

Forschungsdatenmanagement

Fachhochschule für den öffentlichen Dienst,
SS 2019

Michael Franke, Max Planck Digital Library



Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung 4.0 International Lizenz](https://creativecommons.org/licenses/by/4.0/).

Literaturempfehlungen

- 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>
- Universität Konstanz: forschungsdaten.info

Forschungsdaten – Situation normal, alles im Eimer



Data Sharing and Management Snafu in 3 Short Acts



NYU Health Sciences Library

Abonnieren 71

32.692

Wofür Forschungsdatenmanagement?

- Gute wissenschaftliche Praxis (DFG)
- Nachnutzung der Daten
- Wissenschaftliches Renommee
 - z.B. Enhanced Publication
(<http://hdl.handle.net/11858/00-001M-0000-0029-B0C0-8>)



Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

Beispiele für Forschungsdaten



Beispiele für Forschungsdaten

Gravitationswellen

BICEP2

BICEP2 2014 Results Release

1. BICEP2 2014 Release Papers
2. BICEP2 2014 Release Data Products
3. BICEP2 2014 Release Figures from Papers
4. Previous Publications

[Videos for Technical Talk](#) \Rightarrow [and News Conference](#) \Rightarrow
[BICEP2 2014 Release Frequently Asked Questions](#) \Rightarrow
[BICEP2 2014 Release Image Gallery](#) \Rightarrow
[BICEP2 Public Web Pages and News Releases](#) \Rightarrow

1. BICEP2 2014 Release Papers

BICEP2 2014 I: Detection of B-mode Polarization at Degree Angular Scales
The BICEP2 Collaboration, 2014

[PDF](#) / [figures](#) / [arXiv](#) / [ADS](#)

BICEP2 2014 II: Experiment and Three-year Data Set

The BICEP2 Collaboration, 2014

[PDF](#) / [figures](#) / [arXiv](#) / [ADS](#)


2. BICEP2 2014 Release Data Products

| | |
|---|---|
| B2_3yr_rlikelihood_20140314.txt | Text file containing the tabulated likelihood for the tensor-to-scalar ratio, r , computed using the "direct likelihood calculation" described in Section 9.3.1 of Barkats <i>et al.</i> and Section 11.1 of BICEP2 2014 I. |
| B2_3yr_bandpowers_20140314.txt | Text file containing bandpowers and statistical uncertainties, corresponding to Figure 2 of BICEP2 2014 I. |
| B2_3yr_bandpower_window_functions.tgz B2_3yr_bpwf_bin1_20140314.txt B2_3yr_bpwf_bin2_20140314.txt B2_3yr_bpwf_bin3_20140314.txt B2_3yr_bpwf_bin4_20140314.txt | Bandpower window functions for each l bin. |

Beispiele für Forschungsdaten

Gravitationswellen

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

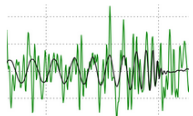
released 2016 June 15:
Event of December 26, 2015: GW151226: total mass 22

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

The screenshot shows the NCBI GenBank database entry for the Homo sapiens class II AP endonuclease (APE) gene, partial CDS. The entry includes various fields such as LOCUS, DEFINITION, ACCESSION, VERSION, KEYWORDS, SOURCE, ORGANISM, REFERENCE, AUTHORS, TITLE, JOURNAL, PUBMED, and FEATURES. The FEATURES section lists source, intron, gene, mRNA, exon, and CDS with their respective coordinates and qualifiers.

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Limits Advanced Help

Display Settings: GenBank Send:

Homo sapiens class II AP endonuclease (APE) gene, partial CDS
GenBank: M99703.1
[FASTA](#) [Graphics](#)

Go to:

LOCUS HUMAPES 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](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3019)
Harrison,L., Ascione,G., Menninger,J.C., Ward,D.C. and Demple,B.
TITLE Human apurinic endonuclease gene (APE): structure and genomic mapping (chromosome 14q11.2-12)
JOURNAL 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)

Change region shown

Customize view

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 j [Am J Clin Pathol. 2014]
Dysregulation of human apurinic/apyrimidinic endonuclease 1 (APE1) x [Er J Ophthalmol. 2014]
PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR [Mol Cell. 2014]
See all...

Reference sequence information
RefSeq alternative splicing
See 6 reference mRNA sequence splice variants for the APEX1 gene.

More about the APEX1 gene
Apurinic/apyrimidinic (AP) sites occur frequently in DNA molecules by spontaneous hydrolysis, by DNA damaging agents or by DNA glycosylases
...
Also Known As: APE, APE1, APEN, APEX, ...

Genbank

Beispiele für Forschungsdaten

Tierbewegungen

Movebank

The screenshot displays the Movebank website interface. At the top, the logo reads "MOVEBANK For Animal Tracking Data". The navigation bar includes "Home", "Tracking Data Map", "Community", "Help", "Tools", "Env-DATA", and "Published Data". A search bar is located on the right side of the navigation bar.

On the left side, there is a "User login" section with fields for "Username:" and "Password:", a "Log in" button, and links for "Create new account" and "Request new password".

The main content area is divided into two columns. The left column contains a search interface with a search bar (containing "stork"), a dropdown for "All Sensor Types", and a checkbox for "Only studies where I can see data". Below this is a "Search result" section with a "Sort by Animal Identifier" dropdown and a list of search results. The right column features a map of Europe with various countries labeled (e.g., Deutschland, France, España, Italia) and several green circular markers indicating tracking data points. A purple line on the map shows a movement path across the Iberian Peninsula and into France.

At the bottom of the page, there is a "Europa" logo and a footer with the text: "Movebank is coordinated by the Max Planck Institute for Ornithology, the North Carolina Museum of Natural Sciences, and the University of Konstanz. For questions or feedback, contact us at support@movebank.org".

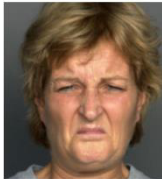
Beispiele für Forschungsdaten

Gesichter/Mimik

FACES



Faces



168_m_f_d_b.jpg

168_m_f_h_a.jpg

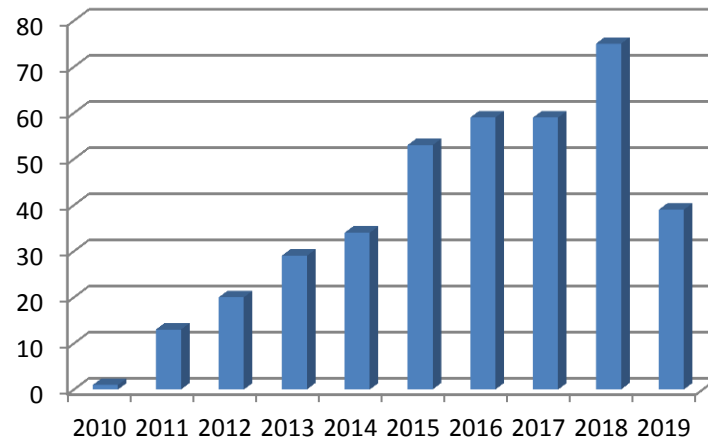
168_m_f_h_b.jpg

168_m_f_s_b.jpg

004_o_m_n_a.jpg

066_y_m_d_a.jpg

Referenzen



Scopus

384 documents have cited:

FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation

Ebner N.C., Riediger M., Lindenberger U.

(2010) Behavior Research Methods, 42 (1), pp. 351-362.

Set feed

REGISTER HELP

72 Items

(12)

(12)

(12)

(12)

(36)

(36)

(24)

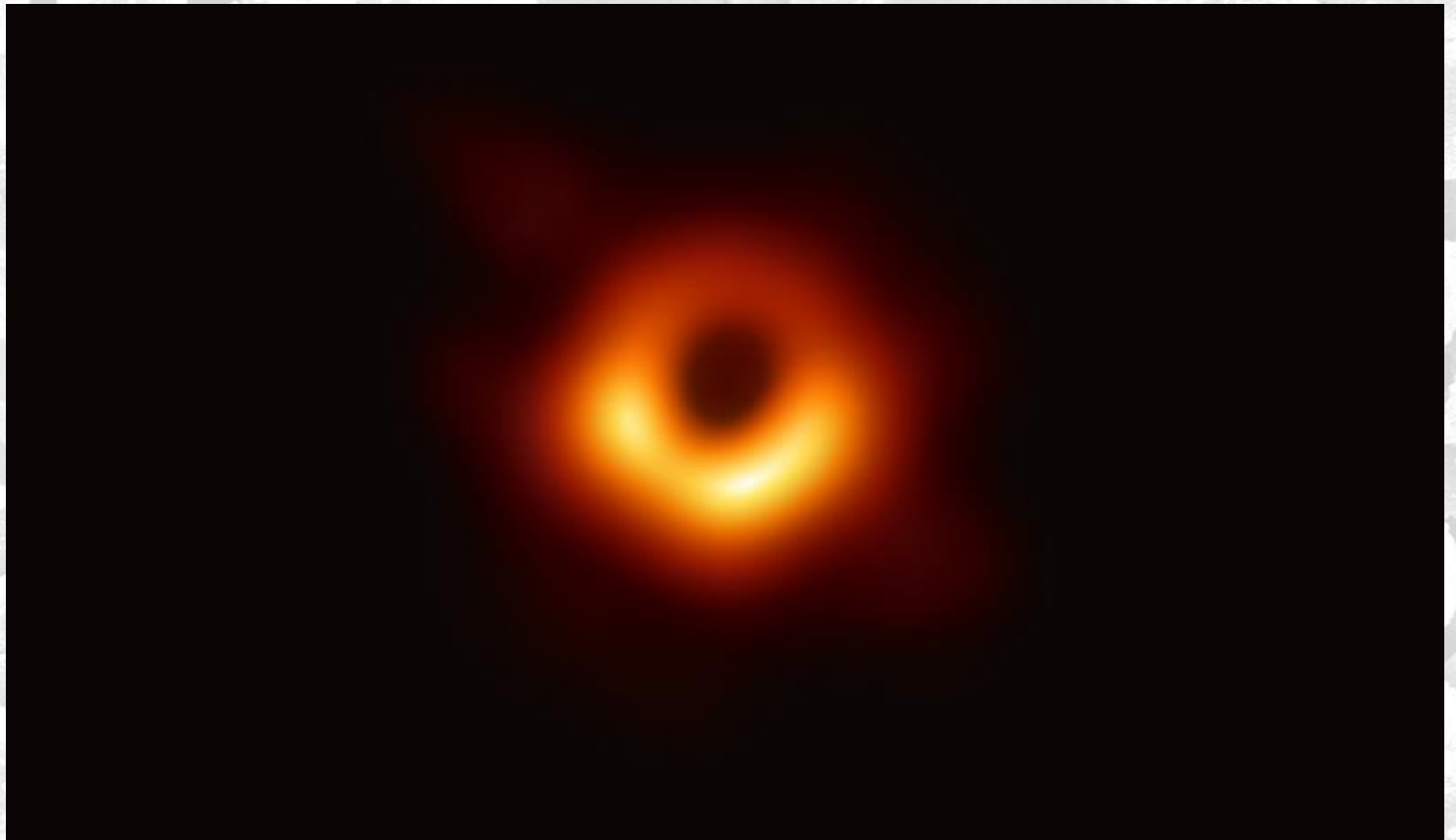
(24)

(36)

(36)

Beispiele für Forschungsdaten

Schwarze Löcher



Definitionsversuche

- „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“ (Uni Heidelberg).
- „Forschungsdaten stellen im weitesten Sinne Primärdaten, Sekundäranalysen, Visualisierungen, Modelle, Analysewerkzeuge, Objektsammlungen oder Produkte dar, die während des wissenschaftlichen Arbeitsprozesses erzeugt und benutzt werden“ (Johanna Vompras, Uni Bielefeld).
- *„Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“ (Allianz AG „Forschungsdaten“)*
- *„(...) researchers know within their projects what research data is“ (Leibniz-Gemeinschaft)*

Worum geht es nicht

- Physische Daten
- Wissenschaftliche Publikationen
- Daten über Wissenschaft

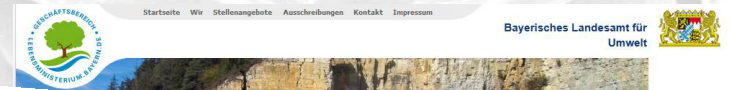
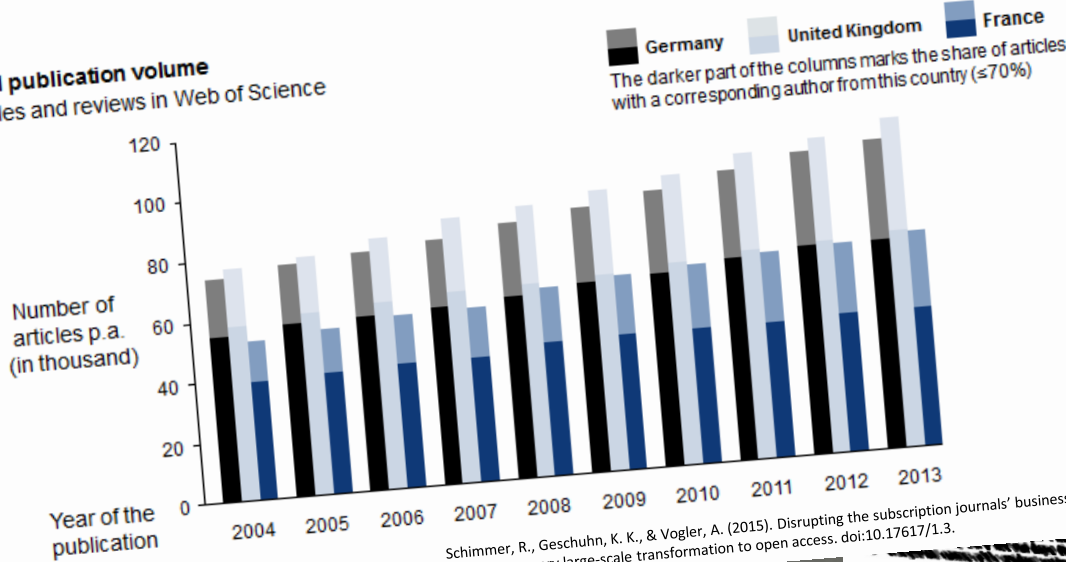
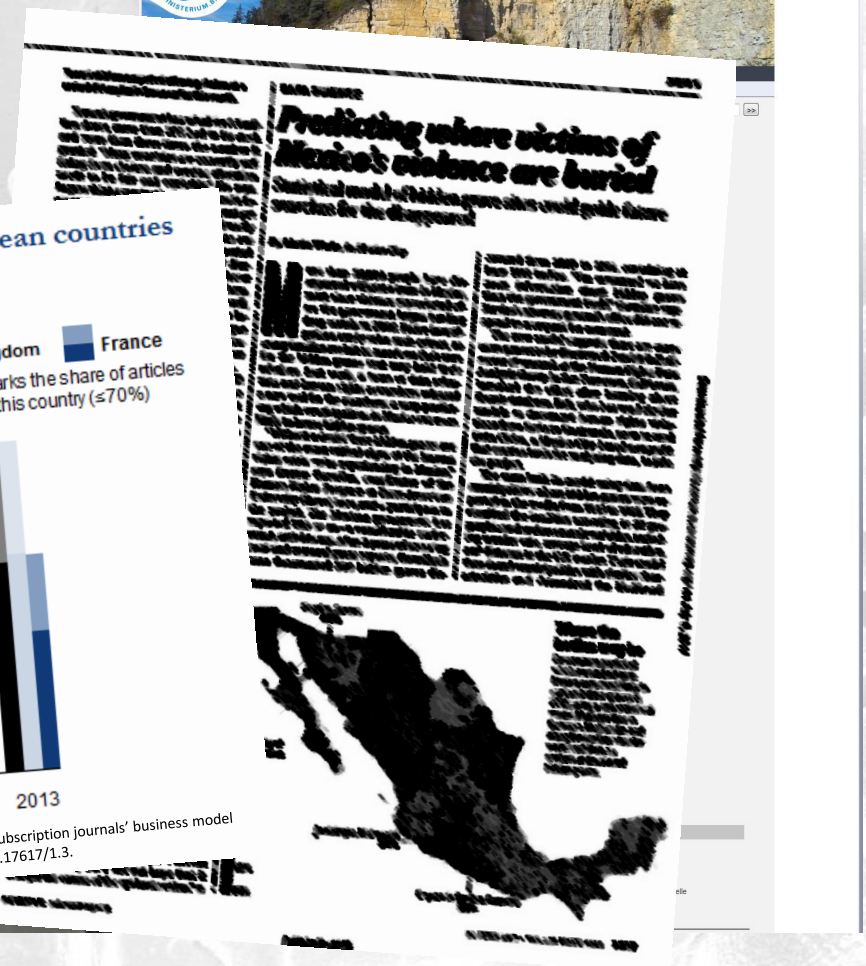


Figure 1: Publication volume with APC relevant share of selected European countries

Total publication volume
Articles and reviews in Web of Science



Schimmer, R., Geschuhn, K. K., & Vogler, A. (2015). Disrupting the subscription journals' business model for the necessary large-scale transformation to open access. doi:10.17617/1.3.



