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

In der Veranstaltungsreihe *Lost in Dissertation?*
an der Staatsbibliothek zu Berlin
8. Oktober 2021

Staatsbibliothek zu Berlin
Preußischer Kulturbesitz

Agenda

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

Einführung ins Forschungsdatenmanagement

Esther Asef
Referentin für Forschungsdatenmanagement

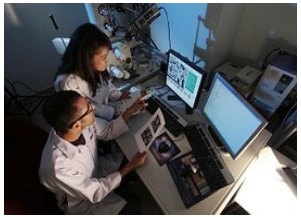
Freie Universität Berlin
Universitätsbibliothek

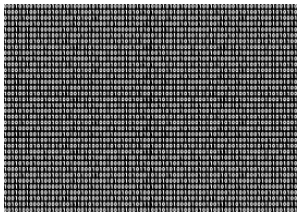
Inhalt

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

DEFINITIONEN UND BEISPIELE

Was sind Forschungsdaten? I





„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? II

- **Messdaten**
- **Laborwerte**
- **audiovisuelle Informationen**
- **Texte**
- **Surveydaten**
- **Objekte aus Sammlungen**
- **Proben**
- **Methodische Testverfahren**
- **Fragebögen**
- **Software**

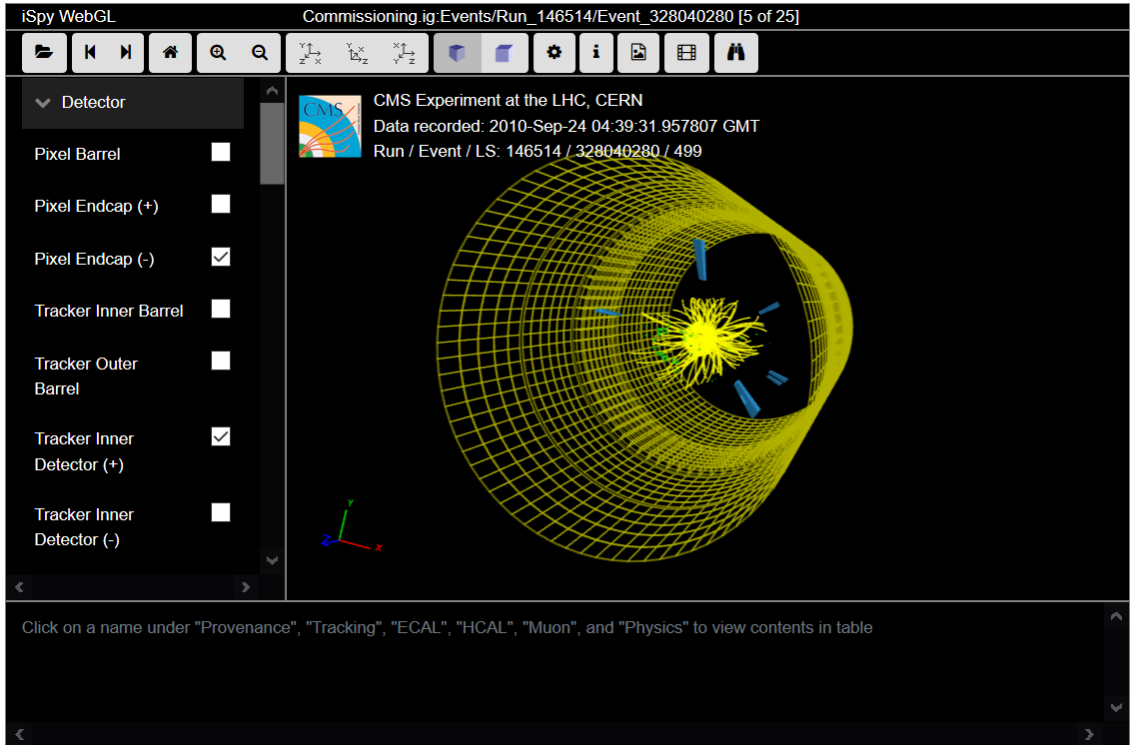
Deutsche Forschungsgemeinschaft (DFG). 2015. „Leitlinien zum Umgang mit Forschungsdaten“. https://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf. Zugriff 27.10.2020

