

Woche der Forschungskompetenzen 2025

Forschungsdaten nutzen und verwalten

07.10.2025

Johanna Vompras

Kompetenzzentrum Forschungsdaten
data@uni-bielefeld.de



Bildquelle: Jørgen Stamp,
<https://digitalbevaring.dk/>

Programm

- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden
- Daten nachnutzen
- Daten organisieren, dokumentieren und sichern
- Abschluss

Programm

- **Verständnis von Forschungsdaten**
 - Empfehlungen und Anforderungen
 - Daten finden
 - Daten nachnutzen
 - Daten organisieren, dokumentieren und sichern
 - Abschluss

Forschungsdaten

DFG (Deutsche Forschungsgemeinschaft):

„Zu Forschungsdaten zählen u.a. **Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten, Objekte aus Sammlungen oder Proben**, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden.“

Methodische Testverfahren, wie Fragebögen, Software und Simulationen **können ebenfalls** zentrale Ergebnisse wissenschaftlicher Forschung darstellen und sollten daher ebenfalls unter den Begriff Forschungsdaten gefasst werden.“ [1]

Weitere allgemeine Definition (University of Leeds Library):

“Research data is any information that has been collected, observed, generated or created to validate original research findings.” [2]

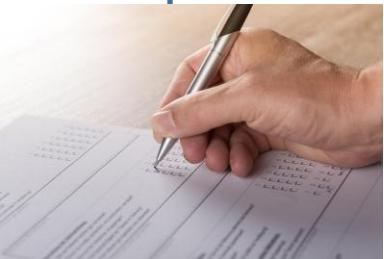


[1] www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

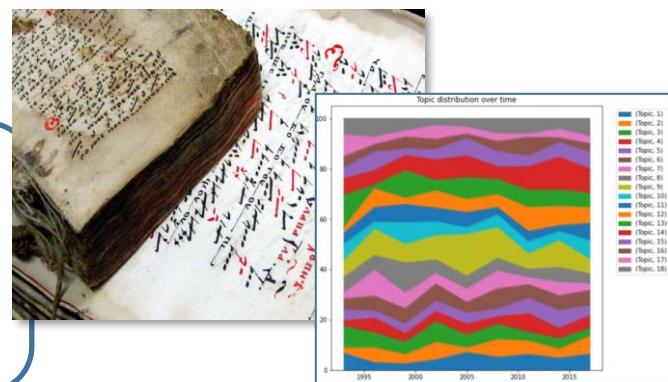
[2] <https://library.leeds.ac.uk/info/14062/research-data-management/61/research-data-management-explained>

Quellen von Forschungsdaten

Beobachtungsdaten,
z. B. Umfragen oder
Verhalten



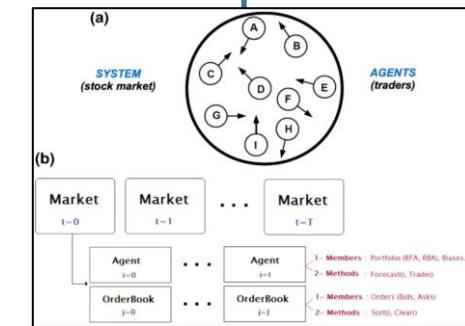
Abgeleitete oder
zusammengestellte
Daten, z. B. Quellen oder
Textmining



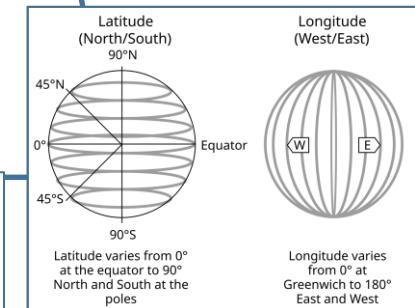
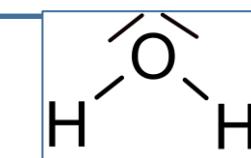
Experimentelle Daten,
z. B. Labor



Simulationsdaten,
z. B. ökonomische
Modelle

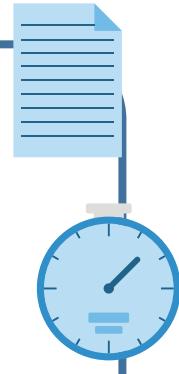


Referenz- oder
kanonische Daten, z. B.
chemische Strukturen



Arten von Forschungsdaten

- Dokumente, Tabellenkalkulationen
- Labornotizbücher, Feldnotizbücher, Tagebücher
- Fragebögen, Protokolle, Codebücher



- Daten-Files
- Inhalt einer Datenbank



- Modelle, Algorithmen, Skripte
- Inhalt einer Anwendung



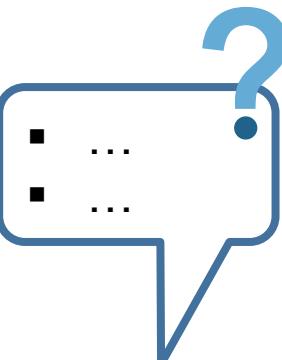
- Audio-/Videodateien/-bänder
- Fotografien, Filme



- Antworten auf Tests
- Dias, Artefakte, Proben, Muster
- Sammlungen von digitalen Ergebnissen



- Methodologien, Arbeitsabläufe
- Standardarbeitsanweisungen und -protokolle



Datenformate (Beispiel: Soziologie)

Verschiedene Quellen/Typen von Daten		Datenformate
Quantitative Daten	Daten aus Erhebungen, Volkszählungen, ... (numerische Daten)	Tabellenkalkulationsformate, z. B. CSV, Excel (.xlsx)
	Erhebungsinstrumente und Fragebögen (textuelle Daten)	Textformate, z. B. PDF, Word (.docx), offene Textformate (.odt) oder einfache Textdateien (.txt)
	Instrumente zur Datenanalyse (statistische Daten)	RData, SPSS, Stata, ... oder allgemein gebräuchliche Formate: CSV, Excel (.xlsx)
Qualitative Daten	Daten, die durch Interviews, Fokusgruppen, ethnografische Studien gesammelt wurden, z. B. Transkripte, Feldnotizen (textuelle Daten)	Textformate, z. B. PDF, Word (.docx), offene Textformate (.odt) oder einfache Textdateien (.txt)
	Aufzeichnungen der oben genannten Daten (Audio-/Videodata)	MP3, MP4, WAV, AVI, ...
	Fotografien oder Bilder (visuelle Daten)	JPEG, PNG, TIFF
Instrumente zur Datenanalyse (qualitative Daten)		MAXQDA (.mx18, .mx19), Atlas.ti (*.atlproj) oder allgemein übliche Formate: CSV, TXT, ...

Datenformate (Beispiel: Chemodiversität von Pflanzen)

Verschiedene Quellen/Typen von Daten	Datenformate
Beschreibende Informationen/ Dokumentation/ Beobachtungen (textuelle Daten)	XML, CSV, TXT, Markdown/txt (Wiki), programmspezifisch (Word-Dateien, ...)
Strukturierte Daten in Tabellenkalkulationen oder Datenbanken (tabellarische Daten)	Datenmatrizen (in XLSX, CSV, SQL-Datenbanken)
Fotografische oder mikroskopische Bilder (Bild daten)	JPEG, PNG, TIFF
Sequenzen und Annotationen (genomische Daten)	FASTA, FASTQ (rohe Sequenzlesungen) und GFF, BED (genomische Annotationen)
Rohe/verarbeitete Daten z. B. aus der Massenspektrometrie (Spektral daten)	mzML, mzXML, oder proprietäre Formate
Daten mit geografischen Komponenten (Geodaten)	Shapefiles (SHP), GeoJSON, KML
Instrumente zur Datenanalyse (statistische Daten)	RData, SPSS, MATLAB, Stata, ... oder allgemein gebräuchliche Formate: CSV, Excel

Komplexität von Forschungsdaten

- **die empirische Basis** für wissenschaftliche Analysen weitet sich kontinuierlich aus
- stetige Zuwachs an verfügbaren Daten, durch **methodische und technische Neuerungen** (z.B. Social Media, stetig wachsende Datenquellen)
- **Komplexität** der Daten deutlich **erhöht**, durch
 - Verknüpfung von Daten anderer Disziplinen
 - unendliche Vielzahl an Weiternutzungsmöglichkeiten