Klassifikation von Forschungsdaten

Nach dem Charakter der Daten

Texte (Quellen, Transkripte)

Bilder (Scans, Fotos, Mikroskopdaten, Teleskopdaten, Satellitendaten)

Multimedia (Audiodaten, Videodaten, 3D-Daten, 4D-Daten)

Zahlenreihen (Umfragedaten, Experimentaldaten, Sensormessreihen, Gensequenzen, Geodaten)

Born-digitals (Simulationsdaten, Algorithmen, Websites)

Qualitative Daten



Quantitative Daten



Klassifikation von Forschungsdaten

the APICS Consortium (Max Planck Institute for Evolutionary Anthropology). APICS supplemental files. <http://edmond.mpg.de/imeji/collection/4WKY1hHhw8iEuNQ4>

Nach dem Format der Daten

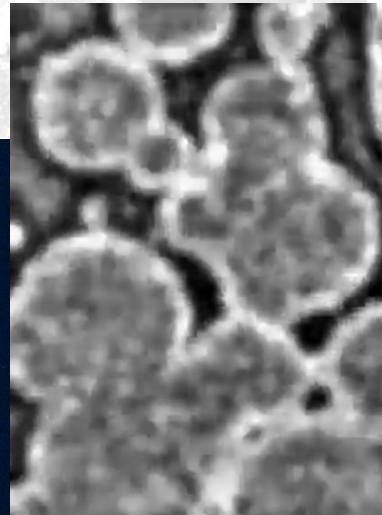
Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Zahlenreihen (CSV, XLSX, ...)

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



```
netcdf tos_01_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" ;
```

