

# Data Curation 1x1 for Engineering Libraries

23.06.2020

Fraunhofer  
Informationszentrum  
Raum & Bau  
Andrea Wuchner M.A.  
CC-BY



# Ihre Referentin



Andrea Wuchner

[Andrea.Wuchner@irb.fraunhofer.de](mailto:Andrea.Wuchner@irb.fraunhofer.de)

Tel: +49 (0)711 / 970-2714

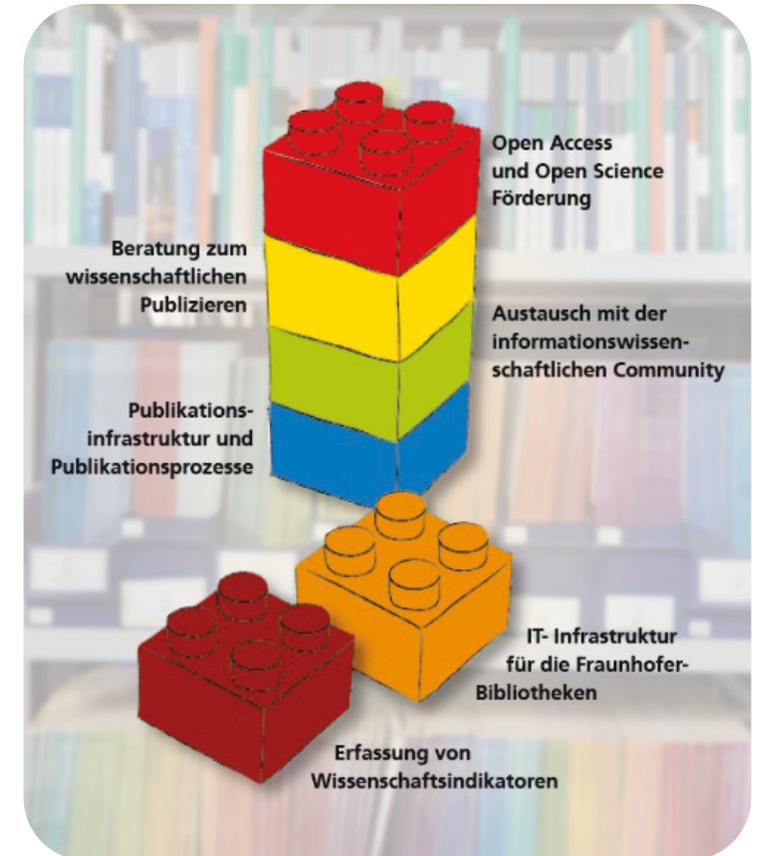
Bibliotheks- und  
Informationsmanagement,  
Master of Arts

- [Competence Center Research Services & Open Science](#)  
am Fraunhofer IRB in Stuttgart
- Aufgaben:
  - Forschungsdatenrepositorium [Fordatis](#)
  - Beratung/Support zum Thema Forschungsdaten  
und Forschungsdatenmanagement
  - Schulung und Informationsveranstaltungen  
Forschungsdatenmanagement
  - Forschungsprojekte: [FDM, Datagovernance](#)



# Fraunhofer Competence Center Research Services & Open Science

- Angesiedelt am Fraunhofer IRB, Stuttgart
- Wissenschaftsservices und Infrastrukturanbieter für 74 Fraunhofer-Institute und die Fraunhofer-Zentrale
- Zentraler Publikationssupport: [publikationssupport@fraunhofer.de](mailto:publikationssupport@fraunhofer.de)
- Fraunhofer-Forschungsdatenrepository Fordatis
- Fraunhofer-Publikationsrepositories [Publica](#), [ePrints](#)
- Fraunhofer-FIS Wissenschaftsindikatoren (intern)
- [Open-Science-Förderung](#) und wissenschaftspolitische Begleitung



# Warum werden Forschungsdaten für Bibliotheken zunehmend wichtiger?

- Die Digitalisierung als Megatrend ermöglicht die Generierung von großen Mengen an digitalen Forschungsdaten
- Politische Vorgaben, wie beispielsweise Förderbedingungen der europäischen Kommission, stellen Anforderungen an den Umgang mit Forschungsdaten
- Output-Kategorien des wissenschaftlichen Prozesses entwickeln sich weiter: neben klassischen Publikationen wird die Publikation von Forschungsdaten und Software in Zukunft an Bedeutung gewinnen
- Nationale und internationale Vereinheitlichung von Forschungsdateninfrastrukturen nehmen Fahrt auf: NFDI auf nationaler, EOSC auf internationaler Ebene, NFDI-Konsortien für Disziplinen

→ Forschungsdaten müssen gemanagt und zugänglich gemacht werden.

→ Das Sammeln, Erschliessen, Bewahren und Vermitteln von Wissensbeständen ist eine klassische Aufgabe von Bibliotheken, die einen großen Erfahrungsschatz auf diesen Gebieten vorweisen können. Es sollte gelingen, diesen auch im Bereich der Dokumentation von digitalisierten Forschungsdaten produktiv zu machen.



# Aufgaben im Forschungsdatenmanagement für Bibliotheken

Beratung	Technik	Standardisierung
Support: Rechtliche Fragen, Datenmanagementpläne, Veröffentlichung von Daten, Austauschformate	Aufbau und Betrieb von Forschungsdaten-Repositoryn	Entwicklung von Metadaten-Standards
Schulung und Training	Aufbau und Betrieb von Langzeitarchiven	Standardisierte Erfassung und Beschreibung von Forschungsdaten
Informationsangebote wie Webseiten, Newsletter	Vergabe von persistenten Identifier (z.B. DOIs)	Qualitätskontrolle der Metadaten



# Herausforderungen beim Aufbau der Angebote in Bibliotheken

- Know-How ist erforderlich
- Zusammenarbeit mit (neuen) Partnern, wie Rechenzentren, IT-Infrastruktur
- Enge Zusammenarbeit mit den Wissenschaftlerinnen und Wissenschaftlern auf Augenhöhe
- Akzeptanz als kompetenter Ansprechpartner im FDM, evtl. auch Konkurrenz zu Datenarchiven
- Anpassung der organisatorischen Rahmenbedingungen
- Proaktiver Aufbau von Angeboten vs. Erwartungshaltung an die Bibliothek



# Forschungsdatenmanagement – Welche Vorteile bietet es?

- Für Forscher\*innen
  - Sicherstellung der Auswertbarkeit
  - Schutz der Daten vor Missbrauch und Diebstahl
  - Erfüllung der Compliance-Anforderungen der Forschungsförderer
  - Bei Veröffentlichung: Klare Zuordnung zum Urheber → Steigerung der Reputation, korrekte Zitate, Kooperationsmöglichkeiten, Verlagsvorgaben
- Für die wissenschaftliche Community
  - Interpretation vorhandener Datensätze unter neuen Fragestellungen
  - Verifizierbarkeit von Forschung
- Für die Gesellschaft / Wirtschaft
  - Möglichkeit der Nachnutzung und Teilhabe
  - Steigerung der Transparenz der Wissenschaft → Vertrauen



# Themen

- Einführung in das Thema „Forschungsdaten“ und „Forschungsdatenmanagement“
- Rechtliche Rahmenbedingungen des Forschungsdatenmanagements
- Planung des Forschungsdatenmanagements
- Forschungsdaten dokumentieren
- Forschungsdaten veröffentlichen
- Forschungsdaten lizenzieren
- Forschungsdaten nachnutzen
- Forschungsdaten und politische Entwicklungen: Nationale Forschungsdateninfrastruktur und European Open Science Cloud



---

# EINFÜHRUNG IN DAS FORSCHUNGSDATENMANAGEMENT

---

## Inhalt

Was sind Forschungsdaten?

Was ist Forschungsdatenmanagement?

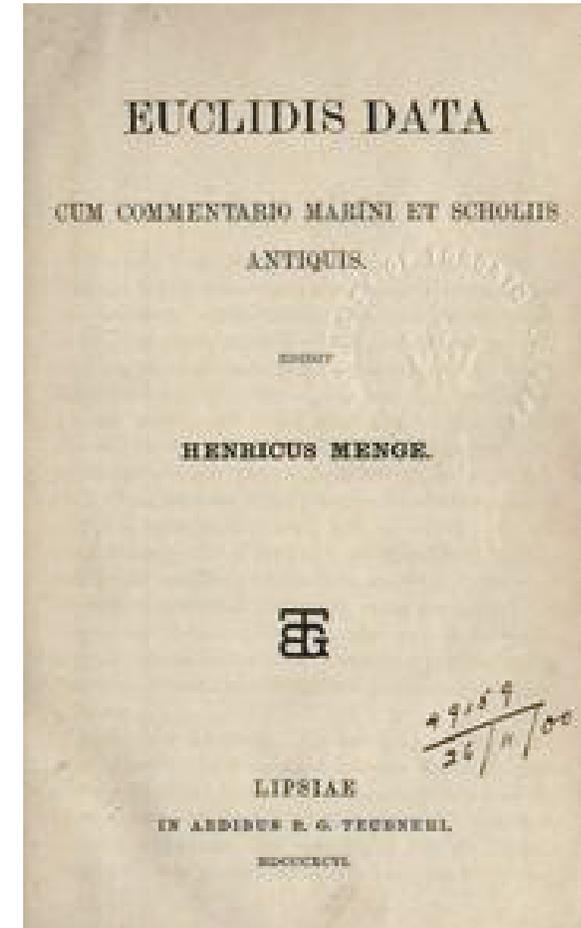
Grundbegriffe des Forschungsdatenmanagements: Who is who?

Aufgaben für Bibliotheken im Forschungsdatenmanagement



# Forschungsdaten - Etymologisches

- Daten: griechisch „ $\Delta\epsilon\delta\omicron\mu\acute{\epsilon}\nu\alpha$ “, Dedomena, das Gegebene
- Vom griechischen Mathematiker Euklid in seinem Werk zur Geometrie erstmals verwendet
- Dann als „Data“ ins Lateinische übersetzt.



# Forschungsdaten - Definition

## Definition der Allianz der deutschen Wissenschaftsorganisationen:

Daten, die im Zuge wissenschaftlicher Vorhaben, z. B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.

### Rohdaten

„Digitale“ Ausgangsdaten

- Reine Messwerte
- Audio-Dateien von Interviews
- Aufnahmen aus bildgebende Verfahren

Forschungs-  
daten

### Primärdaten

Eigene  
Datensammlungen

- Aufbereitete Daten nach Bereinigung falscher Werte
- Text-, Bildannotationen
  - inhaltlich
  - technisch
  - bibliografisch

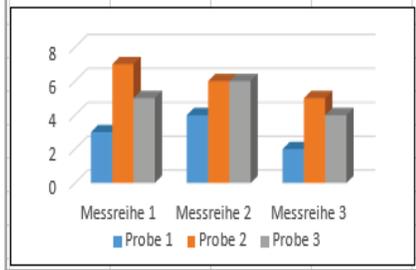
### Sekundärdaten

- Zeitaufwändige Aufbereitung
- Beschränken sich auf vorherige Forschungsfrage
- Problem, wenn Sekundärdaten als Primärdaten interpretiert werden

[Quelle: Uwe Vorher]

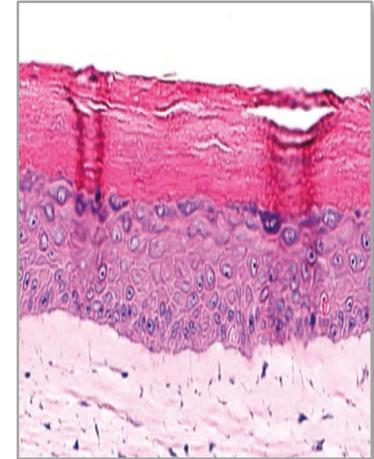
# In welcher Form können Forschungsdaten auftreten ?

	Messreihe 1	Messreihe 2	Messreihe 3
Probe 1	3	4	2
Probe 2	7	6	5
Probe 3	5	6	4



Statistische Daten,  
Laborwerte,  
geräteabhängige  
Messwerte...

Digitalisierte Karte, Fotos,  
Audio- oder  
Videoaufnahmen, ...



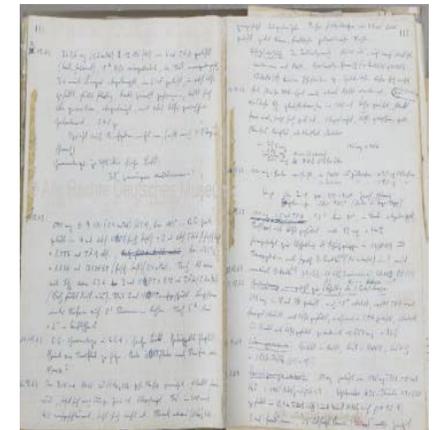
$$f_m = \sum_{k=0}^{n-1} x_{2k} e^{-\frac{2\pi i}{2n} m(2k)} + \sum_{k=0}^{n-1} x_{2k+1} e^{-\frac{2\pi i}{2n} m(2k+1)}$$

$$= \sum_{k=0}^{n-1} x'_k e^{-\frac{2\pi i}{n} mk} + e^{-\frac{\pi i}{n} m} \sum_{k=0}^{n-1} x''_k e^{-\frac{2\pi i}{n} mk}$$

$$= \begin{cases} f'_m + e^{-\frac{\pi i}{n} m} f''_m & \text{falls } m < n \\ f'_{m-n} - e^{-\frac{\pi i}{n} (m-n)} f''_{m-n} & \text{falls } m \geq n \end{cases}$$

Algorithmen, Skripte,  
Programme, Netzwerk-  
analysen, ...

„Digitale“ Labor- oder  
Feldtagebücher,  
Transkripte, Annotationen,  
Fragebögen,...



[Quelle: Uwe Vorher]

# Forschungsdaten - Erhebungsmethoden

- Empirische Beobachtung
- Automatische Messung
- Naturwissenschaftlich-technische Experimente
- Mathematisch-technische Simulationen
- Abbildung von Objekten
- Umfragen und Interviews mit Personen
- Statistiken und Referenzdaten
- Logfiles und Nutzungsdaten



# Datenarchetypen in den Ingenieurwissenschaften

## Archetypes



Quelle: <https://nfdi4ing.de/>

# Forschungsdaten - Qualitativ vs. Quantitativ

## ■ Qualitative Daten

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

## ■ Quantitative Daten

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

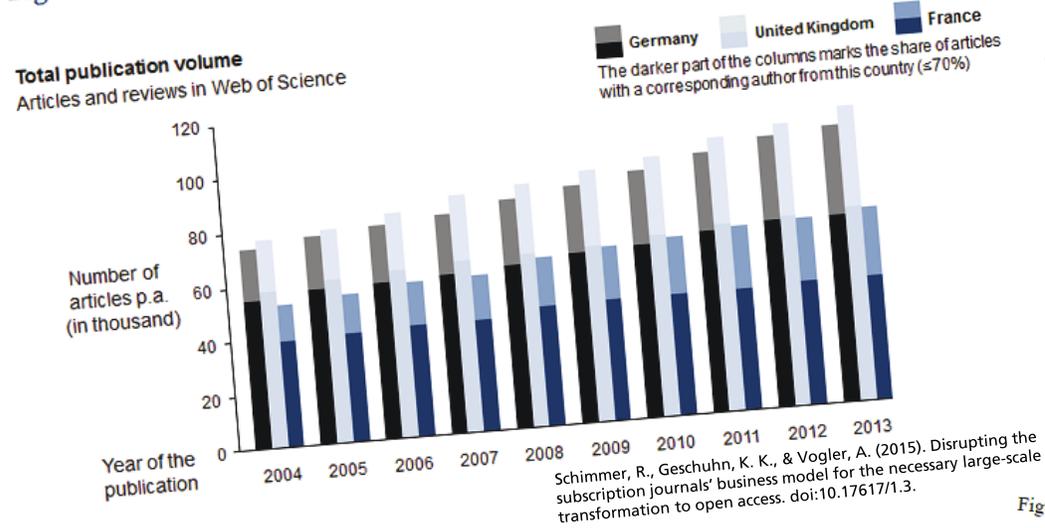
[[CC BY 4.0](#), Michael Franke (2017): Forschungsdatenmanagement, Folie 18]



# Forschungsdaten - Worum geht es nicht?

- Wissenschaftliche Publikationen
- Daten über Wissenschaft

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



Publication volume with APC relevant share of selected European countries

European Journal of Research and Reflection in Management Sciences

Vol. 2 No. 2, 2014  
ISSN 2056-5992

## Planning Approaches

Planning approaches describe the overall basic setting, i.e. guidelines for conducting technology planning. It is widely recognized that planning approaches can be divided into a synoptical [Ansoff, 1990; Albach, 1969] and incremental approach [Learned, 1965; Katz, 1970]. These two approaches depict the extreme ends of the entire range of the approach's characteristics. In practice, most firm's do not apply a pure form of approach but tend to implement a more mixed approach, such as the synoptical-analytical approach, incremental-entrepreneurial etc. [Paine, Anderson, 1977; Wohlgemut, Hess, 2000].

The synoptical planning aims at mastering the complexity of a planning problem in a holistic, comprehensive way without dividing it into sub-problems. This planning approach tries to capture and process as many relevant information as possible [Picot, Lange, 1979]. In a first step one or many objectives are being defined. In a later planning process phase these objectives are being used in order to assess different strategies and to select strategies by the management (top-down). During this process the corporate strategy will be reviewed periodically, revised and forwarded to a lower hierarchy level [Wohlgemut, Hess, 2000].

	Zahn	Szyperski	Rüth	Ansoff	Mintzberg	Hahn	Al-Laham
ontext aspects							
oundings	1						
tegy	0	1	2				
rmal context	1	0	1	2	1		
tellation	0	2	1	1	1	1	0
		1	2	1	2	0	0
				2		2	1
					2	1	0
ational							
n dimensions							
ophy	2	1					
ontrol	2	2	1	1	1	1	0
	2	2	2	1	2	2	2
onal							
imensions							
	1	2	2	1	1	2	2
	2	2	2	1	1	2	2
	1	2	1	1	1	2	2
al							
nensions							
ut	1	0	2	2	1	2	2
	1	2	2	1	1	2	2
	1	2	1	0	2	1	2

Legend: 0 = not processed, 1 = partly processed, 2 = processed  
Figure 1: Overview of the scholarly contribution to the discipline of planning (extract)

Transferability of the synoptical and incremental planning approach to technology planning: a literature review



# Forschungsdatenmanagement - Definition

Forschungsdatenmanagement = Forschungsdatenmanagement (FDM) bezeichnet alle Maßnahmen im Umgang mit digitalen Daten, die während des Forschungsprozesses entstehen oder die beforscht werden. Dazu zählen die einheitliche Datenerhebung, Dokumentation, Benennung und Datenorganisation, die Erteilung von Zugriffsrechten sowie schließlich die sichere Speicherung während des Forschungsprozesses bis zur nachhaltigen Veröffentlichung und Langzeitarchivierung der Daten über den Projektabschluss hinaus.

[FU Berlin <https://www.fu-berlin.de/sites/forschungsdatenmanagement/wiss-praxis/definitionen/index.html> ]

Forschungsdatenmanagement soll sicherstellen, dass die Forschungsdaten während und nach der Projektphase nutzbar sind und es bleiben.



