

Forschungsdatenmanagement

Hochschule für den öffentlichen Dienst in Bayern
Fachbereich Archiv- und Bibliothekswesen
Ausbildungsjahr 2021/2024

Dr. Yves Vincent Grossmann
grossmann@mpdl.mpg.de
Max Planck Digital Library

Willkommen

Dozent

- Yves Vincent Grossmann
- <https://orcid.org/0000-0002-2880-8947>
- Referent für Forschungsdatenmanagement seit Oktober 2020 an der Max Planck Digital Library
- Promotion über die Sozialgeschichte des (west-)deutschen Industriedesign von 1959-1990
- Kontakt: grossmann@mpdl.mpg.de

Max Planck Digital Library

- Amalienstraße 33, 80799 München
- Informationsservices seit 2007
- Seit 2012 eigenständige Serviceeinheit
- rund 80 Mitarbeiter_innen aus Software-Entwicklung, Bibliothek, Wissenschaftsmanagement und Verwaltung
- www.mpdl.mpg.de
- Abteilung Collections für das Thema Forschungsdaten



Das einzige Bücherregal in der MPDL

Max Planck Digital Library

- MPDL versteht sich als eine der vier zentralen IT-Service-Einrichtungen der Max-Planck-Gesellschaft (MPDL, MPCDF, IKT und GWDG)
- MPDL bildet gemeinsam mit den Institutsbibliotheken das Bibliothekswesen der Max-Planck-Gesellschaft (Subsidiaritätsprinzip!)



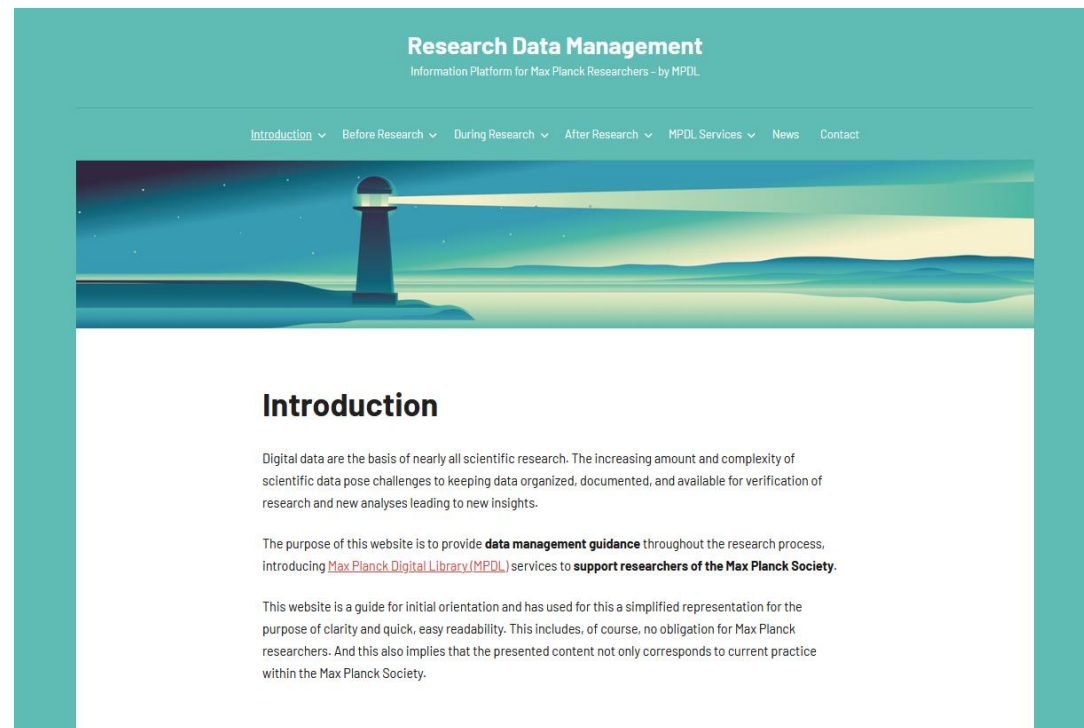
Blick auf das Gebäude mit der MPDL, Ecke Amalienstraße und Theresienstraße in München

MPDL RDM Information Platform

Abschnitte:

- Introduction
- Before
- During
- After
- MPDL Services
- Contact

Seit November 2020



<https://rdm.mpdل.mpg.de>

MPDL RDM Information Platform



RDM Support

Support in the area of research data management



Workshops

Recurring workshops on research data management



Labfolder

An ELN solution by scientists for scientists, licensed for the Max Planck Society



RDMO for MPG

A service organize the data management and the writing of data management plans



Keeper

Data sharing within a project, collaborative working and long-term archiving



Edmond

The open research data repository of the Max Planck Society



DOI-Service

Permanent identifiers for scientific output



Software Licensing Service

An institutionalized basic supply of software and online services

Ziele des Kurses

- Verständnis von Forschungsdaten
- Ziele und Anwendungen von Forschungsdatenmanagement
- aktuelle Entwicklung und Diskussionen

Struktur des Kurses

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Open Science und wissenschaftliche Software

Sitzungstermine

- 16. Dezember 2022
- 25. Januar 2023, **13:30 bis 17:00** Uhr

Erwartungen via Slido

Fragen

1. Was sind Ihre Wünsche?
2. Wie ist ihr FDM-Vorwissen?

Vorgehen

Im Browser www.slido.com öffnen. Code 5029111 eingeben.

Teil 1

Forschungsdaten

Inhalt von Teil 1

1. Beispiele von Forschungsdaten
2. Definitionsversuche
3. Klassifikationen
4. Forschungsdatenzyklen
5. Rechtliche Rahmen
6. Teilen und Probleme

Wofür Forschungsdatenmanagement?

Es ist sinnvoll, weil:

- Wiederverwendung von Daten, zumindest der eigenen, aber auch anderer
- Wissenschaftliches Ansehen
- die Akzeptanz von Daten als eigenständige Publikation wird in Zukunft wahrscheinlich zunehmen
- „passt irgendwie“ hat noch nie einen Nobelpreis gewonnen

Wofür Forschungsdatenmanagement?

Es ist eine Verpflichtung:

- Standard für „Gute wissenschaftliche Praxis“ von Institutionen
- Forschungsdatenpolicies von Institutionen
- Vorgaben von Förderinstitutionen
- Anforderung von Verlagen

The Big Picture*



* which only very schematically presents the handling of research data...

Beispiele für Forschungsdaten

Beispiele für Forschungsdaten

- Gravitationswellen
- LIGO (<https://ligo.org>)

LIGO LIGO Open Science Center
LIGO is operated by California Institute of Technology and Massachusetts Institute of Technology and supported by the U.S. National Science Foundation.

Getting Started
Tutorials
Data
Events
Bulk Data
Timelines
My Sources
Software
GPS → UTC
About LIGO
Data Analysis
Projects
Acknowledgement

Welcome to the LIGO Open Science Center

About LIGO
Get Started with LIGO data
Join the E-mail list for updates
For general information on LIGO, please visit ligo.org
If you have LSC credentials, you may go to the [development site](#)

Discoveries from the LIGO detectors!

released 2017 June 1:
Event of January 4, 2017: GW170104: total mass 50

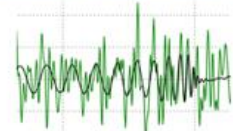
released 2016 June 15:
Event of December 26, 2015: GW151226: total mass 22

released 2016 June 15:
Candidate event of October 12, 2015: LVT151012: total mass 37

released 2016 Feb 11:
Event of September 14, 2015: GW150914: total mass 65

The LIGO Laboratory's Data Management Plan describes the scope and timing of LIGO data releases.

Jupyter notebook
See the new tutorial on signal processing with LIGO data, as a Jupyter (Python) notebook.
[Tutorial on Binary Black Hole Signals in LIGO Open Data](#)



Beispiele für Forschungsdaten

- Gensequenzen
- GenBank
(<https://www.ncbi.nlm.nih.gov/genbank/>)

The screenshot shows the NCBI GenBank entry for the Homo sapiens class II AP endonuclease (APE) gene, partial CDS. The entry includes the following information:

- LOCUS:** HUMAPEB 3019 bp DNA linear PRI 31-DEC-1994
- DEFINITION:** Homo sapiens class II AP endonuclease (APE) gene, partial CDS.
- ACCESSION:** M99703
- VERSION:** M99703.1 GI:178748
- KEYWORDS:** 3' diesterase associated activity; DNA repair protein; DNA-binding; class II AP endonuclease; nuclear localized protein.
- SOURCE:** Homo sapiens (human)
- ORGANISM:** Homo sapiens
- REFERENCE:** Harrison, L., Ascione, G., Menninger, J.C., Ward, D.C. and Demple, B. Human apurinic endonuclease gene (APE): structure and genomic mapping (chromosome 14q11.2-12) Hum. Mol. Genet. 1 (9), 677-680 (1992)
- PUBMED:** 1284593
- FEATURES:** Location/Qualifiers
 - source:** 1..3019 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
 - intron:** 596..778 /gene="APE" /number=1
 - gene:** join(779..904,1115..1302,1869..2061) /gene="APE"
 - mRNA:** join(779..904,1115..1302,1869..2061) /gene="APE"
 - exon:** 779..904 /gene="APE" /number=2
 - CDS:** join(847..904,1115..1302,1869..2061)

The right sidebar contains additional tools and information, including "Analyze this sequence" (Run BLAST, Pick Primers, Highlight Sequence Features, Find in this Sequence), "Articles about the APEX1 gene" (Expression and prognostic significance of APE1/Ref-1 and NPM1, Dysregulation of human apurinic/aprimidinic endonuclease 1 (APE1), PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATI), "Reference sequence information" (RefSeq alternative splicing), and "More about the APEX1 gene" (Apurinic/aprimidinic (AP) sites occur frequently in DNA molecules by spontaneous hydrolysis, by DNA damaging agents or by DNA glycosylases).

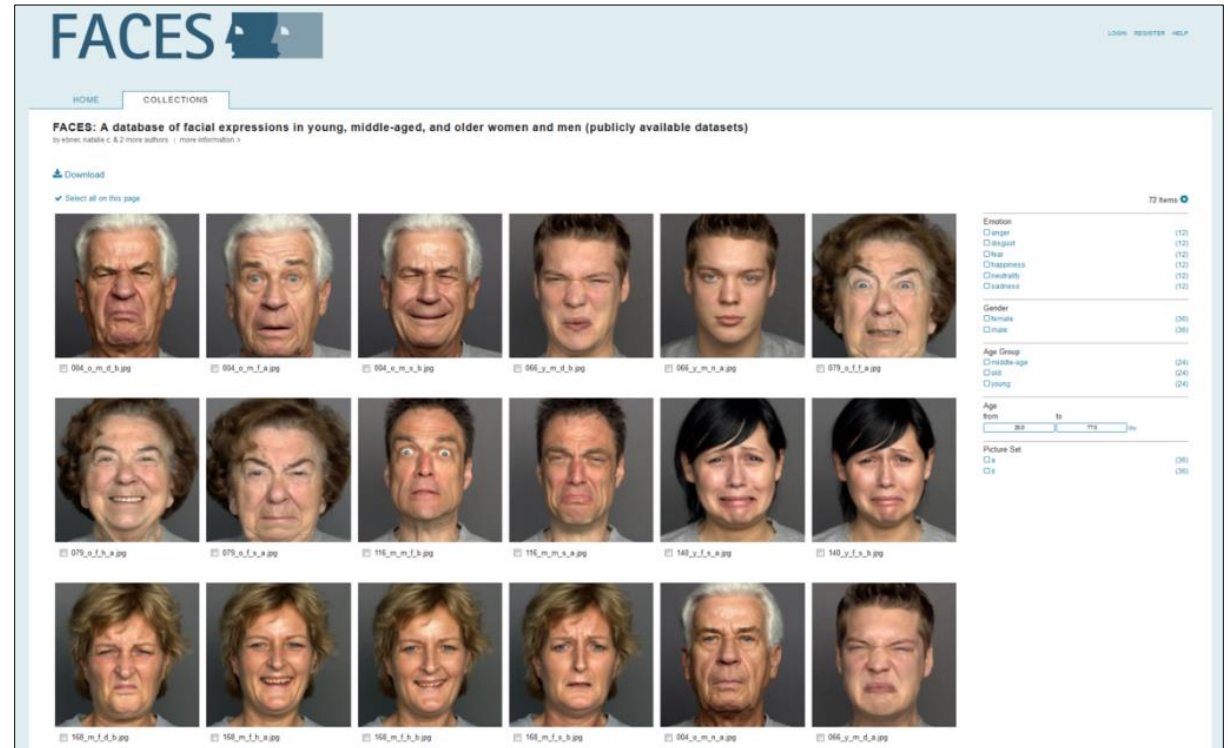
Beispiele für Forschungsdaten

- Tierbewegungen
- Movebank
(<https://www.movebank.org>)

The screenshot displays the Movebank website interface. The top navigation bar includes 'Home', 'Tracking Data Map', 'Community', 'Help', 'Tools', 'New Data', and 'Published Data'. A search bar is located in the top right corner. On the left side, there is a 'User login' section with fields for 'Username' and 'Password', and buttons for 'Log in', 'Create new account', and 'Request new password'. The main content area is divided into a search panel and a map. The search panel shows a search for 'stark' with filters for 'All Sensor Types' and 'Only studies where I can see data'. Below the search bar, a 'Search result' section lists various studies, including 'SBU, Dopes, SICM, 2013', 'Thick-billed Murres, Gichrist, Dopes Island, Canada', 'Turkey Vulture Antipar Center USA GPS', 'Turkey Vulture South Carolina USA', 'Turkey vultures in North and South America (data fr)', 'USPVS California Condor Recovery Program: ARD07', 'USGS Woodcock Migration', 'Vultures (Black & Turkey) Phm Texas', 'Wendelshof (W-13)', 'Amelia (Jm-84), Cicoma cicoma', 'Amelia 1 (Jm-217), Cicoma cicoma', 'Bruno (Jm-84), Cicoma cicoma', 'Dan (Jm-218), Cicoma cicoma', and 'Dorothea (Jm-85), Cicoma cicoma'. The map on the right shows a geographical view of Europe and the Mediterranean region, with several green location markers indicating data points. The footer of the page contains copyright information and a disclaimer.

Beispiele für Forschungsdaten

- Gesichter und Mimik
- **FACES**
(<https://faces.mpdl.mpg.de>)

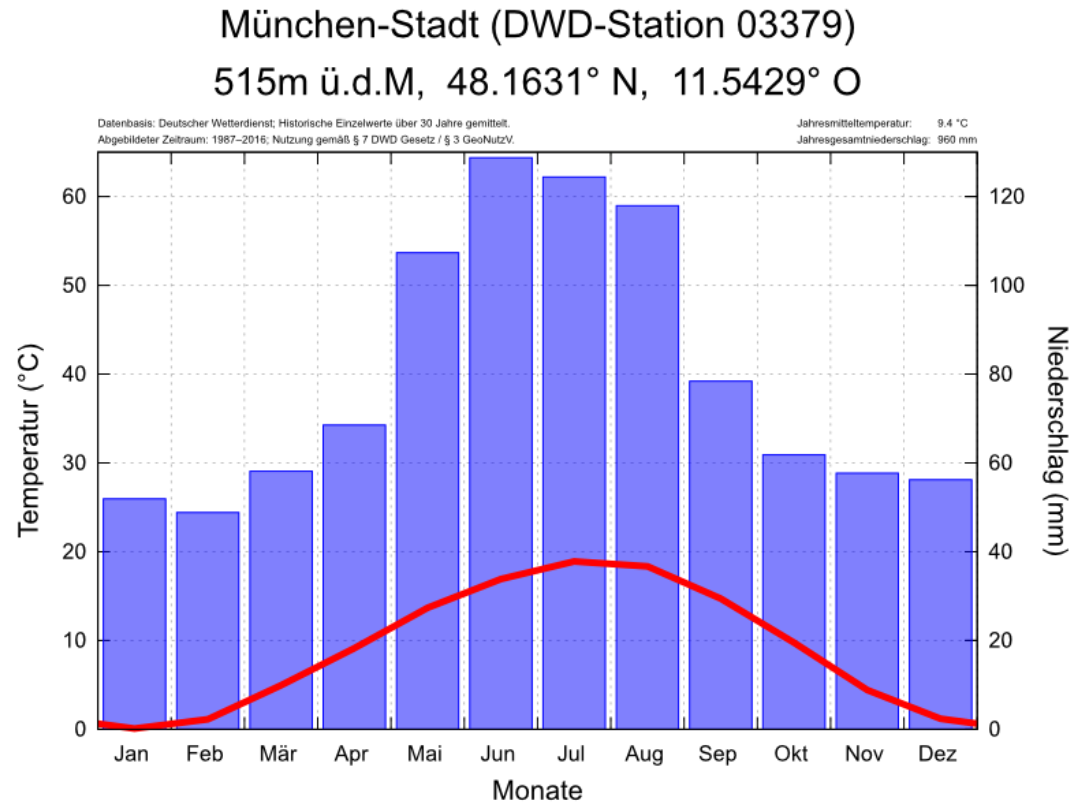


The screenshot displays the FACES database website. The main content area features a grid of 18 facial expression images arranged in three rows and six columns. Each image is accompanied by a small file icon and a filename (e.g., 004_x_m_e_b.jpg). The sidebar on the right contains a list of filters and their corresponding counts:

Filter	Count
Emotion	112
Danger	(12)
Disgust	(12)
Clear	(12)
Chastness	(12)
Cheerfulness	(12)
Disaddress	(12)
Gender	360
Older	(36)
Middle-aged	(36)
Young	(36)
Age Group	360
Older	(36)
Middle-aged	(36)
Young	(36)
Picture Set	360
A	(36)
B	(36)

Beispiele für Forschungsdaten

- Klimadaten
- DKRZ World Data Centre for Climate: <https://cera-www.dkrz.de/WDC/CC/ui/cera-search/>



Der-blaue-elefant, Klimadiagramm München-Stadt DWD 03379 Jahre 1987-2016, CC BY SA 4.0.

Beispiele für Forschungsdaten

- Daten für Psychologie
- PsychData
(www.psychdata.de)

PsychData

Research Data for Psychology

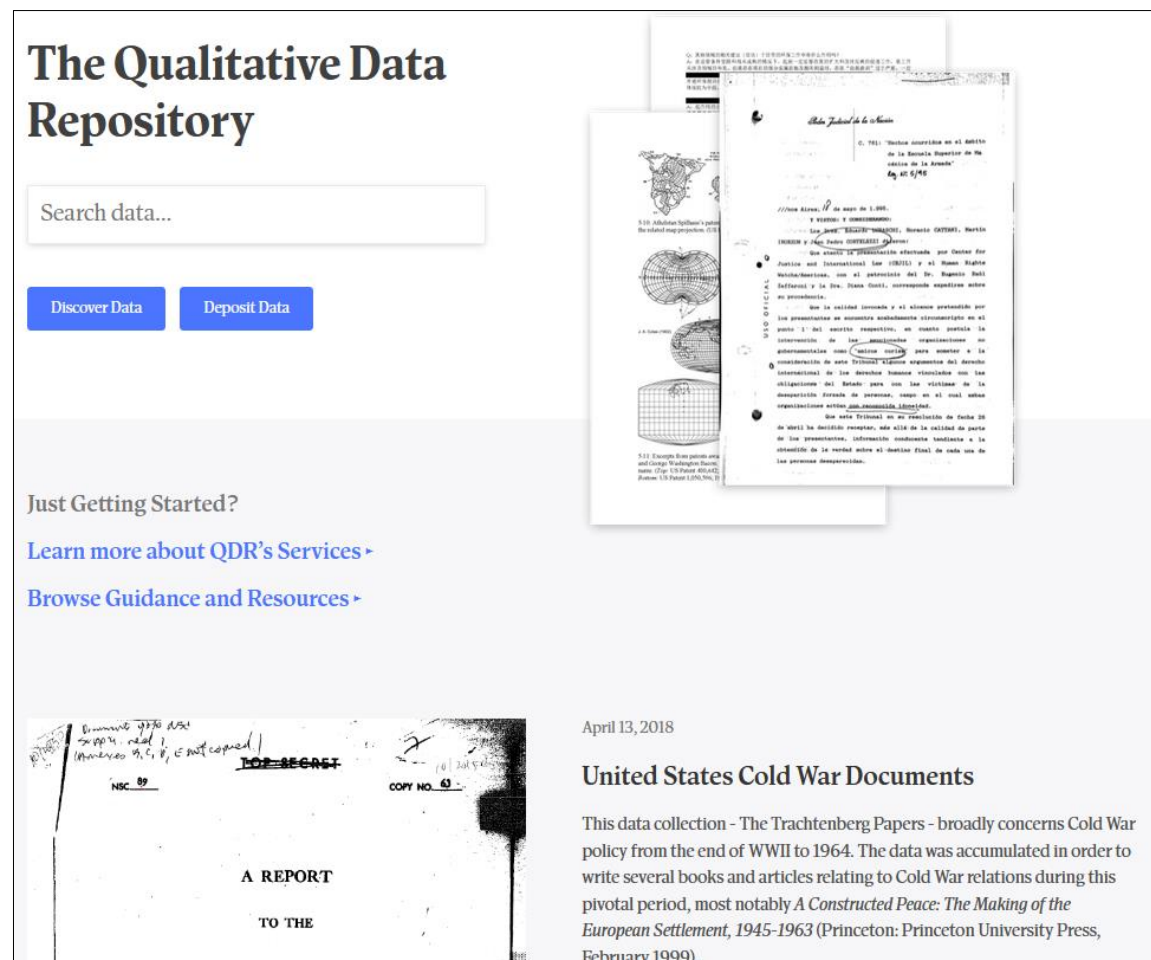
With PsychData, the Leibniz Institute for Psychology (ZPID) has developed a data-sharing platform specialized for psychology research.

PsychData helps researchers to

- provide their research data to the academic community, making their own research more present,
- [access research data](#), which has been released by others for academic use, and use it for their own research.

Beispiele für Forschungsdaten

- Qualitative Daten
- United States Cold War Documents
(<https://qdr.syr.edu>)



The screenshot shows the homepage of The Qualitative Data Repository (QDR). The header features the title "The Qualitative Data Repository" in a large, bold font. Below the title is a search bar with the placeholder text "Search data...". Two buttons, "Discover Data" and "Deposit Data", are positioned below the search bar. To the right of the search bar, there is a preview of a historical document with a map and text. Below the search bar, there is a section titled "Just Getting Started?" with two links: "Learn more about QDR's Services >" and "Browse Guidance and Resources >". At the bottom left, there is a preview of a document titled "A REPORT TO THE" with a "TOP SECRET" stamp and a "COPY NO. 62" stamp. The date "April 13, 2018" is displayed in the bottom right corner, along with the title "United States Cold War Documents".

Just Getting Started?
[Learn more about QDR's Services >](#)
[Browse Guidance and Resources >](#)

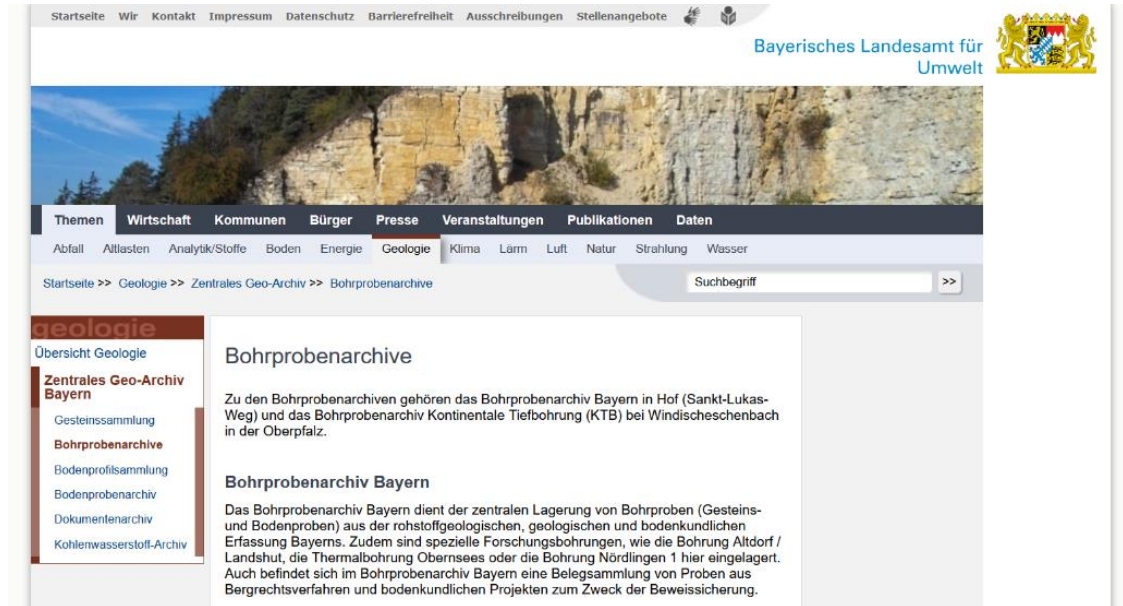
April 13, 2018

United States Cold War Documents

This data collection - The Trachtenberg Papers - broadly concerns Cold War policy from the end of WWII to 1964. The data was accumulated in order to write several books and articles relating to Cold War relations during this pivotal period, most notably *A Constructed Peace: The Making of the European Settlement, 1945-1963* (Princeton: Princeton University Press, February 1999).

