



Staatsbibliothek
zu Berlin
Preußischer Kulturbesitz

Forschungsdatenmanagement an der Freien Universität Berlin, der Humboldt-Universität zu Berlin und der Technischen Universität Berlin

Veranstaltungsreihe *Lost in Dissertation?*
an der Staatsbibliothek zu Berlin
29. November 2022



Agenda

1. Einführung ins Forschungsdatenmanagement (Asef)
2. Forschungsdatenmanagement-Services der Freien Universität Berlin (Asef)
3. Forschungsdatenmanagement-Services der Humboldt-Universität zu Berlin (Jäckel)
4. Forschungsdatenmanagement-Services der Technischen Universität Berlin (Kuberek)
5. Zeit für Fragen

Einführung ins Forschungsdatenmanagement

Esther Asef
Referentin für Forschungsdatenmanagement

Freie Universität Berlin
Universitätsbibliothek

 UNIVERSITÄTS
BIBLIOTHEK

forschungsdaten
management 

Inhalt

1. Definitionen und Beispiele
2. FDM in der Praxis
3. Vorgaben und Kontext



DEFINITIONEN UND BEISPIELE

Was sind Forschungsdaten?

„Unter Forschungsdaten werden alle (digitalen) Daten zusammengefasst, die Gegenstand, Arbeitsschritte oder Ergebnis von Forschungsprozessen sind.“

→ Grasse, Marleen, López, Ania, & Winter, Nina. (2018, March 5). Musterleitlinie für Forschungsdatenmanagement (FDM) an Hochschulen und Forschungseinrichtungen. Zenodo. <http://doi.org/10.5281/zenodo.1149133>

Was sind Forschungsdaten?

Texte
Fragebögen
Objektsammlungen
Simulationen
Protokolle

Messdaten
Software
Texteditionen
Beobachtungsdaten
audiovisuelle Daten

Datenbanken
Befragungsdaten
Digitalisate
Quellcode

Leitfaden, Transkripte und audiovisuelle Daten

→ Thompson, P. (2019). Pioneers of Social Research, 1996-2018. [data collection]. 4th Edition. UK Data Service. SN: 6226, <http://doi.org/10.5255/UKDA-SN-6226-6>
 UK Data Service „Pioneers of Social Research“- Youtube Channel <https://www.youtube.com/channel/UCmK1mj5dCq0XuCWI7DVKx2w/about>

INTERVIEW TRANSCRIPT

I: If we could go back to the beginning. You were born in 1922?

R: That's right.

I: And where was that?

R: In London, in England.

I: Yes. And could you tell me about your family background? I don't know how far back you remember. Your grandparents, for instance?

R: Yes. Well, I knew my maternal grandmother very well, as is quite commonly the case. She lived well into her nineties. My maternal grandfather died when my mother was about three months pregnant, so I never knew him, and she was devastated by this, because she was greatly attached to him.

I: What was his profession, then?

R: I think that he was a salesperson. It's a little bit difficult to tell from what I've read, but he had some kind of business activity, and I had heard, at one time, it was to do with boots and shoes - but the description gets vague. What I am clear about is that the family was Methodist, and my grandmother had converted to Methodism because of my grandfather. I don't know what her lineage had been before. But what I do know is that she was - although I didn't discover this until I had a child

UK Data Archive SN 6226 - Pioneers of Social Research, 1996-2018

Pioneers of Social Research: Interview Guide

Pre-interview
 Look for an outline biography of cv of the person you are recording (Wikipedia, Who's Who, or they may send you a cv)

Try to give a hard look at what seem to be the researcher's key books

For every interview it is essential to establish in the early part

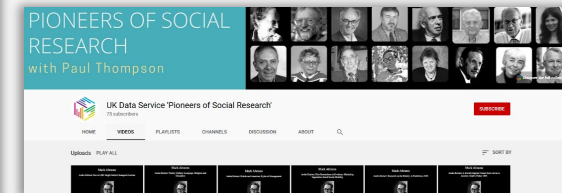
- the date and place of birth of the interviewee, their family background and occupations;
- what led them towards research and university work;
- what were they key influences in leading them towards their pioneering research themes.
-

Opening questions

Date of birth, family and own occupations.