# Forschungsdatenmanagement – Wichtige Oberbegriffe

Forschungsdaten-  
lebenszyklus

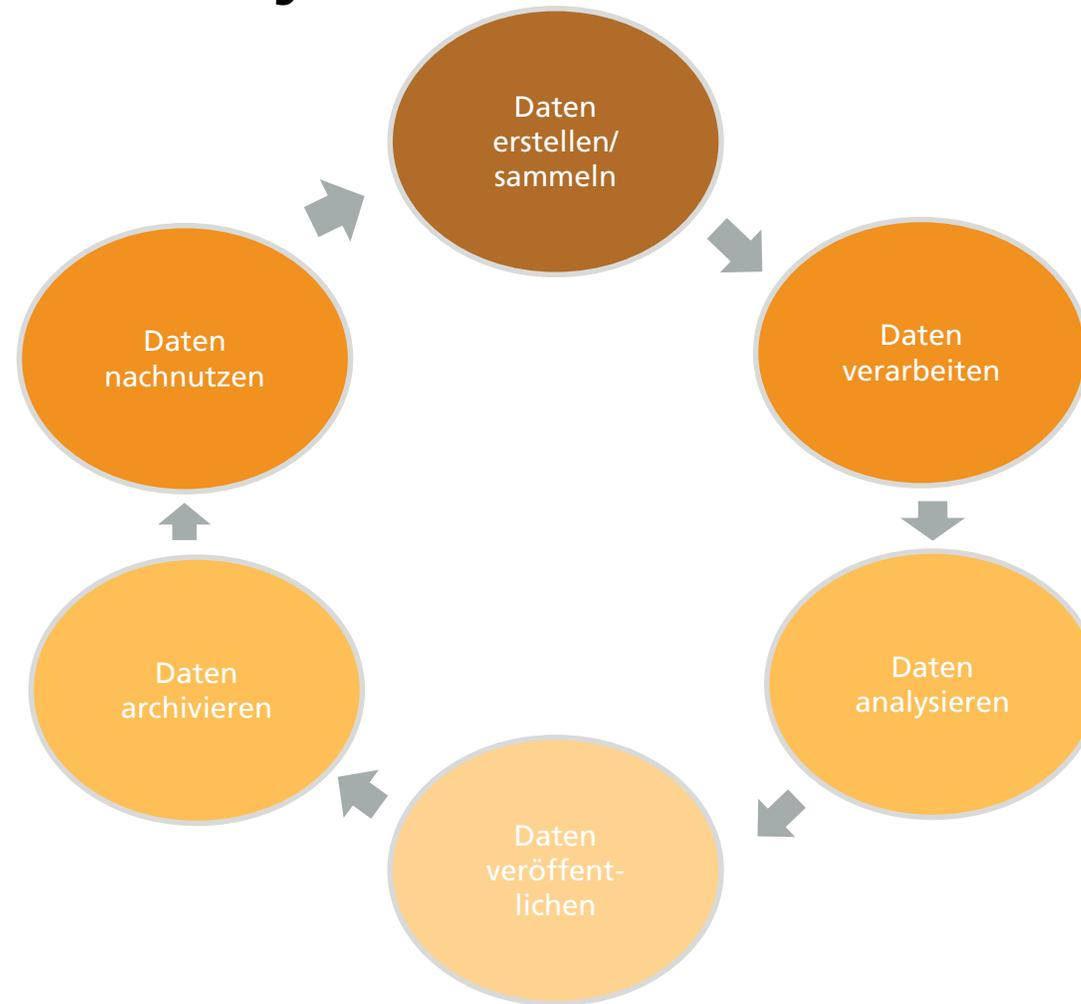
Metadaten

Persistente  
Identifikatoren

Forschungsdaten-  
Repositorien

FAIR-Prinzipien

# Forschungsdatenlebenszyklus



[[Nach Forschungsdatenlebenszyklusmodell, UK Archive Lancaster](#)]

# Daten und Metadaten

Metadaten = Daten/Informationen, die in strukturierter Form analoge oder digitale Objekte beschreiben.

Metadaten

Metadaten

Forschungsdaten

The image shows a screenshot of a research paper abstract and its associated metadata. The abstract is titled "Looking for the 'grey value' in research and geodesy: Is there right or wrong in software framework decisions?" and is authored by Florian Thiery and Clemens Schmid. The abstract text discusses the challenges of making "right" decisions in geodesy research, particularly in the context of software development and version control systems like Git. The abstract is dated June 20, 2018, and is available as a presentation and open access. The abstract is indexed in OpenAIRE. The abstract is also associated with the IGSN (International Geodetic Student Meeting) in Valencia, Spain, from June 23 to 29, 2018. The abstract is published in the Publications of Florian Thiery and linkedgeodesy by Florian Thiery M.Sc. Zenodo. The abstract is also associated with the keywords "git", "Software Development", and "IGSM". The abstract is also associated with the subject "software development" and "Igit". The abstract is also associated with the meeting "XXXI. International Geodetic Student Meeting (IGSM Valencia), Valencia, Spain, 23.06.-29.06.2018". The abstract is also associated with the communities "Publications of Florian Thiery" and "linkedgeodesy by Florian Thiery M.Sc. Zenodo".

Metadaten



# Forschungsdaten-Repositoryen

Forschungsdaten-Repositoryen = Systeme zur Sammlung und öffentlichen Bereitstellung von Forschungsdaten.



# Forschungsdaten veröffentlichen – Repositorien

- Aufgaben
  - Daten aufnehmen
  - Daten bereitstellen
  - Daten auffindbar machen
  - Daten exportieren
- Benutzung oft kostenfrei
- Kein Review-Prozess



# Forschungsdaten veröffentlichen – Repositorien Übersicht

Institutionell

Fachlich

Allgemein

[Open Data LMU](#)

[PsychData](#)

[Figshare](#)

[Edinburgh DataShare](#)

[Pangaea](#)

[DANS Easy](#)

[Dryad](#)

**Fordatis**

[Zenodo](#)

[Radar](#)

[B2SHARE](#)

# Forschungsdaten- Repositorien

## Beispiel: Zenodo

The screenshot displays the Zenodo website's header with the logo, a search bar, and navigation links for 'Upload' and 'Communities'. On the right, there are 'Log in' and 'Sign up' buttons. The main content area features a 'Recent uploads' section with two entries:

- Historical climate model output of ECHAM5-wiso from 1871-2011 at T106 resolution**  
by Nathan J. Steiger  
Metadata: May 18, 2018 (v2), Dataset, Open Access  
Description: Historical climate model simulation of the isotope-enabled ECHAM5-wiso model from the years 1871 to 2011 at T106 (1 degree) resolution. The model code was provided by Martin Werner of AWI. The simulations were designed and run by Nathan Steiger on the Yellowstone supercomputer. The boundary...  
Uploaded on May 18, 2018  
1 more version(s) exist for this record
- Reproducibility Package for "Reproducible research and GIScience: an evaluation using AGILE conference papers"**  
by Nüst, Daniel

On the right side, there are three promotional banners:

- Zenodo now supports DOI versioning!** with a tag icon and a link to read more.
- Using GitHub?** with a GitHub icon and a link to log in.
- Zenodo in a nutshell** with a bullet point: **Research. Shared.** — all research outputs from across all fields of research are



# Forschungsdaten- Repositorien

## Beispiel: Zenodo

Delete Save Publish

### New upload

**Instructions:** (i) Upload minimum one file or fill-in required fields (marked with a red star ). (ii) Press "Save" to save your upload for editing later. (iii) When ready, press "Publish" to finalize and make your upload public.

Files Choose files Start upload

Drag and drop files here

— or —

Choose files

(minimum 1 file required, max 50 GB per dataset - [contact us](#) for larger datasets)

Communities recommended

Start typing a community name...

Upload type required

Publication  Poster  Presentation  Dataset  Image  Video/Audio  Software  Lesson  Other

**Publication type**



# Persistente Identifikatoren

Persistente Identifikatoren = Ein persistenter Identifikator (Persistent Identifier) (PID) ist eine eindeutige Benennung (Referenzierung) einer digitalen Ressource (z.B. Zeitschriftenartikel oder Forschungsdaten) durch Vergabe eines Codes, der im Internet dauerhaft eindeutig referenziert werden kann.

[\[Quelle\]](#)

## Arten von persistenten Identifikatoren

Unique Resource Name	URN	urn:nbn:de:bsz:25-opus-1412
Digital Object Identifier	DOI	doi:10.1088/0004-637X/715/2/1453
Handle	HDL	hdl:11858/00-001M-0000-0019-D20F-6



# Persistente Identifikatoren

The screenshot shows a Zenodo dataset page. At the top, the Zenodo logo is on the left, a search bar in the center, and 'Upload' and 'Communities' on the right. A user profile for 'Andrea.Wuchner@irb.fraunhofer.de' is visible. The main content area features the title 'Intense, directed neutron beams from a laser-driven neutron source at PHELIX' and the author 'Kleinschmidt, Annika'. A summary paragraph describes the laser-driven neutrons. To the right, statistics show 32 views and 7 downloads. A red box highlights the 'Publication date: April 26, 2018' and the 'DOI: DOI 10.5281/zenodo.1230309'. Below this, the license is identified as 'Creative Commons Attribution 4.0 International'. A file list shows 'neutron numbers.xlsx' and 'RCF max proton energy.xlsx' with their respective MD5 hashes. At the bottom, there are options to show only literature, datasets, or software, and a 'Share' button.

zenodo Search Upload Communities Andrea.Wuchner@irb.fraunhofer.de

April 26, 2018 Dataset Open Access

## Intense, directed neutron beams from a laser-driven neutron source at PHELIX

Kleinschmidt, Annika

Laser-driven neutrons are generated by the conversion of the nuclear reactions inside a converter material. We present the campaign at the PHELIX laser at GSI in Darmstadt where protons were accelerated from thin deuterated plastic foils with thickness in the range of 100 nm. The neutrons were generated inside a sandwich-type converter leading to reproducible neutron numbers around  $10^{11}$  neutrons per steradian. The neutron distribution was measured with a high level of detail using a detector array simultaneously. It shows a laser forward directed component of neutrons per steradian, corresponding to a dose of 43 mrem 1 m from the converter.

Files (133.8 kB)

Name
neutron numbers.xlsx
md5:f91ff078e2e09c0c7bb5961c3d0ffa6a
RCF max proton energy.xlsx
md5:dc73a00f978e384adb6ba455145e3f66

**Beta** Citations 0

Show only:  Literature (0)  Dataset (0)  Software

Citations to this version

No citations.

Share

**Publication date:**  
April 26, 2018

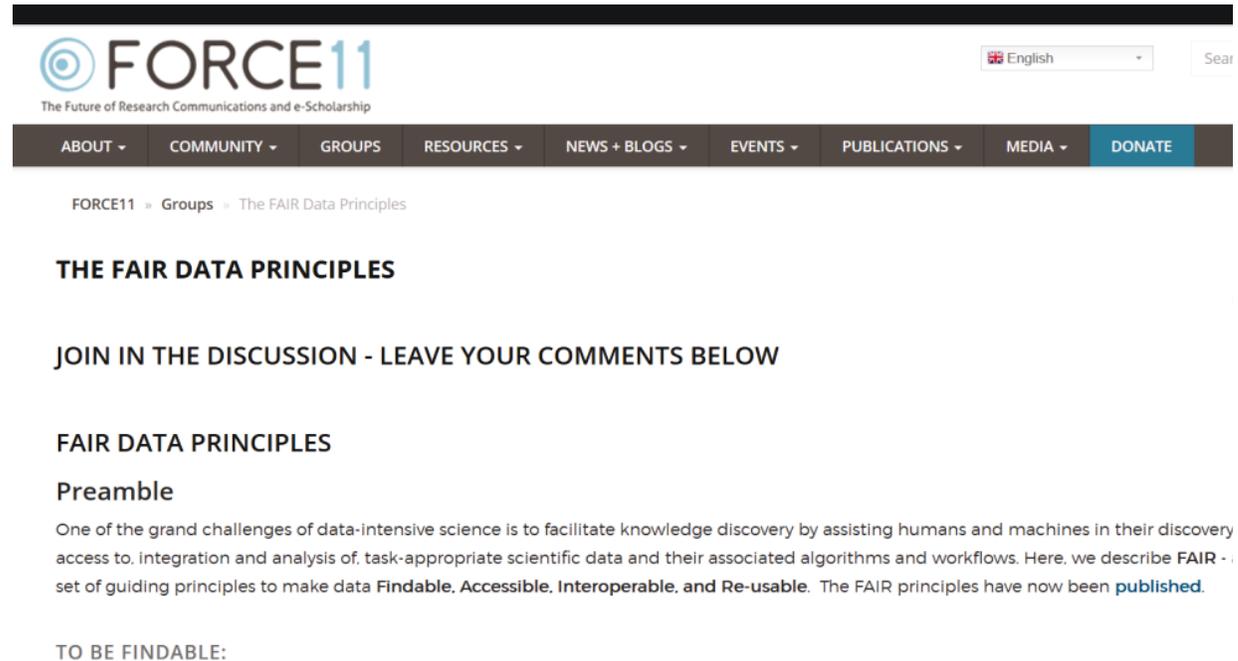
**DOI:**  
DOI 10.5281/zenodo.1230309

**License (for files):**  
Creative Commons Attribution 4.0 International



# Die FAIR Prinzipien

- Findable
- Accessible
- Interoperable
- Re-Usable



The screenshot shows the FORCE11 website header with the logo and tagline "The Future of Research Communications and e-Scholarship". A navigation menu includes links for ABOUT, COMMUNITY, GROUPS, RESOURCES, NEWS + BLOGS, EVENTS, PUBLICATIONS, MEDIA, and DONATE. The breadcrumb trail reads "FORCE11 » Groups » The FAIR Data Principles". The main heading is "THE FAIR DATA PRINCIPLES". Below it is a call to action: "JOIN IN THE DISCUSSION - LEAVE YOUR COMMENTS BELOW". The section title is "FAIR DATA PRINCIPLES" followed by a "Preamble" section. The preamble text states: "One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery 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 Re-usable. The FAIR principles have now been published." Below the preamble is the heading "TO BE FINDABLE:".

Quelle: <https://www.force11.org/group/fairgroup/fairprinciples>



# Die FAIR Prinzipien

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.

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](#).

## To be Findable:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

## TO BE ACCESSIBLE:

- A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
- A1.1 the protocol is open, free, and universally implementable.
- A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
- A2 metadata are accessible, even when the data are no longer available.

## TO BE INTEROPERABLE:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

Quelle: <https://www.force11.org/group/fairgroup/fairprinciples>



# Anforderungen FAIR-Prinzipien an digitalen Forschungsoutput

## WHAT IS FAIR ?

- Persistent Identifier
- Rich Metadata
- Indexierungsstandards
- Standardisierte, offene Protokolle
- Metadatenstandards
- Kontrolliertes Vokabular
- Nachnutzungslizenzen
- Provenienznachweis
- Domänenspezifische Standards

### Findable:

- FM-F1A** **FM-F1B**  
F1 (meta)data are assigned a globally **unique** and **persistent** identifier;
- FM-F2**  
F2 data are described with **rich metadata**;
- FM-F3**  
F3 metadata clearly and explicitly include the **identifier of the data** it describes;
- FM-F4**  
F4 (meta)data are registered or **indexed** in a searchable resource;

### Interoperable:

- FM-I1**  
I1 (meta)data use a formal, accessible, shared, and broadly applicable **language for knowledge representation**.
- FM-I2**  
I2 (meta)data use **vocabularies that follow FAIR principles**;
- FM-I3**  
I3 (meta)data include **qualified references** to other (meta)data;

### Accessible:

- A1 (meta)data are retrievable by their identifier using a standardized communications protocol;
  - FM-A1.1**  
A1.1 the protocol is **open, free, and universally implementable**;
  - FM-A1.2**  
A1.2 the protocol allows for an **authentication and authorization** procedure, where necessary;
- FM-A2**  
A2 metadata are accessible, **even when the data are no longer available**;

### Reusable:

- R1 meta(data) are richly described with a plurality of accurate and relevant attributes;
  - FM-R1.1**  
R1.1 (meta)data are released with a clear and **accessible data usage license**;
  - FM-R1.2**  
R1.2 (meta)data are associated with **detailed provenance**;
  - FM-R1.3**  
R1.3 (meta)data meet domain-relevant **community standards**;

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)

<http://fairmetrics.org>

<https://github.com/FAIRMetrics/Metrics/blob/master/ALL.pdf>



# Weiterführende Informationen

- [Handbuch Forschungsdatenmanagement](#)

Herausgegeben von Stephan Büttner, Hans-Christoph Hobohm, Lars Müller. Bad Honnef : Bock + Herchen, 2011. ISBN 978-3-88347-283-6

- [forschungsdaten.info](#)

Informationsmaterial, Best-Practice-Beispiele, Tools, Forschungsdaten, ihre Organisation, Nachnutzbarkeit und Archivierung

- [Schwerpunktinitiative Digitale Information](#)

Allianz der Wissenschaftsorganisationen, Forschungsdaten

- [Forschungsdaten an der Universität Stuttgart](#)

Planen, Speichern, Teilen, Veröffentlichen:

- [The Digital Curation Centre \(DCC\)](#)

Expert advice and practical help to research organisations wanting to store, manage, protect and share digital research data, How-to Guides, case studies

- [Leitfaden zum Forschungsdaten-Management](#)

Handreichungen aus dem WissGrid-Projekt, Ludwig, Jens; Enke, Harry (Hrsg.)

- [Managing and Data Sharing – Best Practice for Researchers](#)

Herausgegeben von UK Data Archive, 3. Auflage, 2011



# Weiterführende Informationen

## ■ Definitionen

[Definition der Uni Heidelberg](#)

[Definition der Uni Bielefeld](#)

[Allianz der Wissenschaftsorganisationen](#)

[Definition der Leibniz-Gemeinschaft](#)

## ■ Big Data vs. Small Data

[Adam R Ferguson, Jessica L Nielson, Melissa H Cragin et al.: Big data from small data: data-sharing in the 'long tail' of neuroscience, Nature Neuroscience 17, 1442–1447 \(2014\)](#)



# Weiterführende Informationen

- [European Strategy Forum on Research Infrastructures – ESFRI](#)
- [European Open Science Cloud](#)
- [Research Data Alliance](#)
- [Open Knowledge Foundation](#)
- [FAIR-Prinzipien](#)
- [Explanation of the FAIR data principles](#)
- [Erläuterung der FAIR-Prinzipien](#)



# Weiterführende Informationen

- [Stefanie Rümpel: Der Lebenszyklus von Forschungsdaten, in: Handbuch Forschungsdatenmanagement, hrsg. von Stephan Büttner, Hans-Christoph Hobohm, Lars Müller, Bad Honnef: BOCK + HERCHEN Verlag, 2011](#)
- [Sarah Higgins: the Data Continuum Lifecycle Model, in: International Journal of Digital Curation, Vol 3 No 1 2008](#)
- [TRELOAR, Andrew; HARBOE-REE, Cathrine: Data management and the curation continuum: how the Monash experience is informing repository relationships. Proceedings of VALA 2008, 2008.](#)



---

# RECHTLICHE RAHMENBEDINGUNGEN FÜR DAS FORSCHUNGSDATENMANAGEMENT

---

## Inhalt

Rechtsgrundlagen für das Forschungsdatenmanagement

Zuordnungsrecht – Wem gehören die Forschungsdaten?

Datenschutz – Was muss ich im Rahmen des Datenschutzes beachten?