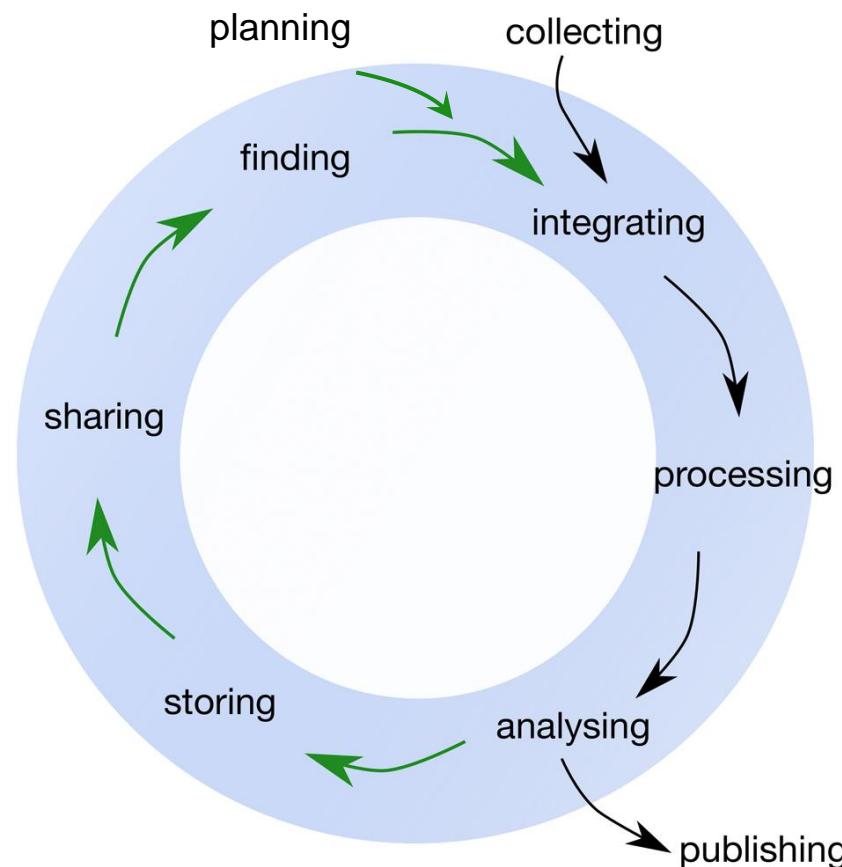


Bildquelle: KI-generiert: Wissenschaftler/in untersucht externe Datenquellen

Programm

- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden
- Daten nachnutzen
- Daten organisieren, dokumentieren und sichern
- Abschluss

Aspekte rund um den Data Life Cycle



Data Life Cycle

Vereinfachte **Repräsentation** eines jeden Forschungsprojektes

Forschungsdatenmanagement (FDM):
die **gesamte Organisation** und
Verwaltung der Prozesse zusammen,
die in Zusammenhang mit der
Entstehung, Verarbeitung,
Dokumentation, Weitergabe und der
nachhaltigen Archivierung von
Forschungsdaten stehen.

Was erwartet/bietet die

- Checkliste „Umgang mit Forschung“
<https://www.dfg.de/resource/blob/174732/for>
 - Fachspezifische Empfehlungen: <https://der-foerderung/forschungsdaten/empfehlung>
 - In vielen Förderprogrammen sind Angaben zu den Daten erforderlich (Aspekte: z. B. Umfang der Daten, Datensicherheit, Rechte und Pflichten bei späteren Wiederverwendung und langfristiger Speicherung)
 - ***Opt Out*** bezüglich der Datenweitergabe ist nach § 24 Abs. 2 DSGVO vorgesehen
→ **Erklärung**
 - **Aspekte, die möglicherweise berücksichtigt werden sollten:**
 - **Opt Out** bezüglich der Datenweitergabe ist nach § 24 Abs. 2 DSGVO vorgesehen
→ **Erklärung**

Fachspezifis	Checkliste zum Umgang mit Forschungsdaten
Im Folgenden finden	1. Datenbeschreibung Auf welche Weise entstehen in Ihrem Projekt neue Daten? Werden existierende Daten wiederverwendet? Welche Datentypen, im Sinne von Datenformaten (z. B. Bilddaten, Textdaten oder Messdaten) entstehen in Ihrem Projekt und auf welche Weise werden sie weiterverarbeitet? In welchem Umfang fallen diese an bzw. welches Datenvolumen ist zu erwarten?
Fachkollegien der DF	
Geistes- und So	2. Dokumentation und Datenqualität Welche Ansätze werden verfolgt, um die Daten nachvollziehbar zu beschreiben (z. B. Nutzung vorhandener Metadaten- bzw. Dokumentationsstandards oder Ontologien)?
	Welche Maßnahmen werden getroffen, um eine hohe Qualität der Daten zu gewährleisten? Sind Qualitätskontrollen vorgesehen und wenn ja, auf welche Weise?
	Welche digitalen Methoden und Werkzeuge (z. B. Software) sind zur Nutzung der Daten erforderlich?
	3. Speicherung und technische Sicherung während des Projektverlaufs Auf welche Weise werden die Daten während der Projektaufzeit gespeichert und gesichert? Wie wird die Sicherheit sensibler Daten während der Projektaufzeit gewährleistet (Zugriffs- und Nutzungsverwaltung)?
	4. Rechtliche Verpflichtungen und Rahmenbedingungen Welche rechtlichen Besonderheiten bestehen im Zusammenhang mit dem Umgang mit Forschungsdaten in Ihrem Projekt? Sind Auswirkungen oder Einschränkungen in Bezug auf die spätere Veröffentlichung bzw. Zugänglichkeit zu erwarten? Auf welche Weise werden nutzungs- und urheberrechtliche Aspekte sowie Eigentumsfragen berücksichtigt? Existieren wichtige wissenschaftliche Kodizes bzw. fachliche Normen, die Berücksichtigung finden sollten?
	5. Datenaustausch und dauerhafte Zugänglichkeit der Daten Welche Daten bieten sich für die Nachnutzung in anderen Kontexten besonders an? Nach welchen Kriterien werden Forschungsdaten ausgewählt, um diese für die Nachnutzung durch andere zur Verfügung zu stellen? Planen Sie die Archivierung Ihrer Daten in einer geeigneten Infrastruktur? Falls ja, wie und wo? Gibt es Sperrfristen? Wann sind die Forschungsdaten für Dritte nutzbar?
Lebenswissensch	6. Verantwortlichkeiten und Ressourcen Wer ist verantwortlich für den adäquaten Umgang mit den Forschungsdaten (Beschreibung der Rollen und Verantwortlichkeiten innerhalb des Projekts)? Welche Ressourcen (Kosten; Zeit oder anderes) sind erforderlich, um einen adäquaten Umgang mit Forschungsdaten im Projekt umzusetzen? Wer ist nach Ende der Laufzeit des Projekts für das Kuratieren der Daten verantwortlich?
Naturwissensch	
	• Fachkollegium "Physik" zum Umgang mit Forschungsdaten im Fachbereich Physik (2023) ↗
	• Fachkollegium "Biodiversitätsforschung" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegium "Pflanzbau" ↗
	• Fachkollegien "Physik" zum Umgang mit Forschungsdaten im Fachbereich Physik (2023) ↗

Forschungsdaten Policy der Uni Bielefeld

<https://www.uni-bielefeld.de/ub/digital/forschungsdaten/policy/>

2011: Grundsätze zum Umgang mit Forschungsdaten

Forschende .. [] .. sollen
Forschungsdaten als
**wertvolles
wissenschaftliches Werk**
behandeln ...

2013: Resolution Forschungsdaten

Die Universität Bielefeld
unterstützt [Forschende], die
Grundsätze .. [] .. im Sinne
**eines verbindlichen und
qualitätsbewussten**
Forschungsdatenmana-
gements **umzusetzen**.

2021:Sicherung guter wiss. Praxis

§12: Herstellung von
öffentliche **Zugang** zu
Forschungsergebnissen
§13: Autorenschaft: .. [] ..
wissenschaftlichen Text-,
Daten- oder
Softwarerepublikation ...“

§16: Archivierung von
Forschungsdaten .. [] ..

Universität Bielefeld

FAIR Data

Subject-specific Recommendations on the Handling of Research Data

Here you will find current subject-specific statements developed in cooperation with DFG committees.