The ZPID lockdown measures dataset

Column1	Column2	Column3	Column4	Column5	Column6	Column7	Column8	Column9	Column10	Column11	Column12	Column13	Column14	Column15	Column16
	stateID	state	Measure	2020-03-08	2020-03-09	2020-03-10	2020-03-11	2020-03-12	2020-03-13	2020-03-14	2020-03-15	2020-03-16	2020-03-17	2020-03-18	2020-03-1
1	1	Baden-Württemberg	leavehome	0	0	0	0	0	0	0	0	0	0	0	0
2	1	Baden-Württemberg	dist	0	0	0	0	0	0	0	0	0	0	0	0
3	1	Baden-Württemberg	msk	0	0	0	0	0	0	0	0	0	0	0	0
4	1	Baden-Württemberg	shppng	0	0	0	0	0	0	0	0	0	0	2	2
5	1	Baden-Württemberg	hcut	0	0	0	0	0	0	0	0	0	0	0	0
6	1	Baden-Württemberg	ess_shps	0	0	0	0	0	0	0	0	0	0	1	1
7	1	Baden-Württemberg	zoo	0	0	0	0	0	0	0	0	0	0	2	2
8	1	Baden-Württemberg	demo	0	0	0	0	0	0	0	0	0	1	1	1
9	1	Baden-Württemberg	school	0	0	0	0	0	0	0	0	0	1	1	1
10	1	Baden-Württemberg	church	0	0	0	0	0	0	0	0	0	0	2	2
11	1	Baden-Württemberg	onefriend	0	0	0	0	0	0	0	0	0	0	0	0
12	1	Baden-Württemberg	morefriends	0	0	0	0	0	0	0	0	0	0	0	0
13	1	Baden-Württemberg	plygrnd	0	0	0	0	0	0	0	0	0	0	2	2
14	1	Baden-Württemberg	daycare	0	0	0	0	0	0	0	0	0	1	1	1
15	2	Bayern	leavehome	0	0	0	0	0	0	0	0	0	0	0	0
16	2	Bayern	dist	0	0	0	0	0	0	0	0	0	0	0	0
17	2	Bayern	msk	0	0	0	0	0	0	0	0	0	0	0	0
18	2	Bayern	shppng	0	0	0	0	0	0	0	0	0	0	2	2
19	2	Bayern	hcut	0	0	0	0	0	0	0	0	0	0	0	0
20	2	Bayern	ess_shps	0	0	0	0	0	0	0	0	0	0	1	1
21	2	Bayern	zoo	0	0	0	0	0	0	0	0	0	2	2	2
22	2	Bayern	demo	0	0	0	1	1	1	1	1	1	2	2	2
23	2	Bayern	school	0	0	0	0	0	0	0	0	1	1	1	1
24	2	Bayern	church	0	0	0	0	0	0	0	0	0	2	2	2
25	2	Bayern	onefriend	0	0	0	0	0	0	0	0	0	0	0	0
26	2	Bayern	morefriends	0	0	0	0	0	0	0	0	0	0	0	0
27	2	Bayern	plygrnd	0	0	0	0	0	0	0	0	0	2	2	2
28	2	Bayern	daycare	0	0	0	0	0	0	0	0	0	1	1	1
29	3	Berlin	leavehome	0	0	0	0	0	0	0	0	0	0	0	0

Steinmetz, H., Batzdorfer, V., & Bosnjak, M. (June, 2020). The ZPID lockdown measures dataset. ZPID Science Information Online, 20(1). http://dx.doi.org/10.23668/psycharchives.3019