Can we begin by talking about your family background and early life, and any early influences which may have led you towards university work?

FAMILY AND EARLY LIFE



Lipid Analytic of Environmental and Cultural Samples

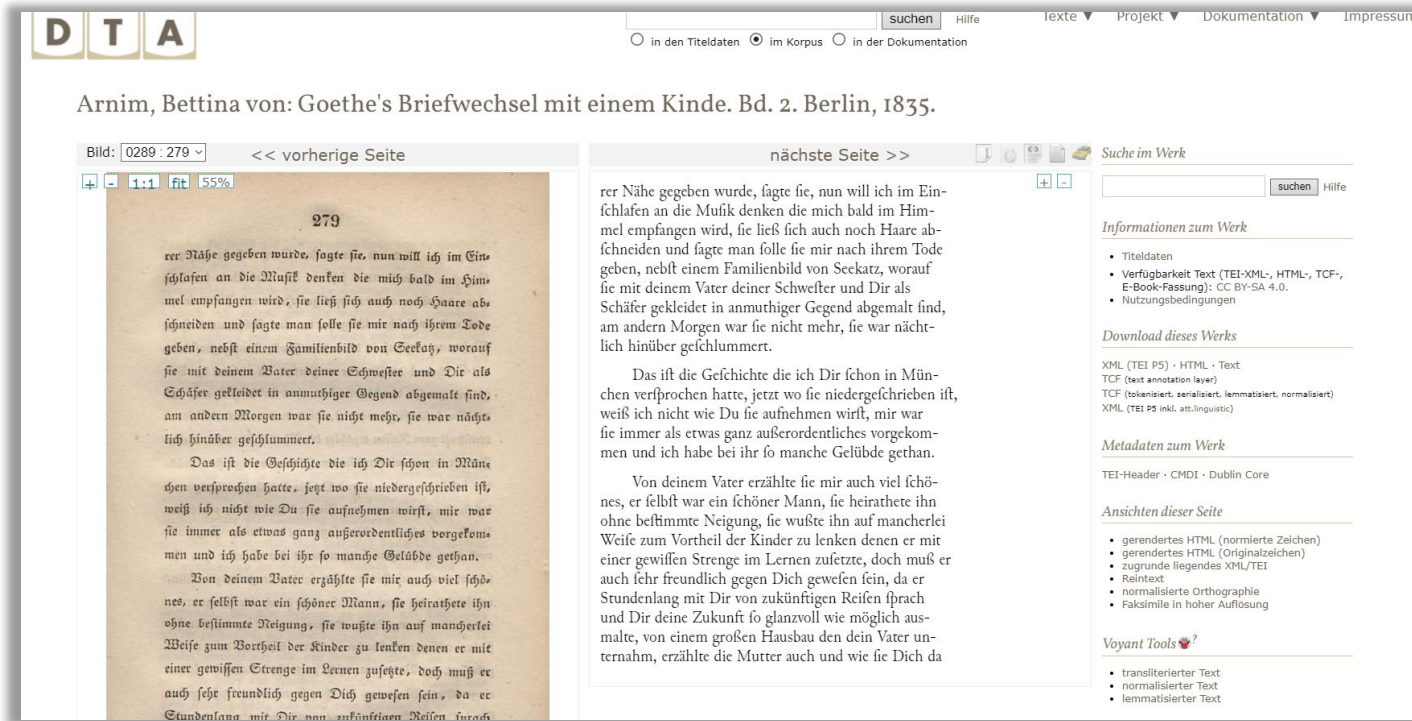
Data

Download dataset as tab-delimited text — use the following character encoding: [UTF-8: Unicode (PANGAEA default)]

