

Forschungsdatenmanagement

Fachhochschule für den öffentlichen Dienst,
WS 2019/2020

Michael Franke, Max Planck Digital Library

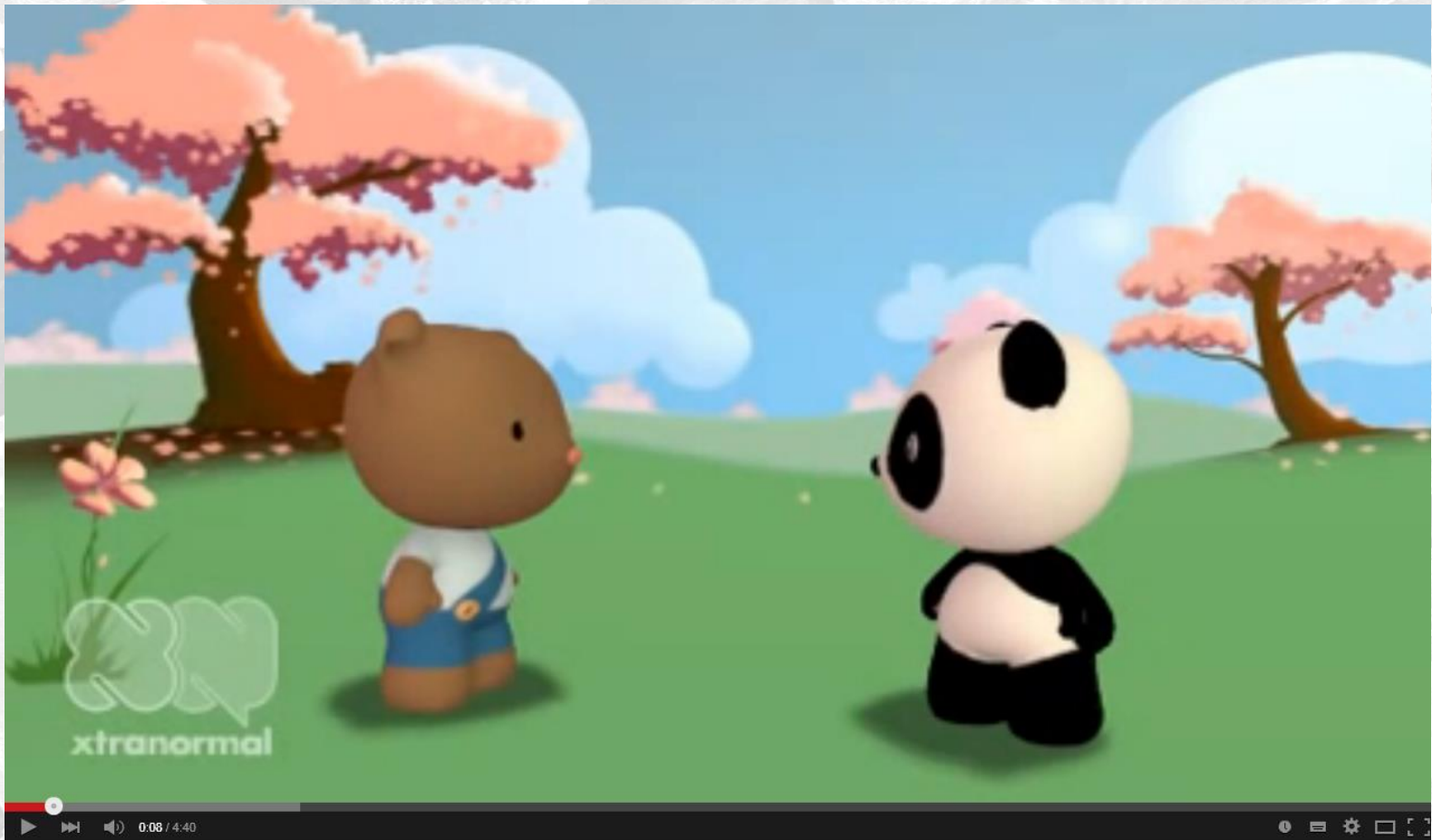
This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



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](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>)



NEU!

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



PhysRevLett.112.241101.pdf

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

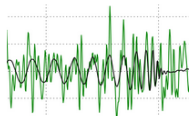
released 2016 June 15:
Event of December 26, 2015: GW151226: total mass 22

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

Genbank

The screenshot shows the NCBI GenBank database entry for the Homo sapiens class II AP endonuclease (APE) gene, partial CDS. The interface includes a search bar at the top, navigation links, and a detailed record for the gene. The record includes fields for 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. On the right side, there are interactive options such as 'Change region shown', 'Customize view', 'Analyze this sequence', 'Articles about the APEX1 gene', and 'Reference sequence information'.

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 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](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3019)
AUTHORS 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)

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 | [Am J Clin Pathol. 2014]
Dysregulation of human apurinic/aprimidinic endonuclease 1 (APE1) | [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/aprimidinic (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, ...

Beispiele für Forschungsdaten

Tierbewegungen

Movebank

The screenshot displays the Movebank website interface. At the top, the navigation bar includes 'Home', 'Tracking Data Map', 'Community', 'Help', 'Tools', 'Env-DATA', and 'Published Data'. The main content area is divided into three sections:

- User login:** A form with fields for 'Username' and 'Password', a 'Log in' button, and links for 'Create new account' and 'Request new password'.
- Search:** A search bar with the text 'stork' and a dropdown menu for 'All Sensor Types'. Below the search bar is a 'Search' button and a checkbox for 'Only studies where I can see data'.
- Search result:** A list of search results sorted by 'Animal Identifier'. The results include various studies such as 'TBMU_Digges_Glchrsr_2013', 'TBMUCOMU GastonMontevicchi.NVAAtlantic', and 'Welfalstrichzug [n=13]'. The 'Welfalstrichzug' entry is highlighted in green.

On the right side of the page, there is a map of Europe showing tracking data points (green circles) and a purple line representing a migration path. The map includes labels for various countries and regions, such as 'Deutschland (Germany)', 'Frankreich (France)', 'Italien (Italy)', and 'Spanien (Spain)'. The map also features a search bar, zoom controls, and a 'Link' button.

At the bottom of the page, there is a 'Drupal' logo and a footer containing 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'.

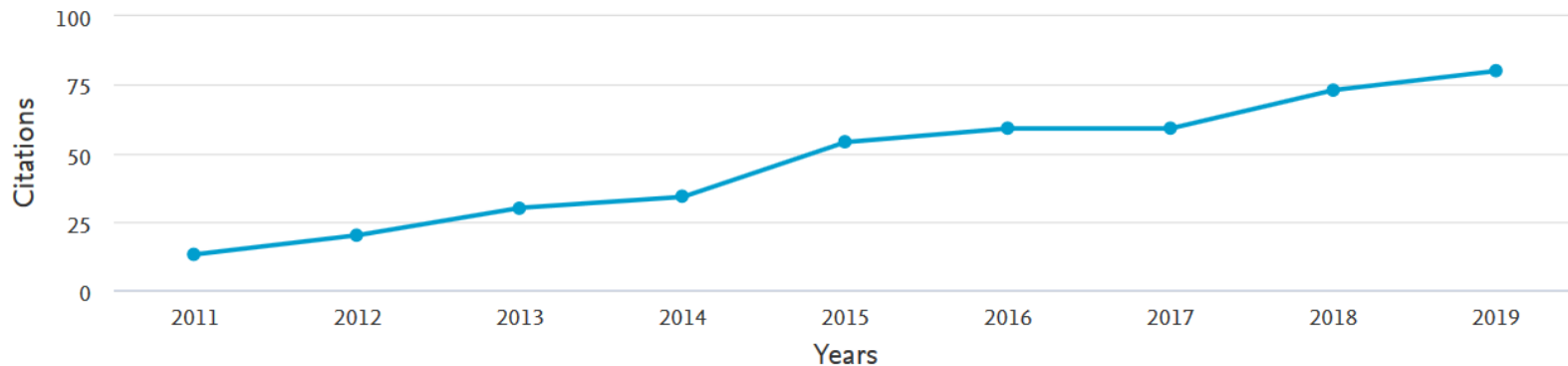
Beispiele für Forschungsdaten

Gesichter/Mimik

FACES

LOGIN REGISTER HELP

422 citations in this date range



72 items

(12)
(12)
(12)
(12)
(12)

(36)
(36)

(24)
(24)



from 200 to 770 go

Picture Set
 a (36)
 b (36)

Faces

FACES—a database of facial expressions in young, middle-aged, and older women and men: Development and validation
(2010) Behavior Research Methods, 42(1), pp. 351-362

