

FROM PLANNING TO PUBLICATION

Effective Management of Research Data

Gernot Deinzer | Constantin Lehenmeier | Sophie Stolzenberger
UR Data Hub, 15.12.2025



<https://epub.uni-regensburg.de/78308>



Why „manage“?

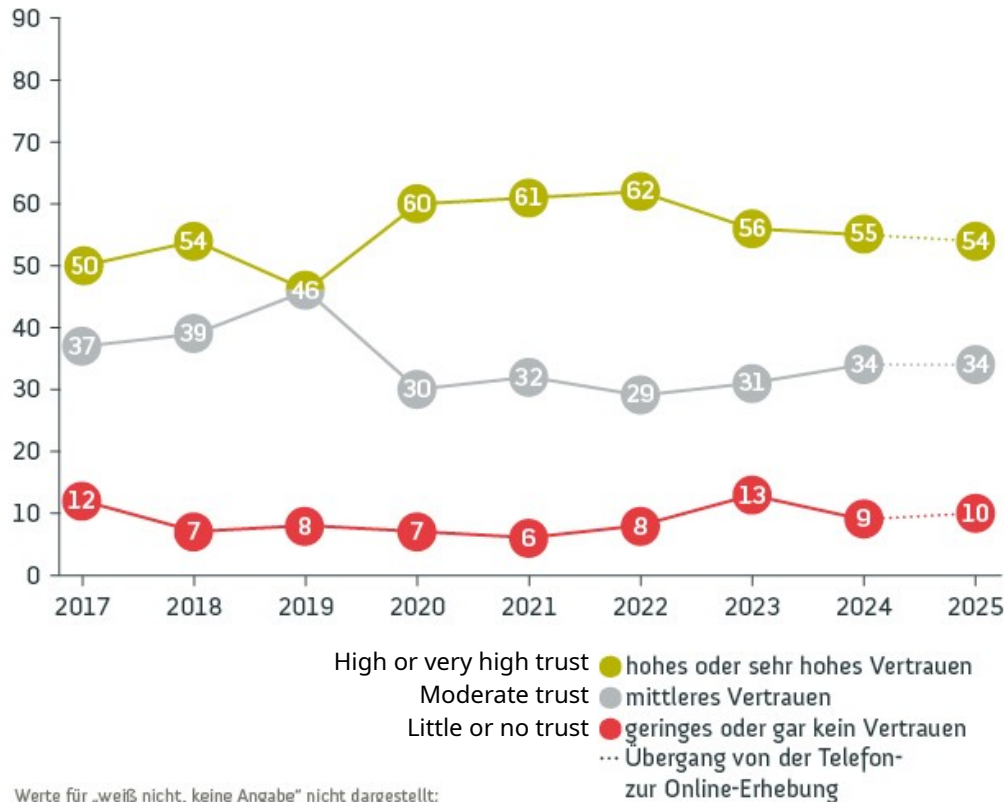
TRANSPARENCY AND REPRODUCIBILITY

Public trust in science (in Germany)



Werte für „weiß nicht, keine Angabe“ nicht dargestellt;
Basis: jeweils mindestens 1.000 Befragte;
Angaben in Prozent – Rundungsdifferenzen möglich
Quelle: Wissenschaftsbarometer – Wissenschaft im Dialog/Verian

Public trust in science (in Germany)



Werte für „weiß nicht, keine Angabe“ nicht dargestellt;
Basis: jeweils mindestens 1.000 Befragte;
Angaben in Prozent – Rundungsdifferenzen möglich
Quelle: Wissenschaftsbarometer – Wissenschaft im Dialog/Verian

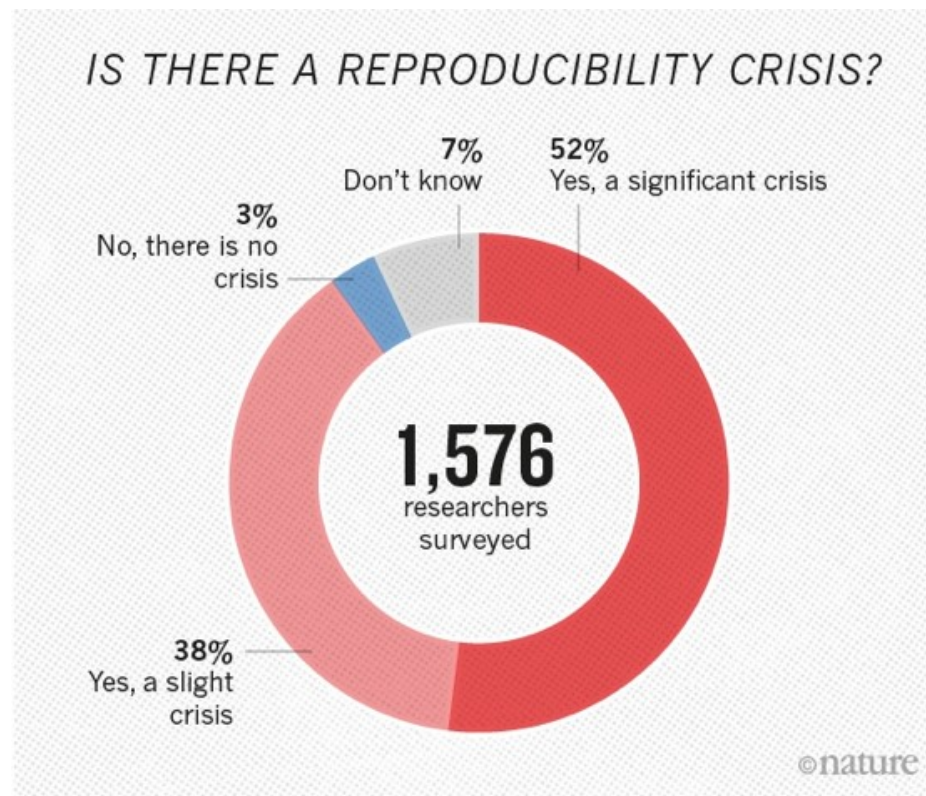
Reasons for trust (2024)

- Scientists' expertise (67% ↑)
- **Rule-based, standards oriented work** (55% ↓)
- Researching in the public interest (43% ↓)

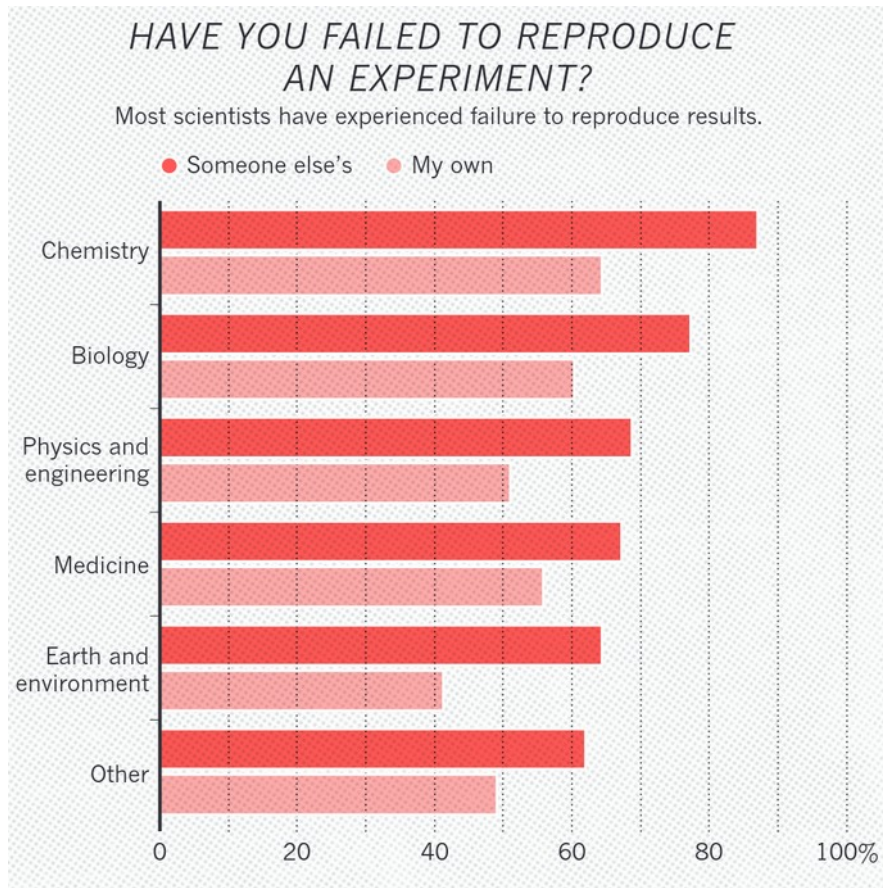
Reasons for mistrust (2024)

- Researchers' dependence on funders (62% ↑)
- Often adjust results to expectations (34% ↑)
- Researchers frequently make mistakes (21% ↑)

Replication crisis



Replication crisis

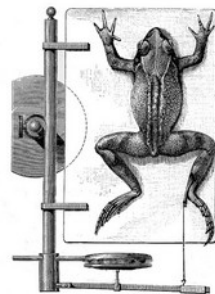
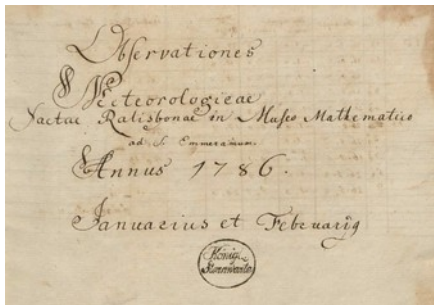


Identified reasons for irreproducible research:

- Methods and code not available
- Raw data not available
- ...