Event	Code	Instrument	Area	Comment	O2 [μmol/l]	Site	Depth water [m]	Vol [l]	Sal	Temp [°C]	GDGT-0 peak area	GDGT-1 peak area	GDGT-2 peak area	GDGT-3 peak area	GDGT-5 Iso peak area	GDGT-9 peak area	GDGT-8 peak area
CB_trap	61	TOF	396	fully oxix	1		45	811.3 >35.0000	22.60	645000	244000	314000	110000	104000	1810000		
CB_trap	62	TOF	396	fully oxix	1		200	869.8 >35.0000	16.80	12500000	4620000	9070000	1120000	3800000	29400000		
CB_trap	63	TOF	396	fully oxix	1		700	1730.1 >35.0000	8.50	13800000	6030000	9050000	4010000	3020000	27100000		
CB_trap	64	TOF	396	fully oxix	1		1500	1757.3 >35.0000	5.00	4930000	2080000	2990000	1150000	913000	9880000		
CB_trap	65	TOF	396	fully oxix	1		2500	1456.3 >35.0000	2.90	4280000	1710000	2130000	92400	686000	7650000		
CB_trap	66	TOF	396	fully oxix	1		3500	2028.2 >35.0000	1.50	6790000	2540000	3080000	168000	1030000	12000000		
CB_trap	67	TOF	396	fully oxix	1		4120	1416.1 >35.0000	0.90	9370000	2020000	1880000	148000	603000	11300000		
CB_trap	69	TOF	396	fully oxix	2		45	410.7 >35.0000	19.50	1930000	595000	502000	158000	134000	4040000		
CB_trap	70	TOF	396	fully oxix	2		150	591.3 >35.0000	16.10	24300000	7540000	7710000	1580000	2670000	34500000		
CB_trap	71	TOF	396	fully oxix	2		300	1271.9 >35.0000	13.10	28700000	12200000	16100000	1150000	5160000	41800000		
CB_trap	72	TOF	396	fully oxix	2		1000	1095.9 >35.0000	6.60	11400000	4340000	4880000	252000	1500000	17800000		
CB_trap	73	TOF	396	fully oxix	2		1900	1182.9 >35.0000	3.90	9270000	3100000	3180000	223000	1050000	14300000		
CB_trap	74	TOF	396	fully oxix	2		2000	1039.8 >35.0000	3.80	11700000	4140000	4130000	274000	1340000	17700000		
CB_trap	75	TOF	396	fully oxix	2		2200	1137.7 >35.0000	3.50	11800000	3950000	4140000	276000	1330000	16900000		
CB_trap	76	TOF	396	fully oxix	2		2620	1102.8 >35.0000	3.00	12000000	3830000	3780000	296000	1180000	16600000		
CB_trap	77	TOF	396	fully oxix	3		45	636.3 >35.0000	18.90	13300000	3030000	2140000	836000	500000	19000000		
CB_trap	78	TOF	396	fully oxix	3		100	450.9 >35.0000	15.70	31400000	11900000	14400000	3010000	5890000	48900000		
CB_trap	79	TOF	396	fully oxix	3		170	984.1 >35.0000	14.10	7860000	2460000	3390000	383000	1320000	14400000		
CB_trap	80	TOF	396	fully oxix	3		300	945.6 >35.0000	12.80	30000000	13500000	17700000	1270000	6300000	42400000		
CB_trap	81	TOF	396	fully oxix	3		600	839.7 >35.0000	9.70	17900000	7670000	10100000	522000	3490000	29500000		
CB_trap	82	TOF	396	fully oxix	3		720	1220.4 >35.0000	8.40	25600000	10000000	11100000	873000	3900000	36200000		
CB_trap	83	TOF	425	fully oxix	4		50	358.2 >35.0000	19.70	2800000	632000	499000	182000	156000	4450000		
CB_trap	84	TOF	425	fully oxix	4		150	902.6 >35.0000	17.20	8780000	2680000	3120000	603000	1280000	15700000		
CB_trap	85	TOF	425	fully oxix	4		400	1128.3 >35.0000	12.20	8300000	3380000	5690000	320000	2010000	17900000		
CB_trap	86	TOF	425	fully oxix	4		1000	220.4 >35.0000	6.70	872000	324000	478000	22800	139000	1730000		
CB_trap	87	TOF	425	fully oxix	4		2200	1176.5 >35.0000	3.50	2560000	920000	1100000	61000	334000	4600000		
CB_trap	89	TOF	425	fully oxix	5		55	484.0 >35.0000	20.20	4940000	1240000	821000	362000	248000	8450000		
CB_trap	90	TOF	425	fully oxix	5		150	1566.5 >35.0000	15.50	15800000	6600000	9880000	873000	4320000	30200000		
CB_trap	91	TOF	425	fully oxix	5		350	870.4 >35.0000	13.30	8790000	3340000	4550000	324000	1700000	17200000		
CB_trap	92	TOF	425	fully oxix	5		1000	820.8 >35.0000	6.70	2960000	1140000	1590000	75900	538000	6080000		
CB_trap	93	TOF	425	fully oxix	6		50	347.7 >35.0000	19.50	3810000	854000	612000	237000	195000	6840000		
CB_trap	94	TOF	425	fully oxix	6		150	1072.1 >35.0000	18.00	12300000	4280000	6110000	750000	2920000	25200000		
CB_trap	95	TOF	425	fully oxix	6		350	1559.3 >35.0000	13.00	3580000	1530000	2660000	176000	1060000	8500000		
CB_trap	96	TOF	425	fully oxix	6		400	979.6 >35.0000	14.50	14500000	6560000	10600000	614000	4050000	27500000		

