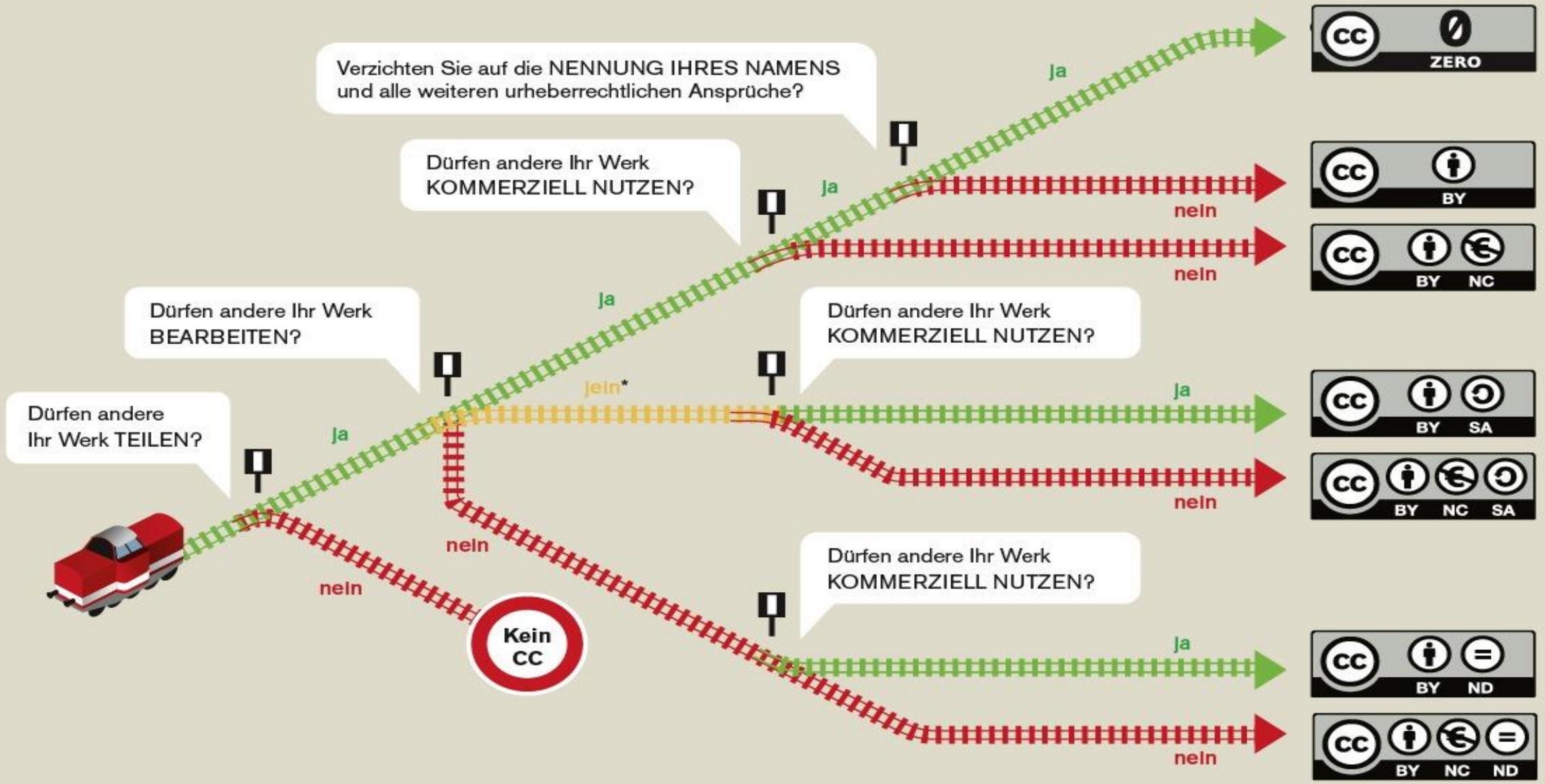
Namensnennung, nicht kommerziell



Namensnennung, nicht kommerziell, Weitergabe unter gleichen Bedingungen



Namensnennung, nicht kommerziell, keine Bearbeitungen



\*jein = ja, solange andere die Bearbeitung unter denselben Bedingungen weitergeben