Worum geht es nicht

- Physische Daten
- Wissenschaftliche Publikationen
- Daten über Wissenschaft
- z.B. [Bohrprobenarchive des Bayerischen Landesamts für Umwelt](#)



The screenshot shows the website of the Bayerisches Landesamt für Umwelt. The header includes navigation links like 'Startseite', 'Wir', 'Kontakt', 'Impressum', 'Datenschutz', 'Barrierefreiheit', 'Ausschreibungen', and 'Stellenangebote'. The main navigation menu has categories like 'Themen', 'Wirtschaft', 'Kommunen', 'Bürger', 'Presse', 'Veranstaltungen', 'Publikationen', and 'Daten'. The 'Geologie' sub-menu is active, showing options like 'Abfall', 'Altlasten', 'Analytik/Stoffe', 'Boden', 'Energie', 'Geologie', 'Klima', 'Lärm', 'Luft', 'Natur', 'Strahlung', and 'Wasser'. The breadcrumb trail reads 'Startseite >> Geologie >> Zentrales Geo-Archiv >> Bohrprobenarchive'. The main content area is titled 'Bohrprobenarchive' and contains text about the archive's location and purpose. A sidebar on the left lists 'Übersicht Geologie' and 'Zentrales Geo-Archiv Bayern' with sub-items like 'Gesteinssammlung', 'Bohrprobenarchive', 'Bodenprofilsammlung', 'Bodenprobenarchiv', 'Dokumentearchiv', and 'Kohlenwasserstoff-Archiv'.

Worum geht es nicht

- Physische Daten
- Wissenschaftliche Publikationen
- Daten über Wissenschaft



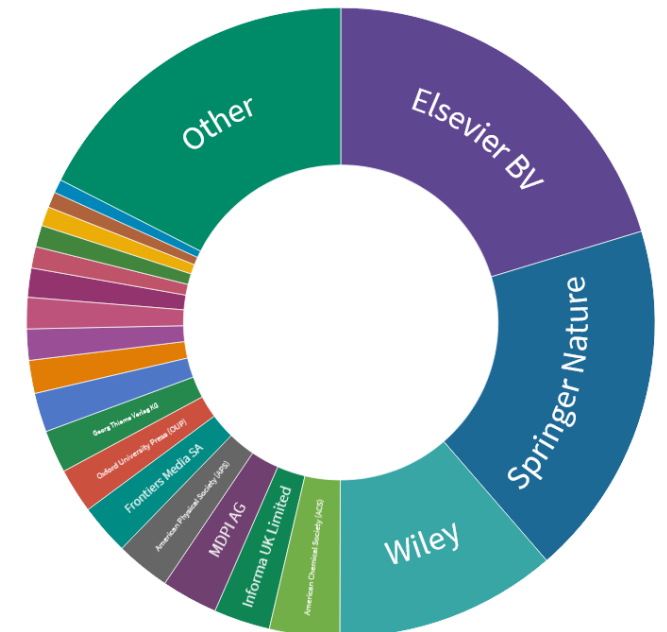
["Paper Weaving"](#) by FeatheredTar is licensed under CC BY 2.0

Worum geht es nicht

- Physische Daten
- Wissenschaftliche Publikationen
- Daten über Wissenschaft

Top 20 publishers based on corresponding author share

Filter:
Germany



<https://esac-initiative.org/market-watch/>

Definitionsversuche

Definitionsversuche: 1

„Ganz allgemein gesprochen sind Forschungsdaten Daten, die im Forschungsprozess erzeugt, gesammelt oder zusammengestellt werden und auf deren Grundlage wissenschaftliche Hypothesen, Modelle oder Theorien gebildet werden.“

Kompetenzzentrum Forschungsdaten der Universität Heidelberg, <https://data.uni-heidelberg.de/faq.html>

Definitionsversuche: 2

„Zu Forschungsdaten zählen u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten oder Beobachtungsdaten, methodische Testverfahren sowie Fragebögen.“

Fragenkatalog der DFG, Stand 10. Juni 2021, S. 1,

https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_de.pdf

Definitionsversuche: 3

Forschungsdaten sind „*Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.*“

Allianz AG „Forschungsdaten“, Forschungsdatenmanagement – Eine Handreichung, 2018, S. 4, https://gfzpublic.gfz-potsdam.de/rest/items/item_3055893_5/component/file_3055894/content

Definitionsversuche: 4

Research data refer “*to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.*”

European Commission, H2020 Programme – Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, v.3.2, 2017, S. 4,

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Definitionsversuche: 5

„Forschungsdaten umfassen alle Daten, die im wissenschaftlichen Arbeitsprozess entstehen und in digitaler Form verarbeitet werden.“

Leitlinie zum Umgang mit Forschungsdaten in der Leibniz-Gemeinschaft, 2018, S. 1, https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/Forschung/Open_Science/Leitlinie_Forschungsdaten_2018.pdf

Gruppenarbeit

Frage:

Mit welchen Forschungsdaten haben Sie bereits gearbeitet oder sind damit in Kontakt gekommen?

Beispiele:

Arbeit in einer Bibliothek, Besichtigungen in anderen Institutionen, Alltag etc.

Methode:

Vier-Augen-Gespräch

Dauer:

3 Minuten

Klassifikation von Forschungsdaten

Klassifikationen: Nach dem Charakter der Daten

Qualitative Daten

- Texte (Quellen, Transkripte)
- Bilder (Scans, Fotos, Mikroskopdaten, Teleskopdaten, Satellitendaten)
- Multimedia (Audiodaten, Videodaten, 3D-Daten, 4D-Daten)

Quantitative Daten

- Zahlenreihen (Umfragedaten, Experimentaldaten, Sensormessreihen, Gensequenzen, Geodaten)
- Born-digitals (Simulationsdaten, Algorithmen, Websites)

Klassifikationen: Nach dem Charakter der Daten

Text (TXT, ODT, PDF, ...)

Glossed text

See p. 12. 13 in:

van der Berg, Margot C. & Smith, Norval S.H. 2014. Early Shona. In: Mchale, Suzanne Maria & Mairor, Philippe & Huzarand, Martin & Huzar, Magana (eds.) *The Survey of African and Creole Languages. Vol. 5: English-based and Dutch-based languages*, 3-11. Oxford: Oxford University Press.

The following lines, taken from the play in van Dyk's (c1765) language manual, present a detailed account of a plantation funeral that is generally considered to be true to life (van Trier-Guicherit 1991).

Zangwe Officer: Tjippe kice nyhi ut go unaira negere
close coffin make 1PL go master black
Black Overseer: 'Close the coffin, let's go, Gentlemen.

omoo wibked unonoi epe di bwanen
2PL six person lift DE:SG woman
you six persons, lift up the woman.

ba ply deu banan kango negere alle
Q place DET.PL woman with black all
Where is the woman with all the blacks?

omoo wakke na bippi kry-man nese
2PL walk LOC heap wail man must
You must walk as a group, the wailers must.

wabbe na fesi omoo no rengi man
walk LOC face 2PL NEG forget one
walk at the front. Don't forget a

sauti kango-kango
thing come-go
thing! Come, come!

Zangwe: Di so ut jurei dible unonoi go ni jure di,
that so 1PL carry dead person go 1SG play
Singers: 'That's how we carry the dead away. I celebrate, We

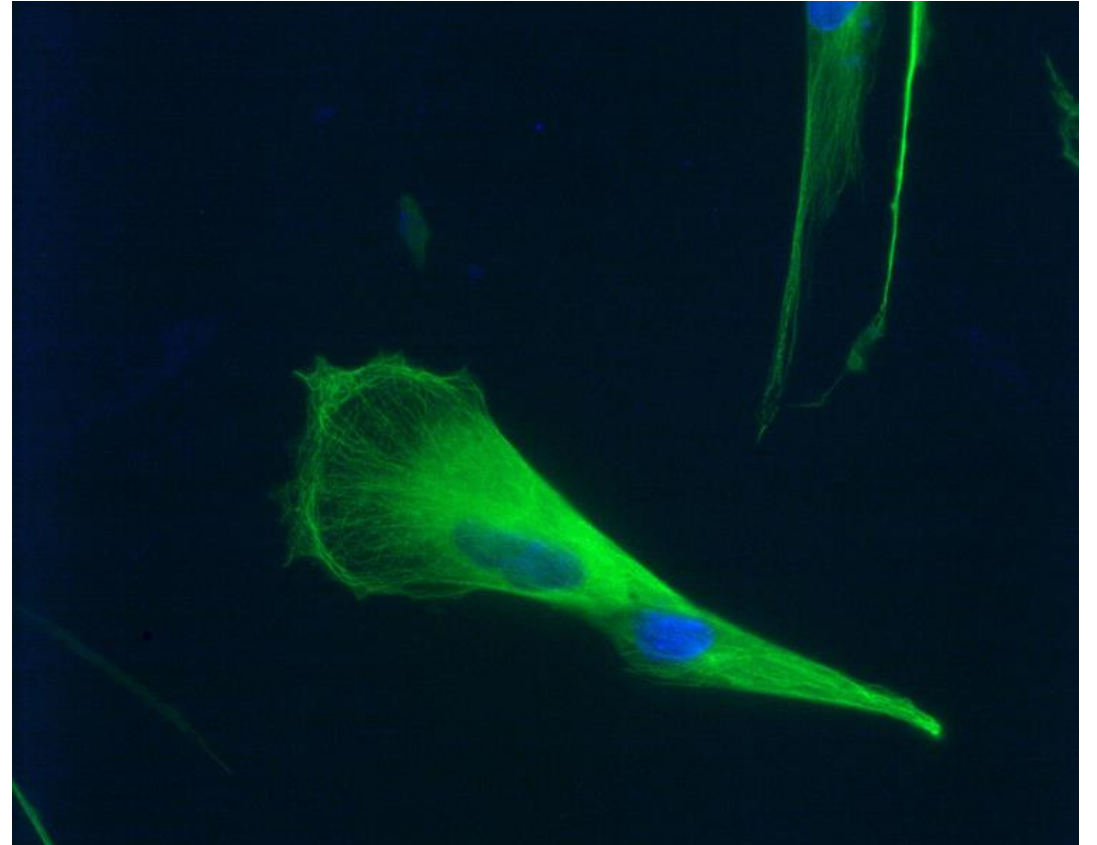
ut zurei fu ju alle di ju go laet en na
1PL regret for 2PL all that 2SG go expire so LOC
are very sorry for you. It is you who died like this.

APiCS Consortium (2014): APiCS supplemental files, CC BY 4.0, <https://doi.org/10.17617/3.LTLK50>.

Klassifikationen: Nach dem Charakter der Daten

Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)



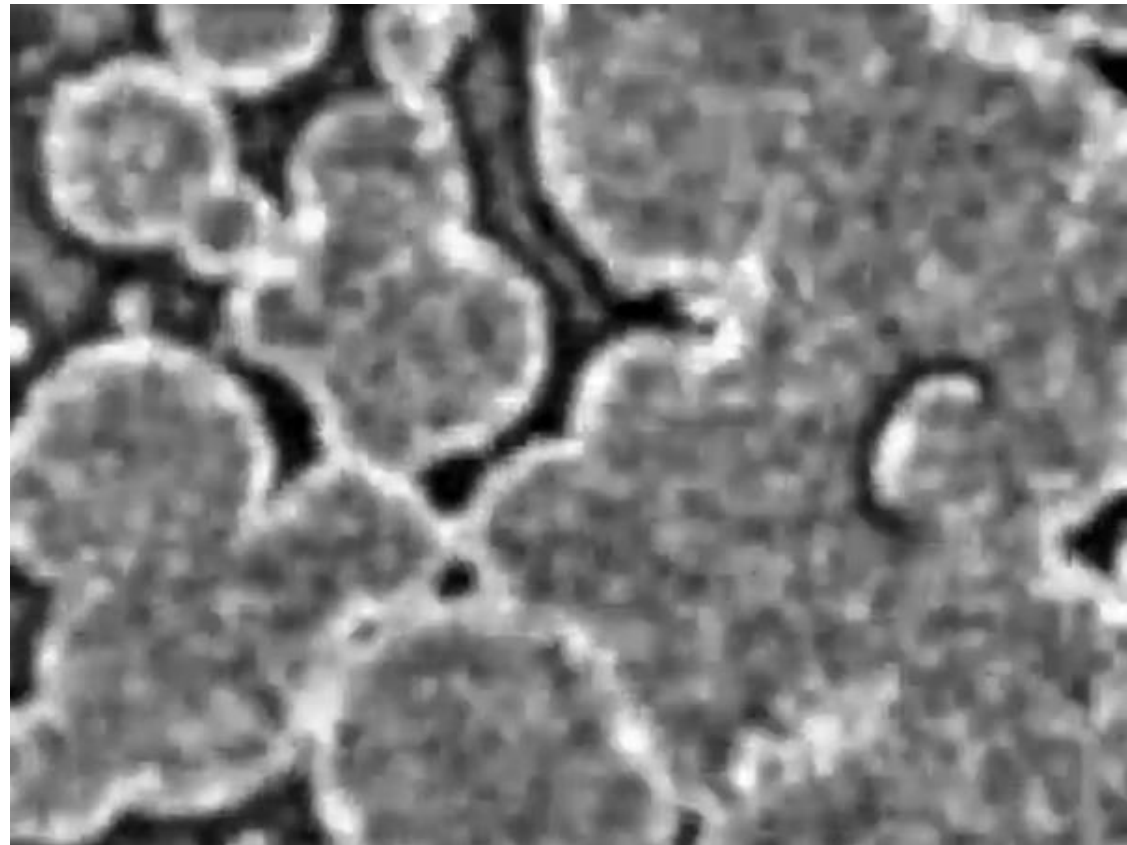
Ries, Albert (Max Planck Institute for Biochemistry), IMR90 fibroblasts, 2014,
CC BY 4.0, <https://doi.org/10.17617/3.ND7FXM>.

Klassifikationen: Nach dem Charakter der Daten

Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)



Kaumudi Prabhakara (Max Planck Institute for Dynamics and Self-Organization),
Spiralwaves, 2015, CC BY 4.0, <https://www.youtube.com/watch?v=DmRZn073Uus>.

Klassifikationen: Nach dem Charakter der Daten

Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Zahlenreihen (CSV, XLSX, ...)

```
HCCT-2010
Schmücke station
MAAP @ interstitial inlet
MPIC Mainz
Contact: Johannes Schneider, johannes.schneider@mpic.de

Data status: Final (1.3.2011)
(unreliable data points have been removed)

Date and Time
Black_Carbon_ng/m3

11.09.2010 18:30:05      144
11.09.2010 18:31:05      117
11.09.2010 18:32:05      157
11.09.2010 18:33:05      165
11.09.2010 18:34:05      198
11.09.2010 18:35:05      226
11.09.2010 18:36:05      230
11.09.2010 18:37:05      216
11.09.2010 18:38:05      232
11.09.2010 18:39:05      232
11.09.2010 18:40:05      249
11.09.2010 18:41:05      241
11.09.2010 18:42:05      449
11.09.2010 18:43:05      329
11.09.2010 18:44:05      252
11.09.2010 18:45:05      215
11.09.2010 18:46:05      119
11.09.2010 18:47:05      173
11.09.2010 18:48:05      215
11.09.2010 18:49:05      235
11.09.2010 18:50:05      256
11.09.2010 18:51:05      230
11.09.2010 18:52:05      252
11.09.2010 18:53:05      238
11.09.2010 18:54:05      244
11.09.2010 18:55:05      215
11.09.2010 18:56:05      196
11.09.2010 18:57:05      162
11.09.2010 18:58:05      188
11.09.2010 18:59:05      216
11.09.2010 19:00:05      197
11.09.2010 19:01:05      213
11.09.2010 19:02:05      243
```

Schneider, Johannes (2017): HCCT2010, CC BY 4.0,
<https://dx.doi.org/10.17617/3.i>

Klassifikationen: Nach dem Charakter der Daten

Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Zahlenreihen (CSV, XLSX, ...)

Born-digitals (netCDF, grib, HTML, ...)

```
netcdf tos_O1_2001-2002 {
dimensions:
  lon = 180 ;
  lat = 170 ;
  time = UNLIMITED ; // (24 currently)
  bnds = 2 ;
variables:
double lon(lon) ;
  lon:standard_name = "longitude" ;
  lon:long_name = "longitude" ;
  lon:units = "degrees_east" ;
  lon:axis = "X" ;
  lon:bounds = "lon_bnds" ;
  lon:original_units = "degrees_east" ;
double lon_bnds(lon, bnds) ;
double lat(lat) ;
  lat:standard_name = "latitude" ;
  lat:long_name = "latitude" ;
  lat:units = "degrees_north" ;
  lat:axis = "Y" ;
  lat:bounds = "lat_bnds" ;
  lat:original_units = "degrees_north" ;
double lat_bnds(lat, bnds) ;
double time(time) ;
  time:standard_name = "time" ;
  time:long_name = "time" ;
  time:units = "days since 2001-1-1" ;
  time:axis = "T" ;
  time:calendar = "360_day" ;
  time:bounds = "time_bnds" ;
  time:original_units = "seconds since 2001-1-1" ;
double time_bnds(time, bnds) ;
float tos(time, lat, lon) ;
  tos:standard_name = "sea_surface_temperature" ;
  tos:long_name = "Sea Surface Temperature" ;
  tos:units = "K" ;
  tos:cell_methods = "time: mean (interval: 30 minutes)" ;
  tos:fill_value = 1.e+20F ;
  tos:missing_value = 1.e+20F ;
  tos:original_name = "sstsst" ;
  tos:original_units = "degC" ;
  tos:history = " At 16:37:23 on 01/11/2005: CMOR altered the data in the following ways: added 2.73150E+02 to yield output units: Cyclical dimension was output
at a different lon" ;
// global attributes:
  :title = "IPSL model output prepared for IPCC Fourth Assessment SRES A2 experiment" ;
  :institution = "IPSL (Institut Pierre Simon Laplace, Paris, France)" ;
  :source = "IPSL-CM4_v1 (2003) : atmosphere : IADC (IPSL-CM4_IPCC, 96x71x19) ; ocean ORCA2 (ipsl_cm4_v1_8, 2x2131) ; sea ice LIM (ipsl_cm4_v" ;
  :contact = "sebastien.denvil, sebastien.denvil@ipsl.jussieu.fr" ;
  :project_id = "IPCC Fourth Assessment" ;
```

Sea surface temperatures collected by PCMDI for use by the [IPCC](https://www.unidata.ucar.edu/software/netcdf/examples/files.html), 2015, <https://www.unidata.ucar.edu/software/netcdf/examples/files.html>

Klassifikation: Nach der Quelle der Daten bzw. der Art der Forschung

Observational [Beobachtung]: data captured in real-time, usually irreplaceable. For example, sensor data, survey data, sample data, neuroimages.

Experimental [Experiment]: Data from lab equipment, often reproducible, but can be expensive. For example, gene sequences, chromatograms, toroid magnetic field data.

Simulation [Simulation]: data generated from test models where model and metadata are more important than output data. For example, climate models, economic models.

Derived or compiled [Ableitung]: data is reproducible but expensive. For example, text and data mining, compiled database, 3D models.

Reference or canonical [Sammlung]: a (static or organic) conglomeration or collection of smaller (peer-reviewed) datasets, most probably published and curated. For example, gene sequence databanks, chemical structures, or spatial data portals.

Klassifikation: Nach dem Grad der Aggregation bzw. Verdichtung

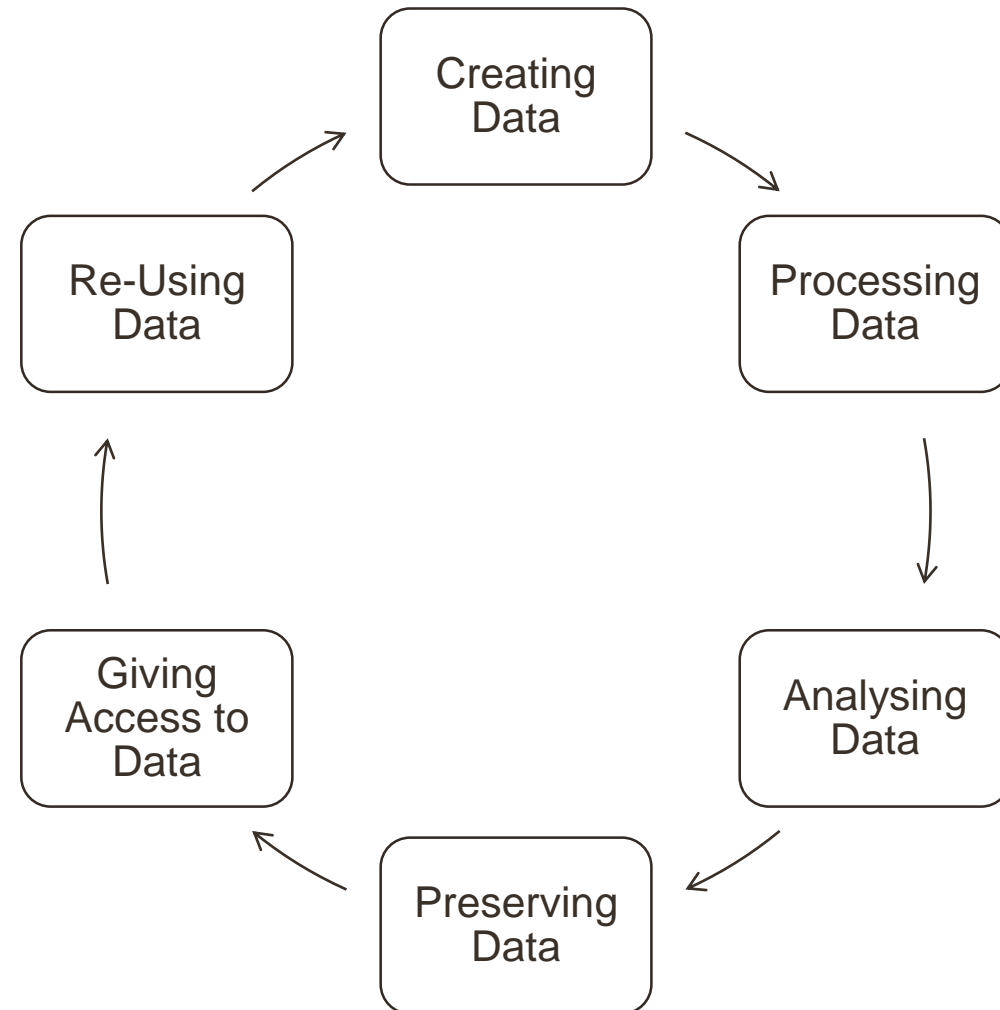
Rohdaten: Als Rohdaten soll die Gesamtmenge aller Daten bezeichnet werden, die empirischer Wissenschaft als Forschungsgrundlage zur Verfügung stehen.

Primärdaten: Primärdaten seien diejenigen Daten, die als Teilmenge der Rohdaten tatsächlich zur Forschung herangezogen werden.

Sekundärdaten: Sekundärdaten seien Daten, die in Prozessschritten aus Primärdaten gewonnen wurden.