Humanities and social sciences

- "Psychology" Review Board on the handling of research data (2020) ↗
- "Educational Research" Review Board on the handling of research data (2020) (in German only) ↗
- "Social Sciences" Review Board on the handling of research data in sociology, political science and communication science (2020) (in German only) ↗
- "Ancient Cultures" Review Board on the handling of research data (2020) (in German only) ↗
- "Social and Cultural Anthropology, Non-European Cultures, Jewish Studies and Religious Studies" Review Board on the handling of research data (2019) (in German only) ↗
- "Economics" Review Board on the handling of research data (2025) ↗
- "Linguistics" Review Board on data standards and tools in the collection of language corpora (2019) (in German only) ↗
- "Literary Studies" Review Board on funding criteria for scholarly editions in literary studies (2015) (in German only) ↗

Life sciences

- "Plant Sciences" and "Zoology" Review Boards on the handling of research data in biodiversity research (2015) (in German only) ↗



Stellungnahme des Fachkollegiums 111 „Sozialwissenschaften“

zum Forschungsdatenmanagement in der Soziologie, der Politikwissenschaft und der Kommunikationswissenschaft

(1) Die Archivierung von Forschungsdaten ist, sofern sie nicht InformantInnen oder Forschungsbeziehungen beschädigt, sinnvoll und notwendig. Archivierung oder Nicht-Archivierung sollten in Forschungsanträgen begründet werden. Im Falle der Archivierung können auch Finanzmittel in angemessenem Umfang beantragt und bereitgestellt werden.

(2) Wissenschaftlich nutzbare Daten sind nützlich für Offenheit und Transparenz vieler Studien. Insbesondere sind sie eine große Unterstützung, um empirische Ergebnisse intersubjektiv nachvollziehbar zu machen (etwa durch Reproduktion der Analysen).

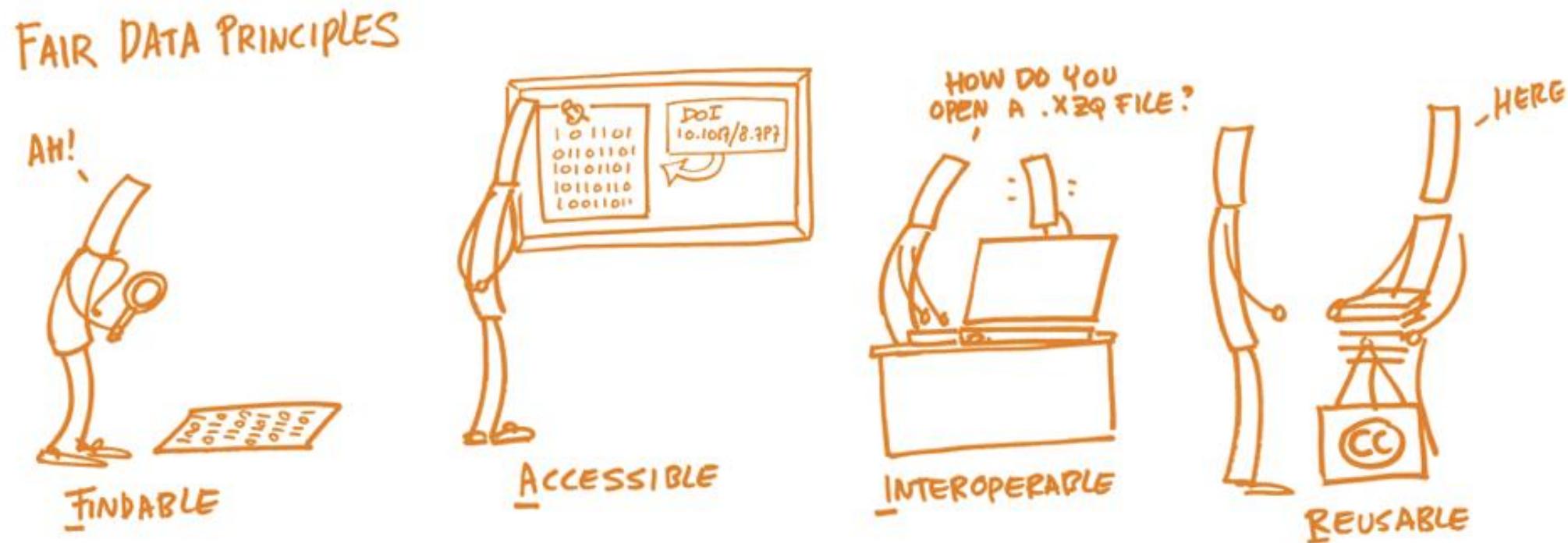
(3) Wenn Forschungsdaten archiviert werden, sind Qualitätsstandards einzuhalten, wie sie in den FAIR-Prinzipien festgehalten sind (findable, accessible, interoperable, reusable).

(4) Nicht immer werden archivierte Daten an andere Forschende weitergegeben werden können. Dies hat keinen Einfluss auf die Frage der Förderwürdigkeit eines Forschungsvorhabens.

„FAIR“ als **Leitprinzip** für das Forschungsdatenmanagement

FAIR Data Principles

Wie sollte man optimalerweise Daten behandeln?
Wie könnte man FAIR Data umsetzen?



Bildquelle: <https://book.fosteropenscience.eu/>

FINDABLE



- Nutzung einer **persistenten Kennung** (z.B. DOI) Beispiel:
10.1038/s41586-020-2649-2



- Prüfung der **Anforderungen zur Auffindbarkeit** der wissenschaftlichen Verlage (z.B. in der „**Data Availability Policy**“)
- Erhöhung der Sichtbarkeit der Daten **durch Verlinkung** (z.B. Daten in eigenen Publikationen oder grauer Literatur, wie „working papers“ oder „technical reports“ referenzieren)
- **Zitation** der eigenen Daten:

Source:
<https://www.fosteropenscience.eu/learning/assessing-the-fairness-of-data/>

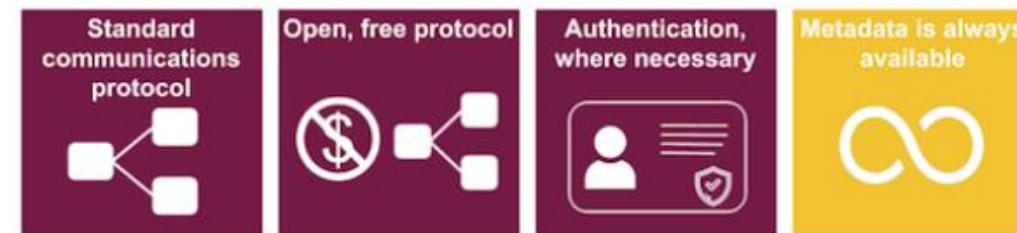
Romer, Christina D., and David H. Romer. 2010. "Replication data for: The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E112357V1>.

ACCESSIBLE



Source:
<https://www.fosteropenscience.eu/learning/assessing-the-fairness-of-data/>

- Ergänzen Sie Ihre Datenpakete um **formale** und **technische** Metadaten!
- **Offene Protokolle** nutzen: Repozitorium mit **HTTPS/FTP** wählen.



- Anbieter mit „**Pay Walls**“ **vermeiden** (zwecks Kostenfreiheit und Barrierefreiheit)
- Zur Einschränkung der **Zielgruppe** oder **Nutzungszwecke**: Repozitorien mit Authentifizierung und Autorisierung nutzen.

Beispiele:

Repositorien: EBI NCBI, GenBank, RePEc, GESIS, usw.

Versionierung/Entwicklung: GitLab, GitHub

INTEROPERABLE

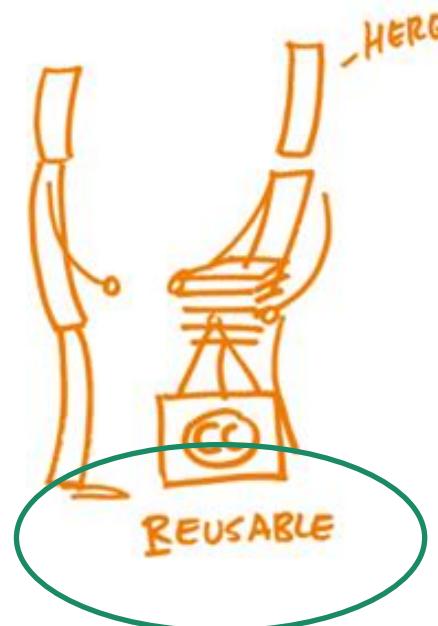


Source:
<https://www.fosteropenscience.eu/learning/assessing-the-fairness-of-data/>