# Vorteile einer Publikation von Forschungsdaten

- ✓ Ermöglicht die Nachnutzung Ihrer Daten durch andere Wissenschaftler:innen
- ✓ Ermöglicht die Reproduzierbarkeit Ihrer Forschungsergebnisse
- ✓ Stärkung des wissenschaftlichen Austauschs
- ✓ Keine Redundanzen in der Datenerhebung
- ✓ Erhöhung der Glaubwürdigkeit der Forschungsergebnisse durch Verfügbarkeit der Primärdaten

# Agenda

1. Begrüßung
2. Was ist FDM und warum das Ganze?
3. FAIRe Daten
4. Der Lebenszyklus von Forschungsdaten
- 5. Rechtliche Rahmenbedingungen**
6. FDM-Anforderungen von Drittmittelgebern
7. Fragerunde und Abschluss

# Rechtliche Rahmenbedingungen - Inhalt

5.1. Datenschutz & Persönlichkeitsrecht

5.2. Urheberrecht

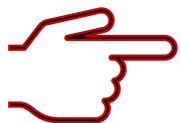
5.3. Weitere relevante Rechtsbereiche

5.4. Praktische Hinweise

# 5.1. Datenschutz & Persönlichkeitsrecht (I)

## Personenbezogene und sensible personenbezogene Daten:

- Europäische Datenschutz-Grundverordnung (GDPR), Bundes- und Landesdatenschutzgesetze, Landeshochschulgesetze (sofern vorhanden)
- ggf. weitere Bestimmungen z.B. im Sozialgesetzbuch, Genetikgesetz, Arzneimittelgesetz, Schulgesetz u.a.m.
- ggf. Richtlinien und Policies auf institutioneller Ebene (z.B. Ethikrichtlinie; FDM-Policy o.ä.)
- Für Mitarbeitende (auch Hiwis) der KH gilt Kirchliches Datenschutzrecht ([KDG](#))



**Grundsätzlich gelten für Forschende die „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ der Deutschen Forschungsgemeinschaft (DFG).**

## 5.1. Datenschutz & Persönlichkeitsrecht (II)

### Rechtsicherer Umgang mit personenbezogenen Daten:

- schriftliche, **informierte Einwilligung** der Betroffenen mit Widerrufsrecht einholen;
- **Anonymisieren** der erhobenen Daten (Pseudonymisierung reicht nicht aus!);
- Zugriffsbeschränkungen mittels **Passwortschutz** und **Vergabe von Zugriffsrechten**;
- Personenbezogene Daten von EU-Bürgern **nur bei entsprechenden Datenschutzgarantien auf Servern außerhalb der EU** speichern!



**Kontakt zum Datenschutzbeauftragten aufnehmen!**

## 5.2. Urheberrecht (UrhG)

### Was ist urheberrechtlich geschützt und was nicht?

- Bestimmte geistige Schöpfungen und Leistungen: “**Werke der Literatur, Wissenschaft und Kunst**” unterliegen dem Urheberrecht ( § 2 Abs. 1 UrhG).
- **Einzelne Forschungsdaten** (z.B. unstrukturierte Messdaten; maschinell erzeugte Daten etc.) unterliegen nur selten dem Urheberrecht, weil die dort geforderte Schaffungshöhe nicht erreicht wird ( § 2, Abs. 2 UrhG).
- Urheberrechtsschutz besteht aber für eigens **entwickelte Software oder Bild- und Tonmaterialien**.
- Bei Daten, die in einer **Datenbank** abgelegt sind, greift das Leistungsschutzrecht der Datenbankherstellenden ( § 87b UrhG). Datensammlungen sind somit ebenfalls geschützt, sofern diese „systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind und deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umfang wesentliche Investition erfordert“.

## 5.3. Weitere relevante Rechtsbereiche

### Leistungsschutzrecht

- schützt den- oder diejenige, der oder die die Investition getätigt hat, um die Daten aufzubereiten. Die reine Erstellung der Daten ist allerdings dadurch nicht geschützt.
- Eine im Zusammenhang mit Forschungsdaten vergebene Open Content-Lizenz (z.B. CC-Lizenz) ist möglicherweise wirkungslos, wenn das Urheberrecht nicht greift. Geschützt sind Auswertungen, Abbildungen etc. die auf Grundlage der Daten angefertigt wurden. Mit der aktuellen Creative-Commons-Lizenz 4.0 können nun auch Datenbanken lizenziert werden.

