

Einführung in das Forschungsdatenmanagement

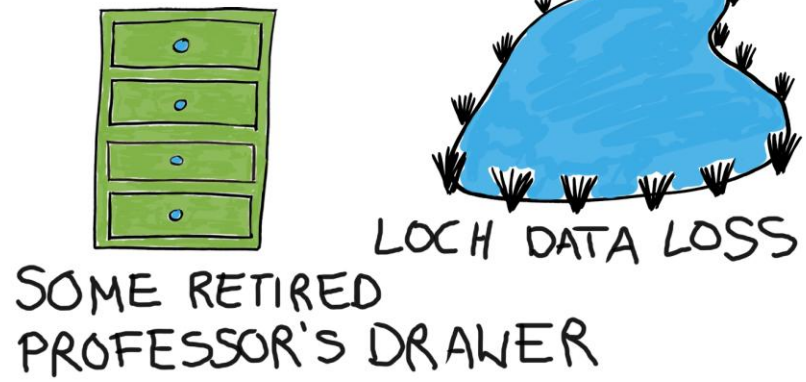
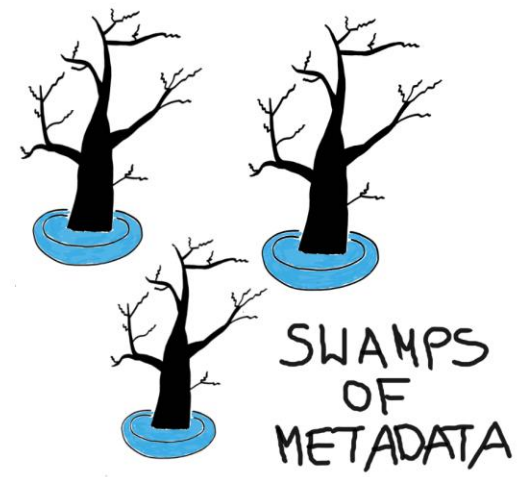
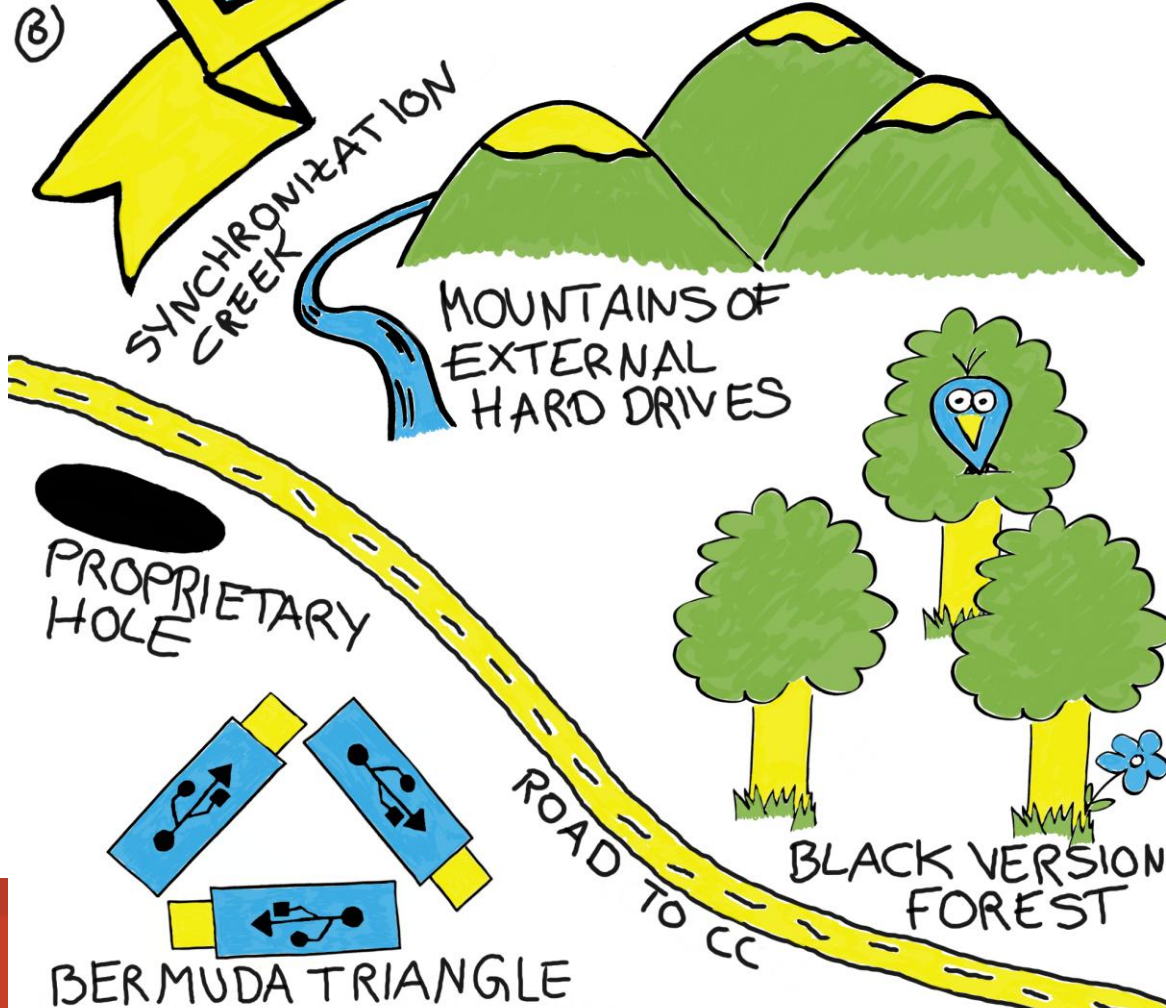
Dr. Anne Vieten; 10.12.2025

JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

JG|U

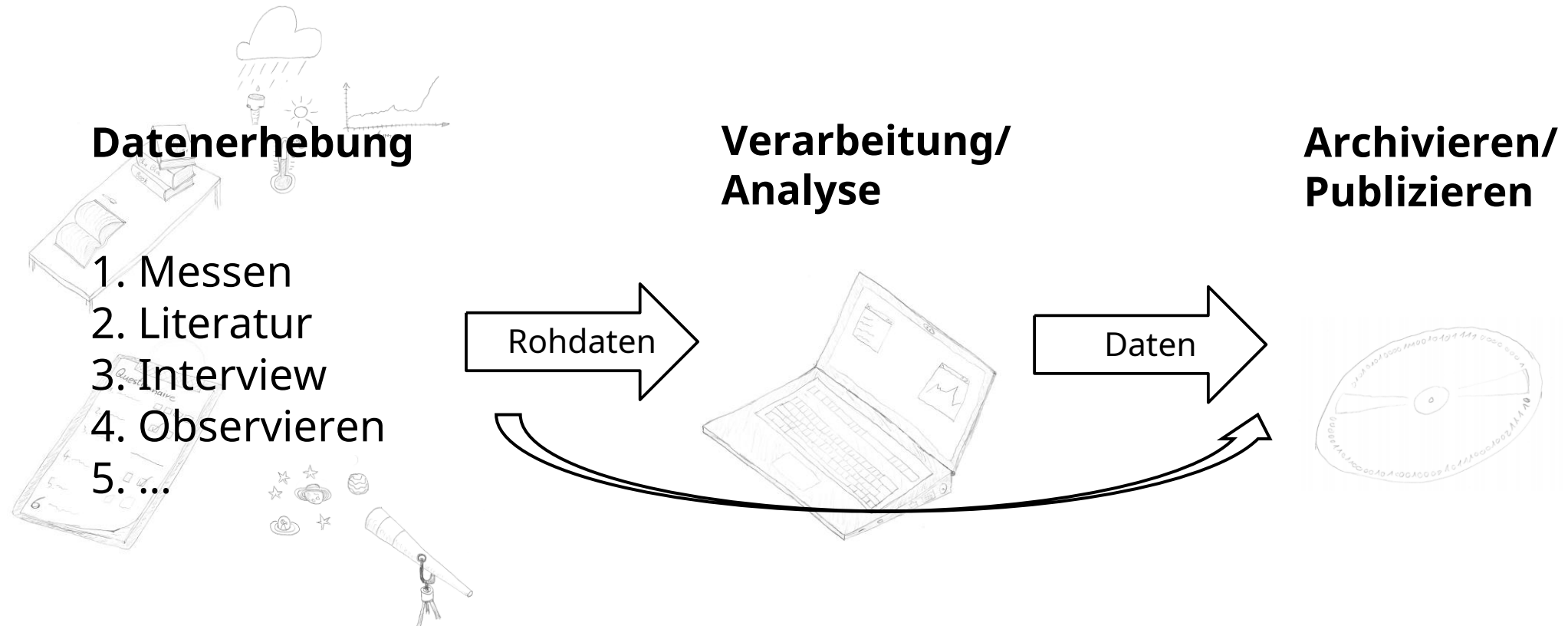


LOST DATA MAP



CC BY-SA 4.0
Franziska Mau
[FranziMachtDas – Sketches, Zeug & der tägliche Wahnsinn](#)

FORSCHUNGSDATEN, EINE DEFINITION



FORSCHUNGSDATENMANAGEMENT (FDM) – WARUM SINNVOLL?