→ Zhu, Chun; Wakeham, Stuart G; Elling, Felix J; Basse, Andreas; Mollenhauer, Gesine; Versteegh, Gerard J M; Könneke, Martin; Hinrichs, Kai-Uwe (2016): Lipid analytic of environmental and cultural samples. PANGAEA, <https://doi.org/10.1594/PANGAEA.845385>

Maschinenlesbarer Korpus: Deutsches Textarchiv



The screenshot shows the interface of the Deutsches Textarchiv. At the top, there are navigation tabs for 'D', 'T', and 'A'. A search bar is present with the text 'suchen' and a dropdown menu for 'Hilfe'. Below the search bar, there are radio buttons for 'in den Titeldaten', 'im Korpus' (selected), and 'in der Dokumentation'. The main content area is titled 'Arnim, Bettina von: Goethe's Briefwechsel mit einem Kinde. Bd. 2. Berlin, 1835.' Below this, there is a navigation bar with 'Bild: 0289 · 279', '<< vorherige Seite', and 'nächste Seite >>'. The left pane shows a scanned image of a page with the number '279' at the top. The text on the page is:

279

rer Nähe gegeben wurde, sagte sie, nun will ich im Einschlafen an die Musik denken die mich bald im Himmel empfangen wird, sie ließ sich auch noch Haare abschneiden und sagte man solle sie mit nach ihrem Tode geben, nebst einem Familienbild von Seekatz, worauf sie mit deinem Vater deiner Schwester und Dir als Schäfer gekleidet in anmuthiger Gegend abgemalt find, am andern Morgen war sie nicht mehr, sie war nämlich hinüber geschlummert.

Das ist die Geschichte die ich Dir schon in Mäntchen versprochen hatte, jetzt wo sie niedergeschrieben ist, weiß ich nicht wie Du sie aufnehmen wirst, mir war sie immer als etwas ganz außerordentliches vorgekommen und ich habe bei ihr so manche Gelübde gethan.

Von deinem Vater erzählte sie mir auch viel schönes, er selbst war ein schöner Mann, sie heirathete ihn ohne bestimmte Neigung, sie wußte ihn auf mancherlei Weise zum Vortheil der Kinder zu lenken denen er mit einer gewissen Strenge im Lernen zusetzte, doch muß er auch sehr freundlich gegen Dich gewesen sein, da er Stundenlang mit Dir von zukünftigen Reisen sprach.