### Gewerbliche Schutzrechte / Arbeits- bzw. Dienstrecht

- z.B. Arbeitsverträge, die nicht den Forschenden, sondern der Institution an der sie tätig sind, das Nutzungsrecht an den Daten einräumen.
- vertragliche Regelungen zu Nutzungsrechten in Förderrichtlinien von Geldgebern oder Richtlinien der beteiligten Gruppen, Abteilungen oder Institutionen

### Vertragsrecht

- Rechte aus bestehenden vertraglichen Vereinbarungen, z.B. Verträge mit Fördermittelgebern und/oder wissenschaftlichen oder wirtschaftlichen Kooperationspartnern; bestehen Absprachen zum geistigen Eigentum an FD?

## 5.4. Praktische Hinweise

- Fremd-Urheberrecht an Sekundärdaten, Software und Datenbanken beachten, die man nicht selbst erstellt hat;
- Verwertungsrechte für zuvor schon publizierte Quellen in den FD beachten;
- Bildrechte (bei Bildquellen)
- Unterschiede Urheberrecht (D) und Copyright (USA, GB etc.) beachten;
- Daten aus Auftragsforschung: wem gehören die Forschungsdaten?

# Agenda

1. Begrüßung
2. Was ist FDM und warum das Ganze?
3. FAIRe Daten
4. Der Lebenszyklus von Forschungsdaten
5. Rechtliche Rahmenbedingungen
- 6. FDM-Anforderungen von Drittmittelgebern**
7. Fragerunde und Abschluss

## 6. Anforderungen der Drittmittelgeber

Bei welchem Drittmittelgeber planen Sie eine Antragstellung?



# Welche Fragen werden in der Regel gestellt?

Beschreibung der Daten

Datentypen, Datenvolumen, Speicherung, Sicherung, Austausch

Rechtliche Aspekte

Datenschutz, Anonymisierung, Pseudonymisierung, Lizenzen

Leitlinien und Standards

z.B. DFG-Leitlinien, FAIR-Prinzipien, disziplinspez. Standards & Umsetzung

Dokumentation

Was wird wie dokumentiert? (z.B. ReadMe-Datei, Codebook, Metadaten)

Archivierung

Was, wo und wie lange? (z.B. 10 Jahre nach DFG-Leitlinie)

Publikation und Nachnutzung

Repositorien, Lizenzen, Embargofrist

Verantwortlichkeiten & Ressourcen

Während und nach Projektlaufzeit

Erstellen eines  
Datenmanagementplans (DMP)

# FDM bei EU – Bund – DFG – Stiftungen

Förderinstitution	FDM-Angaben	Inhalte
Europäische Kommission (Horizon Europe)	verpflichtend	DMP (öffentlich), PID, Lizenzierung, Repository, Dokumentation, FAIR
Bund	uneinheitlich	ggf. FDMP / Verwertungsplan, FAIR
DFG	verpflichtend	FDM im Antrag, ggf. zusätzlich FDM-Verbund-Konzept, fachspez. Standards, Repository, FAIR
VolkswagenStiftung	verpflichtend	DMP, PID, Lizenzierung, fachspez. Standards, Repository, FAIR
Carl-Zeiss-Stiftung	uneinheitlich	Ggf. Forschungsdatenkonzept

Ausschlaggebend ist das Förderprogramm und der jeweilige Ausschreibungstext.