Data culture crisis

Humanists may not be used to thinking of their research materials as "data"



Achaischer Bund



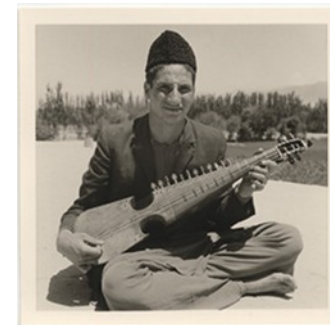
Anemurion



Dyrrhachion



Karthago



Research materials are viewed as "sources/objects" not as data with structure, metadata, and lifecycles

Research Data

DEFINITION

Spectrum of research data

“Research data is an essential foundation for scientific work.” (DFG Guidelines on Handling Research Data, 2015)

Spectrum of research data

“Research data is an essential foundation for scientific work.” (DFG Guidelines on Handling Research Data, 2015)

“Research data might include **measurement data, laboratory values, audiovisual information, texts, survey data, objects from collections, or samples** that were created, developed or evaluated during scientific work.”

“Methodical forms of testing such as **questionnaires, software and simulations** may also produce important results for scientific research and should therefore also be categorised as research data.”

Research Data Management

Research Data Management encompasses all **measures** that ensure the usability of (digital) **research data**

Research Data Management

Research Data Management encompasses all **measures** that ensure the usability of (digital) **research data**



More precisely

Research Data Management refers to "all data **practices, manipulations, enhancements** and **processes** that ensure that research data are of a high quality, well **organised**, clearly **documented, preserved, sustainable, accessible** and **reusable**"

F

Findable



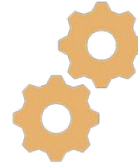
A

Accessible



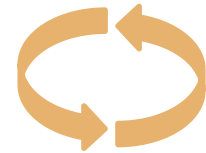
I

Interoperable



R

Reusable



F

Findable



PID; rich **metadata**
indexed in a
searchable resource

A

Accessible



Retrievable via a
standard protocol;
protocol open/free

I

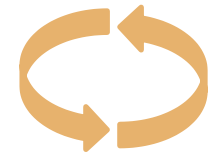
Interoperable



Formal, shared
languages /
standards; FAIR
vocabularies

R

Reusable



Clear/open license;
community
standards

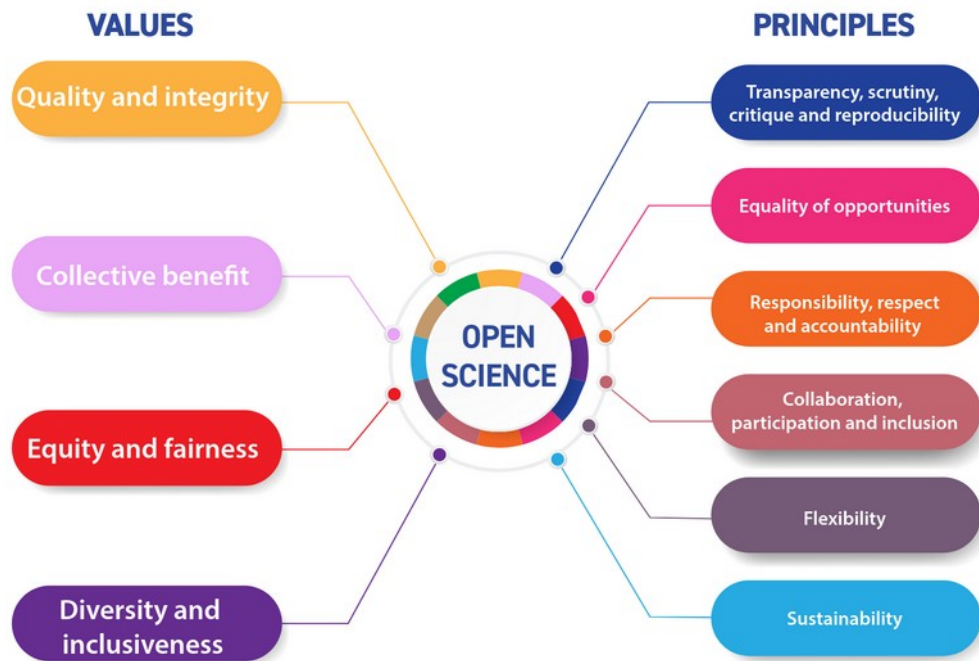
UNESCO Recommendation on Open Science

- Internationally agreed definition of Open Science (193 countries agreed)



UNESCO Recommendation on Open Science

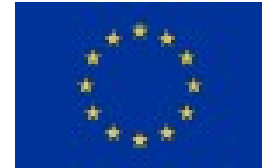
- Internationally agreed definition of Open Science (193 countries agreed)
- Shared values and guiding principles



Research funding

Horizon Europe

- **Mandatory data management** (DMP required)
- Principle: "as open as possible, as closed as necessary"
- Justification required if data are not open



Deutsche Forschungsgemeinschaft DFG Guidelines for applicants:

Emphasis on

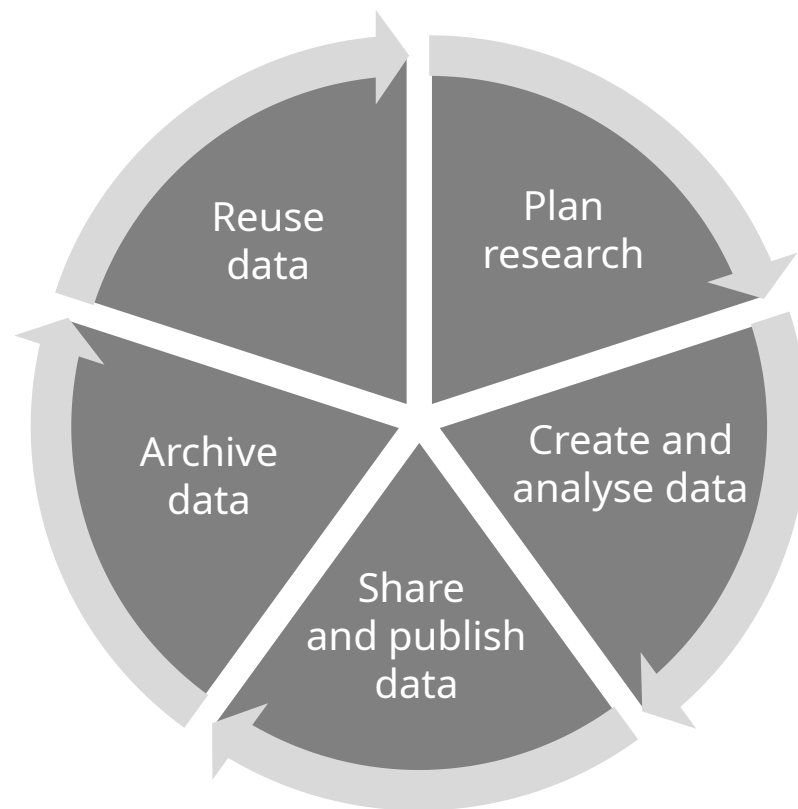
- project planning and submission of proposals
- data accessibility
- long-term preservation



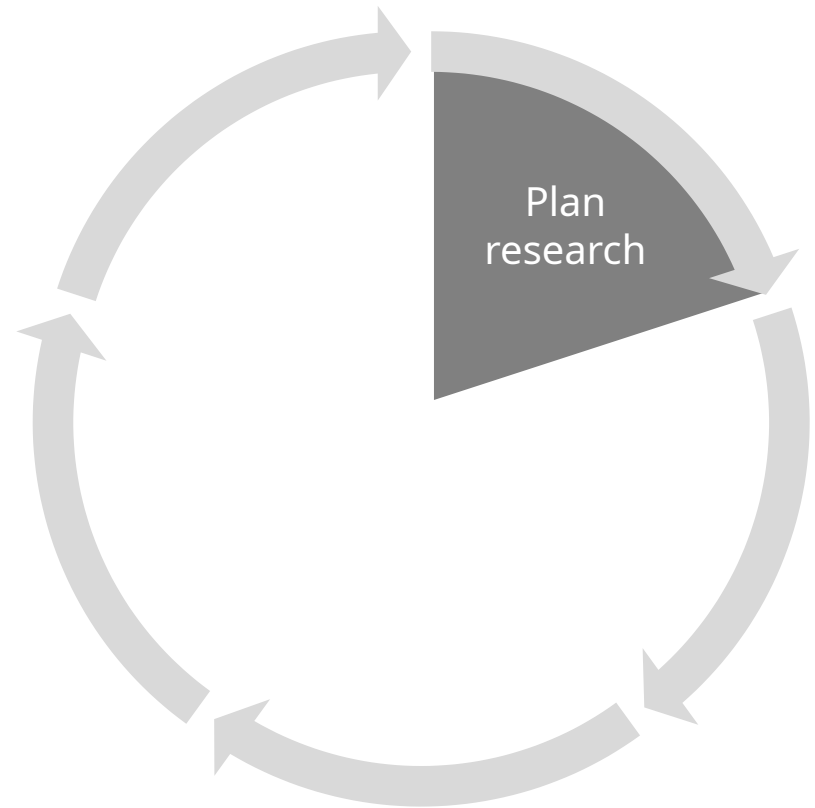
Research Data

RESEARCH DATA MANAGEMENT

Research data lifecycle



Research data lifecycle



Digital tools for support

PLAN RESEARCH

Data Management Plans

Data Management Plans structure the **handling of research data during and after the project period**



DATA MANAGEMENT PLAN

Data Management Plans

Includes, among other things:

- What data will be produced?
- How will the data be created?
- How will the data be documented?
- Which legal conditions (data protection & copyright) apply?
- Where will the data be stored?