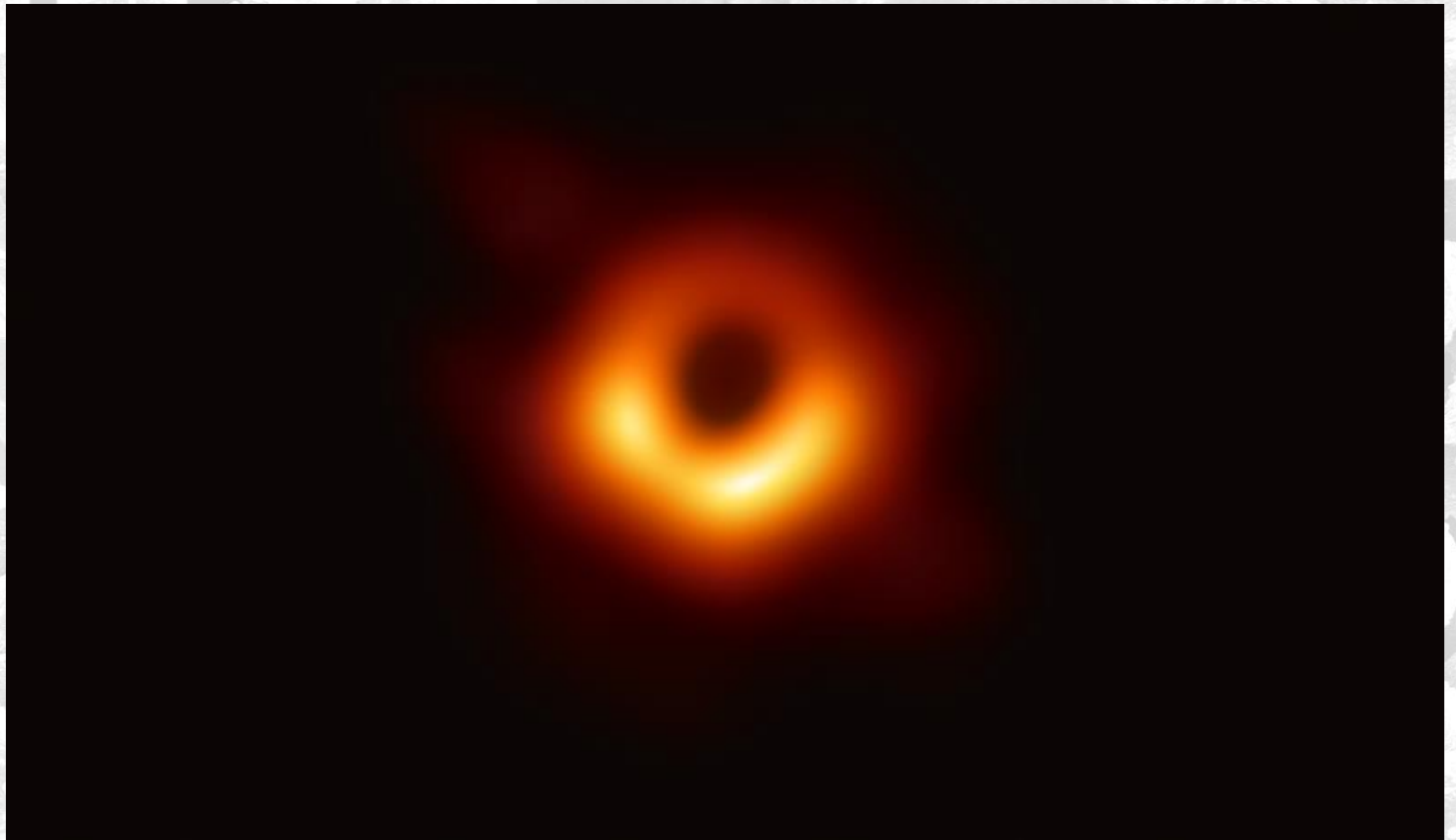
Scopus Metrics

431 Citations

Total number of times this document has been cited in Scopus.

Beispiele für Forschungsdaten

Schwarze Löcher



Definitionsversuche

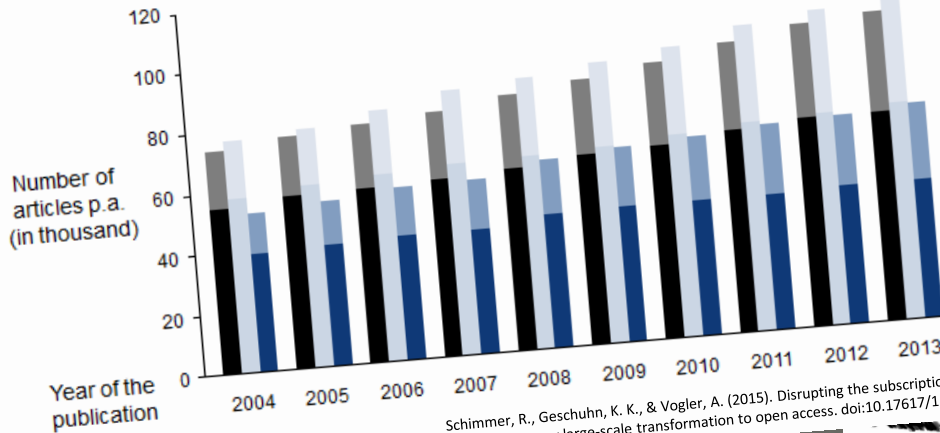
- „Ganz allgemein gesprochen sind Forschungsdaten Daten, die im Forschungsprozess erzeugt, gesammelt oder zusammengestellt werden und auf deren Grundlage wissenschaftliche Hypothesen, Modelle oder Theorien gebildet werden“ (Uni Heidelberg).
- „Forschungsdaten stellen im weitesten Sinne Primärdaten, Sekundäranalysen, 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“)*
- “ (...) information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation” (Europäische Kommission)
- *„(...) researchers know within their projects what research data is“ (Leibniz-Gemeinschaft)*

Worum geht es nicht

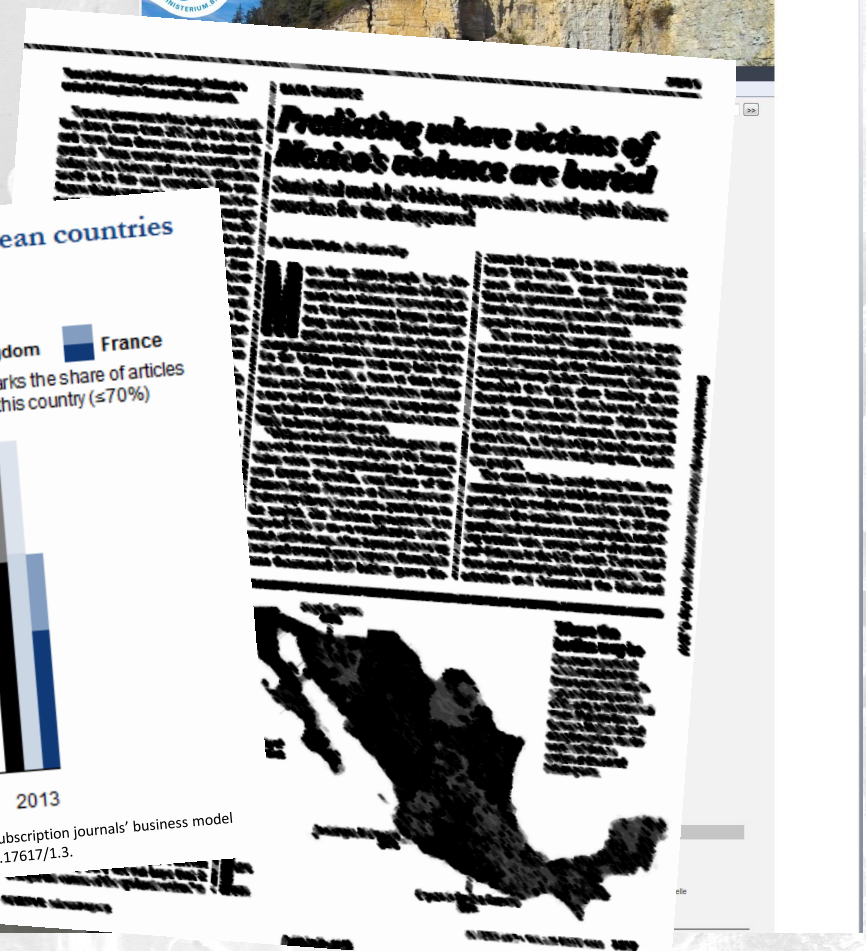
- Physische Daten
- Wissenschaftliche Publikationen
- Daten über Wissenschaft

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

Total publication volume
Articles and reviews in Web of Science



Shimmer, 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

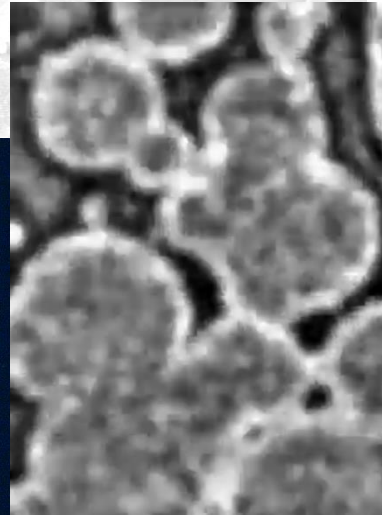
Text (TXT, ODT, PDF, ...)

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Zahlenreihen (CSV, XLSX, ...)

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



```
netcdf tos_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" ;
```