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

```
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 |
| | | 235 |
| | | 256 |
| | | 230 |
| | | 252 |
| | | 238 |
| | | 244 |
| | | 215 |
| | | 196 |
| | | 162 |
| | | 188 |
| | | 216 |
| | | 197 |
| | | 213 |
| | | 243 |
| | | 239 |

<http://dx.doi.org/10.17617/3.i>

Klassifikation von Forschungsdaten

Nach der Quelle der Daten
bzw. der Art der Forschung

Observational: data captured in real time that is usually unique and irreplaceable. For example, remote sensing data, survey data, field recordings, sample data.

Experimental: data captured from lab equipment that is often reproducible. For example, gene sequences, chromatograms, magnetic field data.

Models or simulation: data generated from test models where model and metadata may be more important than output data from the model. For example, climate models, economic models.

Derived or compiled: resulting from processing or combining 'raw' data. For example, text and data mining, compiled databases, 3D models.

Reference or canonical: a static or organic conglomeration or collection of datasets, probably published and curated. For example, gene sequence databanks, collection of letters.

(University of Bristol, Boston University)

Klassifikation von Forschungsdaten

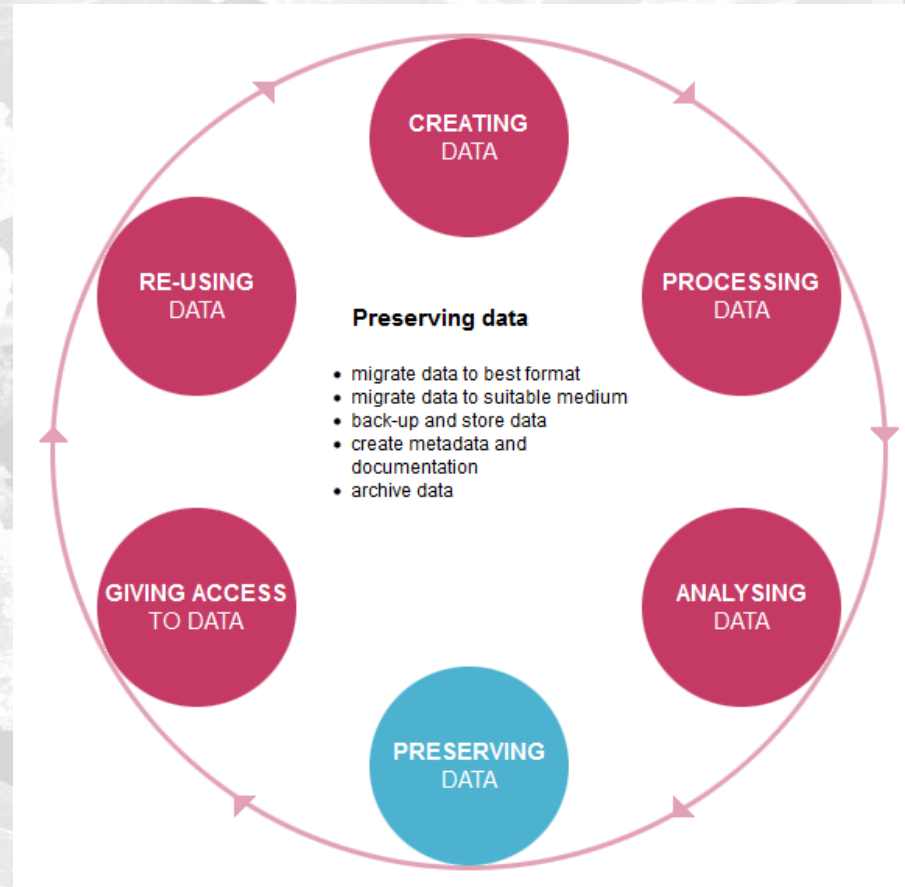
Nach dem Grad der Aggregation/Verdichtung

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.

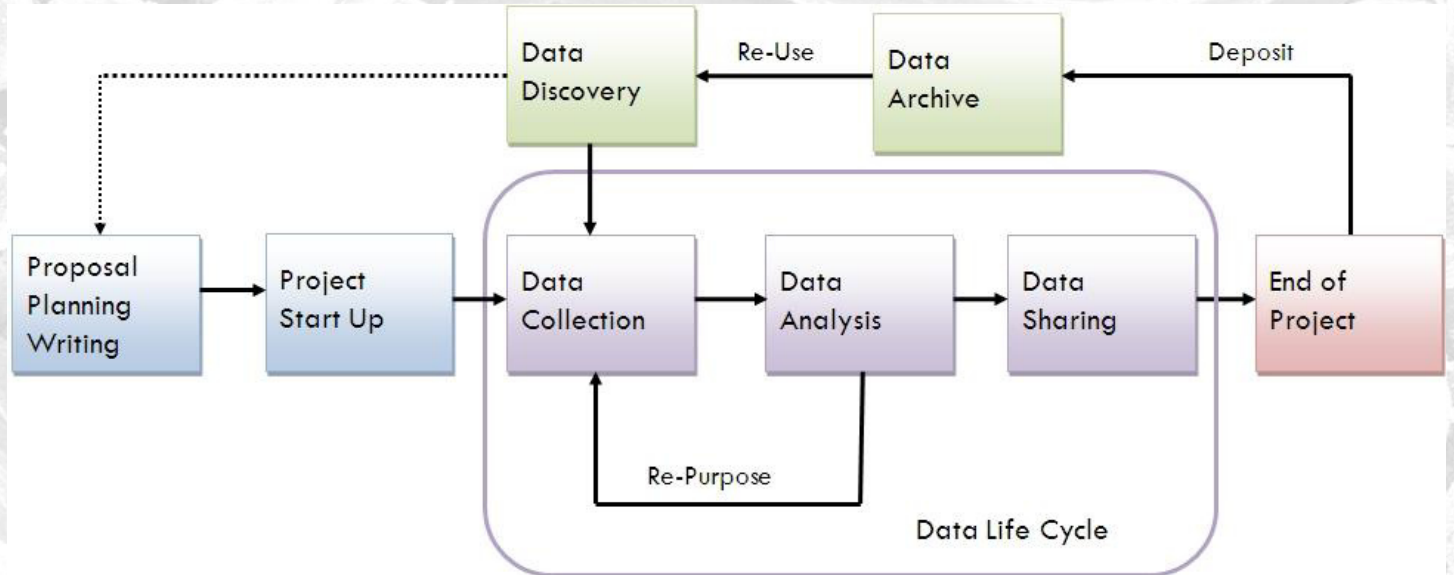
Sekundärdaten: Sekundärdaten seien Daten, die in Prozessschritten aus Primärdaten gewonnen wurden.

Forschungsdatenzyklus



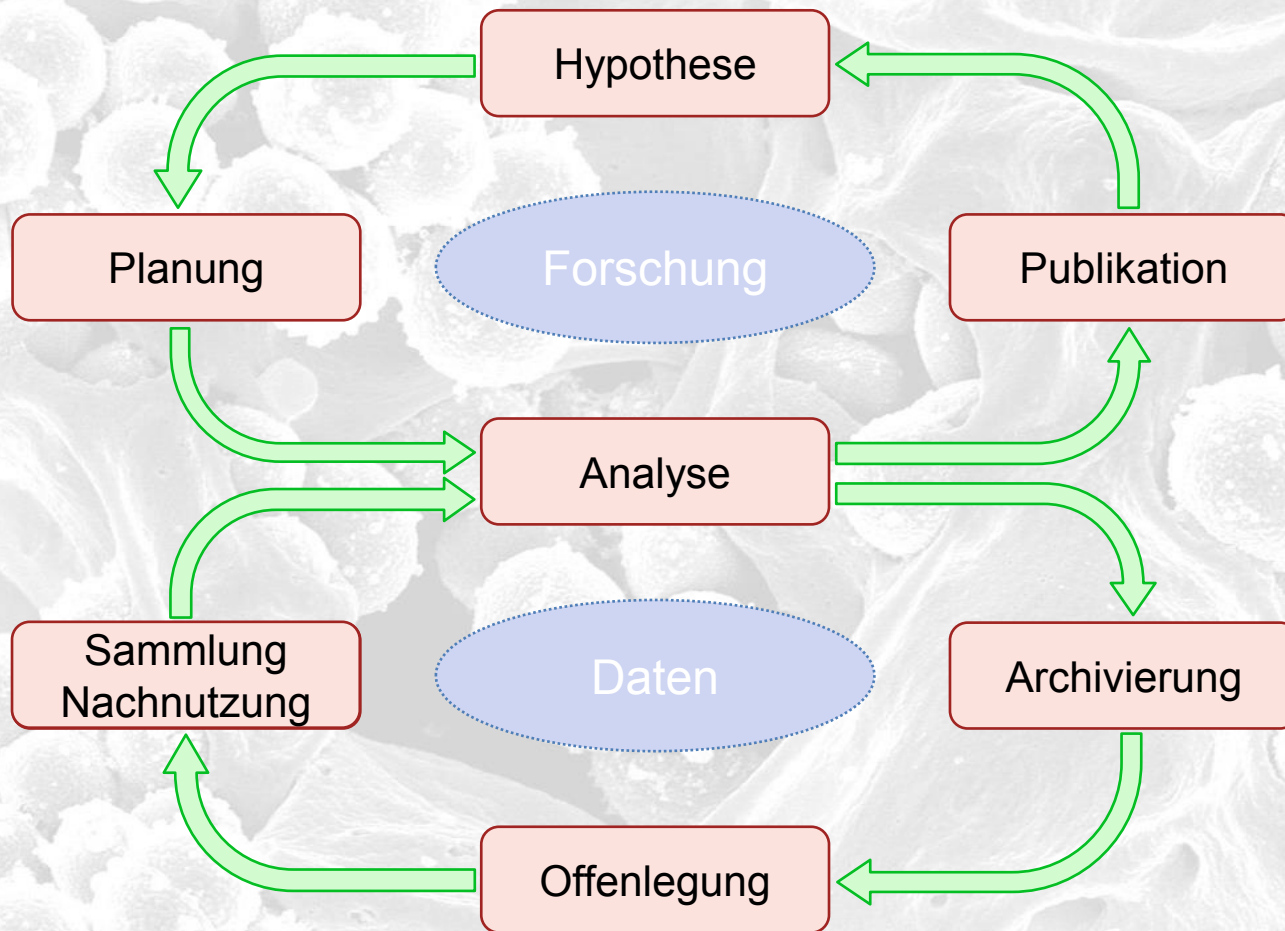
UK Data Archive

Forschungs-Datenzyklus

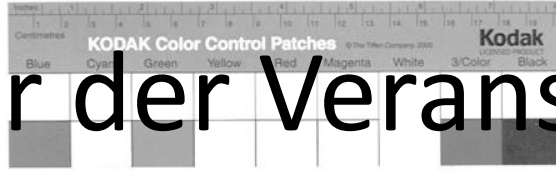


University of Virginia

Forschungs- und Datenzyklus



Struktur der Veranstaltung



Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen



Kodak
White 3/Color Black
ktur?

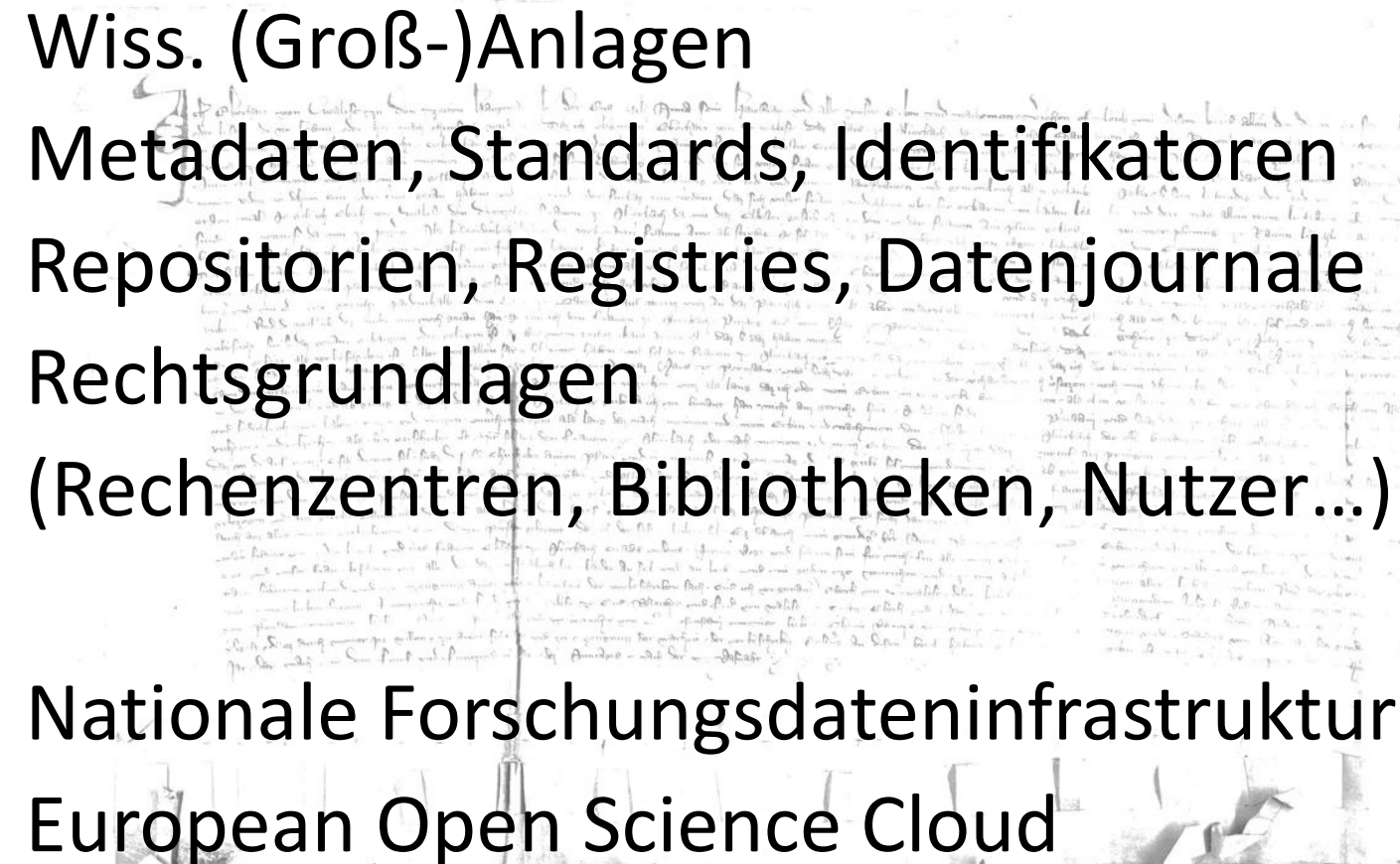
- z.B. Verkehrsinfrastruktur

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





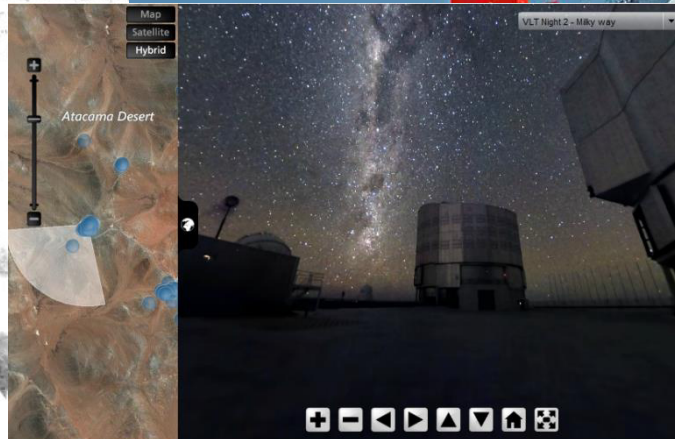
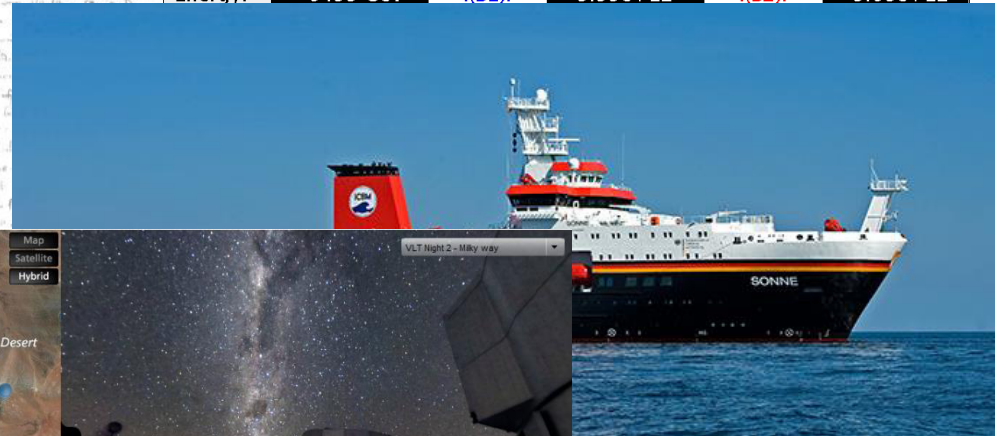
Forschungsdateninfrastruktur

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

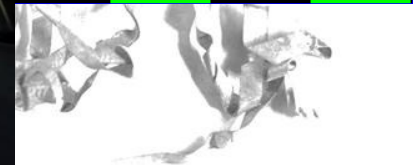
Wissenschaftliche Großanlagen

- z.B. LHC, Sonne,
E-ELT, SKA
Wendelstein 7-X

| | | | | |
|------------------------------|------------|-------------|-----------------|-------------------|
| LHC Page1 | Fill: 5930 | E: 6499 GeV | t(SB): 02:37:13 | 11-07-17 13:32:02 |
| PROTON PHYSICS: STABLE BEAMS | | | | |
| Energy: | 6499 GeV | I(B1): | 5.59e+12 | I(B2): 5.68e+12 |

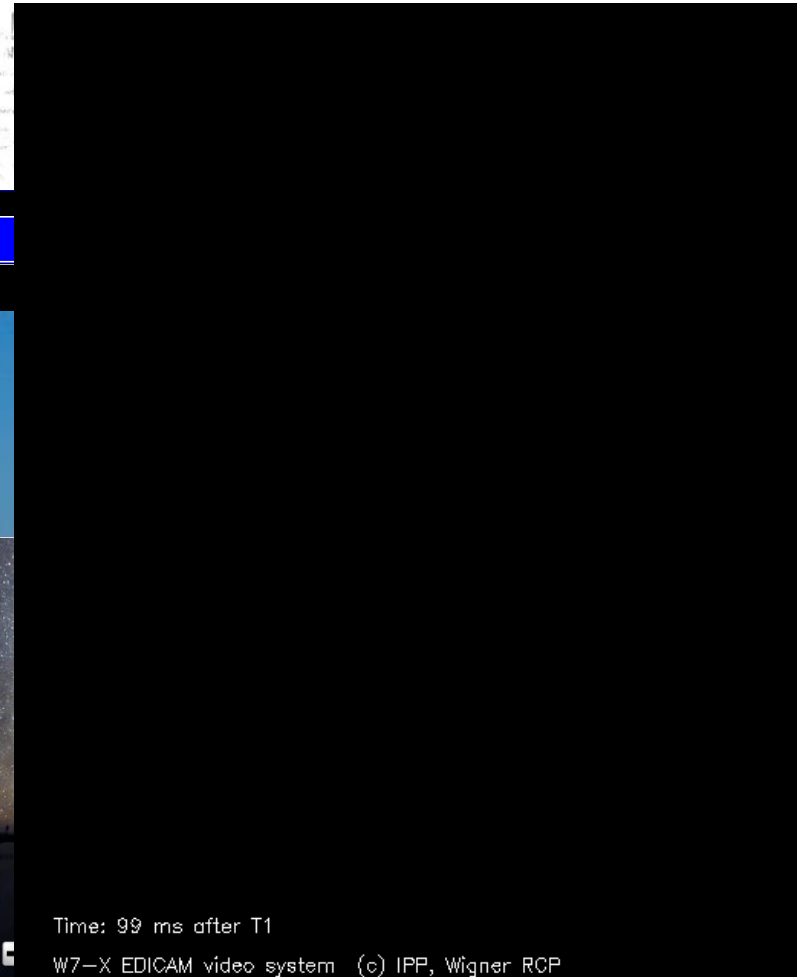
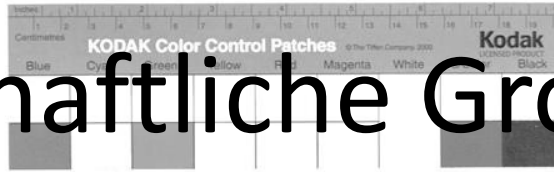


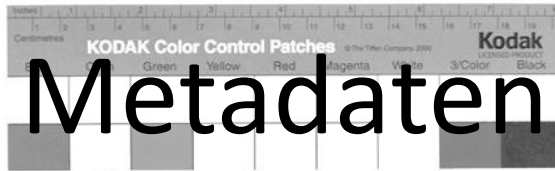
| | | | |
|-----------------------------|---------|--------------|---------|
| Moveable Devices Allowed In | true | true | |
| Stable Beams | true | true | |
| PM Status B1 | ENABLED | PM Status B2 | ENABLED |



Wissenschaftliche Großanlagen

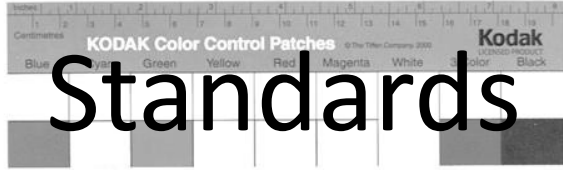
- z.B. [LHC](#), [Sonne](#), [E-ELT](#), [SKA](#), [Wendelstein 7-X](#)
- Erheben/Erzeugen /Verteilen Daten
- Beschleunigen Prozesse
- Zwingen zur Kooperation





Metadaten

- [DC](#) – Dublin Core Metadata Initiative
- [RDF](#) – Resource Description Framework
([Beispiel](#))
- [SKOS](#) – Simple Knowledge Organization System
- [Metadaten-Standards](#)
- ([Metadaten-Registries](#))



- Metadaten
- Daten-/Dateiformate
- Protokolle/Schnittstellen
- Lizenzen

[Pronom](#)

[IANA](#)

HTTP

REST

OAI-PMH

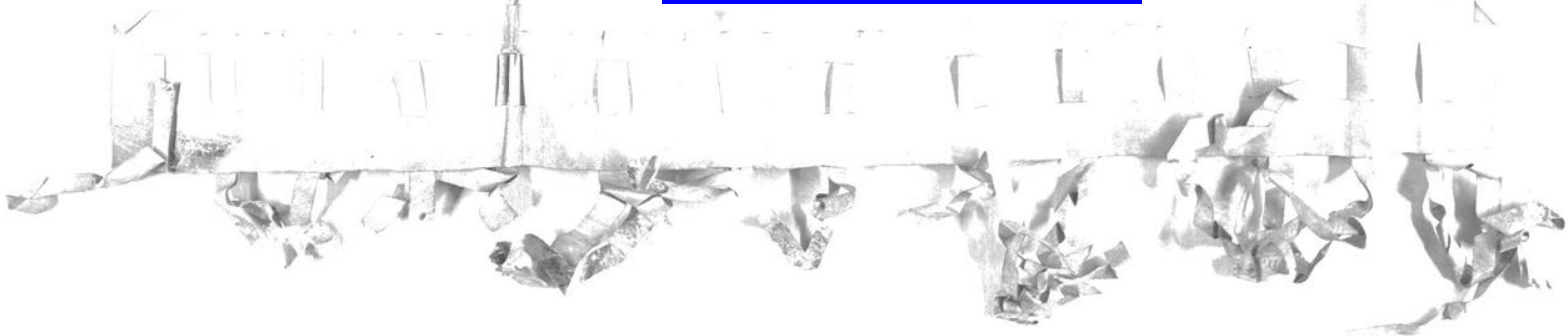
Z39.50

JSON

[Creative Commons](#)

[Open Data Commons](#)

[Datenlizenz Deutschland](#)



Persistente Identifikation

<http://zuse.zib.de/collection/wl3XoEDHO8v0ImCa/item/VqNgKUacPrIhqPKu>

- URL/URI

- URN

<urn:nbn:de:bsz:25-opus-14124>

- Handle

<hdl:11858/00-001M-0000-0019-D20F-6>

- DOI

<doi:10.1088/0004-637X/715/2/1453>

- Researcher-ID

<G-2512-2011>

- ORCID

<0000-0002-2661-8242>

Forschungsdatenrepositorien I

Institutionell

Fachlich

Allgemein

[Open Data LMU](#)

[PsychData](#)

[Zenodo](#)

[Pangaea](#)

[Dryad](#)

[Radar](#)

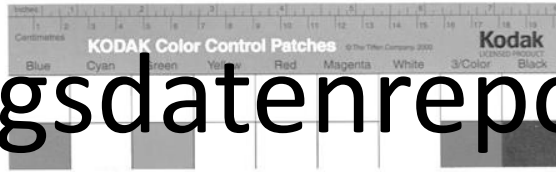
[Edinburgh DataShare](#)

[Edmond](#)

[G-Node](#)

[Figshare](#)

[Mendeley Data](#)



A Kodak Color Control Patches ruler is positioned at the top of the slide. It features a ruler scale from 1 to 18 centimeters. Below the ruler, there are color patches labeled: Continues, Blue, Cyan, Green, Yellow, Red, Magenta, White, 3/Color, and Black. The Kodak logo is visible on the right side of the ruler.

Forschungsdatenrepositorien II

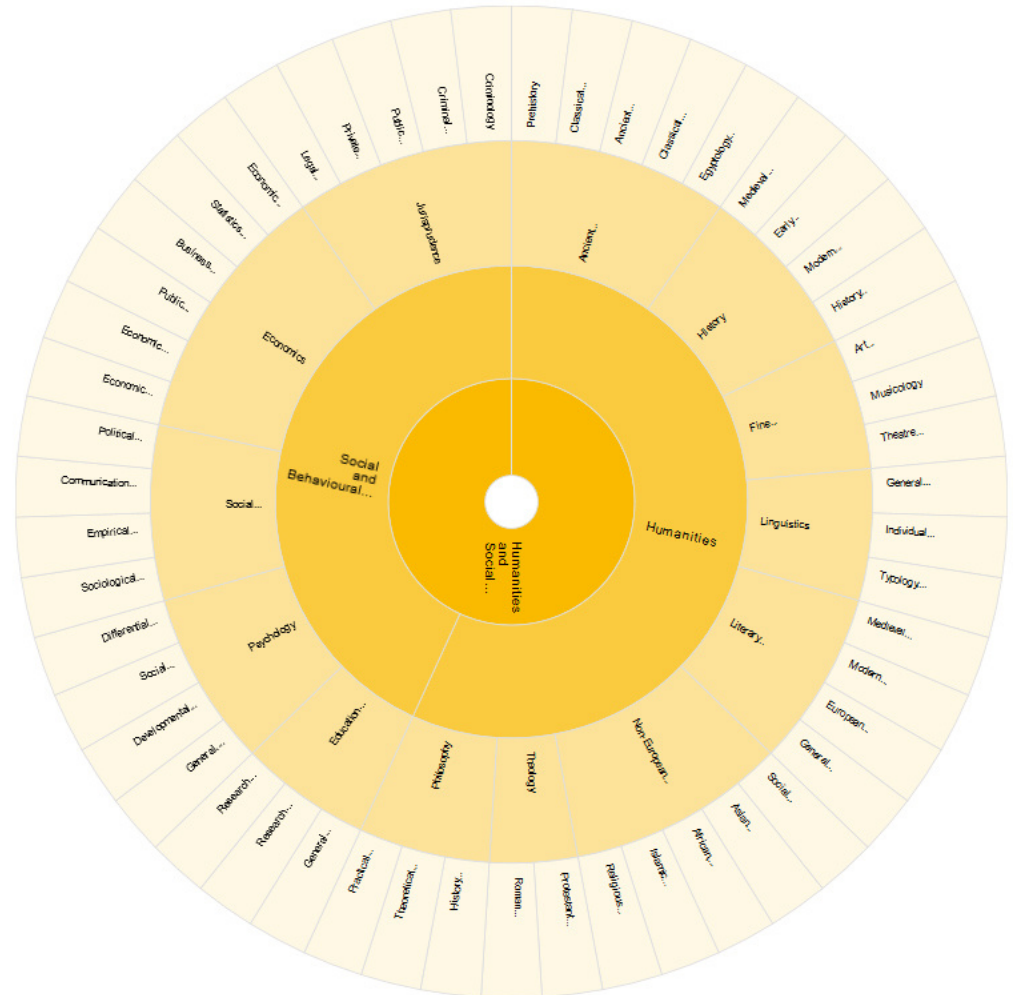
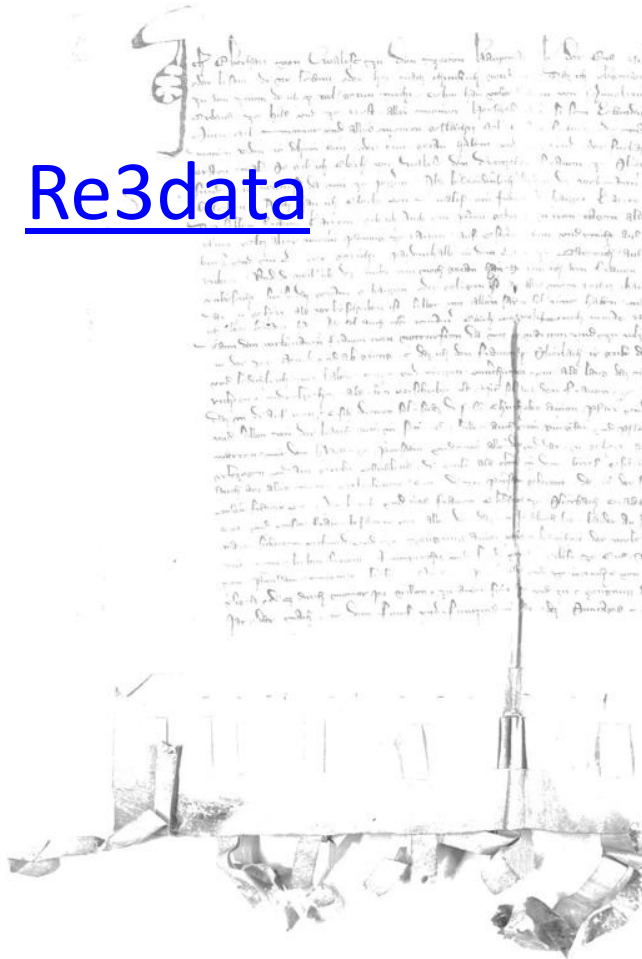
Aufgaben ([Core Trust Seal](#)):

- R1. The repository has an explicit mission to provide access to and preserve data in its domain.
- R2. The repository **maintains all applicable licenses** covering data access and use and monitors compliance.
- R3. The repository has a **continuity plan** to ensure ongoing access to and preservation of its holdings.
- R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in **compliance with disciplinary and ethical norms**.
- R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a **clear system of governance** to effectively carry out the mission.
- R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-house, or external, including scientific guidance, if relevant).
- R7. The repository guarantees the **integrity and authenticity of the data**.
- R8. The repository accepts data and metadata based on defined criteria to **ensure relevance and understandability** for data users.
- R9. The repository applies documented processes and procedures in managing archival storage of the data.
- R10. The repository assumes **responsibility for long-term preservation** and manages this function in a planned and documented way.
- R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.
- R12. Archiving takes place according to defined workflows from ingest to dissemination.
- R13. The repository **enables users to discover the data and refer to them in a persistent way** through proper citation.
- R14. The repository **enables reuse of the data** over time, ensuring that appropriate metadata are available to support the understanding and use of the data.