 The right pane contains a search bar 'Suche im Werk' and a list of 'Informationen zum Werk' including 'Titeldaten', 'Verfügbarkeit Text (TEI-XML-, HTML-, TCF-, E-Book-Fassung): CC BY-SA 4.0', and 'Nutzungsbedingungen'. Below that are 'Download dieses Werks' options: 'XML (TEI P5) · HTML · Text', 'TCF (text annotation layer)', 'TCF (tokenisiert, serialisiert, lemmatisiert, normalisiert)', and 'XML (TEI P5 inkl. att.linguistic)'. There is also a 'Metadaten zum Werk' section with 'TEI-Header · CMDI · Dublin Core'. The 'Ansichten dieser Seite' section lists: 'gerendertes HTML (normierte Zeichen)', 'gerendertes HTML (Originalzeichen)', 'zugrunde liegendes XML/TEI', 'Reintext', 'normalisierte Orthographie', and 'Faksimile in hoher Auflösung'. At the bottom, there is a 'Voyant Tools' section with 'transliterierter Text', 'normalisierter Text', and 'lemmatisierter Text'.

→ Arnim, Bettina von: Goethe's Briefwechsel mit einem Kinde. Bd. 2. Berlin, 1835, S. 279. In: Deutsches Textarchiv http://www.deutschestextarchiv.de/arnimb_goethe02_1835/289, abgerufen am 27.10.2020.

Was ist Forschungsdatenmanagement?

Forschungsdatenmanagement begleitet den gesamten Forschungsprozess und umfasst

- Organisation
- Dokumentation
- Speicherung und Sicherung
- Archivierung
- Teilen und Publikation

von Daten.



FORSCHUNGSDATEN- MANAGEMENT IN DER PRAXIS

Aspekte von Forschungsdatenmanagement



PLANEN



ORGANISIEREN UND
DOKUMENTIEREN



ETHIK, SICHERHEIT
UND RECHT



SPEICHERN



ARCHIVIEREN



TEILEN UND
PUBLIZIEREN



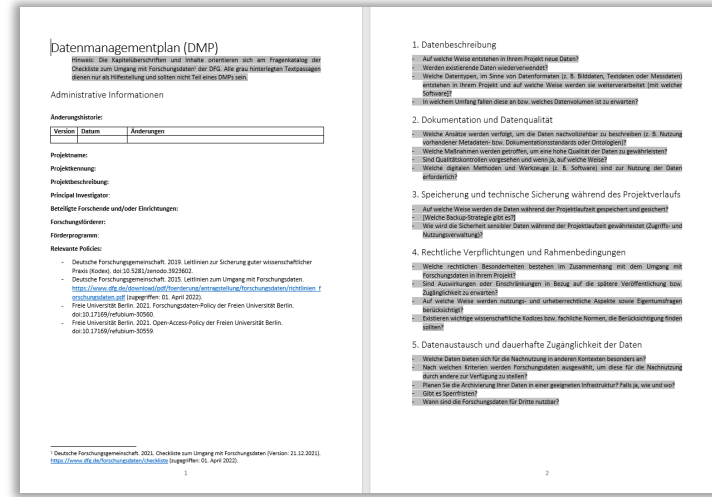
ZITIEREN UND
NUTZEN



PLANEN

Datenmanagementpläne (DMP)

- Strukturierte Dokumentation aller Aspekte der Forschungsdaten in einem Projekt
- Umfasst den Umgang mit Daten im gesamten Projektverlauf sowie die Datenvorhaltung nach dem Projektabschluss
- Ermöglicht Aspekte des Forschungsdatenmanagements frühzeitig und im Team zu planen



➔ Screenshot: DMP-Vorlage (Freie Universität Berlin), <https://www.fu-berlin.de/sites/forschungsdatenmanagement/in-der-praxis/planung/dmp.html>, Zugriff 25.04.2022.