- Reduzierung von Datenverlusten
 - Reduzierung von Fehlern
 - Strukturierte Durchsuchbarkeit von Daten (z.B. für Repositorien)
 - Bessere Auffindbar- und Nachnutzbarkeit der Daten über Datenrepositorien
- **Open Science**
- Zitierung von Datensätzen!
 - Bessere Überprüfbarkeit von Forschungsergebnissen
 - Langfristige Verfügbarkeit von Daten



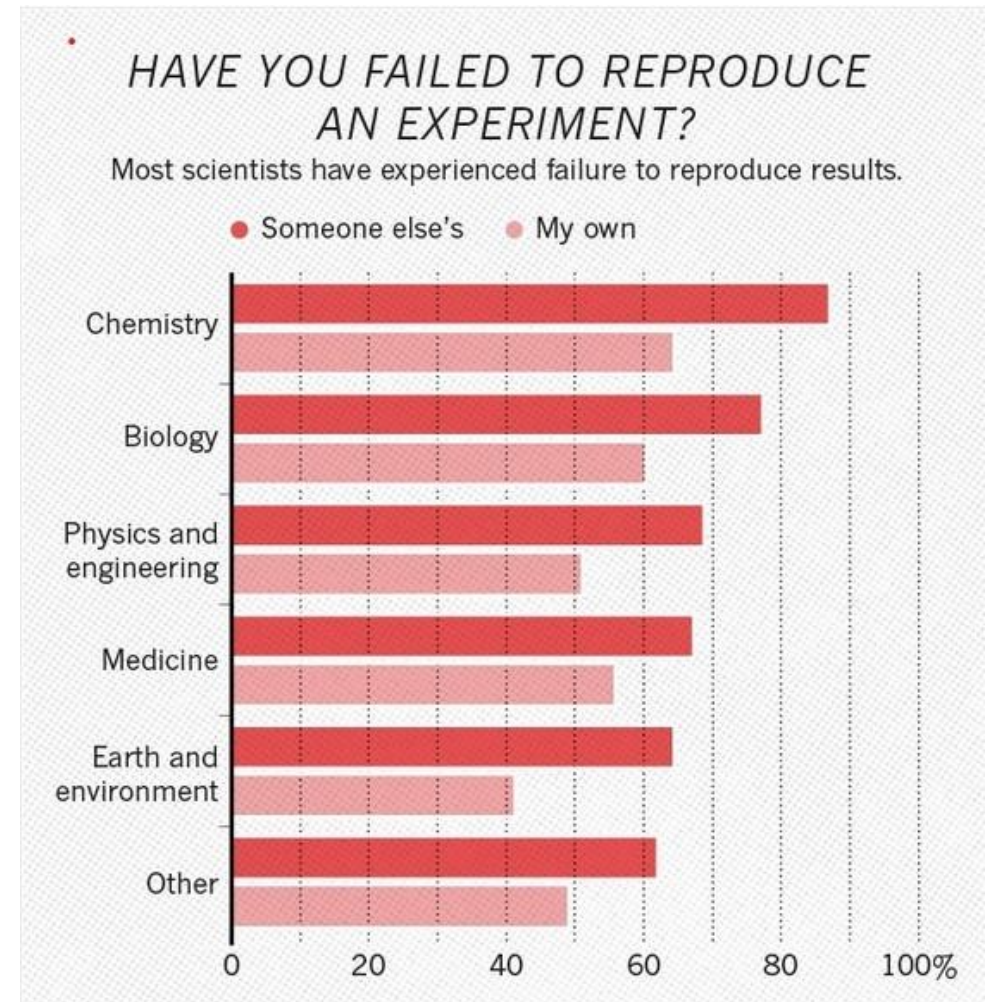
Postcard by:

Landesinitiative NFDI der digitalen Hochschulen NRW - Forschungsdatenmanagement

DIE REPRODUZIERBARKEIT VON FORSCHUNGS- ERGEBNISSEN

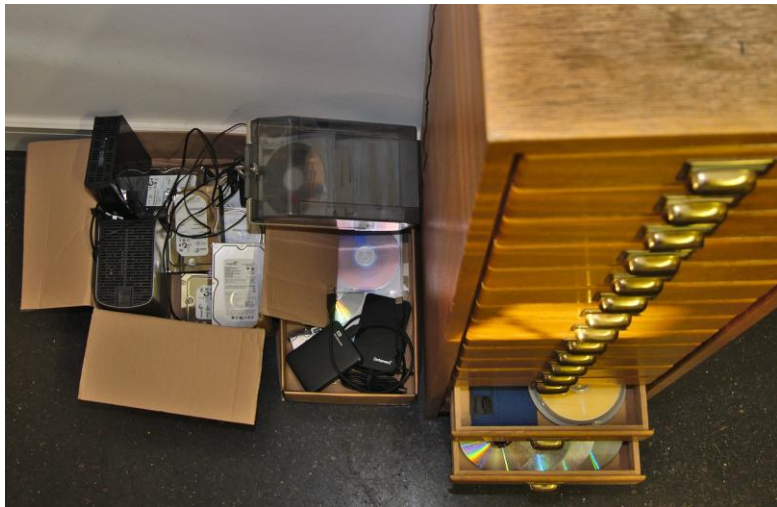
... schlägt oftmals fehl!

[Baker, M.; Nature; May 25th 2016](#)



NUR EIN GENIE KANN MIT CHAOS UMGEGEHEN

... aber was passiert wenn das Genie weggeht?



WEITERE GRÜNDE: VORGABEN DRITTMITTELGEBER UND JGU

- Ordnung zur guten wissenschaftlichen Praxis an der [JGU](#) [Vortrag](#)
- DFG: [Leitlinien zum Umgang mit Forschungsdaten; FDM Checkliste](#)
- Horizont Europa: [Open Science in Horizon Europe \(europa.eu\)](#)
- BMFTR: [Forschungsdaten – BMFTR](#)
- VW-Stiftung:
https://www.volkswagenstiftung.de/sites/default/files/documents/Open_Science_Policy_und_Umsetzung_VolkswagenStiftung.pdf
- Kodex: [„Leitlinie zur guten wissenschaftlichen Praxis der DFG“ \(2019\)](#)
- [Leitlinie zum Umgang mit Forschungsdaten der JGU](#)

FORSCHUNGSDATENLEBENSZYKLUS

Forschungsdatenmanagement:

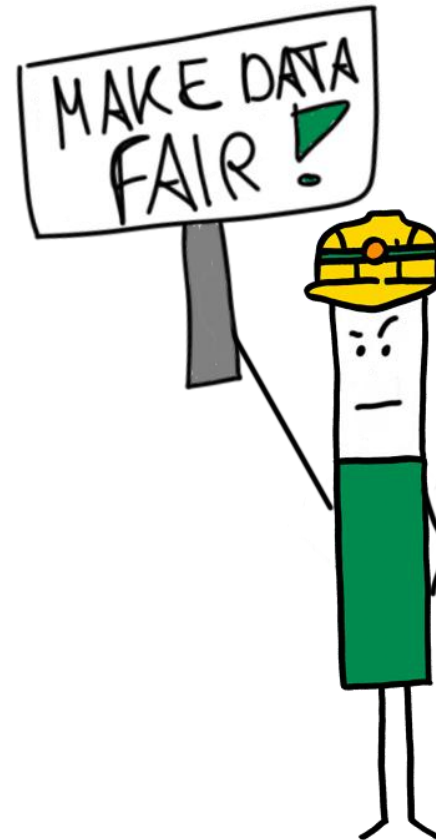
Umfasst alle Forschungsdaten betreffenden Prozesse von der Erstellung, der Verarbeitung, der Analyse, der Archivierung und ggf. auch Löschung sowie die Publikation und anschließende Nachnutzung von Forschungsdaten.



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

UND WIE SETZE ICH DIES UM?

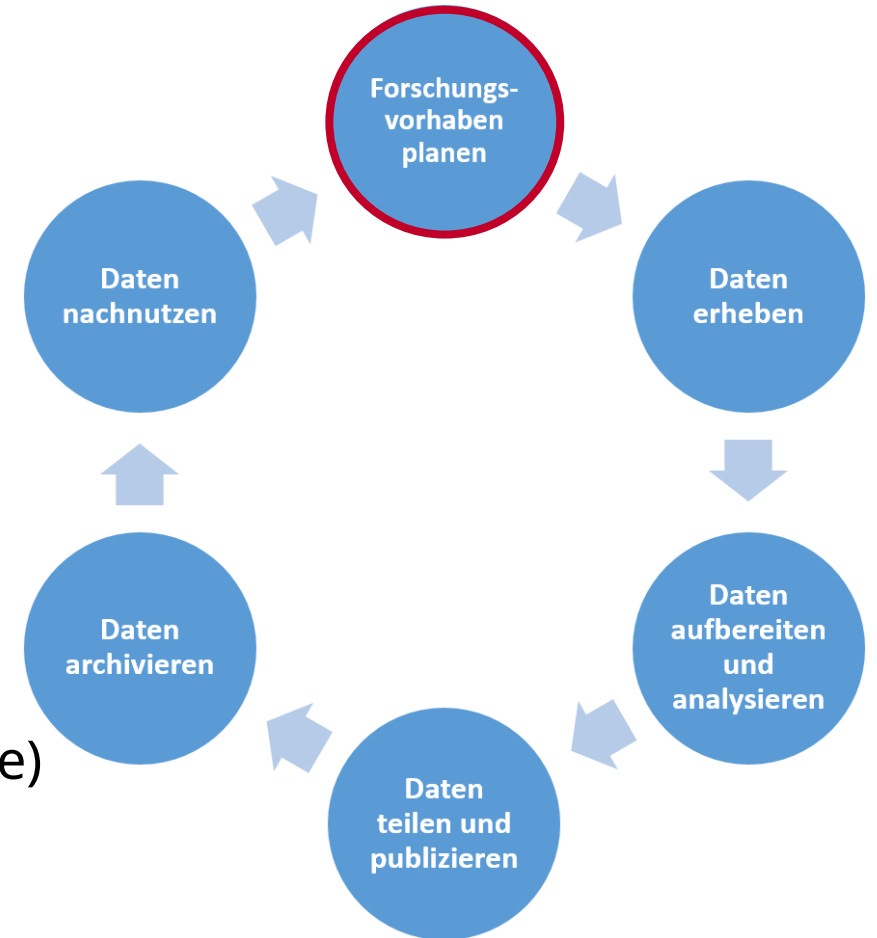
???



[Abb.: Sketchnotes & Sketches – FranzMachtDas _CC BY-SA 4-0.png](#) CC-BY-SA 4.0 Franziska Mau

FORSCHUNGSVORHABEN PLANEN

- Gestaltung/Planung des Forschungsprozesses
- Auswahl Software und Datenformate
- Auswahl des Speicherorts
- Organisation des Backups
- Festlegung der Datendokumentation (Metadaten u.Ä.)
- Regelungen für das Teilen der Daten im Projekt
- Klärung der Rechtssituation (Datenschutz, Urheberrechte)
- etc....