- R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.
- R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Forschungsdatenrepositorien III

Repositorien finden

- [Re3data](#)





Datenjournale

- Lebenswissenschaften
 - [Gigascience](#)
 - [Biodiversity Data Journal](#)
 - [Open Health Data](#)
- Geowissenschaften
 - [Earth System Science Data](#)
- Naturwissenschaften allgemein
 - [Scientific Data](#)



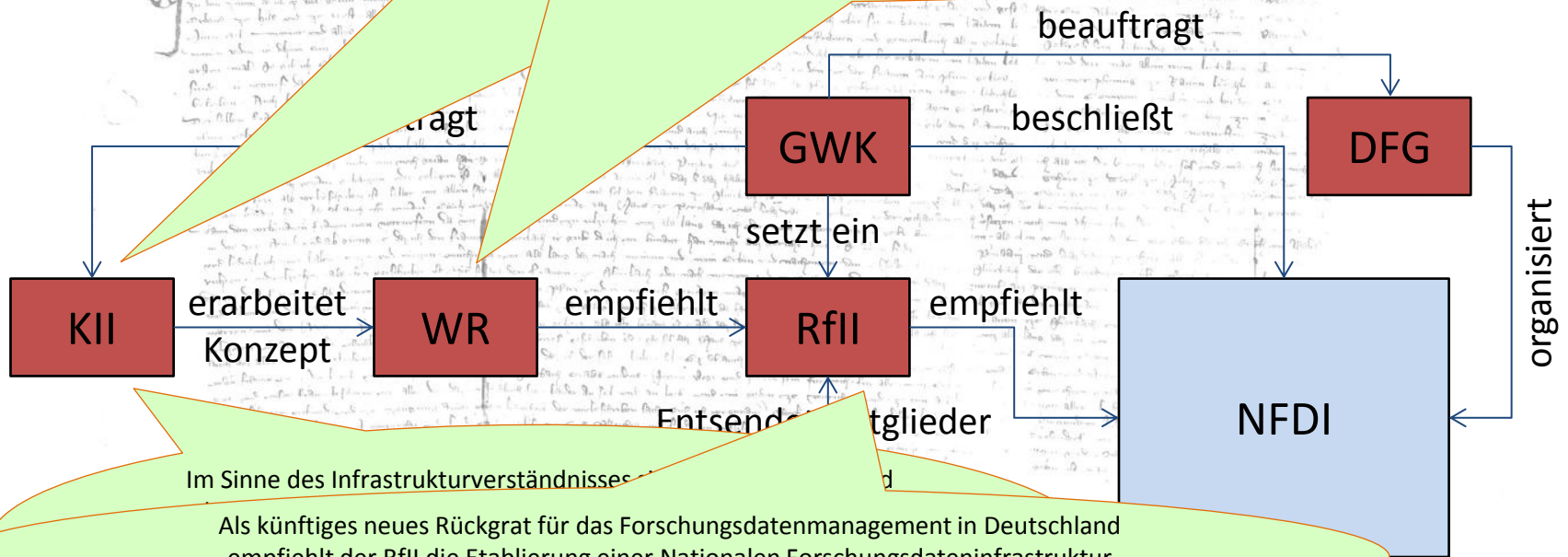
Rechtliche Aspekte

- Grundgesetz, Urheberrecht, Kunsturheberrechtsgesetz, Datenbankrecht, Archivrecht, Datenschutzgrundverordnung, Strafrecht, Kriegswaffenkontrollgesetz, Jugendschutzgesetz, Kooperationsverbot
- PSI-Direktive
- Lizenzen für Forschungsdaten
- Einwilligungserklärungen
- Pre-Registration

Nationale

Nähezu für alle Handlungsfelder die Vernetzung mit weiteren Einrichtungen empfohlen. **Vernetzung bedeutet** a) Aufbau einer stabilen nationalen Struktur b) Verzahnung der Handlungsfelder,

Der Wissenschaftsrat teilt die Einschätzung der KII, dass für eine strategische Weiterentwicklung des Gesamtsystems der Informationsinfrastrukturen in Deutschland ein übergeordnetes Koordinierungs- und Beratungsgremium erforderlich ist, das im Folgenden als Rat für Informationsinfrastrukturen bezeichnet wird.



Im Sinne des Infrastrukturverständnisses...
Als künftiges neues Rückgrat für das Forschungsdatenmanagement in Deutschland empfiehlt der RfII die Etablierung einer Nationalen Forschungsdateninfrastruktur (NFDI).

2011

2012

2014

2016

2018

2019

Bund-Länder-Vereinbarung zur NFDI

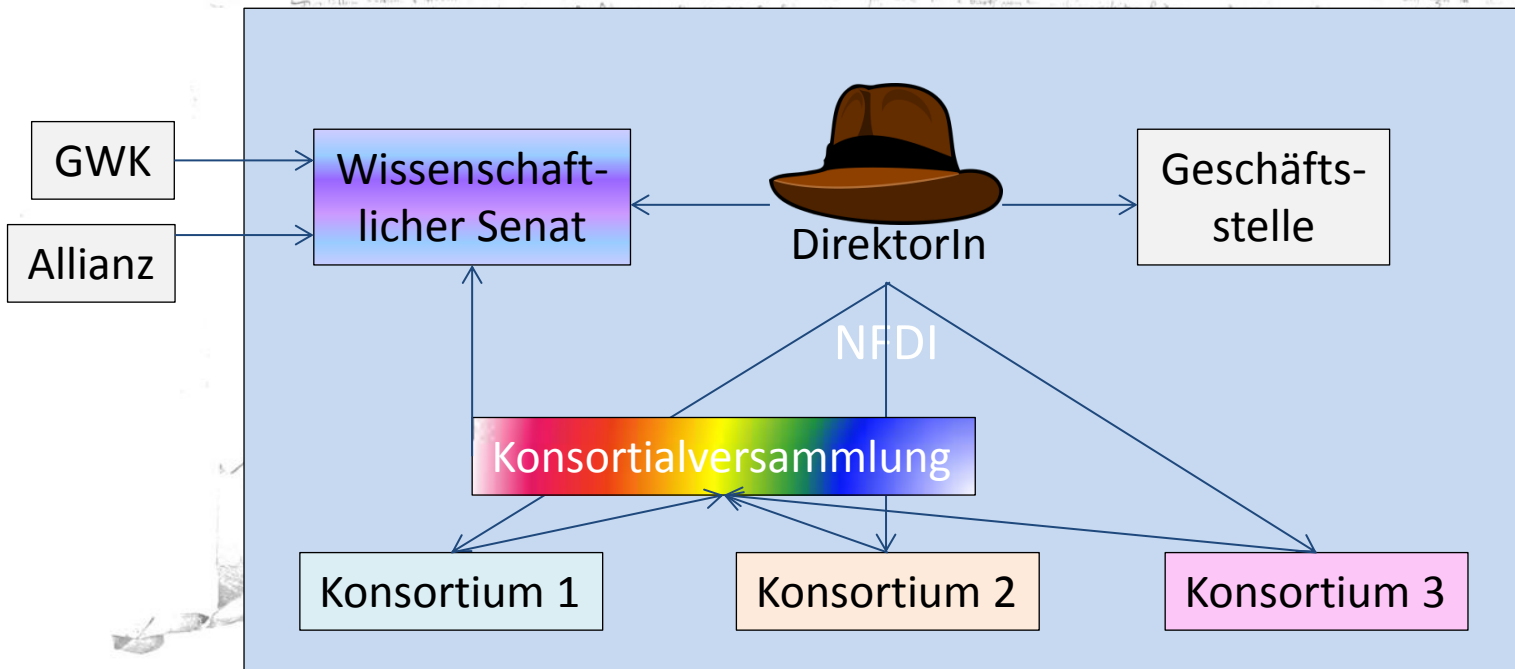
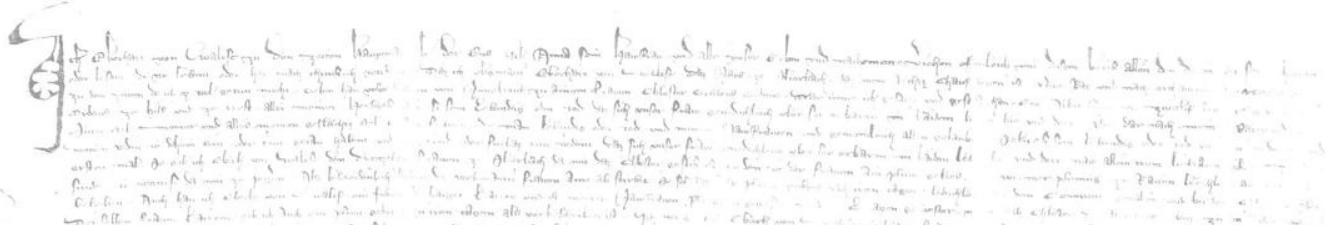


<https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/NFDI.pdf>

- Aufbau einer koordinierten, vernetzten Informationsinfrastruktur zur Entwicklung eines nachhaltigen interoperablen Forschungsdatenmanagements;
- Etablierung von in den wissenschaftlichen Disziplinen akzeptierten Prozessen und Verfahren zum standardisierten Umgang mit Forschungsdaten;
- Schaffung eines verlässlichen und nachhaltigen Dienste-Angebots, welches übergreifende und fachspezifische Bedarfe des Forschungsdatenmanagements in Deutschland abdeckt;
- Entwicklung disziplinübergreifender Metadatenstandards zur flächendeckenden (Nach-) Nutzbarkeit von Forschungsdaten;
- Anbindung der deutschen Forschungsdateninfrastrukturen an europäische und internationale Plattformen;
- Optimierung der Nachnutzbarkeit bereits erhobener Forschungsdaten wie auch der Infrastrukturen, in die sie eingebettet sind; dadurch Generierung zusätzlichen Wissens ohne den hohen Aufwand einer Datenneuerhebung;
- Schaffung einer gemeinsamen Basis für Datenschutz sowie der Souveränität, Integrität, Sicherheit und Qualität von Daten.
- Bund und Länder beabsichtigen, in drei Ausschreibungsrunden die Auswahl von insgesamt bis zu 30 Konsortien vorzunehmen.
- Konsortien sind auf langfristige Zusammenarbeit angelegte Zusammenschlüsse von Nutzern und Anbietern von Forschungsdaten wie staatliche und staatlich anerkannte Hochschulen, außeruniversitäre Forschungseinrichtungen, Ressortforschungseinrichtungen, Akademien und anderen öffentlich geförderten Informationsinfrastruktureinrichtungen. Sie sind in der Regel nach Fachgruppen bzw. Methoden organisiert, ohne Vorgaben für ihre institutionelle Zusammensetzung.

Struktur und Governance der NFDI

| | | | | |
|------|------|--------|----------|-------|
| KOAK | for | Open | Partners | LOGO |
| Büro | Team | Yellow | Red | White |
| | | | | |
| | | | | |





Konsortialanträge

Astro-NFDI

BERD@NFDI

DAPHNE

DataPLANT

FAIRmat

ForumX

GHGA

KonsortSWD

MaRDI

NFDI4Agri

NFDI4BioDiversity

NFDI4cat

NFDI4Chem

NFDI4Crime

NFDI4Culture

NFDI4Earth

NFDI4Health

NFDI4Ing

NFDI4Medicine

NFDI4MobilTech

NFDI4MSE

PAHN-PaN

RWD-DHct

Text+

2linkNFDI

Bridge4NFDI

CompeNDI

DeBioData

Di-Morph

InterdisciplinaryNFDI

MDM-Portal

MOPED

NFDI4AIRR

NFDI4CS4NDFI*

NFDI4Life Umbrella

NFDI4Memory

NFDI4Microbiota

NFDI4NanoSafety

NFDI4NeuroFunction

NFDI4Objects

NFDI4Phys

NFDI4RSE

NFDI4SolidEarth

NFDI-Neuro

NFDI Small Disciplines

NFDI Web

TheoReS

LoReData

NFDI4BIMP

NFDI4Language

European Open Science Cloud

- Europäische Kommission

OPPORTUNITIES

EUROPEAN OPEN SCIENCE CLOUD
BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES

€2 BN IN OVERALL HORIZON 2020 FUNDING TO THE EUROPEAN CLOUD INITIATIVE, WITH ESTIMATED ADDITIONAL PUBLIC AND PRIVATE INVESTMENT OF €4.7 BN REQUIRED TO FURTHER DEVELOP THE EUROPEAN DATA INFRASTRUCTURE.

0.2bn Widening the user base to the public and private sectors

1bn EU-wide Quantum technologies flagship

3.5 bn Data Infrastructure

and seamless access to analyse and use research data

and technology

future

es, industry and SMEs

Bringing benefits to citizens

reduce the cost of big data storage and high-performance analysis

#DigitalSingleMarket

EUROPEAN CLOUD INITIATIVE
UNLOCKING THE POWER OF BIG DATA FOR OPEN SCIENCE

OPPORTUNITIES

- Spurring new solutions in core areas such as research, transport, environment
- Better public services such as diagnostics
- Better science for complex problems
- Commercial opportunities for innovative companies
- Better value for taxpayers - cutting up data produced by projects funded by the Horizon 2020 research and innovation programme. (Protein, Healthcare, Transport data, Research)

EUROPEAN OPEN SCIENCE CLOUD
BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES

A trusted, open environment for sharing scientific data

- Linking data
- Connecting scientists globally
- Open and seamless services to analyse and reuse research data
- Connecting across borders and scientific disciplines
- Improving science
- Long term and sustainable

EUROPEAN DATA INFRASTRUCTURE
UNLOCKING THE POWER OF BIG DATA, DIGITAL BY DEFAULT

- Facilitate access to and re-use of data for researchers, innovators and public sector
- Work in collaboration with national and regional scientific and public data and computing centres
- Reduce the cost of big data storage and high-performance analysis

1 HIGH PERFORMANCE COMPUTING (HPC)
Increase applications based on HPC technology to support the growth of quantum technologies

2 HIGH PERFORMANCE DATA INFRASTRUCTURE
Storage, dissemination, access and management of big data