Datenmanagementpläne (DMP)

Angaben u.a. zu:

- Datentypen, -formate, -umfang
- Methoden
- Metadaten und Standards
- Verantwortlichkeiten
- Kosten

CESSDA
Cooperation of European Social Science Data Archives

DMP

Overview

Title of the project/study	
Date of this plan	
Description of the project	e.g. What is the research question?
Origin of Data	e.g. <ul style="list-style-type: none"> What kind of data will be used during the project? If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated? If you are collecting new data can you clarify why this is necessary?

Organising and documenting your data

Data collection	e.g. <ul style="list-style-type: none"> How will the data be collected? Is specific software or hardware or staff required? Who will be responsible for the data collection? During which period will the data be collected?
-----------------	---

	<ul style="list-style-type: none"> Where will the data be collected?
Date type and size	e.g. <ul style="list-style-type: none"> What type(s) of data will be collected? What is the scope, quantity and format of the material? After the project: What is the total amount of data collected (in MB/GB)?
File format	e.g. <ul style="list-style-type: none"> In what format will your data be? Does the format change from the original to the processed/final data? Will your (final) data be available in an open format?
Folder structure & names	e.g. <ul style="list-style-type: none"> How will you structure and name your folders?
File structure & names	e.g. <ul style="list-style-type: none"> How will you structure and name your files?
Documentation	e.g. <ul style="list-style-type: none"> What documentation will be created during the different phases of the project? How will the documentation be structured?
Metadata	e.g. <ul style="list-style-type: none"> What metadata will be provided with the collected/generated/unused data? How will metadata for each object be created? Is there any program that can be used to document the data?

➔ [Muster und Templates von CESSDA](#)



ORGANISIEREN UND DOKUMENTIEREN

Dateien benennen



Möglichst kurz und aussagekräftig



Zulässige Zeichen: A-Z a-z 0-9 Bindestrich (-) Unterstrich (_)



Unzulässig: Umlaute, sz-Ligatur, Klammern, andere als die oben genannten Satzzeichen und Sonderzeichen



Beachten von Groß- und Kleinschreibung



Datumsangaben JJJJ-MM-TT



Einfache Versionierung durch Suffix `_v01`, `_v02` ...



> Ähnliches gilt für Ordner



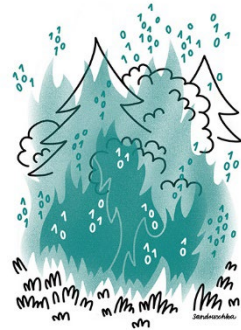
einheitliche Standards innerhalb eines Projektes/einer Arbeitsgruppe



SPEICHERN

Was ist hier mit den Daten passiert?

Daten in Brandgefahr



Der Klimawandel schlägt im Forst doppelt zu



Forschungsdaten Thüringen: Scary Tales NR- 57 <https://forschungsdaten-thueringen.de/geschichten/articles/20221026-daten-in-brandgefahr.html>

Backup und Datensicherheit

3-2-1 Backupregel



→ Quelle: [3-2-1 Backupregel](#) - I. Lang/Bearbeitung E. Böker / [CC BY 4.0](#)

- Regelmäßige Backups zu einem festgelegten Zeitpunkt durchführen
- Systeme der Universität verwenden
- Vor unbefugten Zugriff schützen, z. B. mit Passwortschutz



ARCHIVIEREN

Archivierung

- i. d. R. mindestens 10 Jahre
- Möglichst in unkomprimierten, offenen, nicht proprietären Standardformaten
- Zusammen mit Software und Analyseskripten, Dokumentation und Metadaten
- i. d. R. in Datenarchiven oder Repositorien