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>

```
Climate
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.

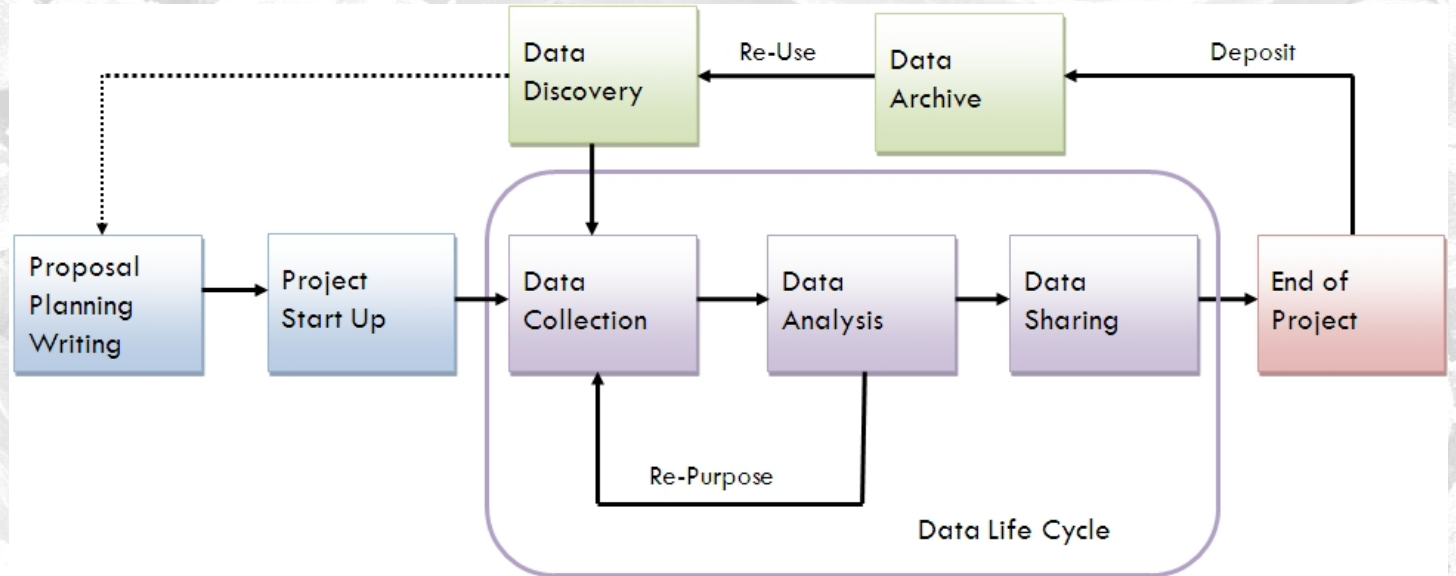
Sekundärdaten: Sekundärdaten seien Daten, die in Prozessschritten aus Primärdaten gewonnen wurden.

Forschungsdatenzyklus



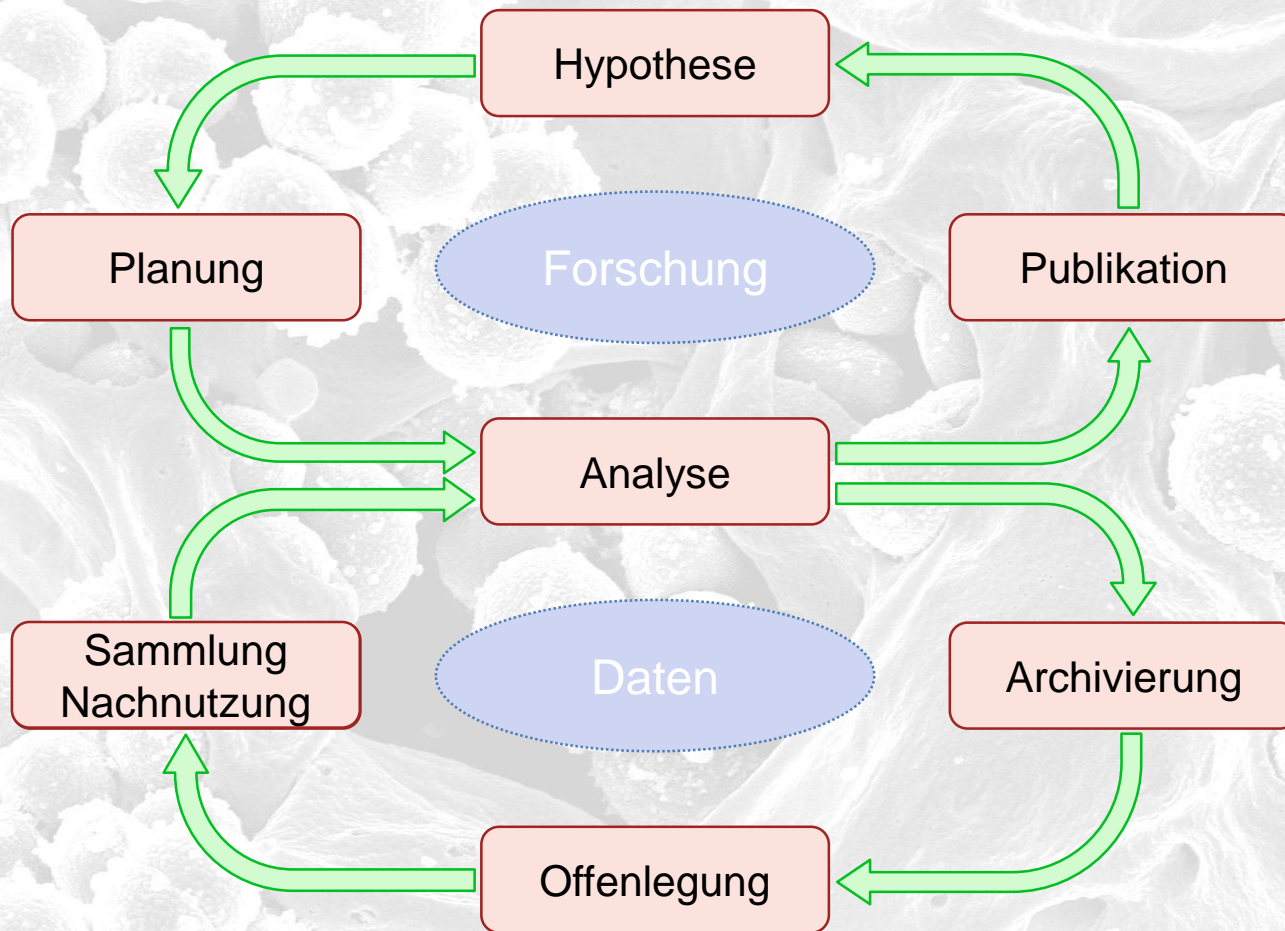
UK Data Archive

Forschungs-Datenzyklus

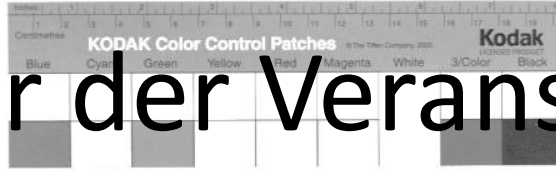


University of Virginia

Forschungs- und Datenzyklus



Struktur der Veranstaltung



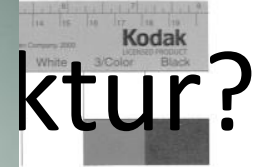
Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

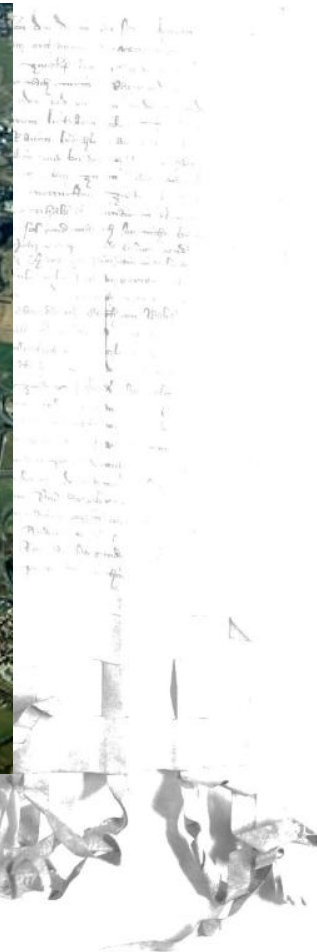




Struktur?

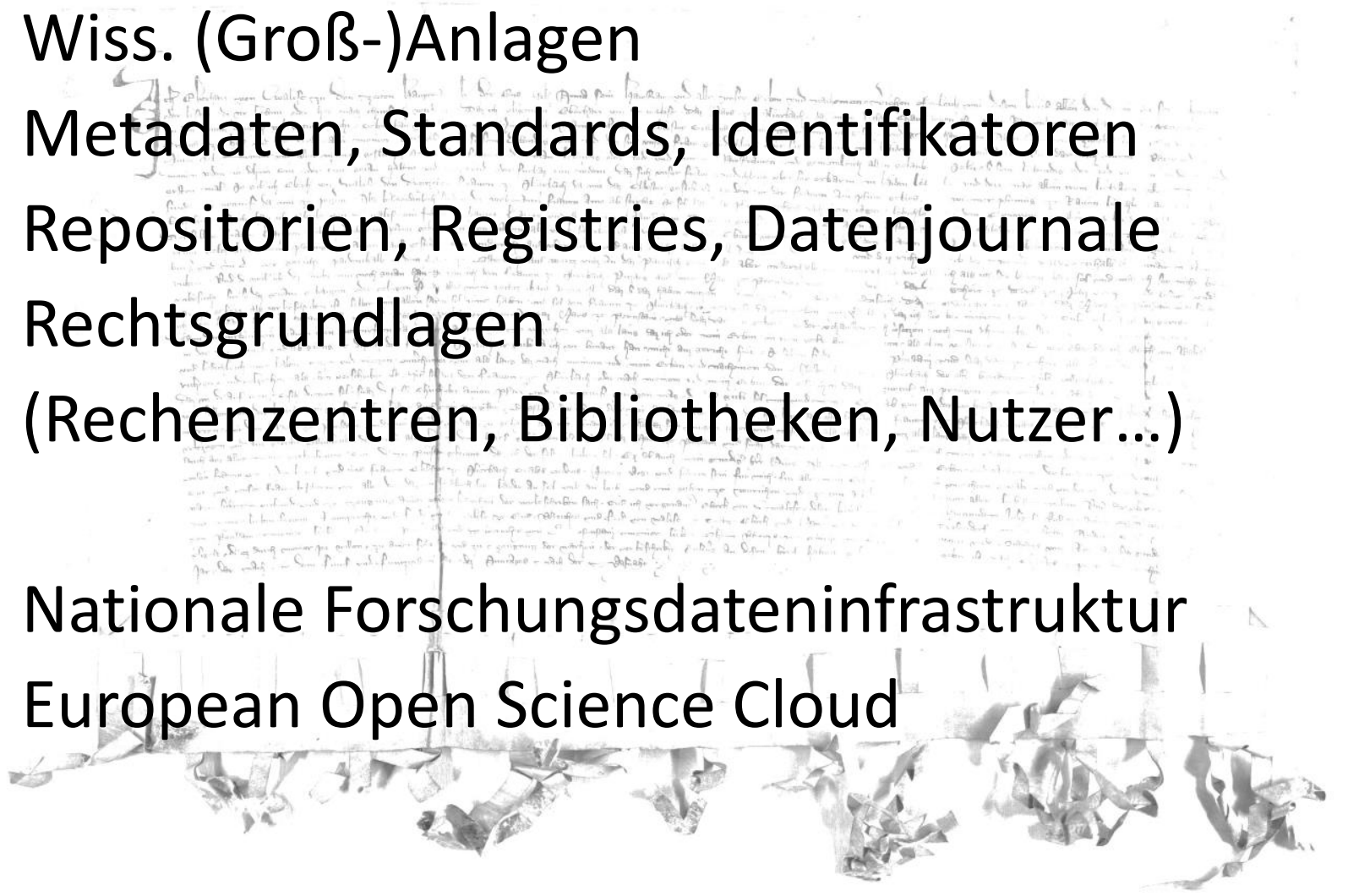
- z.B. Verkehrsinfrastruktur

- Straßen, Brücken
- Schienen, Wasserwege
- Tankstellennetz, Verkehrsfunk, GPS
- Straßenverkehrsordnung, Straßenwacht, ADAC
- Räumdienst, Mautsystem, KEZ-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](#), [Polarstern](#),
[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 RUVINGS: STABLE BEAMS

5.68e+12

IP8: 6.30

Updated: 13:32:01

81	B2
true	true
true	true
false	false
true	true
true	true
true	true
2	ENABLED

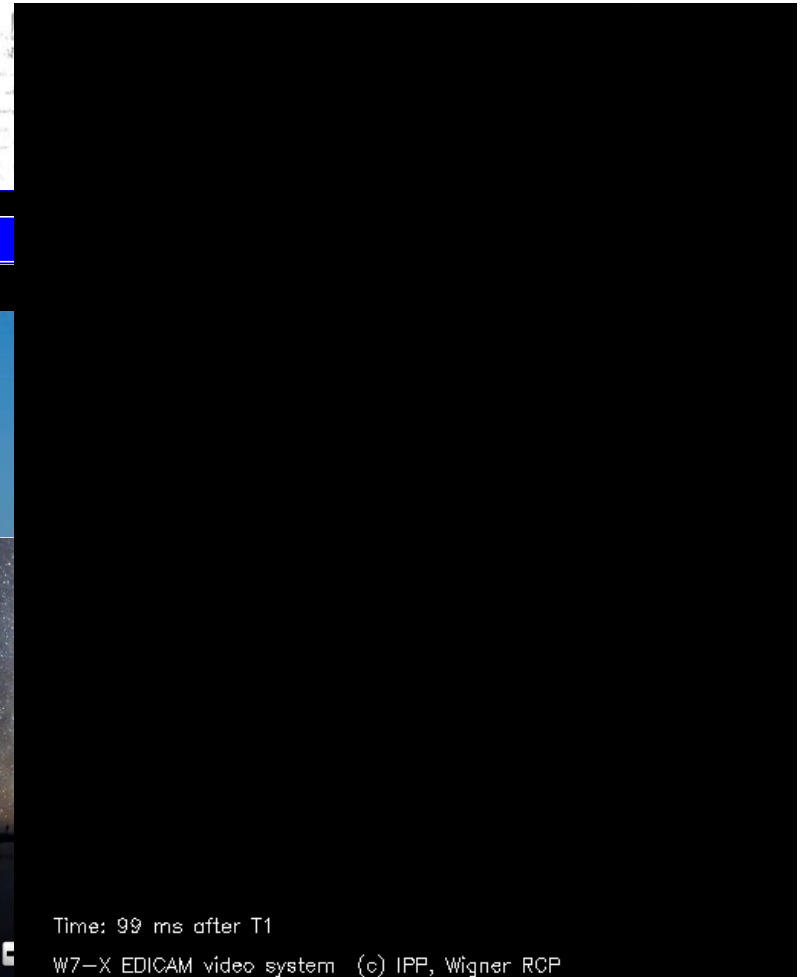
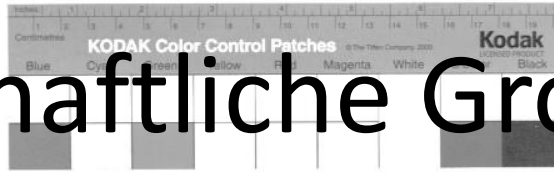
Atacama Desert

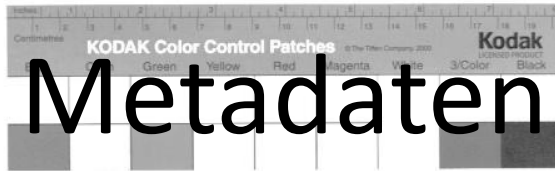
VLT Night 2 - Milky way

08:00 11:00

Wissenschaftliche Großanlagen

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

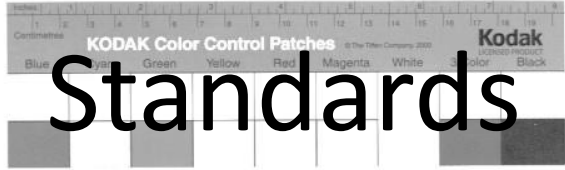




Metadaten

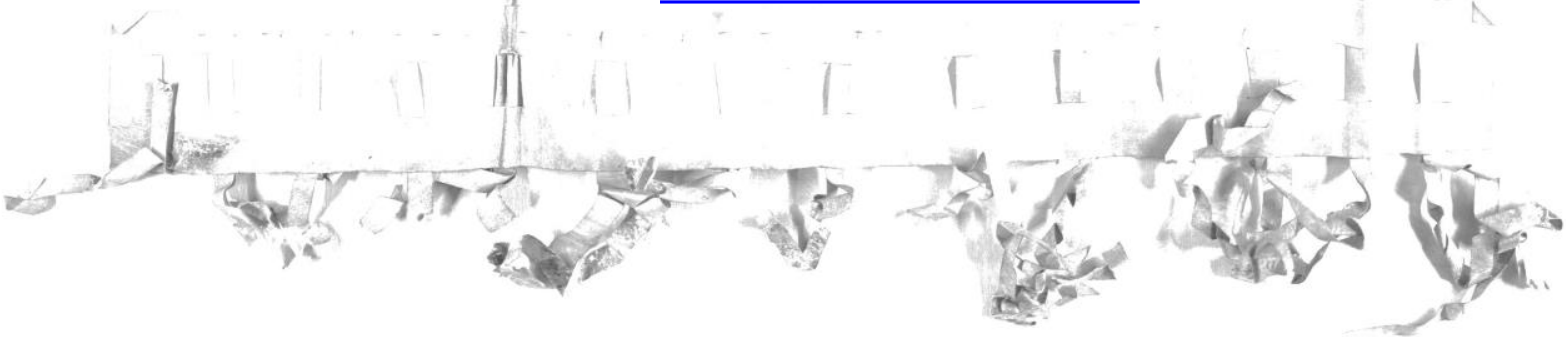
- [DC](#) – Dublin Core Metadata Initiative
- [RDF](#) – Resource Description Framework
([Beispiel](#))
- [SKOS](#) – Simple Knowledge Organization System
- [Metadaten-Standards](#)
- ([Metadaten-Registries](#))





Standards

- 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/wI3XoEDHO8v0ImCa/item/VqNgKUacPrIhqPKu>

- URL/URI

- URN

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

- Handle

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

- DOI

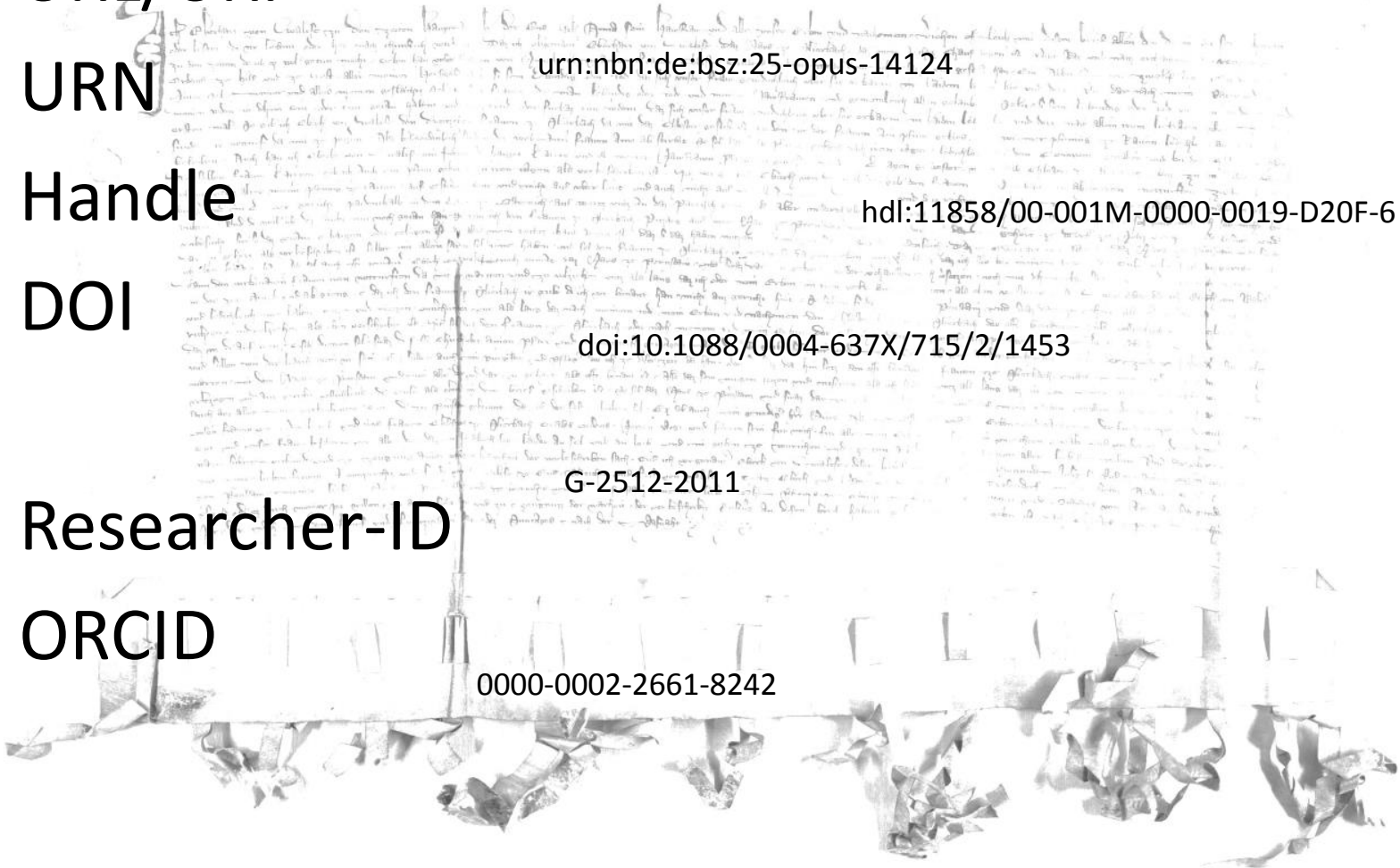
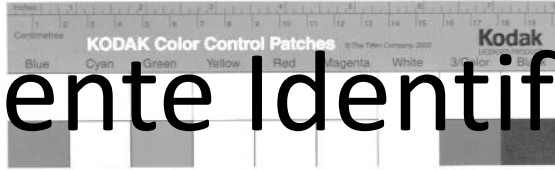
<doi:10.1088/0004-637X/715/2/1453>

- Researcher-ID

<G-2512-2011>

- ORCID

<0000-0002-2661-8242>



Forschungsdatenrepositorien I

Institutionell

Fachlich

Allgemein

[Open Data LMU](#)

[PsychData](#)

[Zenodo](#)