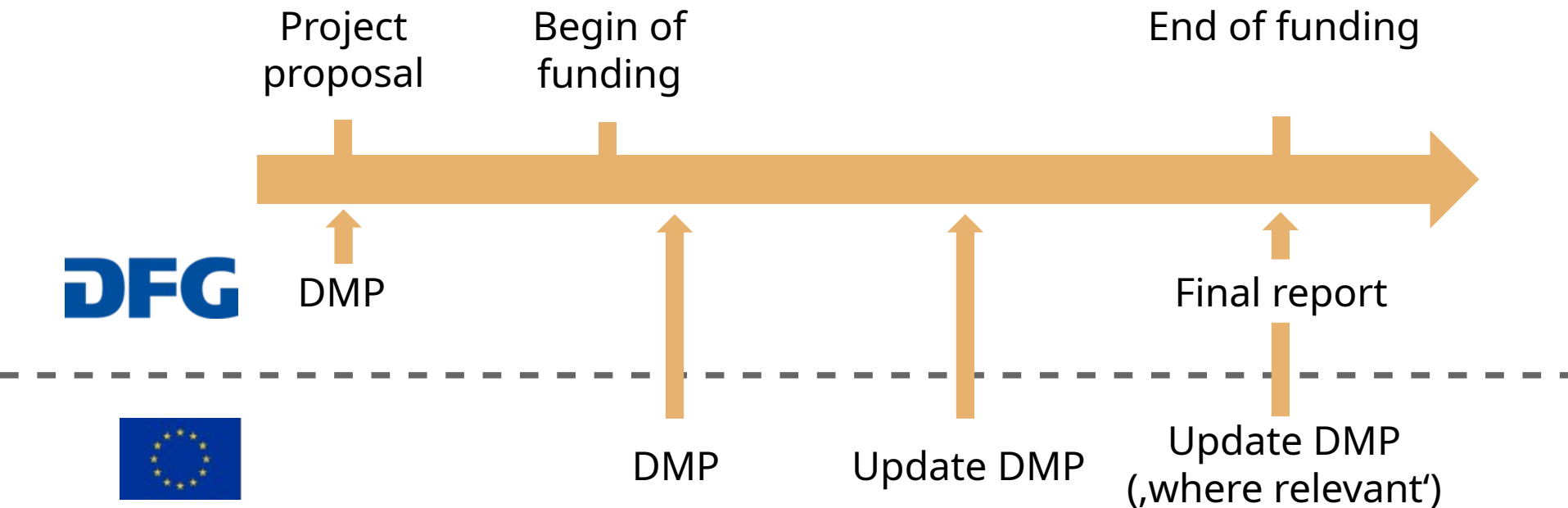
Data Management Plans

Includes, among other things:

- What data will be produced?
- How will the data be created?
- How will the data be documented?
- Which legal conditions (data protection & copyright) apply?
- Where will the data be stored?

→ Defining research data practices in your project and identifying risks and challenges

Data Management Plan for research proposals



Online tool for Data Management Plans



Research Data Management Organiser (RDMO)

- Open-source tool to support the structured planning and management of research data
- Deployed at multiple research institutions
- Multiple question catalogues can be supported
 - DFG
 - Horizon Europe and discipline-specific catalogues are in prospect

RDMO at the UR

RDMO@UR

Language ▾

Login



RDMO

A tool to support the planning, implementation, and organisation of research data management.

Campus der Universität Regensburg / Foto: Universität Regensburg

Welcome to the RDMO instance of the Universität Regensburg

This service is currently under development. Registration is not possible at the moment.

Login (RZ-Account)

Username

Password

RDMO: Create projects

RDMO@UR

Language ▾

John Smith ▾

DFG Projekt "Forschungsdaten"

Description	No description available.	
Catalog	DFG Proposal UR Catalog according to the DFG checklist (December 21, 2021) adapted to the UR infrastructure	

Views

Views are created using the answers given in the project and can then be exported in various formats. Initially, all views are empty. Please answer some questions by visiting **Answer Questions** (at the top of the sidebar).

View	Description	
DFG proposal UR	DFG Template including text modules on basic UR infrastructure	

Members

Here you can see who can access the project and invite additional members. You can use the user roles to manage which rights the benefits have. Unless you are the last owner, you can leave the project with the button next to your name.

User	E-Mail	Role	
John Smith		Owner	

Snapshots

Snapshots allow you to save all responses at a given point in time and preserve a certain stage of the project. Later the snapshot can

Options

[Answer questions](#)[View answers](#)[Update project information](#)[Update project catalog](#)[Update parent project](#)[Update project views](#)[Delete project](#)[Add member](#)[Create snapshot](#)[Back to projects overview](#)

Export

[CSV](#)[JSON](#)

RDMO: Answer questions interactively

RDMO@UR [Back to project](#)

Language ▾

John Smith ▾

[My Projects](#) / [DFG Projekt "Forschungsdaten"](#) / [Data description](#)

Contentwise data description

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.

Interviews

[+ Set](#)

What kind of data is collected or how is it collected?

Please choose from the following categories and briefly describe the data:

☐ observations:☐ surveys:☒ interviews:

We will conduct a series of interviews with 20 participants.

Overview

Project: DFG Projekt "Forschungsdaten"**Catalog:** DFG Proposal | UR[Reload page](#)[Back to my projects](#)

Progress

1 of 29

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

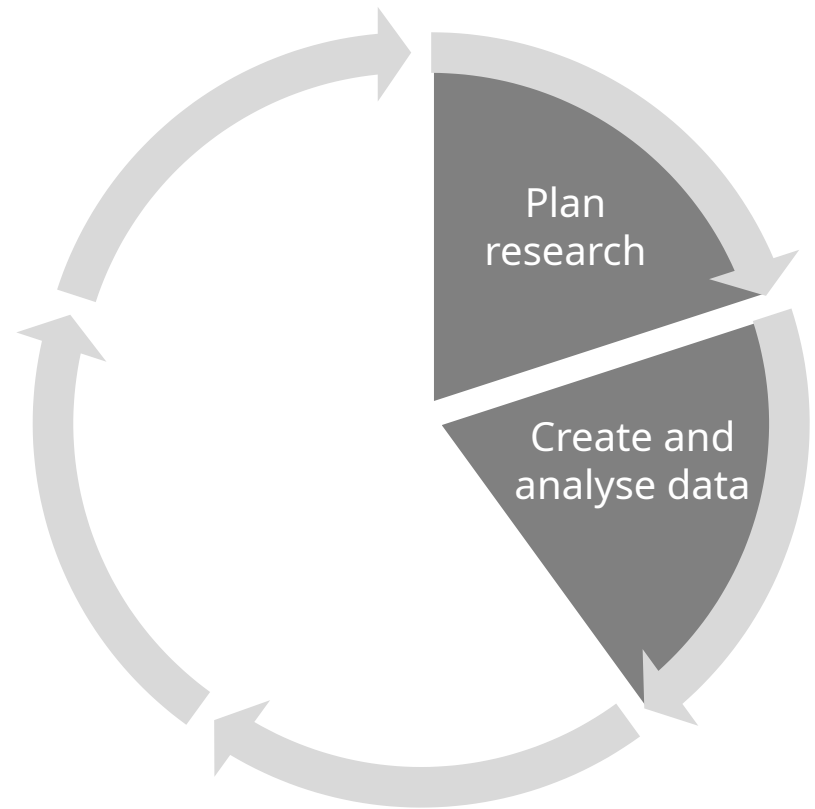
Navigation

Using the navigation will save your input.

Grey entries will be conditionally skipped based on your input.

[Introduction](#)[Data description \(1 of 6\)](#)[→ Contentwise data description \(1 of 4\)](#)[Technical Data Description](#)[Documentation and data quality](#)[Storage and technical protection measu...](#)[Legal obligations and conditions](#)