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

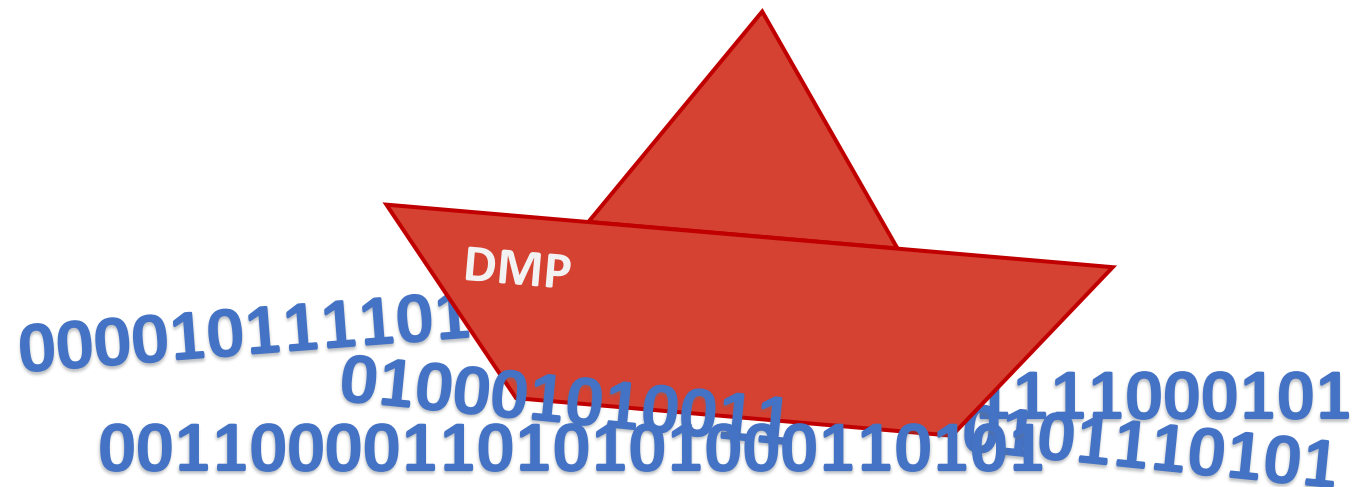
DATENMANAGEMENTPLAN

Typische Inhalte:

1. Administrative Informationen
2. Dokumentation und Metadaten
3. Ethische und rechtliche Fragen
4. Speicherung und Back-Up
5. Datenauswahl und -erhaltung
6. Daten mit anderen teilen
7. Zuständigkeiten und Ressourcen

Aufgaben:

- Verwaltung von Forschungsdaten
- Festlegung von FDM-Schritten
- Dynamisches FDM-Werkzeug



Vergl.: http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf

Data

The following questions collect information on the data that is produced or used in the project. They also help to estimate the value of the data in terms of potential re-use and long-term preservation.

Before data is newly created, it is advisable to check if there is existing data that could be re-used. This way, redundant collection or creation of research data is prevented. This saves efforts and costs. Furthermore, in the case of personal data, the [German Federal Data Protection Act](#) allows the collection of personal data only when there are no other reasonable means to clarify the research question (re-use of existing data would be such a reasonable means). Also, there shall be collected no more information than necessary.

The information regarding the data collected, produced or used in the project is gathered along datasets. The definition of these datasets is an important conceptual decision that has to be made individually and carefully for each project.

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

d [+ Dataset](#)



What kind of dataset is it?

Please briefly describe the data type and / or the method used to create or collect the data, for example:

- quantitative online survey
- 3D model / digital reconstruction of a stone age settlement
- software developed within the project

[Back](#) [Proceed](#)

Overview

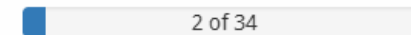
Project: [TestB](#)
Catalog: [Brief questionnaire](#)

[Show management panels](#)

[Reload page](#)

[Back to my projects](#)

Progress



[Back](#)

[Proceed](#)

Navigation

Grey entries will be conditionally skipped based on your input.

[General \(1 of 5\)](#)

[Content classification](#)

→ [Data](#)

[Data origin](#)

[Technical classification](#)

[Data usage](#)

[Metadata and referencing](#)

[Legal and ethics \(1 of 3\)](#)

[Storage and long-term preservation](#)

is for effective
mlandDs

Science Foundation (FCT)

and biodiversity and
assessing trade-offs between
ocial capital underlying such

ext of this plan as they would
redit the creator(s) as the source
: creator(s) endorse, or have

[blic plans](#)

[My Projects](#) / [TestB](#) / [Technical classification](#)

Data size

Please fill in the form for each tab. The same tabs may be used later on other pages. You can add a new tab using the green button. Once created, you can edit or delete tabs using the buttons in the top right corner.

d

f

[+ Dataset](#)

What is the actual or expected size of the dataset?

- Less than 1 GB
- Between 1 GB and 1 TB
- Between 1 TB and 100 TB
- More than 100 TB
- Exact size:

[Back](#)[Proceed](#)

Overview

Project: [TestB](#)Catalog: [Brief questionnaire](#)[Show management panels](#)[Reload page](#)[Back to my projects](#)

Progress

[Back](#)[Proceed](#)

Navigation

Grey entries will be conditionally skipped based on your input.

[General \(1 of 5\)](#)[Content classification](#)[Technical classification](#)[→ Data size](#)[Formats](#)[Tools](#)[Versioning](#)[Data usage](#)[Metadata and referencing](#)[Legal and ethics \(1 of 5\)](#)[Storage and long-term preservation](#)

RECHTLICHES: DATENSCHUTZ

- Nehmen Sie Kontakt zum [Datenschutzbeauftragten](#) auf.
- Beachten Sie die geltenden **Bundes- und das Landesdatenschutzgesetze** sowie die **europäische Datenschutz-Grundverordnung**
- Holen Sie bei Befragungen/Interviews/Patientendaten etc. eine **schriftliche, informierte, Einverständniserklärung** der Betroffenen ein (Widerrufsrecht!).
- **Anonymisieren** (ggf. Pseudonymisieren) Sie die erhobenen Daten.
- Speichern Sie die Rohdaten **zugriffsgeschützt** (z.B. verschlüsselt) und getrennt von den anderen Daten (z.B. Auswertungen).
- Speichern Sie personenbezogene Daten von EU-Bürgern **niemals auf Servern außerhalb der EU!**

DIE CARE PRINZIPIEN

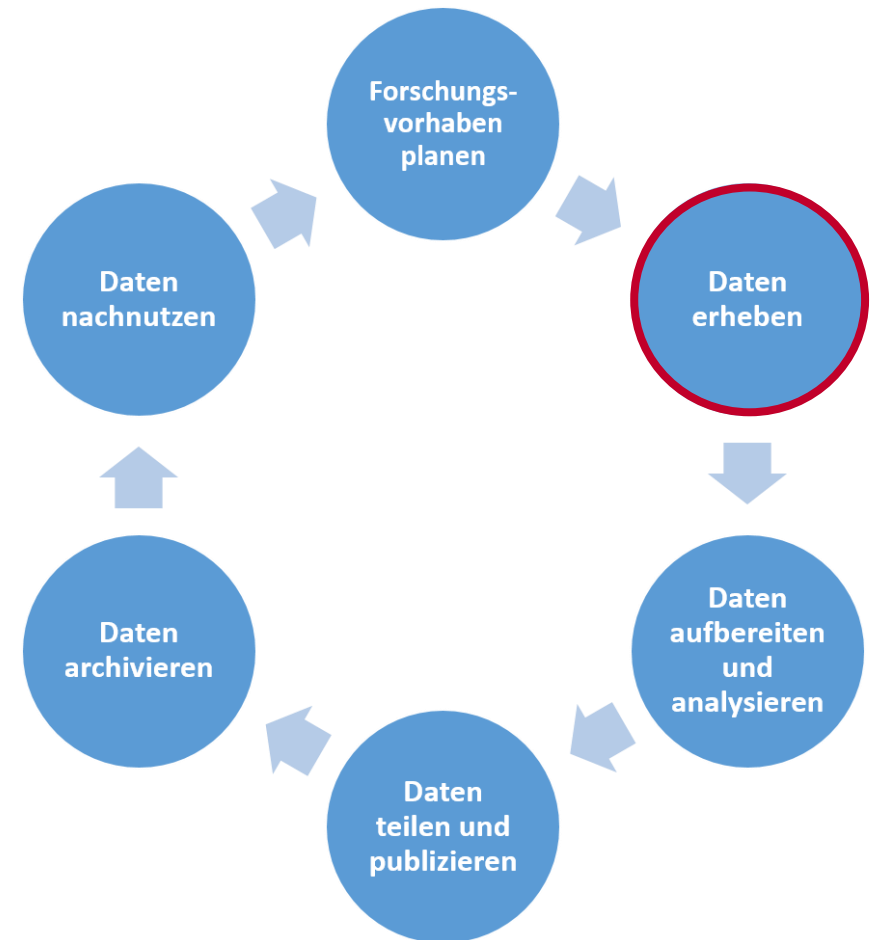
...stehen für einen angemessenen Umgang mit Daten, die **indigene Gemeinschaften** betreffen

- **C**ollective benefit (kollektiver Nutzen): Indigene Völker können Nutzen aus Daten ziehen.
- **A**uthority to control (Recht auf Kontrolle über Daten): Rechte, Interessen und Kontrolle indigener indigener Völker an Ihren Daten müssen sichergestellt sein.
- **R**esponsibility (Verantwortung): Diejenigen, die mit indigenen Daten arbeiten, sind dafür verantwortlich mitzuteilen, wie die Daten verwendet werden.
- **E**thics (Ethik): Die Rechte und das Wohlergehen indigener Völker stehen im Datenlebenszyklus im Vordergrund.