# Europäische Kommission (Horizon Europe)<sup>[1]</sup>

## FDM-Antragsunterstützung

- [Informationen und Inhalte für die Antragsstellung](#) (OpenAIRE)
- [Informationen zu FDM allgemein für Horizon Europe](#) (OpenAIRE)
- [Horizon Europe Programm Guide](#) (PDF)
- [FDM-Template \(RDM-Template\)](#) (PDF)

## FDM-Antragsanforderungen

- Projektantrag: **Stellungnahme zum FDM** (max. 1 Seite)
- Im Projektverlauf: **DMP** entsprechend der DMP-Vorlage verpflichtend:
  - 1. Version innerhalb von 6 Monaten nach Projektstart
  - 2. Version - Aktualisierung: Projektmitte
  - 3. Version: Abschluss

### ALLGEMEIN GILT:

- FAIR-Prinzipien beachten
- Forschungsdaten in Repositorien zugänglich machen
- Nachnutzung ermöglichen:
  - Lizenzierung: mgl. CC-BY oder "Public Domain" CC0
  - Ausführliche Dokumentation (Daten, Erhebungsmethoden, Werkzeuge bzw. Instrumente, inkl. Softwarecode, Algorithmen oder Analyseprotokolle)

# Bundesministerien (u. a. BMBF, BMWi, BMLE)<sup>[1]</sup>

## FDM-Antragsunterstützung

### **Bundesministerium für Bildung und Forschung (BMBF):**

EBF-Rahmenprogramm (empirische Bildungsforschung) => Maßstab und Orientierung

### Küsten-, Meeres- und Polarforschung

- FDM im Verwertungsplan
- Qualität des DMs für Bewertung Projektskizze relevant

### **Bundesanstalt für Landwirtschaft und Ernährung (BLE)**

- [FDMP Skizzenphase \(Stand 01-2024\)](#)
- [FDMP Projektphase \(Stand 01-2024\)](#)
- [Merkblatt FDM \(Stand 08-2023\)](#)

## FDM-Antragsanforderungen

- Keine einheitlichen Vorgaben

Abhängig vom Förderprogramm (Ausschreibungstext beachten!):

- Verwertungsplan mit FDM-Angaben
- DMP / FDMP (Forschungsdatenmanagementplan)

### ALLGEMEIN GILT:

- FDM soll betrieben werden, aber ohne Vorschrift wie
- Man soll mgl. einem fachspezifischen Standard folgen
- Daten sollen mgl. FAIR-nachnutzbar gemacht werden
- Auffindbarkeit & Nachnutzbarkeit wichtig :  
DOI, Repository, Dokumentation (s. BLE)

ENTWICKLUNG DER VORGABEN ORIENTIERT SICH AN NFDI

s. [Aktionsplan BMBF](#)

# Deutsche Forschungsgemeinschaft (DFG)<sup>[1]</sup>

## FDM-Antragsunterstützung

- [Fragenkatalog \(Checkliste PDF\)](#) zum Umgang mit Forschungsdaten (FD)
- [Informationen zu beantragbaren Mitteln](#)  
„Kosten, die für Datenmanagement anfallen und der Nutzbarkeit diese Daten dienen, können beantragt werden“ (keine Grundausstattung!)
- [Förderprogramm](#) - spezifische Informationen
- [Fachspezifische Empfehlungen](#)

## FDM-Antragsanforderungen

### Grundsätzliche Förderkriterien:

- Schilderung Umgang mit FD im (Teil-)Projekt (Abschnitt 2.4 des Antrags, Hilfsmittel: Fragenkatalog der DFG)
  - Konzept zum Erhalt der FD nach Förderende
  - Empfehlung der DFG: frühzeitige Suche nach Repositorium
  - Anschlussfähigkeit an bestehende Strukturen (etwa bei Schaffung von Dateninfrastrukturen und Methoden)
  - Standards der Fachkollegien beachten

### Verbundanträge:

In Skizze & Antrag (Abschnitt 2.3):

Umgang mit FD aufnehmen

=> Entwicklung eines gemeinschaftlichen FDM-Verbund-Konzepts!