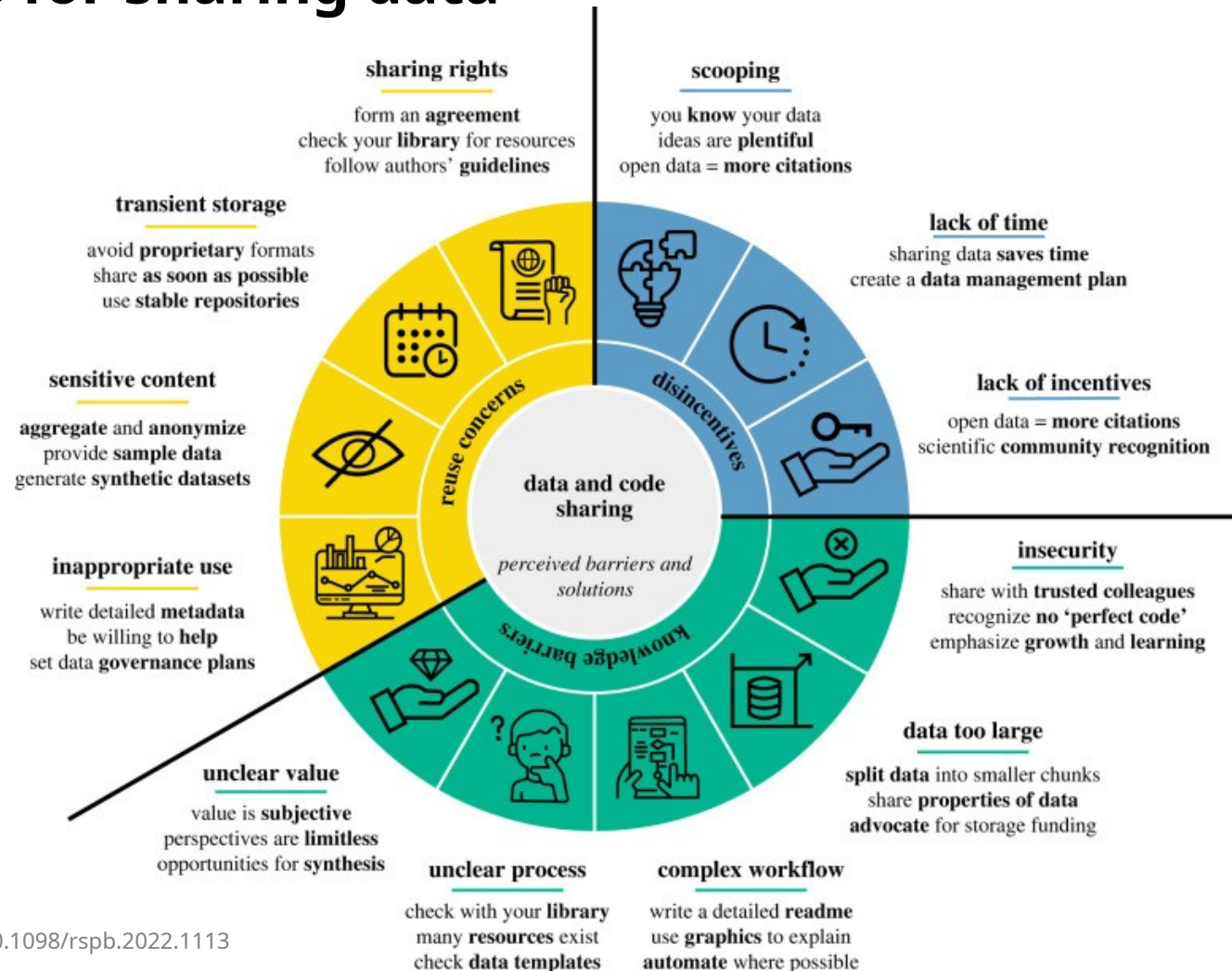
Research data lifecycle



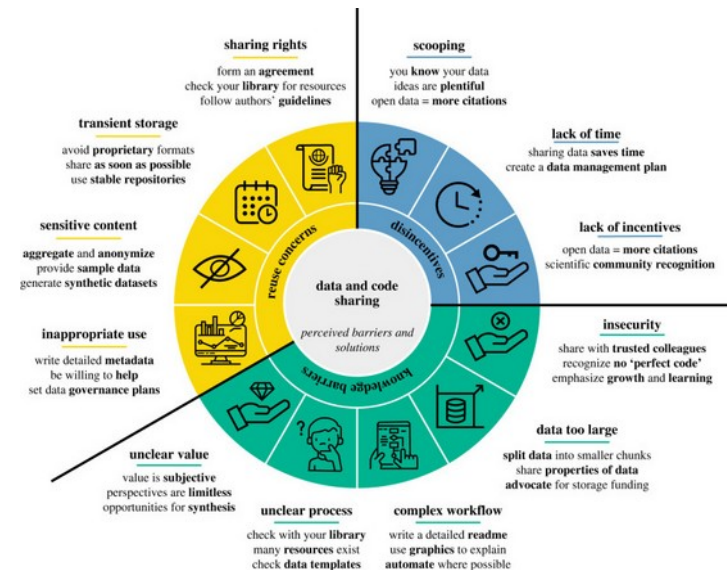
Digital tools for support

RESEARCH PROCESS

Barriers for sharing data



(Digital) Documentation



- **Transparency and traceability** start at data generation
- Documentation of workflows is the **beginning of the scientific value creation**
- Publishing **results alone is often no longer sufficient** today

README files

Overview of data and key contexts needed for understanding

This README.txt file was generated on 2024-02-19 by Naiara Korta Martiartu (naiara.korta@unibe.ch)

----- GENERAL INFORMATION -----

Contributor information:

Name: Naiara Korta Martiartu
Role/Function: Principal Investigator, Data Collector (main contact person)
Institution: Institute of Applied Physics, University of Bern
Address: Sidlerstrasse 5, 3012 Bern, Switzerland
Email: naiara.korta@unibe.ch

Name: Parisa Salemi Yolgunlu
Role/Function: Data Collector
Institution: Institute of Applied Physics, University of Bern
Address: Sidlerstrasse 5, 3012 Bern, Switzerland

Name: Martin Frenz
Role/Function: Sponsor
Institution: Institute of Applied Physics, University of Bern
Address: Sidlerstrasse 5, 3012 Bern, Switzerland

Name: Michael Jaeger
Role/Function: Principal Investigator (alternative contact person)
Institution: Institute of Applied Physics, University of Bern
Address: Sidlerstrasse 5, 3012 Bern, Switzerland
Email: michael.jaeger@unibe.ch

Date of data collection: December 2022 - May 2023

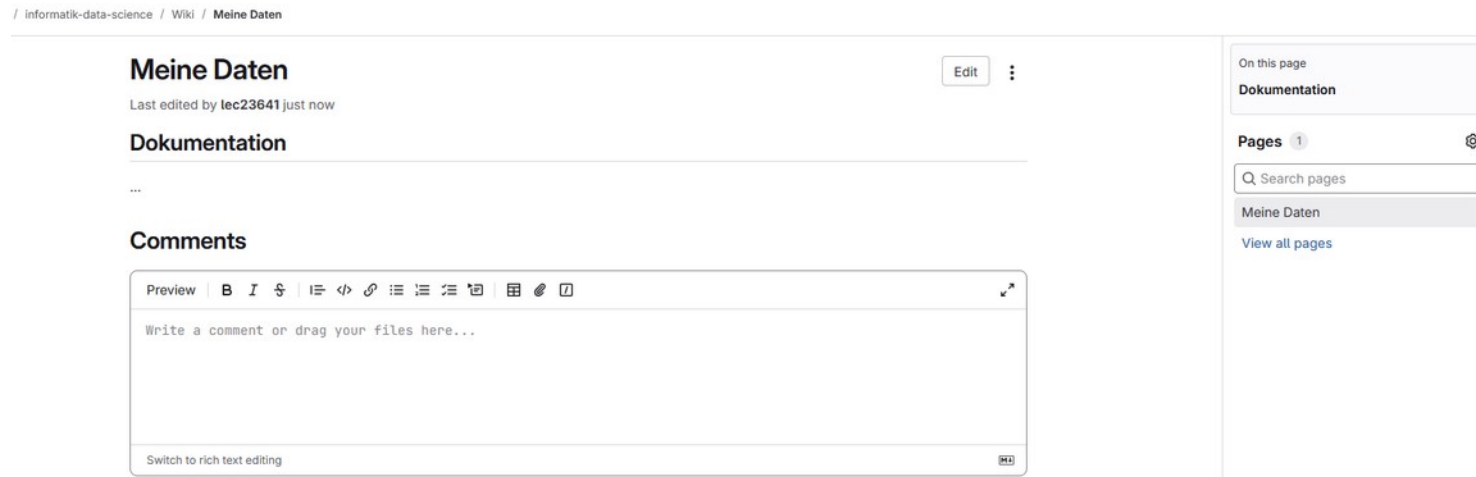
Location of data collection: Institute of Applied Physics, University of Bern, Bern, Switzerland

Keywords: pulse-echo ultrasound, attenuation imaging, tissue characterization, ultrasound tomography

Funding sources: N. K. M. received support of the University of Bern through the UniBe Initial Grant 205320_179038

----- BUILDING ACCESS INFORMATION -----

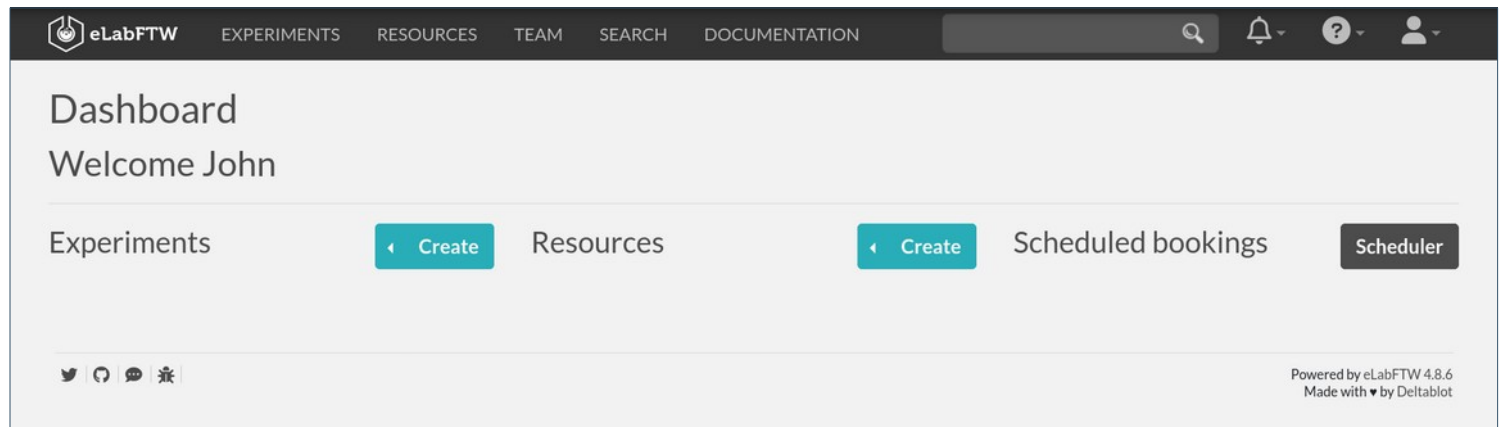
Using a wiki



Using a GitLab Wiki for data documentation

- Low barrier: simple Markdown, previews
- Versioned and reviewable: full history, diffs, approvals
- Team-ready: roles, permissions, collaborations

Digital notebook *eLabFTW*



Experiment and resources

- Structured metadata in a machine-readable format
- Import and export of records

Team management

- Permissions and collaboration controls
- Basic project management

Useful features (e.g., search, tagging, versioning, API)

Create experiments in eLabFTW

TEI XML workflow for digitized letters (1910–1912)



Owned by [REDACTED]

Started on 2025-10-14

Team DEMO

Category NOT SET

Status RUNNING

Tags digital edition letters tei

ID 9845

Visibility + Only members of the team

Can write + Only owner and admins

MAIN TEXT

Objective

Establish a reproducible workflow to convert digitized letters (1910–1912) into TEI-XML with basic entity markup (persons, places)

Sources

- Image scans (TIFF, 600 dpi), shelfmarks: UR-MS-L-1910-001 ... 050

Create experiments in eLabFTW

TEI XML w

← |  

Owned by 

Started on 2025-10-1

Team

DEMO

Category

NOT SET

Status

RUNNING

Tags

digital ec

ID

9845

Visibility

+

Can write

+

MAIN TEXT

Display main text ☒

File Edit View Insert Format Tools Table

Bold

18pt

B

I

U











x^2

x_2









Objective

Establish a reproducible workflow to convert digitized letters (1910–1912) into TEI-XML with basic entity markup (persons, places) and publishable metadata.

Sources

- Image scans (TIFF, 600 dpi), shelfmarks: *UR-MS-L-1910-001 ... 050*

MAIN TEXT

Objective

Establish a reproducible workflow to convert digitized letters (1910–1912) into TEI-XML with basic entity markup (persons, places) and publishable metadata.