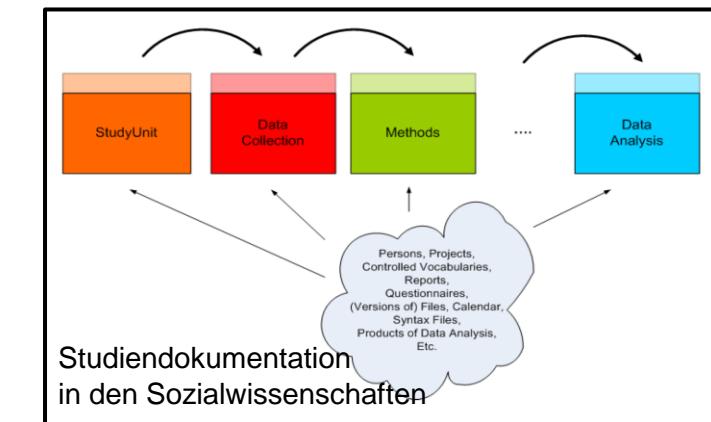
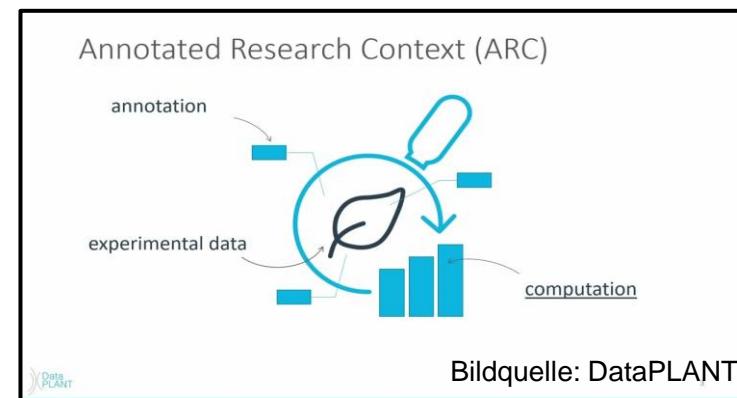
- ... Bedingung für **Datentausch** und **Nutzung durch Tools / Dritte**
- ... möglichst **offene Formate** und Datenformate der **Community** (wenn ja, welche?), wenn **neu definierte Formate**: Definition der Dokumentation beifügen ...
- Dokumentation aller Datenpakete (z.B. Variablen, Ressourcen, Parameter für Berechnungen, etc.) in einer standardisierten Art und Weise:

```
<subjects>
  <subject schemeURI="http://www.nlm.nih.gov/mesh/meshhome.html"
           subjectScheme="Mesh">Neoplasms</subject>
</subjects>
```
- ... auch hier Angabe von **Metadaten** (für die Interoperabilität), z.B. Softwareversion zum Öffnen, Definition zu Schnittstellen (APIs), usw.
- ... Daten in **verschiedenen Formaten anbieten** (um Interoperabilität zu erhöhen)

REUSABLE



- ... sind gut **dokumentiert**,
- ... es gibt einen Plan für die **Veröffentlichung von Forschungsdaten?** (Wann?, Embargo? Patente? rechtliche Aspekte?)
- **Nachnutzbarkeit/Nachvollziehbarkeit:** Alle Quellen, Daten, Dokumente sollten verfügbar gemacht werden und die **(Zwischen-)Schritte** der Datengenerierung und Verarbeitung gut dokumentiert werden.



- ... haben klare Lizenz- und Herkunftsangaben (bestenfalls für die **breite Nachnutzung**)
- ... der gesamte „Forschungsoutput“ ist gut untereinander verlinkt (z.B. Daten, Code, Working Papers, usw.)

Source:
<https://www.fosteropenscience.eu/learning/assessing-the-fairness-of-data/>

FAIR – Zusammenfassung

Generelles Ziel: FAIR soll die **effiziente Datennutzung** in der wissenschaftlichen Gemeinschaft fördern, indem es:

- die **Reproduzierbarkeit** stärkt,
- **Forschungserkenntnisse** beschleunigt,
- die **Transparenz** erhöht,
- und **Zusammenarbeit** erleichtert.



Es wird ein **Ökosystem** geschaffen, in dem Daten nutzbarer sind, und somit

- **bessere Forschungsergebnisse**,
- **schnellere wissenschaftliche Fortschritte**,
- **mehr Vertrauen der Öffentlichkeit in Forschung** ermöglicht werden sollen.

Programm

- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden**
- Daten nachnutzen
- Daten organisieren, dokumentieren und sichern
- Abschluss

Daten finden

Datenquellen

Abhängig vom **Forschungsgebiet**
möglicherweise unterschiedliche Arten

Einige Beispiele:

1. Klimadaten
2. Model databases, e.g. CEDA
3. European fact data (e.g. Flash Eurobarometer 315:
Attitudes of European entrepreneurs towards eco-innovation)

<https://re3data.org>:

Registry of **certified** and **reliable** research data repositories

The screenshot shows the re3data.org website interface. At the top, there's a header with the university logo and the text "Flash Eurobarometer 315: Attitudes of European entrepreneurs towards eco-innovation". Below this, it says "Publisher: Directorate-General for Communication" and "Updated: [date]". The main content area displays a list of data volumes: Volume A: Countries; Volume AA: Groups of countries; Volume A' (AP): Trends; Volume AA' (AAP): Trends of groups of countries; Volume B: EU/socio-demographics; Volume B' (BP): Trends of EU/ socio-demographics; and Volume C: Country/socio-demographics. It also mentions that researchers can contact GESIS - Leibniz Institute for the Social Sciences. On the right side, there's a sidebar with categories like Agriculture, Fisheries, Forestry & Foods; Health; Technology; and Transport. At the bottom, there's a footer with the re3data.org logo and search bar.

Daten finden

<https://re3data.org>:

Verzeichnis zertifizierter und vertrauenswürdiger Forschungsdaten-Repositorien

The screenshot shows the re3data.org website interface. The top navigation bar includes links for Search, Browse, Suggest, Resources, and Contact. A search bar contains the term "economics". On the left, there is a sidebar with a "Filter" section and a "Subjects" section listing various academic fields like Humanities and Social Sciences, History, and Linguistics. The main content area displays search results for "economics", showing 30 results found. One result is highlighted: "The National Archives Discovery". This entry includes fields for Subject(s) (Humanities and Social Sciences, Humanities, History, Social and Behavioural Sciences, Social Sciences, Political Science, Economics, Economic and Social History, Jurisprudence, Principles of Law and Jurisprudence), Repository type(s) (institutional), Provider type(s) (dataProvider), and Country (United Kingdom). A detailed description of The National Archives is provided at the bottom.

re3data.org

Search

Browse ▾

Suggest

Resources ▾

Contact

economics

Search

← Previous 1 2 Next →

Found 30 result(s)

The National Archives

The National Archives Discovery

Subject(s)

Humanities and Social Sciences, Humanities, History, Social and Behavioural Sciences, Social Sciences, Political Science, Economics, Economic and Social History, Jurisprudence, Principles of Law and Jurisprudence

Repository type(s)

institutional

Provider type(s)

dataProvider

Country

United Kingdom

The National Archives is home to millions of historical documents, known as records, which were created and collected by UK central government departments and major courts of law. Data of the former National Digital Archive of Datasets (NDAD) collection, which was active from 1997 to 2010 and preserves and provides online access to archived digital datasets and documents from UK central government departments, is integrated. Access to records held by The National Archives and more than 2,500 other archives.

Daten finden – Beispiele

Fachspezifische Portale für offene Forschung

- [KonsortSWD](#) – Suche nach Sozial- und Wirtschaftsforschungsdatenzentren
- [The Economics Network](#) – Quellen für frei verfügbare Daten
- [GenBank](#) – Die NIH-Datenbank für genetische Sequenzen
- Sammlung aller öffentlich verfügbaren DNA-Sequenzen
- [GESIS](#): Leibniz-Institut für Sozialwissenschaften
- [IEDB](#): Immune Epitope Database & Tools

Portale mit “Open Government Data”

- [Australia](#), [Canada](#), [EU](#), [Japan](#), [USA](#), [Netherlands](#)
- Daten des Statistischen Bundesamtes [Destatis](#)
- Daten der Europäischen Kommission [Eurostat](#)
- [Open Government Data Germany](#) – offizielle Staatliche Datenplattform
- [Open Government Data EU](#)

Kommerzielle Faktendatenbanken (lizenziert)

- [Refinitiv](#) (Finanz- und Marktdaten)
- [OECD iLibrary](#) (Publikationen und Datensätze der OECD)
- Embase (medizinische Forschungsdatenbank, Elsevier)

→ Fragen Sie Ihre Fachreferentin/Ihren Fachreferenten an Ihrer Einrichtung nach Verfügbarkeit ...