Zyklen von Forschungsdaten

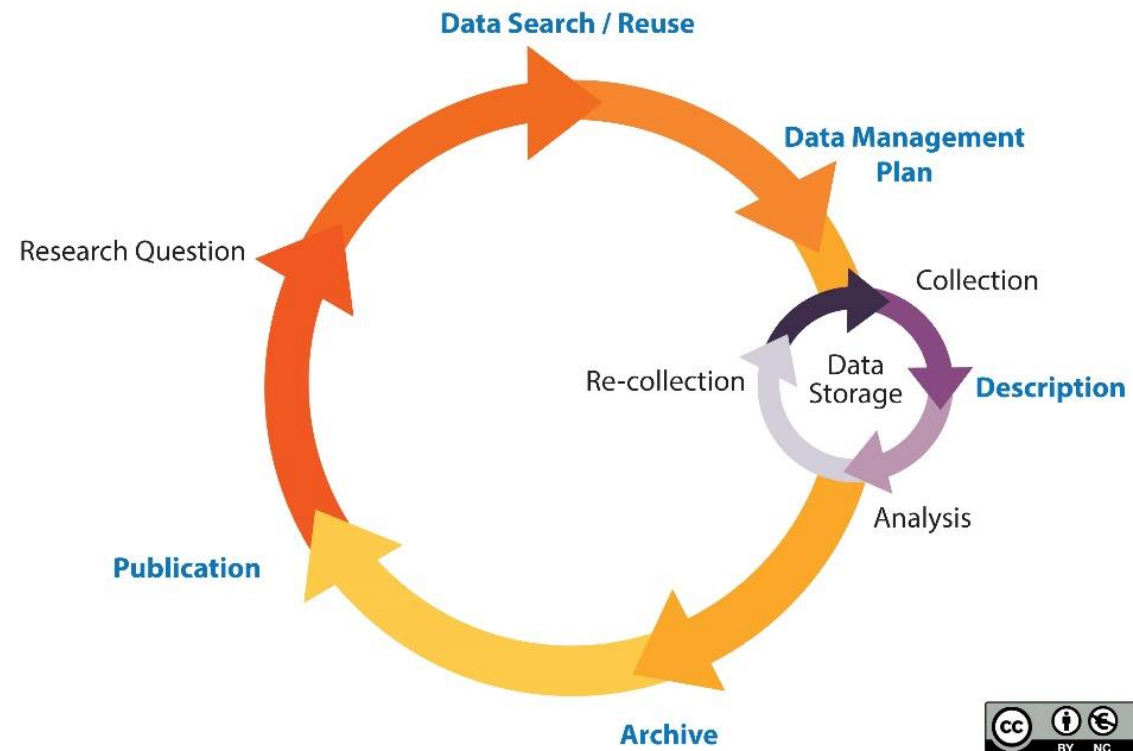
Forschungsdatenzyklus



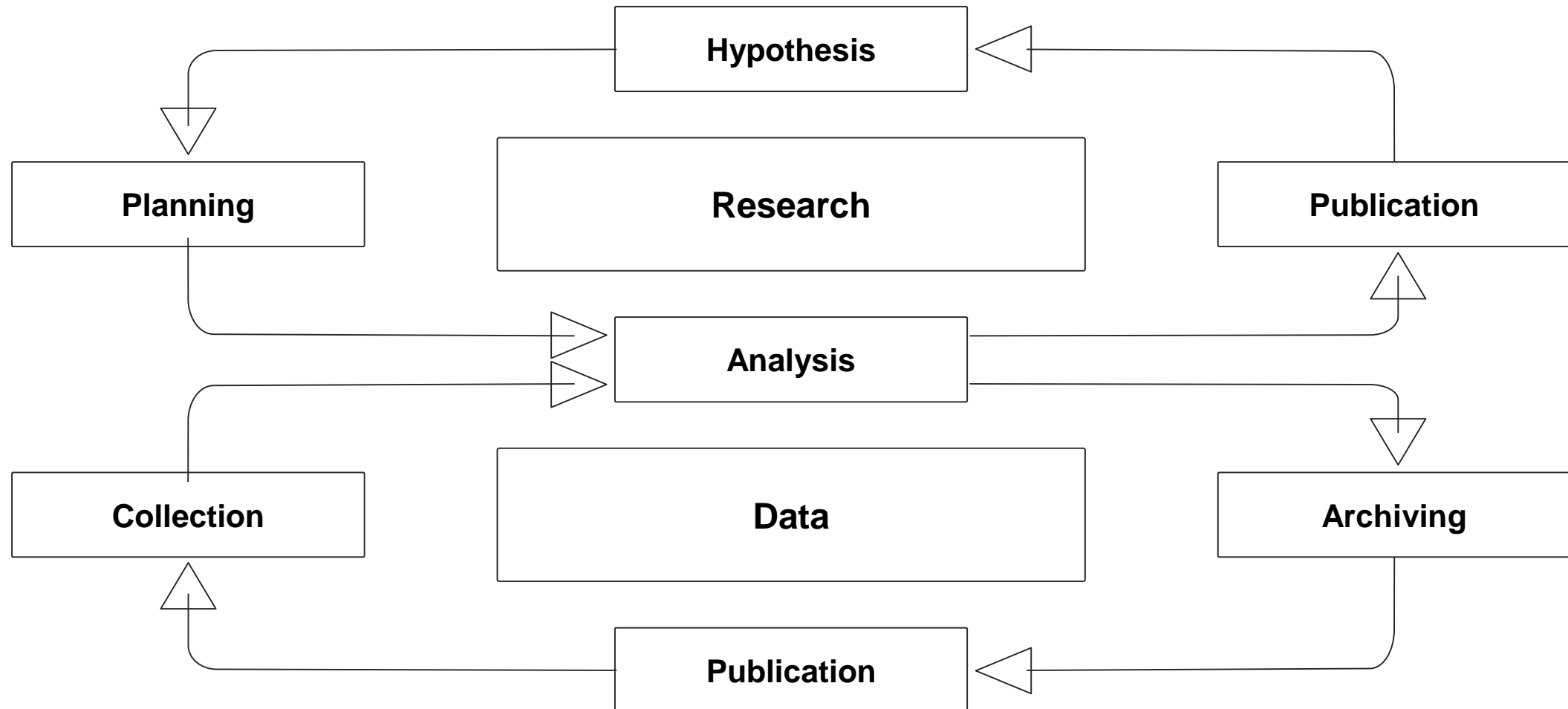
UK Data Service,
<https://www.ukdataservice.ac.uk/manage-data/lifecycle>

Forschungs-Datenzyklus

The Research Data Management Lifecycle



Forschungs- und Datenzyklus



Rechtliche Rahmen

Rechtliche Vorgaben

- Standard für „Gute wissenschaftliche Praxis“ von Institutionen
- Forschungsdatenpolicies von Institutionen
- Vorgaben von Förderinstitutionen
- Anforderung von Verlagen
- ...

Max-Planck-Gesellschaft: Gute Wissenschaftliche Praxis

„Vorhalten von Forschungsdaten

Die Dauer der Aufbewahrungsfristen ergibt sich aus rechtlichen Anforderungen, organisationsinternen Regelungen und Rechenschaftspflichten. In der Regel beträgt diese Frist 10 Jahre.“

(<https://www.mpg.de/197494/rulesScientificPractice.pdf>, S. 13)

DFG's Code of Conduct

“Safeguarding Good Research Practice”

Guideline 13: Providing public access to research results

“[...] If it has been decided to make results available in the public domain, **researchers describe them clearly and in full**. Where possible and reasonable, this includes making the research data, materials and information on which the results are based, as well as the methods and software used, available and fully explaining the work processes. [...]“
(<http://doi.org/10.5281/zenodo.6472827>, p. 17)

Guideline 17: Archiving

“When scientific and academic findings are made publicly available, the research data (generally raw data) on which they are based are generally archived in **an accessible and identifiable manner for a period of ten years at the institution** where the data were produced or in cross-location repositories. [...]“ (<http://doi.org/10.5281/zenodo.6472827>, p. 20)

European Commission

ERC Work Programme 2021:

”Finally, as from 2021 it is no longer possible for applicants to opt out of the submission of Research Data Management plans.”
(European Commission Decision C(2021) 930 of 22/02/2021, p. 4)

Marie Skłodowska-Curie Actions:

“[...] data management plan submitted at mid-term and an update towards the end of the project if needed” (European Commission Decision C(2021) 930 of 22/02/2021, p. 82)

Horizon Europe

“The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- **establish a data management plan** ('DMP') (and regularly update it)
- as soon as possible and within the deadlines set out in the DMP, **deposit the data** in a trusted repository [...].
- as **soon as possible** and within the deadlines set out in the DMP, ensure **open access** – via the repository – to the deposited data [plus CC0, CC BY or equivalent].“

Horizon Europe and Euratom: General Model Grant Agreement, Version 1.0, 1st June 2021, p. 109,
https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-auratom_en.pdf

Teilen und Probleme

Teilen von Forschungsdaten

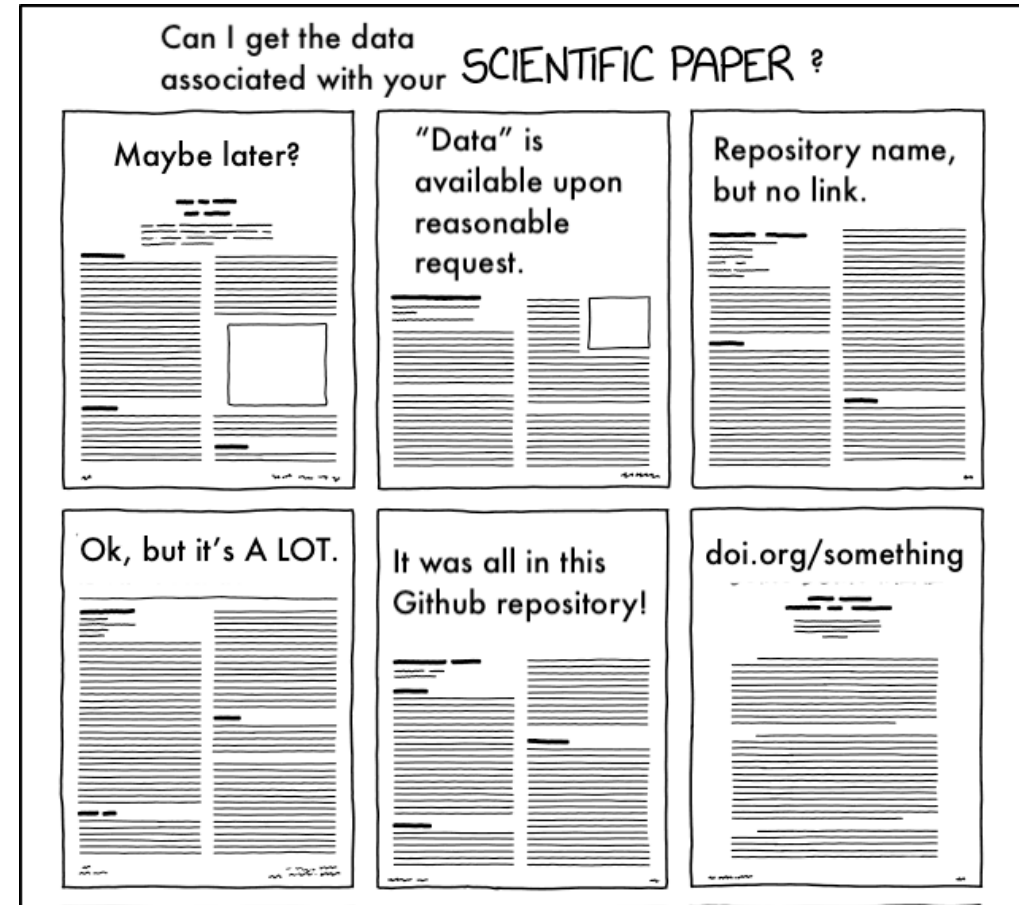
- Community-Standards
- erwünscht von Institutionen
- erwünscht von Geldgebern
- größere Sichtbarkeit der eigenen Forschung
- Datenpublikation als eignes Genre
- ...

Probleme bei Zugängen

Idealer sollten die Forschungsdaten für eine Veröffentlichung frei zugänglich sein

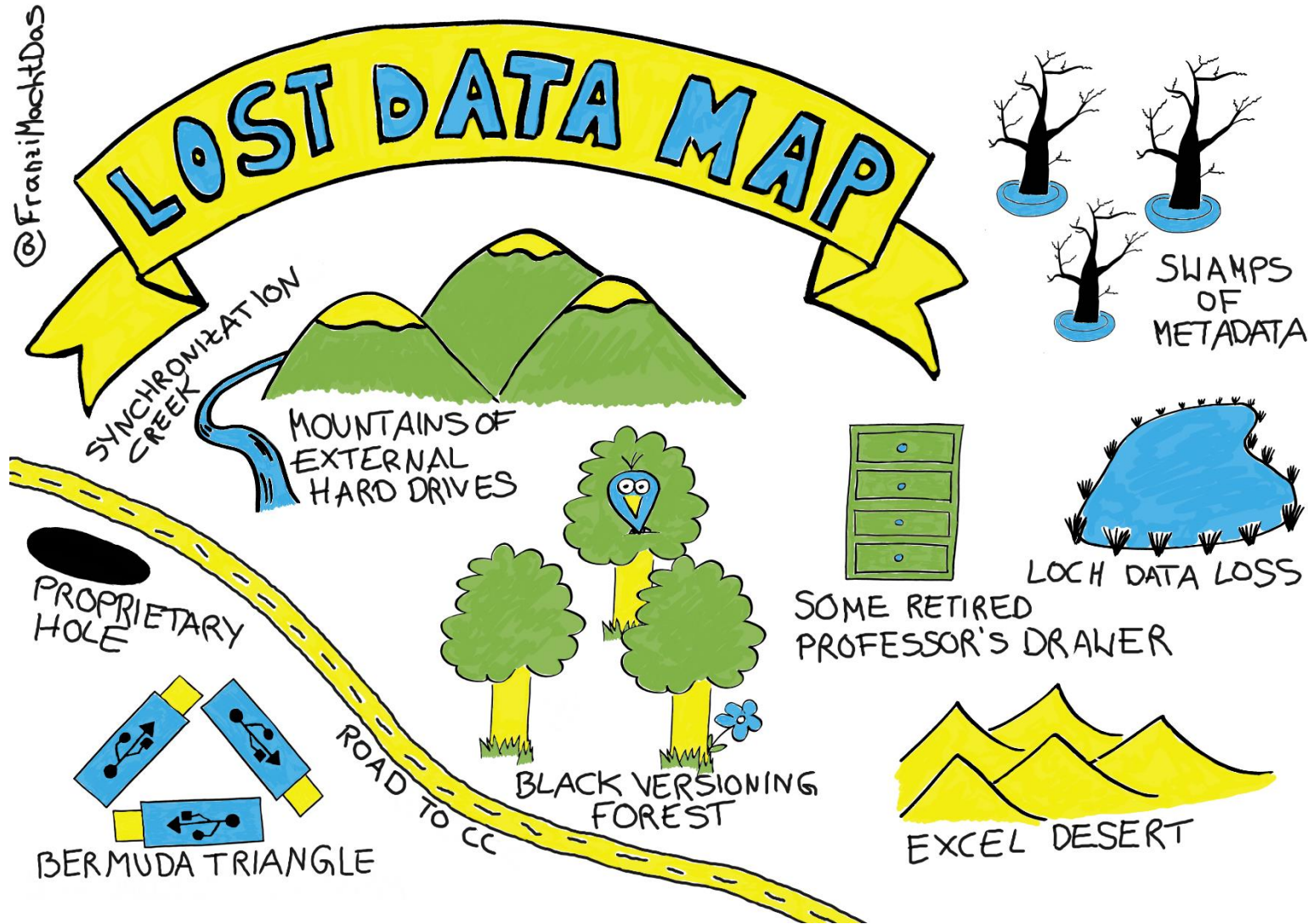
vs.

„available upon reasonable request“



<https://twitter.com/JohnBorghini/status/1388873617002299394/photo/1>,
CC BY-NC 2.5.

Probleme beim Management



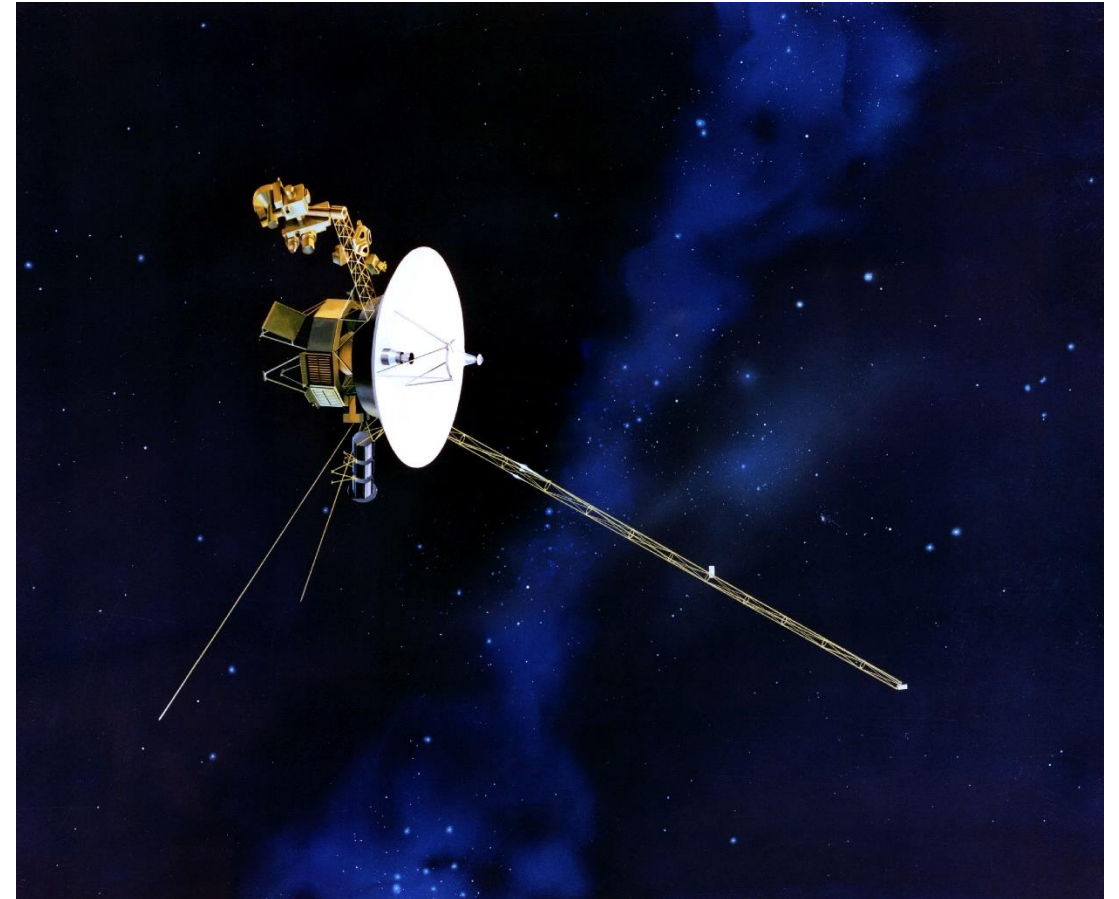
Franziska Mau, CC BY-SA 4.0,
<https://franzimachtdas.wordpress.com/sketches/>.

Probleme bei Forschungsdaten

„Fehlerhafte Daten von Voyager 1: Zähe Suche nach technischer Dokumentation

Seit zwei Monaten sendet Voyager 1 falsche Daten. Die Suche nach der Ursache wird dadurch erschwert, dass technische Dokumente zu der Sonde verschollen sind.“

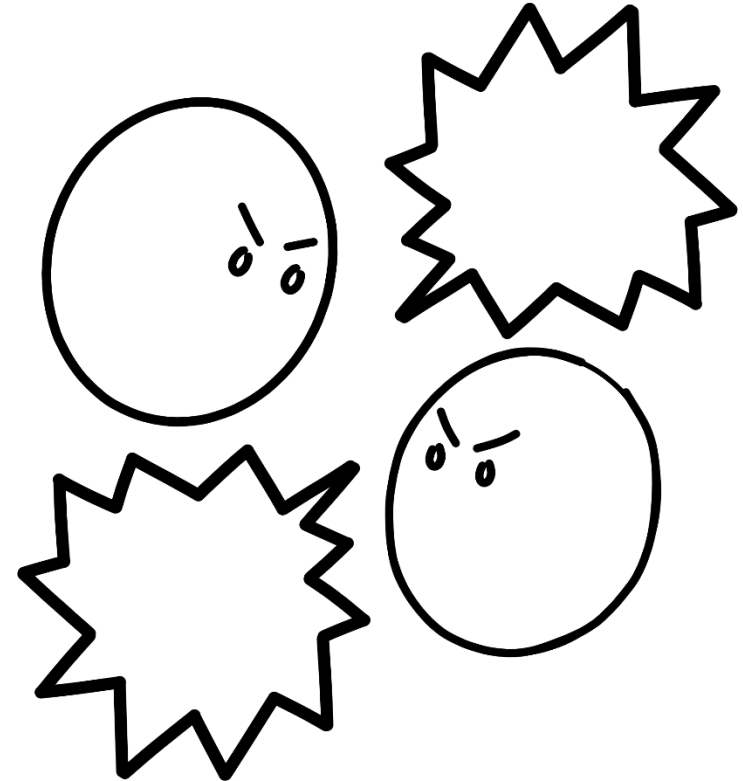
(<https://www.heise.de/news/Fehlerhafte-Daten-von-Voyager-1-Zaehe-Suche-nach-technischer-Dokumentation-7170603.html>)



NASA/JPL, Artist's concept of Voyager in flight, D=2194, gemeinfrei.

Probleme bei Forschungsdaten

- „Manipulierte“ Alzheimer-Daten könnten die Forschung seit 2006 in die Irre geführt haben
- <https://www.telegraph.co.uk/news/2022/07/21/manipulated-alzheimers-data-may-have-misled-research-16-years/>



Julian Kücklich, CC0,
<https://doi.org/10.5281/zenodo.2581783>

Allgemeine Literaturempfehlungen

- Putnings, Markus, Heike Neuroth, und Janna Neumann (Hrsg.) 2021: Praxishandbuch Forschungsdatenmanagement, <https://doi.org/10.1515/9783110657807>.
- Büttner, Stephan; Enke, Harry; Helbig, Kerstin; Hobohm, Hans-Christoph; Kindling, Maxi; Ludwig, Jens; Neumann, Janna; Pampel, Heinz; Schwarz, Karin (2016): Lehrbuch Forschungsdatenmanagement; https://handbuch.tib.eu/w/Lehrbuch_Forschungsdatenmanagement
- Büttner, Stephan; Hobohm, Hans-Christoph; Müller, Lars (2011): Handbuch Forschungsdatenmanagement; <https://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/208>

Linkempfehlungen

- <https://www.forschungsdaten.info> (Universität Konstanz)
- <https://www.forschungsdaten.org> (freie Community)
- <https://rdm.mpdl.mpg.de> (MPDL-Service)

Offene Fragen Ihrerseits?

Teil 2

Forschungsdateninfrastruktur

Was fällt Ihnen zu diesem Bild ein?

T-1					4 of 5				
me Cruise School Lighter					Whistle Station Buoy 2				
L.	W.	H.	T	L	W	L _n	L _w	L _H	L _w
15.0	5.6	8.5	29.1	51.5	19.2	29.2	268	1.76	1.52
13.0	4.0	7.8	24.8	52.5	16.1	31.4	325	1.67	1.95
14.5	4.1	7.3	25.9	56.0	15.8	28.2	354	1.99	1.78
12.3	4.4	7.6	24.3	50.6	18.1	31.3	280	1.62	1.73
13.6	4.5	7.3	25.4	53.5	17.7	28.8	302	1.86	1.62
11.5	4.7	6.2	22.4	51.3	21.0	27.6	245	1.85	1.32
11.6	3.9	6.0	21.5	54.0	18.1	27.9	298	1.93	1.54
13.2	5.5	8.4	27.1	48.7	20.3	31.0	240	1.57	1.53
10.4	3.6	5.6	19.6	53.1	18.3	28.6	289	1.86	1.56
11.1	4.5	6.1	21.7	51.2	20.7	28.1	247	1.82	1.35
10.0	3.0	5.0	18.0	55.6	16.6	27.8	333	2.00	1.66
9.6	3.4	5.6	18.6	51.5	18.3	30.1	242	1.71	1.65
8.5	3.4	4.1	16.0	53.0	21.2	25.7	250	2.07	1.21
8.5	3.5	5.5	17.5	48.5	20.0	21.4	242	1.54	1.57
7.2	2.2	4.5	13.9	51.7	15.8	32.4	328	1.60	2.04
8.4	2.6	4.6	15.0	53.8	16.6	29.5	213	1.83	1.77
8.2	3.5	4.4	16.1	50.9	21.7	27.2	234	1.86	1.25
7.5	2.2	4.5	14.2	52.8	15.5	31.7	340	1.66	2.04
6.6	2.1	3.3	12.0	55.0	17.5	27.5	314	2.00	1.57
6.1	2.1	3.2	11.4	53.5	18.4	28.1	290	1.91	1.52

Louis W. Hutchins, Buoy fouling survey, research data, Smithsonian Institution Archives, no known copyright restrictions.

Inhalt von Teil 2

1. Forschungsdateninfrastruktur
2. Forschungsdatenrepositorien
3. Datenmanagementpläne
4. FAIR-Prinzipien
5. Metadaten und Lizenzen

Forschungsdateninfrastruktur

Infrastruktur?

- z.B. Verkehrsinfrastruktur
 - Straßen, Brücken
 - Schienen, Wasserwege
 - Tankstellennetz, Verkehrsfunk, GPS
 - Straßenverkehrsordnung, Straßenwacht, ADAC
 - Räumdienst, Mautsystem, KFZ-Steuer
 - Autoindustrie, Verkehrsministerium, ÖPNV



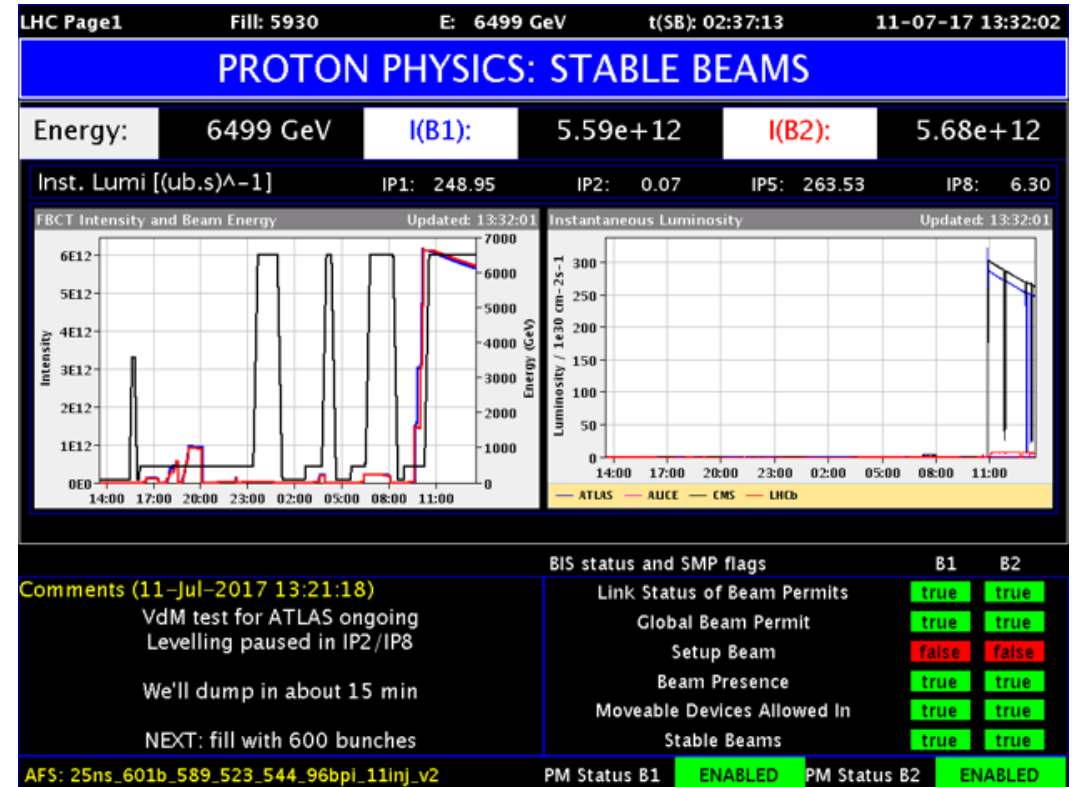
Bundesarchiv, Bild 183-F0307-0001-046, Straßenverkehr in Leipzig, 1967, CC BY SA 3.0.