Visualisierung und Daten einer Partikel-Kollision



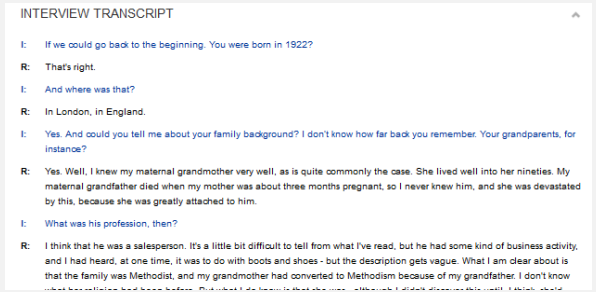
McCauley, Thomas; (2014). Event display file derived from /Commissioning/Run2010B-Apr21ReReco-v1/AOD. CERN Open Data Portal. DOI:10.7483/OPENDATA.CMS.ZUH9.ZSZW

Verwendet einen Datenauszug aus:

CMS collaboration (2014). Commissioning primary dataset in AOD format from RunB of 2010 (/Commissioning/Run2010B-Apr21ReReco-v1/AOD). CERN Open Data Portal. DOI:10.7483/OPENDATA.CMS.VTD5.V7Y3

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?

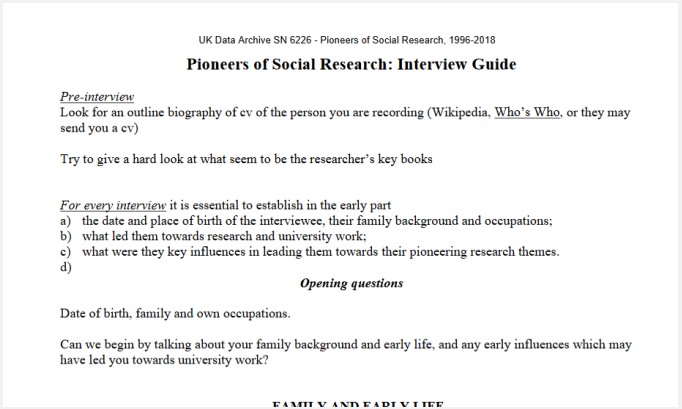
R: In London, in England.

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


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

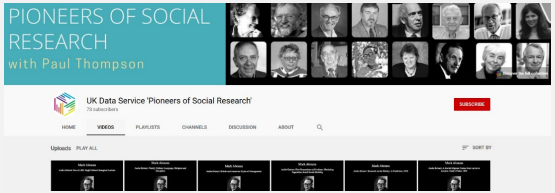
For every interview it is essential to establish in the early part
a) the date and place of birth of the interviewee, their family background and occupations;
b) what led them towards research and university work;
c) what were they key influences in leading them towards their pioneering research themes.
d)