The screenshot shows a web interface for the German Federal Institute for Statistics (Destatis). At the top, there's a navigation bar with links for 'Login', 'German', 'Contact', 'FAQ', 'Watchlist (0)', 'Services', 'Research', and 'Institute'. Below the navigation is a search bar with placeholder text 'What are you looking for?'. The main content area displays a table titled 'EU-SILC - Persons in primary residence households: Germany, years, household type, poverty or social exclusion'. The table includes columns for Year, Household type, Total population, Persons at risk of poverty or exclusion, Persons at risk of poverty, Severely materially and socially deprived persons, and Persons in primary residence households. The data is presented for the year 2024, showing various household types like 'Households without children', 'Persons living alone', etc., with their respective counts.

Daten finden – Bibliotheksbasierter Zugang

Database Information Systems (DBIS) – Ein Zugang zu wissenschaftlichen Datenbanken

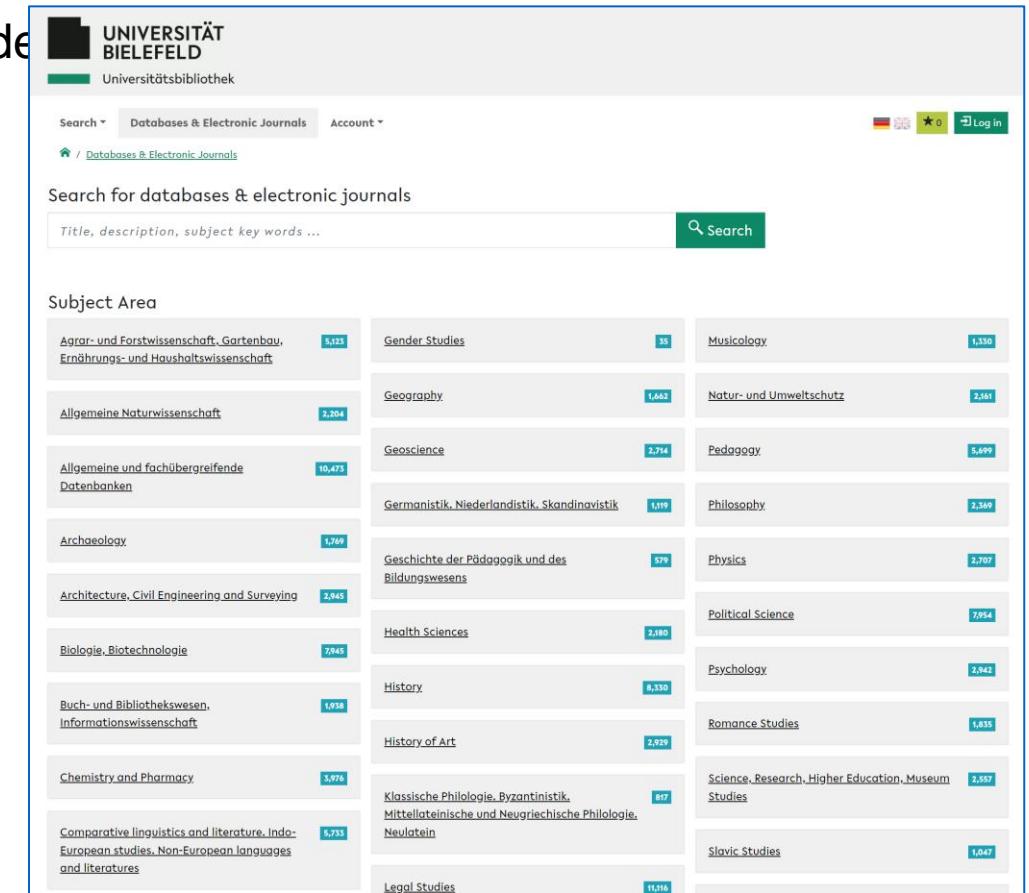
Die **DBIS** (bereitgestellt von Bibliotheken) sind ein hervorragender Zugang zu wissenschaftlichen Datenbanken:

Suche nach Fachgebiet oder gezielt nach Datenbanken:

<https://ub-bielefeld.digibib.net/eres/subject-area/>

Beispiele der dort gelisteten Informationsressourcen:

- **Statista:** Statistiken zu Wirtschaft, Gesellschaft & Medien.
<https://www.statista.com/>
- **OECD iLibrary:** Eine der größten Online-Bibliotheken für Daten und Analysen zu Wirtschaft, Gesellschaft und Umwelt.
<https://www.oecd-ilibrary.org/>
- **ViFa Recht (Virtual Library Law):** Ein Instrument für die rechtliche Online-Recherche und Zugang zu wissenschaftlichen Informationen zu Recht und Rechtsgeschichte.
<https://vifa-recht.de/>



The screenshot shows the DBIS search interface. At the top, there's a header with the University of Bielefeld logo and a search bar. Below the header, a section titled "Subject Area" displays a grid of 24 categories, each with a small thumbnail, the category name, and a numerical count. The categories include: Agrar- und Forstwissenschaft, Gartenbau, Ernährungs- und Haushaltswissenschaft (5,123); Gender Studies (35); Musicology (1,530); Geography (1,652); Natur- und Umweltschutz (2,161); Allgemeine Naturwissenschaft (2,204); Geoscience (2,714); Pedagogy (3,699); Archaeology (1,769); Germanistik, Nederlandistik, Skandinavistik (1,119); Philosophy (2,349); Allgemeine und fachübergreifende Datenbanken (10,473); History (8,350); Physics (2,707); Architecture, Civil Engineering and Surveying (2,645); Health Sciences (2,180); Political Science (7,954); Biologie, Biotechnologie (7,845); History of Art (2,029); Psychology (2,942); Buch- und Bibliothekswesen, Informationswissenschaft (1,978); Chemistry and Pharmacy (3,676); Romance Studies (1,835); Comparative linguistics and literature, Indo-European studies, Non-European languages and literatures (5,751); Klassische Philologie, Byzantinistik, Mittelalterliche und Neugriechische Philologie, Neulatein (817); Science, Research, Higher Education, Museum Studies (2,557); Slavic Studies (1,047); Legal Studies (11,116).

Daten finden – Allgemeine Suchmöglichkeiten

(Disziplin-)übergreifende Portale und Metasuchmaschinen

- [DataCite search portal](#) ist eine internationale Suchmaschine für Forschungsergebnisse, mit der nach registrierten Datensätzen gesucht werden kann.
- [European Union Open Data Portal](#) – Open-Data-Portal der Europäischen Union
- [Research Data Australia](#) – Australiens Open Research Data Portal
- [Google Dataset Search](#) – (Kommerzielle) Suchmöglichkeit für Forschungsdaten
- [gesisDataSearch](#) – Suche nach Daten zu Sozial- und Wirtschaftsforschung in Datenrepositorien und Metadaten-Diensten.
- [OpenAIRE](#) – eine internationale Suchmaschine für Forschungsergebnisse
- [Zanran](#) – Suchmaschine für Daten, mit der man nach Grafiken und Tabellen in Dokumenten suchen kann
- [Mendeley Data](#) – (kommerziell & multidisziplinär, Elsevier): Suche in einer Sammlung von Datenrepositorien
- [Data Portals](#) ist eine umfassende [Liste](#) von Open-Data-Portalen aus aller Welt.
- [Open Data Impact Map](#) eine öffentliche Datenbank von Organisationen weltweit, die Open Government Data nutzen



Daten finden – In Zeitschriften

Beispiel aus der Mikrobiologie

- Verweis über eine lokale ID (Zugangsnummer) eines Repositoriums/Datenbank
 - Datenverfügbarkeit im Artikel prüfen, z.B. im Unterkapitel zu „**data availability**“

Data availability. The genome assembly of *Paenibacillus marchantiae* has been deposited under the NCBI GenBank accession number [CP118270.1](#). The SRA, BioProject, and BioSample accession numbers are [SRP421713](#), [PRJNA933122](#), and [SAMN33224160](#), respectively.