[Die CARE-Prinzipien für indigene Data Governance CARE+Principles_One+Pagers+FINAL_Oct_17_2019.pdf](#)

DATEN ERHEBEN

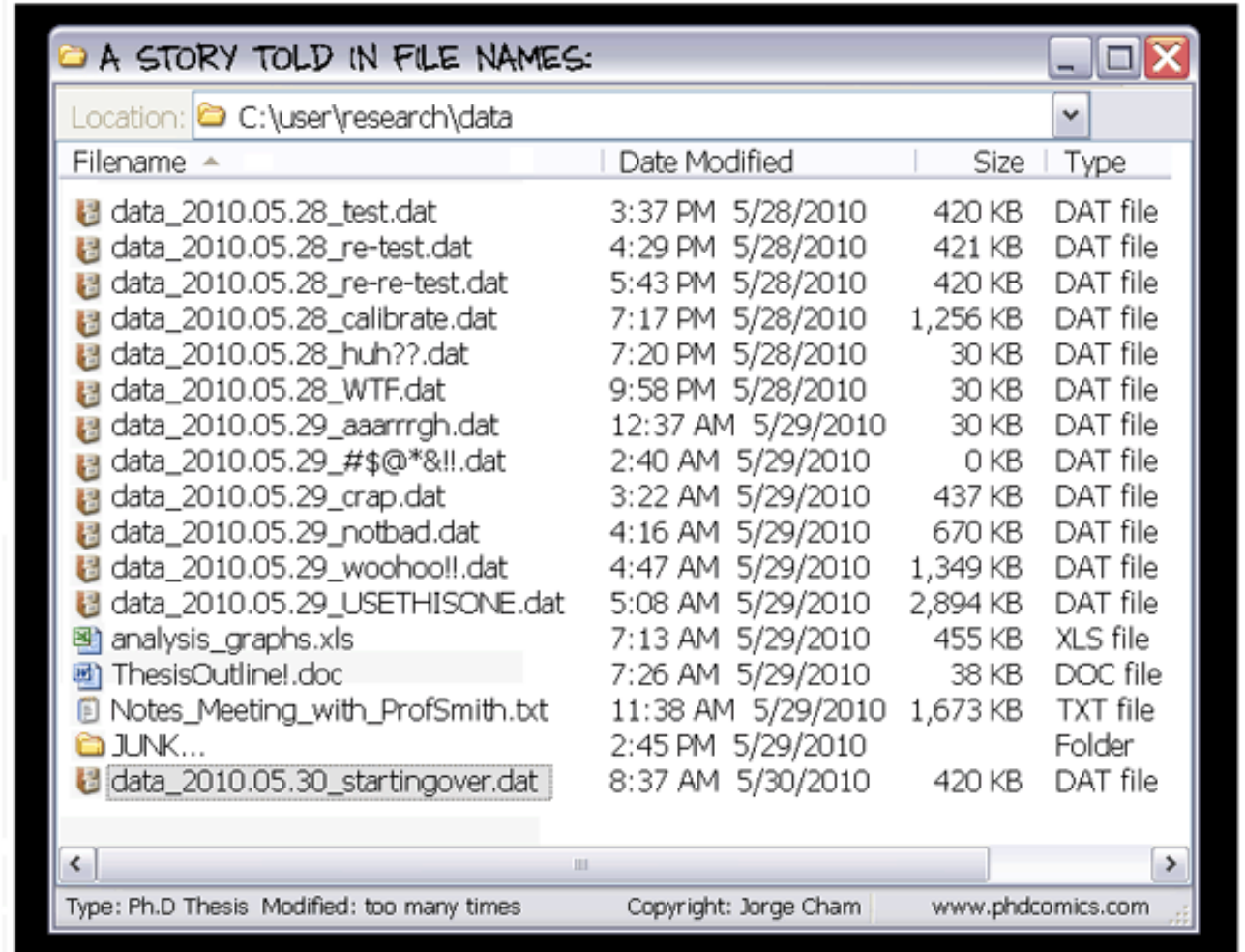
- Dateneingabe und -übertragung
- ggf. Digitalisierung von Medien
- Datenprüfung und-bewertung
- Anonymisierung von Daten, falls notwendig
- Erstellung von Dokumentationen/Metadaten
- Datenorganisation und standardisierte Dateibenennung
- Maßnahmen zur Versionierung
- Backup
- Zugriffsrechte vergeben
- ...



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

DATEN ORGANISIEREN

...oder die Übersicht verlieren?



Jorge Cham: [PHD Comics: A story in file names](http://www.phdcomics.com)

BESSER SO!

Zentral ist die initiale Planung:

Entwicklung einer Benennungskonvention für Dateien und Ordner → Dokumentation!

Dateibenennung:

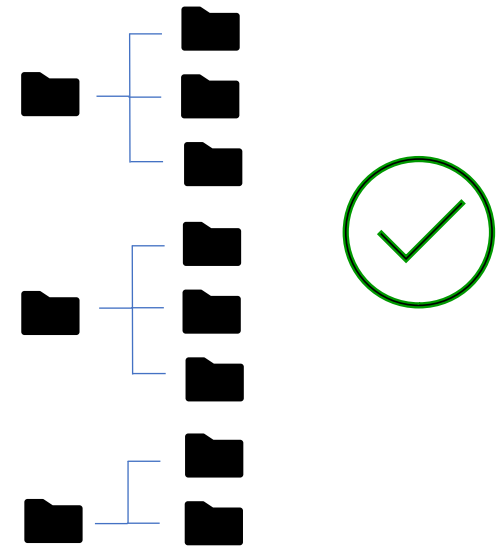
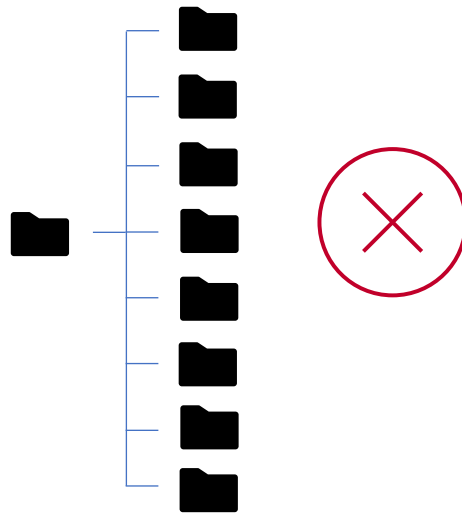
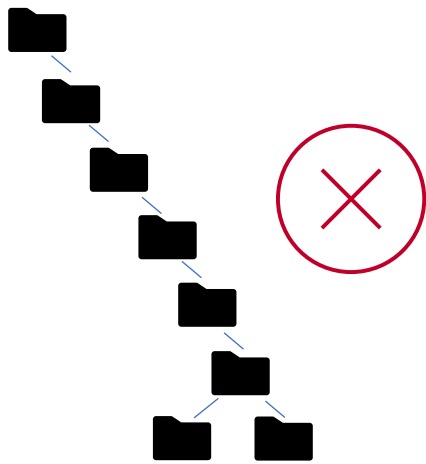
- Sinnhaft
- Sortierbar (z.B. festgelegte Reihenfolge der Namensbestandteile, chronologische Sortierung YYYYMMDD)
- Benennung auch bei Umzug der Datensätze in andere Ordner sinnvoll
- Mögl. Elemente: Versionsnummer, Datum, Inhaltsbeschreibung, Projektnummer etc.; ≤ 32 Zeichen
- So kurz wie möglich und so lang wie nötig
- Keine Sonderzeichen wie % oder ? sowie keine Leerzeichen
- Versionierung berücksichtigen und eindeutig kenntlich machen

WARUM?

- | | |
|--|---------|
| 1. Fg23Spat6ev.mp3 | nein |
| 2. Projekt005_Probe060_UV_VIS_20250803.csv | ja |
| 3. Vortrag_Forschungsdatenmanagement.pptx | ja/nein |
| 4. Probe51 2504 Petra+Uwe neu! Korrigiert.docx | nein |
| 5. 2020_09_18_JGU_Aktionstag_Forschungsdaten.pdf | ja |

ANLAGE VON ORDNERN

- Ordnernamen nicht redundant anlegen
- Benennungskonvention folgen
- Balance zwischen Verschachtelung und Anzahl der (Unter-)Ordner finden



VERSIONIERUNG

Möglichkeiten:

1. Versionsmanagement als Teil der Dateibenennungskonvention

- Versionsnummern in der Dateibenennung z.B. V01, V02 etc.
- Verwendung der aktuellen Daten der Version in der Dateibenennung

2. Verwendung von Versionierungswerkzeugen

- Z.B. Git basierte Software, wie GitLab (eigentlich für Softwarecode, aber auch für andere Texte anwendbar)

METADATEN

Bedeutung: Metadaten beschreiben/definieren andere Daten näher.

- Beispiel: Metadaten zu einem Buch sind z.B. Angaben über den Autor, das Erscheinungsjahr, Verlag etc.
- Notwendig für:
 1. Vertieftes Verständnis der eigentlichen Daten
 2. Durchsuchbarkeit von Repositorien/Datenbanken.
- Standards: z.B. Dublin Core, DataCite →
- Repositorien haben meist eigene Standards

Data Cite Basic Scheme

- Identifier
- Creators
- Title
- Publisher
- Publication year
- Subject
- Language
- Resource type
- Descriptions

BEISPIELE FÜR METADATENWERKZEUGE

<https://dhvlab.gwi.uni-muenchen.de/datacite-generator/>

[The Simple Dublin Core Generator!](#)



Verschiedene FDM-Werkzeuge:

[All tools and resources | RDMkit](#)

The screenshot shows a web form titled "Input" with a yellow header. It contains several input fields for metadata, each with a label and a "(+)[-]" icon. The fields are: Title?, Creator?, Subject?, Description? (with a larger text area and a small icon in the bottom right), Publisher?, Contributor?, Date?, Type?, Format?, Identifier?, Source?, Language?, Relation?, Coverage?, and Rights?. Each field has a light gray border and a small icon in the bottom right corner.