# Rechtsgrundlagen für das Forschungsdatenmanagement

- Europäische Datenschutzverordnung
- Grundgesetz
- Urheberrecht
- Arbeitsrecht
- Datenbankrecht
- Archivrecht
- Persönlichkeitsrechte
- Strafrecht
- Kriegswaffenkontrollgesetz
- Jugendschutzgesetz
- Kooperationsverbot
- Lizenzen für Forschungsdaten
- Einwilligungserklärungen



# Gutachten „Rechtliche Rahmenbedingungen des Forschungsdatenmanagements“

- Veröffentlicht 2018, aus dem BMBF-Projekt DataJus
- Behandelt rechtliche Fragen zum Forschungsdatenmanagement
  - Zuordnungsrecht
  - Datenschutz
  - Haftungsfragen
  - Fragen der Nachnutzung



[https://tu-dresden.de/gsw/jura/igetem/jfbimd13/ressourcen/dateien/publikationen/DataJus\\_Zusammenfassung\\_Gutachten\\_12-07-18.pdf](https://tu-dresden.de/gsw/jura/igetem/jfbimd13/ressourcen/dateien/publikationen/DataJus_Zusammenfassung_Gutachten_12-07-18.pdf)

# Rechtliche Rahmenbedingungen – Zuordnungsrecht



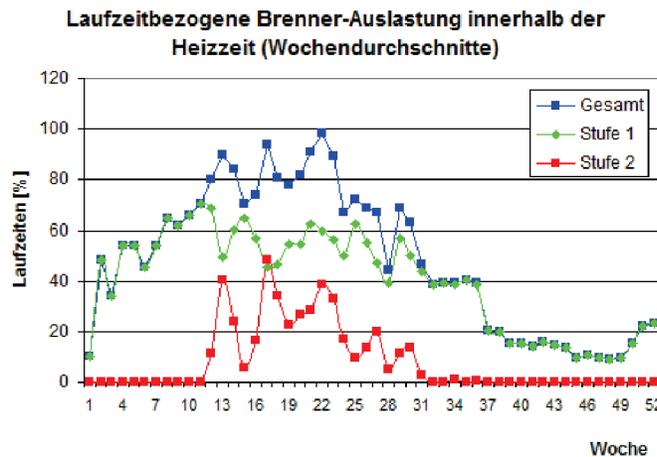
# Zuordnungsrecht – Wem „gehören“ die Forschungsdaten?

■ Wer kann über den Umgang mit Forschungsdaten und insbesondere über ihre Veröffentlichung entscheiden?

→ Derjenige, dem die Daten rechtlich zugeordnet sind.

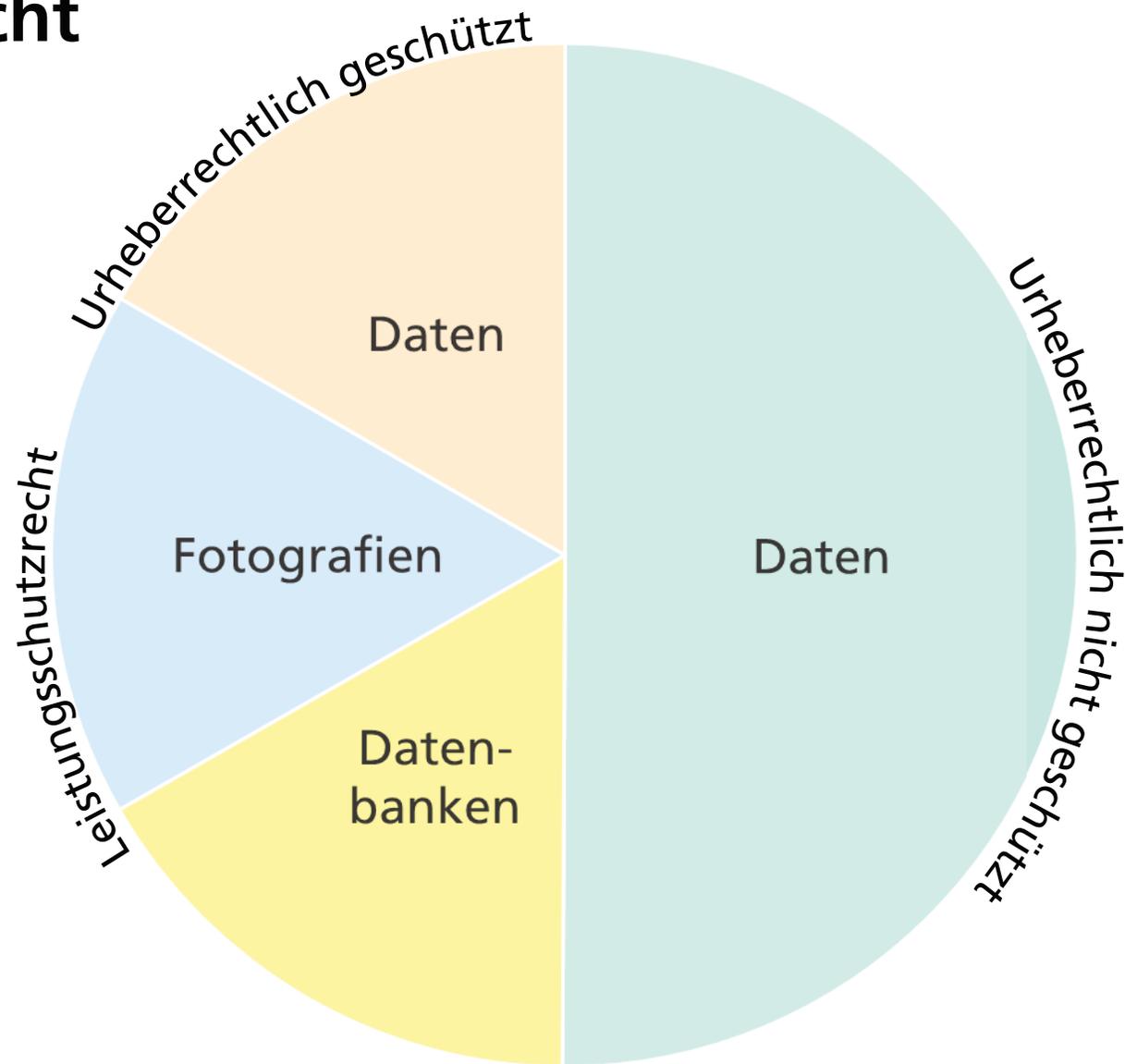


## §§ Rechtliche Zuordnung

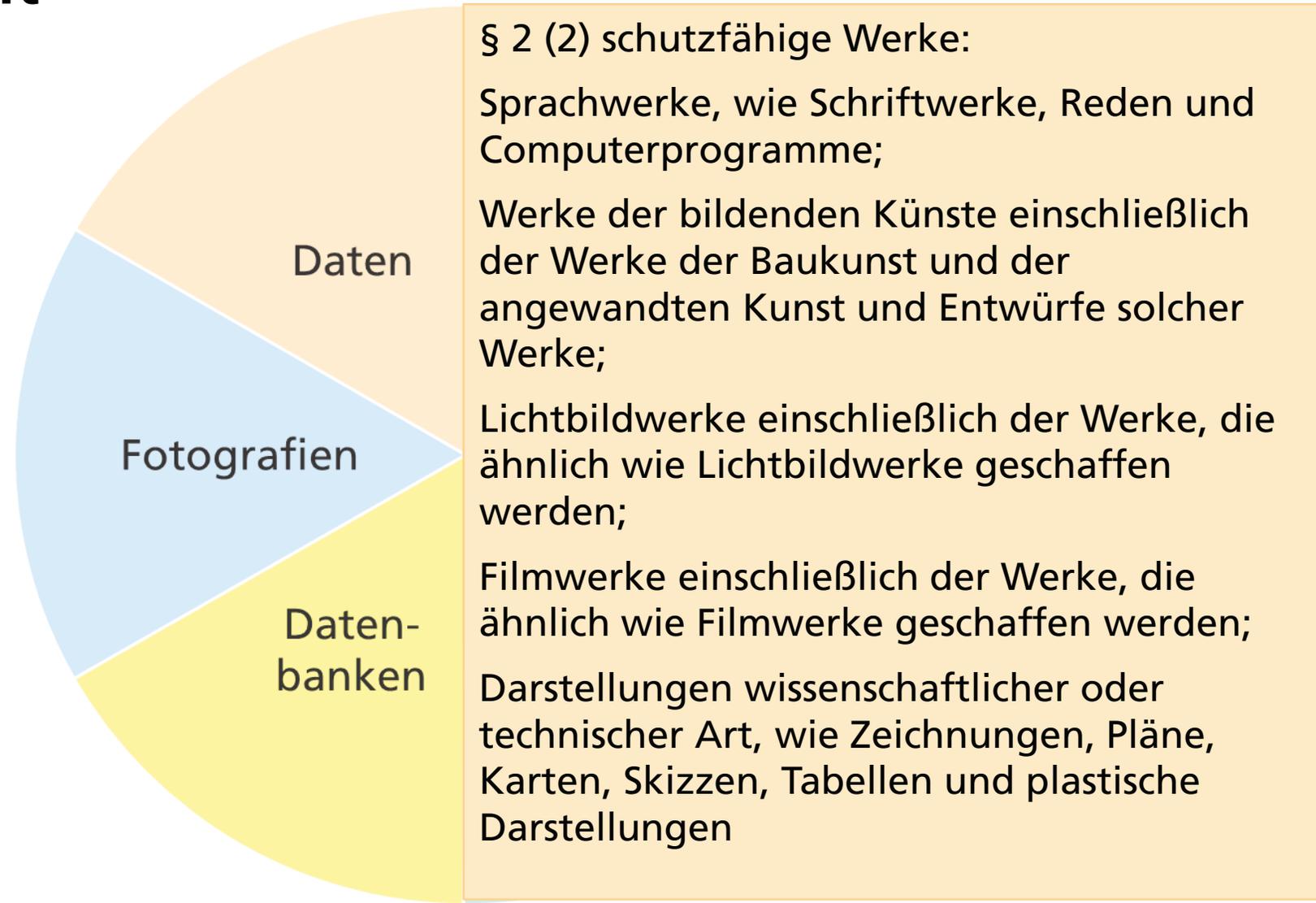


- § Urheberrecht
- § Leistungsschutzrechte
- § Arbeitsrecht
- § Patentrecht
- § ...