European big data centre for high-repacity cloud services

3 CONNECTIVITY
Interconnect high-speed networks and secure connectivity and big data storage for EU-wide HPC access
Integration of European public services networks

WIDENING ACCESS; BUILDING TRUST

Widening the user base to the public and private sectors

A safe, secure environment with privacy and data protection built on trusted data and intellectual property protected by highest standards of quality, reliability and confidentiality

WHO IS IT FOR?

- 1.7 million researchers
- 70 million professionals in science and technology
- Open up in the future to public services, industry and SMEs

Bringing benefits to citizens

€2 BN IN OVERALL HORIZON 2020 FUNDING TO THE EUROPEAN CLOUD INITIATIVE, WITH ESTIMATED ADDITIONAL PUBLIC AND PRIVATE INVESTMENT OF €4.7 BN REQUIRED TO FURTHER DEVELOP THE EUROPEAN DATA INFRASTRUCTURE

- 0.2bn Widening the user base to the public and private sectors
- 1bn EU-wide Quantum technologies flagship
- 3.5 bn Data Infrastructure

Horizon 2020

EUROPEAN COMMISSION

Metadaten bei Publikationen und Forschungsdaten

Grundsätzlich ist es sinnvoll, zwei Arten von Metadaten klar voneinander zu unterscheiden: bibliographische bzw. administrative Daten sowie inhaltsbeschreibende bzw. fachliche Daten. Erstere geben Informationen zur Verwaltung der Daten, wohingegen letztere einzelne Aspekte oder Datensätze genauer beschreiben oder zusätzliche Informationen zu diesen bieten.

Inhaltsbeschreibende Metadaten sind, je nach Disziplin, sehr verschieden aufgebaut. Mit den bibliographischen Metadaten werden hingegen Informationen zur Entstehung der Gesamtheit der Daten gegeben, die eine allgemeinere Natur haben und weitaus weniger community-spezifisch sind.

Struktur der Veranstaltung



Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

Warum?

- Neue Technologien
- Neue Anforderungen
- Neue kommerzielle Dienste
- Wahrnehmung
- → Krise der Bibliotheken?

| Von | Betreff | Erhalten | Gr... | Katego... |
|-------------------------------------|--|---------------|-------|-----------|
| Datum: Donnerstag | | | | |
| Peeral Nicolas Becker via InetBio | [InetBio] DSpace Anwendertreffen 2013 | Do 19.07.2013 | 20... | |
| Datum: Mittwoch | | | | |
| Marin Bittl via InetBio | [InetBio] Stellenausschreibung Uni Lüneburg: MitarbeiterIn Forschungsdaten -management und Open Access (IG 13 TV-L) | Mi 18.07.2013 | 18... | |
| Datum: Letzte Woche | | | | |
| Hilmar Wassermann via Inet... | [InetBio] e-bib: Heft 2/2013 erschienen | Fr 13.07.2013 | 30... | |
| Sofka, Kerstin via InetBio | [InetBio] 4. RADAR-Workshop: Zusammenfassung und Präsentationen | Do 12.07.2013 | 19... | |
| Birkler, Helena via InetBio | [InetBio] FOSTER Open Science Kurs bis Ende Juli für Kommentar und Feedback verfügbar | Do 12.07.2013 | 22... | |
| Stelaris, Ursula via InetBio | [InetBio] Einladung zur DSpace 4.0.0.1000-Workshop am 25. und 26. Juni 2013 in Karlsruhe | Di 09.07.2013 | 17... | |
| Jacob, Boris via InetBio | [InetBio] Stellenausschreibung Leibniz-Gesellschaft Wiss. Referat/in Open Science und Digitalisierung, Geschäftsstelle Berlin | Di 09.07.2013 | 18... | |
| Petra Ziemannmann via InetBio | [InetBio] Fortbildungsangebot neue Dienstleistungen in wissenschaftlichen Spezialbibliotheken / Forschungsdaten -management | Di 09.07.2013 | 14... | |
| Racum, Matthias via InetBio | [InetBio] Stellenausschreibungen von IIC Karlsruhe - 3 Software-Entwicklerinnen (m/w/d) gesucht | Mo 08.07.2013 | 28... | |
| Datum: Vorletzte Woche | | | | |
| Lutz, Jan via InetBio | [InetBio] Neuer DataBank für mehr Transparenz und Fortschritt in der Wissenschaft | Sa 07.07.2013 | 42... | |
| Walther, Claudia via InetBio | [InetBio] DBN-Pressemitteilung (04.07.2013): Open Science – Thema der DBN-Jahresagung 2013 | Do 05.07.2013 | 21... | |
| Dickel, Julia via InetBio | [InetBio] Erinnerung: Stellenausschreibung Wiss. Bibliothekarin/Bibliothekar in Forschungsdaten , E 13 TV-L, Tilo Hannover, zunächst befristet | Mo 04.07.2013 | 20... | |
| Datum: Vorvorletzte Woche | | | | |
| Durwen, Anna via InetBio | [InetBio] Pressemitteilung: Bundesministerium für Ernährung und Landwirtschaft fördert Projekt mit 28-MED-Beteiligung | Do 28.06.2013 | 28... | |
| Datum: Letzten Monat | | | | |
| Nick, Christa via InetBio | [InetBio] Stellenausschreibung LE Erlangen-Nürnberg, E: 13 TV-L | Fr 22.06.2013 | 23... | |
| Bambay, Denis via InetBio | [InetBio] Stellenausschreibung: Projektkoordinator/in Verbund Forschungsdaten Bildung | Mo 20.06.2013 | 14... | |
| Schulte-Hermann, Stefan via In... | [InetBio] Stellenausschreibung 22 MED | Mo 20.06.2013 | 19... | |
| Toni Becker via InetBio | [InetBio] BA-Studienangabe Bibliothek und digitale Kommunikation an der TH Köln | Di 19.06.2013 | 18... | |
| Roland Brandmann via InetBio | [InetBio] Stellenausschreibung Fothalm | Mo 18.06.2013 | 17... | |
| Peeral Nicolas Becker via InetBio | [InetBio] Call for participations (mit besonderer DSpace Anwendertreffen 2013 | Di 12.06.2013 | 20... | |
| Sofka, Kerstin via InetBio | [InetBio] Einweisung 4. RADAR-Workshop am 25. und 26. Juni 2013 in Karlsruhe - Anmeldefrist endet am 15. Juni | Di 12.06.2013 | 20... | |
| Wolke, Carsten via InetBio | [InetBio] MACCall auf dem 107. Bibliothekstag 2013 in Berlin | Mo 11.06.2013 | 20... | |
| Kerstin Hebig via InetBio | [InetBio] Erinnerung: Stellenausschreibung Forschungsdaten -management an der HU Berlin (Stellenausschreibung) | Mo 11.06.2013 | 21... | |
| Siegrind, Doreen via InetBio | [InetBio] Internationales GO-FAB-Büro eröffnet: Open Science-Projekt erhalten Unterstützung aus Hamburg, London und Paris | Do 07.06.2013 | 28... | |
| Hackenberg, Eva Maria via Inet... | [InetBio] Safe the Data 3.1.13: Vernetzungsworkshop Preatory Publishing bei 28 MED | Mo 06.06.2013 | 22... | |
| Julian Schmitz via InetBio | [InetBio] Neues FDS-Projekt: "Wissenschaften - Interdisziplinär" | Di 05.06.2013 | 22... | |
| Dickel, Julia via InetBio | [InetBio] Stellenausschreibung Wiss. Bibliothekarin/Bibliothekar in Forschungsdaten , E 13 TV-L, Tilo Hannover, zunächst befristet | Di 05.06.2013 | 21... | |
| Katharina Pappenbergler via Ine... | [InetBio] Stellenausschreibung Uni Konstanz: Fachreferat Fachwissenschaft (DIN, TV 13) | Mo 04.06.2013 | 23... | |