Forschungsdateninfrastruktur

- Wiss. (Groß-)Anlagen
 - Metadaten, Standards, Identifikatoren
 - Repositorien, Registries, Datenjournale
 - Rechtsgrundlagen
-
- Nationale Forschungsdateninfrastruktur
 - European Open Science Cloud

Wissenschaftliche Großanlagen

- Large Hadron Collider
- Teilchenbeschleuniger am Europäischen Kernforschungszentrum CERN
- home.web.cern.ch/about/accelerators/large-hadron-collider



Wissenschaftliche Großanlagen

Polarstern

Forschungsschiff des Alfred-Wegener-Institut für Polar- und Meeresforschung

<https://www.awi.de/expedition/schiffe/polarstern.html>

Data Flow Framework:

<https://www.awi.de/en/about-us/service/computing-centre/data-flow-framework.html>



Janek Uin: MOSAIC-Expedition, Polarwinter 2019/2020, CC BY 4.0.

Wissenschaftliche Großanlagen

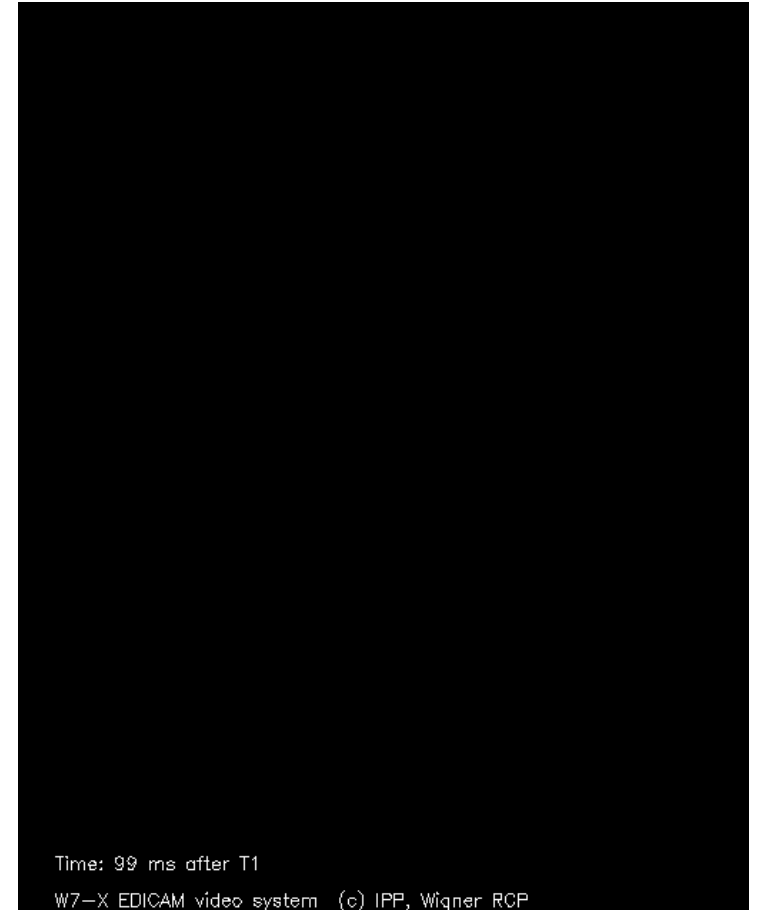
- Antennen des Atacama Large Millimeter/submillimeter Array (ALMA), auf dem Chajnantor Plateau in den chilenischen Anden



ESO/C. Malin, The Atacama Large Millimeter/submillimeter Array (ALMA) by night, under the Magellanic Clouds, 2013, CC BY 4.0, <https://www.eso.org/public/images/ann13016a/>.

Wissenschaftliche Großanlagen

- Wendelstein 7-X
- weltweit größte Fusionsanlage vom Typ „Stellarator“
- Ziel: Kraftwerkseignung dieses Bautyps untersuchen
- <https://www.ipp.mpg.de/wendelstein7x>



IPP/Wigner RCP

Wissenschaftliche Großanlagen

- Erheben/Erzeugen/Verteilen Daten
- Erweitern Möglichkeiten
- Beschleunigen Prozesse
- Zwingen zur Kooperation

Forschungsdatenrepositorien

Einteilung von Forschungsdatenrepositorien

1. Fachlich

- [PsychData](#)
- [Pangaea](#)
- [G-Node](#)

2. Institutionell

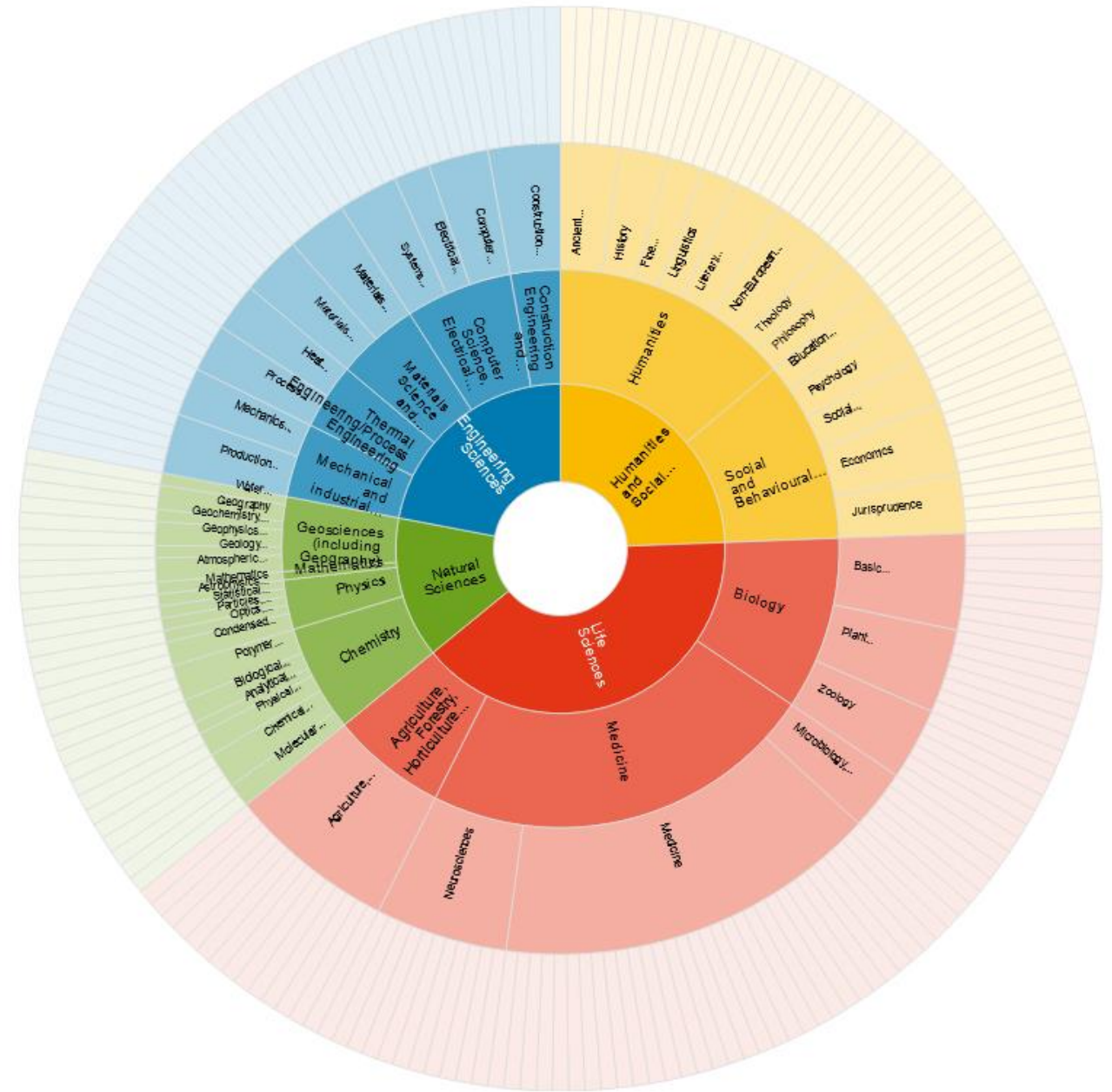
- [Open Data LMU](#)
- [Edinburgh DataShare](#)
- [Edmond](#)

3. Allgemein

- [Zenodo](#)
- [Dryad](#)
- [Radar](#)
- [Figshare](#)
- [Mendeley Data](#)

Forschungsdaten-repositorien

- Repositorien finden:
- re3data (<https://www.re3data.org>)
- Bild: Search by Subject, CC BY 4.0



Zertifizierungen

- CoreTrustSeal (<https://www.coretrustseal.org>)
- CLARIN certificate B (<https://www.clarin.eu>)
- DIN 31644 (<https://www.beuth.de/en/standard/din-31644/147058907>)
- DINI Certificate (<https://dini.de/dienste-projekte/dini-zertifikat/>)
- RatSWD (<https://www.konsortswd.de/ratswd/>)
- ...

Recherche-Aufgabe

Frage:

Sind die Forschungsdatenrepositorien an Ihrer Institution zertifiziert?
Wenn ja, wie wird damit geworben?

Falls kein Datenrepositorium vorhanden, wählen Sie ein interessantes Repository aus [re3data](https://re3data.org/) aus und recherchieren dazu.

Methode:

Online-Recherche

Dauer:

5 Minuten

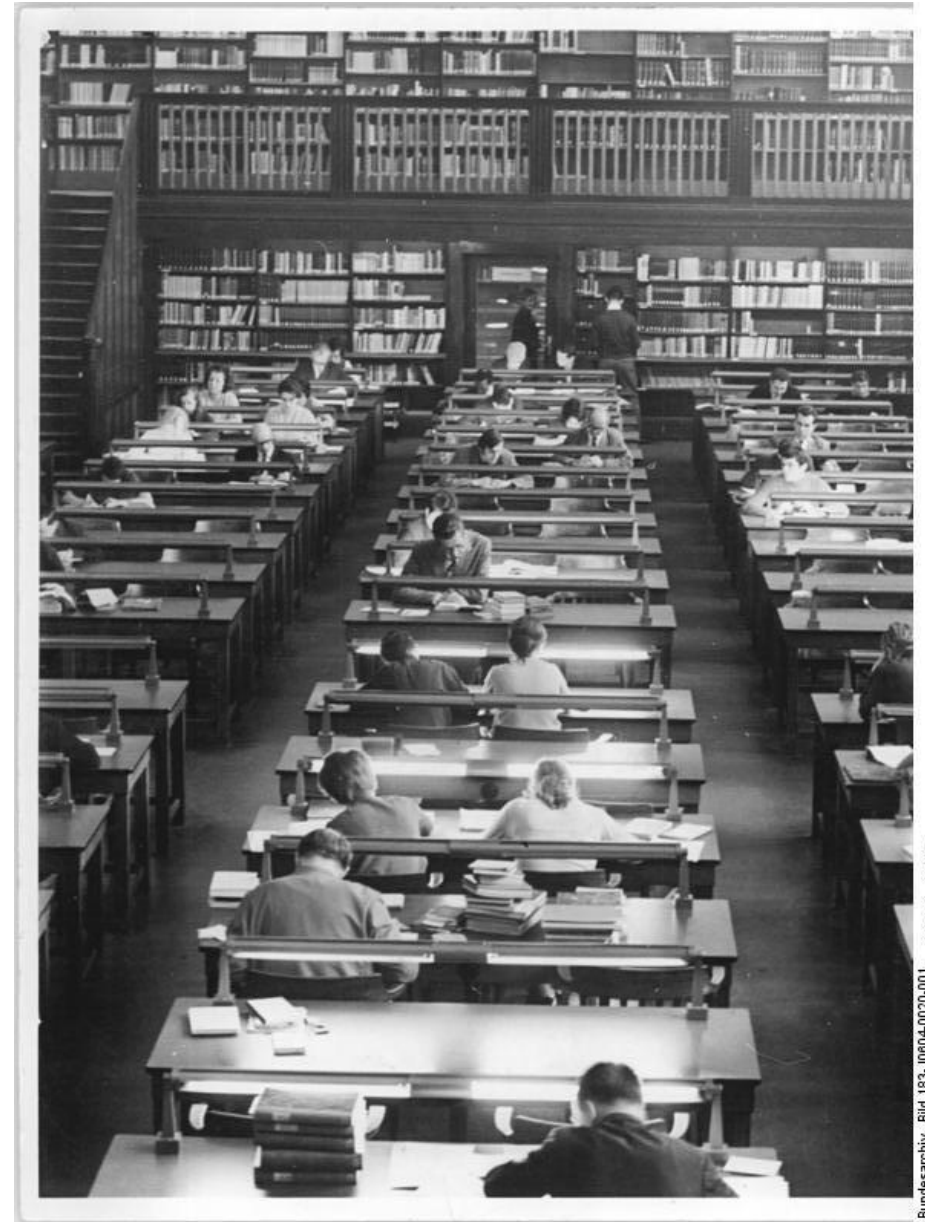
Auswertung:

Gemeinsame Diskussion der Ergebnisse

Datenmanagementpläne

Wieso überhaupt einen DMP schreiben?

- zunehmend von Förderorganisationen gefordert (z.B. Horizon Europe)
- Kosten für das Datenmanagement (Software, Hardware, technisches Fachwissen) sind häufig förderfähig
- Eigenes Projektmanagement
- Teil eines Beratungsgesprächs



Bundesarchiv, Bild 183-J0604-0020-001 / Raphael (verehel. Grubitzsch), Waltraud / CC BY SA 3.0.

Bundesarchiv - Bild 183-J0604-0020-001
Foto: Raphael verehel. Grubitzsch, Waltraud | 4. Juni 1970

Wie schreibt man einen DMP?

1. Brainstorming
2. Vorlage nutzen
3. Digitales Werkzeug nutzen

Vorlagen für DMPs

- [Horizon Europe DMP Template by the European Commission, 2021](#)
- [Horizon 2020 DMP Template by the European Commission, 2018](#)
- [ERC Grants DMP Template by the European Research Council, 2021](#)
- [DMP Template by the Swiss National Science Foundation, 2017](#)

How to DMP

- Argos (<https://argos.openaire.eu>)
 - Data Stewardship Wizard (<https://ds-wizard.org>)
 - DMPTool (<https://dmptool.org>)
 - DMPOnline (<https://dmponline.dcc.ac.uk>)
 - RDMO (<https://rdmo.aip.de>)
-
- Helbig, K., Paul-Stüve, T., & Rex, J. (2021): DMP-Toolguide, [doi:10.5281/zenodo.4632308](https://doi.org/10.5281/zenodo.4632308).

Beispiele für DMPs

- Digital Curation Centre Example DMPs and Guidance:
<http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>
- LIBER Europe DMP Catalogue:
<https://libereurope.eu/dmpcatalogue/>
- Examples for Horizon 2020 DMPs by the University of Vienna:
<https://phaidra.univie.ac.at/search#?page=1&pagesize=10&collection=o:1140797>

Aspekte eines DMPs

Die DMP umfassen in der Regel die folgenden Aspekte:

1. Beschreibung und Erhebung von Daten
2. Dokumentation und Metadaten
3. Speicherung und Sicherung
4. Rechtliche und ethische Anforderungen
5. Gemeinsame Nutzung und langfristige Aufbewahrung von Daten
6. Zuständigkeiten und Ressourcen

Mögliche Fragen eines DMPs

Die folgenden Fragen können in einem DMP beantwortet werden:

- Welche Daten werden erhoben/genutzt?
- Wie werden die Daten verarbeitet?
- Welche Standards werden verwendet?
- Wie werden die Daten dokumentiert?
- Wo werden die Daten (offen) gespeichert und wann?
- Was, wenn überhaupt, spricht gegen eine Veröffentlichung?
- Welche Kosten entstehen durch das Forschungsdatenmanagement?

Checkliste von der ETH Zürich (https://www.dlcm.ch/download_file/force/66/371)

Live-Demonstration RDMO

<https://rdmo.forschungsdaten.info>

Sign-up: <https://rdmo.forschungsdaten.info/account/signup/>

Kosten für Datenmanagement

- OpenAIRE: <https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>
- Data Management Costing Tool by the UK Data Archives: <https://www.ukdataservice.ac.uk/manage-data/plan/costing>

FAIR-Prinzipien

FAIR-Prinzipien

Findable

Accessible

Interoperable

Reusable

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016).

<https://doi.org/10.1038/sdata.2016.18>.

Findable

- F1. (Meta)data are assigned a globally unique and persistent identifier
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

<https://www.go-fair.org/fair-principles/>

Accessible

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

A1.1 The protocol is open, free, and universally implementable

A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

A2. Metadata are accessible, even when the data are no longer available

<https://www.go-fair.org/fair-principles/>

Interoperable

1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
2. (Meta)data use vocabularies that follow FAIR principles
3. (Meta)data include qualified references to other (meta)data

<https://www.go-fair.org/fair-principles/>

Reusable

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

R1.1. (Meta)data are released with a clear and accessible data usage license

R1.2. (Meta)data are associated with detailed provenance

R1.3. (Meta)data meet domain-relevant community standards

<https://www.go-fair.org/fair-principles/>

Test: F-Uji



<https://www.f-uji.net>

i.e. <https://doi.org/10.5281/zenodo.5566761>

Identifler aus Recherche-Repository in F-Uji

Metadaten und Lizenzen

Metadaten

Author ?	Giorgetta, Marco   Max Planck Institute for Meteorology 
Study Type ?	simulation/modelling
Description ?	This dataset provides additional information related to the publication "The ICON-A model for direct QBO simulations on GPUs (version icon-cscs:baf28a514)" submitted to Geoscientific Model Development (GMD). Details for the content are given in the README.pdf file.
Subject ?	Geosciences
Keyword ?	Global atmospheric modelling storm resolving GPU port ICON atmosphere model weak and strong scaling quasi-biennial oscillation
Depositor ?	Giorgetta, Marco
Deposit Date ?	2022-03-28
Related Publication ?	Giorgetta, M. A., Sawyer, W., Lapillonne, X., Adamidis, P., Alexeev, D., Clément, V., Dietlicher, R., Engels, J. F., Esch, M., Franke, H., Frauen, C., Hannah, W. M., Hillman, B. R., Kornblueh, L., Marti, P., Norman, M. R., Pincus, R., Rast, S., Reinert, D., Schnur, R., Schulzweida, U., and Stevens, B.: The ICON-A model for direct QBO simulations on GPUs (version icon-cscs:baf28a514), submitted to Geosci. Model Dev., 2022.

Metadaten von <https://doi.org/10.17617/3.5CYUFN>

Metadaten

- Bibliographische Metadaten
 - [DC](#) – Dublin Core Metadata Initiative
 - [SKOS](#) – Simple Knowledge Organization System
- Fachliche Metadaten
 - [RDF](#) – Resource Description Framework
 - [Metadaten-Standards](#)
 - ([Metadaten-Registries](#))
- Technische Metadaten, Lizenzmetadaten, Provenienzmetadaten

Persistente Identifikation

- URL/URI: <http://zuse.zib.de/collection/wl3XoEDHO8v0ImCa/item/VqNgKUacPrIhqPKu>
- URN: urn:nbn:de:bsz:25-opus-14124 (Resolver z.B. <https://nbn-resolving.org>)
- Handle: <hdl:11858/00-001M-0000-0019-D20F-6>
- DOI: <doi:10.1088/0004-637X/715/2/1453>

- ORCID: [0000-0002-2880-8947](https://orcid.org/0000-0002-2880-8947)

- ROR: [0061msm67](https://ror.org/0061msm67)
- ~~GRID: [grid.431460.4](https://www.grid.ac/details/grid.431460.4)~~

Digital Object Identifier

- für digitale Objekte
- <https://www.doi.org>
- DOIs für Forschungsdaten durch DataCite (im Gegensatz zu Crossref)



Public domain

Open Researcher and Contributor ID

- für Personen
- <https://orcid.org>



Public domain, CC0

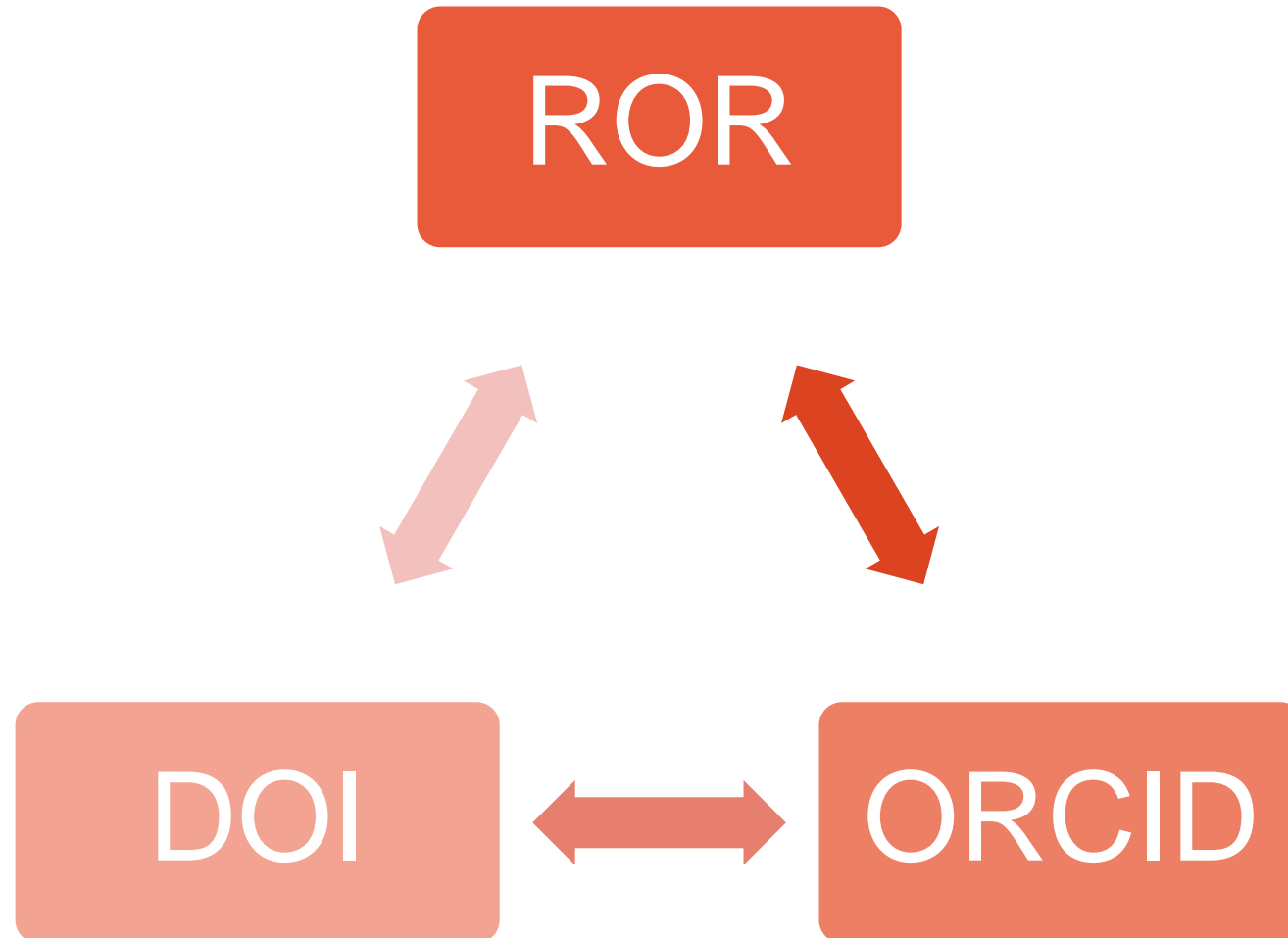
Research Organization Registry

- für Organisationen
- <https://ror.org>



Research Organization Registry, CC BY 4.0,
<https://doi.org/10.5281/zenodo.4701802>

Identifier Network



Live Demonstration

- Beispiel Ludwig-Maximilians-Universität München:
 - <https://ror.org/05591te55>
 - <https://commons.datacite.org/ror.org/05591te55>

Daten-Lizenzen

- Creative Commons (<https://creativecommons.org>)
 - CC 0
 - CC BY
- Open Data Commons (<https://opendatacommons.org>)
 - ODC-By
 - PDDL
- Datenlizenz Deutschland

FDM-Tabu

Teilnehmende:

2 Mannschaften

Spielregeln:

Eine Person bekommt einen FDM-Begriff und muss ihn erklären, ohne den Begriff oder Bestandteile zu verwenden. Die beiden Mannschaften müssen den Begriff erraten.