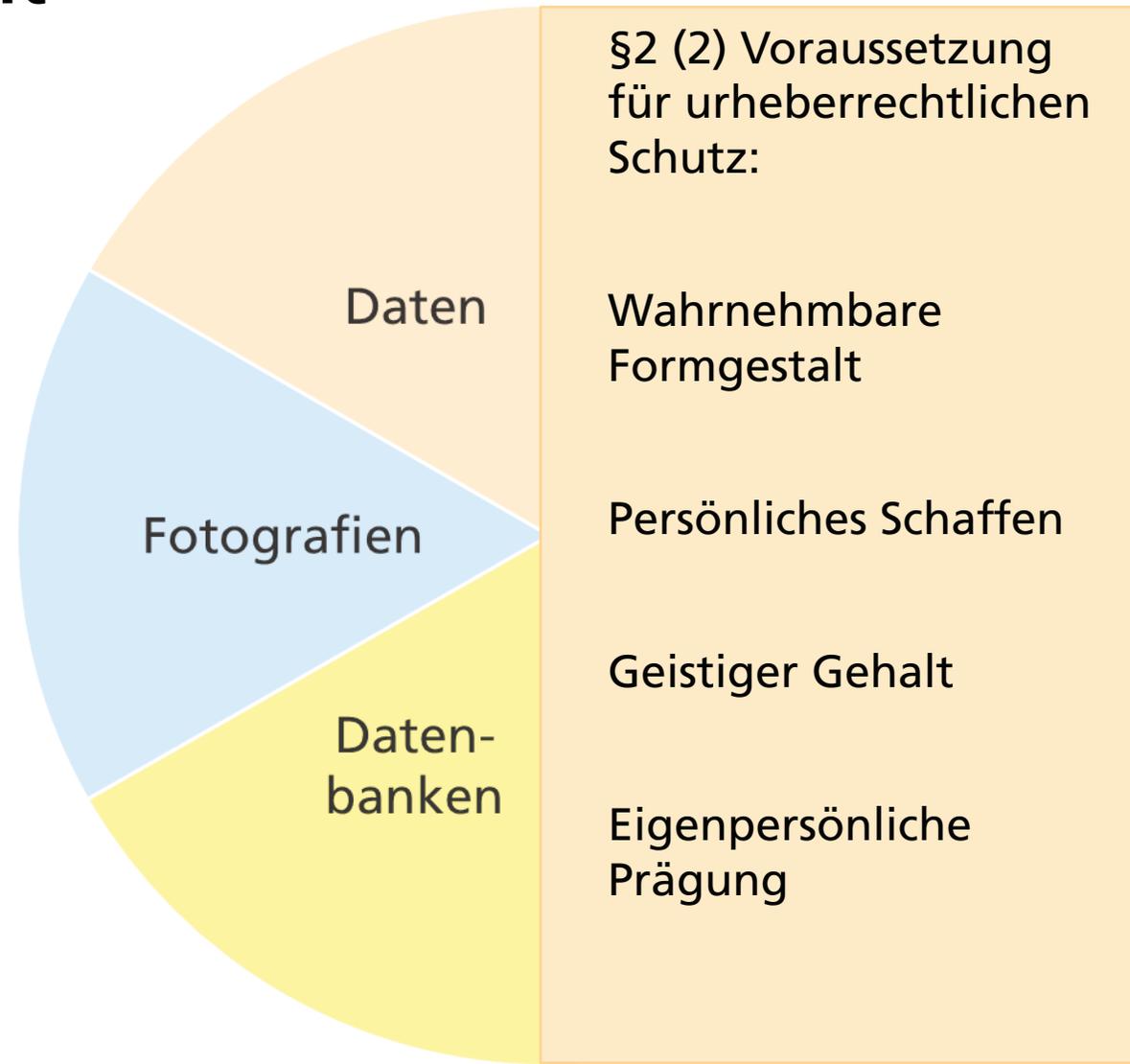
# Zuordnungsrecht



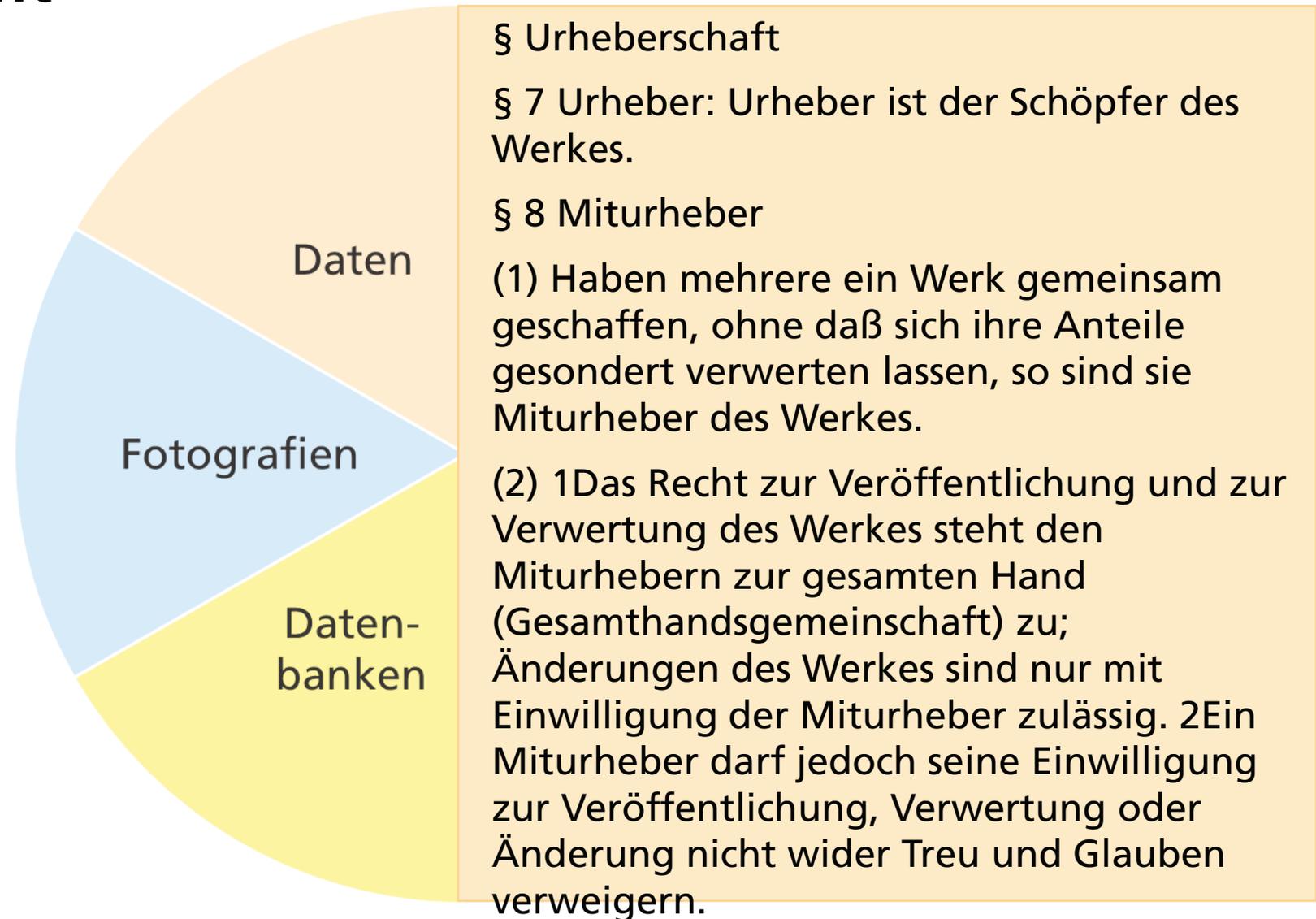
# Zuordnungsrecht



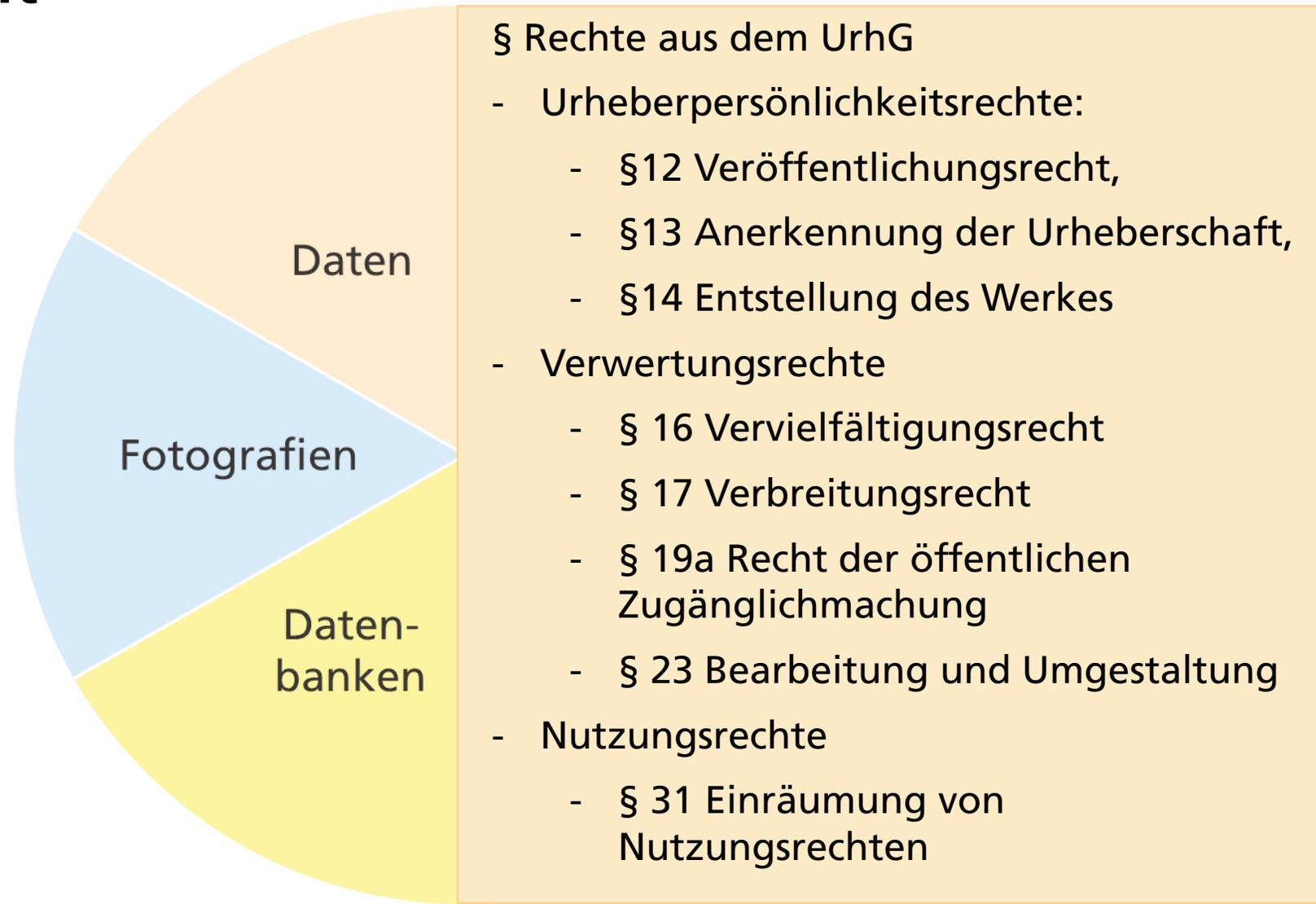
# Zuordnungsrecht



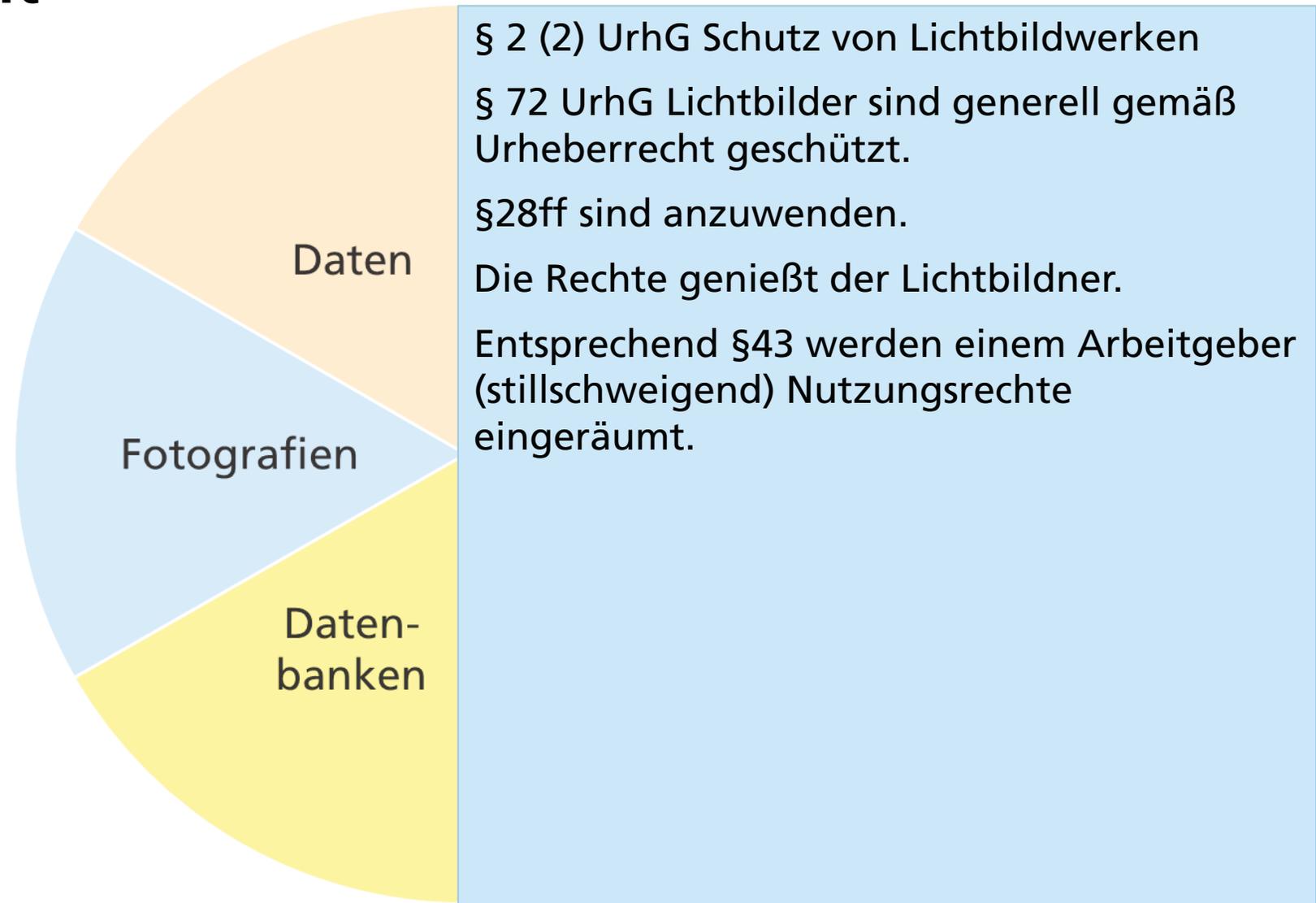
# Zuordnungsrecht



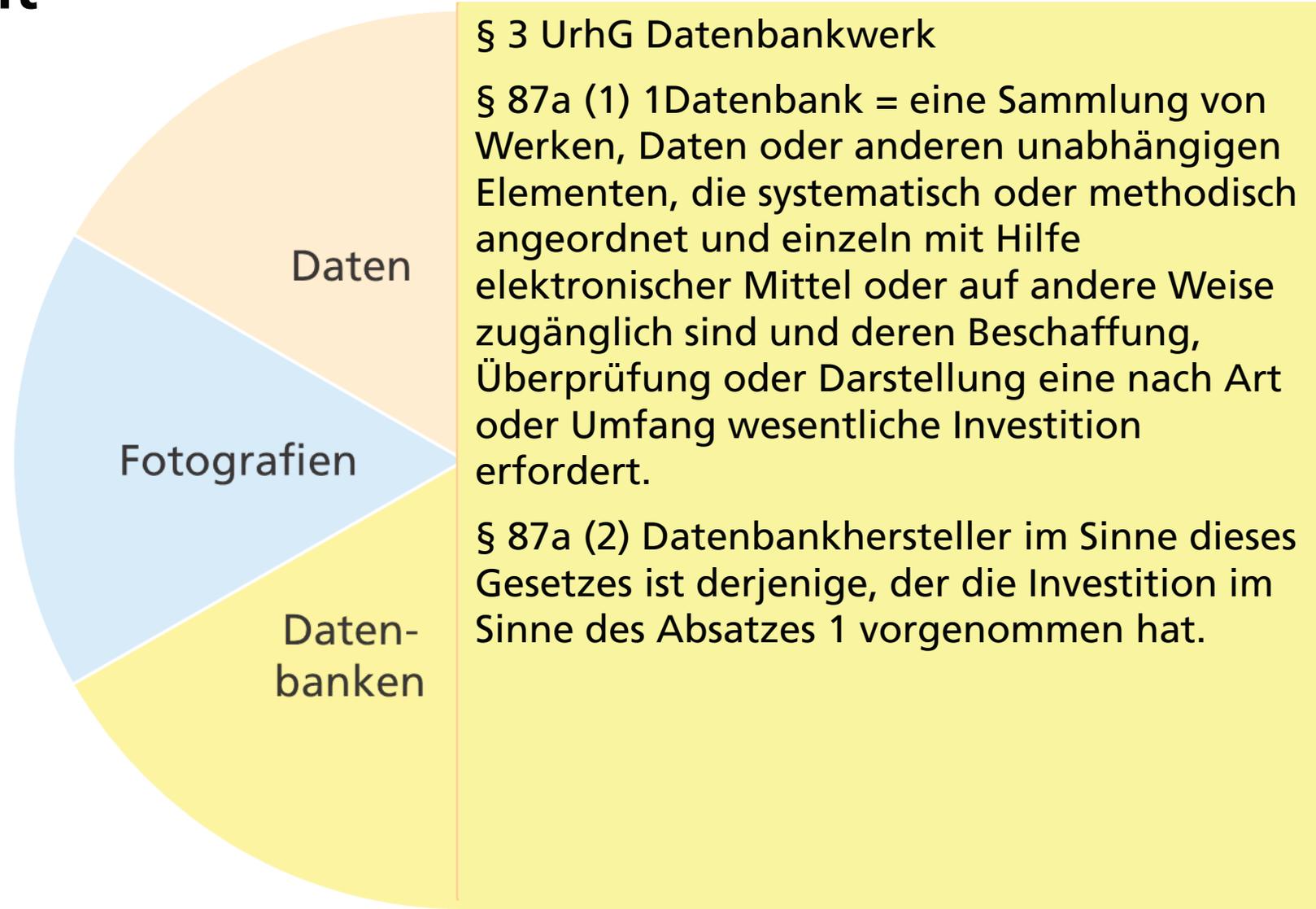
# Zuordnungsrecht



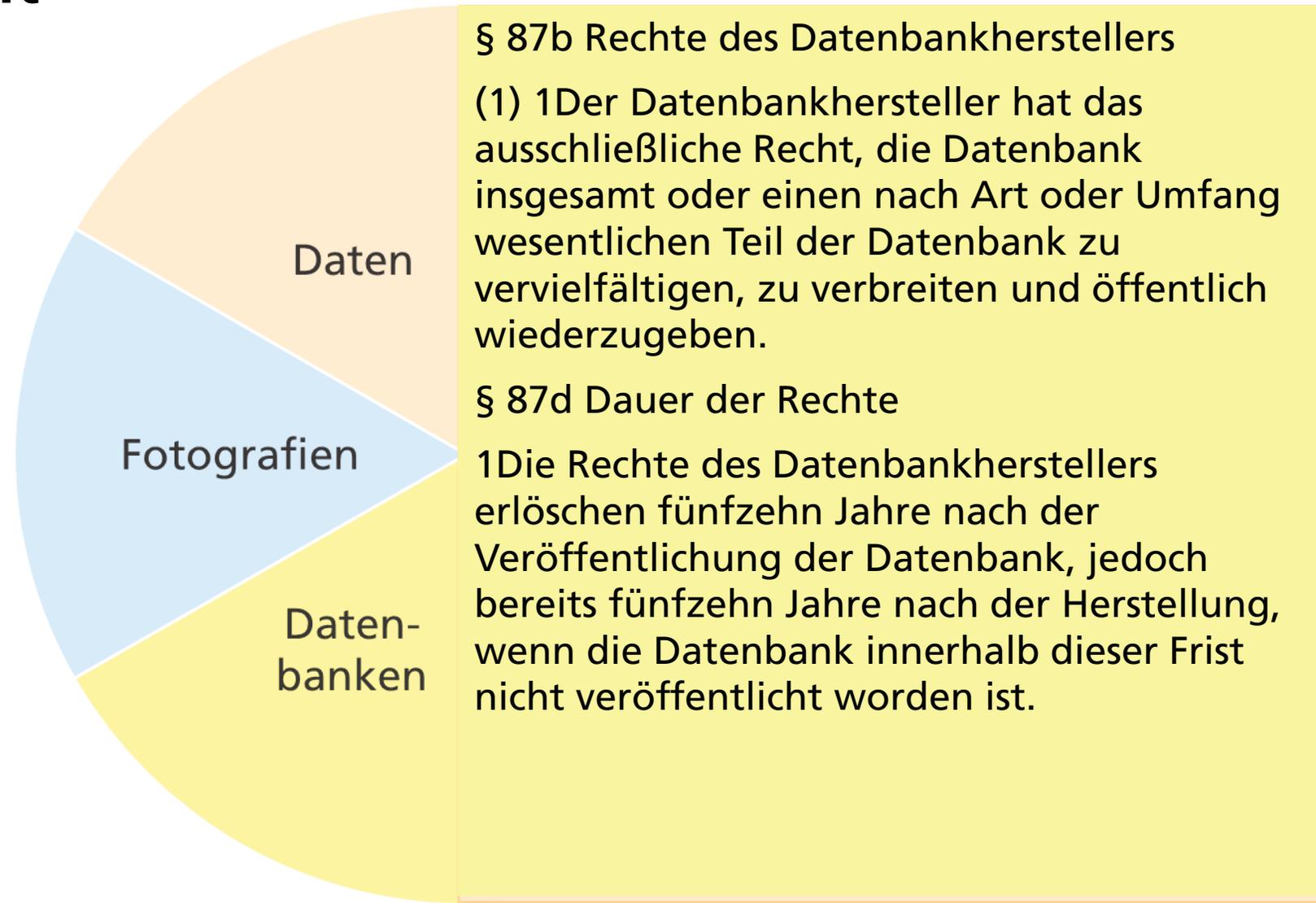
# Zuordnungsrecht



# Zuordnungsrecht



# Zuordnungsrecht



# Zuordnungsrecht

Sind die Voraussetzungen nicht erfüllt: Daten genießen keinen Urnehberschutz.

DataJus-Gutachten:

Zuordnung: Weitgehend rechtlich ungeklärt.

Zuordnung zum Wissenschaftler aufgrund seines

„Wissenschaftler-Persönlichkeitsrechts“

Zuordnung ergibt sich auch aus den „Grundsätzen zur Sicherung der guten wissenschaftlichen

Praxis“ bzw. „Leitlinien zur Sicherung der guten wissenschaftlichen Praxis“.

Daten

# Zuordnungsrecht

„Die Nutzung (der Forschungsdaten) steht insbesondere der Wissenschaftlerin und dem Wissenschaftler zu, die/der sie erhebt.“  
(Erläuterung zur Leitlinie 10)

Wissenschaftlerinnen und Wissenschaftler entscheiden [...] ob, wie und wo sie ihre Ergebnisse öffentlich zugänglich machen. Ist eine Entscheidung, Ergebnisse öffentlich zugänglich zu machen, erfolgt, beschreiben Wissenschaftlerinnen und Wissenschaftler diese vollständig und nachvollziehbar. Dazugehört es auch, soweit dies möglich und zumutbar ist, die den Ergebnissen zugrunde liegenden Forschungsdaten [...] verfügbar zu machen und Arbeitsabläufe umfänglich darzulegen. (Leitlinie 13)

Daten

# §43 Urheber in Arbeits- oder Dienstverhältnissen

- Nutzungsrechte aufgrund von Arbeits- und Dienstvertrages: Soweit die Schaffung urheberrechtlich geschützter Werke zu den arbeitsvertraglichen Pflichten des Arbeitnehmers gehören, werden dem Arbeitgeber Nutzungsrechte an „Pflichtwerken“ eingeräumt. Dies kann auch stillschweigend erfolgen.
- Dies gilt grds. auch für in der Forschung tätige Arbeitnehmer.
- Im Rahmen dieser stillschweigenden Einräumung überlässt der Wissenschaftler dem Arbeitgeber auch das Recht zu bestimmen, ob und wie das Werk veröffentlicht wird. Namensnennungsrechte dagegen können nur im Einzelfall eingeschränkt werden und bleiben insbesondere im wissenschaftlichen Kontext für gewöhnlich bestehen.



# §43 Urheber in Arbeits- oder Dienstverhältnissen

- Hochschullehrer: Einschränkung der Nutzungsrechtseinräumung auf den Arbeitgeber durch die Freiheit der Forschung (Art. 5 Abs. 3 GG): Urheberrechtlich geschützte Forschungsdaten, die von Hochschullehrern geschaffen wurden, sind nur dann der Hochschule zugewiesen, wenn ausdrückliche vertragliche Abreden, z.B. bei Drittmittelprojekten, bestehen. Andernfalls übt der Hochschullehrer die Verwertungs- und Veröffentlichungsrechte aus.
- Wissenschaftlichen Assistenten und Mitarbeiter: Wenn und soweit die wissenschaftliche Arbeit weisungsfrei erfolgt, z.B. da sie durch die Forschungsfreiheit gemäß Art. 5 Abs. 3 GG privilegiert wird, findet § 43 UrhG auch hier keine Anwendung. Erfolgt die Forschung dagegen weisungsabhängig, ist eine stillschweigende Nutzungsrechtseinräumung anzunehmen.
- Auch bei von Studierenden und externen Promovierenden geschaffenen Werken findet grds. keine Nutzungsrechtseinräumung an die Hochschule statt, da diese keine Arbeitnehmer sind. Auch können lediglich diese Personen entscheiden, ob die Daten veröffentlicht werden. Jedoch können abweichende vertragliche Vereinbarungen getroffen.





# Datenschutz - Allgemeines

- Zweck: Soll Personen und deren personenbezogene Daten schützen
- Spielt im Zusammenhang mit Forschungsdaten eine Rolle, wenn im Rahmen wissenschaftlicher Untersuchungen personenbezogene Daten erhoben werden, z.B. bei Interviews, Umfragen, Videoaufnahmen
- Rechtsgrundlagen: Europäische Datenschutzverordnung, Bundesdatenschutzgesetz, Landesdatenschutzgesetze



# Datenschutz – Personenbezogene Daten nach Artikel 4 DSGVO

Personenbezogene Daten sind „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person [...] beziehen.“

In der Praxis fallen darunter also sämtliche Daten, die auf jedwede Weise einer Person zugeordnet werden oder zugeordnet werden können.



# Datenschutz – Personenbezogene Daten nach Artikel 9 DSGVO

Eine besondere Kategorie personenbezogener Daten, stellen folgende Daten dar: Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person.

# Datenschutz – Verbotsprinzip mit Erlaubnisvorbehalt Art. 6 DSGVO

Das Erheben, Verarbeiten und/oder Nutzen von personenbezogenen Daten ist grundsätzlich verboten, es sei denn



Ein Gesetz erlaubt die Verarbeitung §§

Die betroffene Person hat in die Verarbeitung eingewilligt 👍

# Datenschutz – Einwilligung

## Einwilligung

- eindeutige, bestätigende Handlung,
  - mit der freiwillig,
  - für den konkreten Fall,
  - in informierter Weise und
  - Unmissverständlich bekundet wird,
- dass die betroffene Person mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist

ohne eine physische oder psychische Beeinflussung (§ 51 BdsG)

keine konkrete Form → Bei Nachnutzung und Veröffentlichung jedoch schriftlich (Art. 7 Abs. 1 DSGVO)

welche seiner persönlichen Daten wie, für was, von wem und wie lange verwendet werden sollen (Art. 7 Abs. 2 S. 1 und Art. 12 Abs. 1 S. 1 DSGVO)

die Informationen in präziser, transparenter, verständlicher und leicht zugänglicher Form in einer klaren und einfachen Sprache an den Betroffenen übermittelt werden (Art. 7 Abs. 2 S. 1 und Art. 12 Abs. 1 S. 1 DSGVO)

die Verantwortlichen, den Verarbeitungszweck, die Freiwilligkeit, die Folgen, die Speicherdauer, mögliche sonstige Datenempfänger im In- und Ausland, den zuständigen Datenschutzbeauftragten und die Aufsichtsbehörde sowie die bestehenden Betroffenenrechte informiert wird (vgl. Art. 13 und 14 DSGVO)

# Datenschutz - Datenminimierung

- Datenminimierung (Art 5 lit. C. DSGVO)
- Beschränkung auf das notwendige Maß (Erforderlichkeit)
- Daten müssen für die vorgesehene Zwecke erheblich sein.
- Erhebung muss im Verhältnis zum Zweck angemessen sein.
- Kontrollfragen:
  - Ist anzuerkennen, dass die Daten erforderlich sind, um den Zweck zu erreichen?
  - Kann der Zweck der Verarbeitung nicht in zumutbarer Weise durch andere Mittel erreicht werden?