[Edinburgh DataShare](#)

[Pangaea](#)

[Dryad](#)

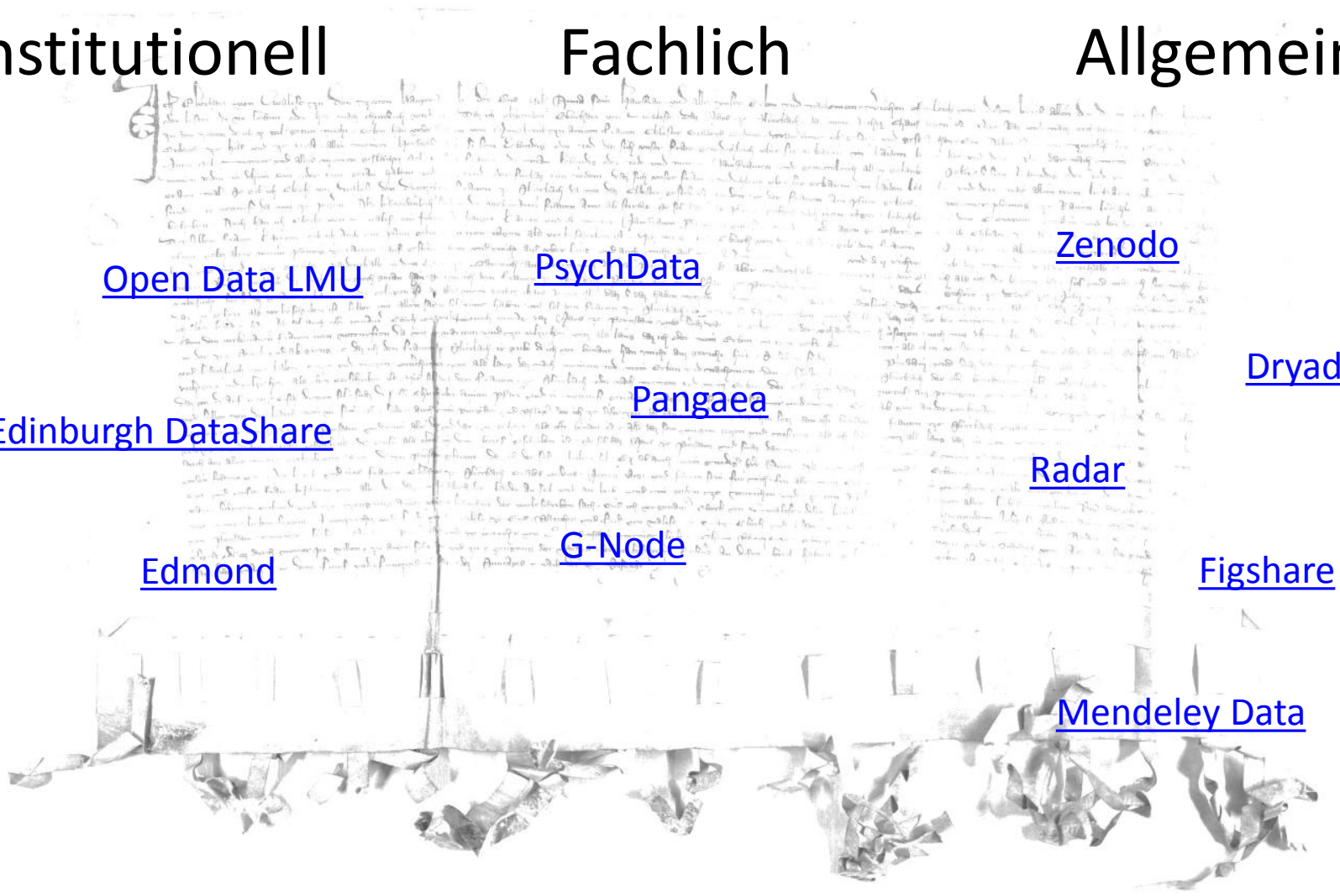
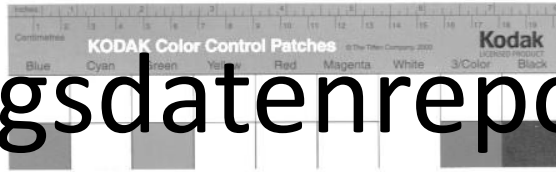
[Edmond](#)

[G-Node](#)

[Radar](#)

[Figshare](#)

[Mendeley Data](#)



A Kodak Color Control Patches ruler is positioned at the top of the slide. It features a ruler scale in centimeters and millimeters, with color patches labeled: Blue, Cyan, Green, Yellow, Red, Magenta, White, 3/Color, and Black. The text 'KODAK Color Control Patches' and '© The Yellow Corporation 2000' is visible on the ruler.

Forschungsdatenrepositorien II

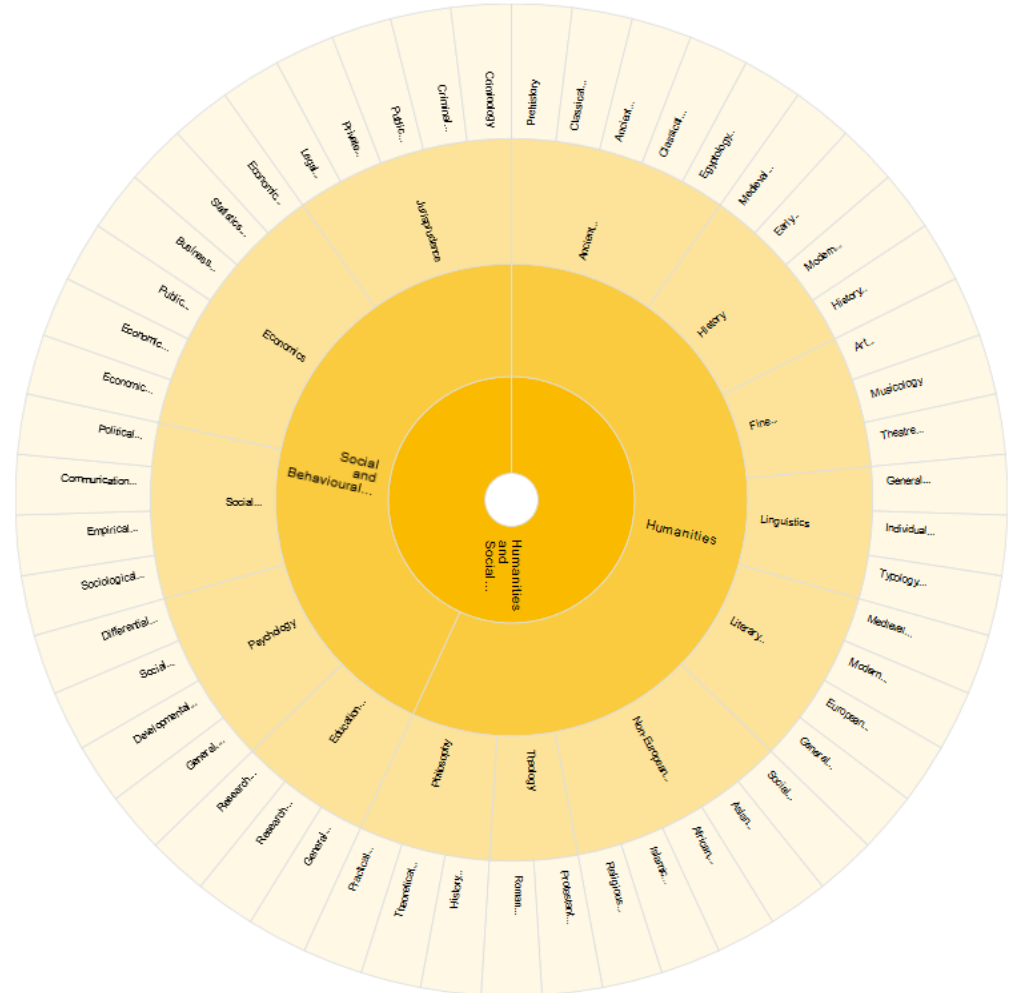
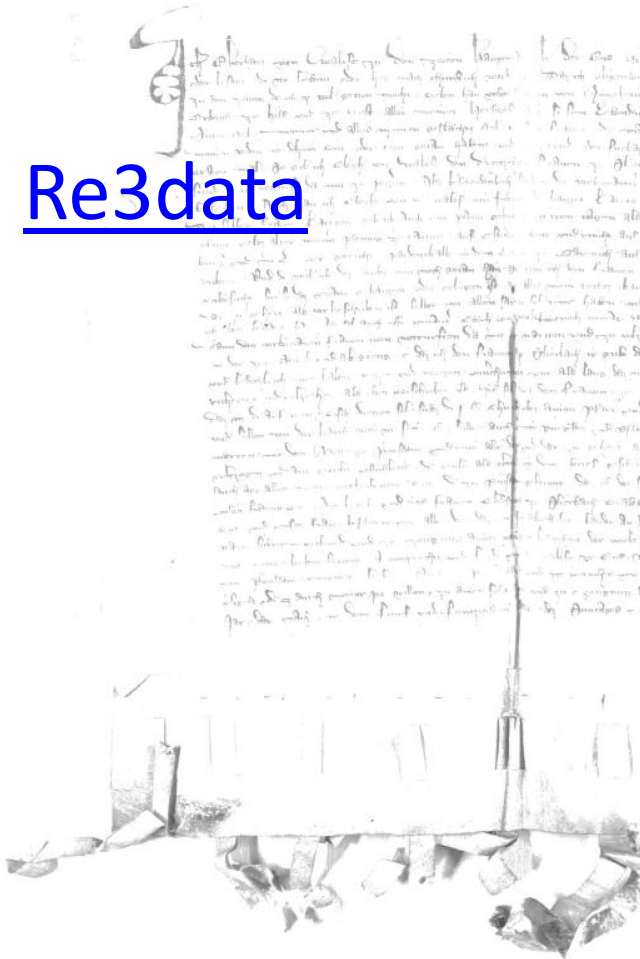
Aufgaben ([Core Trust Seal](#)):

- 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](#)



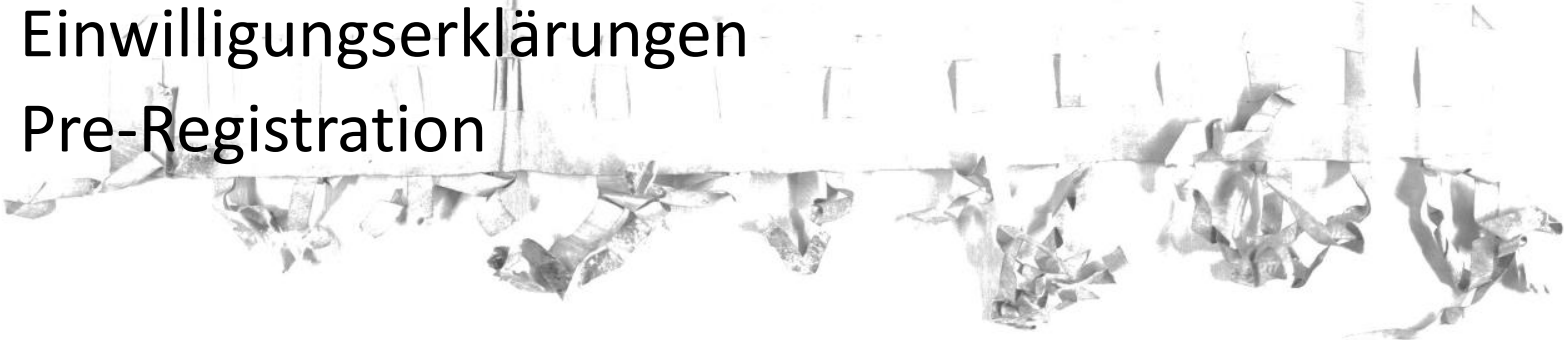


Datenjournale

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



Normative Aspekte

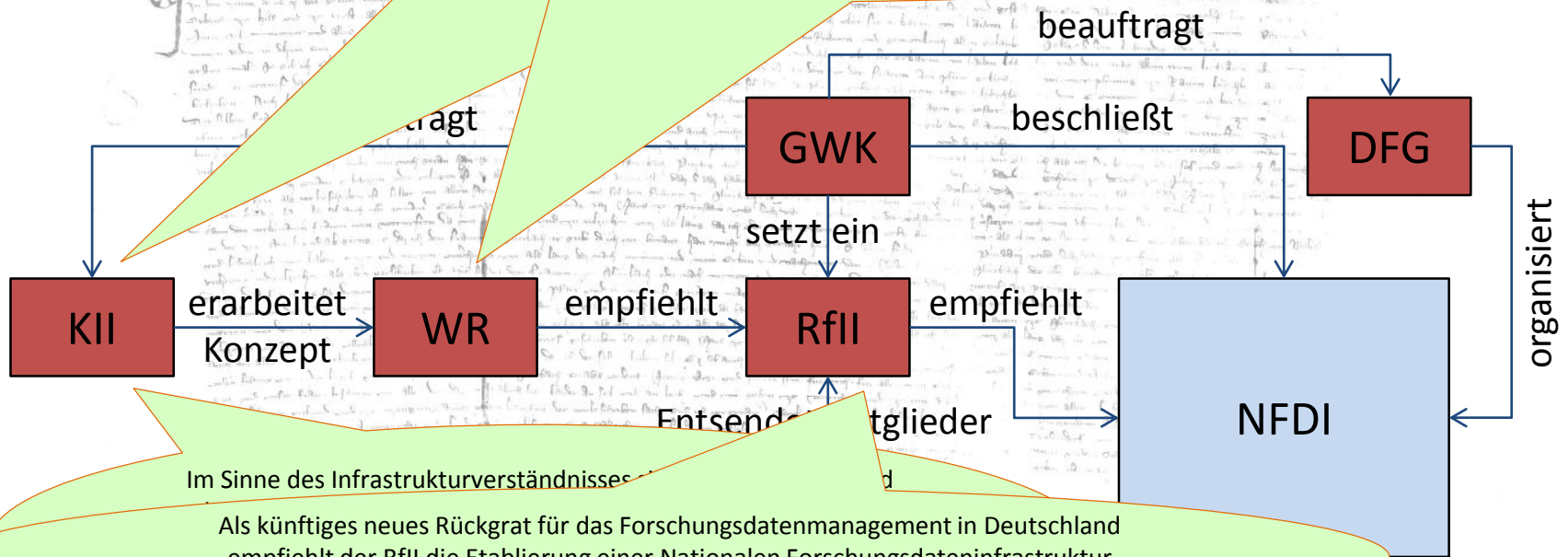
- Grundgesetz, Urheberrecht, Kunsturheberrechtsgesetz, Datenbankrecht, Archivrecht, Datenschutzgrundverordnung, Strafrecht, Kriegswaffenkontrollgesetz, Jugendschutzgesetz
 - PSI-Direktive
 - Gute wissenschaftliche Praxis
 - Lizenzen für Forschungsdaten
 - Einwilligungserklärungen
 - Pre-Registration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

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

A Kodak Color Control Patches strip is positioned at the top center of the slide. It features a ruler at the top with markings in centimeters and millimeters. Below the ruler, the text 'KODAK Color Control Patches' is printed, along with '© The First Company 2000' and the 'Kodak' logo. The strip contains several color patches labeled 'Blue', 'Cyan', 'Green', 'Yellow', 'Red', 'Magenta', 'White', '3/Color', and 'Black'.

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.



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

en and seamless
ces to analyse and
se research data

€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 technology

Bringing benefits to citizens

future
es,
industry and SMEs

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 science for complex problems
- Better public services such as diagnostics
- Commercial opportunities for innovative companies