Ziel:

Die Mannschaft mit den meisten erratenen Begriffen gewinnt.

Literaturempfehlungen

- allgemein: Fotis Jannidis, Hubertus Kohle, und Malte Rehbein (Hrsg.): Digital Humanities: eine Einführung, Stuttgart 2017.
- zu Repositorien: Friedrich Summann und Andreas Czerniak: „Die deutsche und europäische Repositorien-Landschaft im Überblick“, gehalten auf den Open Access Tagen 2020 (OAT2020), Bielefeld, 17. September 2020 <https://doi.org/10.5281/zenodo.4034075>.
- zu Metadaten: Jana Baur: „FDM erklärt – Metadaten & Metadatenstandards“, 13. August 2021. <https://blog.rwth-aachen.de/forschungsdaten/2021/08/13/fdm-erklaert-metadaten-metadatenstandards/>.
- zu Lizenzen: Brettschneider et al. (2021): Offene Lizenzen für Forschungsdaten: Rechtliche Bewertung und Praxistauglichkeit verbreiteter Lizenzmodelle. O-Bib. Das Offene Bibliotheksjournal 8(3), S. 1–22, <https://doi.org/10.5282/o-bib/5749>.

Teil 3

Forschungsdatenmanagement in Bibliotheken

Inhalt von Teil 3

1. FDM an Bibliotheken
2. LIBER Recommendations
3. Institutionelle Selbstevaluierung
4. Science Clouds
 1. Nationale Forschungsdateninfrastruktur
 2. European Open Science Cloud
 3. GOSC & GAIA-X

FDM an Bibliotheken, Warum?

- Neue Technologien
- Neue Anforderungen
- Neue kommerzielle Dienste
- Wahrnehmung
- Krise der Bibliotheken? Bibliotheken im Wandel?

Kommission Virtuelle Bibliothek (KVB) im BVB: Die Rolle der wissenschaftlichen Bibliotheken Bayerns beim Forschungsdatenmanagement – Herausforderungen, Aufgabenfelder, Handlungsempfehlungen, 2016, https://www.bib-bvb.de/documents/11301/1476009/PositionspapierKVB_August_2016.pdf/369b9b2c-0513-45e9-b3eb-72819613de9d.

LIBER

- Ligue des bibliothèques européennes de recherche (<http://libereurope.eu>)
- 450 Forschungsbibliotheken
- 62 deutsche Bibliotheken
- München: UB TU, BSB, UB UniBwM, MPDL
- <https://libereurope.eu/liber-participants/>

Liber Recommendations for Libraries

Final report of the LIBER working group
on E-Science / Research Data
Management, 2012,
[https://libereurope.eu/wp-
content/uploads/2020/11/The-research-
data-group-2012-v7-final.pdf](https://libereurope.eu/wp-content/uploads/2020/11/The-research-data-group-2012-v7-final.pdf).



Ten recommendations for libraries to get started with research data management

Final report of the LIBER working group on E-Science / Research Data Management

Birte Christensen-Dalsgaard et al
4-7-2012

Liber Recommendations for Libraries

1. Offer research data management support, including data management plans for grant applications, intellectual property rights advice and information materials. Assist faculty with data management plans and the integration of data management into the curriculum.

2. Engage in the development of metadata and data standards and provide metadata services for research data.

3. Create Data Librarian posts and develop professional staff skills for data librarianship.

4. Actively participate in institutional research data policy development, including resource plans. Encourage and adopt open data policies where appropriate in the research data life cycle.

5. Liaise and partner with researchers, research groups, data archives and data centers to foster an interoperable infrastructure for data access, discovery and data sharing.

6. Support the lifecycle for research data by providing services for storage, discovery and permanent access.

7. Promote research data citation by applying persistent identifiers to research data.

8. Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.

9. Get involved in subject specific data management practice.

10. Offer or mediate secure storage for dynamic and static research data in co-operation with institutional IT units and/or seek exploitation of appropriate cloud services:

Final report of the LIBER working group on E-Science / Research Data Management, 2012, <https://libereurope.eu/wp-content/uploads/2020/11/The-research-data-group-2012-v7-final.pdf>.

LIBER Recommendation: 1

“Offer research data management support, including data management plans for grant applications, intellectual property rights advice and information materials. Assist faculty with data management plans and the integration of data management into the curriculum.”

Mögliche Angebote einer Bibliothek:

- Unterstützung bei DMPs
- Beratung bei Datenlizenzen
- FDM-Schulungen
- ...

LIBER Recommendation: 2

Engage in the development of metadata and data standards and provide **metadata services** for research data.

Mögliche Angebote einer Bibliothek:

- Beratung
- Metadaten-Services
 - Bibliographische Metadaten
 - Transformationen
 - Kontrollierte Vokabulare

LIBER Recommendation: 3

Create Data Librarian posts and develop professional staff skills for **data librarianship**.

- [OpenBiblioJobs](#)
- Hapke, H. (2016) Data Librarian: Das moderne Berufsbild; b.i.t. online, 19 (2016) Nr. 2, S. 159-164, <https://www.b-i-t-online.de/heft/2016-02-fachbeitrag-hapke.pdf>.
- Georgy, Ursula, und Marvin Lanczek: „Big Data in der Bibliothek bewältigen: Der ZBIW-Zertifikatskurs Data Librarian“. Leipzig, 18. März 2019, <https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/16475>.
- Thomas, Ashley, und Elaine R. Martin (2020): Developing a Community of Practice: Building the Research Data Management Librarian Academy, Medical Reference Services Quarterly 39 (4), S. 323-33. <https://doi.org/10.1080/02763869.2020.1826185>.

LIBER Recommendation: 4

Actively participate in institutional research data policy development, including resource plans. Encourage and adopt open data policies where appropriate in the **research data life cycle**.

- Rahmen für den Umgang mit Forschungsdaten an einer Institutionen
- Zentrales Beispiel für Deutschland: [DFG „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“](#)
- umfangreiche Auflistung: https://www.forschungsdaten.org/index.php/Data_Policies

LIBER Recommendation: 5

Liaise and partner with researchers, research groups, data archives and data centers to **foster an interoperable infrastructure for data access, discovery and data sharing.**

- Siehe FAIR-Prinzipien Teil 2

LIBER Recommendation: 6

Support the lifecycle for research data by **providing services for storage, discovery and permanent access.**

LIBER Recommendation: 7

Promote research data citation by **applying persistent identifiers to research data.**

LIBER Recommendation: 8

Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.

- siehe auch re3data in Teil 2

LIBER Recommendation: 9

Get involved in **subject specific data management** practice.

LIBER Recommendation: 10

Offer or mediate secure storage for dynamic and static research data in co-operation with institutional IT units and/or seek exploitation of **appropriate cloud services.**

Offene Fragen Ihrerseits?

Nächste Sitzung via Slido

Fragen

1. Welche Themen hätten Sie gerne in der nächsten Sitzung besprochen?

Vorgehen

Im Browser www.slido.com öffnen. Code 5029111 eingeben.

Recherche für die nächste Stunde

Gibt es Richtlinien oder Standards an Ihrer Institution, welche sich auf Forschungsdaten beziehen?

Forschungsdatenmanagement

Hochschule für den öffentlichen Dienst in Bayern
Fachbereich Archiv- und Bibliothekswesen
Ausbildungsjahr 2021/2024

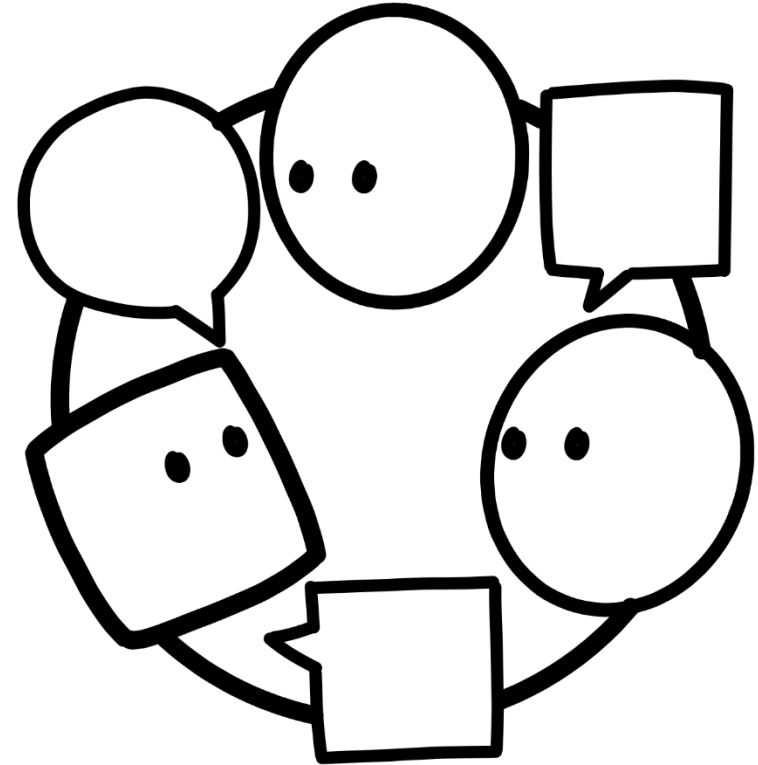
Dr. Yves Vincent Grossmann
grossmann@mpdl.mpg.de
Max Planck Digital Library

Willkommen zurück

Fragen zur letzten Sitzung?

Wheel of Names zur letzten Sitzung

- Gute Wissenschaftliche Praxis
- Probleme
- Wissen. Großanlagen
- FAIR-Prinzipien
- Qualitative Forschungsdaten
- Metadaten
- re3data
- ROR
- DOI
- ORCID
- Datenmanagementplan (DMP)
- Quantitative Forschungsdaten
- Forschungsdatenrepositorien
- LIBER Recommendations



Julian Kücklich, CC0,
<https://doi.org/10.5281/zenodo.2581783>

Teil 3

Forschungsdatenmanagement in Bibliotheken

Inhalt von Teil 3

1. FDM an Bibliotheken
2. LIBER Recommendations
3. Institutionelle Selbstevaluierung
4. Science Clouds
 1. Nationale Forschungsdateninfrastruktur
 2. European Open Science Cloud
 3. GOSC & GAIA-X

Institutionelle Selbstevaluierung

RISE

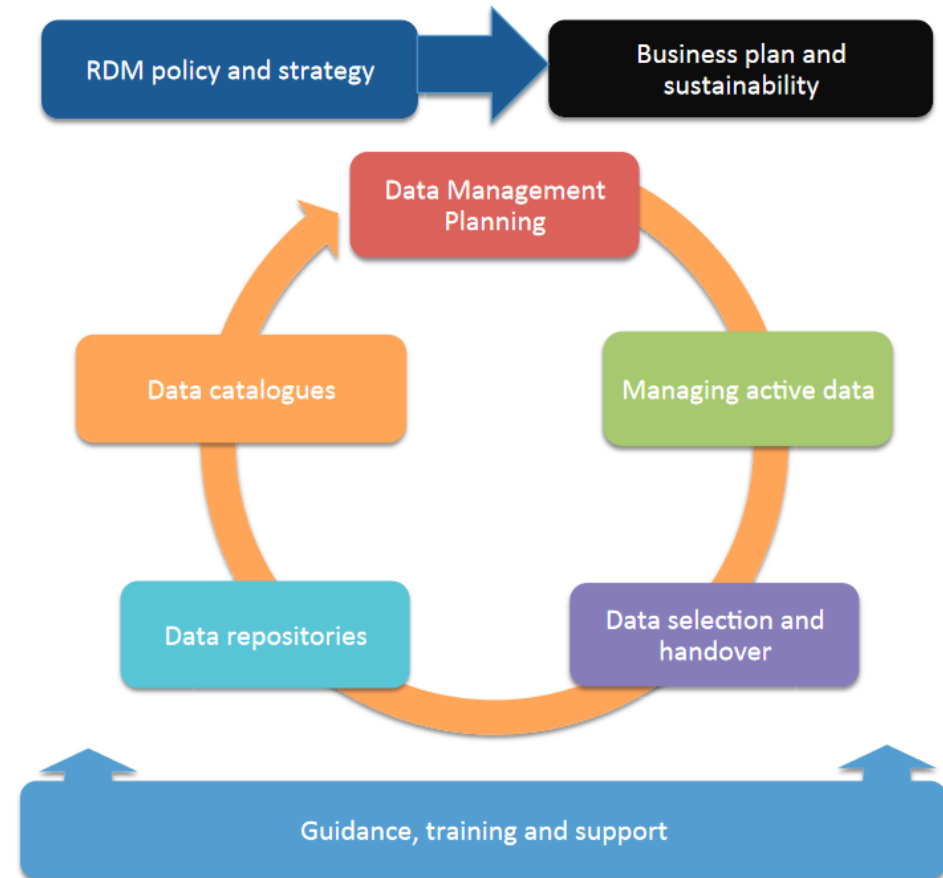
- Research Infrastructure Self-Evaluation Framework (RISE v1.1) des UK Digital Curation Centre (DCC) von 2017
- <https://www.dcc.ac.uk/guidance/how-guides/RISE>



Rans, J and Whyte, A. (2017). 'Using RISE, the Research Infrastructure Self-Evaluation Framework' v.1.1 Edinburgh: Digital Curation Centre. Available online: www.dcc.ac.uk/guidance/how-guides, CC BY 4.0.

RISE

- Komponenten der Unterstützungsdienste für das Forschungsdatenmanagement



Jones, S., Pryor, G. & Whyte, A. (2013). 'How to Develop Research Data Management Services – A guide for HEIs'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: [/guidance/how-guides](#), CC BY 4.0, p. 5.

RISE-DE

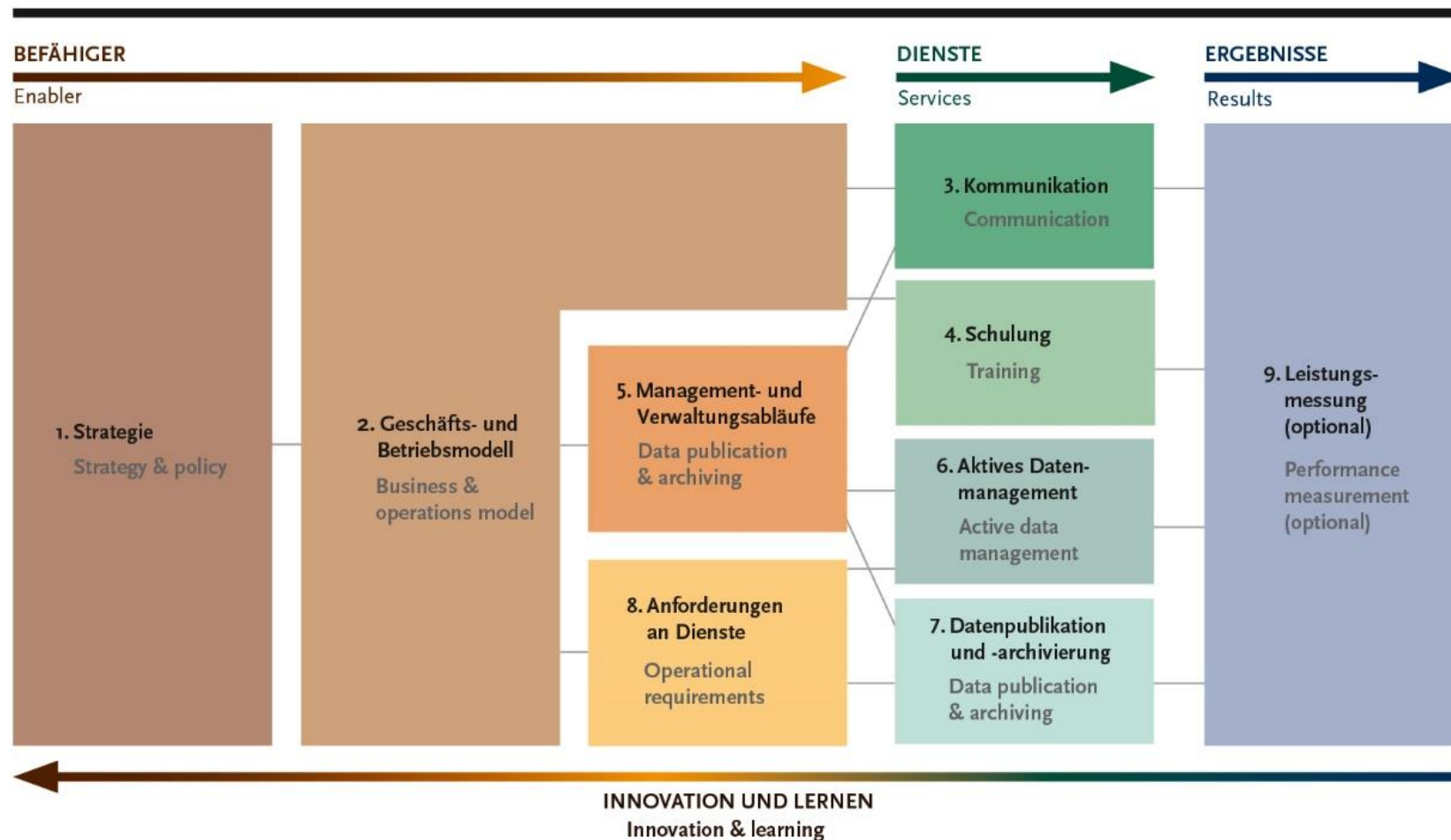
- <https://zenodo.org/record/3585556>
- RISE-DE orientiert sich an dem Research Infrastructure Self-Evaluation Framework (RISE v1.1)
- Anpassung an bundesdeutsche Spezifika in der Bildungslandschaft



Hartmann, Jacob und Weiß (2019). RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement (1.0), CC BY 4.0, <https://doi.org/10.5281/zenodo.3585556>.

RISE-DE

RISE-DE Referenzmodell



Hartmann, Niklas K., Jacob, Boris, & Weiß, Nadin. (2019). RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement (1.0). Zenodo. <https://doi.org/10.5281/zenodo.3585556>, CC BY 4.0, p. 9.

Themenfelder von RISE-DE

1 Strategy

- 1.1 Strategy development
- 1.2 Strategic goals
- 1.3 Commitment and incentives
- 1.4 RD Policy

2 Business and Operating Model

- 2.1 Governance and organisational development
- 2.2 Human resource investments
- 2.3 Technology investments
- 2.4 Cost modelling

3 Communication

- 3.1 Information
- 3.2 Advice
- 3.3 Public Relations

4 Training

- 4.1 Training materials
- 4.2 Training events

Themenfelder von RISE-DE

5 Management processes and administrative procedures

- 5.1 Integration into project planning
- 5.2 Integration with research information
- 5.3 Integration with legal and ethical processes

6 Active research data management

- 6.1 Basic IT services
- 6.2 Services for Cooperative Work

7 Data publication and archiving

- 7.1 Publication services
- 7.2 Archiving services
- 7.3 Risk management

8 Requirements for services

- 8.1 Fail-safety
- 8.2 Data security
- 8.3 Metadata and indexing
- 8.4 Standard conformity of publication and LTA service

Stufen der RISE-DE-Bewertungsskala

Level	Definition	Typical Requirement
0	No (significant) activity	None
1	Minimal activity	External Requirements, Preservation of Third party funding
2	Committed activity	Needs of the researchers of the institution
3	Outstanding activity	National or international industry leader

Hartmann, Niklas K., Jacob, Boris, & Weiß, Nadin. (2019). RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement (1.0). Zenodo. <https://doi.org/10.5281/zenodo.3585556>, CC BY 4.0, p. 10.

RISE-DE Evaluation Tool

- <https://zenodo.org/record/3585556/files/RISE-DE%20V1.0%20Evaluations-Tool.xlsx?download=1>
- ein xlsx-Dokument

1.1 Strategieentwicklung	
Das Thema Strategieentwicklung beschreibt die Qualität und den Umfang des Prozesses zur Selbstevaluation und Strategieentwicklung.	
Stufe 0	Aktivitäten ohne strategische Planung
	Die Aktivitäten der Zentralen Einrichtungen in Bezug auf Forschungsdaten, insbesondere Auf- und Ausbau von Diensten, erfolgen ohne strategische Planung.
Stufe 1	Einfache Strategieentwicklung
	Die Strategieentwicklung erfolgt durch von der Leitung beauftragte Personen. Als Basis der Selbstevaluation dienen Informationen über die Forschungsdatenpraxis an der eigenen Einrichtung, die aus Verwaltungsdaten und Befragungen gewonnen werden. Die Ziele werden von der Leitung festgelegt und ggf. durch Gremien legitimiert.
Stufe 2	Strukturierte Strategieentwicklung unter Einbindung einzelner Akteure
	Die Strategieentwicklung erfolgt durch von der Leitung beauftragte Personen, die gezielt Kontakt zu relevanten Akteuren aufnehmen, um diese einzubinden und weitere Informationen zu gewinnen. Die Strategieentwicklung wird durch Verwendung strukturierter Instrumente zur Selbstbewertung und Zielbestimmung unterstützt. Die Selbstbewertung wird von den relevanten Gremien zur Kenntnis genommen und die Zielbestimmung durch diese legitimiert.
Stufe 3	Partizipative strukturierte Strategieentwicklung
	Zur Strategieentwicklung werden Selbstevaluation und Zielbestimmung im Rahmen eines strukturierten, Teilorganisations- und Hierarchieebenen-übergreifenden Stakeholder-orientierten, durch Leitung und Gremien legitimierten Prozesses durchgeführt.

Hartmann, Jacob und Weiß (2019). RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement (1.0), CC BY 4.0, <https://doi.org/10.5281/zenodo.3585556>.

Empfehlungen für FDM- und Strategieprozess-Einsteiger

Bei der Bewertung des institutionellen RDM einer Einrichtung und der Entwicklung einer Strategie zur Einrichtung oder Verbesserung von Diensten ist es sinnvoll, in mehreren Schritten vorzugehen:

1. Grundsatzentscheidungen, Vorüberlegungen und Initiierung
2. Evaluierung mit Beteiligung der Betroffenen
3. Entwicklung und Vereinbarung eines Textentwurfs
4. Verabschiedung und Veröffentlichung

Weitere Evaluierungsmodell: DIAMANT

Lemaire, Marina / Gerhards, Lea / Kellendonk, Stefan / Blask, Katarina / Förster, André:
Das DIAMANT-Modell 2.0. Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft.
(Universität Trier eSciences Working Papers, Nr. 05), Trier 2020,
<https://doi.org/10.25353/ubtr-xxxx-f5d2-fffb>, CC BY 4.0.



Literaturempfehlung

- Petersen et al. (2022): Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards, <https://doi.org/10.5281/zenodo.7034478>.
- DINI AG Forschungsinformationssysteme (2022): Management von Forschungsinformationen in Hochschulen und Forschungseinrichtungen, <https://doi.org/10.18452/25440>.
- Helbig, Kerstin, Katarzyna Biernacka, Petra Buchholz, Dominika Dolzycka, Niklas Hartmann, Thomas Hartmann, Beate Maria Hiemenz, u. a. (2019): Lösungen und Leitfäden für das institutionelle Forschungsdatenmanagement, in: o-bib. Das offene Bibliotheksjournal 6 (3): S. 21–39, <https://doi.org/10.5282/o-bib/2019H3S21-39>.

Antworten zur eigenen Recherche

Gibt es Richtlinien oder Standards an Ihrer Institution, welche sich auf Forschungsdaten beziehen?