TEILEN UND PUBLIZIEREN

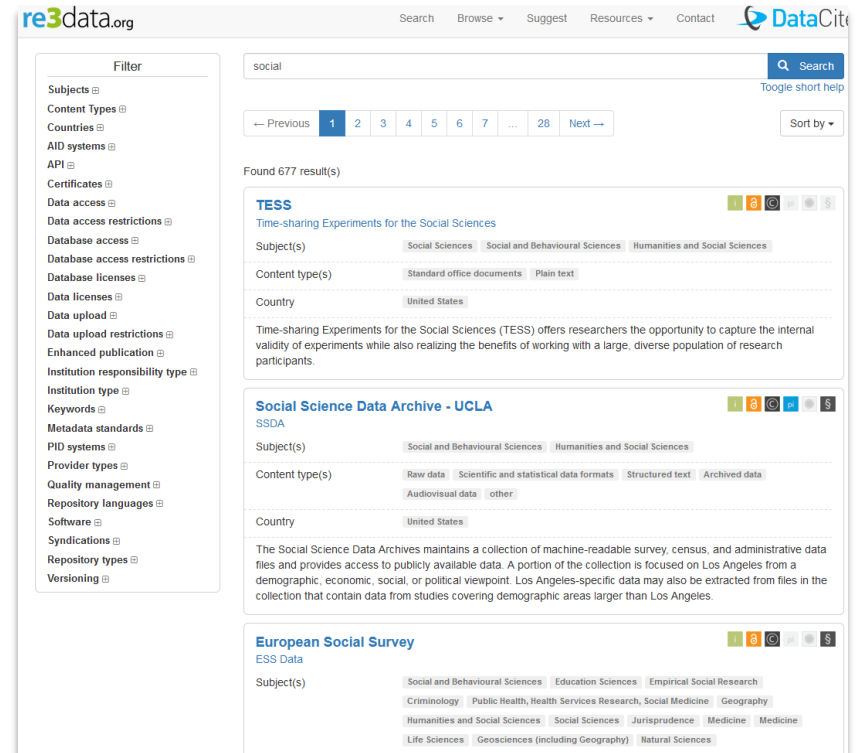
Forschungsdaten publizieren

Als eigenständige Publikation in Forschungsdaten-Repositoryn:

- fachspezifische Repositoryn, z.B. [PANGAEA](#)
- generische Repositoryn, z. B. [ZENODO](#)
- institutionelle Repositoryn, z. B. [DepositOnce](#), [edoc-Server](#), [Refubium](#)

Repositorien finden: re3data.org

- Weltweites Register
- Berücksichtigt alle Disziplinen
- Repositorien, die dauerhafte Speicherung und Veröffentlichung der Daten anbieten
- Verschiedene Kriterien als Filtermöglichkeit



The screenshot shows the re3data.org search interface. The search term 'social' has yielded 677 results. Three results are visible:

- TESS** (Time-sharing Experiments for the Social Sciences):
 - Subject(s): Social Sciences, Social and Behavioural Sciences, Humanities and Social Sciences
 - Content type(s): Standard office documents, Plain text
 - Country: United States
 - Description: Time-sharing Experiments for the Social Sciences (TESS) offers researchers the opportunity to capture the internal validity of experiments while also realizing the benefits of working with a large, diverse population of research participants.
- Social Science Data Archive - UCLA** (SSDA):
 - Subject(s): Social and Behavioural Sciences, Humanities and Social Sciences
 - Content type(s): Raw data, Scientific and statistical data formats, Structured text, Archived data, Audiovisual data, other
 - Country: United States
 - Description: The Social Science Data Archives maintains a collection of machine-readable survey, census, and administrative data files and provides access to publicly available data. A portion of the collection is focused on Los Angeles from a demographic, economic, social, or political viewpoint. Los Angeles-specific data may also be extracted from files in the collection that contain data from studies covering demographic areas larger than Los Angeles.
- European Social Survey** (ESS Data):
 - Subject(s): Social and Behavioural Sciences, Education Sciences, Empirical Social Research, Criminology, Public Health, Health Services Research, Social Medicine, Geography, Humanities and Social Sciences, Social Sciences, Jurisprudence, Medicine, Medicine, Life Sciences, Geosciences (including Geography), Natural Sciences

Checkliste Veröffentlichung

Veröffentlichungsort

- ✓ Passend für Datentyp und Fachgebiet
- ✓ Vertrauenswürdige Anbieter, nachhaltig
- ✓ Vergabe eines persistenten Identifikators
- ✓ Klare Nutzungsbedingungen für Datenautor*innen sowie Nutzer*innen
 - z. B. Lizenzen, Kosten