# Volkswagenstiftung<sup>[1]</sup>

## FDM-Antragsunterstützung

- [Basis-DMP](#) mit 7 Fragen
- [FAQs für Antragsstellende](#)

Förderprogramm: [Data-Reuse](#)

Mittel für FDM bis 6 Monate nach Projektende

## FDM-Antragsanforderungen

- Datenmanagementplan (DMP) bei datenintensiven Forschungsanträgen (Basis-DMP inkl. ggf. Vorgaben des Zielrepositoriums)
- **Skizze des Datenmanagements im Antrag sinnvoll, auch wenn nicht im Förderprogramm erwähnt!**

Hinweise zum DMP – erwünscht ist:

- Beachtung FAIR-Prinzipien & community-spezifische Standards
- Vergabe von persistenten Identifikatoren (z. B. DOI)
- Lizenzvergabe (während und nach Projekt)
- Repositoriumsüberlegung/-auswahl

# Carl-Zeiss-Stiftung<sup>[1]</sup>

## FDM-Antragsunterstützung

- [Übersicht der Fördermöglichkeiten für HAW](#)

## FDM-Antragsanforderungen

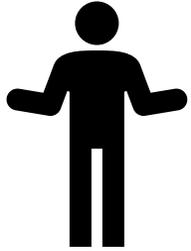
abhängig vom Förderprogramm

Ggf. Forschungsdatenkonzept (keine Aktualisierung):

- Datenschutz
- Open Source
- Open Access
- gemeinsame Datenbanken
- Reproduzierbarkeit

# Agenda

1. Begrüßung
2. Was ist FDM und warum das Ganze?
3. FAIRe Daten
4. Der Lebenszyklus von Forschungsdaten
5. Rechtliche Rahmenbedingungen
6. FDM-Anforderungen von Drittmittelgebern
- 7. Fragerunde und Abschluss**



# Fragen und Antworten



# FDM@HAW.rlp



## Vielen Dank für Ihre Teilnahme!



# Links zum Thema

Umfassende Info-Plattform Forschungsdaten.info

<https://www.forschungsdaten.info>

Infonetzwirk: Forschungsdaten.org

<https://www.forschungsdaten.org/index.php/Hauptseite>

Services für die Bildungsforschung von kooperierenden Bildungsdatenzentren

<https://www.forschungsdaten-bildung.de/>

FDM-FAQ

[https://www.fdmhawrlp.de/?page\\_id=46](https://www.fdmhawrlp.de/?page_id=46)

# Sekundärliteratur zum Thema

Handbuch Forschungsdatenmanagement / hrsg. von Stephan Büttner ... – Bad Honnef : Bock + Herchen, 2011. - 223 S. E-book (lizenzfrei)

<http://opus.kobv.de/fhpotsdam/volltexte/2011/241/pdf/HandbuchForschungsdatenmanagement.pdf>

# Nützliche Tools

- Wie FAIR sind Ihre Daten: <https://satisfyd.dans.knaw.nl/>
- Interaktiver Assistent, wie ich Datenschutz in meinem Projekt beachten sollte: [https://wiki.bib.uni-mannheim.de/xerte/play.php?template\\_id=217#page1](https://wiki.bib.uni-mannheim.de/xerte/play.php?template_id=217#page1)

# Quellennachweise (I)

**Folie 14:** [1] Foto von Alex Green bei [pexels](#); [2] Credit: [NIAID](#), Jeff Dahl, Public domain, [via Wikimedia Commons](#), [Giovanni Boccardi](#), CC BY-SA 3.0 IGO; [Lucas Taylor](#), [CERN](#); [3] Bild: [Jens P. Raak](#) from [Pixabay](#).

**Folie 15:** [1] Abbildung nach Soßna, Volker, Glasenapp, Yvana, Renziehausen, Anna-Karina: [Digitale Forschungsdaten managen – Grundlagen, Tipps und Tricks](#), Leibniz Universität Hannover, TIB.

**Folie 16:** Abbildung nach Soßna, Volker, Glasenapp, Yvana, Renziehausen, Anna-Karina: [Digitale Forschungsdaten managen – Grundlagen, Tipps und Tricks](#), Leibniz Universität Hannover, TIB.

**Folie 17:** [1] [Arthur, Charles \(23 August 2013\): "Tech giants may be huge, but nothing matches big data". The Guardian. Retrieved 30 April 2019.](#)

**Folie 18:** [1] Abbildung nach [The Open Traits Network: Using Open Science principles to accelerate trait-based science across the Tree of Life – Scientific Figure on ResearchGate](#).

**Folie 22:** [1] Abbildung nach: All Research Data - Sarah Jones (DCC), [University of Capetown – Digital Library Services](#) [What are Data?](#).