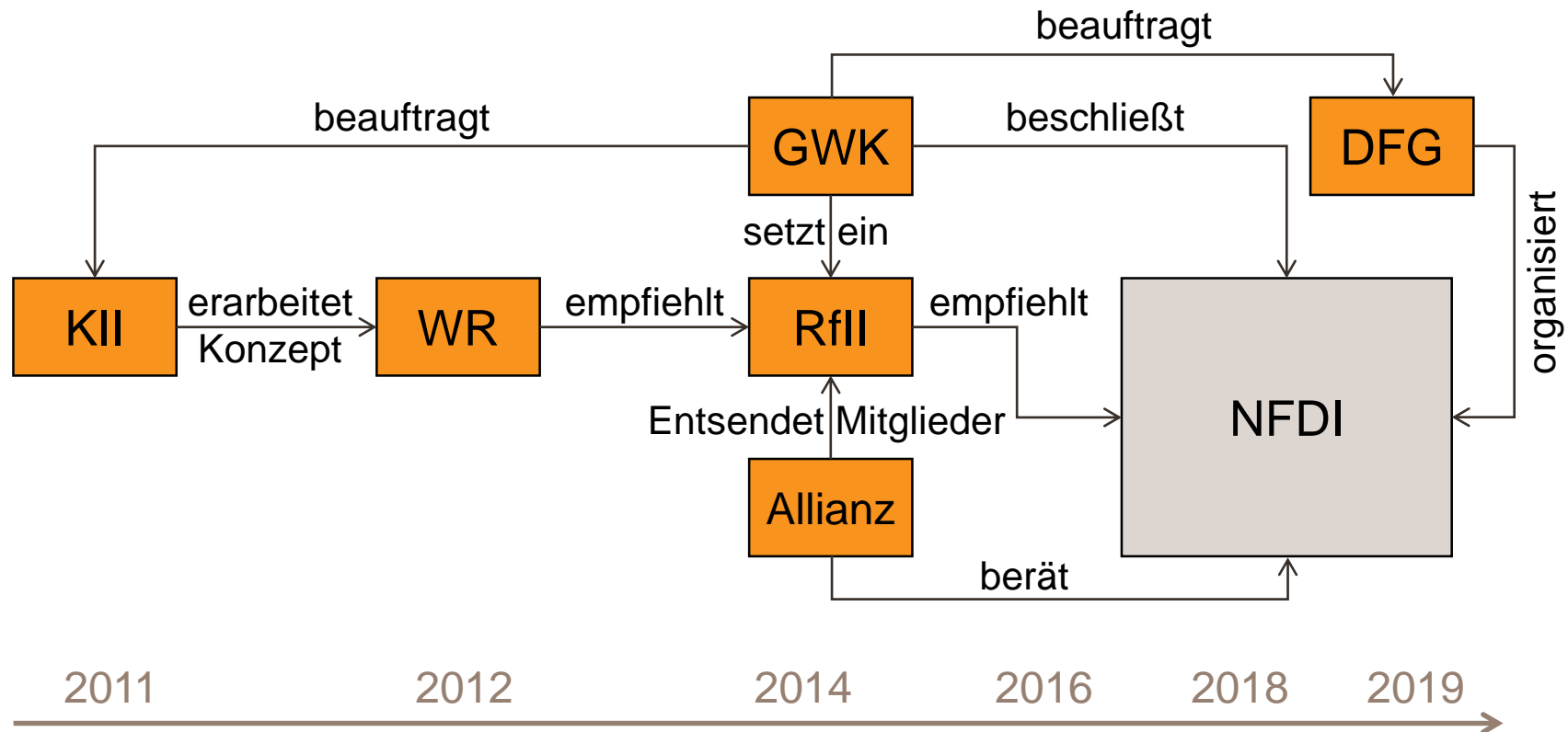
Nationale Forschungsdateninfrastruktur

Nationale Forschungsdateninfrastruktur

„Die nationale Forschungsdateninfrastruktur (NFDI) soll die Datenbestände von Wissenschaft und Forschung systematisch erschließen, nachhaltig sichern und zugänglich machen sowie (inter-)national vernetzen. Sie wird in einem aus der Wissenschaft getriebenen Prozess als vernetzte Struktur eigeninitiativ agierender Konsortien aufgebaut werden.“ (DFG-Website)

- NFDI: <https://www.nfdi.de>
- DFG: <https://www.dfg.de/foerderung/programme/nfdi/index.html>

Nationale Forschungsdateninfrastruktur



DFG: Deutsche Forschungsgemeinschaft

GWK: Gemeinsame Wissenschaftskonferenz

KII: Kommission Zukunft der Informationsinfrastruktur

RfII: Rat für Informationsinfrastruktur

WR: Wissenschaftsrat

siehe auch <https://www.nfdi.de/verein/#historie>

Werbevideo zur NFDI



M-Create, Düsseldorf für DFG, CC BY 4.0, <https://youtu.be/x3Cvn1vNQ98>

1. Runde der NFDI-Konsortien 2019/20

Akronym	Name	weiterführende Informationen
DataPlant	Data in PLANT research	http://nfdi4plants.de/
GHGA	German Human Genome-Phenome Archive	https://ghga.dkfz.de/
KonsortSWD	Konsortium für die Sozial-, Bildungs-, Verhaltens-und Wirtschaftswissenschaften	https://www.konsortswd.de/
NFDI4BioDiversity	NFDI4BioDiversity: Biodiversity, Ecology & Environmental Data	https://www.nfdi4biodiversity.org/
NFDI4Cat	NFDI for Catalysis-Related Sciences	http://gecats.org/NFDI4Cat.html
NFDI4Chem	Fachkonsortium Chemie für die Nationale Forschungsdateninfrastruktur	https://www.nfdi4chem.de/
NFDI4Culture	Consortium for research data on material and immaterial cultural heritage	https://nfdi4culture.de/
NFDI4Health	National Research Data Infrastructure for Personal Health Data	https://www.nfdi4health.de/
NFDI4Ing	Nationale Forschungsdateninfrastruktur für die Ingenieurwissenschaften	https://nfdi4ing.de/

2. Runde der NFDI-Konsortien 2020/21

Akronym	Name	weiterführende Informationen
BERD@NFDI	Business, Economic and Related Data @ NFDI	https://www.berd-nfdi.de/
DAPHNE4NFDI	Data from Photon and Neutron Instruments for NFDI	https://www.sni-portal.de/de/daphne-nfdi
FAIRmat	FAIR Data Infrastructure for Condensed-Matter Physics and the Chemical Physics of Solids	https://www.fair-di.eu/fairmat/fairmat_consortium
MaRDI	Mathematical Research Data Initiative	https://www.mardi4nfdi.de/
NFDI4DataScience	National Research Data Infrastructure for Data Science	https://www.nfdi4datascience.de
NFDI4Earth	NFDI Consortium Earth System Science	http://www.nfdi4earth.de
NFDI4Microbiota	National Research Data Infrastructure for Microbiota Research	http://nfdi4microbiota.de
NFDI-MatWerk	National Research Data Infrastructure for Materials Science & Engineering	https://nfdi-matwerk.de/
PUNCH4NFDI	Particles, Universe, NuClei and Hadrons for the NFDI	https://www.punch4nfdi.de/
Text+	Text+: Language- and Text-Based Research Data Infrastructure	https://www.text-plus.org

3. Runde der NFDI-Konsortien 2021/22

Akronym	Name	weiterführende Informationen
BASE4NFDI	Basisdienste für die NFDI	https://base4nfdi.de
FAIRAgro	FAIR für Agrosystemforschung	https://fairagro.net
NFDI4BioImage	National Research Data Infrastructure for Bio-Imaging	https://nfdi4bioimage.de
NFDI4Energy	National Research Data Infrastructure for the Interdisciplinary Energy System Research	https://nfdi4energy.uol.de
NFDI4Immuno	National Research Data Infrastructure for Immunology	https://www.nfdi4immuno.de
NFDI4Memory	NFDI4Memory – The Consortium for the Historically Oriented Humanities	https://4memory.de
NFDI4Objects	Research Data Infrastructure for the Material Remains of Human History	https://www.nfdi4objects.net
NFDIxCS	National Research Data Infrastructure for and with Computer Science	https://nfdixcs.org

NFDI-Statistiken der DFG

DFG: Statistische Übersicht zu den Förderentscheidungen in beiden Ausschreibungsrunden, June 2021,

https://www.dfg.de/download/pdf/foerderung/programme/nfdi/nfdi_auswertung_2021.pdf.

DFG: NFDI – Statistische Übersichten zum Antragsingang, November 2019,

https://www.dfg.de/download/pdf/foerderung/programme/nfdi/nfdi_auswertung_2019.pdf.

DFG: Statistische Übersichten zum Antragsingang – Zweite Ausschreibungsrunde, September 2020,

https://www.dfg.de/download/pdf/foerderung/programme/nfdi/nfdi_auswertung_2020.pdf.

Beispiele für Dienste in NFDI-Konsortien

KonsortSWD: Data access
via data centres
(<https://www.konsortswd.de/en/konsortswd/tasks/data-access/>)

NFDI4Ing: Vocabularies and
Ontology Service
(<https://terminology.nfdi4ing.de/ts4ing/ontologies>)

NFDI4Culture: Software
Consulting and Tools
Software Consulting Agency
(<https://nfdi4culture.de/what-we-do/task-areas/task-area-3.html>)

Kurzrecherche

Frage:

Welches NFDI-Konsortium könnte für Ihren Bereich von Relevanz sein?
Wieso?

Methode:

Individuelle Online-Recherche über

https://www.dfg.de/foerderung/programme/nfdi/gefoerderte_konsortien/index.html

Dauer:

5 Minuten

European Open Science Cloud

European Open Science Cloud

*For the European Commission the EOSC “is a **trusted digital platform for the scientific community**, providing seamless access to data and interoperable services that address the whole research data cycle, from discovery and mining to storage, management, analysis and re-use across borders and scientific disciplines”.*

European Commission, 2019, European open science cloud: A new paradigm for innovation and technology. Publications Office, S. 2, <https://data.europa.eu/doi/10.2759/016783>.

European Open Science Cloud

*The central goal of the EOSC is nothing less than **global EU leadership in research data management**. This goes hand in hand to ensure that European researchers have access to all the benefits of data-driven research.*

Achim Streit und Jos van Wezel (2021): Deutschland in der European Open Science Cloud, in: M. Putnings, H. Neuroth, & J. Neumann (Hrsg.), Praxishandbuch Forschungsdatenmanagement, S. 32, <https://doi.org/10.1515/9783110657807-003>.

European Open Science Cloud

EOSC is not a cloud “made in Brussels”, it should help Europe to build a “Schengen Area for digital Data”.

Budroni, P., Claude-Burgelman, J., & Schouppe, M. (2019): Architectures of Knowledge: The European Open Science Cloud, in: ABI Technik 39(2), S.140, <https://doi.org/10.1515/abitech-2019-2006>.

European Open Science Cloud

*From a more technical perspective EOSC is an **integration layer**. It aims at establishing interoperability among the existing and forthcoming European Research Infrastructures. So, by federating existing research data infrastructures, the EOSC leverages national investments and adds value in terms of scale, interdisciplinary and faster innovation.*

European Commission, <https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud>.

European Open Science Cloud

Sie ist:

- nicht rein europäisch, da auch außereuropäische Partner teilnehmen
- nicht nur offen, da es auch Dienste mit Zugangsbeschränkungen gibt
- nicht nur wissenschaftlich sondern beinhaltet auch Datenservice für z.B. die Industrie
- keine Cloud, sondern eher ein föderiertes Infrastruktursystem

EOSC Catalogue and Marketplace

- <https://marketplace.eosc-portal.eu>
- Integrierte Plattform zur Übersicht
- einfacher Zugang zu zahlreichen Diensten für verschiedene Forschungsbereiche sowie zu Daten
- Suche nach wissenschaftlichem Bereich, aktueller Forschungstätigkeit oder Anbieter

GOSC & GAIA-X

Global Open Science Cloud

“The Global Open Science Cloud (GOSC) initiative will encourage cooperation, alignment, and ultimately interoperability, between existing and emerging Open Science Clouds”

CODATA, S. 3, <https://codata.org/wp-content/uploads/2021/06/GOSC-Introduction-Event-Objectives-and-Agenda.pdf>.

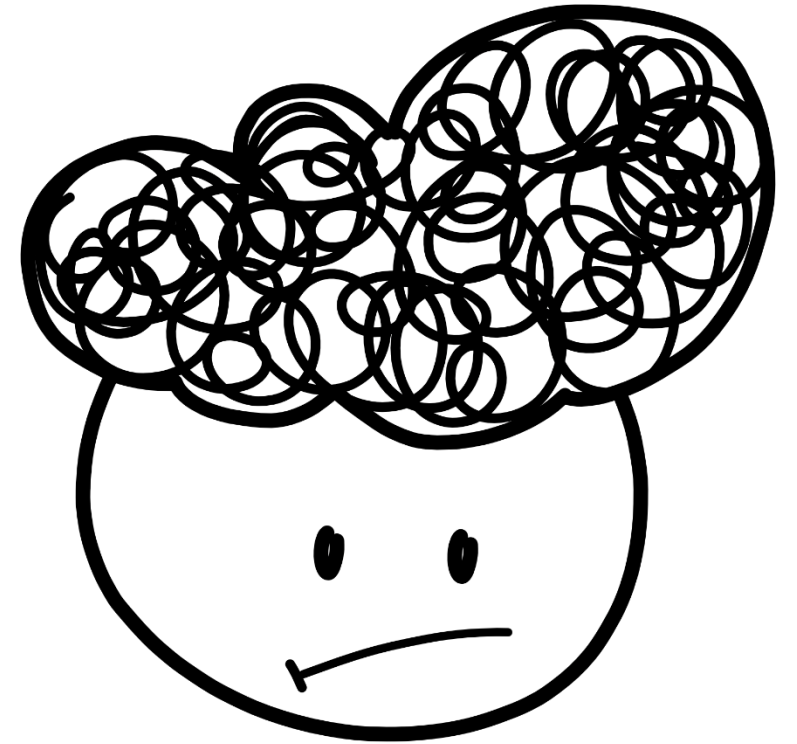
GAIA-X

- [Deutsch-Französische Initiative 2020 für eine Data Cloud, mit besonderem Schwerpunkt auf die Wirtschaft](#)
- Erklärtes Ziel von GAIA-X ist es, die digitale Souveränität von Wirtschaft, Wissenschaft, Regierung und Gesellschaft zu stärken, indem die Entwicklung von Innovationsökosystemen gefördert wird.
- Außerdem geht es bei GAIA-X auch um große Politik und Marketing: Es geht um europäische Werte, um Datenschutz und vor allem um Unabhängigkeit.

And More

- [China Science and Technology Cloud \(CSTCloud\)](#)
- [African Open Science Platform](#)
- [LA Referencia](#) for Latin America
- ...

Offene Fragen Ihrerseits?



Literaturempfehlung

- Zu NFDI: Kammerer, Dietmar, Schrade Torsten, und Stellmacher Martha (2021): NFDI4Culture: Konsortium für Forschungsdaten zu materiellen und immateriellen Kulturgütern, in: Bausteine Forschungsdatenmanagement, Nr. 2 (Juli), S. 23-33, <https://bausteine-fdm.de/article/view/8341>.
- Zu EOSC: Achim Streit, und Jos van Wezel (2021) Deutschland in der European Open Science Cloud, in: Praxishandbuch Forschungsdatenmanagement, herausgegeben von Markus Putnings, Heike Neuroth, und Janna Neumann, S. 31-52, <https://doi.org/10.1515/9783110657807-003>.
- Zu GAIA-X: Mahn, Jan (2020): Gaia-X: Wie Europa in der Cloud unabhängig werden soll, c't Magazin, 25. Juni 2020. <https://www.heise.de/ct/artikel/Gaia-X-Wie-Europa-in-der-Cloud-unabhaengig-werden-soll-4783596.html>.

Kurzes Audio-Quiz

Was versteckt sich hinter dieser Audio-Datei eines Forschungsdatensatzes?



Auflösung: Cecilia Durojaye, 2020, „Recodings – Nigerian Talking drum“, file D1s,Edmond, <https://doi.org/10.17617/3.IUMMLP>, CC BY 4.0.

Teil 4

Weiterführende Überlegungen

Inhalt von Teil 4

1. Beispiele für Datenrepositorien
2. Open (Research) Data
3. Exkurs: Forschungsdaten in den Geisteswissenschaften
4. FDM-Beratung
5. Open Science
6. Wissenschaftliche Software
7. Fortbildungsmöglichkeiten

Beispiele für Datenrepositorien

Einteilung von Forschungsdatenrepositorien

1. Fachlich

- [PsychData](#)
- [Pangaea](#)
- [G-Node](#)

2. Institutionell

- [Open Data LMU](#)
- [Edinburgh DataShare](#)
- [Edmond](#)

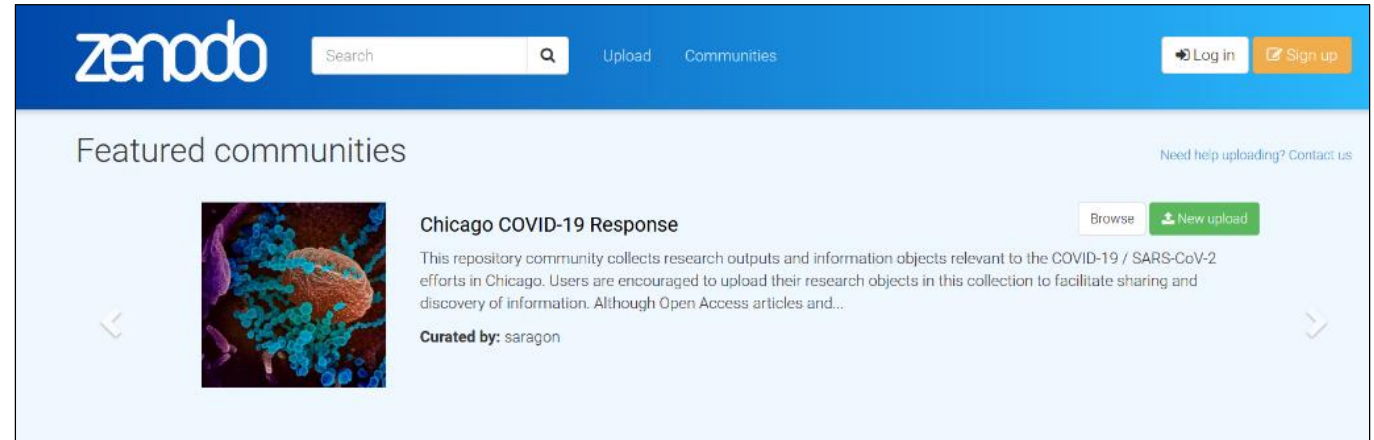
3. Allgemein

- [Zenodo](#)
- [Dryad](#)
- [Radar](#)
- [Figshare](#)
- [Mendeley Data](#)

Repositories: Example 1

Zenodo

- <https://zenodo.org>
- vom CERN
- generisches Repositorium



CERN Data Centre & Invenio, CC BY 4.0, <https://zenodo.org>.

Repositories: Example 2

Discuss.Data

- <https://discuss-data.net>
- Repository für die post-sowjetische Region
- betrieben durch die Universität Bremen und die SUB Göttingen
- fachspezifisches Repository



Maidan Oral History Archive

I. Title of the Data Collection

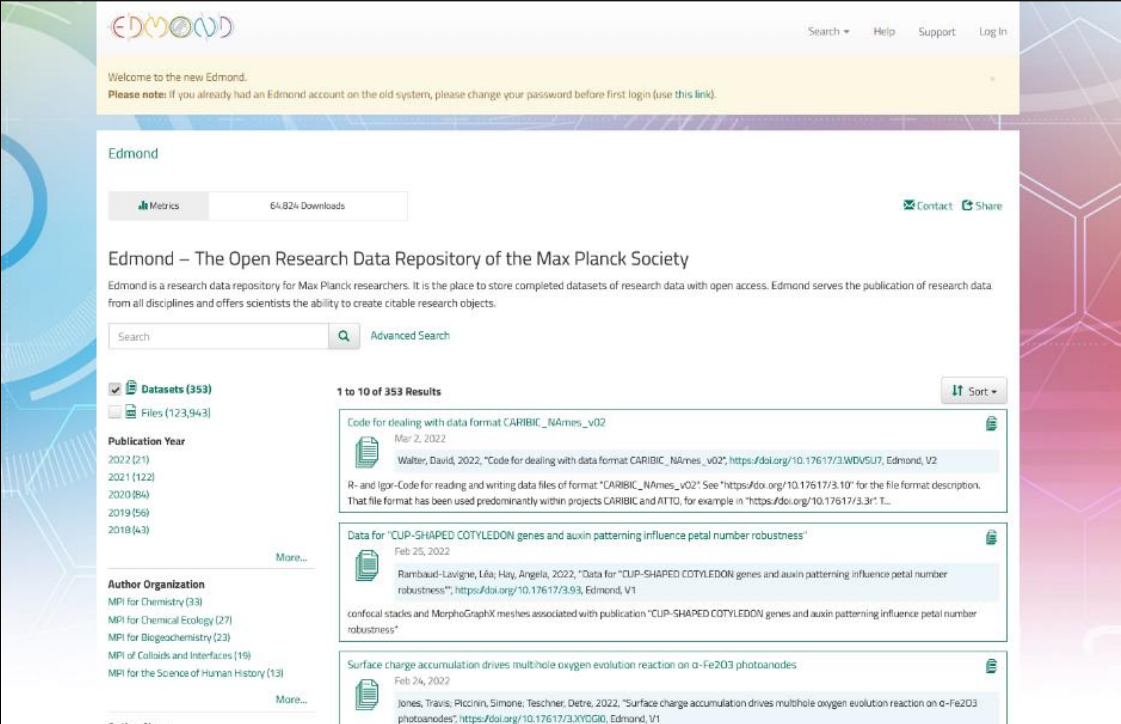
Maidan Oral History Archive

The Foundation to Preserve the History of Maidan (FPHM) (2021): Maidan Oral History Archive, v. 1.0, Discuss Data, <https://doi.org/10.48320/F5509D39-B2D4-409D-AD37-9E72570206AD>

Repositories: Example 3

Edmond

- <https://edmond.mpdl.mpg.de>
- durch die Max Planck Digital Library für die Max-Planck-Gesellschaft
- Institutionelles Repositorium

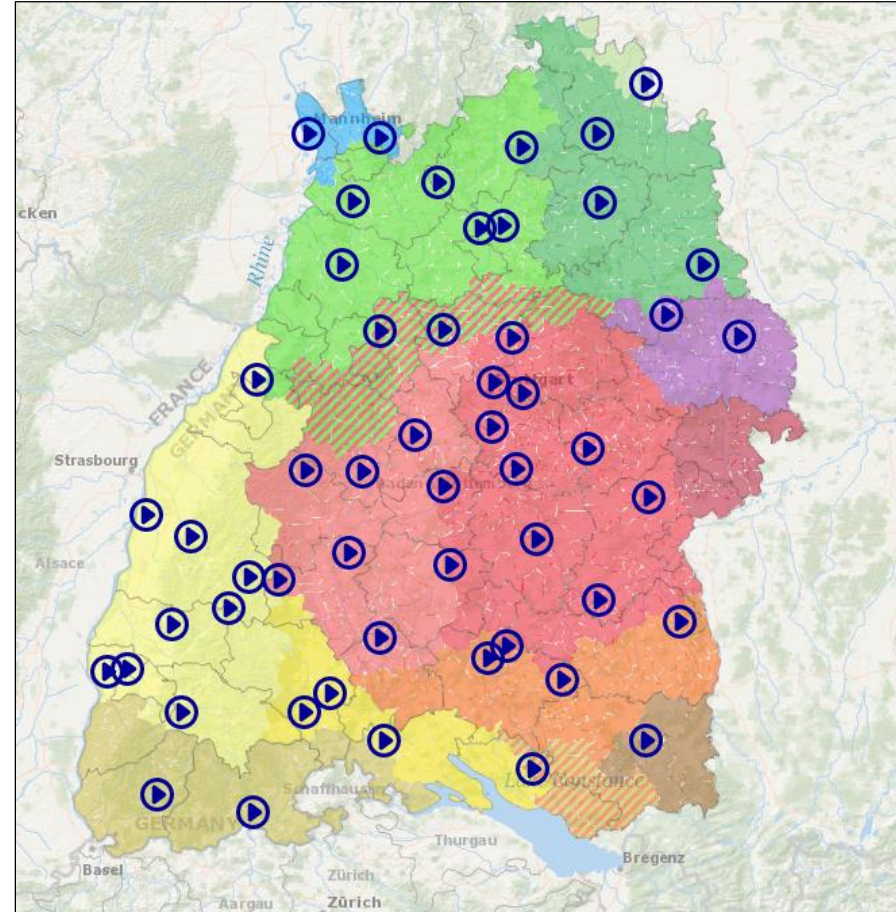


The screenshot displays the Edmond website interface. At the top, there is a navigation bar with the Edmond logo, a search bar, and links for Help, Support, and Log In. Below the navigation bar, a yellow banner welcomes users to the new Edmond system and provides a note for existing users. The main content area features a search bar and a "Metrics" section showing 64,824 Downloads. The page title is "Edmond – The Open Research Data Repository of the Max Planck Society". Below the title, there is a brief description of Edmond and a search bar. The search results section shows 1 to 10 of 353 results. The first result is a dataset titled "Code for dealing with data format CARIBIC_NArmes_v02" by David Walter, dated Mar 2, 2022. The second result is "Data for 'CUP-SHAPED COTYLEDON genes and auxin patterning influence petal number robustness'" by Rambaud-Lavigne, Léa; Hai, Angela, dated Feb 25, 2022. The third result is "Surface charge accumulation drives multihole oxygen evolution reaction on α -Fe₂O₃ photoanodes" by Jones, Travis; Piccini, Simone; Teschner, Detre, dated Feb 24, 2022. The left sidebar contains filters for "Publication Year" (2022 (2), 2021 (122), 2020 (84), 2019 (56), 2018 (43)) and "Author Organization" (MPI for Chemistry (33), MPI for Chemical Ecology (27), MPI for Biogeochemistry (23), MPI of Colloids and Interfaces (19), MPI for the Science of Human History (13)).