- Better value for taxpayers - counting up data produced by projects funded by the Horizon 2020 research and innovation programme. (Protein, healthcare, transportation data, bioenergy)

EUROPEAN OPEN SCIENCE CLOUD
BRINGING TOGETHER CURRENT AND FUTURE DATA INFRASTRUCTURES

A trusted, open environment for sharing scientific data

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

EUROPEAN DATA INFRASTRUCTURE
UNLOCKING THE VALUE 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)

relocate applications based on its performance (up to 10x faster)

bring users together to share the power of quantum technologies

2
FOR COMPUTATION-LINE AND SOFTWARE INFRASTRUCTURE

storage, elimination, access and management of big data

European big data centre for high-repacity cloud solutions

3
CONNECTIVITY

interconnect high-speed, reliable and secure connectivity and big data storage for EU-wide HPC access

integration of European public services networks

WIDENING ACCESS; BUILDING TRUST

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

Widening the user base to the public and private sectors

WHO IS IT FOR?

- 1.7 million researchers
- 70 million professionals in science and technology
- Open up to 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

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?

Die Rolle der wissenschaftlichen Bibliotheken Bayerns beim
Forschungsdatenmanagement

- Herausforderungen, Aufgabenfelder, Handlungsempfehlungen -

vorgelegt von der Kommission Virtuelle Bibliothek (KVB) im BVB

LIBER

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

The background of the slide is a grayscale, high-magnification microscopic image of biological tissue. It shows a complex network of cells and fibers. In the foreground, there are several large, roughly spherical cells with a granular or porous internal structure. These cells are interconnected by a dense, fibrous network that forms a mesh-like pattern. The overall appearance is that of a biological structure, possibly a cross-section of a plant or animal tissue, viewed under a scanning electron microscope.

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.**

1

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+2020+forschungsdatenmanagement>
- HfÖD
- [TUM](#)
- [LMU](#)

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