**Folie 24:** Abbildung [1]: eigene Darstellung nach: <https://forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>

**Folien 32, 40:** vgl. Biernacka, Katarzyna, Buchholz, Petra, Danker, Sarah Ann, Dolzycka, Dominika, Engelhardt, Claudia, Helbig, Kerstin, Jacob, Juliane, Neumann, Janna, Odebrecht, Carolin, Petersen, Britta, Slowig, Benjamin, Trautwein-Brunns, Ute, Wiljes, Cord, Wuttke, Ulrike: Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement. Version 4.0. Berlin, 2021. <https://doi.org/10.5281/zenodo.5773203>

**Folie 31:** Abbildung [1] : [https://phdcomics.com/comics/archive\\_print.php?comid=1323](https://phdcomics.com/comics/archive_print.php?comid=1323); [2] JorgeCham(2012).„FINAL“.doc. <https://phdcomics.com/comics/archive.php?comid=1531>. [3] <https://www.buero-kaizen.de/ordnerstruktur/>

**Folie 32:** vgl. Biernacka, Katarzyna, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes, Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020. <https://doi.org/10.5281/zenodo.3938533>

**Folie 37:** vgl.: Biernacka, Katarzyna, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes, Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020, S. 59 ff. <https://doi.org/10.5281/zenodo.3938533>

**Folie 38 Abb.:** vgl. Biernacka, Katarzyna, Buchholz, Petra, Danker, Sarah Ann, Dolzycka, Dominika, Engelhardt, Claudia, Helbig, Kerstin, Jacob, Juliane, Neumann, Janna, Odebrecht, Carolin, Petersen, Britta, Slowig, Benjamin, Trautwein-Brunns, Ute, Wiljes, Cord, Wuttke, Ulrike: *Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement*. Version 4.0. Berlin, 2021, S. 59. <https://doi.org/10.5281/zenodo.5773203>

# Quellennachweise(II)

**Folie 38:** vgl.: Biernacka, Katarzyna, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020, S. 59 ff. <https://doi.org/10.5281/zenodo.3938533>

**Folie 39:** Vgl.: <https://www.forschungsdaten-bildung.de/dateien-benennen>

**Folie 42:** Foto: privat

**Folie 43:** \*vgl.: Piwowar, H. A., Day, R. S., Fridsma, D. B. (2007). [Sharing detailed research data is associated with increased citation rate](#). PloS one, 2 (3), e308.  
vgl.: Swan, A. (2010): The Open Access citation advantage: Studies and results to date. <https://eprints.soton.ac.uk/268516/>

**Folie 44:** Quelle: <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards/> (30.10.2023)

**Folie 45:** vgl.: <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation/>

**Folie 49:** Tabelle: eigene Darstellung

**Folie 51: Tabelle** in Anlehnung an: Biernacka, Katarzyna, Buchholz, Petra, Danker, Sarah Ann, Dolzycka, Dominika, Engelhardt, Claudia, Helbig, Kerstin, Jacob, Juliane, Neumann, Janna, Odebrecht, Carolin, Petersen, Britta, Slowig, Benjamin, Trautwein-Bruns, Ute, Wiljes, Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020. <https://doi.org/10.5281/zenodo.3938533>

**Folie 58:** vgl.: [Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., McGillivray, B. \(2020\): The citation advantage of linking publications to research data. PLoS ONE 15\(4\): e0230416. https://doi.org/10.1371/journal.pone.0230416](#)

**Folie 60:** Creative Commons Germany: <https://de.creativecommons.net/was-ist-cc/>

**Folie 61:** Grafik: Barbara Klute und Jöran Muuß-Merholz für [wb-web](#) unter [CC BY SA 3.0](#)

**Folie 68:** vgl. <https://forschungsdaten.info/themen/rechte-und-pflichten/recht-und-forschungsdaten-ein-ueberblick/>

**Folien 77 - 81:** Tabellen nach Boelter, Sarah, Schimmer, Thomas, Richter, Manuela, Schmidt, Diana, Schneidenbach, Esther (February 2024): [Und jetzt noch FDM?! Ein Erfahrungsaustausch zur Antragsberatung an HAW/FH](#), Love Data Week 2024.