# Datenschutz – Anonymisierung und Pseudonymisierung

- Die Anonymisierung ist das Verändern personenbezogener Daten derart, dass diese Daten nicht mehr einer Person zugeordnet werden können. → Absolute Anonymisierung
- Werden die Daten derart verändert, dass ein Personenbezug nur mit unverhältnismäßig großem Aufwand möglich ist, spricht man von → faktischer Anonymisierung
- Durch Anonymisierung werden Daten von den Restriktionen des Datenschutzes befreit.
- Einschränkungen der datenschutzrechtlichen Vorgaben zugunsten der Forschung: Für die interne Nutzung können eine faktische Anonymisierung (z.B. Klassenbildung) oder bereits eine Pseudonymisierung als ausreichend angesehen werden.



# Datenschutz – Wie funktioniert Anonymisierung?

- Direkte Identifikationsmöglichkeiten wie Name oder Adresse entfernen.
- Pseudonyme verwenden
- Informationen aggregieren oder reduzieren, z.B. Altersgruppen statt Geburtsdatum
- Begrenzung von oberen und unteren Werten um Ausreiser zu verstecken, z.B. ein Löhnen
- Generalisierung von Text, z.B. Ersetzen eines bestimmten Spezialgebiets eines Arztes mit einem allgemeineren Gebietgeneralisierung
- Bei Audio-Visuellen Medien: Stimmen verzerren, Auflösung heruntersetzen



# Datenschutz – Verarbeitungsverzeichnis

- Muss für die automatische Verarbeitung personenbezogener Daten angefertigt werden. → Trifft auch auf personenbezogene Forschungsdaten zu. (DSGVO Art. 30)
- Meldung an Datenschutzbeauftragten
- Vorlage und weitere Informationen im Fraunhofer-Intranet verfügbar.



# Datenschutz – Verarbeitungsverzeichnis

Um weitere Hauptauftragnehmer hinzuzufügen, klicken Sie bitte hier:

## Zweckbestimmung und Rechtsgrundlage der Datenverarbeitung

Zweckbestimmung

Erhebung der Daten zum Zweck der Erfassung der Arbeitszeiten und zur Öffnung der Türen

Erläuterung

Rechtsgrundlage

Vertrag o. Vertragsanbahnung mit dem Betroffenen  Einwilligung des Betroffenen  Vorrangige Rechtsvorschriften  
 Interessenabwägung  Sonstiges

Bitte ankreuzen soweit zutreffend und ggf. erläutern.

Betriebsvereinbarung "Zeiterfassung und Zutrittskontrolle", § 32 BDSG

Erläuterung Rechtsgrundlage

## Art der Daten und Personengruppen

Datenkategorie der gespeicherten Daten

Personaldaten  Kundendaten  Identifikations- und Adressdaten  
 Vertragsstammdaten  Kontokorrentdaten  IT-Nutzungsdaten  
 Sonstiges

Arbeitszeiten

Erläuterung

Betroffene Personengruppe

Mitarbeiter  Berater  Private Kunden  
 Lieferanten  Patienten  Sonstiges

Verarbeitung besonderer Arten von Daten

Gesundheit  Sexualität  Rassistische- ethnische Herkunft  
 Gewerkschaftszugehörigkeit  Religiösen o. philosophischen Überzeugungen  Keine dieser Daten

Quelle: Fraunhofer Intranet



# Weiterführende Informationen

- [Rechtliches zu Forschungsdaten auf \*\*forschungsdaten.info\*\*](#)
- [Urheberrechtsgesetz](#)
- [Bundesdatenschutzgesetz](#)
- [EU-Datenschutzgrundverordnung](#)
- [Gutachten zu den rechtlichen Rahmenbedingungen von Forschungsdatenmanagement](#)



---

# ANFORDERUNGEN DER FÖRDERORGANISATIONEN

---

## Inhalt

Anforderungen der Europäischen Kommission in Horizon2020 und Horizon Europe

Anforderungen der Deutschen Forschungsgemeinschaft

Anforderungen des BMBF

# Forschungsdaten in Horizon2020

- Open Research Data Pilot gilt seit 2017 für alle beantragten Projekte
- Anforderungen dokumentiert in den „Guidelines on FAIR Data Management in Horizon2020“ und im Abschnitt 29.3 des Grant Model Agreements



[http://ec.europa.eu/research/participants/data/ef/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ef/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)



# Forschungsdaten in Horizon2020 - Projektantrag

- Angaben im Abschnitt „Research Data Management“ des Projektantrags
  - Welche Daten werden im Projekt generiert?
  - Wie sollen die Daten ausgewertet werden?
  - Wie und wo sollen die Daten veröffentlicht werden?
  - Wie werden Daten kuratiert und archiviert?
  - Welche Kosten fallen an?



# Forschungsdaten in Horizon2020 - Opt-Out-Möglichkeiten

- Vollständiges Opt-Out: Bei einem „vollständigen“ Opt-Out ist Art. 29.3 des Grant Agreement für das gesamte Projekt nicht anwendbar. Daher besteht weder eine Verpflichtung zur Erstellung eines Datenmanagement-Plans noch zur Veröffentlichung von Forschungsdaten.
- Teilweises Opt-Out für einzelne Datensätze durch Angabe der jeweiligen Opt-Out-Gründe im Datenmanagement-Plan (die Verpflichtungen für Daten bleiben aufrecht).



# Forschungsdaten in Horizon2020 - Valide Opt-Out-Gründe

- Unvereinbarkeit mit dem Schutz der Ergebnisse
- Geheimhaltung/Sicherheitsgründe
- Datenschutzgründe
- Unvereinbarkeit mit dem Hauptziel des Projekts
- Im Projekt werden keine Daten erzeugt/gesammelt
- Sonstige legitime Gründe



# Forschungsdaten in Horizon2020 - Datenmanagementplan

- Einreichung eines Datenmanagementplans (bis 6 Monate nach Projektstart)
- Fortschreibung des Datenmanagementplans zur Projektmitte und zum Abschluss
- Ausrichtung an den FAIR-Prinzipien: Guidelines on FAIR Data Management in Horizon2020
- Template vorhanden



# Forschungsdaten in Horizon2020 – Template des European Research Council

- Template für den Datenmanagementplan des European Research Council
- Template für den Datenmanagement basierend auf der Vorlage des ERC, angepasst auf Fordatis
- 6 Abschnitte

erc  
European Research Council  
Established by the European Commission

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

Project Acronym	Project Number	SEP
Oscar	210510196 (Proposal ID)	

Template for the ERC Open Research Data Management Plan (DMP). The following sections should describe how you plan to make the project data Findable, Accessible, Interoperable and Reusable (FAIR). Each of the following five issues should be addressed with a level of detail appropriate to the project.

+

SUMMARY (dataset<sup>1</sup> reference and name; origin and expected size of the data generated/collected; data types and formats)

- Interview data
- Format: Audio: mp3 with high transfer rate (320 kbps) either with mobile and app (e.g. Smart Recorder for Android or Dictaphone for Apple) or with a recorder.
- Transcriptions as Word/Excel-Formats transferred to PDF/A
- Workshop documentation: pictures in jpg created with mobiles,
- transcripts of results as tables as PDF/A
- Expected size:
- Audio: 1 h = 60 MB, 20 Interviews with 1,5 h Interview = 2 GB
- Pictures: 100 photos = 400 MB
- Transcriptions as PDF/A:
- 20 pages pro interview +
- 100 Seiten = 650 KB 20 Interviews with 20 pages = 2,6 MB
- Overall storage:
- During project the generated data is stored at the central content server of the Fraunhofer-Gesellschaft: [livelink](#). Only project members gain access to the server.
- The use of computer-assisted interview software will be considered.

<sup>1</sup> Based on: [Guidelines on FAIR Data Management in Horizon2020](#), v4.0, 26.07.2018, Annex 1  
<sup>2</sup> Several datasets may be included into a single DMP.

Page 2 of 6

ERC OPEN RESEARCH DATA MANAGEMENT PLAN (DMP)

1. MAKING DATA FINDABLE (dataset description: metadata, persistent and unique identifiers e.g., DOI)

Dataset description will be done by an added .txt-File that describes the data by a text, so that a potential re-user gets context and can evaluate the data. Additionally the metadata that are part of the Fordatis Application profile will be documented in a .txt-File "Metadata". Relevant additional metadata should get entered, too.

By the time of publication the data(sets) will get an DOI as unique identifier.

2. MAKING DATA OPENLY ACCESSIBLE (which data will be made openly available and if some datasets remain closed; the reasons for not giving access; where the data and associated metadata, documentation and code are deposited (repository?); how the data can be accessed (are relevant software tools/methods provided?))

Data that are made openly available within this project are the interview-records in mp3-Format, the transcriptions of the interviews and the documentations of the workshops. If the interviewed persons don't agree, the interview data will not get published. Pictures that are taken to document workshop results will not get published as own data publications because the complete transcriptions are more meaningful.

The MP3-Data can get opened by each mp3-player, the PDF/A-documents can get opened by each PDF-Reader. No special, proprietary software is needed to open the data.

The generated data will be published by the end of the project via the Fraunhofer Repository "Fordatis" (currently under construction, expected to be available at the beginning of 2019).

The datasets will get described by the Fordatis application profile, based on the datacite-standard. This standard contains 16 mandatory fields and 23 optional fields to describe the data in a standardized way. Additionally there is a field to describe the software that is needed to open the data.

Page 3 of 6

[https://erc.europa.eu/sites/default/files/document/file/ERC\\_info\\_document-Open\\_Research\\_Data\\_and\\_Data\\_Management\\_Plans.pdf](https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf)



# Forschungsdaten in Horizon2020

- Veröffentlichung der Daten in einem kostenfreien Repository  
→ später mehr



# Anforderungen in Horizon Europe

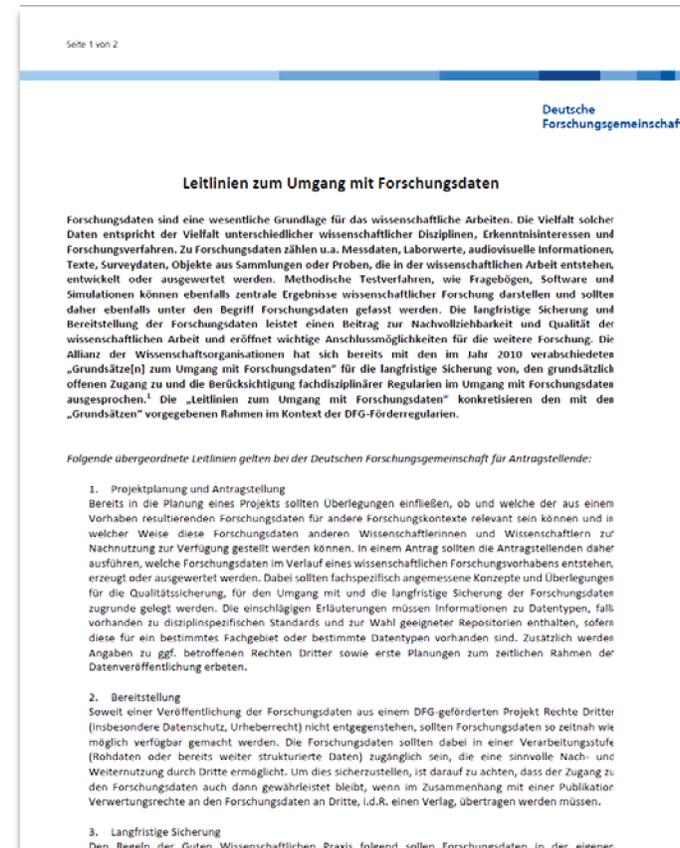


- Open Data by Default
- Verpflichtend: Das Veröffentlichen von Forschungsdaten (so offen wie möglich, so geschlossen wie nötig)
- In Übereinstimmung mit den FAIR-Prinzipien
- Aufmerksamkeit soll auch auf die Langzeitarchivierung gelegt werden.
- Datenmanagementpläne verpflichtend.



# Anforderungen der DFG

## ■ DFG-Leitlinien zum Umgang mit Forschungsdaten



[https://www.dfg.de/download/pdf/foerd-erung/antragstellung/forschungsdaten/r-ichtlinien\\_forschungsdaten.pdf](https://www.dfg.de/download/pdf/foerd-erung/antragstellung/forschungsdaten/r-ichtlinien_forschungsdaten.pdf)



# Anforderungen der DFG – Projektplanung und Antragstellung

Der Antrag soll folgende Informationen enthalten:

- Angaben über die Forschungsdaten (Informationen zu Datentypen), die im Verlauf eines wissenschaftlichen Forschungsvorhabens entstehen, erzeugt oder ausgewertet werden. Falls vorhanden zu disziplinspezifischen Standards und zur Wahl geeigneter Repositorien, sofern diese für ein bestimmtes Fachgebiet oder bestimmte Datentypen vorhanden sind.
- Dabei sollen fachspezifische, angemessene Konzepte und Überlegungen für die Qualitätssicherung, für den Umgang mit und die langfristige Sicherung der Forschungsdaten zugrunde gelegt werden.



# Anforderungen der DFG – Bereitstellung der Daten

- Wann?: Daten sollen so zeitnah wie möglich verfügbar gemacht werden.
- Wie?: Daten sollen in einer Verarbeitungsstufe zur Verfügung gestellt werden, die eine sinnvolle Nach- und Weiternutzung durch Dritte ermöglicht.
- Ausnahme: wenn die Rechte Dritter (Datenschutz, Urheberrecht) die Veröffentlichung nicht verhindern.
- Wo?: Fachliche oder institutionelle Repositorien



# Anforderungen der DFG – Langfristige Sicherung

„Den Regeln der Guten Wissenschaftlichen Praxis folgend, sollen Forschungsdaten in der eigenen Einrichtung oder in einer fachlich einschlägigen überregionalen Infrastruktur für **mindestens 10 Jahre** archiviert werden.“

[https://www.dfg.de/foerderung/grundlagen\\_rahmenbedingungen/gwp/](https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/)



# Anforderungen der DFG - Unterstützungsangebote

- Unterstützung und Beratung
- Finanzielle Unterstützung:
  - Kosten, die durch **Aufbereitung** von Forschungsdaten für eine Anschlussnutzung bzw. für die **Überführung in existierende Infrastrukturen** entstehen, können mit dem Antrag eingeworben werden.
  - Es können Mittel zur Finanzierung derjenigen Kosten eingeworben werden, die für die **Nutzung einschlägiger Infrastrukturen** entstehen. Gefördert werden können Personalkosten, projektspezifische Hard- und Software sowie Nutzungsgebühren.

# Anforderungen der DFG – Appell an die Fachcommunities

- Appell an die Fachcommunities
  - Erarbeitung disziplinspezifischer Regularien
  - Anerkennung der Leistung der Verfügbarmachung von Forschungsdaten

# Anforderungen des BMBF

- Anforderungen an das Forschungsdatenmanagement werden in den einzelnen Ausschreibungen bekannt gegeben.
- Übergreifende Anforderungen gibt es bezüglich des Forschungsdatenmanagements nicht.



# Weiterführende Informationen

- [Guidelines on FAIR Data Management](#)
- [Template für den Datenmanagementplan des European Research Council](#)
- [DFG-Leitlinie zum Umgang mit Forschungsdaten](#)
- [Denkschrift "Sicherung guter wissenschaftlicher Praxis"](#)



---

# PLANUNG DES FORSCHUNGSDATENMANAGEMENTS

---

## Inhalt

Welche Aspekte muss ich bei der Planung berücksichtigen?

Datenmanagementplan



# Welche Aspekte müssen bei der Planung bedacht werden?

- Daten: Welche Daten werden mit welcher Methode erhoben? Wie sollen sie analysiert werden? Welche Software ist hierzu notwendig? Welches Format haben meine Daten? Welche Größe haben meine Daten?
- Zugriff: Welche Personen sollen in welcher Projektphase Zugriff auf die Daten haben? Wo sollen die Daten abgelegt werden?
- Dokumentation: Wie sollen meine Daten dokumentiert und beschrieben werden? Welche Standards werden hierfür eingesetzt?
- Veröffentlichung: Wie und wo sollen Daten veröffentlicht werden? Unter welcher Lizenz?
- Langzeitarchivierung: Wo sollen die Daten archiviert werden? Ist eine Konvertierung in ein anderes Format notwendig? Wie sieht die Back-Up-Strategie aus?
- Kosten: Welche Kosten entstehen für das Forschungsdatenmanagement?



# Welche Aspekte müssen bei der Planung bedacht werden?

- Rechtliche und ethische Einschränkungen: Datenschutz, gute wissenschaftliche Praxis, IPR-Regelungen
- Festlegung von Rollen und Verantwortlichkeiten
- Anforderungen der Forschungsförderer



# Datenmanagementplan

- Ein Datenmanagement ist ein **formales Dokument**, das beschreibt, wie in **Projekten** mit **Forschungsdaten umgegangen** wird, um die Arbeitsergebnisse während der Laufzeit zu sichern und zur weiteren Nutzung durch die Forschergemeinschaft bereitzustellen.
- **Ziel** ist es alle Aspekte des Forschungsdatenmanagements möglichst frühzeitig zu planen, um so gutes Datenmanagement sicherzustellen.
- Datenmanagementpläne werden im Verlauf eines Projekts **fortgeschrieben**: in den einzelnen Projektphasen können unterschiedliche Inhalte festgehalten werden. Der Datenmanagementplan wird fortwährend angepasst.
- Einige Förderorganisationen, wie EU und DFG fordern Datenmanagementpläne als **Deliverables** ein.



# Core Requirements for Data Management Plan by Science Europe

**Introduction to the Core Requirements for Data Management Plans**

Research funding organisations and research organisations increasingly require researchers to develop data management plans. These plans support the researcher in considering all relevant aspects of data management from the very beginning of a research project. A DMP should stimulate researchers to think about optimal handling, organising, documenting, and storing of their data.

Currently, there is a lot of variation in research data management policies. Many research funding organisations, research organisations, and research communities have developed their own rules and templates. This can be confusing for researchers and is especially problematic as many researchers acquire their funding from different sources; they are increasingly confronted with different grant requirements and institutional policies. There is an urgent need to align data management policies in order to provide more clarity for researchers. DMPs should not be a bureaucratic burden for them, but a useful means of support when planning and conducting a research project.

The following list presents six topics that should be covered in DMPs, each specified with several guiding questions. These topics and questions for setting up a DMP form the core requirements that every research funding organisation should request in order for the researcher to develop a useful DMP. The order of the core requirements can be changed according to specific needs and organisational focal points. However, all six core requirements need to be addressed in a DMP.