TYPEN VON METADATEN

- **Deskriptiv:** Titel, Autor, Schlagworte, Erscheinungsjahr, etc.
- **Administrativ-rechtlich:** Rechteinhaber, Lizenzbedingungen, Zugangsregeln, etc.
- **Strukturell:** physische oder logische Beziehungen zwischen Objekten (z.B. zwischen Tabellen in einer Datenbank)
- **Administrativ- bzw. erhaltungsbezogen:** Konvertierungs- und Erhaltungsmaßnahmen, etc.
- **Technisch:** Informationen zu Soft- oder Hardware, Dateiformate, Zeichen-kodierung, Parameter von Messgeräten, etc.

BEISPIELE FÜR DATENDOKUMENTATIONEN/- BESCHREIBUNGEN

- Laborbücher (papierbasiert/digital z.B. [eLabFTW](#))
- Codebücher
- Datenbankschemata
- Versuchsprotokolle
- Softwaresyntax
- Readme.txt Datei zur Beschreibung der Datensätze
- ...

WAS MUSS ICH ALLES DOKUMENTIEREN?

All das, was Dritte zum Verständnis des Datensatzes benötigen!

DATENSICHERUNG ALLGEMEIN

- Zugriffsrechte klären!
- Datenschutz berücksichtigen!
- Versionierung beachten (z.B. Verwendung von GitLab)
- Regelmäßige Backups durchführen (z.B. Nutzung Uniserver)
- Mehrere Speichermedien verwenden!

BACKUP 3-2-1 REGEL

3 = 3 Kopien

2 = 2 verschiedene Speichermedien

1 = 1 Kopie an einem anderen Standort als die anderen Kopien

DATEN AUFBEREITEN UND ANALYSIEREN

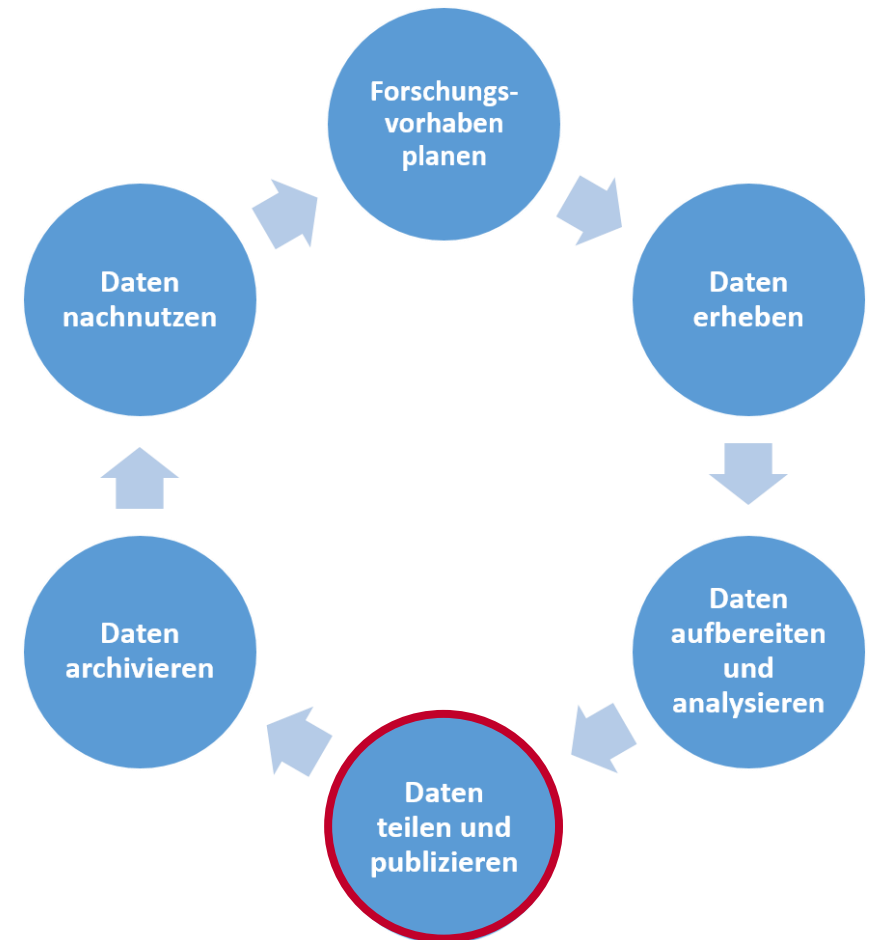
- Interpretation von Daten
- Vernetzung von Daten
- Maßnahmen zur Plausibilität und Reproduzierbarkeit von Daten
- Erstellung von Dokumentationen/Metadaten
- Produktion von Forschungsergebnissen
- Qualitätssicherung
- ...



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

DATEN TEILEN UND PUBLIZIEREN

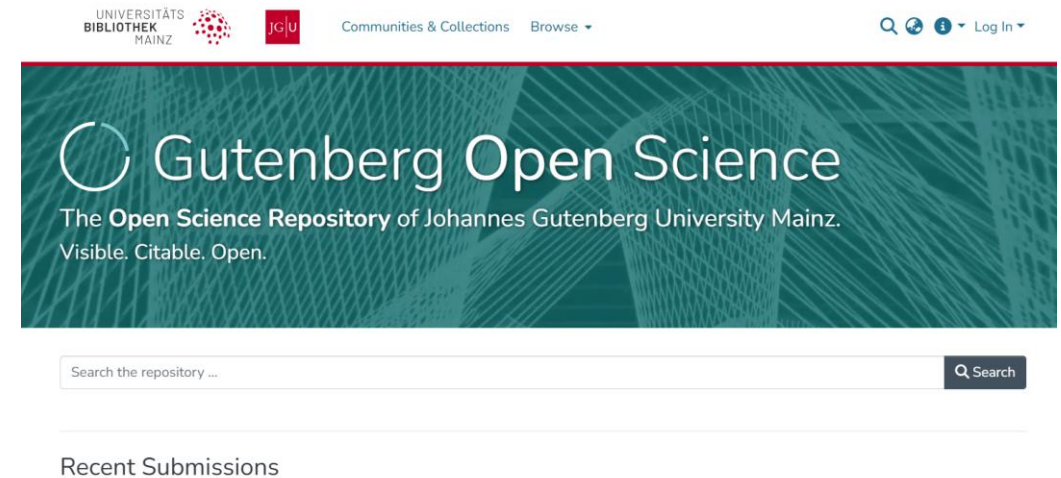
- Erstellung von Publikationen in Zeitschriften, Büchern, etc.
- Veröffentlichung von Daten (z.B. über Repositorien/ Datenbanken)
- Zugangskontrolle/Rechtmanagement zu Daten
- Urheberrechtlichen Schutz der Daten (wenn mögl.) bzw. Nutzungsbedingungen (z.B. Creative Commons) sicherstellen/festlegen
- Schlagwörter für die Durchsuchbarkeit in Repositorien
- Festlegung Offener oder beschränkter Zugang
- Datenschutz berücksichtigen ...



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

GUTENBERG OPEN SCIENCE JGU

- Institutionelles Open Access Repository
- Pre- und Postprints, Doktorarbeiten und Forschungsdaten (bis 10 gb pro Datensatz)
- Möglichkeit der Bündelung von Publikationen und dazugehörigen Forschungsdaten
- Standardisierte, von UB kuratierte Metadaten (Dublin Core)
- Single-Sign-On („ZDV-Konto“)
- Zeitliche Embargos begrenzt möglich
- Übertragung von Metadaten in Gutenberg Research Elements möglich



[Gutenberg Open Science :: Home](#)

ZENODO

Zenodo <https://zenodo.org/>

- offen für alle Wissenschaftsfelder
- Publikationen und Forschungsprimärdaten
- DOI für alle Uploads
- Gehostet und betreut durch das CERN
- Flexible Lizenzierung
- Upload bis 50 GB pro Datensatz (kostenlos)
- Zugangsbeschränkung möglich (kurzzeitig)

The screenshot shows a Zenodo record page for 'PDSI_for_LPJ-GUESS: initial release' by Jörg Steinkamp. The page includes a search bar, user profile, and various action buttons like 'Edit', 'Upgrade to versioned record', and 'Remove'. A file preview window is open, showing a directory structure for 'joergsteinkamp-PDSI_for_LPJ-GUESS-256f700' with files 'LICENCE.md', 'README.md', and 'pdsi.h'. Below the preview, a table lists the main file 'PDSI_for_LPJ-GUESS-initial.zip' with a size of 16.7 kB and a download link. The 'Citations' section shows no citations. The right sidebar contains statistics (41 views, 1 download), availability in GitHub and OpenAIRE, and publication details (September 23, 2014) with a DOI of 10.5281/zenodo.11843.

zenodo Search Upload Communities steinkamp.joerg@gmail.com

September 23, 2014 Software Open Access Edit

Upgrade to versioned record

PDSI_for_LPJ-GUESS: initial release

Jörg Steinkamp

Initial release

Preview

PDSI_for_LPJ-GUESS-initial.zip

- joergsteinkamp-PDSI_for_LPJ-GUESS-256f700
 - LICENCE.md 35.1 kB
 - README.md 292 Bytes
 - pdsi.h 10.0 kB

Files (16.7 kB)

Name	Size	Preview	Download
PDSI_for_LPJ-GUESS-initial.zip	16.7 kB	<input type="checkbox"/>	<input type="checkbox"/>

md5:6588ff39afd77fbd9ed6e1d830d7e94

Citations 0

Show only Literature (0) Dataset (0) Software (0) Unknown (0) Citations to this version

No citations.