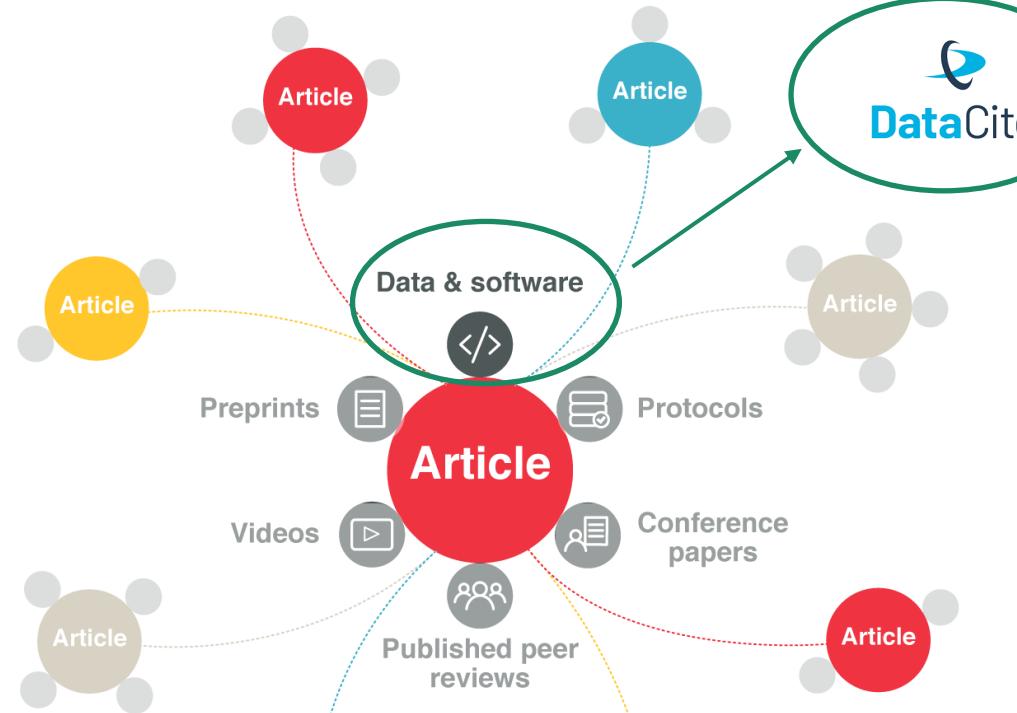
DN) were used for polishing, and Bacrora (v0.2.3; <https://github.com/CheckM/bacrora>) was run for overlap trimming. The assembly resulted in one circular contig (6,983,959 bp; GC content of 46.02%) (Table 1). Benchmarking Universal Single-Copy Orthologs (BUSCO) (v5.4.3; database “*bacillales_odb10*” (17) and CheckM (v1.2.2) (18) were applied to check assembly quality resulting in 0.097% contamination and 99.84% assembly completeness. A total of 6,195 protein-coding genes

July 2023 Volume 12 Issue 7

10.1128/mra.00354-23 1

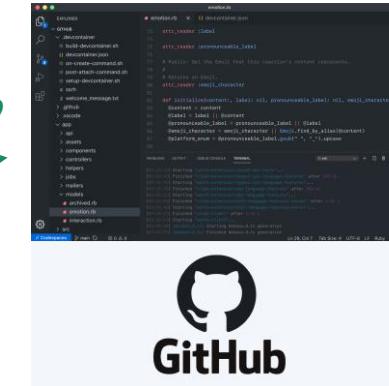
Downloaded from https://journals.asm.org/journal/mra on 10 August 2023 by 141.48.67.87.

Eigene Daten auffindbar machen



Persistente Referenzierung durch DOIs

zB <https://doi.org/10.4119/unibi/123456789>



Optionen für Veröffentlichung von Daten

4.2 Project data will be published:

- Immediately upon collection
- Upon publication of findings.
- At the end of project.
- After embargo period.
- Upon request.
- Upon IP issues have been addressed.

Wann?

Source: crossref.org

Daten teilen – PUB an der Universität Bielefeld



<https://pub.uni-bielefeld.de>

✓ ORCID iD present!

[Connect your ORCID iD](#)

[ORCID iDs in PUB](#)

[+ Add new Publication](#)

pub.uni-bielefeld.de/librecat/record/new?type=research_data

Home My Publication List Admin Deutsch vompr

Research Data

Save Cancel Delete

Details

Publication Type: Research Data (circled)

DOI: e.g. 10.4119/unibi/2695161 I want to register a new DOI

Creator*: First Name: e.g. John Last Name: e.g. Appleseed extern UnibI? [+](#) [-](#)

Department: Departments will be filled in automatically. [+](#) [-](#)

Project: Start typing, or use the key (arrow down) to choose a project from the list. [+](#) [-](#)

Research Group: Start typing, or use the key (arrow down) to choose a research group from the list. [+](#) [-](#)

Title*: e.g. Handbook of eBook Publishing

Publisher*: Bielefeld University

Year of Publication*: e.g. 2021 Published in the same year as the record

Weitere Publikationsmöglichkeiten

Publikationsmöglichkeiten

- Als **Anhang** (supplementary material) (Publikationsort: Verlag, Beispiele: *Integrating Data* oder *American Economic Review*)
- Als **eigenständig** (Publikationsort: Datenbank, Beispiele: PANGEA oder SowiDataNet)
- **Data Journal:** Journal für “Daten in Forschung und Praxis” zur Wiederverwendbarkeit zu fördern. Beispiele: Scientific Data (Nature), Biodiversity Data Journal
- Sonstige generische Repositorien: Zenodo, Open Science Framework, Figshare

PLOS BIOLOGY

BROWSE PUBLISH ABOUT SEARCH advanced search

Introduction
Minimal Data Set Definition
Acceptable Data Sharing Methods
Acceptable Data Access Restrictions
Unacceptable Data Access Restrictions
FAQs
PLOS Data Advisory Board

AMERICAN ECONOMIC ASSOCIATION

ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are openly available in [Repository Name] at [Persistent Link to data in Repository, e.g., DOI, Accession Number].

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/XXXX>.

Experimental procedures and characterization (PDF)

PLOS believes in sharing data freely under that policy.

Data and code have been deposited in a trusted repository.

Please complete ONE option below:

[AEA Data and Code Repository](#)

Project Number: _____

I have followed the AEA Repository [deposit instructions](#).

Other [trusted repository](#)

Deposit DOI: <https://doi.org/> 1314-2828 (online)

Was ist zu beachten, wenn ich Daten teile?

- Alle Ressourcen für die **Nachnutzung verfügbar** und **zugänglich** machen!
 - Vertrauenswürdige Repositorien nutzen | Persistente Identifikatoren verwenden | Offener oder eingeschränkter Zugang | Ethische und rechtliche Vorgaben einhalten
- **Daten** und **Datenverarbeitung** sauber dokumentieren!
 - Metadaten | Verarbeitungsschritte ebenfalls beschreiben
- **Nachvollziehbarkeit** sichern: Setup, Code, Skripte, Trainingsdaten, Parameter, Outputs dokumentieren
 - Standardisierte und offene Formate | Best Practices der Community
- Publikationen mit Daten **verlinken** – z. B. per DOI
 - Zitationsformate | Daten sichtbar machen



Programm

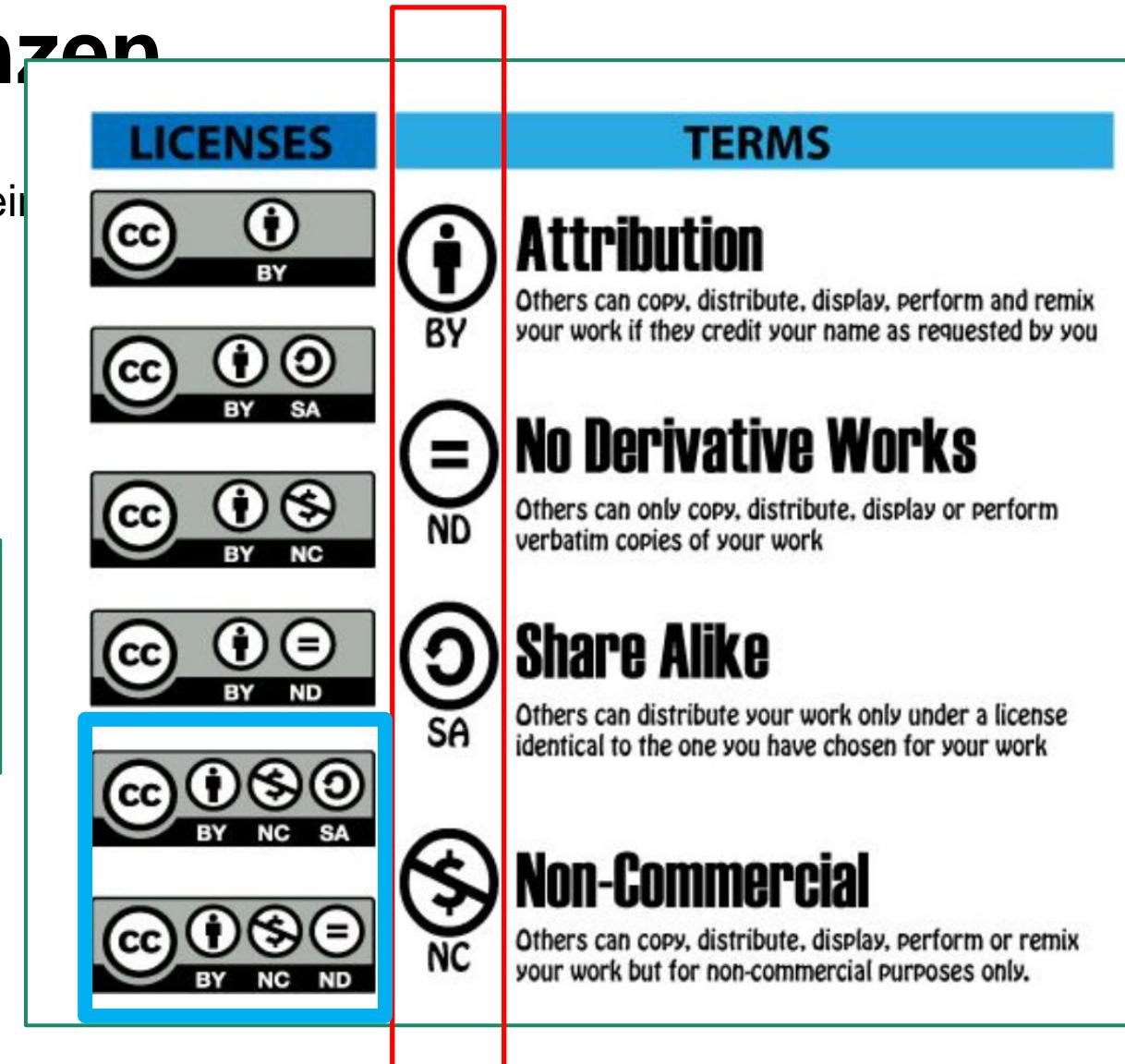
- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden
- Daten nachnutzen**
- Daten organisieren, dokumentieren und sichern
- Abschluss