| Bene, Helmut Jürgen via InetBio | [InetBio] Bibliothekstag: Virtual Call der Bibliothekarischen Arbeitskreise Praxis- und Sozialwissenschaften | Mo 04.06.2013 | 20... | |
| Kotler-Hadziag, Sabrina via Ine... | [InetBio] Veröffentlichung des neuer Positionspapier | Mo 04.06.2013 | 21... | |
| Datum: Aber | | | | |
| Peeral Nicolas Becker via InetBio | [InetBio] Call for participations (mit besonderer DSpace Anwendertreffen 2013 | Di 29.05.2013 | 21... | |
| Christmann, Diana via InetBio | [InetBio] Bibliothekstag: Virtual Call für wissenschaftliche Bibliothekare:innen in Ausbildung | Di 29.05.2013 | 20... | |
| Racum, Matthias via InetBio | [InetBio] Einladung zum 4. RADAR-Workshop am 25. und 26. Juni 2013 in Karlsruhe | Mo 28.05.2013 | 28... | |
| Kerstin Hebig via InetBio | [InetBio] Stellenausschreibung Forschungsdaten -management an der HU Berlin (Stellenausschreibung) | Fr 25.05.2013 | 18... | |
| Hess, Volker via InetBio | [InetBio] Stellenausschreibung der Universität Siegen: Forschungsdaten -management | Fr 25.05.2013 | 14... | |
| Barbara Bauer via InetBio | [InetBio] Stellenausschreibung UB Bochum: Bibliothekarin im Geschäftsbereich Forschungsinformation (A 9 E 9) - Verlängerung der Bewerbungsfrist | Fr 10.05.2013 | 29... | |
| IBBIO Sekretariat (DU@IBBIO) via... | [InetBio] Stellenausschreibung TU München, MitarbeiterIn Forschungsdaten -management, E13 TV-L, befristet | Do 17.05.2013 | 25... | |
| Mecherintzky, Leifung via Ine... | [InetBio] 107. Bibliothekstag: Themen-Casestudy, Vorträge und Anträge zum praktischen Umgang mit Daten im Berufsalltag | Do 17.05.2013 | 22... | |
| Spletzer, Sina via InetBio | [InetBio] Stellenausschreibung TU Wien Bibliothek: Data Librarian (m, w) (Befristet 20 Wochenstunden/Berufstitel auf 1 Jahr) | Do 17.05.2013 | 24... | |
| Spletzer, Sina via InetBio | [InetBio] Stellenausschreibung TU Wien Bibliothek: Informationsmanager, m/ w für permanente Identifikatoren (RDF) (Vollzeit 40 Wochenstunden/Berufstitel auf 2 Jahre) | Do 17.05.2013 | 25... | |
| Schubert, Anja via InetBio | [InetBio] Stellenausschreibung der Technischen Informationsbibliothek (TIB) Hannover - Projektleiterin/Bibliothekar Forschungsdaten Mobilität und Verkehr (E 13 TV-L) | Mo 16.05.2013 | 26... | |
| Hellke, Theda via InetBio | [InetBio] Stellenausschreibung MitarbeiterIn im Bereich Digitale Forschungsdaten (E 13 TV-L) UB Frankfurt/M. | Di 15.05.2013 | 31... | |
| Carolina Konrad via InetBio | [InetBio] Viro-Geschichten und Forschungsdaten -management: Fortschrittskonferenz beim DSpace | So 13.05.2013 | 32... | |
| Walther, Claudia via InetBio | [InetBio] DBN-Pressemitteilung (03.05.2013): DBN legt Thesen zur Digitalisierung von Forschung und Lehre vor | Do 08.05.2013 | 23... | |
| Kötzing, Ernst (IBB) via InetBio | [InetBio] Stellenausschreibung der KT-Bibliothek, BibliothekarIn / Informationswissenschaftlerin/m gesucht | Mo 07.05.2013 | 20... | |
| Kleider Siegel via InetBio | [InetBio] Bibliothekstag: in-ent (Schwerpunkt Forschungspraxis und Forschungsdaten) | Do 05.05.2013 | 19... | |
| Nick, Christa via InetBio | [InetBio] Stellenausschreibung für Erlangen-Nürnberg, MitarbeiterIn für die Informationsbibliothek, E9 TV-L | Do 05.05.2013 | 23... | |
| Engelhard, Ines via InetBio | [InetBio] Open Science in Berlin und Hamburg - Die Wissenschaft öffnet ihre Tore, Kooperation mit Dr. Christina Sievermeier v. Konstanzer Ochofay auf 05.05.2013 | Do 05.05.2013 | 22... | |
| Sofka, Kerstin via InetBio | [InetBio] See the Data - 4. RADAR-Workshop am 25. und 26. Juni 2013 in Karlsruhe | Di 29.04.2013 | 21... | |
| Katja Friedenbergler via InetBio | [InetBio] Stellenausschreibung Universität Erfurt: Universitätsbibliothek, wissenschaftliches MitarbeiterIn für das Kompetenzzentrum Forschungsdaten -management | Di 24.04.2013 | 25... | |
| Barbara Bauer via InetBio | [InetBio] Stellenausschreibung UB Bochum: Bibliothekarin im Geschäftsbereich Forschungsinformation (A 9 E 9) - Verlängerung der Bewerbungsfrist | Do 19.04.2013 | 29... | |
| via InetBio | [InetBio] Inklusionsprojekte - Bibliotheken als Arbeitgeber | Do 19.04.2013 | 25... | |
| Spletzer, Sina via InetBio | [InetBio] Stellenausschreibung TU Wien Bibliothek: Leitung des Zentrums für Forschungsdaten -management (40 Wochenstunden/Berufstitel auf 2 Jahre) | Do 19.04.2013 | 25... | |

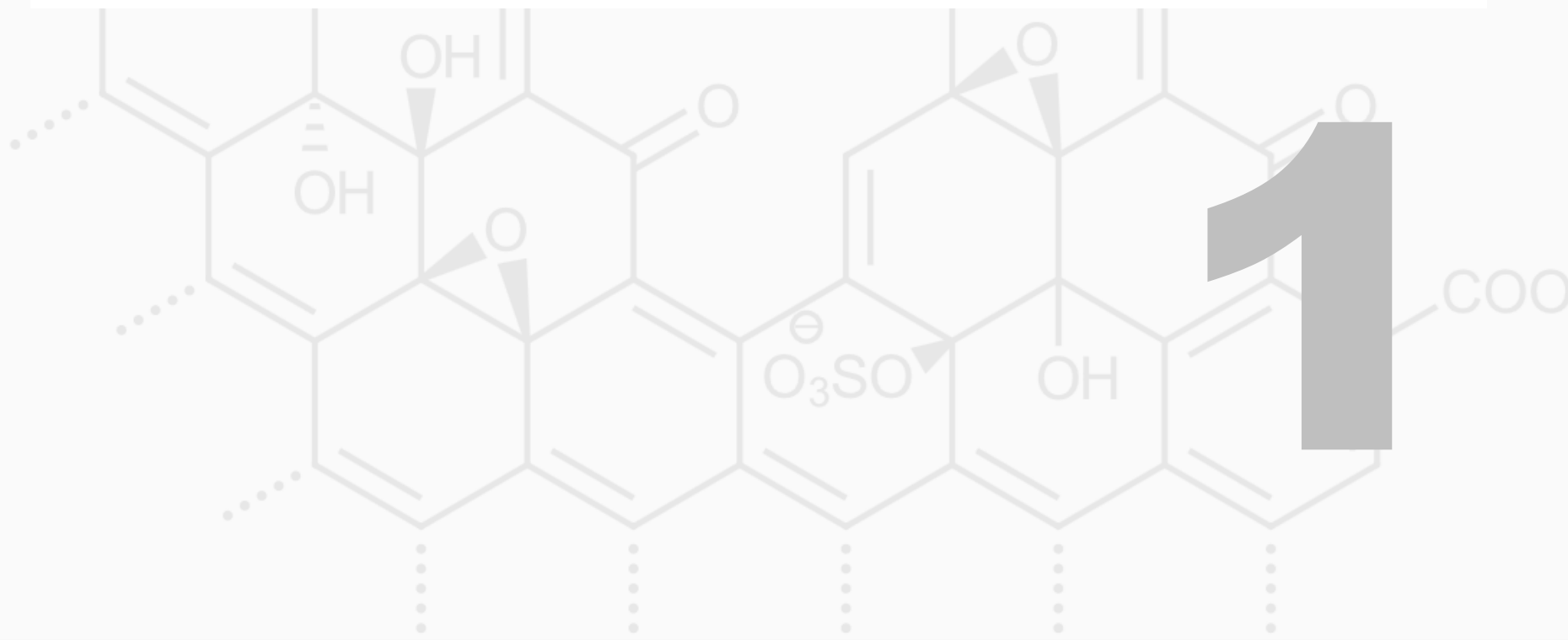
LIBER

- [Ligue des bibliothèques européennes de recherche](#)
- 450 Forschungsbibliotheken
- 54 deutsche Bibliotheken
- München: ~~UB TU~~, BSB, UB UniBwM, MPDL

Ten recommendations for libraries to get started with research data management

[Final report](#) of the LIBER working group on E-
Science / Research Data Management, 2012

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.**



Datenmanagementpläne I

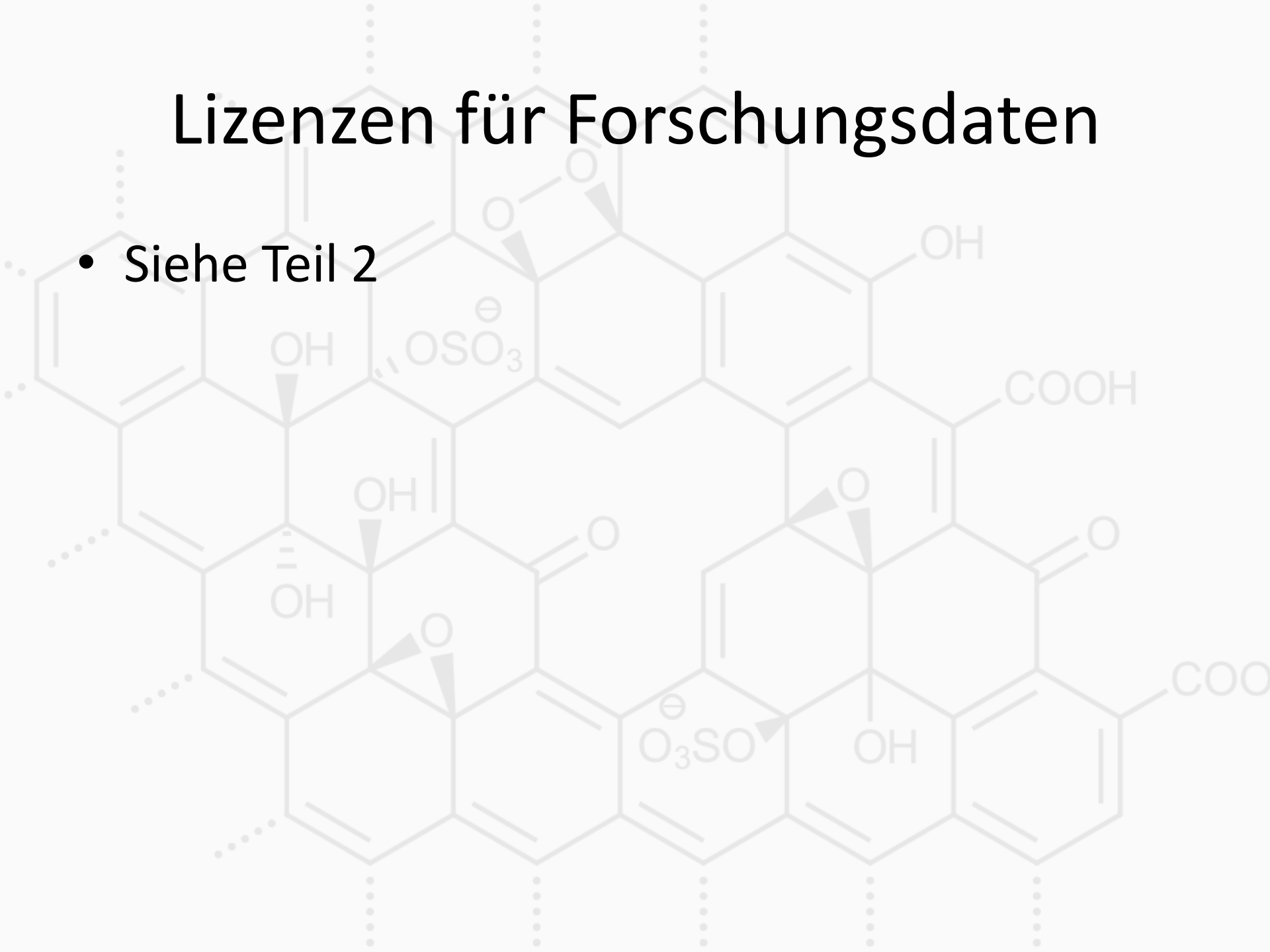
- Welche Daten werden erhoben/verwendet?
- Wie werden die Daten verarbeitet?
- Welche Standards werden benutzt?
- Wie werden die Daten dokumentiert?
- Wo werden die Daten wann (offen) abgelegt?
- Was spricht ggf. gegen eine Veröffentlichung?
- Welche Aufwände entstehen durch das Forschungsdatenmanagement?

Datenmanagementpläne II

- Im Vorfeld vs. während eines Projekts
- [DFG](#)
- [EU \(Beispiele\)](#)
- [DMPTool](#), [DMPOnline](#), [RDMO](#)
- [Beispiele](#)
- [Science Europe](#)

Lizenzen für Forschungsdaten

- Siehe Teil 2



Kurse für Forschungsdatenmanagement

- <https://www.google.com/search?q=semester+2019+forschungsdatenmanagement>
- HfÖD
- [TUM](#)
- [LMU](#)

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

- Beispiel [RDA](#), [DC](#)
- Beispiel [DDI](#)
- Metadaten-Services:
 - Bibliographische Metadaten
 - Transformationen
 - Kontrollierte Vokabulare

2

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

[Google](#)

Hapke, H. (2016) Data Librarian: Das moderne Berufsbild; [b.i.t. online](#)

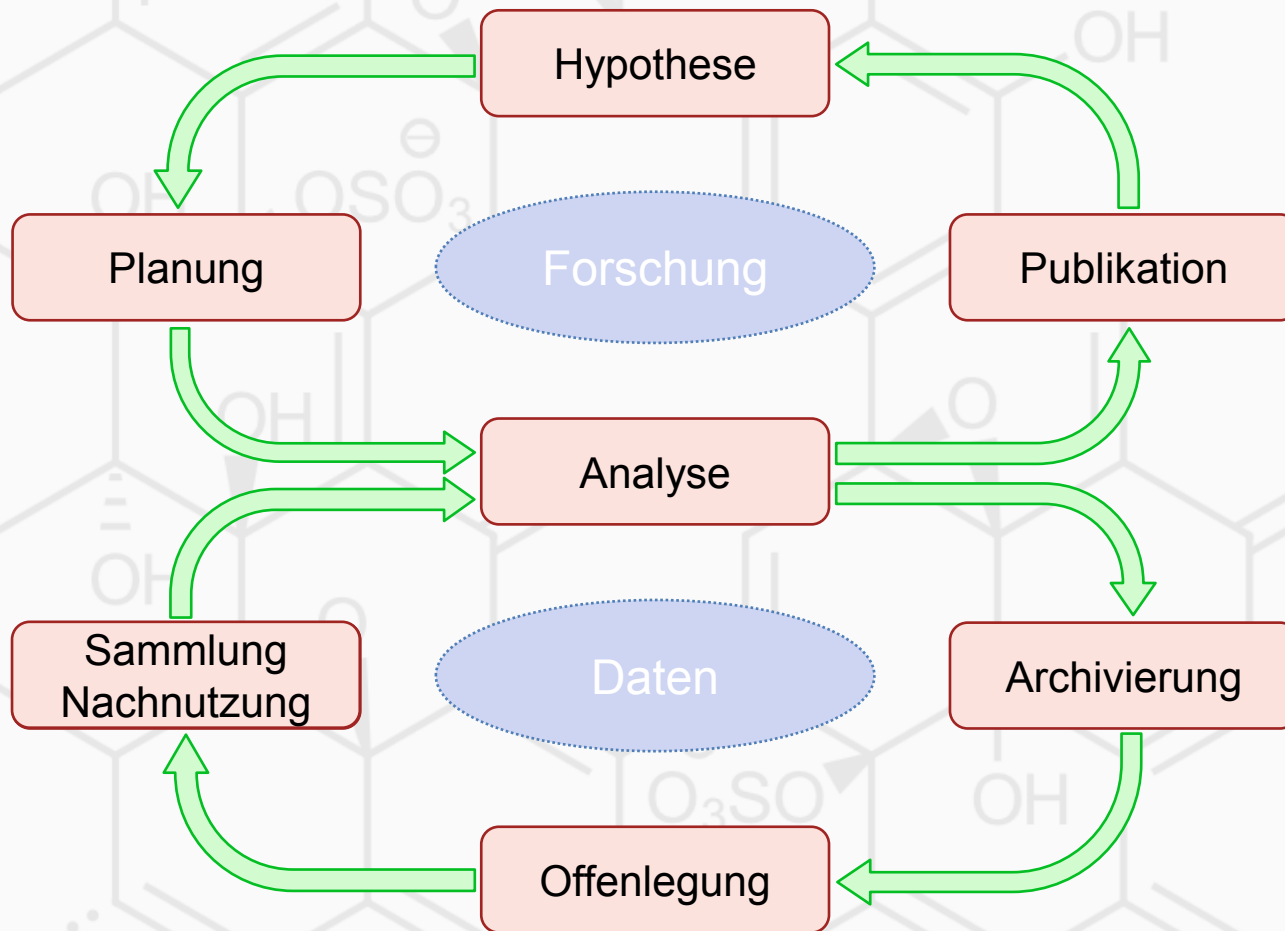
3

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**.



4

Forschungs- und Datenzyklus



Leitlinien für FDM

- Who owns the data?
- What Requirements are Imposed By Others?
- Which Data Should Be Retained?
- For How Long Should Data Be Maintained?
- How Should Digital Data Be Preserved?
- Are there Ethical Considerations?
- How are Data Accessed?
- How Open Should the Data Be?