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

2

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

Monster

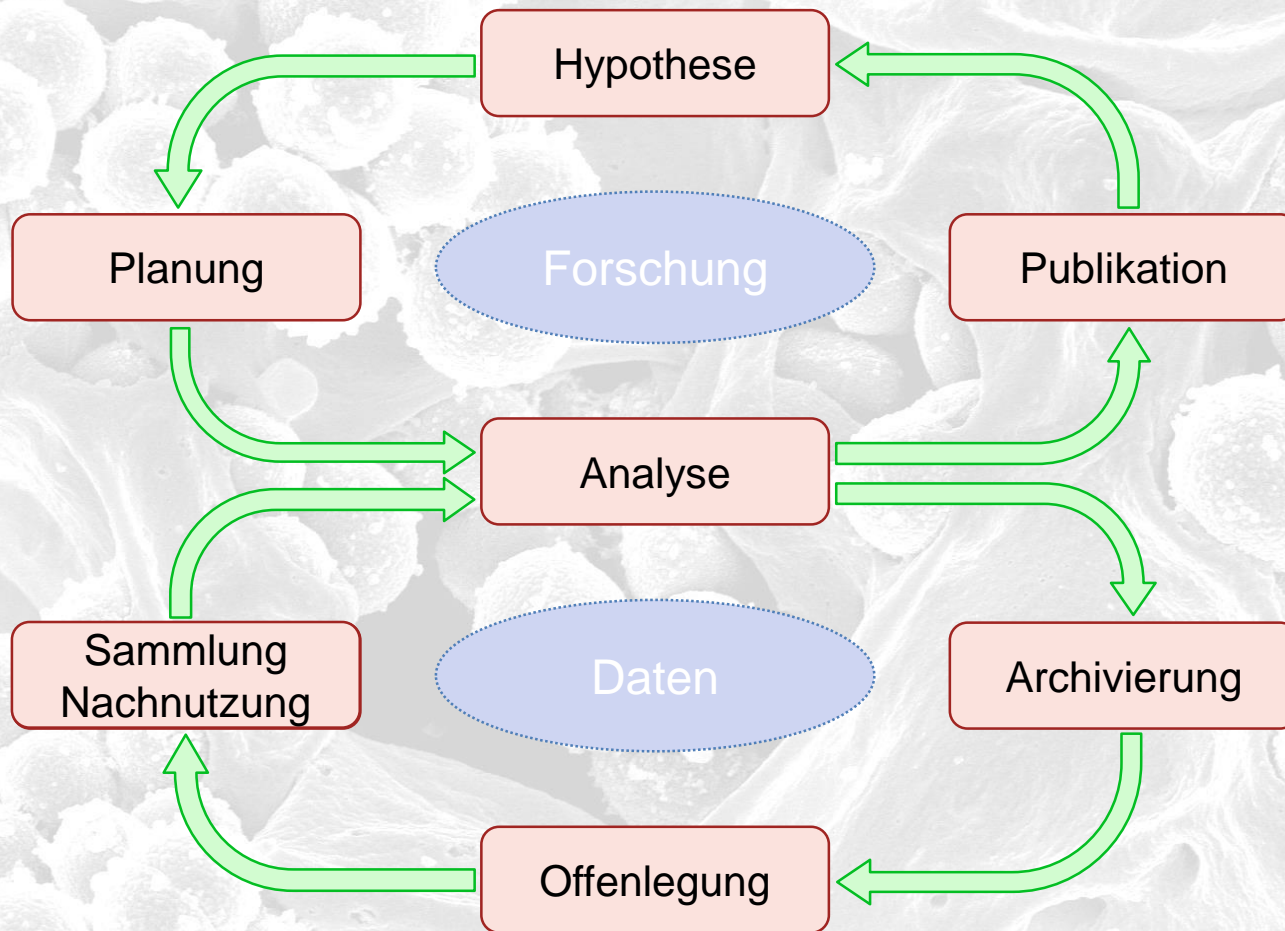
Hapke, H. (2016) Data Librarian: Das moderne Berufsbild; [b.i.t. online](#)

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)
- ...
- [Stiftung Universität Hildesheim](#) (2020)

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.**

6

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.

9

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?

Sorbonne-Deklaration

DIE ZEIT vom 30.01.2020

DIE ZEIT

DIE POSITION

- [Download](#)

Teilt euren Datenschatz!

Forschung lebt vom Austausch. Ihn zu stärken, versprechen Spitzenuniversitäten in der Sorbonne-Erklärung

Von Georg Krausch und Jan Wöpping

Daten sind das neue Öl, heißt es oft. Doch anders als Öl sind Daten erneuerbar, und ihr Umfang wächst täglich. Prognosen zufolge verdoppelt sich das weltweite Datenvolumen sogar alle 18 Monate. Ein riesiger Schatz, dessen Potenzial bisher kaum genutzt wird. Das muss sich ändern. Denn ob bei künstlicher Intelligenz, personalisierter Medizin, Elektromobilität oder industrieller Fertigung: Künftig geht nichts mehr ohne kluge Datenpolitik. Auch Bundesregierung und EU haben das erkannt und arbeiten an entsprechenden Grundsätzen.

Für die Innovationsfähigkeit unserer Gesellschaft sind Forschungsdaten von überragender Bedeutung. Sie bilden die Grundlage für bahnbrechende Entdeckungen in Medizin, Klima- und Materialforschung und eröffnen neue Forschungsfelder im Bereich der Gesundheits-, Sozial- und Kulturwissenschaften. Allerdings nur, wenn die Datenberge für Forschungszwecke zugänglich und nutzbar gemacht werden. Eine solche Öffnung würde die Verbindung von bisher getrennten Wissensfeldern erlauben und gerade in der Zweit- oder Drittverwendung von Daten enormes Innovationspotenzial freisetzen.

Um den Datenaustausch zu befördern, haben Anfang der Woche in Paris neun Verbände international führender Uni-

versitäten die »Sorbonne-Erklärung für offene Forschungsdaten« verabschiedet. Zusammen repräsentieren sie mehr als 160 der weltweit stärksten Forschungsuniversitäten, von Oxford bis Kapstadt, von Paris bis Tokio, von Berlin bis Toronto. Zu ihnen gehört auch der deutsche Verband forschungsstarker Universitäten, die German U15.