Available in

GitHub

Indexed in

OpenAIRE

Publication date: September 23, 2014

DOI: [10.5281/zenodo.11843](https://doi.org/10.5281/zenodo.11843)

Related identifiers: Supplement to https://github.com/joergsteinkamp/PDSI_for_LPJ-GUESS/tree/initial

Communities: Zenodo

License (for files): [Other \(Open\)](#)

DATENVERÖFFENTLICHUNG/-NACHNUTZUNG

Datenrepositorien: <http://www.re3data.org/>

Registry of Research Data Repositories -
Verzeichnis von über 3000 Datenrepositorien



http://www.forschungsdaten.org/index.php/Datei:Re3data_icons.jpg

Browse by subject

Graphical Text

click to zoom into subjects or to select a bottommost subject in the hierarchy as filter for the re3data search page
ctrl + click on a top subject to select it as filter



re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

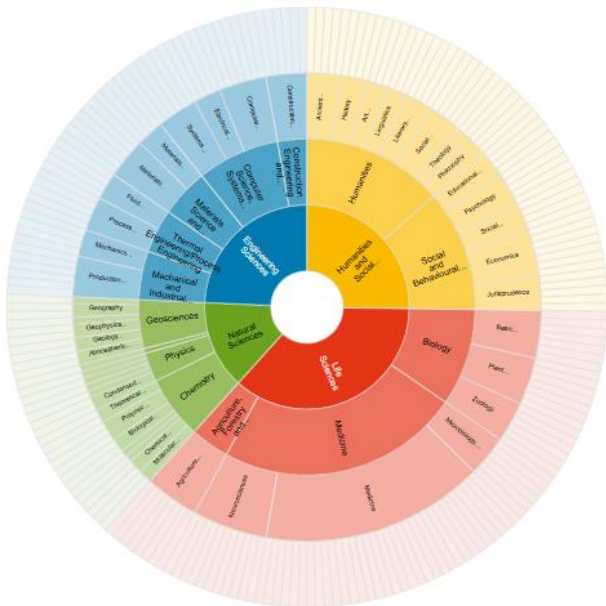
Geology

Search

Browse by subject

Graphical Text

click to zoom into subjects or to select a bottommost subject in the hierarchy as filter for the re3data search page
shift + click on a top subject to select it as filter



Browse by subject

Graphical Text

click to zoom into subjects or to select a bottommost subject in the hierarchy as filter for the re3data search page
shift + click on a top subject to select it as filter



d. Geosciences

I. Atmospheric Science, Oceanography and Climate Research

1. Atmospheric Science
2. Oceanography

II. Geology and Palaeontology

1. Geology
2. Palaeontology

III. Geophysics and Geodesy

1. Geophysics
2. Geodesy, Photogrammetry, Remote Sensing, Geoinformatics, Cartography

IV. Mineralogy, Petrology and Geochemistry

1. Mineralogy, Petrology and Geochemistry

V. Geography

1. Physical Geography
2. Human Geography

VI. Water Research

1. Hydrogeology, Hydrology, Limnology, Urban Water Management, Water Chemistry, Water Resources Management

Filter

[Reset all](#)

- [Subjects](#) ⊞
- [Content Types](#) ⊞
- [Countries](#) ⊞
- [AID systems](#) ⊞
- [API](#) ⊞
- [Certificates](#) ⊞
- [Data access](#) ⊞
- [Data access restrictions](#) ⊞
- [Database access](#) ⊞
- [Database access restrictions](#) ⊞
- [Database licenses](#) ⊞
- [Data licenses](#) ⊞
- [Data upload](#) ⊞
- [Data upload restrictions](#) ⊞
- [Enhanced publication](#) ⊞
- [Institution responsibility type](#) ⊞
- [Institution type](#) ⊞
- [Keywords](#) ⊞
- [Metadata standards](#) ⊞
- [PID systems](#) ⊞
- [Profiles](#) ⊞
- [Provider types](#) ⊞
- [Quality management](#) ⊞
- [Repository languages](#) ⊞
- [Software](#) ⊞
- [Syndications](#) ⊞
- [Repository types](#) ⊞
- [Versioning](#) ⊞

Geology

← Previous **1** 2 3 4 5 6 7 Next →

Found 175 result(s)

New Mexico Bureau of Geology & Mineral Resources Data Repository

Subject(s)	Natural Sciences Geosciences Geology and Palaeontology Geophysics and Geodesy
Repository type(s)	disciplinary
Provider type(s)	dataProvider
Country	United States

Archiving data and housing geological collections is an important role the Bureau of Geology plays in improving our understanding of the geology of New Mexico. Aside from our numerous publications, several datasets are available to the public. Data in this repository supplements published papers in our publications. Please refer to both the published material and the repository documentation before using this data. Please cite repository data as shown in each repository listing.

Index to Marine & Lacustrine Geological Samples

Subject(s)	Natural Sciences Geosciences Atmospheric Science, Oceanography and Climate Research Oceanography Geology and Palaeontology Geology Geophysics and Geodesy Geography
Repository type(s)	disciplinary Institutional
Provider type(s)	dataProvider serviceProvider
Country	United States

The Index to Marine and Lacustrine Geological Samples is a tool to help scientists locate and obtain geologic material from sea floor and lakebed cores, grabs, and dredges archived by participating institutions around the world. Data and images related to the samples are prepared and contributed by the institutions for access via the IMLGS and long-term archive at NGDC. Before proposing research on any sample, please contact the curator for sample condition and availability. A consortium of Curators guides the IMLGS, maintained on behalf of the group by NGDC, since 1977.

Missouri Geological Survey Program

Subject(s)	Natural Sciences Geosciences Geology and Palaeontology Geology Water Research Hydrogeology, Hydrology, Limnology, Urban Water Management, Water Chemistry, Integrated Water Resources Management
Repository type(s)	Institutional
Provider type(s)	dataProvider
Country	United States

We conduct integrated earth resource evaluation that incorporates sustainable economic vitality and environmental and public health protection.

Geo Big Data Open Platform

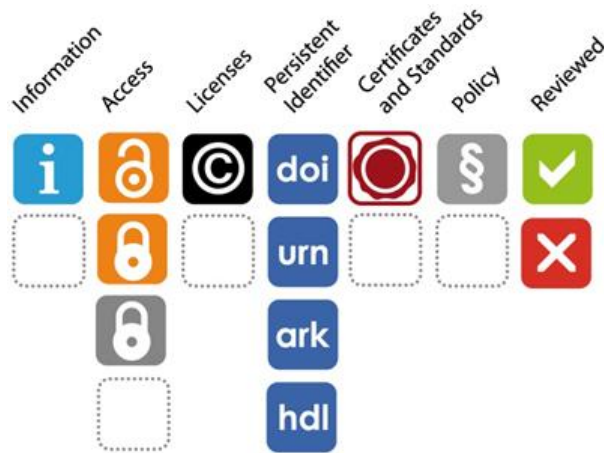
Subject(s)	Natural Sciences Geosciences Geology and Palaeontology Geology Geophysics and Geodesy Geophysics Geodesy, Photogrammetry, Remote Sensing, Geoinformatics, Cartography Mineralogy, Petrology and Geochemistry Mineralogy, Petrology and Geochemistry
Repository type(s)	disciplinary
Provider type(s)	dataProvider
Country	Korea, Republic of

The Geo Big Data Open Platform of the Korea Institute of Geological Resources is a data-based repository that allows anyone to easily access the latest geological resource information scattered in Korea. It was established for the purpose of quickly organizing and providing domestic and foreign geological resource research information pouring out of a super-gap society to utilize the solution of national social problems and create an open science research ecosystem in the geological resource field.

Alberta Geological Survey Open Data Portal

Subject(s)	Natural Sciences Geosciences Geology and Palaeontology Mineralogy, Petrology and Geochemistry Water Research
Repository type(s)	other

Example PANGAEA



http://www.forschungsdaten.org/index.php/Datei:Re3data_icons.jpg

Repository details



PANGAEA

General Institutions Terms Standards

Name of repository	PANGAEA
Additional name(s)	Data Publisher for Earth and Environmental Science
Repository URL	https://www.pangaea.de/
Subject(s)	Life Sciences Biology Natural Sciences Geosciences Atmospheric Science, Oceanography and Climate Research Oceanography Geology and Palaeontology Geology Geophysics and Geodesy Geophysics Mineralogy, Petrology and Geochemistry Mineralogy, Petrology and Geochemistry
Description	PANGAEA - Data Publisher for Earth & Environmental Sciences has an almost 30-year history as an open-access library for archiving, publishing, and disseminating georeferenced data from the Earth, environmental, and biodiversity sciences. Originally evolving from a database for sediment cores, it is operated as a joint facility of the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI) and the Center for Marine Environmental Sciences (MARUM) at the University of Bremen. PANGAEA holds a mandate from the World Meteorological Organization (WMO) and is accredited as a World Radiation Monitoring Center (WRMC). It was further accredited as a World Data Center by the International Council for Science (ICS) in 2001 and has been certified with the Core Trust Seal since 2019. The successful cooperation between PANGAEA and the publishing industry along with the correspondent technical implementation enables the cross-referencing of scientific publications and datasets archived as supplements to these publications. PANGAEA is the recommended data repository of numerous international scientific journals.
Contact(s)	administrative: https://www.pangaea.de/contact/ support: info@pangaea.de
Content type(s)	dataset recorded data
Certificate(s)	CoreTrustSeal Start date 2023-04-20 End date 2026-04-19
Keyword(s)	FAIR agriculture atmosphere biology biosphere cryosphere earth science ecology environmental science fisheries land surface lithosphere oceans paleontology
Persistent identifier(s) of the repository	FAIRsharing: 10.25504/FAIRsharing.6yw6cp
Repository type(s)	disciplinary
Mission statement for designated community	https://www.pangaea.de/about/
Research data repository language(s)	English
Data and/or service provider	dataProvider

RECHTLICHES: URHEBERRECHT

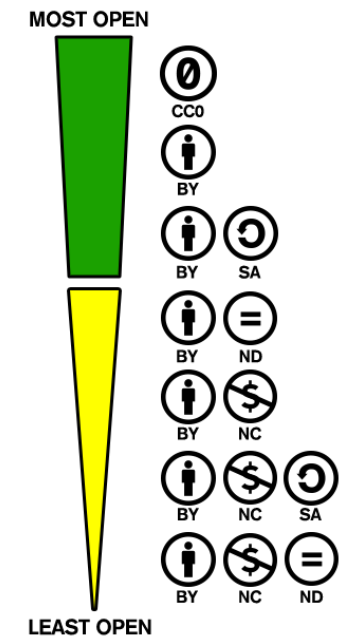
- Schützt **geistige Schöpfungen** (Elemente: Wahrnehmbare Formgestalt, persönliches Schaffen, geistiger Gehalt, eigenpersönliche Prägung)
- Einzelne Forschungsdaten unterliegen oft nicht dem Urheberrecht
- I.d.R. kein Urheberrecht auf unstrukturierte Messdaten
- Klären Sie rechtliche Fragen im Voraus (z.B. Rechte an Daten in Kooperationsprojekten)
- Datenveröffentlichung: Vergabe von Creative Commons Lizenzen
- Bei Benutzung von Fremddaten/-programmen:
 - Existieren Einschränkungen zur Verwendung der Daten?
 - Dürfen Daten/Programme in Archivierung/Veröffentlichung einbezogen werden?