- How Will Costs Be Managed?
- What are the Alternatives to Local Data Management?

Forschungsdaten-Policies

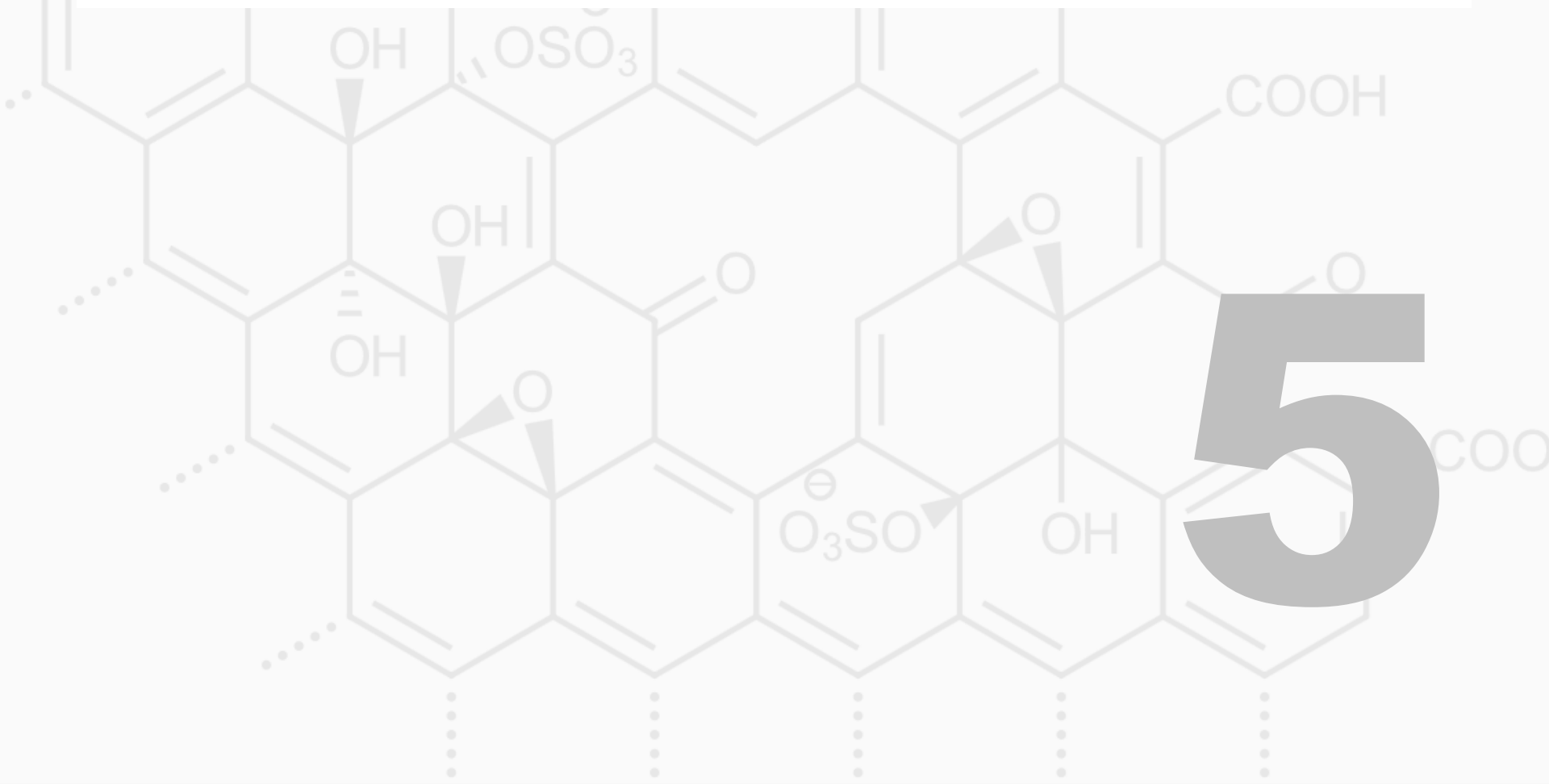
- [Universität Bielefeld](#) (2011)
- [HU Berlin](#) (2014)
- [Universität Leipzig](#) (2019)



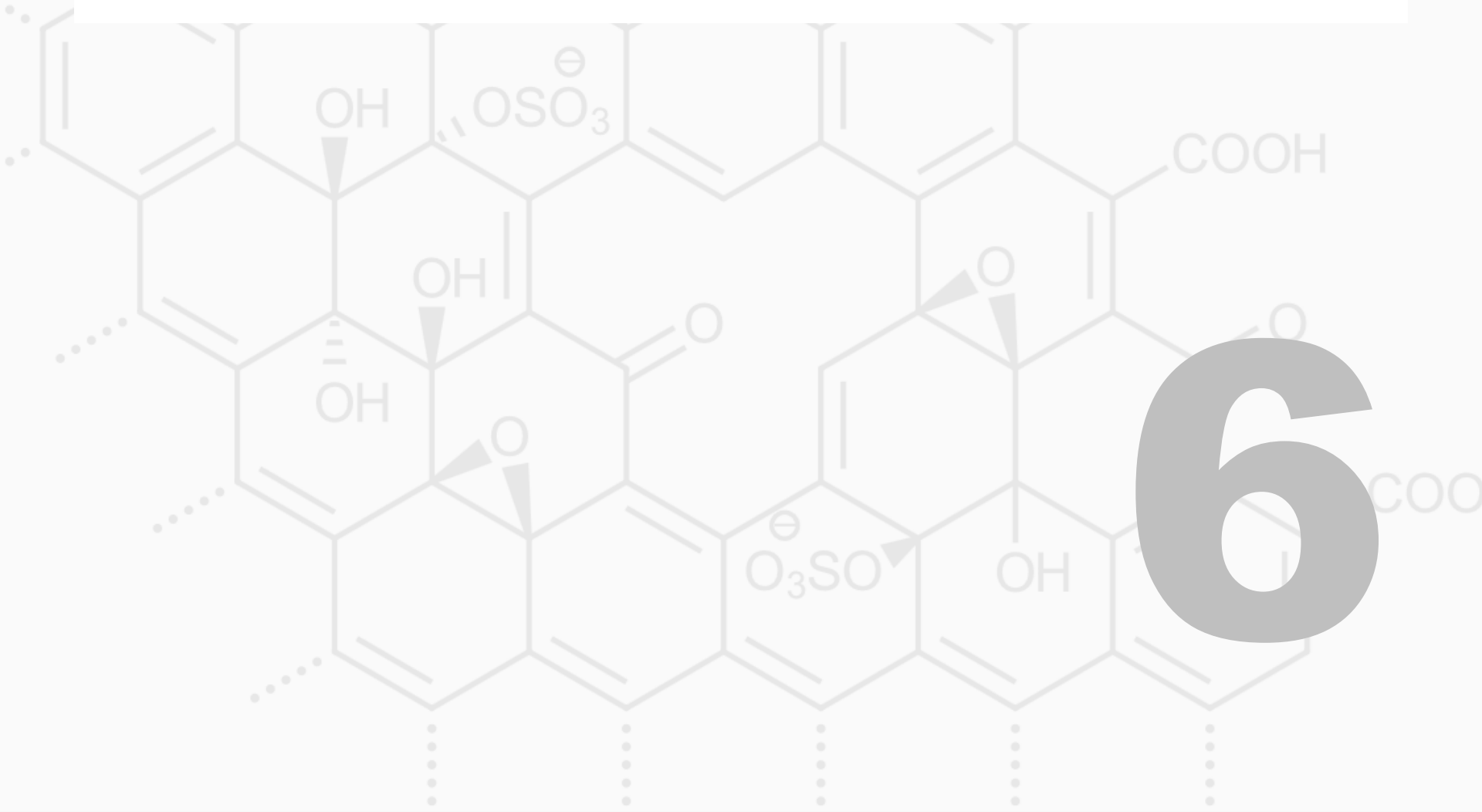
Liaise and partner with researchers, research groups, data archives and data centers to **foster an interoperable infrastructure for data access, discovery and data sharing.**



5



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



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

- URNs → [DNB](#) – Deutsche Nationalbibliothek
- Handles → [GWDG](#) – Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen
- DOIs
 - [TIB](#) – Technische Informationsbibliothek Hannover
 - [DaRa](#) – Registrierungsagentur für Sozial- und Wirtschaftsdaten

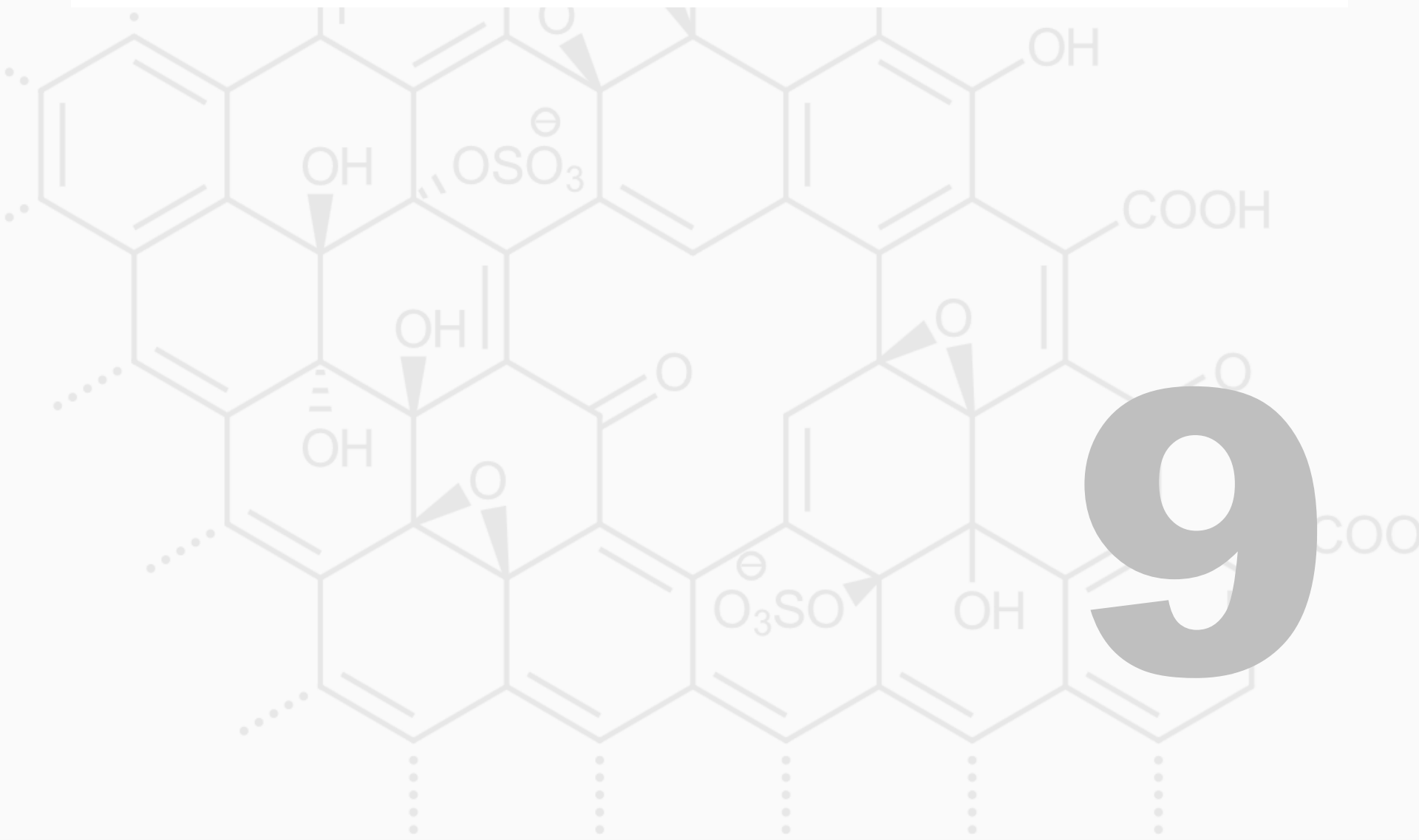


Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.

- [Re3Data](#)

8

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



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.**



10

Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

Wissenschaftler

- Erheben Daten
- Verwenden Daten
- Nutzen Daten nach
- Archivieren Daten?
- Beschreiben Daten?
- Veröffentlichen Daten?

| | | | |
|----------------------------------|------------------------------------|--------------------------------------|------------------------|
| There's no API | We will get too many enquiries | I don't mind, but someone else might | Thieves will use it |
| People may misinterpret the data | It's too complicated | Data Protection | Terrorists will use it |
| It's not very interesting | We might want to use it in a paper | It's too big | We'll get spam |
| Lawyers want a custom License | What if we want to sell it later | There's already a project to... | Poor Quality |

"Open Data Excuse" Bingo, JISC

https://docs.google.com/document/d/1nDtHpnIDTY_G32EMJniXaOGBufjHCCk4VC9WGO7jK4

<https://www.open3.at/2013/05/einwaende-gegen-die-veroeffentlichung-von-open-data-und-bewaehrte-antworten-darauf>

➔ Motivation?

Spieltheorie

Exkurs: Gefangenendilemma

„Wenn wir dichthalten kommen wir mit zwei Jahren davon“

„Wenn Sie auspacken und ihr Komplize schweigt, machen wir Sie zum Kronzeugen und nach einem Jahr sind Sie draußen“

„Wenn Sie leugnen und ihr Komplize packt aus, dann wandern Sie für vier Jahre ein“

„Wenn Sie und ihr Komplize reden, dann erhalten Sie ihre gerechte Strafe von drei Jahren“

Spieltheorie

Exkurs: Gefangenendilemma

| | | B | |
|----------|----------|---------------------------------|-------------------------------|
| | | C _{Cooperation} | D _{Defection} |
| A | C | -2 -2 | -4 -1 |
| | D | -1 -4 | -3 -3 |

Spieltheorie

Exkurs: Gefangenendilemma

Pareto-Optimum

A

B

C Cooperation **D** Defection

| | | |
|----------|---------|---------|
| C | -2 -2 | -4 -1 |
| D | -1 -4 | -3 -3 |

Nash-Gleichgewicht

Wissenschaftler vs. Community

| | | Community | |
|-----------------|---|-----------|------------------|
| | | Open Data | Do Not Open Data |
| Wissenschaftler | O | 4 | 4 |
| | N | 5 | 3 |

Payoff matrix for the game between Wissenschaftler (W) and Community (C). The matrix shows the payoffs for each player based on their choice of Open Data (O) or Do Not Open Data (N).

The matrix is structured as follows:

| | | Community | |
|-----------------|---|-----------|------------------|
| | | Open Data | Do Not Open Data |
| Wissenschaftler | O | 4 | 4 |
| | N | 5 | 3 |

The payoffs are: (W, C) = (4, 4) for (O, O), (4, -1) for (O, N), (5, 0) for (N, O), and (3, 0) for (N, N).

Wissenschaftler vs. Community

| | | Community | |
|-----------------|---|-----------|------------------|
| | | Open Data | Do Not Open Data |
| Wissenschaftler | O | 4 4 | -1 1 |
| | N | 5 3 | 0 0 |

Wissenschaftler vs. Community

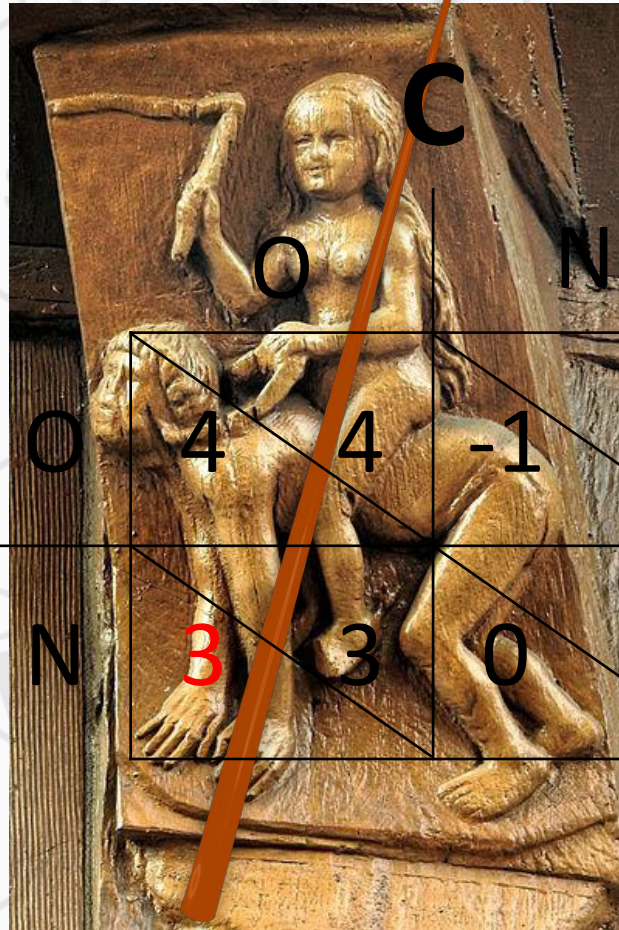
| | | Community | |
|-----------------|---|-----------|------------------|
| | | Open Data | Do Not Open Data |
| Wissenschaftler | O | 4 | 4 |
| | N | 5 | 3 |

A red arrow points from the value 3 in the bottom-right cell of the matrix to the value 0 in the bottom-right cell of the matrix.

Wissenschaftler vs. Community

| | | Community | |
|-----------------|---|-----------|------------------|
| | | Open Data | Do Not Open Data |
| Wissenschaftler | O | 4 4 | -1 1 |
| | N | 5 3 | 0 0 |

Möglichkeit 1: Sanktionen



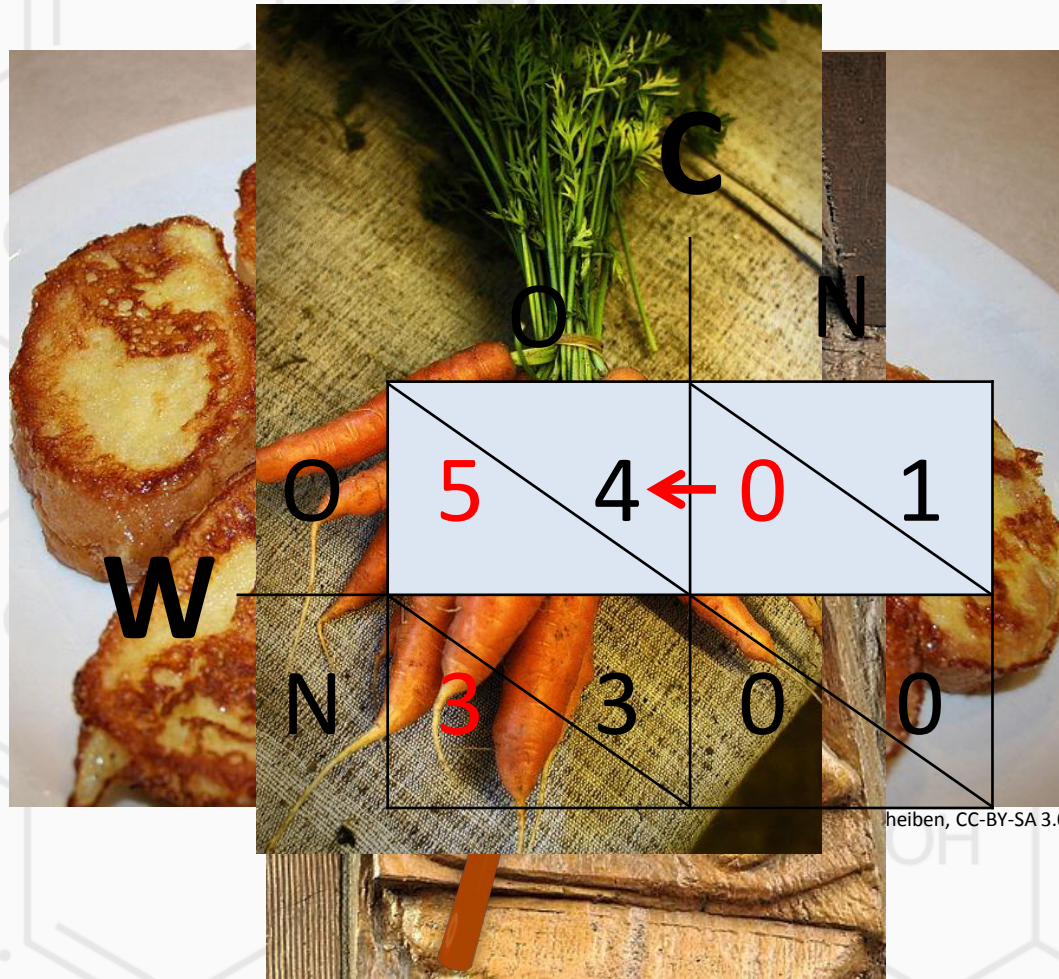
Holger Uwe Schmitt: Der Butterhanne-Comic, CC-BY-SA

Möglichkeit 2: Belohnung



Quettescheiben, CC-BY-SA 3.0

Möglichkeit 3: Kombination



heiben, CC-BY-SA 3.0

Holger Uwe Schmitt: Der Butterhane-Comic, CC-BY-SA

Open Science

- Transparency in experimental methodology, observation, and collection of data.
- Public availability and reusability of scientific data.
- Public accessibility and transparency of scientific communication.
- *Using web-based tools to facilitate scientific collaboration.*
 - ([What, exactly, is Open Science?](#), Dan Gezelter, 2009)
- [LMU Open Science Center](#)

Open Science

The conduction of science in a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, with terms that allow reuse, redistribution and reproduction of the research.

([FOSTER](#))

Open Science

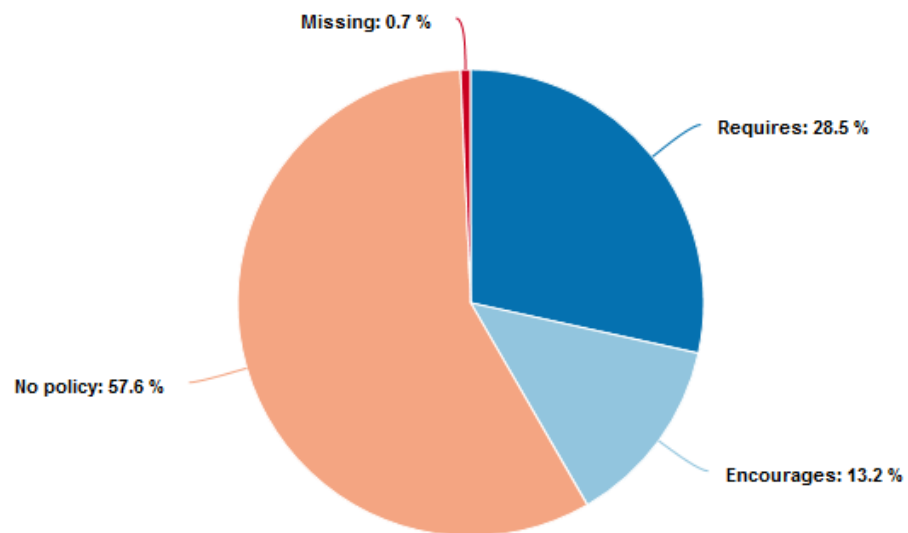
meint

- den öffentlichen Zugang zu Forschungsergebnissen in Form von Publikationen (**Open Access**)
- die freie Verfügbarkeit von Forschungsdaten/Rohdaten (**Open Research Data**)
- die Zugänglichkeit und Vernetzung von öffentlichen Datenbeständen (**Linked Open Data**)
- die Suche nach Alternativen zu den traditionellen Begutachtungsverfahren der Zeitschriften-Verlage, um eine höhere Transparenz in diesem Bereich zu gewährleisten (**Open Review**)
- die Entwicklung und den Einsatz quelloffener Software für die Wissenschaft (**Open Source**)
- die Einbeziehung nicht-professioneller Wissenschaftler in den wissenschaftlichen Prozess (**Citizen Science**)
- freien Zugang zu digitalen Lehrmaterial und (Aufzeichnungen von) Lehrveranstaltungen (**Open Education**)

Open Science Monitor

Number of open data policies, by type of mandate

Source: Sherpa-Juliet - Reference date: April 15th 2018



Science 2.0

- Facebook, ResearchGate
- Twitter
- Altmetrics
- Dropbox