Daten nachnutzen - Lizenzen

Wenn Sie externe Daten in Ihre eigene Forschung einbinden möchten:

→ Achten Sie auf die Lizenz, unter der die Daten bereitgestellt werden!

Eine Lizenz ist eine erteilte **Erlaubnis/Recht**:
Sie legt fest, was Sie mit den Daten tun dürfen
(z. B. teilen, verändern, wiederverwenden)
und auch unter welchen Bedingungen.



Daten nachnutzen – richtig zitieren

Wenn Sie externe Daten in Ihre eigene Forschung einbinden oder nutzen möchten:

→ **Genutzte Quellen, Daten und Ressourcen zitieren!**

Elemente einer Zitation von Daten/Statistiken:

- Autor(en)/Urheber
- Titel: "Daten aus"
- Publikationsdatum: Das Veröffentlichungs- bzw. Freigabedatum der Statistik/des Datensatzes (nicht das Erhebungs- oder Abdeckungsdatum)
- Herausgeber/Verlag: Datenzentrum/Repository
- **Sonstige Identifikatoren (inkl. Edition oder Version)**
- **Verfügbarkeit und Zugriff: URL / DOI oder andere Ortsangabe**

Daten nachnutzen – Beispiele von Zitationen

Zitationen (externer) Daten / Datenbanken / Ressourcen:

Bureau of Transportation Statistics, Bureau of Transportation Statistics (2021-07-25). 15 Minute or More Delay: 15 Minute or More Delay | , 01/01/2010 - 02/28/2021. Data Planet™ Statistical Datasets: A SAGE Publishing Resource. ([Dataset](#)). Dataset-ID: 007-001-004

Food and Agriculture Organization of the United Nations. (2021). Production of Bananas: top 10 producers. [FAOSTAT](#). Rome, Italy: FAO. Retrieved August 30, 2021 from <http://www.fao.org/faostat/en/#home>

Vobejda, C. (2025). Data: Stability of heart rate at maximal lactate steady state in recreational sportsmen. Bielefeld University. DOI: <https://doi.org/10.4119/unibi/3000793>

Daten nachnutzen – Zitationsstile

Am besten Empfehlungen der **Datenportale oder Repositorien** beachten ...



Citation requirements

The Core Scientific Team of the ESS requests that references to ESS data and the Data Documentation Reports should use the form of words listed below.

To ensure that such source citations must appear in the

Citation of data

- ESS Round 11: European Social Survey European Research Infrastructure (ESS ERIC) (2024) ESS11 - integrated file, edition 2.0 [Data set]. Sikt - Norwegian Agency for Shared Services in Education and Research. <https://doi.org/10.21338/ess11-2024>

Citation of documentation

- ESS Round 11: European Social Survey European Research Infrastructure (ESS ERIC) (2024) ESS round 11 - 2023. Social inequalities in health, Gender in contemporary Europe. Sikt - Norwegian Agency for Shared Services in Education and Research. doi.org/10.21338/ess11-2023.

Citation requirements for

Citation of documentation

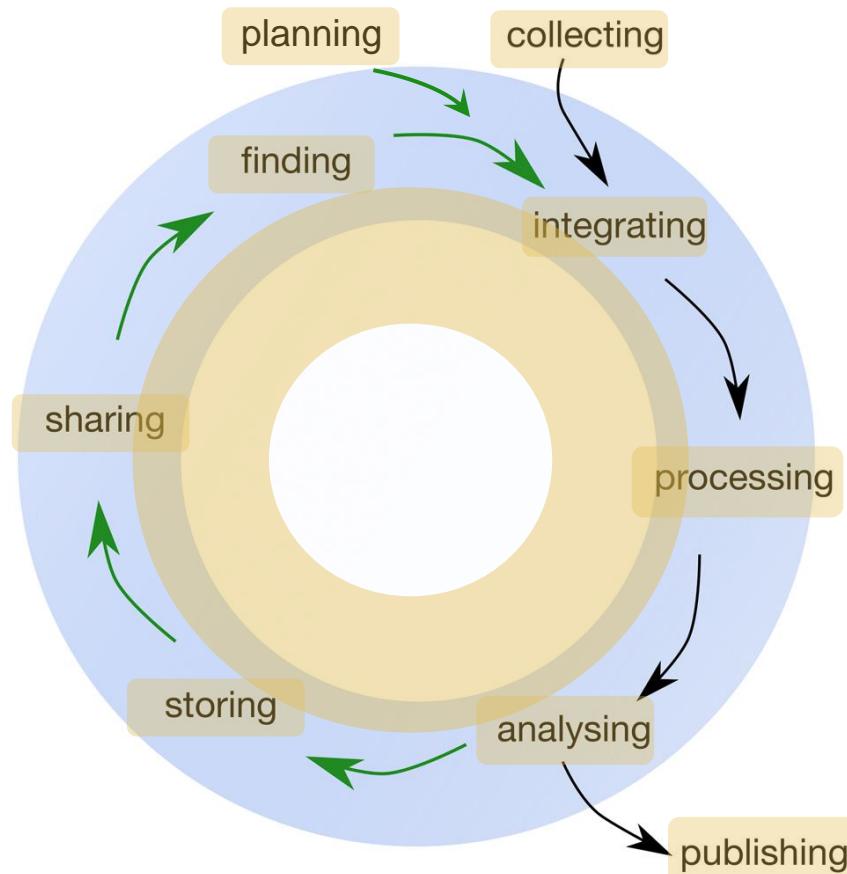
Citation of documentation

- ESS Round 9: European Social Survey European Research Infrastructure (ESS ERIC) (2023) ESS round 9 - 2020. Social inequalities in health, Gender in contemporary Europe. Sikt - Norwegian Agency for Shared Services in Education and Research, Norway for ESS ERIC. [doi:10.21338/NSD-ESS9-2020](https://doi.org/10.21338/NSD-ESS9-2020).

Programm

- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden
- Daten nachnutzen
- Daten organisieren, dokumentieren und sichern**
- Abschluss

Datenorganisation

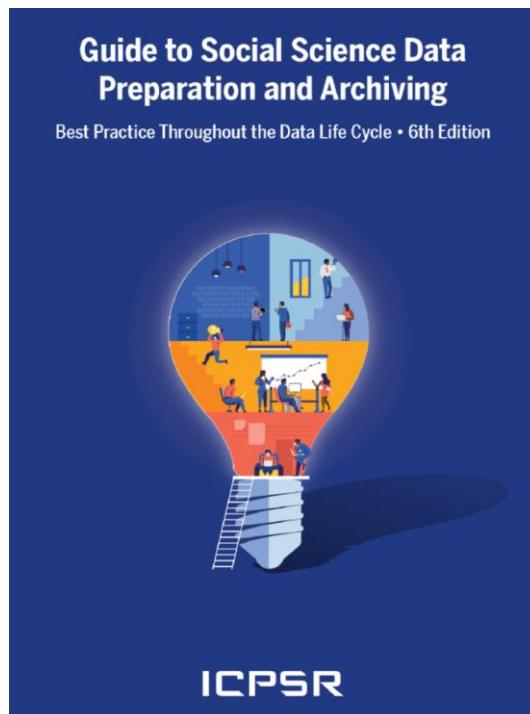


Aspekte der Datenorganisation:

Am besten im Rahmen eines Data Management Plans (DMP) in Vorfeld klären/festlegen:

- Daten – und Dateistruktur
- Ordnerstruktur / Gruppierung
- Benennung von Dateien
- Benennung Variablen, etc.
- Versionierung

Datenorganisation



Namenskonvention

Minimalanforderung:

- Sollten korrekte Informationen vermitteln,
- Benennung sollte eindeutig sein (Codebuch erstellen)
- Länge kompatibel mit technischen Anforderungen.