Opening questions

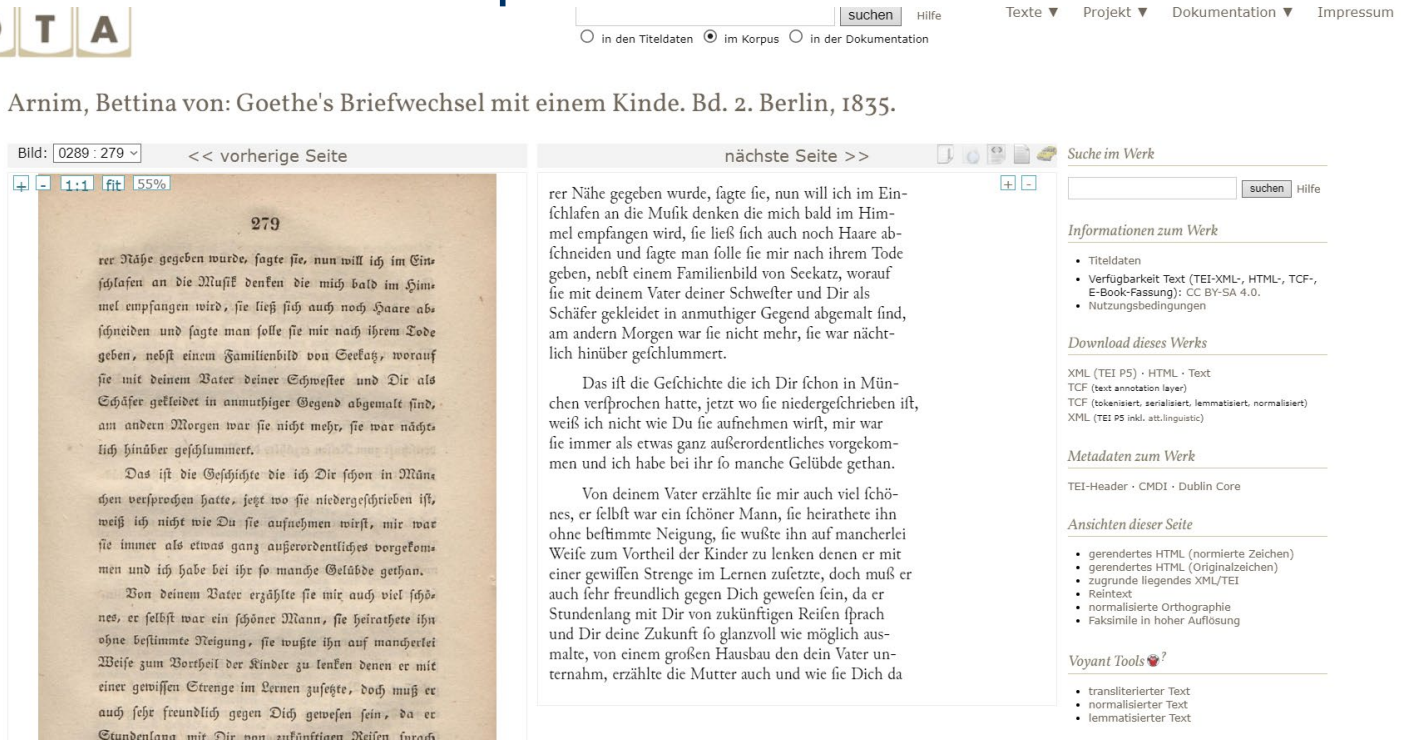
Date of birth, family and own occupations.