Die Sorbonne-Erklärung ist ein starkes Plädoyer für eine globale Kultur frei zugänglicher Forschungsdaten: Diese »sollen, soweit es nur geht, offen geteilt und wiederverwendet werden«. Zugleich hält die Erklärung fest, dass private, personenbezogene Daten geschützt und Urheberrechte respektiert werden müssen.

Was aber braucht es, um eine neue Kultur des Austauschs von Forschungsdaten umfassend und nachhaltig zu etablieren?

1. Die wissenschaftliche Gemeinschaft muss den komplexen Wandel in Richtung offener Forschungsdaten vorantreiben und gestalten. Sie muss auf hoher Datenqualität bestehen. Und sie muss bei der Gestaltung künftiger Clouds und Apps zum Teilen der Daten nicht nur auf technische Machbarkeit, sondern vor allem auf die Nutzerfreundlichkeit achten. Damit sie auch tatsächlich Akzeptanz unter den Wissenschaftlerinnen und Wissenschaftlern finden.

2. Freie Daten gibt es nicht umsonst. Es braucht Investitionen in bestehende und neue Infrastruktur. Vorhaben wie die

Nationale Forschungsdateninfrastruktur und die European Open Science Cloud sind vielversprechende erste Ansätze. Zugleich müssen Politik und Fördermittelgeber aber auch Geld für das Forschungsdatenmanagement bereitstellen. 3. Das Teilen von Forschungsdaten muss politisch durch kluge Regeln und Gesetze unterstützt werden. Dabei ist es besonders wichtig zu verhindern, dass Abhängigkeitsverhältnisse zwischen kommerziellen Anbietern oder Plattformen und Forschungseinrichtungen entstehen.

4. Am wichtigsten ist jedoch: Eine offene Datenkultur braucht das Vertrauen der beteiligten Akteure und einen kooperativen Wettbewerb. Das sind die größten Herausforderungen, nicht die technischen Aspekte. Wenn es um das Teilen von Daten geht, herrscht oft eine Pattsituation, die an das spieltheoretische Gefangenendilemma erinnert: Obwohl alle Akteure von einer Öffnung profitieren würden, überwiegen Skepsis und Misstrauen. Die Sorbonne-Erklärung will auch hier ein Zeichen setzen, indem internationale Spitzenuniversitäten gemeinsam für eine offene Forschungskultur eintreten.

Nach der Berliner Erklärung von 2003, die heute als Meilenstein der Open-Access-Bewegung gilt, ist die Öffnung der Forschungsdaten der nächste Schritt.