Sources

- Image scans (TIFF, 600 dpi), shelfmarks: *UR-MS-L-1910-001 ... 050*

Browser platform for programming: *JupyterHub*

The screenshot displays the JupyterHub web interface. On the left, a sidebar shows a file browser with a search bar and a list of files and folders. The main area is divided into several panes:

- Launcher**: Shows the current notebook, 'Lorenz.ipynb', and other open tabs like 'README.md', 'Terminal 1', 'Console 1', and 'Data.ipynb'.
- Code**: The main editor area for the notebook. It contains the title 'The Lorenz Differential Equations', an introductory paragraph, a code cell with imports, and a plot of the Lorenz attractor.
- Output View**: A pane showing the output of the code cell, including a 3D plot of the Lorenz attractor and a table of parameters.
- lorenz.py**: A pane showing the source code for the 'lorenz.py' file.

The code in the 'Code' pane is as follows:

```
[1]: %matplotlib inline
from ipywidgets import interactive, fixed

We explore the Lorenz system of differential equations:


$$\dot{x} = \sigma(y - x)$$

```

The 'lorenz.py' file contains the following code:

```
1 from matplotlib import pyplot as plt
2 from mpl_toolkits.mplot3d import Axes3D
3 import numpy as np
4 from scipy import integrate
5
6 def solve_lorenz(sigma=10.0, beta=8./3, rho=28.0):
7     """Plot a solution to the Lorenz differential
8     equations."""
9
10    max_time = 4.0
11    N = 30
12
13    fig = plt.figure()
14    ax = fig.add_axes([0, 0, 1, 1], projection='3d')
15    ax.axis('off')
```

The 'Output View' pane shows the following parameters:

Parameter	Value
sigma	10.00
beta	2.67
rho	28.00

The 3D plot shows the Lorenz attractor, a complex, chaotic trajectory in a three-dimensional space.

At the bottom of the interface, there is a status bar showing 'Simple', '1', '3', 'Python', and 'Ln 1, Col 1 Spaces: 4 lorenz.py'.

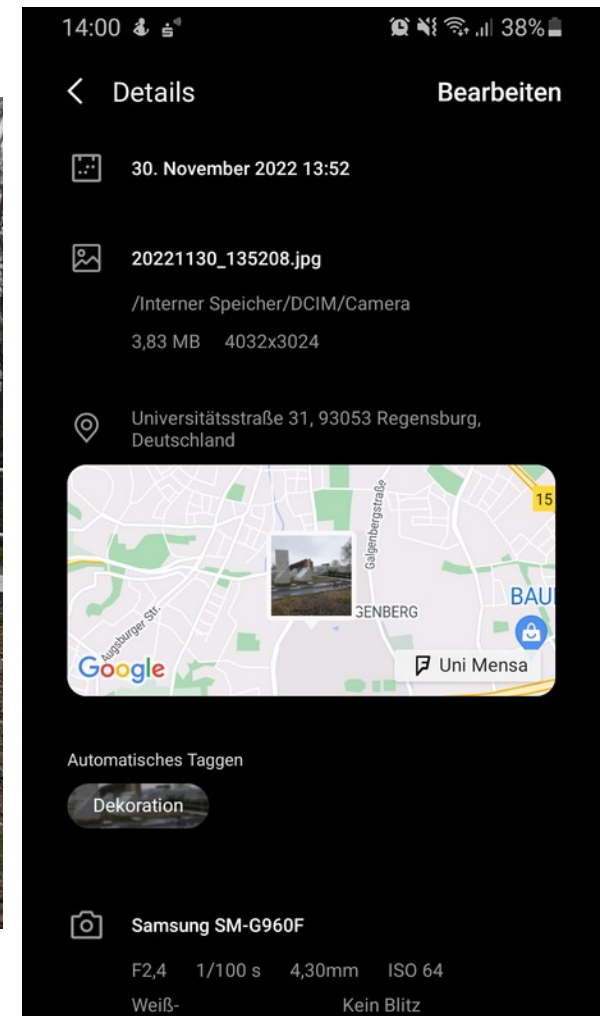
Metadata

DESCRIPTION OF THE DATA

What can you see on this picture?



What can you see on this picture?



Metadata

„Metadata is data about data“

Metadata is necessary for:

- Findability of data
- Accessibility
- Interoperability & exchangeability
- Reusability & long-term preservation



Metadata

„Metadata is data about data“

Metadata is necessary for:

- Findability of data
- Accessibility
- Interoperability & exchangeability
- Reusability & long-term preservation



Metadata for a digital photo:

- Creator (photographer)
- Description / caption
- Location
- Date
- Rights (copyright, licence)
- Image size / resolution
- File format
- ...

Metadata formats

Different **standard formats** offer **generic as well as subject specific, rich and structured description of data**

<i>Text annotation</i>	Text Encoding Initiative (TEI)
<i>Image data</i>	IIIF
<i>Vocabulary</i>	GND, VIAF, AAT
<i>Geo data</i>	GeoJSON
<i>Bio data</i>	Darwin Core
<i>Data collection</i>	Digital Documentation Initiative (DDI)



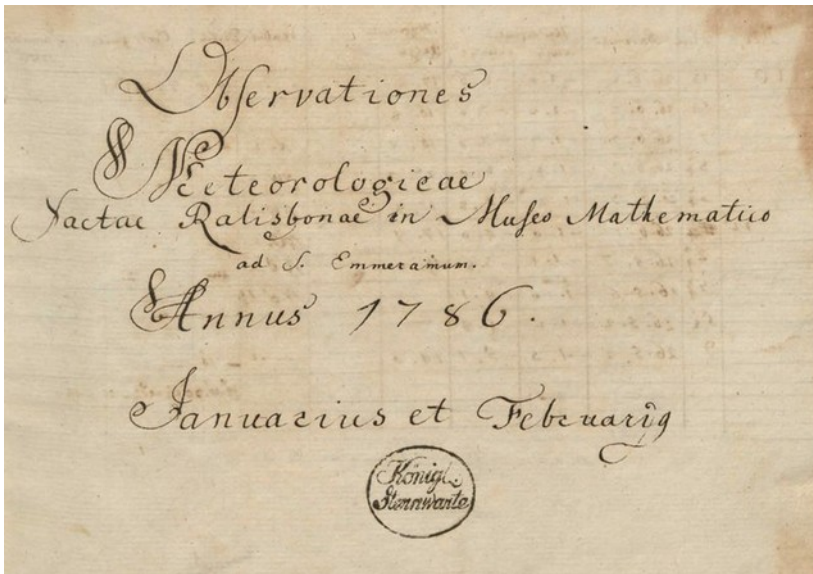
Chose formats that are known to support rich and structured metadata (*e.g. the format recommended by your discipline's community*)

Text annotation

Text Encoding Initiative (TEI)

Goal: A standardized, open format for scholarly text editions, encoding and exchanging texts

- Metadata (*title, authors*)
- Document structure (*chapters, paragraphs*)
- Logical & semantic markup (*person names, places, events*)

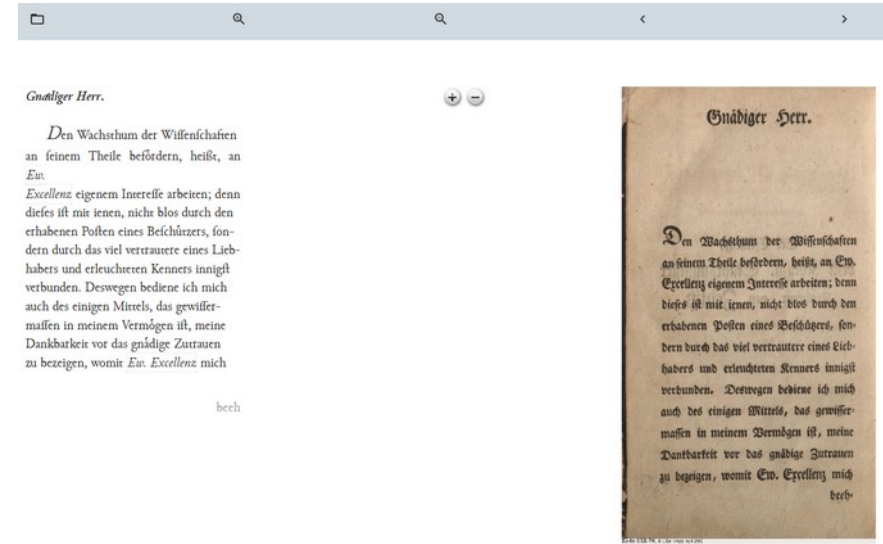
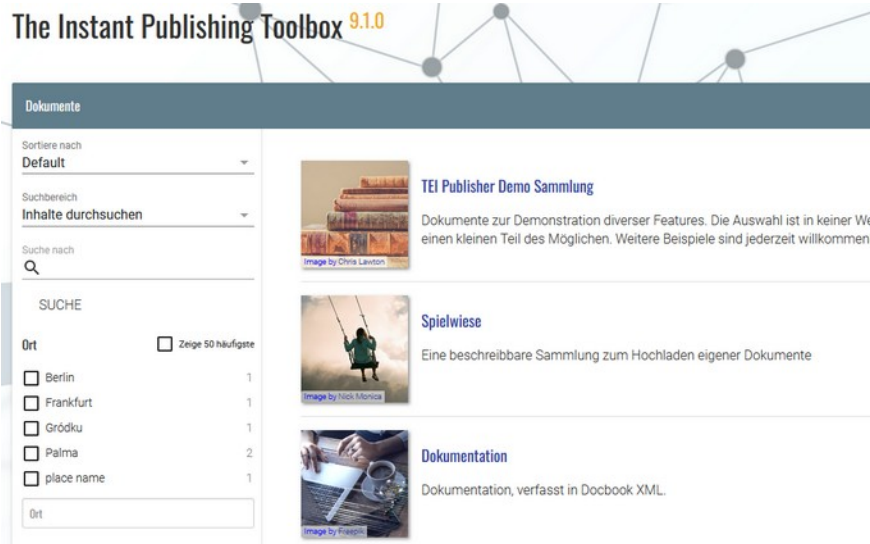


```
1 <TEI xmlns="http://www.tei-c.org/ns/1.0">
2   <teiHeader>
3     <fileDesc>
4       <titleStmt>
5         <title type="main">Observationes meteorologicae</title>
6       </titleStmt>
7       <publicationStmt>
8         <publisher>Universitätsbibliothek Regensburg</publisher>
9       </publicationStmt>
10      <sourceDesc>
11        <p>document creator: constantin.lehenmeier@ur.de</p>
12      </sourceDesc>
13    </fileDesc>
14  </teiHeader>
15  <facsimile>
16    <surface ulx="0" uly="0" lrx="183" lry="2582" xml:id="facs_1">
17      <lb n="1"/>Observationes
18      <lb n="2"/>Meteorologicae
19      <lb n="3"/>Factae <placeName ref="https://explore.gnd.network/en/gnd/4048989-9">
Ratisbonae</placeName> in Museo Mathematico
20      <lb n="4"/>ad S. Emmeramum.
21      <lb n="5"/>Annus 1786.
22      <lb n="6"/><date when="1786-01-01">Ianuarius et Februario</date>
23      <lb n="7"/>Königl.
24      <lb n="8"/>Sternwarte
25    </surface>
26  </facsimile>
27 </TEI>
```