An example template providing guidance on which aspects to further consider in a DMP can be found on Page 17 of this guide.

**CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS**

When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:

- 1. Data description and collection or re-use of existing data**
  - How will new data be collected or produced and/or how will existing data be re-used?
  - What data (for example the kinds, formats, and volumes) will be collected or produced?
- 2. Documentation and data quality**
  - What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
  - What data quality control measures will be used?
- 3. Storage and backup during the research process**
  - How will data and metadata be stored and backed up during the research process?
  - How will data security and protection of sensitive data be taken care of during the research?
- 4. Legal and ethical requirements, codes of conduct**
  - If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
  - How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
  - How will possible ethical issues be taken into account, and codes of conduct followed?

8

9

- Datenbeschreibung und Sammlung/Nachnutzung von Daten
- Dokumentation und Datenqualität
- Speicherung und Backup während der Projektphase
- Rechtliche und ethische Anforderungen, Codes of Conduct
- Data Sharing und Langzeitarchivierung
- Verantwortlichkeiten im FMD, Ressourcen

<https://www.scienceurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/>



# Datenmanagementplan

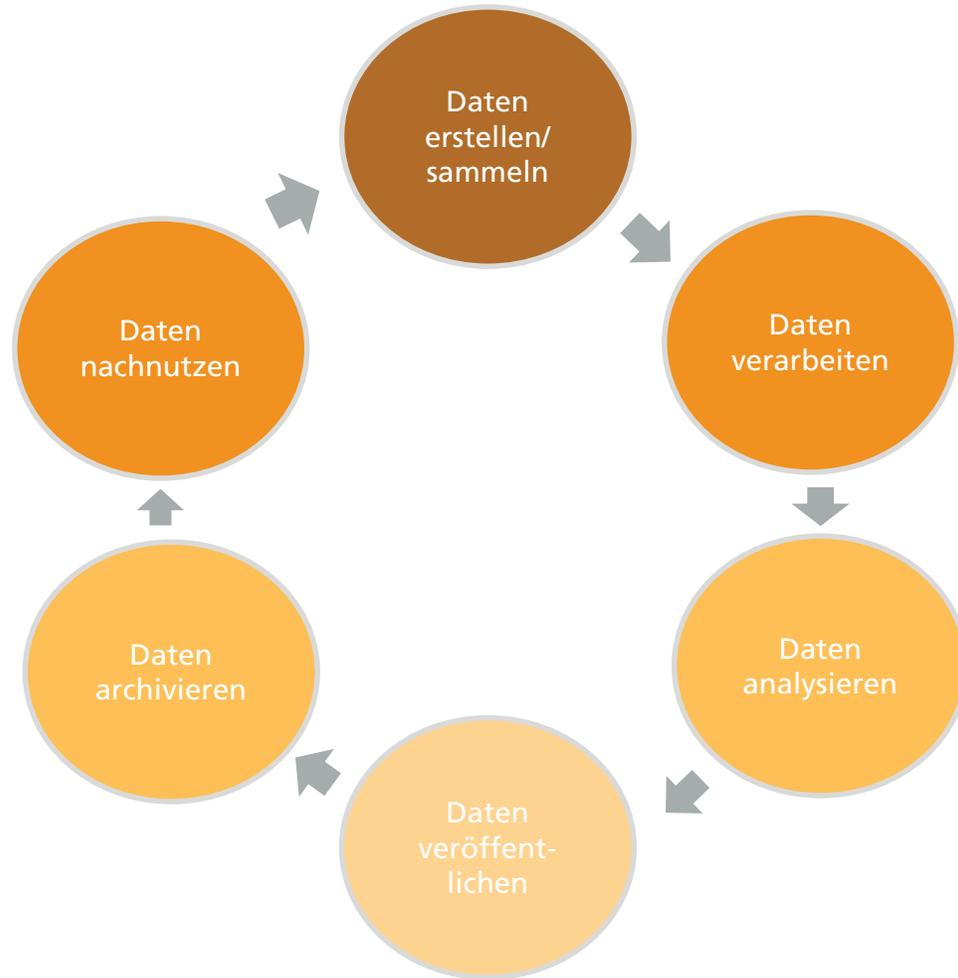
**Phasen-orientiert**

**Aktivitäts-orientiert**

**FAIR-Prinzipien-orientiert**

**IP-Protection-orientiert**

# Datenmanagementplan



## Phasen-Orientiert

Wie werden Daten erstellt?

Wie werden Daten analysiert?

Wie und wo werden Daten veröffentlicht?

Wie werden Daten archiviert?

Wie können die Daten nachgenutzt werden?

# Datenmanagementplan

## Table of Contents

1. Data collection and documentation .....	4
1.1 What data will you collect, observe, generate or re-use? .....	4
1.2 How will the data be collected, observed or generated? .....	6
What standards, methodologies or quality assurance processes will you use? .....	6
How will you organize your files and handle versioning? .....	6
1.3 What documentation and metadata will you provide with the data? .....	8
2. Ethics, legal and security issues.....	10
2.1 How will ethical issues be addressed and handled? .....	10
2.2 How will data access and security be managed? .....	14
2.3 How will you handle copyright and Intellectual Property Rights issues? .....	16
3. Data storage and preservation .....	18
3.1 How will your data be stored and backed-up during the research? .....	18
3.2 What is your data preservation plan?.....	20
4. Data sharing and reuse .....	22
4.1 How and where will the data be shared? .....	22
4.2 Are there any necessary limitations to protect sensitive data? .....	24
4.3 Are there any necessary limitations to protect sensitive data? .....	25
4.4 I will choose digital repositories maintained by a non-profit organisation .....	25
5. External useful resources .....	26

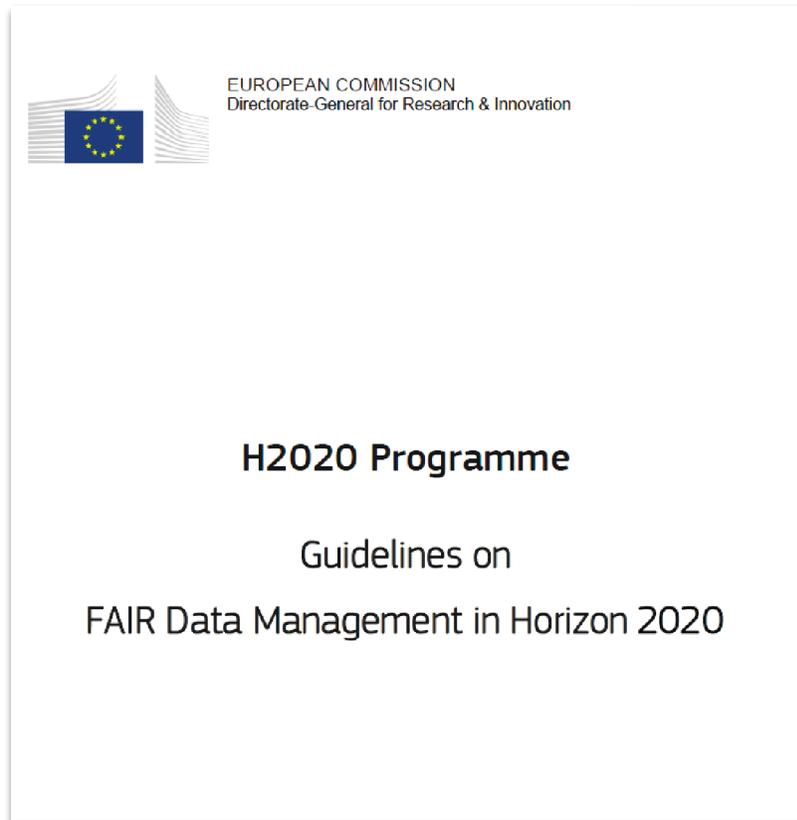
## Aktivitäts-Orientiert

Welche Aktivitäten fallen im Forschungsdatenmanagement an?

Wie sollen diese ausgeführt werden?



# Datenmanagementplan



## FAIR-Prinzipien-Orientiert

Wie werden Daten auffindbar gemacht?

Wie werden Daten zugänglich gemacht?

Wie werden Daten interoperable gemacht?

Wie werden Daten nachnutzbar gemacht?

---

# EXKURS: KOSTEN

---

## Inhalt

Relevante Aspekte

Wie schätze ich Kosten für das Forschungsdatenmanagement?

Kostenschätzungstool des UK Data Archives



# Kosten

- Kosten für Forschungsdatenmanagement müssen von vornerein in Projekten miteingeplant werden.
- Notwendige Ressourcen = Menschen, Fähigkeiten, Ausrüstung, Infrastruktur, Tools um Daten zu managen, zu dokumentieren, zu organisieren, zu speichern und Zugriff zu ermöglichen



# Kosten – Wie schätze ich Kosten ab?



# Kosten – UK Data Archive Data Management Costing Tool

ACTIVITY	COMMENTS AND SUGGESTIONS	✓	COST
<b>Data description</b> <ul style="list-style-type: none"> <li>Are data in a spreadsheet or database clearly marked with variable and value labels, code descriptions, missing value descriptions, etc.?</li> <li>Are labels consistent?</li> <li>Do textual data like interview transcripts need description of context, e.g. included as a heading page?</li> </ul>	<ul style="list-style-type: none"> <li>if data description is carried out as part of data creation, data input or data transcription – low or no additional cost</li> <li>if needed to be added afterwards – higher cost</li> <li>codebooks for datasets can often be easily exported from software packages</li> </ul>		
<b>Data cleaning</b> <ul style="list-style-type: none"> <li>Do quantitative data need to be cleaned, checked or verified before sharing, e.g. check validity of codes used, check for anomalous values?</li> <li>Will data match documentation, e.g. same number of variables, cases, records, files?</li> <li>Does textual information in data need to be spell-checked?</li> </ul>	<ul style="list-style-type: none"> <li>if carried out as part of data entry and preparation before data analysis – low or no additional cost</li> <li>if needed afterwards – higher cost</li> </ul>		
<b>Documentation</b> <ul style="list-style-type: none"> <li>Do you have documentation for the data that describes the context and methodology of how data were gathered, created, processed and quality controlled?</li> </ul>	<ul style="list-style-type: none"> <li>often essential contextual and methods documentation will be written up in publications and reports</li> <li>if all data creation steps are well documented and documentation is kept well organised during research – low or no additional cost</li> <li>if documentation to be written or compiled specifically afterwards – higher cost</li> </ul>		
<b>Metadata</b>	- completing a UK Data Archive deposit form		

<https://www.ukdataservice.ac.uk/manage-data/plan/costing>



# Weiterführende Informationen

- [Kostenplanungstool für das Datenmanagement des UK Data Archives](#)
- [Open Research Data Pilot in Horizon2020](#)
- [Grant Model Agreement in Horizon2020 \(Version 5, Oktober 2017\)](#)
- [Hilfestellungen bei Datenmanagementplänen in Horizon2020](#)
- [Fraunhofer Template für Datenmanagementpläne in Horizon2020 \(deutsch u. englisch\)](#)
- [Opt-Out-Gründe in Horizon2020](#)
- [DFG-Leitlinien zum Umgang mit Forschungsdaten](#)



# Weiterführende Informationen

- [Anforderungen für DFG-Projektanträge](#)
- [Fraunhofer Geheimhaltungsvereinbarung](#)
- [Template für den Datenmanagementplan](#)
- [RDMO – Research Data Management Organiser](#)
- [DMP-Tool](#)
- [UK Data Archive Costing Tool](#)
- [Practical Guide to the International Alignment of Research Data Management](#)



---

# ÜBERGREIFENDE AUFGABEN – DATEN DOKUMENTIEREN

---

## Inhalt

Daten dokumentieren – Warum?

Metadaten als Instrument zur strukturierten Beschreibung

Best-Practice-Empfehlung



# Forschungsdaten dokumentieren – Warum?

- Dokumentation ist Voraussetzung für die (Nach)nutzung von Daten:
  - Daten können während der Projektphase gut gemanagt werden.
  - Daten können im richtigen Kontext verstanden werden.
  - Daten können wiedergefunden werden.
  - Daten können entsprechend ihrer Lizenz nachgenutzt werden.
  - Daten können mit geeigneter Software bearbeitet und ausgewertet werden.

→ Forschungsdaten stehen zur Erst- und Nachnutzung zur Verfügung.



# Forschungsdaten dokumentieren – Metadaten

Metadaten = Daten/Informationen, die in strukturierter Form analoge oder digitale Forschungsdaten (Objekte) beschreiben.

Metadatenschema = Zusammenstellung von Metadaten-Attributen, um eine Ressource zu beschreiben. Es enthält Definitionen darüber, welche Information bei den einzelnen Attributen angegeben werden sollen und kann festgelegte Werte zur Beschreibung enthalten.



# Forschungsdaten dokumentieren – Metadaten

Metadatenstandard = Definierte und gut dokumentierte Zusammenstellung von Metadaten, die innerhalb einer Community breit eingesetzt wird.

Generische Standards

[Dublin Core](#)

[Data Cite](#)

[Mets](#)

Fachspezifische Standards

Sozialwissenschaften: [DDI](#)

Geisteswissenschaften: [TEI](#)

Ingenieurwissenschaften: NeXus

# Forschungsdaten dokumentieren – Daten beschreiben

- Dokumentation vom Zeitpunkt der Datenerhebung an
- .txt-Datei mit folgenden Angaben:
  - Deskriptive Metadaten: Titel, Autoren, Angaben zum Projekt etc.
  - Abstract: Beschreibung des Projekts und der Datensätze, Erhebungs- und Analysemethoden
  - Technische Dokumentation: Technische Angaben zu den Datensätzen: Verwendete Software/Hardware, Parameter
  - Fachspezifische Metadaten
- → die Dokumentation wird bei Veröffentlichung relevant, dort sind Metadaten mitanzugeben.



# Forschungsdaten dokumentieren – Versionierung

- Festlegung, wie viele Versionen einer Datei gespeichert werden sollen, für wie lange und wie die Versionen organisiert werden sollen.
- Identifikation von Meilenstein-Versionen, die behalten werden sollen
- Versionen sollen anhand einer Namenskonvention eindeutig identifizierbar sein
  - Datum im Namen oder innerhalb der Datei
  - Versionsnummer im Dateinamen (v1, v2, v3 or 00.01, 01.00)
  - Versionsbeschreibung im Dateinamen (draft, final)
- Dateihistorie, Versionierungstabellen oder Notizen können in die Datei eingebettet werden
- Versionierung kann weiterhin erfolgen durch:
  - Versionierung innerhalb der verwendeten Software
  - durch die Verwendung von Versionierungstools wie Subversion (SVN)
  - Durch Verwendung von File Sharing Diensten wie Dropbox, Google Docs oder Amazon S3



# Forschungsdaten dokumentieren – Versionierung

- Es sollte dokumentiert werden, was genau gemacht wird, wenn eine neue Datenversion erhoben wird.
- Sollte zwischen mehreren Versionen ein Datei Verknüpfung bestehen, sollten diese dokumentiert werden
- Die Speicherorte sollten dokumentiert werden, es sollte regelmäßig eine Synchronisation erfolgen
- Stammdateien sollten definiert und in einem geeigneten Dateiformat gespeichert werden
- Identifizieren Sie einen einzelnen Speicherort für die Speicherung von Meilensteinen und Master-Versionen



# Forschungsdaten dokumentieren – Versionierung

VERSION CONTROL TABLE FOR A DATA FILE			
<b>Title:</b>	Vision screening tests in Essex nurseries		
<b>File Name:</b>	VisionScreenResults_00_05		
<b>Description:</b>	Results data of 120 Vision Screen Tests carried out in 5 nurseries in Essex during June 2007		
<b>Created By:</b>	Chris Wilkinson		
<b>Maintained By:</b>	Sally Watsley		
<b>Created:</b>	04/07/ 2007		
<b>Last Modified:</b>	25/11/ 2007		
<b>Based on:</b>	VisionScreenDatabaseDesign_02_00		
VERSION	RESPONSIBLE	NOTES	LAST AMENDED
00_05	Sally Watsley	Version 00_03 and 00_04 compared and merged by SW	25/11/2007
00_04	Vani Yussu	Entries checked by VY, independent from SK	17/10/2007
00_03	Steve Knight	Entries checked by SK	29/07/2007
00_02	Karin Mills	Test results 81-120 entered	05/07/2007
00_01	Karin Mills	Test results 1-80 entered	04/07/2007

Quelle: <https://ukdataservice.ac.uk/media/622417/managingsharing.pdf>



# Weiterführende Informationen

- [Metadaten-Standards für Forschungsdaten](#)
- [Leitmodell für die Langzeitarchivierung von Forschungsdaten \(OAIS\)](#)
- [TRAC Kriterien und Checkliste für ein vertrauenswürdigen Archiv](#)
- [DRAMBORA \(Digital Repository Audit Method\)](#)
- [Data Seal of Approval](#)
- [Nestor Kriterienkatalog für Langzeitarchivierung](#)
- [PREMIS \(Preservation Metadata: Implementation Strategies\)](#)



# Weiterführende Informationen

- [FORCE11-Prinzipien zur Datenzitierung](#)
- [Datenzitation nach da | ra](#)
- [LMER-Standard](#)
- [METS \(Metadata Encoding and Transmission Standard\)](#)
- [Metadata-Universe](#)



---

# DATEN VERÖFFENTLICHEN

---

## Inhalt

Hinderungsgründe und Nutzen

Wo? Repositorien und Datenjournale

Wie? Persistente Identifikatoren und Lizenzen

Wie? Umsetzung der FAIR-Prinzipien



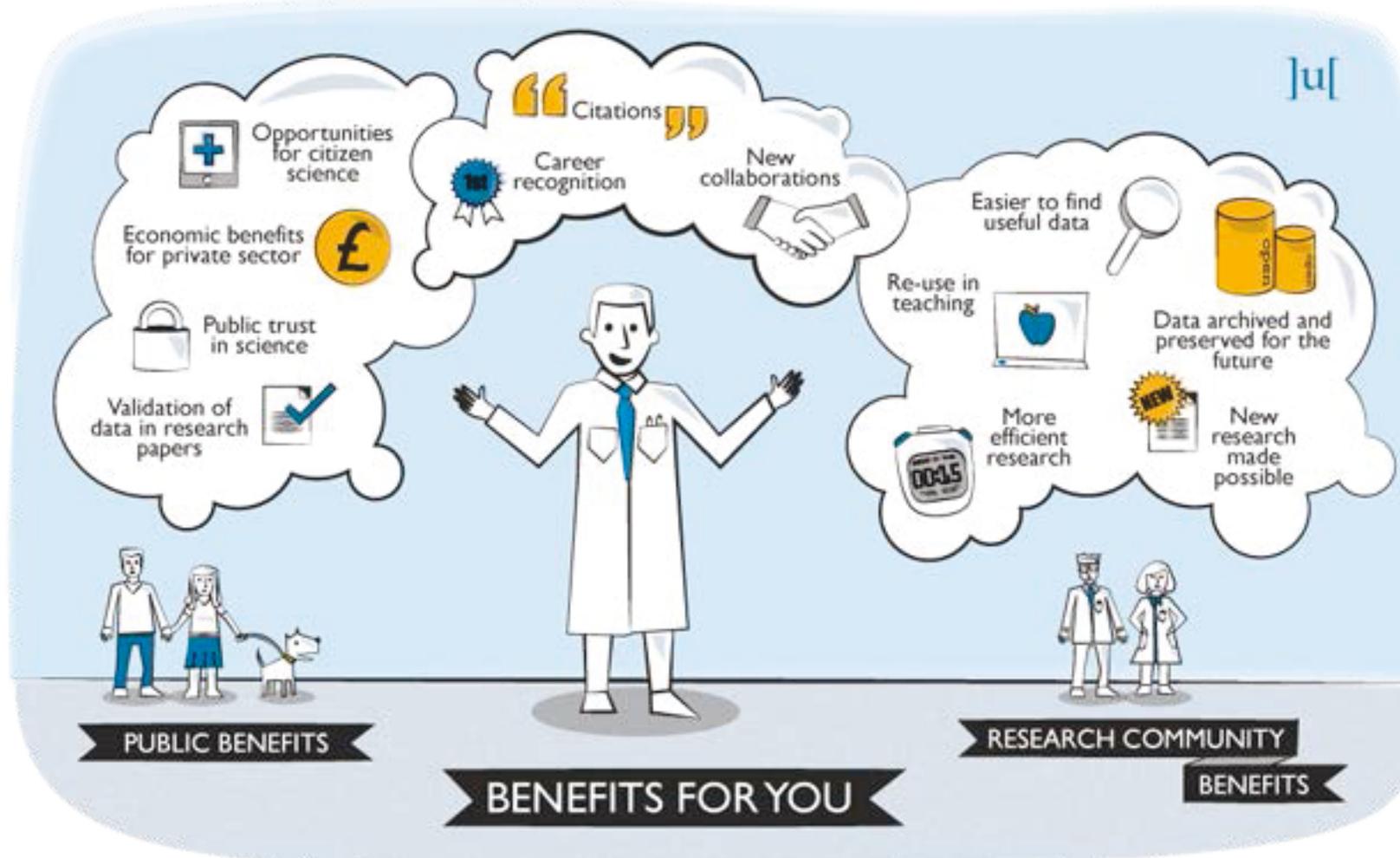
# Daten veröffentlichen – mögliche Hinderungsgründe