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 _{ooperation}	D _{efection}
A	C	-2 / -2	-4 / -1
	D	-1 / -4	-3 / -3

Spieltheorie

Exkurs: Gefangenendilemma

A

Pareto-Optimum

		B	
		C _{ooperation}	D _{efection}
C	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	-1 / 1
	N	5 / 3	0 / 0

Wissenschaftler vs. Community

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

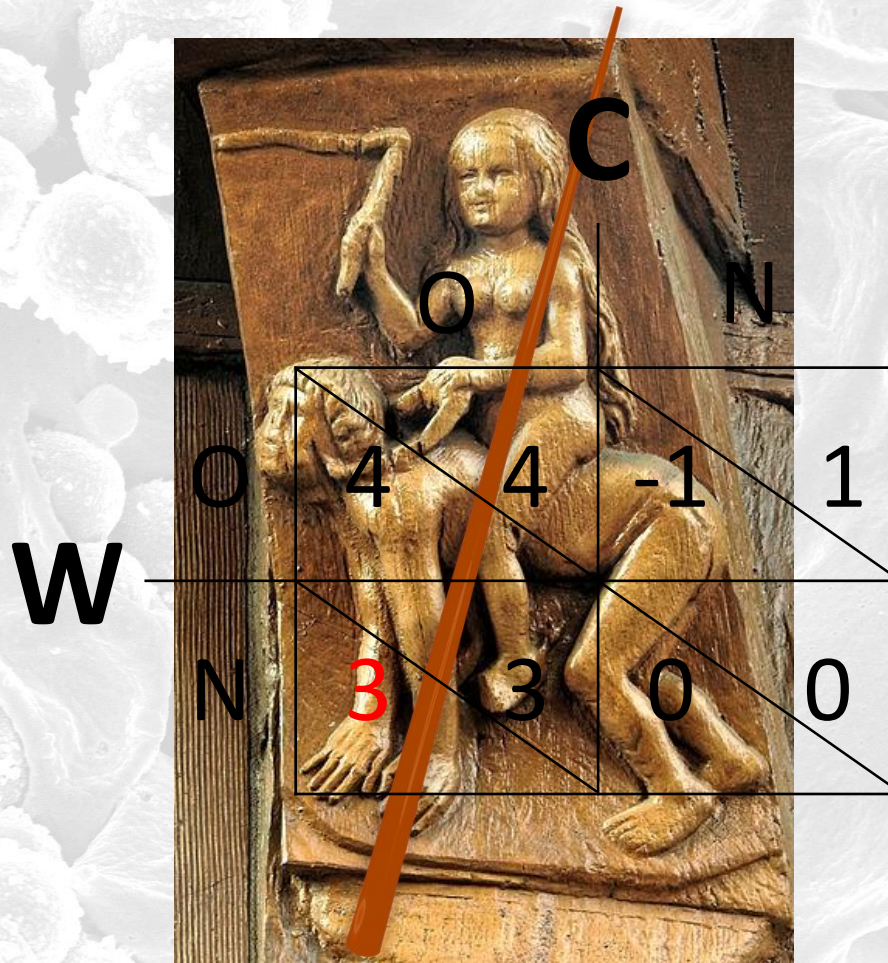
Wissenschaftler vs. Community

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

Wissenschaftler vs. Community

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

Möglichkeit 1: Sanktionen



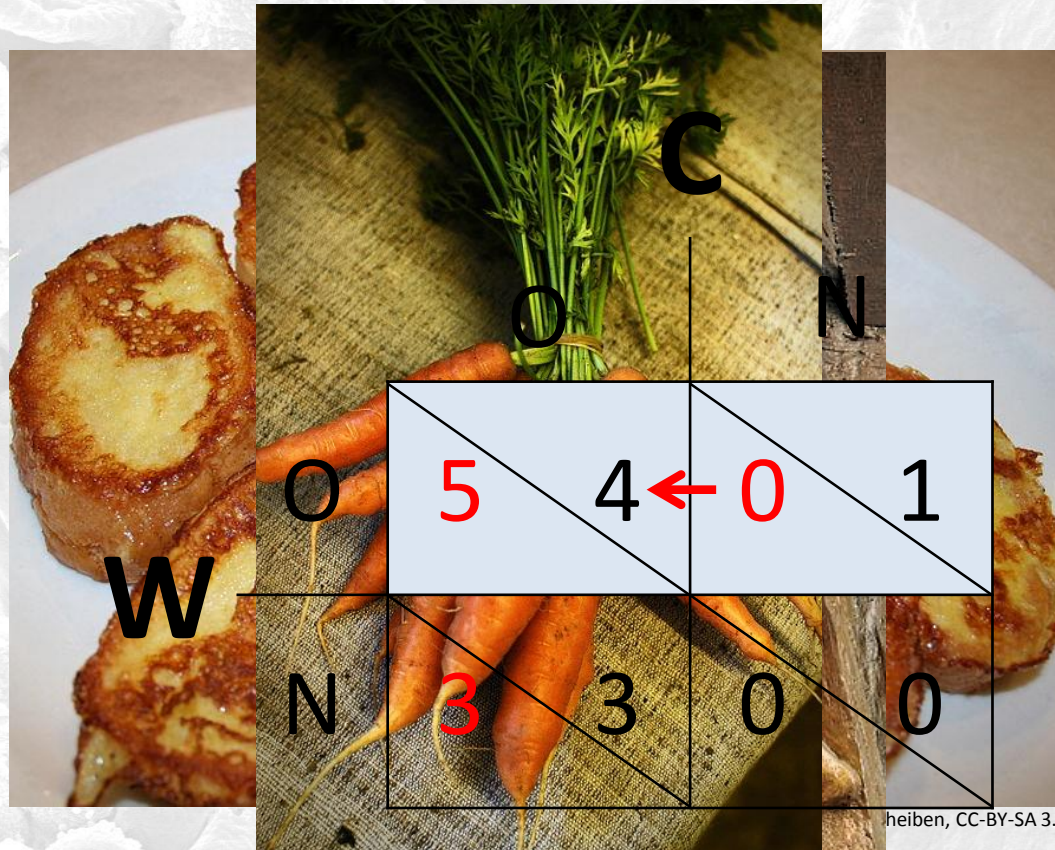
Holger Uwe Schmitt: Der Butterhane-Comic, CC-BY-SA

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)

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

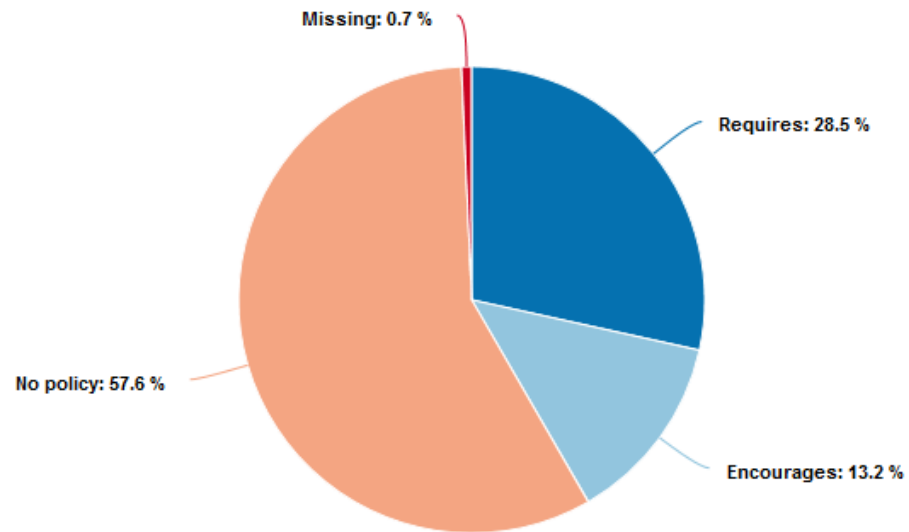
meint

- 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

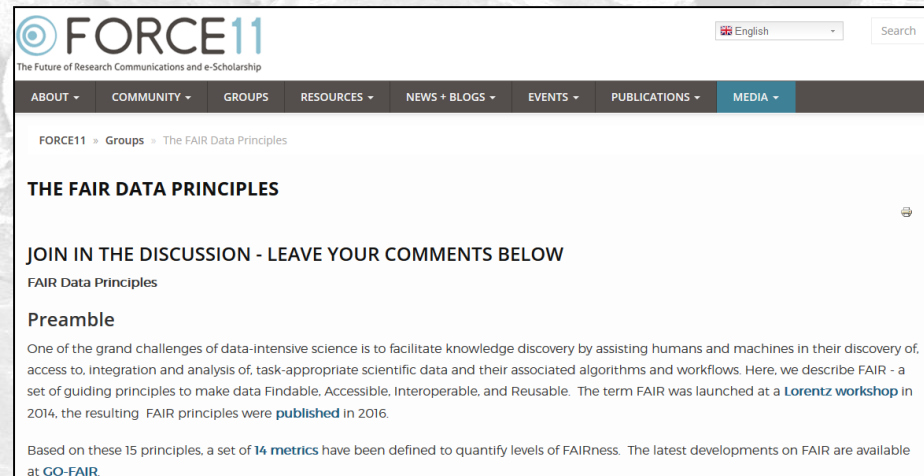
- Blogs
- Wikis

➔ 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 links for ABOUT, COMMUNITY, GROUPS, RESOURCES, NEWS + BLOGS, EVENTS, PUBLICATIONS, and MEDIA. The main content area shows the breadcrumb "FORCE11 » Groups » The FAIR Data Principles" followed by the heading "THE FAIR DATA PRINCIPLES". Below this is a section titled "JOIN IN THE DISCUSSION - LEAVE YOUR COMMENTS BELOW" with the 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

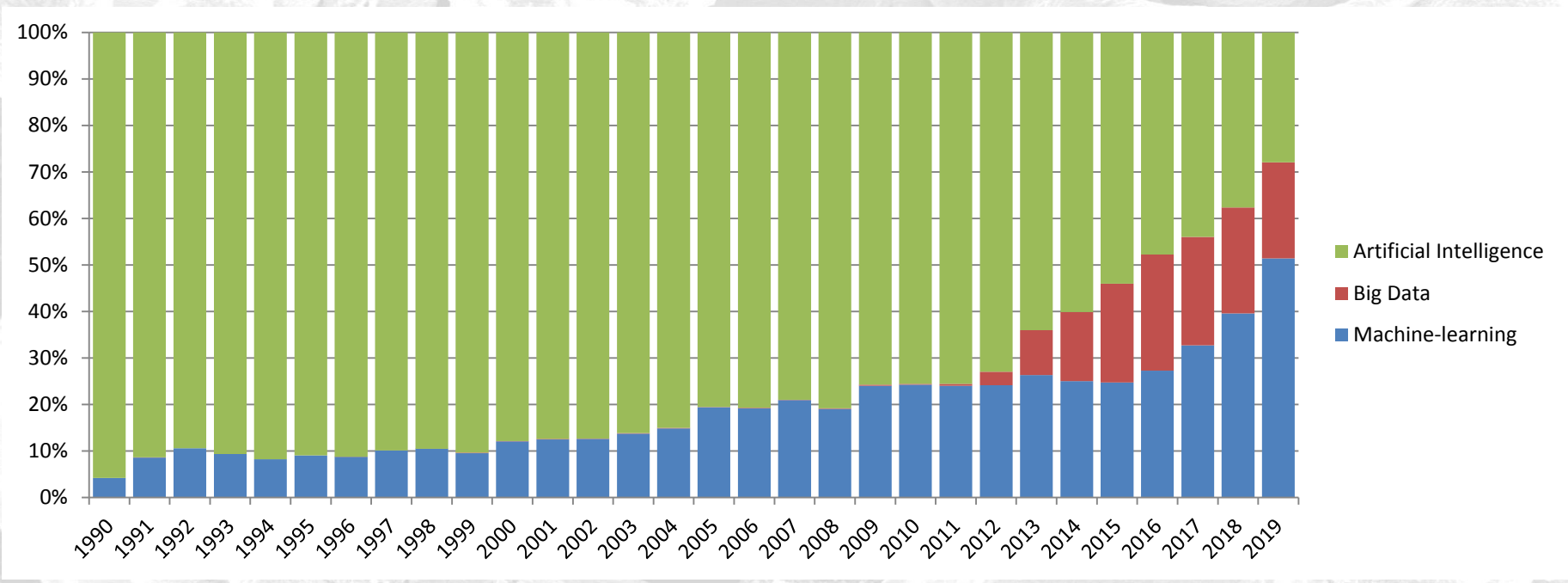


[Kausalitätsproblem](#)

[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



FORCE11
The Future of Research Communications and e-Scholarship

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) initial set of software citation principles (https://www.force11.org/group/software-citation-implementation-working-group). The members of the Software Citation Implementation Working Group are: archivists, funders, repository developers, other

Goals

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. Contribute to the activities.

Progress/Activities

Deliverables and Timeline

Group Email: software@force11.org

Working Group Participation

Subscribe to group

65 Member(s)



deRSE 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 Schwerpunktkonferenz 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.