Publish TEI

TEI Publisher

Web application for publishing and annotating TEI data (*soon to be hosted by the UR Data Hub*)



More Tools: Galaxy Europe

Web-based toolbox for FAIR data analysis and presentation

- Process and analyse data
- Create custom workflows or reuse existing workflows
- Visualization tools

The screenshot displays the Galaxy Europe web interface. The top navigation bar includes the 'Galaxy Europe' logo and a search bar. Below the navigation bar, there are tabs for 'My workflows', 'Workflows shared with me', and 'Public workflows'. The 'Public workflows' tab is selected, showing a grid of workflow cards. Each card contains the workflow name, the user who created it, a brief description, the tools used (e.g., unicycler, assembly, annotation, prokka, snippy, mutants, miRNA, Quast_Busco, VGP6b), and a 'Run' button. The workflows are sorted by 'Name' and 'Update time'. The left sidebar contains navigation links for 'Upload', 'Tools', 'Workflows', 'Visualization', 'Histories', and 'Pages'.

Research data lifecycle



Text and data
PUBLISH

Prepare the data

Formats

- Open formats
- Check whether file conversion is possible if proprietary formats have to be used during the project

Meaningful names and structure

- Use vocabulary and terminology from the discipline

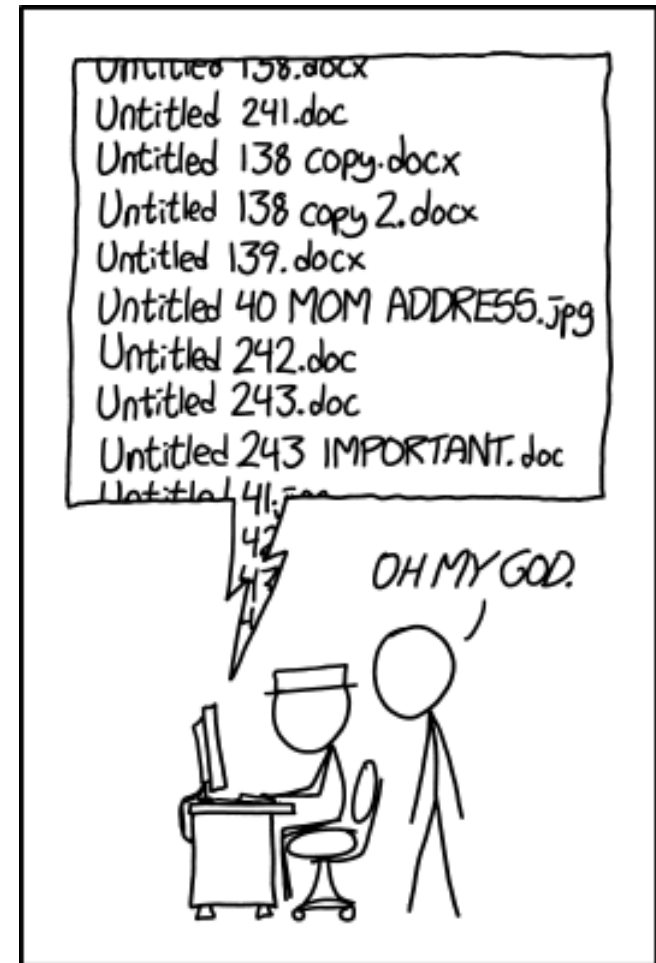
Formats

Use the most open formats possible to ensure future usability

Data type	Recommendation	To avoid
<i>Tables</i>	CSV, TSV, ODS, SQL, (XLST)	XLS, SPSS
<i>Text</i>	TXT, ODT, HTML, PDF/A, (RTF, DOCX, PTTX)	DOC, PPT
<i>Multimedia</i>	(MP4, AVI, MOV), WAV	Flash, WMV, MP3
<i>Images</i>	TIFF, PNG, JPEG2000, SVG	GIF, JPG, PSD
<i>Exchange</i>	XML, RDF, JSON	ODB, MDB, DBF
<i>Software</i>	Source code	Compiled code

Meaningful naming

- Descriptive names
- Consistent scheme
- Avoid automatic filenames
- Include dates for chronological sorting (YYYY-MM-DD)
- Avoid spaces and special characters
- Separate parts with underscores or *CamelCase*
- Keep filenames as short as possible, as long as necessary



PRO TIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.

Where to publish

Discipline-specific repositories

- Current listings at *re3data*

General-purpose repositories

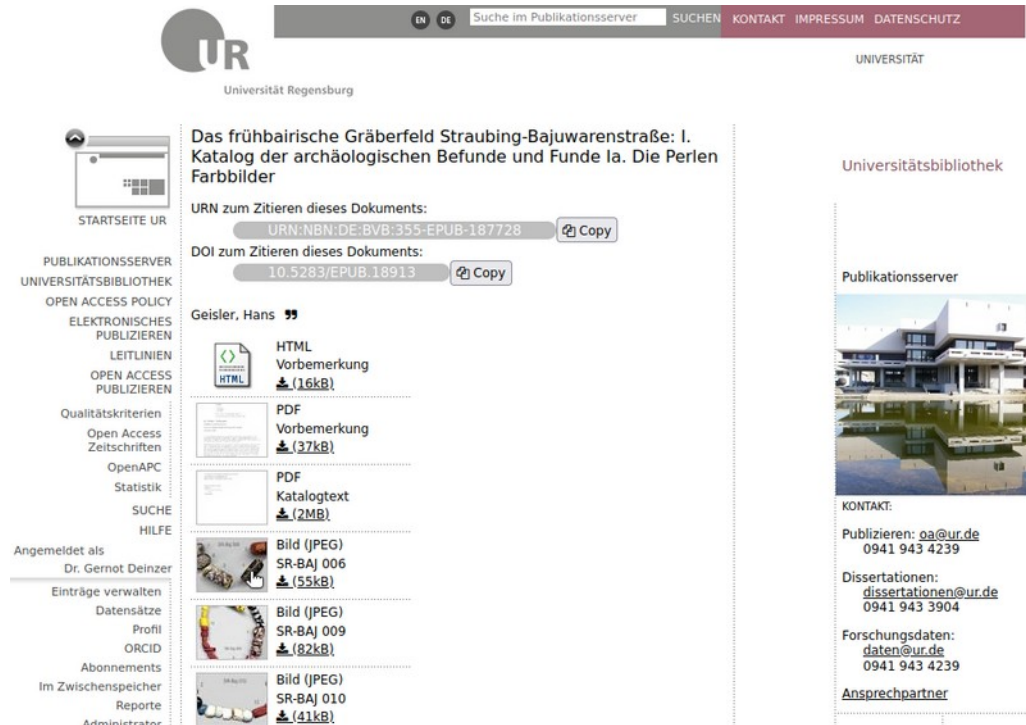
- e.g., *Zenodo*

Institutional repositories

- e.g., *UR Publication Server*

UR Publication Server

Provides all researchers at the University of Regensburg with a central platform for publishing and archiving data and documents.



The screenshot displays the UR Publication Server interface. At the top, there is a navigation bar with the UR logo, a search bar labeled 'Suche im Publikationsserver', and links for 'SUCHEN', 'KONTAKT', 'IMPRESSUM', and 'DATENSCHUTZ'. Below the navigation bar, the main content area is divided into several sections. On the left, a sidebar contains links to 'STARTSEITE UR', 'PUBLIKATIONSSERVER', 'UNIVERSITÄTSBIBLIOTHEK', 'OPEN ACCESS POLICY', 'ELEKTRONISCHES PUBLIZIEREN', 'LEITLINIEN', 'OPEN ACCESS PUBLIZIEREN', 'Qualitätskriterien', 'Open Access Zeitschriften', 'OpenAPC', 'Statistik', 'SUCHE', 'HILFE', 'Angemeldet als Dr. Gernot Deinzer', 'Einträge verwalten', 'Datensätze', 'Profil', 'ORCID', 'Abonnements', 'Im Zwischenspeicher', 'Reporte', and 'Administrator'. The main content area features a document titled 'Das frühbairische Gräberfeld Straubing-Bajuwarenstraße: I. Katalog der archäologischen Befunde und Funde Ia. Die Perlen Farbbilder'. Below the title, there are links for 'URN zum Zitieren dieses Dokuments: URN:NBN:DE:BVB:355-EPUB-187728' and 'DOI zum Zitieren dieses Dokuments: 10.5283/EPUB.18913', both with 'Copy' buttons. The document is by 'Geisler, Hans'. Below the author name, there are three download options: 'HTML Vorbemerkung (16kB)', 'PDF Vorbemerkung (37kB)', and 'PDF Katalogtext (2MB)'. At the bottom, there are three image thumbnails with captions: 'Bild (JPEG) SR-BAJ 006 (55kB)', 'Bild (JPEG) SR-BAJ 009 (82kB)', and 'Bild (JPEG) SR-BAJ 010 (41kB)'. On the right side of the interface, there is a section for 'Universitätsbibliothek' and 'Publikationsserver', which includes contact information: 'KONTAKT: Publizieren: pa@ur.de 0941 943 4239', 'Dissertationen: dissertationen@ur.de 0941 943 3904', 'Forschungsdaten: daten@ur.de 0941 943 4239', and 'Ansprechpartner'.