Variablenamen:
ausgelegt auf eine
langfristige Nutzung
durch ein **breites Publikum**

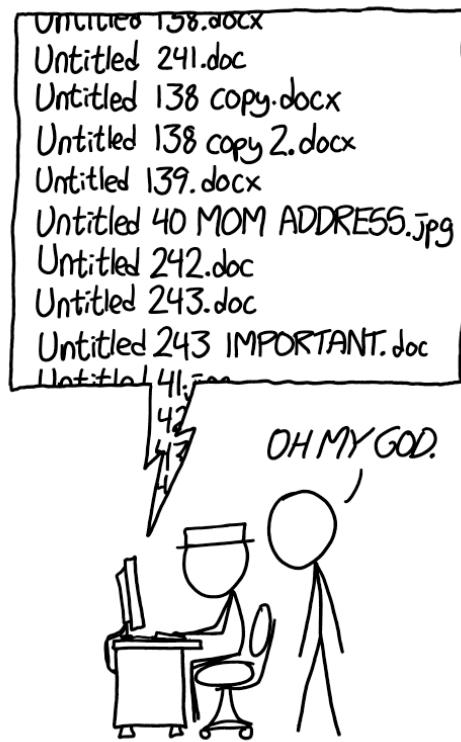
Datenorganisation

Systeme für die Konstruktion von Variablennamen.

Einige Beispiele:

- 1. Zahlenkombinationen:** Variable 1 ... n: **v0001, v0002, ...**
- 2. Fragenummern:** Variable 1 ... n: **Q1, Q2a, Q2b, ...**
- 3. Mnemonische Namen:** z.B. **B_EKOM** (*Einkommen des Befragten*)
- 4. Präfix, Stamm und Suffix:** z.B. Stamm: **ED** für Bildung
Höchster Bildungsstand der Mutter: **MOED**, Vater: **FAED**
Suffix für ergänzende Informationen
Beispiele: **A)** Welle: **MOED_W3** oder **MOED3**
B) Landesspezifizität: **MOED_DE**

Datenorganisation



Bildquelle: <https://www.forschungsdaten.uni-bonn.de/>

Anwendung von Namenskonventionen auf Dateien und Ordner

1. Verschiedene Ordnungsmethoden

- Teilprojekte (TN1, TN2, ...)
- Zeit: YYYY-MM
- Inhaltsbasiert: Dateiinhalt

Historie / Version	
1.0.	(10.3.2021)
1.1.	(17.3.2021)
1.2.	(19.3.2021)
2.0.	(19.6.2017)

- ### 2. Dateiversionierung (größere Anzahl von Dateien)

Changed access method of video elements

Showing 1 changed file with 2 additions and 2 deletions [Show diff stats](#)

[Inline Diff](#) [Side-by-side Diff](#)

synchro.js Wrap text Diff comments [View file @bd9e51b](#)

```
...   ... @@ -22,8 +22,8 @@ function init(array) {  
22     arr = array;  
23     video = document.getElementById("video1");  
24     videoSlides = document.getElementById("video2");  
25     - span1 = document.getElementById("spanvid1");  
26     - span2 = document.getElementById("spanvid2");  
25     + span1 = document.getElementByName("spanvid1");  
26     + span2 = document.getElementByName("spanvid2");  
27     27     quizSolved = false;  
28     28  
29     29     video.addEventListener("timeupdate", function () {  
...   ...
```

Weitere Informationen:

<https://www.forschungsdaten-bildung.de/dateien-benennen#Dateiorganisation-Ordnerstruktur>

Data Crunch handout "DIY: File naming": <https://zenodo.org/records/10275946>

Datensicherung und Backup

Warum Backups?

- Datenverlust kann **jeden** treffen!
- Ihre Daten und Dokumente sind **das Gold** Ihrer Forschung!



Bestandteile von Backups

- Backup-Strategie
- Dedizierte Backup-Hardware (Speicherorte)
- Backup-Software

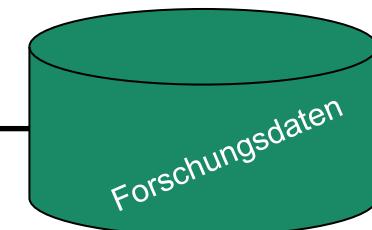
Datensicherung und Backup

Ersetzbarkeit / Wiederherstellbarkeit von Daten

Leicht	Mittel/Schwer	Unmöglich
Betriebssystem, kostenfreie Anwendungen	Anwendungseinstellungen / Betriebssystemeinstellungen	Selbst gemachte Fotos, Videos
Fotos, die von anderen gemacht wurden	Adressdatenbanken	(Abschluss-)Arbeiten, Steuerunterlagen
gekauft digitale Inhalte wie Musik, eBooks, Filme; privat genutzte Software	Verlauf Kommunikation (Messengerdienste); Email-Kommunikation	generell: Erzeugnisse eigener Arbeit
Lizenzen für beruflich genutzte Software oder alles was frei im Internet zur Verfügung steht (Publikationen, Daten)		

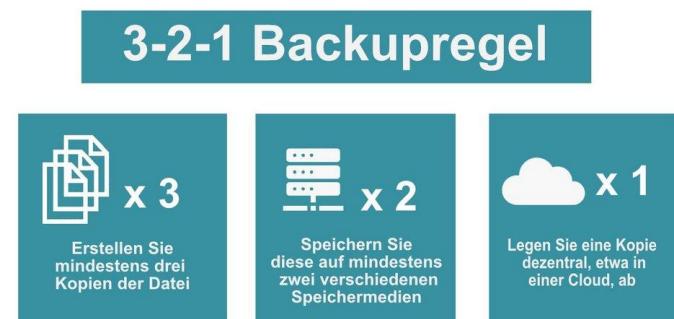
Vor allem kritisch:

- mühselig erhobene Daten, Experimentaldaten, ...
- Weitere Hilfsdokumente (Dokumentationen, Skripte), ...
- Drafts von Publikationen, usw.



Datensicherung und Backup

- Backup machen! Backup jetzt erstellen!
- Anfangen – z. B. mit einer Kopie wichtiger Ordner und Dateien auf einem USB-Stick / externen Laufwerk
- Die 3-2-1-Regel: 3 Kopien auf 2 Datenträgern, davon 1 außer Haus



Beispiel: Eine Kopie auf der lokalen Platte, zwei auf zwei USB-Datenträgern, eine davon bei Verwandten, Eltern, Arbeitsplatz zusätzlich in der Cloud

- Kontrolle des Backups – Restore-Fall testen!

Programm

- Verständnis von Forschungsdaten
- Empfehlungen und Anforderungen
- Daten finden
- Daten nachnutzen
- Daten organisieren, dokumentieren und sichern
- **Abschluss**

Was wir Ihnen mitgeben möchten



Denken Sie frühzeitig an **das Datenmanagement!**

Nutzen Sie die **Beratungsangebote** der Universität Bielefeld

Ein gutes Datenmanagement ist ein **Qualitätssiegel** Ihrer Forschung. Es hilft dabei, die **offene Wissenschaft** zu leben (Open Science)

Es legt auch die Weichen für:

- Ermöglichung der Weitergabe und **Wiederverwendung** von Daten
- Gute wissenschaftliche Praxis: **Überprüfung** von wissenschaftlichen Ergebnissen möglich!
- Effizienzsteigerung: **Vermeidung von doppelter Arbeit**, (z.B. schnelles Onboarding neuer KollegInnen)

vielen Dank!



Markus Rump (BITS & UB), Johanna Vompras (UB), Anne Gärtner (UB)

**Kompetenzzentrum
Forschungsdaten**

Tel.: 0521-106-3967

E-Mail: data@uni-bielefeld.de

<https://data.uni-bielefeld.de>

sprechen Sie uns an!