Datenvorbereitung

- ✓ Möglichst offene Formate und Lizenzen
- ✓ Detaillierte Metadaten und Dokumentation



ZITIEREN

Daten zitieren

DataCite Schema:

Creator (Publication Year): Title. **Version. Publisher/Data Repository or Archive. Resource Type. Identifier**

- DataCite Metadata Working Group. (2011). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 2.1. DataCite e.V. <http://doi.org/10.5438/0003>
- Principles and further examples: Data Citation Synthesis Group (2014). Joint Declaration of Data Citation Principles. Martone M. (ed.). San Diego CA: FORCE11. <https://doi.org/10.25490/a97f-egyk>



VORGABEN UND KONTEXT

Einige Vorteile von FDM

- Schutz vor Datenverlust
- Durchdachte Ordnung und Dokumentation macht es leichter, Daten zu finden und zu verstehen
- Es ist einfacher, Daten mit anderen zu teilen
- Daten können publiziert, nachgenutzt und zitiert werden
- Erfüllung von Vorgaben und Empfehlungen von Förderern
- Beitrag zur Öffnung der Wissenschaft (Open Science/Research)
- ...

Mögliche Vorgaben und Leitlinien

- **Institutionelle Policies**
- Vorgaben von **Förderinstitutionen**
- Vorgaben aus der **Fachdisziplin**
- [Fachspezifische Empfehlungen](#) (gesammelt von der DFG)
- Vorgaben von **Verlagen**

Leitlinien zur Sicherung guter wissenschaftlicher Praxis

- Leitlinie 7: Phasenübergreifende Qualitätssicherung
- Leitlinie 10: Rechtliche und ethische Rahmenbedingungen, Nutzungsrechte
- Leitlinie 11: Methoden und Standards
- Leitlinie 12: Dokumentation
- Leitlinie 13: Herstellung von öffentlichem Zugang zu Forschungsergebnissen
- Leitlinie 17: Archivierung



→ Deutsche Forschungsgemeinschaft. (2022). Leitlinien zur Sicherung guter wissenschaftlicher Praxis. <https://doi.org/10.5281/zenodo.6472827>

FAIR-Prinzipien



→ Quelle der Grafik : [Australian National Data Service](https://www.ands.org.au/working-with-data/fairdata/training) [ANDS] FAIR resources graphic
<https://www.ands.org.au/working-with-data/fairdata/training> CC BY 4.0

→ Weitere Infos:
<https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/faire-daten/>

Weiterführende Information

- Forschungsdaten.info <https://www.forschungsdaten.info/>
- CESSDA Data Management Expert Guide
<https://www.CESSDA.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide>

Danke für Ihre Aufmerksamkeit!

Esther Asef

Referentin für Forschungsdatenmanagement
Freie Universität Berlin

E-Mail: forschungsdaten@fu-berlin.de

Web: <https://www.fu-berlin.de/forschungsdatenmanagement>

Twitter: [fdm_FUBerlin](https://twitter.com/fdm_FUBerlin)

 UNIVERSITÄTS
BIBLIOTHEK

forschungsdaten
management 



Dieses Werk ist unter einer [Creative Commons Namensnennung 4.0 International Lizenz](https://creativecommons.org/licenses/by/4.0/) lizenziert
(mit Ausnahme der verwendeten Logos und Bilder [Attribution auf entsprechenden Folien eingefügt]).

Alle Screenshots wurden am 12.10.2020/6.9.2021 aufgenommen und bilden die verlinkten Webseiten ab. Bildquellen sind auf jeder Folie genannt, Piktogramme aus Microsoft Powerpoint ohne Urheberrecht <https://support.microsoft.com/de-de/office/einfügen-von-piktogrammen-in-microsoft-office-e2459f17-3996-4795-996e-b9a13486fa79> [Zugriff 16.11.2022]