UR Publication Server

Persistent identifiers & metadata

- Assignment of DOIs/URNs
- Entry of metadata (per DataCite)

Access & usage rights

- Configure access permissions
- Upload/download individual files

Use & monitoring

- Download and usage statistics

Long-term preservation

- Secure storage

SORCK-Modell: Verhaltensanalyse als Ausgangspunkt für eine Förderplanung

URN zum Zitieren dieses Dokuments:

URN:NBN:DE:BVB:355-EPUB-533867

Copy

DOI zum Zitieren dieses Dokuments:

10.5283/EPUB.53386

Copy

Brodersen, Gunnar ; Grabowski, Frederike ; Castello, Armin



Lizenz: [Creative Commons Namensnennung-NichtKommerziell-Weitergabe unter gleichen Bedingungen 4.0 International](#)
PDF - Veröffentlichte Version
[\(100kB\)](#)

Veröffentlichungsdatum dieses Volltextes: 15 Dez 2022 07:49

Verknüpfung von Datensätzen

Details

Vorschau

Bibliographische Daten exportieren

Metadaten zuletzt geändert: 15 Dez 2022 07:49

Nur für Besitzer und Autoren: [Kontrollseite d](#)

Downloadstatistik

Weitere Literatur (mittels CORE)

„Catch-All” repository example: Zenodo

Free, interdisciplinary repository by the EU and CERN

The screenshot displays the Zenodo website interface. At the top, there is a blue header with the Zenodo logo, a search bar, and navigation links for 'Communities' and 'My dashboard'. On the right of the header are 'Log in' and 'Sign up' buttons. Below the header, the 'Featured communities' section highlights 'AAS Journals' with a 'Browse' button. The 'Recent uploads' section lists two items: a dataset titled 'Intensivkapazitäten und COVID-19-Intensivbettenbelegung in Deutschland' and a software item titled 'Escalated begging does not compromise nestling health'. A sidebar on the right titled 'Why use Zenodo?' lists benefits such as safety, trust, citability, no waiting time, open or closed access, versioning, GitHub integration, and usage statistics. At the bottom, there is a cookie consent banner.

zenodo Search records... Communities My dashboard Log in Sign up

Featured communities

AAS Journals
Collection of records for AAS Journal (ApJ, AJ, ApJL, ApJS, BAAS, PSJ, RNAAS) articles. Draft review and community submissions are welcome and will be curated by the AAS Data Editors. More information at <https://journals.aas.org/data-guide/>

Recent uploads

January 17, 2025 (2025-01-17) Dataset Open
Intensivkapazitäten und COVID-19-Intensivbettenbelegung in Deutschland
Robert Koch-Institut
Das Intensivregister (www.intensivregister.de) ist eine digitale Plattform zur Echtzeiterfassung von intensivmedizinischen Behandlungs- und Bettenkapazitäten sowie den Fallzahlen intensivmedizinisch behandelter COVID-19-Patient:innen der etwa 1.300 Akut-Krankenhäuser Deutschlands. Bis 12 Uhr ist die tägliche Meldung laut Verordnung für die...
Uploaded on January 17, 2025
Part of Robert Koch Institute
769 more versions exist for this record 52668 41828

January 17, 2025 (v1) Software Open
Escalated begging does not compromise nestling health
Parejo-Pulido, Daniel Redondo, Tomás; Casquero, Silvia

Why use Zenodo?

- **Safe** — your research is stored safely for the future in CERN's Data Centre for as long as CERN exists.
- **Trusted** — built and operated by CERN and OpenAIRE to ensure that everyone can join in Open Science.
- **Citeable** — every upload is assigned a Digital Object Identifier (DOI), to make them citable and trackable.
- **No waiting time** — Uploads are made available online as soon as you hit publish, and your DOI is registered within seconds.
- **Open or closed** — Share e.g. anonymized clinical trial data with only medical professionals via our restricted access mode.
- **Versioning** — Easily update your dataset with our versioning feature.
- **GitHub integration** — Easily preserve your GitHub repository in Zenodo.
- **Usage statistics** — All uploads display standards compliant usage statistics

This site uses cookies. Find out more on how we use cookies Accept all cookies Accept only essential cookies

„Catch-All” repository example: Zenodo

Persistent identifiers & metadata

- DOI assignment
- Dataset versioning

Integration

- GitHub integration

Uploads & hosting

- Up to 50 GB per dataset (up to 100 files)
- Support for large binary files

Long-term preservation

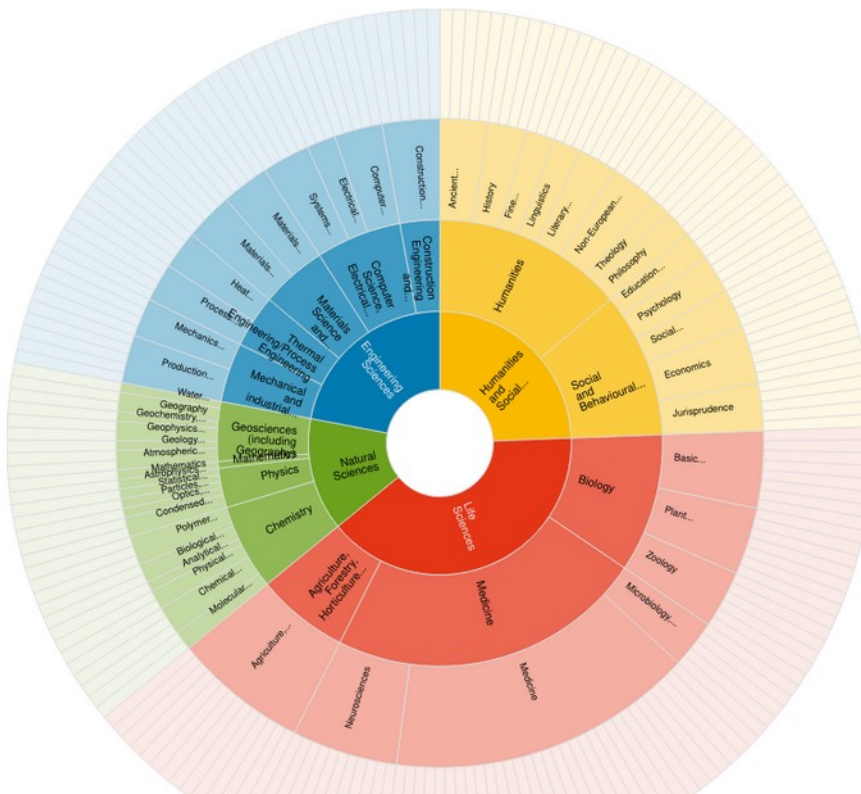
- Backed by CERN infrastructure

The screenshot shows a Zenodo dataset page. At the top, it says 'Published May 13, 2023 | Version v1' with 'Other' and 'Open' buttons. The title is 'HTR model for German manuscripts trained from several datasets' by 'Weil, Stefan'. There is a 'Show affiliations' button. A description states: 'This German handwriting recognition model has been trained with ground truth from different sources. It is based on a model for printed text. Details of the training process can be found here: <https://github.com/UB-Mannheim/kraken/wiki/Training-German-Handwriting>'. Under 'Files', there are two items: 'metadata.json' and 'Files (16.4 MB)'. The 'Citations' section shows '2K VIEWS' and '5K DOWNLOADS' with a 'Show more details' link. The 'Versions' section shows 'Version v1' published on 'May 13, 2023' with DOI '10.5281/zenodo.7933463'. It includes a 'Cite all versions?' section. The 'External resources' section shows 'Indexed in OpenAIRE'. The 'Communities' section shows 'OCR/HTR model repository'. The 'Keywords and subjects' section lists 'kraken_pytorch', 'HTR', 'Handwritten text recognition', and 'Manuskript'. The 'Details' section shows the DOI '10.5281/zenodo.7933463', 'Resource type: Other', and 'Publisher'.

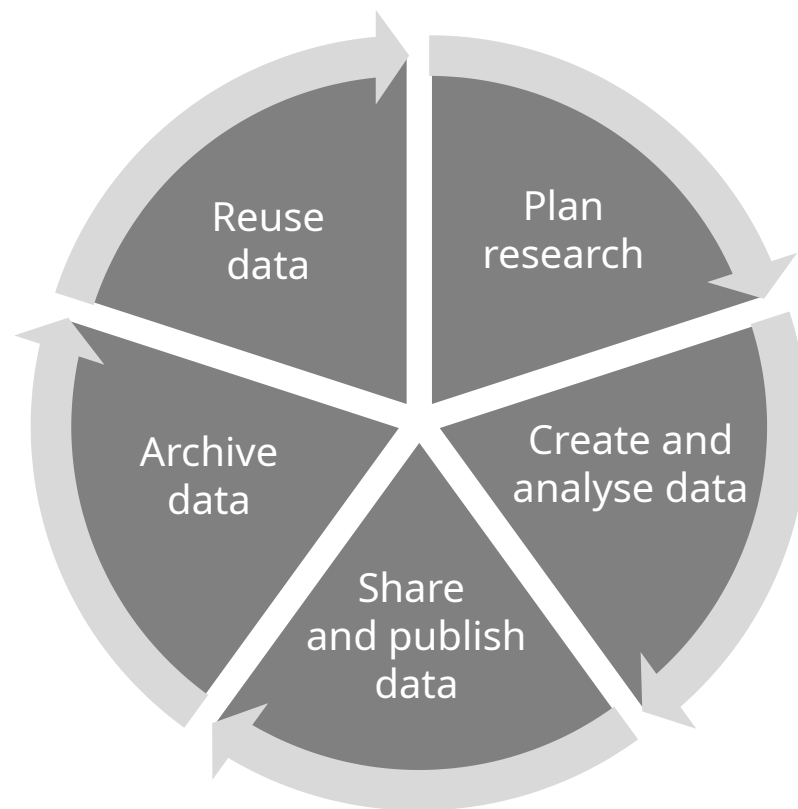
Searching for research data and repositories

Example re3data:

- Global registry with 2,500+ research data repositories across disciplines
- Search by discipline, institution, open-access policies, and more



Research data lifecycle



Research data **REUSE**

Licensing

- Creative Commons (CC) / for texts, images and data
- GNU General Public Licence (GPL) / for software
- Open Data Commons (ODC) / for databases
- Community Data Licence Agreement / for data

Lizenzmodell	Bearbeitungen verboten	Kommerzielle Nutzung verboten	Copyleft	Namensnennung	Urheberrechts- verzicht / Verzicht auf Rechtsdurch- setzung
Creative Commons	CC BY NC ND CC BY ND	CC BY NC SA CC BY NC	CC BY SA	CC BY	CC0
Open Data Commons			ODbL DbCL (für Datenbankinhalte)	ODC-By	PDDL
Community Data License Agreement			CDLA-Sharing	CDLA-Permissive (Beibehaltung v. Urheberrechtshin- weisen)	
restriktiv			offen		

Definition offene Lizenzen: Open Knowledge Foundation
([Version 2.1](#))

Licensing Options: UR Publication Server

Document upload

File

From URL

Browse...