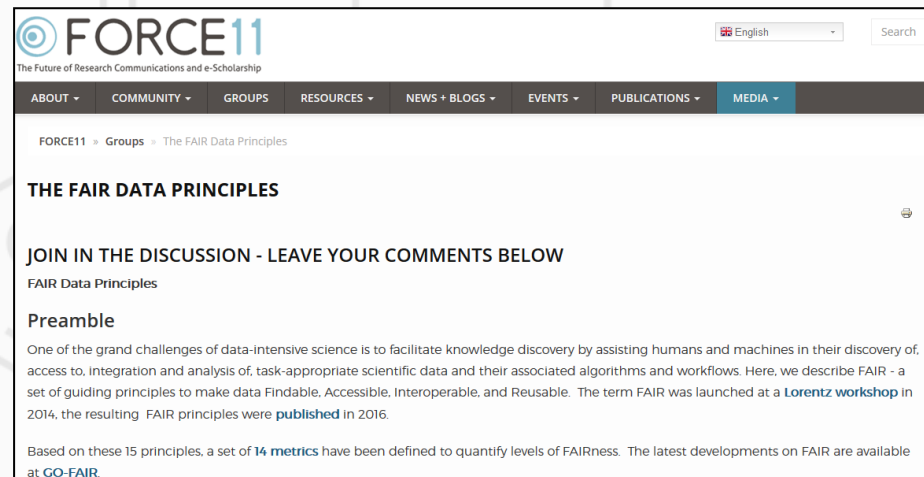
- Blogs
- Wikis

➔ kein soziales Dilemma,

➔ meistens kommerzielle Anbieter

FAIR Data

- Findable – Auffindbar
- Accessible – Zugreifbar (nicht unbedingt offen)
- Interoperable – Interoperabel
- Reusable – Wiederverwendbar



The screenshot shows the FORCE11 website interface. At the top left is the FORCE11 logo with the tagline 'The Future of Research Communications and e-Scholarship'. To the right is a language dropdown menu set to 'English' and a search box. Below the logo is a navigation menu with items: ABOUT, COMMUNITY, GROUPS, RESOURCES, NEWS + BLOGS, EVENTS, PUBLICATIONS, and MEDIA. The main content area shows a breadcrumb trail: FORCE11 > Groups > The FAIR Data Principles. The title 'THE FAIR DATA PRINCIPLES' is displayed in bold. Below the title is a section titled 'JOIN IN THE DISCUSSION - LEAVE YOUR COMMENTS BELOW' with a sub-heading 'FAIR Data Principles'. The 'Preamble' section begins with the text: 'One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery of access to, integration and analysis of, task-appropriate scientific data and their associated algorithms and workflows. Here, we describe FAIR - a set of guiding principles to make data Findable, Accessible, Interoperable, and Reusable. The term FAIR was launched at a Lorentz workshop in 2014, the resulting FAIR principles were published in 2016.' At the bottom, it states: 'Based on these 15 principles, a set of 14 metrics have been defined to quantify levels of FAIRness. The latest developments on FAIR are available at GO-FAIR.'

Big Data/Data-Driven Research

- [Definition](#) (Hasso Plattner, Enzyklopödie der Wirtschaftsinformatik)
- Induktiver Forschungsansatz

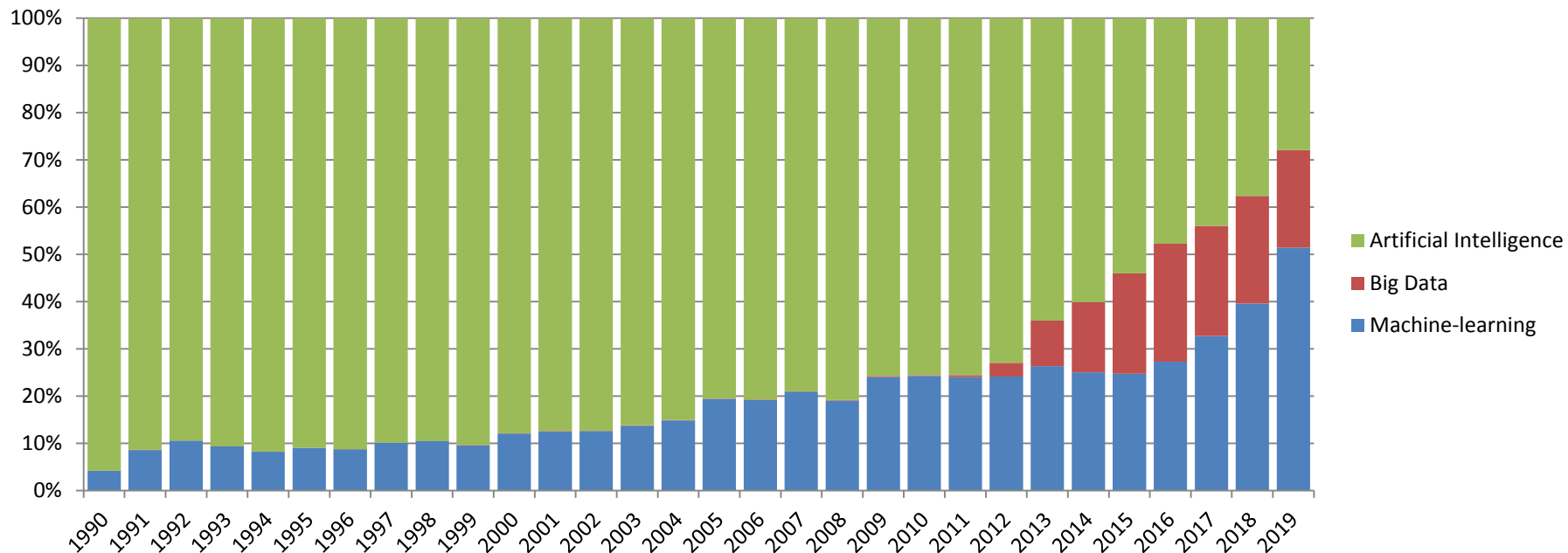


[Kausalitätsproblem](#)

[Statistisches Problem](#)

- Tipp: [Hans Rosling „The Joy of Stats“](#)

Big Data vs. Artificial Intelligence vs. Machine-learning



Open ~~Research~~ Data

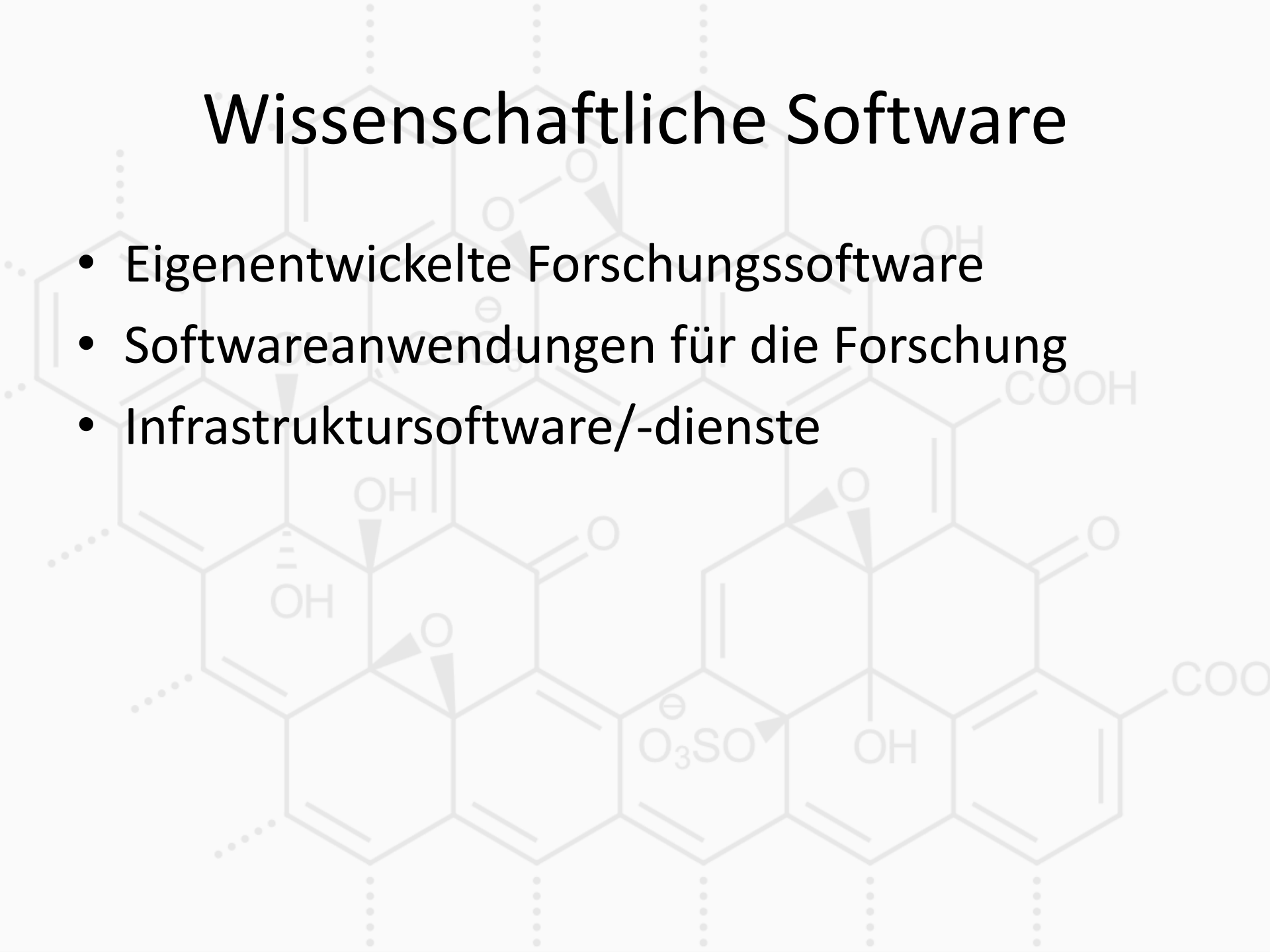
- [Open Data Portal München](#)
- [Open Data Portal der Bundesregierung \(DCAT-AP.de\)](#)
- [Open Data Portal der EU](#)

Forschungsdaten-Discovery

- [Research Data Australia](#)
- [Jisc Research Data Discovery Service \(Großbritannien\)](#)
- [Portage Network \(Kanada\)](#)
- [CESSDA \(Europa\)](#)
- [GeRDI \(Deutschland\)](#)
- Data Citation Index (Clarivate)

Wissenschaftliche Software

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



Wissenschaftliche Software

Gemeinsamkeiten mit 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 mit 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 Know How für ein adäquates Management

Wissenschaftliche Software

Unterschiede 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

Wissenschaftliche Software



ABOUT + COMMUNITY + GROUPS RESOURCES + NEWS + BLOGS + EVENTS + PUBLICATIONS + MEDIA + DONATE

- GROUP MENU
- Group Home
- Members
- Links + Files
- Google Forum
- Subscribe to group

FORCE11 > Groups > Software Citation Implementation Working Group

SOFTWARE CITATION IMPLEMENTATION WORKING GROUP

Description

This group builds on the previous Software Citation Implementation Working Group (https://www.force11.org/group/software-citation-implementation-working-group) and its initial set of software citation principles (https://www.force11.org/group/software-citation-implementation-working-group/principles). The group's activities are: archivists, funders, repository developers, other

1. endorse the principles
2. develop sets of guidelines for implementing the principles
3. help implement the principles
4. test specific implementations of the principles. C from the activities.

Goals

Progress/Activities

Deliverables and Tim

Group Email: softwa

Working Group Participa

Subscribe to group

65 Member(s)

GROUP LEADER
Neil Chue Hong



GROUP LEADER
Martin Fenner



GROUP LEADER
Daniel S. Katz



Ziele

Mitmachen

Termine

Karte

Blog

Verein

RSE4NFDI

deRSE19 Konferenz

Englisch/English

Research Software Engineers (RSEs) - verantwortlich für wissenschaftliche Software

Softwareentwicklung ist ein elementarer, unverzichtbarer Bestandteil der Forschungstätigkeit. Wissenschaftliche Software unterstützt zunehmend die Gewinnung, Verarbeitung und Auswertung von empirischen Daten, aber auch die Modellierung und Simulation von komplexen Prozessen. Damit hat Software einen maßgeblichen Einfluss auf die Qualität der erzielten Forschungsergebnisse. Das Britische Software Sustainability Institut (SSI) hat dafür den Slogan "Better Software - Better Research" geprägt.

Jedoch spiegelt der aktuelle Umgang mit Software und die Bewertung der Software-Entwicklung die Bedeutung dieser Arbeit im Forschungsprozess nicht in angemessener Weise wider. Die Ursachen dafür sind breit gefächert und äußern sich in vielfältigen Befunden, zum Beispiel:

- Fehlende Anerkennung für Software-Entwicklung als wissenschaftliche Leistung
- Fehlende Verankerung im wissenschaftlichen Reputationssystem
- Eingeschränkte Verfügbarkeit und Nutzbarkeit wissenschaftlicher Software
- Mehrfach- und Parallelentwicklung

Software ist heute ein essentielles Werkzeug im Prozess des wissenschaftlichen Arbeitens und des Transfers wissenschaftlicher Resultate. Mit der voranschreitenden Digitalisierung von Forschung und Lehre steigt die Abhängigkeit von Software-Lösungen. Die grundlegenden Prinzipien der „guten wissenschaftliche Praxis“ wie Nachvollziehbarkeit, Reproduzierbarkeit, Transparenz und Qualitätssicherung müssen auch bei der Entwicklung und Nutzung von wissenschaftlicher Software angewandt werden. Darüber hinaus sind Nachnutzbarkeit und Verwertung von Software zentrale Herausforderungen.

Das Steuerungsgremium der Schwerpunktkinitiative hat im Mai 2016 die Einsetzung einer Ad-hoc-AG zum Themenfeld „Zugang zu und Nachnutzung von wissenschaftlicher Software“ beschlossen. Aufgabe dieser AG ist es, das Thema weiter zu erschließen und zu definieren, welche Aufgaben in welchen Formaten und mit welchen voraussichtlichen Ergebnissen konkret bearbeitet werden könnten. Dabei werden u. a. die folgenden Themenfelder bearbeitet: Leitlinien, Infrastrukturen, Qualitätssicherung, Publikationsstrategien, offene Lizenzen für den Quellcode von Software, Organisations- und Geschäftsmodelle und Anreizsysteme.



Software Sustainability Institute