Repositories: Example 4

Sprachatlas Baden-Württemberg

- <https://escience-center.uni-tuebingen.de/escience/sprachatlas/#8/48.402/7.498>
- Eberhard Karls Universität Tübingen
- Audio-Daten mit Karte zu gesprochenen Sprachen in Baden-Württemberg



Nutzung von Data Repositories

Bearbeiten Sie eine der folgenden Fragen:

1. Wieviel Publikationen der Hochschule für den öffentlichen Dienst in Bayern gibt es auf Zenodo (<https://zenodo.org>)?
2. Wieviel Publikationen befinden sich auf Pangaea (<https://www.pangaea.de>) mit Bezug zu den Bayrischen Alpen? Wo fanden deren Messungen statt?
3. Welcher Datensatz auf Open Data LMU (<https://data.ub.uni-muenchen.de>) gibt Informationen zu historischen Gaststätten in München? Welche Informationsquellen wurden für diesen Datensatz zusammengeführt?

Open Research Data

Open Research Data

Vier Schritte zu offenen Daten von der Europäischen Kommission:

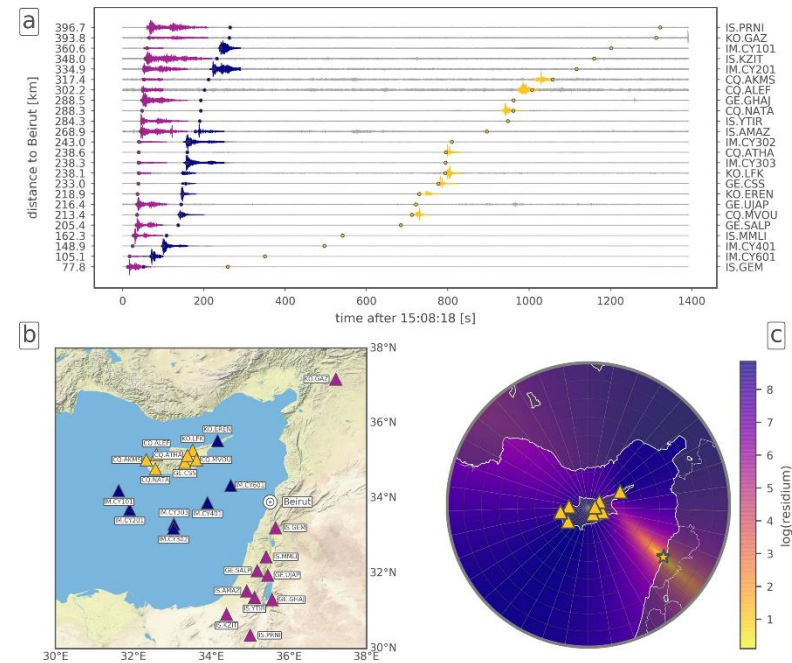
1. Bereiten Sie Ihre Daten für die gemeinsame Nutzung vor
2. Wählen Sie ein Repository
3. Fügen Sie eine Datenverfügbarkeitserklärung zu Ihrem Artikel hinzu
4. Verknüpfen Sie Ihre Datensätze mit Ihrem Artikel

https://think.f1000research.com/open-research-europe-submission/opendataguide/?utm_source=CPB&utm_medium=cms&utm_campaign=JQC19229

Beispiel für Open Data

Explosion im Hafen von Beirut

- Kombination von verschiedenen offenen Datenquellen zur Rekonstruktion der Explosionen
- Pilger, C., Gaebler, P., Hupe, P. et al. (2021): Yield estimation of the 2020 Beirut explosion using open access waveform and remote sensing data, Scientific Reports 11, 14144, <https://doi.org/10.1038/s41598-021-93690-y>.



Pilger, C., Gaebler, P., Hupe, P. et al., 2021, CC BY 4.0, <https://www.nature.com/articles/s41598-021-93690-y/figures/1>

Spezifikationen für Open Data

- Open Data ↔ Open Research Data
- Open Data ↔ FAIR Data

- Open Data Portal München
- [Kompetenzzentrum für Open Data des Bundes](#)
- [Open Data Portal der EU](#)
- ...

Nutzung von Open Data

Bearbeiten Sie eine der folgenden Fragen:

1. Was war der durchschnittliche Bierpreise auf den Oktoberfesten in den Jahren 1985-2019?
Lösen Sie mit dem Datensatz im Open Data Portal der Stadt München (<https://s.gwdg.de/mkqiye> bzw. Identifizier)
2. Zu welcher Uhrzeit haben Forscher_innen in ihrer Beobachtung die Ringeltaube (engl. „wood pigeon“) durchschnittlich beobachtet?
Nutzen Sie den Datensatz <https://s.gwdg.de/ruQteP> bzw. <https://doi.org/10.17617/3.EGLBLP> in Edmond, dem Open Data Repository der Max-Planck-Gesellschaft.
3. Wieviel Munition aus beiden Weltkriegen wurde zwischen 2015-19 in der Kieler Förde gemeldet?
Beziehen sie sich hierbei auf den Datensatz der Zentralen Meldestelle für Munition im Meer der Küstenländer im Open Data Portal der EU-Kommission (<https://s.gwdg.de/iawr0Q> bzw. Identifizier)

Diskussion: Entwicklungen Rund um Datenökosystemen

- Open Sharing vs. Quasi-Kommerzialisierung von Datenrepositorien
 - Existiert eine solche Dualität eigentlich wirklich?
- Lässt sich mit Forschungsdaten Gewinn erwirtschaften?
- Begrenzung von Wissenschaft

- Politische Perspektive: Sind offene Daten in Konkurrenzsituationen gegenüber Russland und China ein Nachteil?

Leseempfehlung: Putnings (2021): Datenökosystem, In: Praxishandbuch Forschungsdatenmanagement, hrsg. Von Putnings, Neuroth und Neumann, De Gruyter Saur, S. 7–10, <https://doi.org/10.1515/9783110657807-001>.

Exkurs* Forschungsdaten in den Geisteswissenschaften

* auf Nachfrage

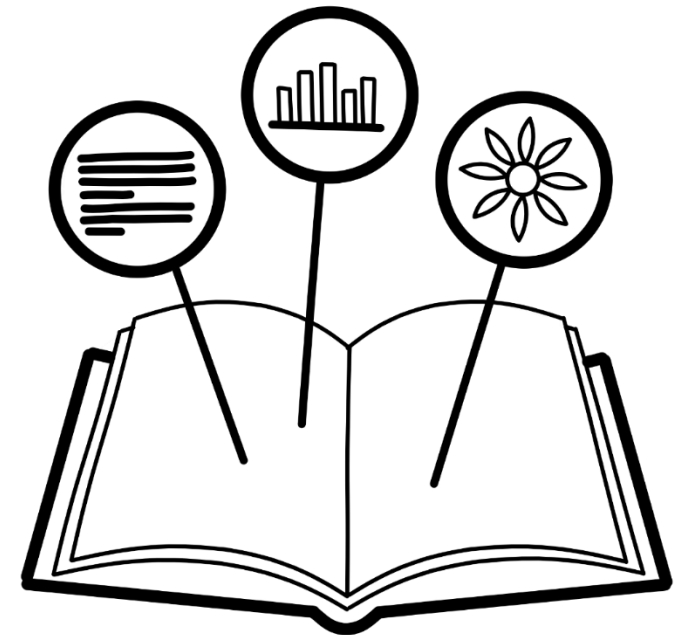
Problemlagen

- *„Was beispielsweise für Natur- oder Sozialwissenschaftler*innen zum täglichen Sprachgebrauch gehört, ist bei Geisteswissenschaftler*innen eher fremd, denn Bilder, Texte, Aufzeichnungen, werden, auch wenn sie digital vorhanden sind, kaum als Daten wahrgenommen.“*
- *„Bis vor wenigen Jahren stand vor allem in den Geisteswissenschaften am Ende eines Forschungsprozesses eine Publikation, meist in Form einer Monographie oder als Beitrag in einem Sammelband, bzw. einer wissenschaftlichen Zeitschrift. Der Endbericht für den Fördergeber belegte zwar den Prozess dorthin, beschrieb die Methoden und die Quellen, aber nach den zugrundeliegenden Materialien fragte nach Projektende im Regelfall niemand mehr. Aufzeichnungen, Kopien, Scans, Bilder, Transkripte, Audio- oder Videoaufnahmen wurden entweder entsorgt oder in privaten oder dienstlichen Arbeitszimmern aufbewahrt“.*

Susanne Blumesberger (2021): Forschungsdaten in den Geisteswissenschaften: Bereits selbstverständlich oder doch noch etwas exotisch?. O-Bib – Das Offene Bibliotheksjournal 8(4), S. 1–8. <https://doi.org/10.5282/o-bib/5739>.

Kontext Digital Humanities

- Nutzung von digitalen Werkzeugen, z.B. Computer, ist mittlerweile gängige Praxis
- Humanities mit digitalen Methoden oder eigene Disziplin „Digital Humanities“?
- Curricula und Weiterbildungsangebote adressieren den Aufbau digitaler Kompetenzen
- Digitaler Kulturwechsel im Fach? Oder doch nur quantitative Beschleunigung ohne qualitativer Entwicklung?



<https://doi.org/10.5281/zenodo.5608845>

Forschungsdaten veröffentlichen

- Forschungsdaten in den Geisteswissenschaften publizieren ist (aktuell) nur in Teilbereich gängige Praxis
- Beispiele:
 - Editorik bei der Digitalisierung von historischen Briefkorrespondenzen an der BBAW
 - Diplomantik mit openMGH (<https://www.mgh.de/de/mgh-digital/openmgh>) an der BSB München

Jean-Paul-Edition der BBAW (<https://www.jean-paul-edition.de>), CC BY SA 4.0.

Forschungsdaten veröffentlichen

Beispiel OstData

- Angebot der BSB München mit Kooperationspartner
- Spezialisiert auf Forschungsdaten für die Ost-, Ostmittel- und Südosteuropaforschung
- <https://www.osmikon.de/servicemenue/ueber-uns/ueber-ostdata>
- Ziele:
 - Schaffung von Möglichkeiten für Publikation und Archivierung von Forschungsdaten
 - Suchmöglichkeit zu Forschungsdaten
 - Nachnutzbarkeit von Forschungsdaten

Services:

- Beratung
- Datenrepositorium (Betastatus): <http://www.ostdata.de>
- Checkliste Forschungsdaten publizieren: <https://zenodo.org/record/6951564>
- Überlegungen und Vorschläge hinsichtlich der Nutzung von Review-Verfahren: <https://doi.org/10.5281/zenodo.7248022>
- ...

Aktuelle Entwicklungen in der NFDI

Frage: Welche NFDI-Konsortien können den Geisteswissenschaften zugeordnet werden?

- KonsortSWD (Bildungs und Sozialwissenschaften)
- NFDI4Culture (Kunst- und Kulturwissenschaften)
- NFDI4Memory (Geschichtswissenschaften)
- NFDI4Objects (materielle Hinterlassenschaften der Menschheitsgeschichte)
- Text+ (Sprachwissenschaften)

Literaturempfehlung

- Blumesberger (2021): Forschungsdaten in den Geisteswissenschaften: Bereits selbstverständlich oder doch noch etwas exotisch?. O-Bib – Das Offene Bibliotheksjournal 8(4), S. 1–8. <https://doi.org/10.5282/o-bib/5739>.
- Cremer, Klaffki und Steyer (2018): Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften, In: O-Bib – Das Offene Bibliotheksjournal 5(2), S. 142–162. <https://doi.org/10.5282/o-bib/2018H2S142-162>.
- Ibid. (2019): Redaktionssache Forschungsdaten – Ein Servicekonzept zur Forschungsdatenpublikation in den Geisteswissenschaften, In: Zeitschrift Bibliothek Forschung und Praxis, <https://doi.org/10.1515/bfp-2019-2018>.
- Feidicker (2022): Ist die Geschichtswissenschaft eine datenverarbeitende Disziplin? Gedanken nach dem Datathon, <https://vergleichen.hypotheses.org/872>.
- Lehmkuhl und Lemaire (2021): Data Literacy Kompetenzprofil für historisch arbeitende Geisteswissenschaftler*innen: Welche Kompetenzen benötigen wir? <https://doi.org/10.5281/zenodo.5379679>.

FDM-Beratung

Forschungsnahe Dienste durch Bibliotheken

Positionspapier der Kommission für forschungsnahe Dienste des VDB:

- Forschungsdatenmanagement (FDM)
- Publikationsdienste: Open Access, Repositorien, Academic Networking
- Publikationsdatenmanagement: Affiliation und Autor*innen-Identität
- Bibliometrien, Evaluierung und Leistungsmessung
- Systematic Reviews: Recherche bei komplexen Forschungsprojekten
- Digitale Integration von (Forschungs-) Infrastrukturen

Stille et al. (2021): Forschungsunterstützung an Bibliotheken: Positionspapier der Kommission für forschungsnahe Dienste des VDB, In: o-bib. Das offene Bibliotheksjournal 8(2), <https://doi.org/10.5282/o-bib/5718>, S. 3-8.

Empfehlungen

- „Support is user experience“
- „Problem first, solution second“
- Zuhören, verstehen, Fragen stellen
- Ergebnisse oder Aufgaben notieren
- Schnelle Lösungen sind nicht immer die Besten
- ggf. auch andere Kolleg_innen und/oder Expert_innen hinzuziehen

Spiel zu Themen in Beratungsgesprächen

Research Data ScaryTales (<https://doi.org/10.5281/zenodo.7276802> oder <https://forschungsdaten-thueringen.de/fdm-scarytales/articles/ueberblick.html>)

Spielanleitung:

1. Eine Person wird als Spielleitung festgelegt. Sie bestimmt, welche der Geschichten gewählt wird und führt die Teilnehmer*innen durch das Spiel.
2. Die Spielleitung öffnet die Lösung und liest diese.
3. Die Mitspielenden stellen nun reihum Ja/Nein-Fragen, um herauszufinden, was passiert ist.
4. Die Spielleitung antwortet mit "Ja" oder "Nein".
5. Die Runde ist beendet, sobald erraten wurde, was passiert ist.
6. Zum Schluss werden die Lösung und das FDM-Fazit vorgelesen.

Dauer:

15min

Ziel:

Präsentation von konkreten FDM-Beispielen und deren Lösungen.

Open Science

Aspekte von Open Science

- **Open Access:** Öffentlicher Zugang zu Forschungsergebnissen in Form von Veröffentlichungen
- **Open Research Data:** Freie Verfügbarkeit von Forschungsdaten/Rohdaten
- **Linked Open Data:** Zugänglichkeit und Vernetzung von öffentlichen Repositorien
- **Open Peer Review:** Suche nach Alternativen zu den traditionellen Begutachtungsverfahren der Zeitschriftenverlage, um eine größere Transparenz in diesem Bereich zu gewährleisten

Aspekte von Open Science

- **Open Source:** Entwicklung und Nutzung von Open-Source-Software für die Wissenschaft
- **Citizen Science:** Einbindung von Nicht-Wissenschaftlern in den wissenschaftlichen Prozess
- **Open Educational Resources:** Kostenloser Zugang zu digitalem Lehrmaterial und (Aufzeichnungen von) Kursen
- **Pre-registration:** Hinterlegung und Begutachtung eines Projekt- und Methodenplans mit Hypothesen, die auf einer geeigneten Plattform getestet werden sollen

Fragen via Slido

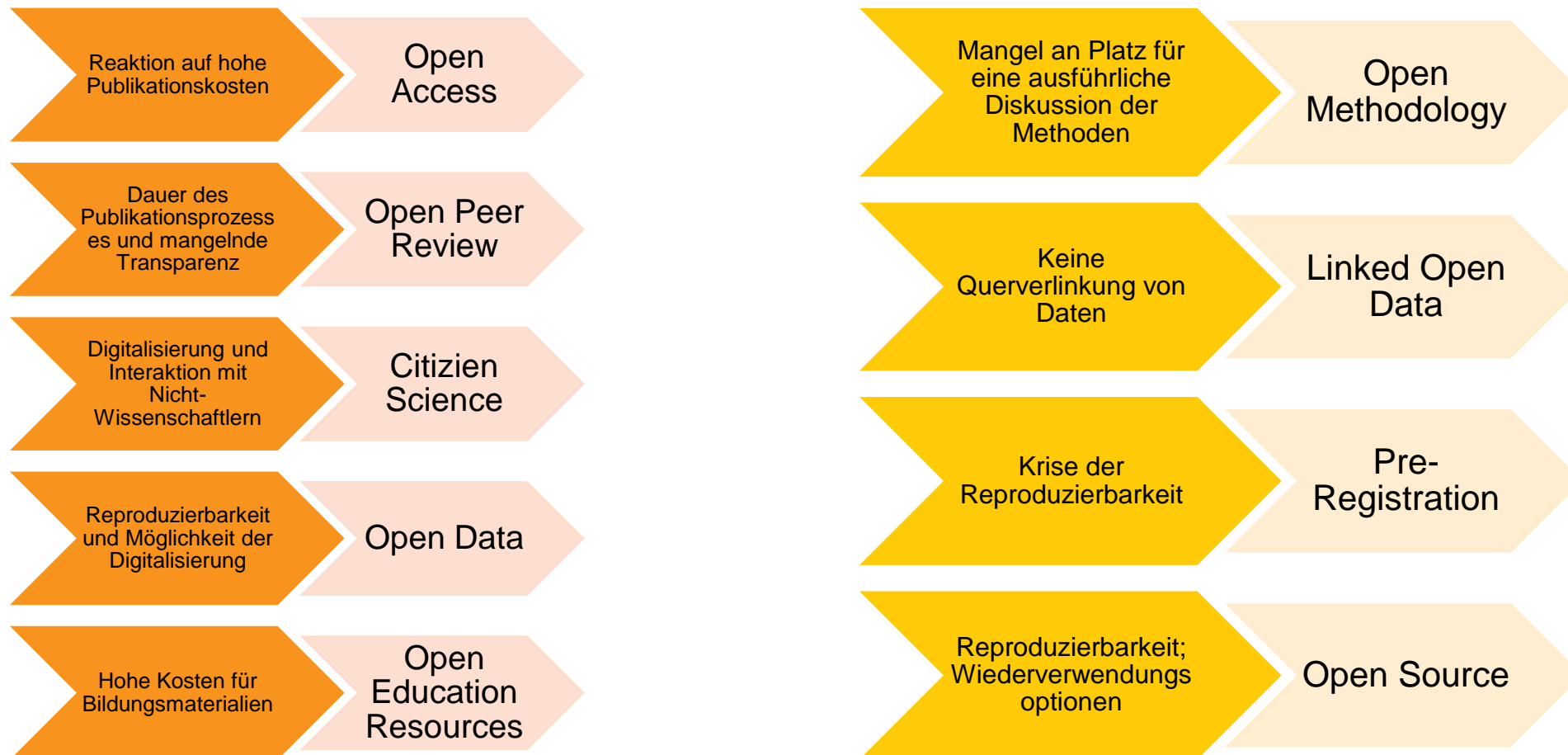
Fragen

Von welchen Open Science-Aspekten hatten Sie im Vorfeld bereits gehört?

Vorgehen

Im Browser www.slido.com öffnen. Code 5029112 eingeben.

Hypothesis: OS als Reaktion auf ...



Open Science Communities

Example The Netherlands

- <https://www.osc-nl.com>
- <https://www.openscience.nl>



<https://www.osc-nl.com>, CC BY 4.0

Openness in Libraries

“The Commission for Research-Related Services sees the promotion of openness as a central task of libraries. In this sense, it is strongly committed to the establishment and further development of innovative library services, and committed to the principles of openness.”

Openness in Libraries
Position Paper of the Commission for Research-Related Services of the VDB

Alexander Berg-Weiß
University Library of the Ludwig-Maximilians-Universität München
<https://orcid.org/0000-0002-7435-8676>


Sibylle Hermann
Information and Communication Center of the University of Stuttgart
<https://orcid.org/0000-0001-9239-8789>

Miriam Kötter
University of Duisburg-Essen, University Library
<https://orcid.org/0000-0002-7253-2933>

Caroline Leiß
Universitätsbibliothek der TUM
<https://orcid.org/0000-0002-2792-2625>

Christoph Müller
Library of the Ibero-Amerikanisches Institut Stiftung Preußischer Kulturbesitz
<https://orcid.org/0000-0001-9399-6676>

Annette Strauch-Davey
University Library of Hildesheim
<https://orcid.org/0000-0003-2355-325X>



PDF

HTML


Berg-Weiß, A., Hermann, S., Kötter, M., Leiß, C., Müller, C., & Strauch-Davey, A. (2022): Openness in Libraries: Position Paper of the Commission for Research-Related Services of the VDB, In: o-bib. Das offene Bibliotheksjournal 9(2), S. 1–4. <https://doi.org/10.5282/o-bib/5834>, CC BY 4.0

LIBER Strategy 2023-2027

“D. ADVANCING OPEN SCIENCE:
By 2027, in collaboration with
researchers, research libraries
stimulate, facilitate, co-develop, and
manage infrastructures and
practices designed to take Open
Science to the next level.”

LIBER Strategy 2023-2027, https://libereurope.eu/wp-content/uploads/2022/01/LIBER_STRAT_A5_digital-final-1.pdf, S. 3

Strategy is a
style of thinking,
a conscious and
deliberate process,
an intensive
implementation
system, the
science of
insuring
future success.



Pete Johnson
Strategy
2023 - 2027

LIBER 2023-2027
Strategy, CC BY-NC-ND,
https://libereurope.eu/wp-content/uploads/2022/01/LIBER_STRAT_A5_digital-final-1.pdf.

Offene Fragen zu Open Science?

Brainwriting Pool

Frage:

Wie wird Open Science die deutschen Bibliotheken zukünftig prägen?

Methode:

Gemeinsames Schreiben an drei Textabschnitte

Einteilung:

Reihe 1 übernimmt Teil A, Reihe 2 B und Reihe 3 + 4 C

Werkzeug:

<https://s.gwdg.de/VTB0Ny> in Markdown

Dauer:

15 Minute (alle 5 Minuten Wechsel A → B → C)

Ziel:

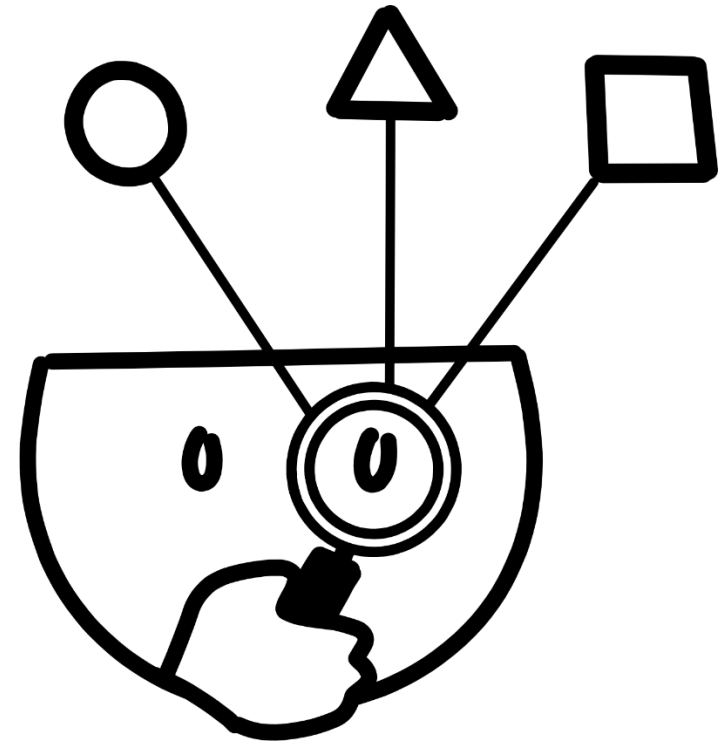
Kreative Annäherung an das Thema, gemeinsamer Text

Und was schreibt ChatGPT dazu?

siehe

<https://pad.gwdg.de/s/fWwIEK5>

[zJ](#)



<https://doi.org/10.5281/zenodo.3674561>

Wissenschaftliche Software

Three Pillars of Research Output?!

Text

- Journal Article
- Book
- Poster Presentation
- ...

Data

- Research Data
- Laboratory Data
- ...

Code

- Software
- Software Documentation
- ...

Second French Plan for Open Science

Path Three “Opening up and promoting source code produced by research”

French Ministry of Higher Education,
Second French Plan for Open
Science. Generalizing open science
in France 2021-2024, July 2021,

https://www.ouvrirlascience.fr/wp-content/uploads/2021/10/Second_French_Plan-for-Open-Science_web.pdf

pp. 1 and 16-19, CC BY 4.0.