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

<http://data.dev8d.org/devbingo/bingo.php?n=1&w=4&h=4&title=%22Open+Data+Excuse%22+Bingo&tag=#openDataExcuses&statements=Terrorists+will+use+it>



# Daten veröffentlichen – Wo ist der Benefit für mich?



Why share archaeology data? Benefits for data sharing/ Journal of Open Archaeology Data (JOAD)/ Published by Ubiquity Press CC By 3.0

# Daten veröffentlichen - Welche Daten darf ich nicht veröffentlichen?

Daten, die personenbezogene Daten enthalten	Nein Ja: Person stimmt der Veröffentlichung zu Ja: nach absoluter Anonymisierung
Daten aus Projekten mit Geheimhaltungsvereinbarungen	Nein
Daten, die patentierbare Erfindungen beschreiben, die auch patentiert werden sollen.	Nein
Daten, deren Veröffentlichung nicht alle Mitautoren zustimmen	Nein



# Daten veröffentlichen – Wo?

Datenjournale

Daten-Repositoryn

Daten-Supplement



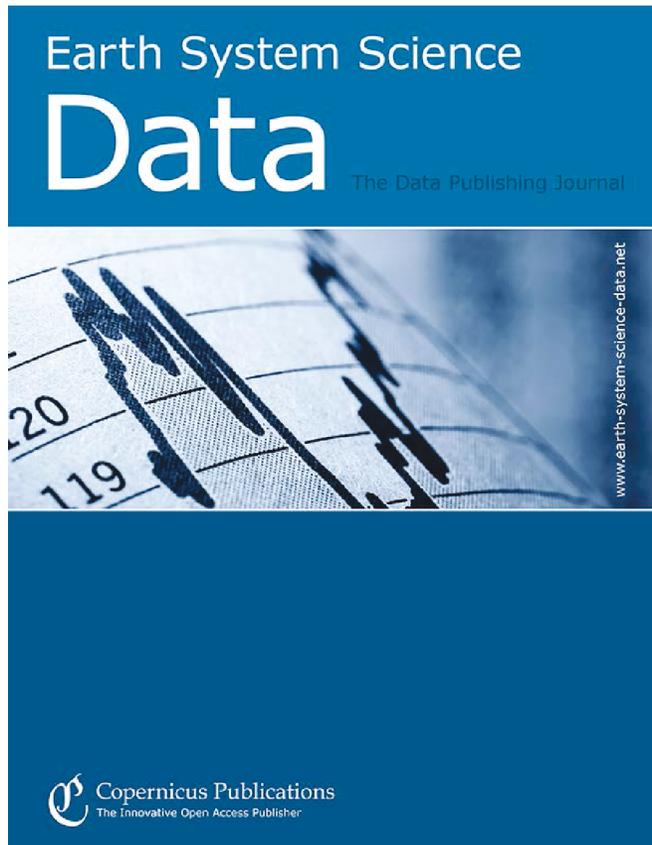
# Daten veröffentlichen - Datenjournale

- Datenjournale werden von klassischen Verlagen und Fachverbänden herausgegeben, auch die Rolle der Selbstpublikation steigt.
- Meist wird nur die Dokumentation des Datensatzes (Information zu Datenerhebung, Charakteristiken und Funktionen sowie potenzielle Nachnutzungsmöglichkeiten) im Journal publiziert. Die Forschungsdaten selbst liegen in einem externen Repository. Data Journal und Repository sind getrennt.
- Wissenschaftliche Entdeckungen sind für Datenveröffentlichungen keine Voraussetzung, die hochqualitativen Daten stehen im Vordergrund.
- Möglichkeit der Qualitätssicherung, als (interaktive) (Open) Peer-Reviews



# Daten veröffentlichen - Datenjournale

- Beispiel „Earth System Science“



<https://www.earth-system-science-data.net/>



# Daten veröffentlichen - Datenjournale

Earth Syst. Sci. Data, 11, 881-894, 2019  
<https://doi.org/10.5194/essd-11-881-2019>  
© Author(s) 2019. This work is distributed under the Creative Commons Attribution 4.0 License.

ESSD | Articles | Volume 11, issue 2

Copernicus Publications  
The Innovative Open Access Publisher

Search articles  
Search  
Author

Article Assets Peer review Metrics Related articles

21 Jun 2019

Data description paper

## A dataset of 30 m annual vegetation phenology indicators (1985–2015) in urban areas of the conterminous United States

Xuecao Li<sup>1</sup>, Yuyu Zhou<sup>1</sup>, Lin Meng<sup>1</sup>, Ghassem R. Asrar<sup>2</sup>, Chaohun Lu<sup>3</sup>, and Qiusheng Wu<sup>4</sup>

<sup>1</sup>Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA, 50011, USA  
<sup>2</sup>Joint Global Change Research Institute, Pacific Northwest National Lab, College Park, MD, 20740, USA  
<sup>3</sup>Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA, 50011, USA  
<sup>4</sup>Department of Geography, University of Tennessee, Knoxville, TN, 37996, USA

Received: 13 Jan 2019 – Discussion started: 14 Feb 2019 – Revised: 31 May 2019 – Accepted: 05 Jun 2019 – Published: 21 Jun 2019

Abstract

Medium-resolution satellite observations show great potential for characterizing seasonal and annual dynamics of vegetation phenology in urban domains from local to regional and global scales. However, most previous studies were conducted using coarse-resolution data, which are inadequate for characterizing the spatiotemporal dynamics of vegetation phenology in urban domains. In this study, we produced an annual vegetation phenology dataset in urban ecosystems for the conterminous United States (US), using all available Landsat images on the Google Earth Engine (GEE) platform. First, we characterized the long-term mean seasonal pattern of phenology indicators of the start of season (SOS) and the end of season (EOS), using a double logistic model. Then, we identified the annual variability of these two phenology indicators by measuring the difference of dates when the vegetation index in a specific year reaches the same magnitude as its long-term mean. The derived phenology indicators agree well with in situ observations from the PhenoCam network and Harvard Forest. Comparing with results derived from the moderate-resolution imaging spectroradiometer (MODIS) data, our Landsat-derived phenology indicators can provide more spatial details. Also, we found the temporal trends of phenology indicators (e.g., SOS) derived from Landsat and MODIS are consistent overall, but the Landsat-derived results from 1985 offer a longer temporal span compared to MODIS from 2001 to present. In general, there is a spatially explicit pattern of phenology indicators from the north to the south in cities in the conterminous US, with an overall advanced SOS in the past 3 decades. The derived phenology product in the US urban domains at the national level is of great use for urban ecology studies for its medium spatial resolution (30m) and long temporal span (30 years). The data are available at <https://doi.org/10.6084/m9.figshare.7685645.v5>.

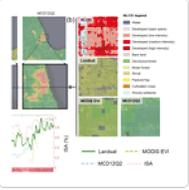
How to cite. Li, X., Zhou, Y., Meng, L., Asrar, G. R., Lu, C., and Wu, Q.: A dataset of 30 m annual vegetation phenology indicators (1985–2015) in urban areas of the conterminous United States, Earth Syst. Sci. Data, 11, 881-894, <https://doi.org/10.5194/essd-11-881-2019>, 2019.

Download  
PDF  
FULL-TEXT XML  
Supplement (897 KB)

Short summary  
We generated a long-term (1985–2015) and medium-resolution (30 m) product of phenology...  
Read more

Citation  
BibTeX  
EndNote

Share  
ResearchGate  
CrossRef  
ORCID  
Twitter  
Facebook  
LinkedIn

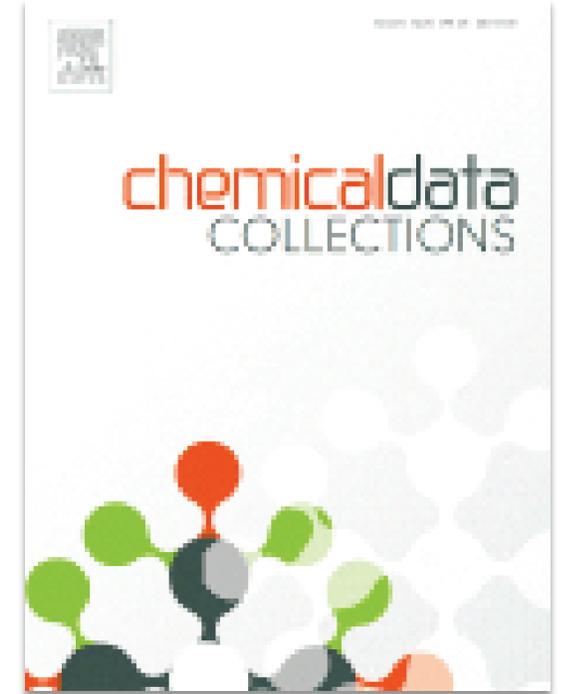


<https://essd.copernicus.org/articles/11/881/2019/>



# Daten veröffentlichen – Datenjournale Listen

- [Datenjournalliste](#) auf [forschungsdaten.org](http://forschungsdaten.org) (21 Einträge)
- [Servicetools](#) der Universitätsbibliothek Michigan (ca. 70 Einträge)
- [Humboldt-Universität](#) zu Berlin (28 Einträge)
- [Finnisches Komitee für Forschungsdaten](#) (17 Einträge)



<https://www.journals.eisevier.com/chemical-data-collections>



# Forschungsdaten veröffentlichen – Verzeichnis von Repositorien

- Verzeichnis von Repositorien re3data: <https://www.re3data.org/>
- Suche und Browsing nach Repositorien möglich
- Umfangreiche Beschreibung jedes Repositoriums

The screenshot shows the re3data.org search interface. On the left is a 'Filter' sidebar with categories like Subjects, Content Types, Countries, AID systems, API, Data access, Data access restrictions, Database access, Database access restrictions, Database licenses, Data licenses, Data upload, Data upload restrictions, Enhanced publication, Institution responsibility type, and Institution type. The main search area has a search bar containing 'figShare' and a 'Search' button. Below the search bar are navigation buttons for 'Previous', '1', and 'Next', and a 'Sort by' dropdown. The search results show 'Found 20 result(s)' and a detailed entry for 'University of Adelaide figshare'. This entry includes subject tags (Humanities and Social Sciences, Life Sciences, Natural Sciences, Engineering Sciences), content type tags (Standard office documents, Images, Structured graphics, Audiovisual data, Scientific and statistical data formats, Source code, Software applications), and a country tag (Australia). A descriptive paragraph follows: 'Figshare has been chosen as the University of Adelaide's official data and digital object repository with unlimited local storage. All current staff and HDR students can access and publish research data and digital objects on the University of Adelaide's Figshare site. Because Figshare is cloud-based, you can access it anywhere and at any time.'

Quelle: <https://www.re3data.org/>



# Forschungsdaten veröffentlichen – Verzeichnis von Repositorien

General	Institutions	Terms	Standards
Name of repository	figshare		
Repository URL	<a href="http://figshare.com/">http://figshare.com/</a>		
Subject(s)	<a href="#">Humanities and Social Sciences</a> <a href="#">Life Sciences</a> <a href="#">Natural Sciences</a> <a href="#">Engineering Sciences</a>		
Description	figshare allows researchers to publish all of their research outputs in an easily citable, sharable and discoverable manner. All file formats can be published, including videos and datasets. Optional peer review process. figshare uses creative commons licensing.		
Contact	<a href="http://figshare.com/contact">http://figshare.com/contact</a> info@figshare.com		
Content type(s)	<a href="#">Standard office documents</a> <a href="#">Images</a> <a href="#">Structured graphics</a> <a href="#">Audiovisual data</a> <a href="#">Raw data</a> <a href="#">Plain text</a> <a href="#">Archived data</a> <a href="#">Scientific and statistical data formats</a> <a href="#">Source code</a>		
Keyword(s)	<a href="#">multidisciplinary</a> <a href="#">data collection platform</a>		
Persistent Identifier(s) of the repository	RRID:SCR_004328 RRID:nlx_34569		
Repository type(s)	other		
Mission statement for designated community	<a href="http://figshare.com/about">http://figshare.com/about</a>		
Research data repository language(s)	eng		
Data and/or service provider	dataProvider serviceProvider		

[Back to search](#) [Submit a change request](#) [Get a badge](#)

Quelle: <https://www.re3data.org/>

# Forschungsdaten veröffentlichen – Verzeichnis von Repositorien

General	Institutions	Terms	Standards
Persistent identifier system(s)		DOI	
Name of the repository software		other	
Versioning		yes	
Data citation guideline		<a href="https://knowledge.figshare.com/articles/item/how-to-share-cite-or-embed-my-data">https://knowledge.figshare.com/articles/item/how-to-share-cite-or-embed-my-data</a>	
Author identifier system(s)		ORCID	
Enhanced Publication		yes	
Quality management		yes	
<b>Application programming interfaces (1)</b>			
API type		OAI-PMH	
URL		<a href="https://api.figshare.com/v2/oai">https://api.figshare.com/v2/oai</a>	
<b>Metadata standards (2)</b>			
Metadata standard name		DataCite Metadata Schema	
Metadata standard scheme		DCC	
Metadata standard URL		<a href="http://www.dcc.ac.uk/resources/metadata-standards/datacite-metadata-schema">http://www.dcc.ac.uk/resources/metadata-standards/datacite-metadata-schema</a>	
Metadata standard name		Dublin Core	
Metadata standard scheme		DCC	
Metadata standard URL		<a href="http://www.dcc.ac.uk/resources/metadata-standards/dublin-core">http://www.dcc.ac.uk/resources/metadata-standards/dublin-core</a>	
<b>Remarks</b>			
Remarks		figshare is covered by Thomson Reuters Data Citation Index.c. figshare is partner of the	

Quelle: <https://www.re3data.org/>



# Veröffentlichung als Datensupplement zu wissenschaftlichen Artikeln

- Interpretativen Text-Publikationen mit können durch Datensupplemente angereichert werden.
- Die Publikation bezieht sich dabei auf Daten, die insbesondere der Unterstützung der interpretativen Publikation dienen.
- Bei diesem Publikationsmodell sind die Forschungsdaten am engsten an den traditionellen Publikationsprozess geknüpft.
- Traditionelles Modell: Ursprünglich wurden supplementäre Daten als Anhang zum interpretativen Artikel in Form von Abbildungen, Tabellen oder detaillierteren Angaben zur Methodik publiziert. Dabei wurden die Daten auf der gleichen Plattform wie auch die Textpublikation veröffentlicht, aber nicht individuell persistent und standardisiert adressiert. Bei diesem Modell sind die Daten Bestandteil des Artikels.
- Neues Modell: Hierbei können Forschungsdaten aber auch als eigenständige Objekte auf der Verlagsplattform oder zeitlich parallel zum interpretativen Artikel in einem externen Datenrepositorium publiziert werden. Die Referenzierbarkeit wird über Persistent Identifier gewährleistet. Im interpretativen Artikel wird dann diese Identifizierung angegeben. Zum Teil gibt es Kooperationen zwischen Verlagen und Datenrepositorien (z.B. Elsevier und PANGEA).



---

# DATEN LIZENZIEREN

---

## Inhalt

Lizenzierung

Creative Commons Lizenzen

Open Data Commons Lizenzen



# Lizenzierung



Eine Lizenzierung legt fest, wie Forschungsdaten nach der Veröffentlichung benutzt werden dürfen. Es besteht ein Lizenzvertrag zwischen dem Datenproduzenten und dem Nutzer.

# Lizenzierung



Eine Lizenzierung legt fest, wie Forschungsdaten nach der Veröffentlichung benutzt werden dürfen. Es besteht ein Lizenzvertrag zwischen dem Datenproduzenten und dem Nutzer.

# Lizenzierung

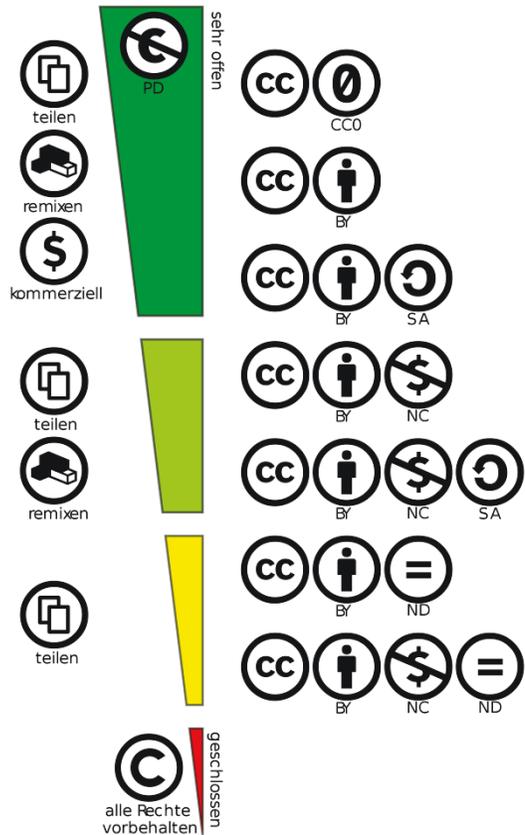
## Lizenz-Familien

Creative Commons-Lizenzen

Open Data Commons-Lizenzen



# Creative Commons-Lizenzen

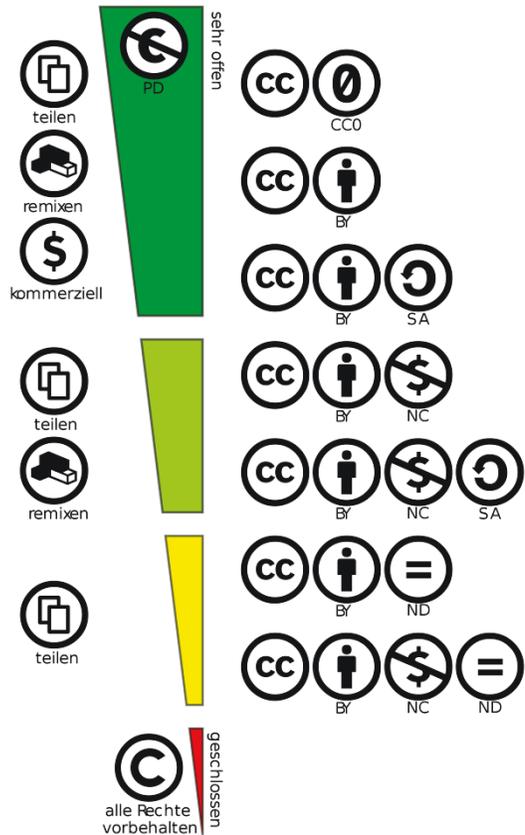


**CC-O:** Ein unter CC-O geschütztes Werk wird in die Gemeinfreiheit - auch genannt Public Domain – entlassen. Weltweit wird auf alle urheberrechtlichen und verwandten Schutzrechte verzichtet, soweit das gesetzlich möglich ist.