Text

Plai

5B

Content:

★ Format:

Description:

★ Visible to:

License:

Embargo expiry date:

Year:

Month: Unspecified

Day: ?

★ Language:

English

Update Metadata

UNSPECIFIED

Publishing license for publications excluding print on demand

Allianz- bzw. Nationallizenz

Creative Commons: Attribution-No Derivative Works 3.0

Creative Commons: Attribution 3.0

Creative Commons: Attribution-Noncommercial 3.0

Creative Commons: Attribution-Noncommercial-No Derivative Works 3.0

Creative Commons: Attribution-Noncommercial-Share Alike 3.0

Creative Commons: Attribution-Share Alike 3.0

Creative Commons Attribution No Derivatives 4.0

Creative Commons Attribution 4.0

Creative Commons Attribution Non-commercial 4.0

Creative Commons Attribution Non-commercial No Derivatives 4.0

Creative Commons Attribution Non-commercial Share Alike 4.0

Creative Commons Attribution Share Alike 4.0

Creative Commons: Public Domain Dedication

Software: GNU GPL 2.0

Software: GNU LGPL 2.1

Evaluate your research data

Assess the quality of digital data via the FAIR check

F

A

I

R

0 of 26 questions completed

R

F

I

A

FAIRness
level

Reset questionnaire

Download PDF

Findable

The first step to reusing data by humans and computers is to be able to find it easily. Machine-readable rich metadata is essential for finding data automatically. Data and metadata need a globally unique and persistent identifier and are registered or indexed in a searchable resource.

Do your **data** (records) have a globally unique and permanently **persistent identifier** (PID) so that they can be cited? (E. g. a DOI or URN. A pure http URL address is not a PID).

☐ Yes

☐ No

Does your **metadata** have a globally unique and **persistent identifier** (e. g. a DOI, URN)?

Research data management at the UR

UR DATA HUB

Research data policy

Guidelines on Research Data Management

- Provide information on responsible and sustainable RDM (e.g., DMPs, data storage, etc.)
- Offer best practices
- Set the course for structural developments at UR (University of Regensburg)

Timeline



07/2021

Regulations on ensuring good scientific practice

09/2022

Guidelines on handling research data

Research data policy

Guidelines on Research Data Management

- Provide information on responsible and sustainable RDM (e.g., DMPs, data storage, etc.)
- Offer best practices
- Set the course for structural developments at UR (University of Regensburg)

**UR Data Hub as central unit for RDM
at the UR**

Timeline

07/2021

Regulations on ensuring good scientific practice

09/2022

Guidelines on handling research data

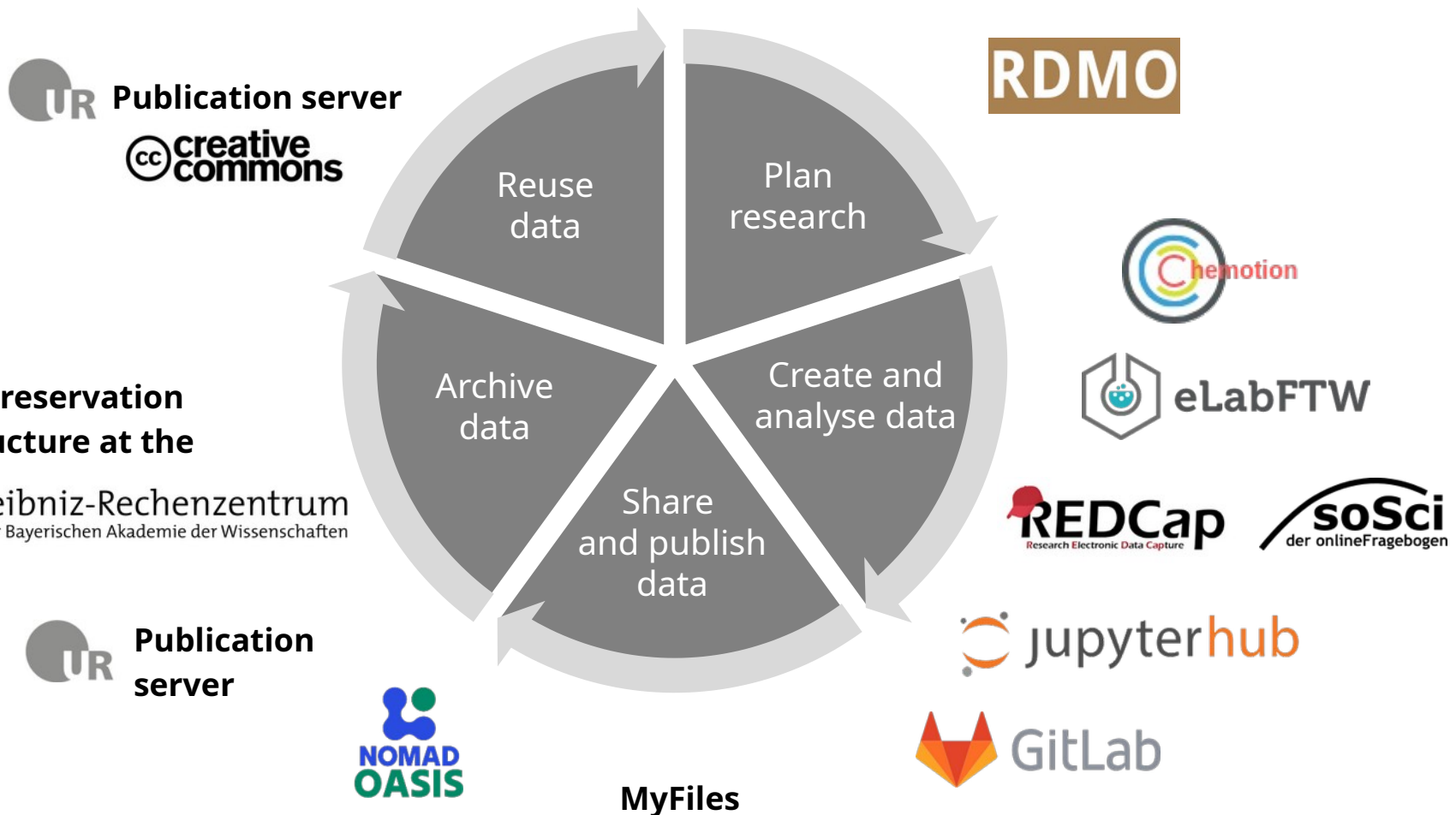
07/2023

Regulations concerning the UR Data Hub

Services of the UR Data Hub

- **Consulting on research data management needs:** before, during, and after the project
- Support in **creating Data Management Plans** (DMPs)
- Setup of **digital services** (e.g., electronic lab notebook) and other tailored virtual environments
- Information on **storage and long-term preservation** of research data
- **Training and information events** (updates via the UR Data Hub newsletter)

Central RDM services at the University of Regensburg



NFDI consortia

Goals: Manage scientific and research data, provide long-term data storage as well as accessibility to the data

Humanities and Social Sciences



Engineering Sciences



Life Sciences



Natural Sciences



Information and links

Tools and services

- RDMO: <https://rdmo.ur.de>
- Chemotion: <https://chemotion.ur.de>
- eLabFTW: <https://elab.ur.de>
- GitLab: <https://git.ur.de>
- JupyterHub: <https://jupyter.ur.de>
- MyFiles: <https://myfiles.ur.de>
- UR Publication Server: <https://epub.ur.de>

Long-term archiving

<https://lzv-bayern.de>

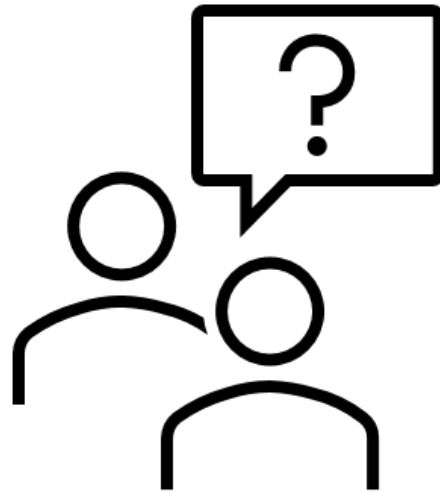
Newsletter

<https://www.listserv.dfn.de/sympa/info/ur-dh-newsletter>

Data Protection Office

<https://www.uni-regensburg.de/universitaet/organisation/beauftragte/datenschutzbeauftragte>

Contact



Dr. Gernot Deinzer
gernot.deinzer@ur.de

Constantin Lehenmeier
constantin.lehenmeier@ur.de

Dr. Sophie Stolzenberger
sophie.stolzenberger@ur.de

UR Data Hub
datahub@ur.de
<https://datahub.ur.de>