<https://www.forschungsdaten.info/themen/rechte-und-pflichten/urheberrecht/>

RECHTLICHES: FREIE LIZENZEN

Bieten die Möglichkeit eine Bedingung für die Nachnutzung von Daten festzulegen

- CC bieten standardisierte Lizenzverträge
 - CC 0: Keine Rechte vorbehalten
 - CC BY: Namensnennung (gebräuchlich in Wissenschaft)
 - CC BY SA: Namensnennung - Share alike
 - CC BY ND: Namensnennung – Keine Bearbeitungen
- GNU General Public License (GPL) → Software
- Open Data Commons (ODC) → Datensammlungen



creativecommons.org, CC BY 4.0

FORSCHUNGSDATEN SOLLEN FAIR SEIN

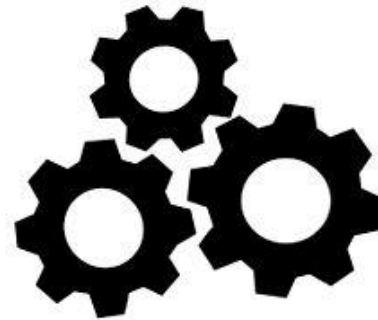
F
Findable



A
Accessible



I
Interoperable



R
Reusable




[FAIR Principles - GO FAIR](#)

WAS HEIßT FAIR IM EINZELNEN?



Findable: Leicht zu finden für Menschen und Computer

- Forschungsdatensätze mit Metadaten beschreiben
- Metadaten sind in einem durchsuchbaren Verzeichnis bzw. Repository indiziert
- Ggf. Publikation von Forschungsdaten
- Forschungsdaten (Metadaten) mit persistenten Identifikatoren (DOI, URN etc.) versehen

PANGAEA. 

ECOLOGY

18121 datasets found on search in topic Ecology


< 1 2 3 4 5 6 7 8 9 10 >

- Chouvelon, T; Munsch, C; Bruzac, S et al. (2021):** Stable isotopes, energy density and elemental composition of deep pelagic crustaceans and fish sampled in October 2017 in the Bay of Biscay, North-East Atlantic
Related to: **Duhamel, E; Pawlowski, L; Garren, F (2017):** EVHOE 2017 cruise, Thalassa R/V. *Sismer*
Size: 2 datasets
<https://doi.pangaea.de/10.1594/PANGAEA.937345> - Score: 1.8
- Verdugo, J; Damm, E (2021):** Biogeochemical data during POLARSTERN cruise PS106/1 (ARK-XXXI/1.1)
Related to: **Macke, A; Flores, H (2018):** The Expeditions PS106/1 and 2 of the Research Vessel POLARSTERN to the Arctic Ocean in 2017. *Berichte zur Polar- und Meeresforschung = Reports on Polar and Marine Research*
Nikolopoulos, A; Linders, T; Rohardt, G (2018): Physical oceanography during POLARSTERN cruise PS106/1 (ARK-XXXI/1.1). *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven*
Rohardt, G (2018): CTD Data Processing Report - RV POLARSTERN Cruise PS106.
Size: 2 datasets
<https://doi.pangaea.de/10.1594/PANGAEA.937242> - Score: 1.8

WAS HEIßT FAIR IM EINZELNEN?

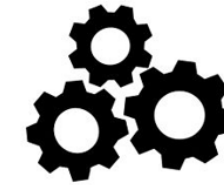


Accessible: Zugänglichkeit.

Coverage:	Median Latitude: 47.196250 * Median Longitude: 8.468551 * South-bound Latit
License:	 Creative Commons Attribution 4.0 International (CC-BY-4.0)
Size:	73 datasets

- Langfristige Speicherung (z.B. in Repositorien/Archiven),
- Definierte Lizenz- und Zugriffsbedingungen
- Leicht zugänglich (Metadaten und/oder eigentliche Daten)
- Metadaten auch dann vorhanden, wenn eigentliche Daten nicht zugänglich sind

WAS HEIßT FAIR IM EINZELNEN?



Interoperable: Die Daten liegen in einer Form vor, in der sie durch Menschen oder Computer mit anderen Datensätzen kombiniert werden können

- Verwendung allgemein üblicher und möglichst nicht-proprietärer, offener und beschriebener Formate
- Kontrolliertes Vokabular wie Schlagworte
- Verweise auf verwandte (Meta-)Daten

Empfohlene Formate

Tabellen: CSV, TSV, SPSS portable

Text: TXT, HTML, RDF, PDF/A

Bild: TIFF, JPEG2000, PNG

Multimedia: AVI, WAV, MP4

WAS HEIßT FAIR IM EINZELNEN?




Reusable: Die Daten liegen in einer Form vor, in der sie für weitere Forschung und die Verarbeitung mit rechnergestützten Methoden genutzt werden können

- nicht-proprietäre, offene Formate
- Definierte Nutzungslizenz
- Meta- und Forschungsdaten entsprechen Fachstandards

Empfohlene Formate
<u>Tabellen:</u> CSV, TSV, SPSS portable
<u>Text:</u> TXT, HTML, RDF, PDF/A
<u>Bild:</u> TIFF, JPEG2000, PNG
<u>Multimedia:</u> AVI, WAV, MP4

Coverage: Median Latitude: 47.196250 * Median Longitude: 8.468551 * South-bound Latit

License:  Creative Commons Attribution 4.0 International (CC-BY-4.0)

Size: 73 datasets

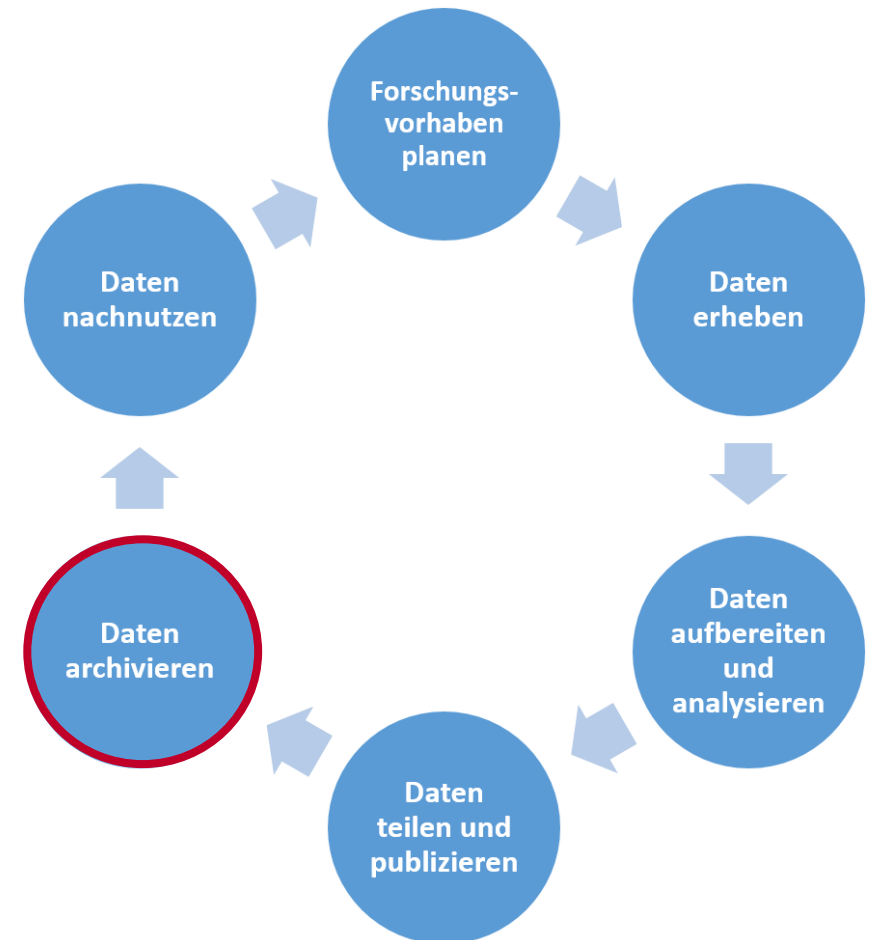
Data Cite (Metadata)

Basic Scheme

- Identifier
- Creators
- Title
- Publisher
- Publication year
- Subject
- Language
- Resource type
- Descriptions

DATEN ARCHIVIEREN

- Migration von Daten in das passendste Archivierungsformat (nicht proprietär)
- Metadaten und Datendokumentation (besser direkt bei Erstellung)
- Übergabe der Daten an Experten für die Langzeitarchivierung übergeben (Bibliotheken, Rechenzentren, Fach-Repositoryen)
- Dauerhafte Datenarchivierung in professioneller Archivierungsinfrastruktur



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

BACKUP VS. ARCHIVIERUNG

	Backup	Archivierung
Ziel	Aktiv genutzte Daten	Finale Daten
Zweck	Schutz und Auffindbarkeit	Erhalten und Auffinden von Information
Zeitraum	Kurzzeit	Langzeit



FORSCHUNGSDATENARCHIV DER JGU

Basiert auf der Software iRODS

Was ist/kann iRODS?