**CC-BY:** Das Material darf in jedwedem Format oder Medium vervielfältigt und weiterverbreitet werden. Das Material darf gemischt und verändert werden und zwar für beliebige Zwecke, sogar kommerziell. Voraussetzung ist, das der Name des Urhebers genannt wird.

[CC BY 4.0, Shaddim: original CC license symbols by Creative Commons via Wikimedia Commons](#)

# Creative Commons-Lizenzen

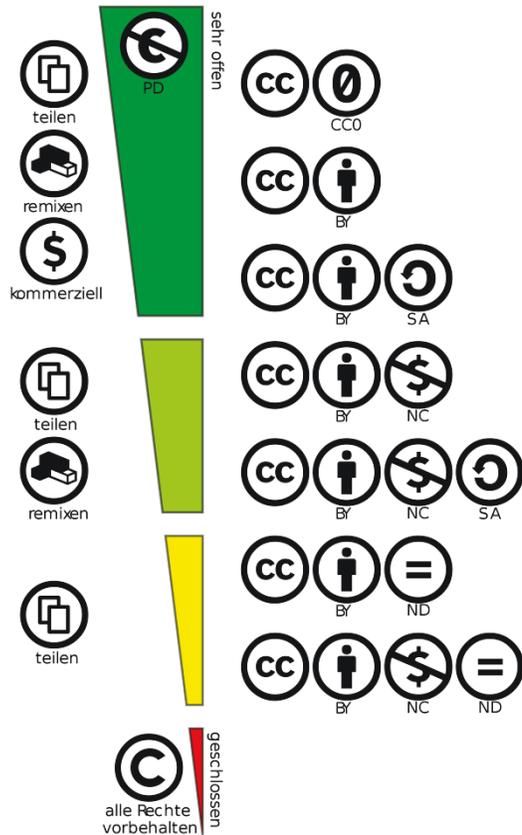


**CC-BY-SA:** Weitergabe von abgewandeltem Material nur unter dieser Lizenz.

**CC-BY-NC:** Schränkt die Nutzung von lizenziertem Material auf nicht-kommerzielle ein. Wirft Schwierigkeiten in der Auslegung (hinsichtlich Drittmittel- oder Auftragsforschung) auf und erscheint zu unbestimmt.

[CC BY 4.0, Shaddim: original CC license symbols by Creative Commons via Wikimedia Commons](#)

# Creative Commons-Lizenzen



**CC-BY-ND:** schließt Weitergabe „abgewandelten“ Materials aus. Attribution ist zu restriktiv, um eine breite Nutzung zu ermöglichen.

[CC BY 4.0, Shaddim: original CC license symbols by Creative Commons via Wikimedia Commons](#)

# Open Data Commons-Lizenzen

Open Data Commons-Lizenzen sind speziell auf Daten ausgelegt.

**PDDL** Public Domain Dedication and Licence: Daten werden in die Public Domain entlassen und somit „gemeinfrei“.

**ODC-By** Attribution for Data/Databases: Daten dürfen geteilt, weiterverarbeitet und verändert werden, solange der Urheber genannt wird.

**ODbL** Open Data Commons Open Database License: Daten dürfen geteilt, weiterverarbeitet und verändert werden, solange der Urheber genannt wird. Daraus entstandene Werke müssen unter gleicher Lizenz veröffentlicht werden.

<https://opendatacommons.org/>



# Weiterführende Informationen

- [FAIR-Prinzipien](#)
- [Creative Commons Lizenzen](#)
- [Open Data Commons Lizenzen](#)
- [Open Science](#)
- [Open Science in Horizon2020](#)
- [Verzeichnis der \(Forschungsdaten-\)Repositorien](#)
- [Projekt „FORDATIS“](#)
- [Repository „Zenodo“](#)
- [Repository „RADAR“](#)
- [Datenjournaliste auf forschungsdaten.org](#)



# Weiterführende Informationen

- [B2Share](#)



---

# DATEN ARCHIVIEREN

---

## Inhalt

Anforderungen

Kriterien geeigneter Dateiformate

Empfohlene Dateiformate



# Daten archivieren – Anforderungen Leitlinie der DFG

„Den Regeln der Guten Wissenschaftlichen Praxis folgend sollen Forschungsdaten in der eigenen Einrichtung oder in einer fachlich einschlägigen, überregionalen Infrastruktur für mindestens 10 Jahre archiviert werden.“

[[https://www.dfg.de/foerderung/grundlagen\\_rahmenbedingungen/gwp/](https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/)]



# Daten archivieren – Kriterien geeigneter Dateiformate

	Maschinen Lesbarkeit	Von Menschen lesbar	Langzeit Stabilität	Beschreibende Daten
<b>sehr gut</b>	mit weit verbreiteter offener Software	ja und Lesbarkeit ohne Spezialsoftware	genormter Standard	vollständig enthalten
<b>gut</b>	mit gut verbreiteter und dokumentierter Software	nach Standard-Verfahren komprimiert	schon lang oder großflächig etabliert	technische Angaben sind enthalten

<http://www.forschungsdaten.info/>



# Daten archivieren – Kriterien geeigneter Dateiformate

	Maschinen Lesbarkeit	Von Menschen lesbar	Langzeit Stabilität	Metadaten
<b>mittelmäßig</b>	proprietäres Standard-format	mit offener Software (zuverlässig?) in höhere Klasse konvertierbar	relativ neues Format	einige wichtige (z. B. Einheiten) sind enthalten
<b>schlecht</b>	Selbst entwickelte Lese-Software	nein	Format, das gerade erst erfunden wurde	keine beschreibende Daten

<http://www.forschungsdaten.info/>



Datei-Typ	Empfohlene Datei-Formate für das Teilen, für die Wiederverwendung und Erhaltung
<p><b>Quantitative tabellarische Daten mit umfangreichen Metadaten</b></p> <p>Ein Dataset mit variablen und definierten Bezeichnungen, fehlende Werte zusätzlich zur Datenmatrix</p>	<ul style="list-style-type: none"> <li>• SPSS portable format (.por)</li> <li>• Begrenzte Text- und Befehlsdatei (Setup-Datei) (SPSS, Stata, SAS usw.) mit Metadateninformationen</li> <li>• Beliebiger strukturierter Text oder eine Markierungsdatei mit Metadateninformationen, z.B. DDI-XML-Datei</li> </ul>
<p><b>Quantitative tabellarische Daten mit minimalen Metadaten</b></p> <p>Eine Datenmatrix mit oder ohne Spaltenüberschriften oder Variablennamen, aber ohne andere Metadaten oder Beschriftungen</p>	<ul style="list-style-type: none"> <li>• CSV-Datei (CSV = Comma Separated Values)</li> <li>• tabulatorgetrennte Datei (.tab) Einschließlich des durch Trennzeichen getrennten Texts des angegebenen Zeichensatzes und gegebenenfalls der SQL-Datendefinitionsanweisungen</li> </ul>
<p><b>Geodaten</b></p> <p>Vektor- und Rasterdaten</p>	<ul style="list-style-type: none"> <li>• ESRI-Shapefile (wesentlich: .shp, .shx, .dbf; optional: .prj, .sbx, .sbn)</li> <li>• georeferenziertes TIFF (.tif, .tiff)</li> <li>• CAD Daten (.dwg)</li> <li>• tabellarische GIS-Attributdaten</li> </ul>

Datei-Typ	Empfohlene Datei-Formate für das Teilen, für die Wiederverwendung und Erhaltung
<b>Qualitative Daten</b>  Text	<ul style="list-style-type: none"> <li>• eXtensible Markup Language (XML) -Text gemäß einer geeigneten Dokumenttypdefinition (DTD) oder einem entsprechenden Schema (.xml)</li> <li>• Rich Text Format (.rtf)</li> <li>• Klartextdaten, ASCII (.txt)</li> </ul>
<b>Digitale Bilddaten</b>	<ul style="list-style-type: none"> <li>• TIFF Version 6 unkomprimiert (.tif)</li> </ul>
<b>Digitale Audiodaten</b>	<ul style="list-style-type: none"> <li>• Freier verlustfreier Audio-Codec (FLAC) (.flac)</li> </ul>
<b>Digitale Videodaten</b>	<ul style="list-style-type: none"> <li>• MPEG-4 (.mp4)</li> <li>• Motion JPEG 2000 (.jp2)</li> </ul>
<b>Dokumentation</b>	<ul style="list-style-type: none"> <li>• Rich Text Format (.rtf)</li> <li>• PDF/A oder PDF (.pdf)</li> <li>• OpenDocument-Text (.odt)</li> </ul>

# Weiterführende Informationen

- [Datenformate](#)

- [Dateiformate](#)



---

# DATEN NACHNUTZEN

---

## Inhalt

Wo finde ich Daten?

Was darf ich mit den Daten?

Daten zitieren



# Daten nachnutzen – Wo finde ich Daten?

## Forschungsdaten

- [Research Data Australia](#)
- [JISC Research Data Discovery](#)
- [Portage Network Canada](#)
- [CESSDA](#)
- [Repositorien](#)
- [Forschungsdatenzentren](#)

## Verwaltungsdaten

- [Open Data Portal München](#)
- [Open Data Portal Bundesregierung](#)
- [Open Data Portal der EU](#)



# Daten nachnutzen – Was darf ich mit den Daten machen?

- Zu beachten sind:
  - Nutzungsbestimmungen des Archivs/Repositoriums/Datenzentrums
  - Lizenz des Datensatzes
  - Schutz durch Patentrecht ist zu beachten
  - Datenschutzbestimmungen bei personenbezogenen Daten sind zu beachten
  - Entsprechend den Grundsätzen guter Wissenschaftlicher Praxis: Datenautoren nennen



# Daten nachnutzen – Daten zitieren

- Daten sind legitime, zitierbare Produkte wissenschaftlicher Tätigkeit. Das Zitieren von Forschungsdaten ist genauso unerlässlich wie das Zitieren von Publikationen oder anderen Ergebnissen wissenschaftlicher Arbeit.
- Daten sollten so zitiert werden, dass alle Personen, die zur Datenerstellung beigetragen haben, berücksichtigt werden.
- Die Zitierung von Daten sollte eine persistente Methode zur Identifikation beinhalten, die durch Maschinen ausführbar und global einzigartig ist sowie in großem Umfang von einer Community genutzt wird. → Angabe des Identifikators



# Daten nachnutzen – Daten zitieren – Bsp.: da|ra

da|ra Registrierungsagentur für Sozial- und Wirtschaftsdaten

Mein da|ra Für Datenzentren Für Forscher Für Verlage Über uns

## Datenzitation

Daten recherchieren  
**Daten zitieren**  
Daten archivieren

Die Nutzung und Analyse von Forschungsdaten sowie die Veröffentlichung von Forschungsergebnissen erfordern Informationen zu Urheber, Standort und Identifikation der Daten.

Eine richtige und vollständige Zitierung einer Ressource kann auf der Basis der da|ra Metadaten mit den folgenden Elementen erstellt werden:

**Creator (Veröffentlichungsdatum): Titel. Version. Publikationsagent. Genereller Ressourcentyp. Identifikator**

Der Identifikator (Identifikator) steht für den Digital Object Identifier (DOI). Er wird in der Zitierung als URL-Link inklusive dem notwendigen Resolver sowie ohne „doi“ angezeigt.

Einige Beispiele für die Zitierung:

GESIS Panel Team (2014): GESIS Panel - Standard Edition. Version: 2.0.0. GESIS Datenarchiv Dataset.  
Ⓞ <http://doi.org/10.4232/1.11947>

Markowski, Radoslaw; Gebethner, Stanislaw; Grabowska, Mirosława; Grzelak, Pawel; Jasiewicz, Kysztof et. al. (2006): Polish National Election Study 2000 (PGSW). Version: 1.0.0. GESIS Data Archive: Dataset.  
Ⓞ <http://doi.org/10.4232/1.4334>

Autoren (Veröffentlichungsdatum): Titel. Version.  
Publikationsagent. Genereller Ressourcentyp. Identifikator

<https://www.da-ra.de/de/fuer-forscher/daten-zitieren/>



# Daten nachnutzen – Daten zitieren – Data Citation Formatter

## DOI Citation Formatter

Paste your DOI:

For example 10.1145/2783446.2783605

Select Formatting Style:

Begin typing (e.g. Chicago or IEEE.) or use the drop down menu.

Select Language and Country:

Begin typing (e.g. en-GB for English, Great Britain) or use the drop down menu.

Format

Do you want to integrate this service? Check the [Documentation](#)

<https://crosscite.org/>



# Weiterführende Informationen

- Quellen für Forschungsdaten
  - [Research Data Australia](#)
  - [JISC Research Data Discovery](#)
  - [Portage Network Canada](#)
  - [CESSDA](#)
- Quellen für Verwaltungsdaten
  - [Open Data Portal München](#)
  - [Open Portal Bundesregierung](#)
  - [Open Data Portal der EU](#)
- [FORCE11-Prinzipien zur Datenzitation](#)
- [Datenzitation nach da | ra](#)



---

# FORSCHUNGSDATENDATEN UND POLITISCHE ENTWICKLUNGEN

---

## Inhalt

Nationale Forschungsdateninfrastruktur NFDI

European Open Science Cloud



# Nationale Forschungsdateninfrastruktur NFDI



<http://www.rfii.de/?p=1998>

▼ Strukturelemente der NFDI

▼ NFDI-Expertengremium

▼ Eckpunkte zur Antragstellung

▼ Ablauf des Verfahrens

▼ NFDI-Konferenzen

▼ Zeitplan

▼ Ansprechpersonen

NFDI4Ing

NFDI4Chem

NFDI4Life

NFDI4...

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

# European Open Science Cloud

My Services Profile Contact Us Login

 EUROPEAN OPEN SCIENCE CLOUD

About Governance **Services & Resources** Policy 

EOSC in Practice Med **Sharing & Discovery**

Processing & Analysis

Data Management

Compute

Storage

Networking

Training & Support

Security & Operations

Home » Services & Resources

## Data Management

**Filter by**

- Any -

**The National Open Access Research**

<https://www.eosc-portal.eu/services-resources/data-management>



# Weiterführende Informationen

- [Wissenschaftliche Software in der Initiative „Digitale Information“ der Allianz der Wissenschaftsorganisationen](#)
- [Codecademy](#)
- [Free Software Foundation](#)
- [GitHub](#)
- [GitLab](#)
- [Jupyter](#)



# Wissenschaftliches Publizieren bei Fraunhofer - Informationsveranstaltung

- Sichtbarkeit von Publikationen, wissenschaftliches Identitätsmanagement mit ORCID
- Bibliometrische Grundlagen, Fraunhofer-Wissenschaftsindikatorik
- Open Access, Open Data Open Science, Projektanforderungen bei EU, BMBF,...
- Gute wissenschaftliche Praxis, Rechtliche Fragen, Verlagsverträge
- Verlage einschätzen, Predatory Publishing
- Patentieren und Publizieren
- Weitere Themen auf Anfrage...



Mehr unter:

[\[Intranet Wissenschaftliches Publizieren\]](#)

[\[Publikationsberatung im Wiki\]](#)

[\[Fraunhofer Bildungsprogramm 2018\]](#)

# Diskussion

Wie können Bibliotheken Aufgaben im Forschungsdatenmanagement übernehmen?

Welche Bedenken gibt es?

Welche Hindernisse gibt es?

Welche Aufgaben sind geeignet?

Welcher Unterstützungsbedarf besteht



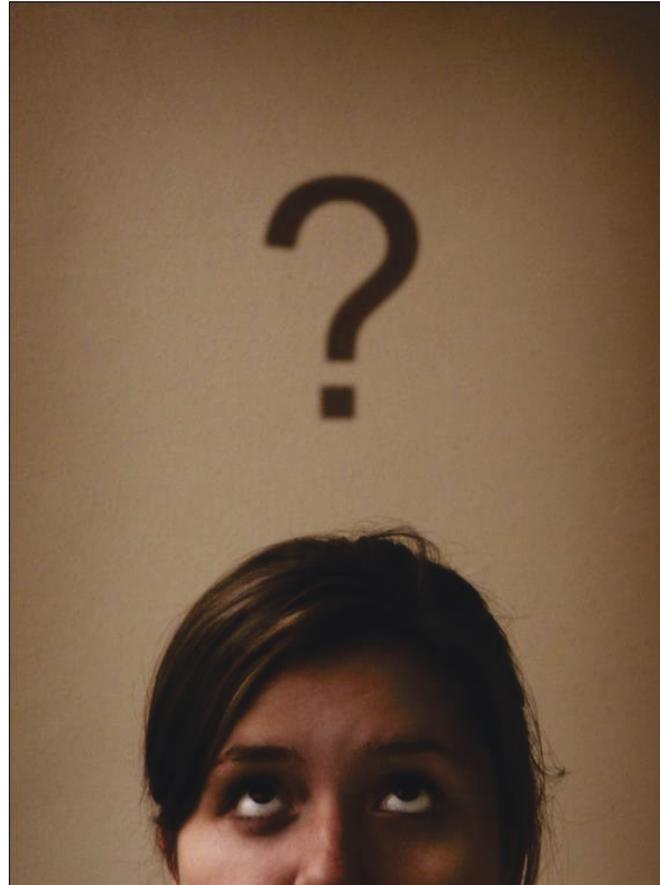
# Veranstaltungshinweis

Veranstaltung „Zukunft  
Wissenschaftlicher Bibliotheken?“ der  
ZBMED im Mai

→ [https://www.youtube.com/channel/UCKJnC4S0ZTU9q5pmwf\\_TSbg/%20nach  
zuerleben.](https://www.youtube.com/channel/UCKJnC4S0ZTU9q5pmwf_TSbg/%20nach%20zuerleben)



# Fragen?



[CC BY-NC-ND 2.0, Scabeater via flickr](#)



# Geschafft!



[CC BY-NC-SA 2.0, Ralf Appelt via flickr](#)

# Kontakt



Andrea Wuchner

[Andrea.Wuchner@irb.fraunhofer.de](mailto:Andrea.Wuchner@irb.fraunhofer.de)

Tel: +49 (0)711 / 970-2714

## **Competence Center Research Services & Open Science**

Fraunhofer-Informationszentrum Raum und Bau (IRB)

Publikationsmanagement

Nobelstraße 12

D-70569 Stuttgart