Wissenschaftliche Software

Begrifflichkeit:

- Eigenentwickelte Forschungssoftware
- Softwareanwendungen für die Forschung
- Infrastruktursoftware/-dienste

Wissenschaftliche Software

Gemeinsamkeiten von Software und Forschungsdaten im **Forschungsprozess:**

- beides spielt eine immer bedeutendere Rolle in der Forschung
- beides ist notwendig, um Forschungsergebnisse reproduzierbar zu machen
- beides stellt potentiell einen Wert für Nachnutzer dar
- beides wird bisher nicht (oder nur in geringem Maße) als eigenständige Forschungsleistung anerkannt

Wissenschaftliche Software

Gemeinsamkeiten von Software und Forschungsdaten im **Management**:

- beides kann mit Metadaten versehen und archiviert werden
- beides muss kuratiert werden, um die Nutzbarkeit über längere Zeiträume zu gewährleisten
- bei beidem fehlt oft das notwendige Wissen für ein adäquates Management

Wissenschaftliche Software

Unterschiede von Software zu Forschungsdaten:

- die Kuratierung von Software ist wesentlich aufwändiger
- [Metadaten für Software](#) sind wesentlich homogener als die für Forschungsdaten
- bei Software gibt es viel Vorerfahrung mit offenen/kommerziellen [Lizenzen](#)
- eine genaue Versionierung spielt bei Software eine wesentlich größere Rolle

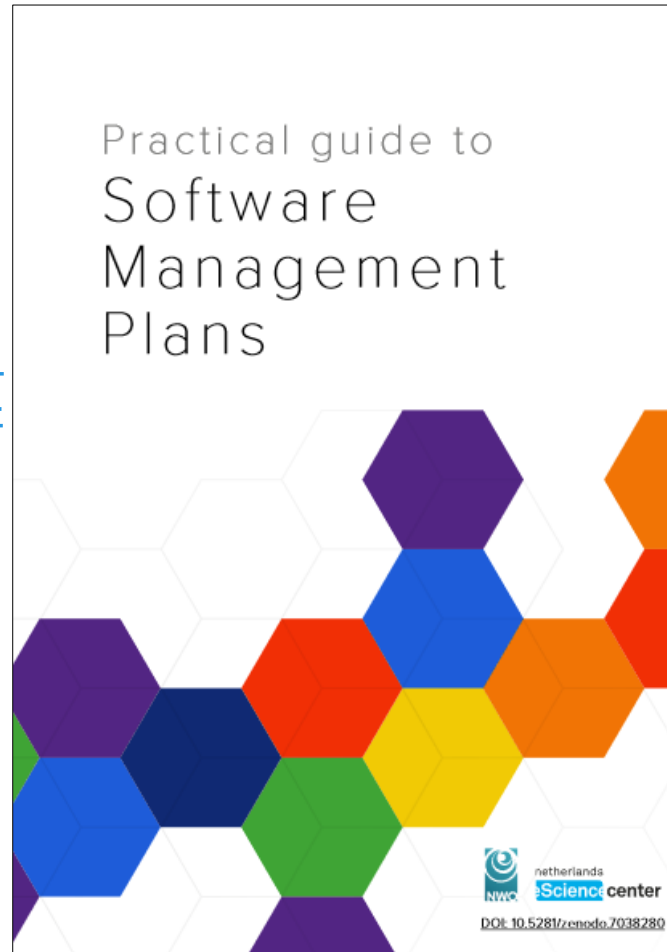
Wissenschaftler_innen schreiben wissenschaftliche Software

Beobachtungen:

- häufig keine ausgebildeten Software-Entwickler_innen, eher Autodidakten
- Funktionalität geht vor Dokumentation geht vor langfristiger Erhaltung
- erst Textpublikation, dann lange nichts, und dann vielleicht Datenpublikation und Software-Publikation
- Software wird von Generation zu Generation „vererbt“
- ...

SMP aktuell in Diskussion

Martinez-Ortiz et al. (27.10.2022):
Practical guide to
Software
Management
Plans, v1.0,
<https://doi.org/10.5281/zenodo.7248877>,
CC BY 4.0.




DFG: **Call for Proposals to Increase the Usability of Existing Research Software** durch „Research Software – Quality Assured and Re-usable“, (Information für die Wissenschaft Nr. 85 | 8. November 2022)
http://www.dfg.de/en/research_funding/announcements_proposals/2022/info_wissenschaft_22_85.

FAIR Principles for Research Software (FAIR4RS Principles)

- Adaption der FAIR Data Principles für Forschungssoftware
- von der Community unterstützte Grundsätze

FAIR Principles for Research Software
(FAIR4RS Principles)



RDA Recommendation

DOI: 10.15497/RDA00068

Authors: Neil P. Chue Hong*, Daniel S. Katz*, Michelle Barker*, Anna-Lena Lamprecht, Carlos Martinez, Fotis E. Psomopoulos, Jen Harrow, Leyla Jael Castro, Morane Gruenpeter, Paula Andrea Martinez, Tom Honeyman, Alexander Struck, Allen Lea, Axel Loewe, Ben van Werkhoven, Catherine Jones, Daniel Garijo, Esther Plomp, Françoise Genova, Hugh Shanahan, Joanna Leng, Maggie Hellström, Malin Sandström, Manodeep Sinha, Mateusz Kuzak, Patricia Herterich, Qian Zhang, Sharif Islam, Susanna-Assunta Sansone, Tom Pollard, Udayanto Dwi Atmojo, Alan Williams, Andreas Czerniak, Anna Niehues, Anne Claire Fouilloux, Bala Desinghu, Carole Goble, Céline Richard, Charles Gray, Chris Erdmann, Daniel Nüst, Daniele Tartarini, Elena Rangelova, Hartwig Anzt, Ilian Todorov, James McNally, Javier Moidon, Jessica Burnett, Julián Garrido-Sánchez, Khalid Belhajjame, Laurents Sesink, Lorraine Hwang, Marcos Roberto Tovani-Palona, Mark D. Wilkinson, Mathieu Servillat, Matthias Liffers, Merc Fox, Nadica Miljković, Nick Lynch, Paula Martinez Lavanchy, Sandra Gesing, Sarah Stevens, Sergio Martinez Cuesta, Silvio Feroni, Stian Soiland-Reyes, Tom Bakker, Tovo Rabemanantsoa, Vanessa Sochat, Yo Yehudi, [FAIR4RS WG](#)

(*) lead authors with equal contributions

Published: 24th May 2022

Version: 1.0

Abstract: To improve the sharing and reuse of research software, the FAIR for Research Software (FAIR4RS) Working Group has applied the FAIR Guiding Principles for scientific data management and stewardship to research software, bringing together existing and new community efforts. Many of the FAIR Guiding Principles can be directly applied to research software by treating software and data as similar digital research objects. However, specific characteristics of software — such as its executability, composite nature, and continuous evolution and versioning — make it necessary to revise and extend the principles.

This document presents the first version of the FAIR Principles for Research Software (FAIR4RS Principles), and includes explanatory text to aid adoption. It is an outcome of the FAIR for Research Software Working Group (FAIR4RS WG) based on community consultations that started in 2019.

The FAIR for Research Software Working Group was jointly convened as a Research Data Alliance (RDA) Working Group, FORCE11 Working Group, and Research Software Alliance (ReSA) Task Force.

Going forward, the RDA Software Source Code Interest Group is the maintenance home for the principles. Concerns or queries about the principles can be raised at RDA plenary events organized by the SSC IG, where there may be opportunities for adopters to report back on progress. The full maintenance and retirement plan for the principles can be found on the RDA website.

Chue Hong et al.
(2022): FAIR
Principles for
Research Software
(FAIR4RS Principles,
Zenodo, CC BY 4.0,
<https://doi.org/10.15497/RDA00068>).

Unterschiede FAIR Data ↔ FAIR4RS

Findable: Software, and its associated metadata, is easy for both humans and machines to find

- (=) F1 Software is assigned a globally unique and persistent identifier
 - (new) F1.1 Components of the software representing levels of granularity are assigned distinct identifiers
 - (new) F1.2 Different versions of the software are assigned distinct identifiers
- (=) F2 Software is described with rich metadata
- (=) F3 Metadata clearly and explicitly include the identifier of the software they describe
- (=) F4 Metadata are FAIR, searchable and indexable

Accessible: Software, and its metadata, is retrievable via standardised protocols

- (=) A1 Software is retrievable by its identifier using a standardised communications protocol
 - (=) A1.1 The protocol is open, free, and universally implementable
 - (=) A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- (=) A2 metadata are accessible, even when the software is no longer available

Interoperable: Software interoperates with other software by exchanging data and/or metadata, and/or through interaction via application programming interfaces (APIs), described through standards

- (≠) I1. Software reads, writes and exchanges data in a way that meets domain-relevant community standards
- ~~I2.~~ (meta)data use vocabularies that follow FAIR principles
- (=) I2~~3~~. Software includes qualified references to other objects

Reusable: Software is both usable (can be executed) and reusable (can be understood, modified, built upon, or incorporated into other software)

- (=) R1. Software is described with a plurality of accurate and relevant attributes
 - (=) R1.1. Software is given a clear and accessible license
 - (=) R1.2. Software is associated with detailed provenance
- (new) R2 Software includes qualified references to other software
- (=) R1~~3~~. Software meets domain-relevant community standards

Policies für wissenschaftliche software

- kann helfen, Standardverfahren zu etablieren
- kann bei der Einrichtung von Veröffentlichungsprozessen helfen
- kann bei der Auswahl von Lizenzen helfen
- kann bei der Klärung der Eigentumsverhältnisse helfen
- kann bei der Kommerzialisierung der Software helfen
- kann bei der Wartung der Software helfen
- kann helfen, die Software in eine Gemeinschaft einzubringen
- kann die Akzeptanz der Veröffentlichung von Software als wissenschaftlicher Wert verbessern

Beispiel für eine Policy

TU Delft

<https://doi.org/10.5281/zenodo.4629635>

- Licensing
- Registration
- Commercialisation

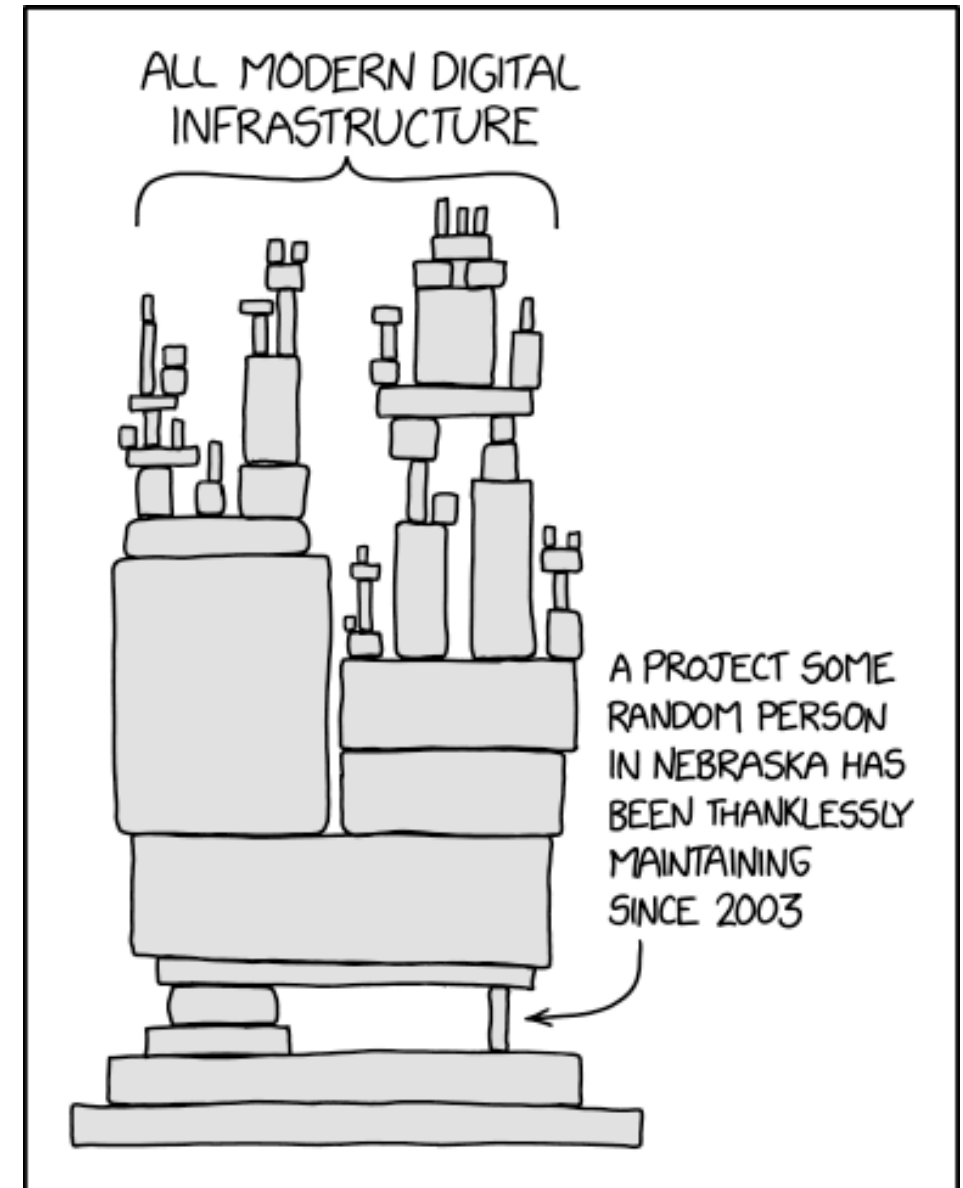


Bazuine et al. (2021): TU Delft Guidelines on Research Software: Licensing, Registration and Commercialisation, CC BY 4.0, <https://doi.org/10.5281/zenodo.4629635>

Open Source Software

Mögliche Vorteile von Open Source:

1. Software ist überprüfbar
2. Flexibilität bei der Entwicklung (eigenes Hosting, lokal oder als Cloud)
3. reduzierte Kosten
4. offen
5. Veröffentlichung der Software



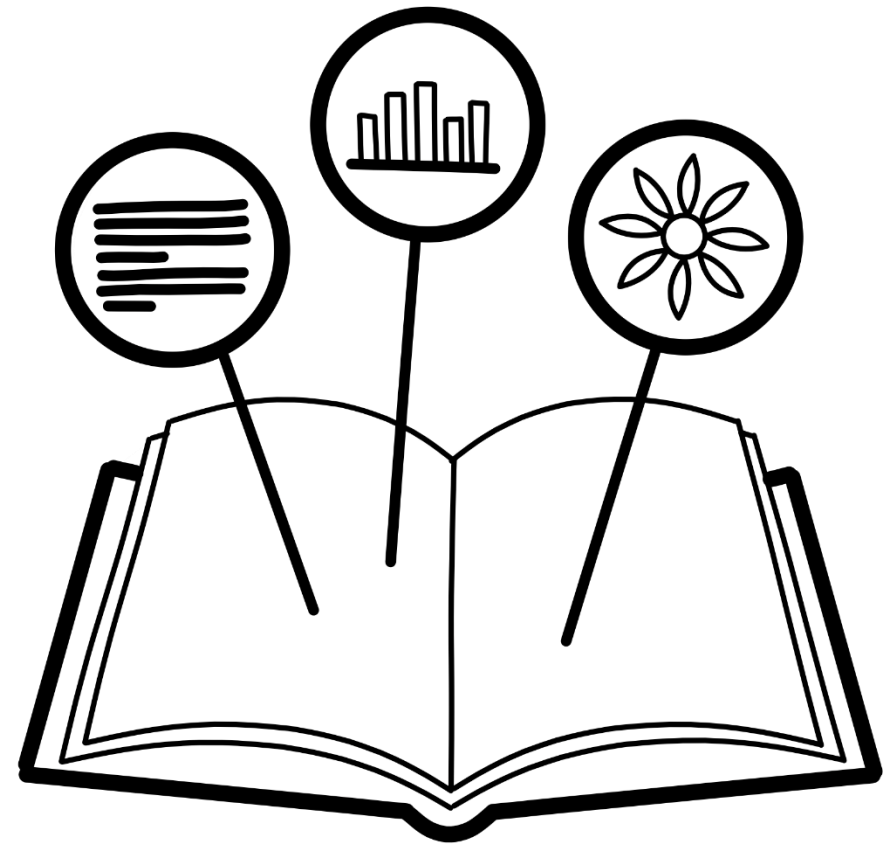
Literaturempfehlung

- Frank Löffler, Christian Busse, und Stephan Janosch: Forschungssoftware in Deutschland, gehalten auf der RDA Deutschland Tagung 26. Februar 2021, <https://doi.org/10.5281/zenodo.4564161>.
- Messerschmidt et al. (2021): Checkliste zur Unterstützung der Helmholtz-Zentren bei der Implementierung von Richtlinien für nachhaltige Forschungssoftware, Potsdam, <https://doi.org/10.48440/os.helmholtz.031>.
- Pampel et al. (2022): Helmholtz Open Science Briefing. Helmholtz Open Science Forum Forschungssoftware. Report, Potsdam, <https://doi.org/10.48440/os.helmholtz.045>.

Jupyter Notebook

Ju Pyt e R

- name Jupyter refers to the three programming languages Julia, Python and R. The file name extension is .ipynb.
- open source web application



<https://doi.org/10.5281/zenodo.5608845>

Jupyter Notebook

- allows you to create, work and share documents for live code, equations, visualisations and narrative text
- easy-to-use platform for data analyses
- Jupyter document is a .json document with a versioned schema consisting of a list of input and output cells and markdown text

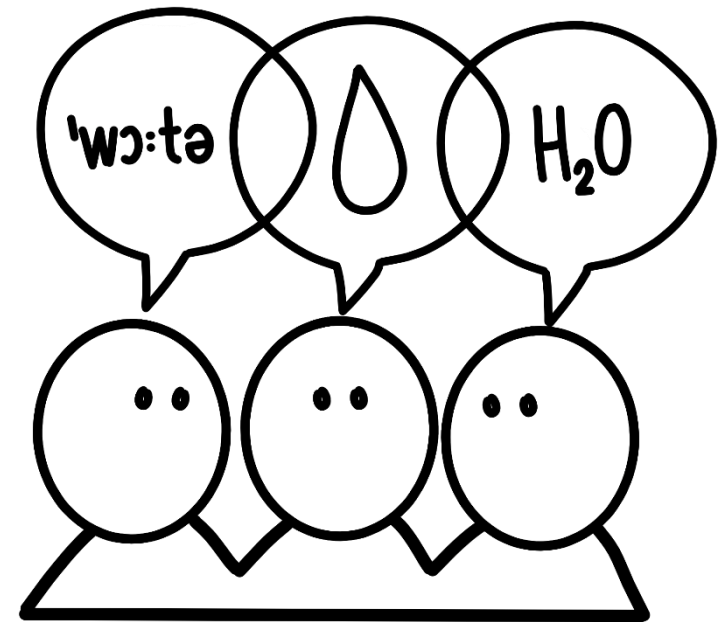


<https://doi.org/10.5281/zenodo.5608845>

Jupyter Notebook

Advantages:

- enables you as researcher to work with text editors, terminals and custom components in a flexible, integrated and extensible way
- easy to use
- does not need to be installed, which is very practical in teaching



<https://doi.org/10.5281/zenodo.3674561>

Markdown + Python = Reproducibility

5/4/21 60563 131327 122999 13316 27284 1232 3047417 217407 29865 624

469 rows × 192 columns

Exploring the data: Plotting and filtering

Plotting

Appending `.plot()` on a dataset, plots all column values vs. the index column.

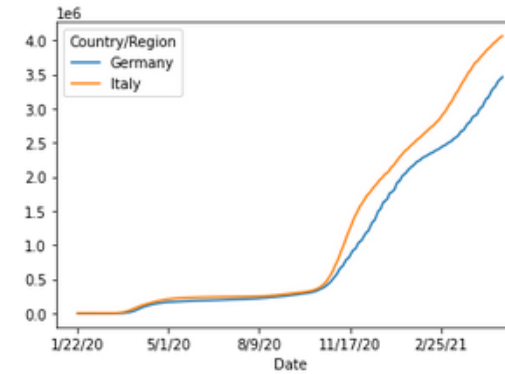
```
[97]: cases.plot()
```

...

plot only selected countries

```
[64]: cases.loc[:,["Germany", "Italy"]].plot()
```

```
[64]: <AxesSubplot: xlabel='Date'>
```



Filtering

Selecting rows and columns from a dataset, based on index or column headers is facilitated through the `.loc[]` slice

Filter on dates (index).

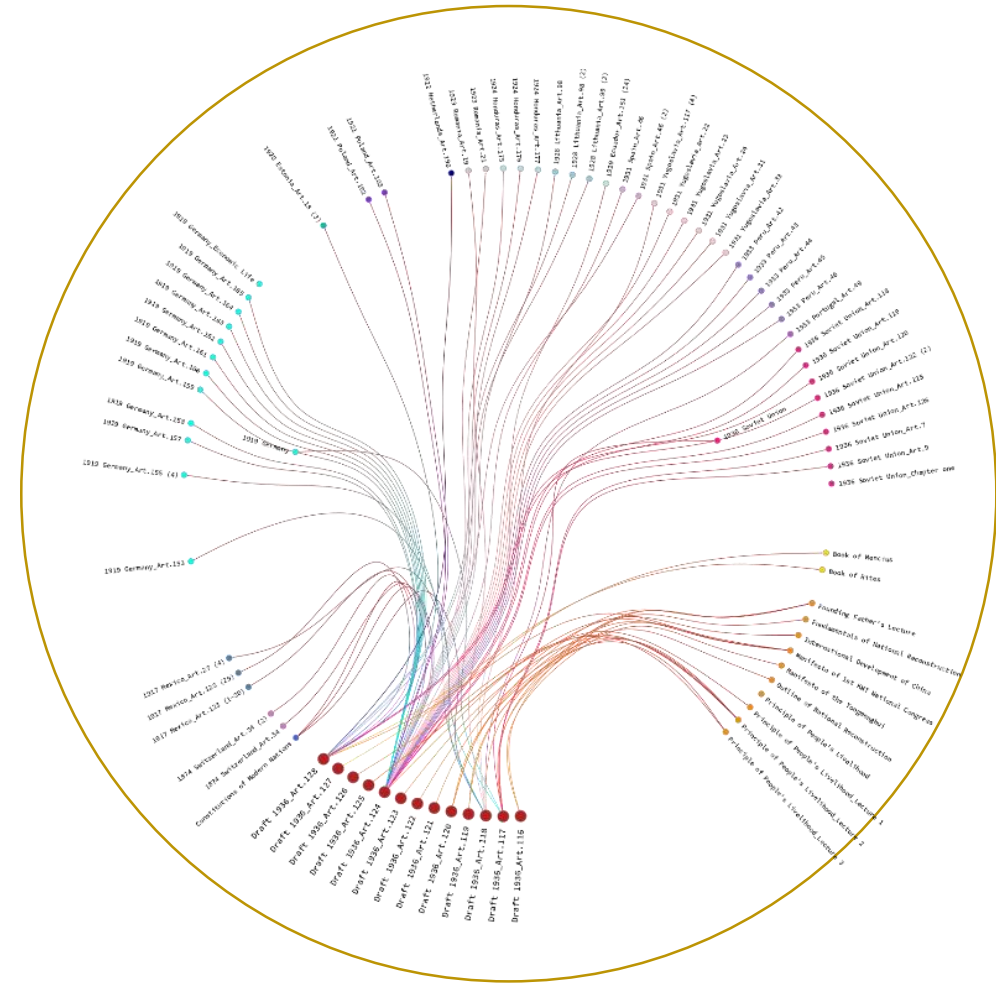
```
[65]: # Get cases on April 1 2020 (Row filter)
cases.loc["4/1/20"]
```

```
[65]: Country/Region
Afghanistan      197
Albania          259
Algeria          847
```

Jupyter and Research Data in the Humanities

Wagner, Andreas, & Li, Fupeng. (2020): rg-mpg-de/fupeng: Citation Network Visualisation of two Drafts (1923 and 1936) of the Chinese Constitution (v1.0). Zenodo.

<https://doi.org/10.5281/zenodo.3699154>.



HPC Clusters and Jupyter

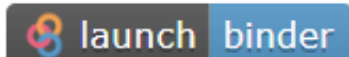
Some examples

- Pangeo: <https://gallery.pangeo.io/contributing.html> (Jupyter documents)
- NeuroLibre: <https://www.neurolibre.com/about/> (Jupyter documents)

Live Demonstration Jupyter Notebook

DataCite_Value-Parser

A short Jupyter Notebook to parse DataCite Values via REST API for Statistics

 launch binder

DOI 10.5281/zenodo.7257106

https://mybinder.org/v2/gh/yvgrossmann/DataCite_Value-Parser/HEAD

Aufgabe:

Öffnen Sie das jypnb in Mybinder und lassen es mit einer ROR-ID Ihrer Institution laufen

Fortbildungsmöglichkeiten

Software Carpentry

- Vermittlung grundlegender Kenntnisse für Forschungssoftware
- Teilnehmen: <https://software-carpentry.org/workshops/>
- Workshop anfordern: <https://software-carpentry.org/workshops/request/>

Data Carpentry

- Data Carpentry: Datenkenntnisse im wissenschaftlichen Kontext (<https://datacarpentry.org>)
- Beispiel: „Data Analysis and Visualization with Python for Social Scientists“ <https://datacarpentry.org/python-socialsci/>

Library Carpentry

- Library Carpentry: Software im bibliothekarischen Kontext (<https://librarycarpentry.org/>)
- Beispiel: „Library Carpentry @ Uni Innsbruck ONLINE – ZB MED - Information Centre for Life Sciences” https://zbmed.github.io/2022-11-14-LC_ZBMED-online/

Offene Fragen Ihrerseits?

Abschließendes

- Folien werden zugänglich gemacht
- Evaluation: Frau Schindler meldet sich diesbezüglich bei Ihnen

Ende

grossmann@mpdl.mpg.de