- Virtuelles Dateisystem
- Durchsuchbare Metadaten
- Publikationsmöglichkeit (ohne PID)
- Fein granulierte Zugangskontrolle
- Automatisierte Arbeitsabläufe
- Commandline Werkzeuge

[Archivierung von Forschungsdaten mit iRODS | Zentrum für Datenverarbeitung](#)

[iRODS Wiki – Forschungsdatenmanagement](#)

The screenshot shows the website of the Zentrum für Datenverarbeitung (ZDV) at the JGU. The page title is "Archivierung von Forschungsdaten mit iRODS". The header includes the ZDV logo and navigation links for HOME, SUCHEN, INDEX, THEMEN, KONTAKT, ENGLISH, SUPPORT, ÜBER UNS, and GLOSSAR. A search bar is visible with the text "SUCHE" and "Suche" and a "LOS" button. The main content area is divided into sections: "Zugang", "Experimenteller Web-Zugang", "ZDV unabhängige Systeme", and "Preisübersicht". Each section contains detailed information about the iRODS system, including access methods, system requirements, and pricing. The "Zugang" section mentions that access is currently only possible via Linux and requires Kerberos authentication. The "Experimenteller Web-Zugang" section lists three points: 1. The web interface can be changed but is not stable. 2. It is only for one file at a time. 3. Files must be smaller than 5GB. The "ZDV unabhängige Systeme" section notes that only centrally managed Debian systems are supported. The "Preisübersicht" section lists a "Kostenfreies Angebot" of 10 TB per working group and a "Mehrbedarf" of 50 EUR per year for 1 TB additional. Contact information "Kontakt: unix@uni-mainz.de" is provided at the bottom.

(LANGZEIT-)ARCHIVIERUNG - VORAUSSETZUNGEN

- Verwendung weit verbreiteter Formate
- Verwendung von Formaten mit offenem, dokumentierten und nicht proprietären Standard
- Vermeidung von Kompression
- Beschreibung mit Metadaten
- Migration der Daten in aktuelle Datenformate
- Nutzung von professionellen Infrastrukturen für die Archivierung

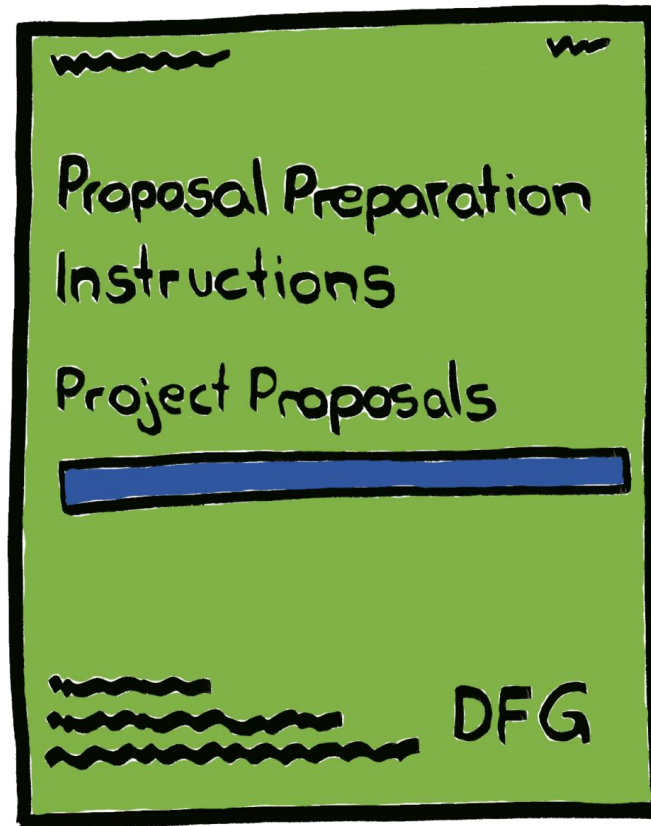
DATEN NACHNUTZEN

- Neue Forschungsfragen unter Nutzung vorhandenerer Daten
- Replikation von Forschungsergebnissen
- Hinterfragung von Ergebnissen
- Erstellung Forschungsrückblicken
- Datennutzung für Unterrichten und Lernen
- ...



[Bild: Forschungsdatenmanagement | Open Science | Unterseiten | Kommunikations-, Informations-, Medienzentrum \(KIM\)](#)

FDM-ANFORDERUNGEN DER DRITTMITTELGEBER



CC BY-SA 4.0
Franziska Mau
[FranziMachtDas –
Sketches, Zeug & der
tägliche Wahnsinn](#)

DFG - CHECKLISTE FORSCHUNGSDATEN

1. Datenbeschreibung
2. Dokumentation und Datenqualität
3. Speicherung und technische Sicherung während des Projektverlaufs
4. Rechtliche Verpflichtungen und Rahmenbedingungen
5. Datenaustausch und dauerhafte Zugänglichkeit der Daten
6. Verantwortlichkeiten und Ressourcen

Orientierung an den Standards und Richtlinien der jeweiligen Fachdisziplinen!

Beschreibung des Umgangs mit Daten verpflichtend.

HORIZONT EUROPA - FDM

- FDM in Übereinstimmung mit den **FAIR-Prinzipien**
- **Datenmanagementplan (DMP)** zu M6 in der Mitte und Ende des Projekts
- Ablage von (Meta)daten so früh wie möglich nach Produktion, Prozessierung o. Qualitätskontrolle
- Ablage in Vertrauenswürdigen Repositorien. Offene Daten so früh wie möglich (DMP)
 - „Daten so offen wie möglich, so geschlossen wie nötig“, d.h. es kann es Ausnahmen vom offenen Zugang zu Forschungsdaten geben. (Datenschutz, Patentierung etc.)
- Datenzugang geschlossen, wenn nötig. Metadaten unter CC0 Lizenz veröffentlicht
- **Daten unter offenen Lizenzen, vorzugsweise CC0 oder CC-BY**
- Detaillierte Information/Dokumentation zu den Daten → Software Nutzung ...
- Kosten für FDM sind förderfähig

Informationen: [Open Research Europe publishing platform](#) [Guides \(openaire.eu\)](#)

BMFTR - AKTIONSPLAN FORSCHUNGSDATEN

- Berücksichtigung der **FAIR-Prinzipien**
- Anträge/ Drittmittelprojekte: Bei Förderung von Datenerhebungen und -sammlungen, Verpflichtung die im Projekt erhobenen Daten systematisch, möglichst standardbasiert zu verarbeiten und aufzubewahren
- Berücksichtigung von Datenpublikationen nach den FAIR-Standards als wissenschaftliche Leistung bei der Begutachtung von Förderanträgen
- Notwendige Ausgaben für FDM sind förderfähig
- **DMPs teilweise** verpflichtend (v.a. Bildungsforschung)

VOLKSWAGENSTIFTUNG

- DMP als Antragsanlage erwartet, wenn viele Daten erzeugt werden.
https://www.volkswagenstiftung.de/de/faqservice#download_formulare_berichte
- Data Reuse: Förderprogramm zur Veröffentlichung von Forschungsdaten aus VW-Stiftungsprojekten
<https://www.volkswagenstiftung.de/de/foerderung/foerderangebot/data-reuse-zusaetzliche-mittel-fuer-die-aufbereitung-von-forschungsdaten>
- Open Science Policy
https://www.volkswagenstiftung.de/sites/default/files/documents/Open_Science_Policy_und_Umsetzung_VolkswagenStiftung.pdf

IHRE ANSPRECHPARTNER:INNEN AN DER JGU

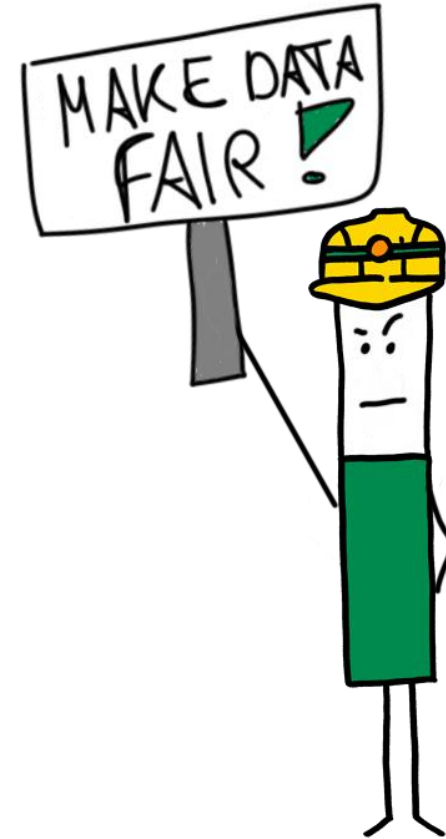
Kompetenzteam Forschungsdaten: forschungsdaten@uni-mainz.de

- **Dr. Anne Vieten** (Abteilung Forschung und Technologietransfer – Allg. FDM-Aspekte, Anforderungen der Drittmittelgeber, DMPs...)
- **Dr. Jörg Steinkamp** (Zentrum für Datenverarbeitung – Technische FDM-Aspekte, GitLab, Datenarchivierung in iRODs ...)
- **Sarah Wettermann** (Zentrum für Datenverarbeitung – RDMO-DMPs, Technische FDM-Aspekte,...)
- **Karin Eckert** (Universitätsbibliothek - Metadaten, Gutenberg Open Science,...)
- **Esther Reinecke und Tina Rotzal** (Universitätsbibliothek- Gute wissenschaftliche Praxis, akademische Integrität)

<https://www.forschungsdaten.uni-mainz.de/>

VIELEN DANK FÜR IHRE AUFMERKSAMKEIT!

Fragen?



[Abb.: Sketchnotes & Sketches – FranzMachtDas_CC BY-SA 4-0.png](#) CC-BY-SA 4.0 Franziska Mau

ZUSATZINFORMATIONEN

Erklärvideos in englischer Sprache der UGent:

[Knowledge clip: The research data lifecycle](#)

[Knowledge clip: What is Research Data Management \(RDM\)?](#)

[Knowledge clip: Data Management Plans \(DMPs\)](#)

[Knowledge clip: Keeping research data organized](#)

[Knowledge clip: file formats](#)

[Knowledge clip: Data Security](#)

[Knowledge clip: Data Documentation](#)

[Knowledge clip: Metadata](#)

[Knowledge clip: Preserving data](#)

[Knowledge clip: FAIR data principles](#)

[Knowledge clip: Data repositories](#)