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

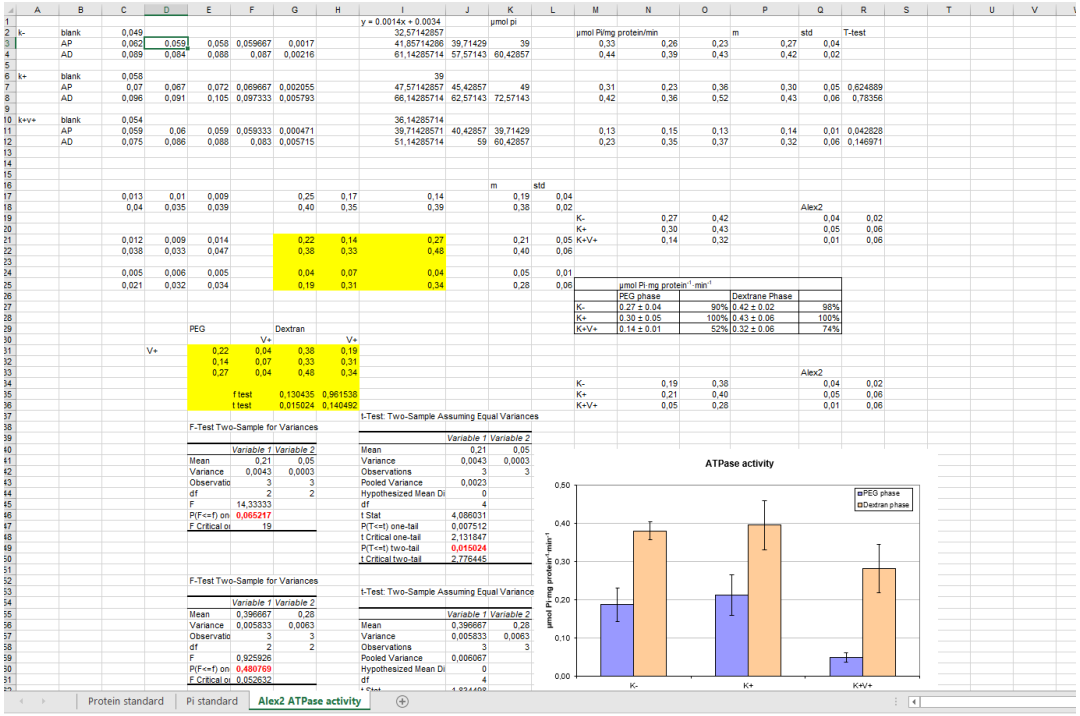


Maschinenlesbarer Korpus: Deutsches Textarchiv



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.

Messungen und Analyse von lytischen Enzymen



Ma, Haiyan; Krock, Bernd; Tillmann, Urban; Bickmeyer, Ulf; Graeve, Martin; Cembella, Allan (2011): Files of figures 2, 3 and table. PANGAEA, https://doi.org/10.1594/PANGAEA.769889

Supplement to: Ma, H et al. (2011): Mode of action of membrane-disruptive lytic compounds from the marine dinoflagellate Alexandrium tamarense. Toxicon, 58(3), 247-258, https://doi.org/10.1016/j.toxicon.2011.06.004

Daten aus der Verwaltung

GOVDATA
Das Datenportal für Deutschland

https://www.govdata.de/
Letzter Zugriff 3.9.2021

Stöbern Sie in diesen Kategorien

- Bevölkerung und Gesellschaft (10954)
- Bildung, Kultur und Sport (4687)
- Energie (4630)
- Gesundheit (1833)
- Internationale Themen (256)
- Justiz, Rechtssystem und öffentliche Sicherheit (4770)
- Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel (8294)
- Regierung und öffentlicher Sektor (14064)
- Regionen und Städte (10111)
- Umwelt (14402)
- Verkehr (8984)
- Wirtschaft und Finanzen (9547)
- Wissenschaft und Technologie (5192)

Was ist Forschungsdatenmanagement?

Forschungsdatenmanagement begleitet den gesamten Forschungsprozess und umfasst

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

von Daten.

Datenlebenszyklus

Quelle der Grafik:
https://www.forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/ zuletzt Besucht 03.09.2021

VORTEILE, VORGABEN UND KONTEXT

Einige Vorteile von FDM

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

Mögliche Vorgaben und Leitlinien

- **Institutionelle Policies**
- Vorgaben von **Förderinstitutionen**
  - Deutsche Forschungsgemeinschaft (DFG): Sicherung guter wissenschaftlicher Praxis (1998, 2013, 2019), Leitlinie zum Umgang mit Forschungsdaten (2015)
  - Europäische Kommission: Horizon 2020 / Open Research Data Pilot (in englischer Sprache)
- Vorgaben aus der **Fachdisziplin**
  - Fachspezifische Empfehlungen
- 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 (2019) Leitlinien zur Sicherung guter wissenschaftlicher Praxis. https://doi.org/10.5281/zenodo.3923602

7 Phasenübergreifende Qualitätssicherung

„Art und Umfang von im Forschungsprozess entstehenden Forschungsdaten werden beschrieben. Der Umgang mit ihnen wird, entsprechend den Vorgaben im betroffenen Fach, ausgestaltet. Der Quellcode von öffentlich zugänglicher Software muss persistent und dokumentiert sein.“ (Leitlinie S. 14-15)

Deutsche Forschungsgemeinschaft (2019) Leitlinien zur Sicherung guter wissenschaftlicher Praxis. https://doi.org/10.5281/zenodo.3923602

13 Herstellung von öffentlichem Zugang zu Forschungsergebnissen

Freie Universität Berlin

„[…] soweit dies möglich und zumutbar ist, die den Ergebnissen zugrunde liegenden Forschungsdaten, Materialien und Informationen, die angewandten Methoden sowie die eingesetzte Software verfügbar zu machen [...].“ (Leitlinie 13, S. 19)

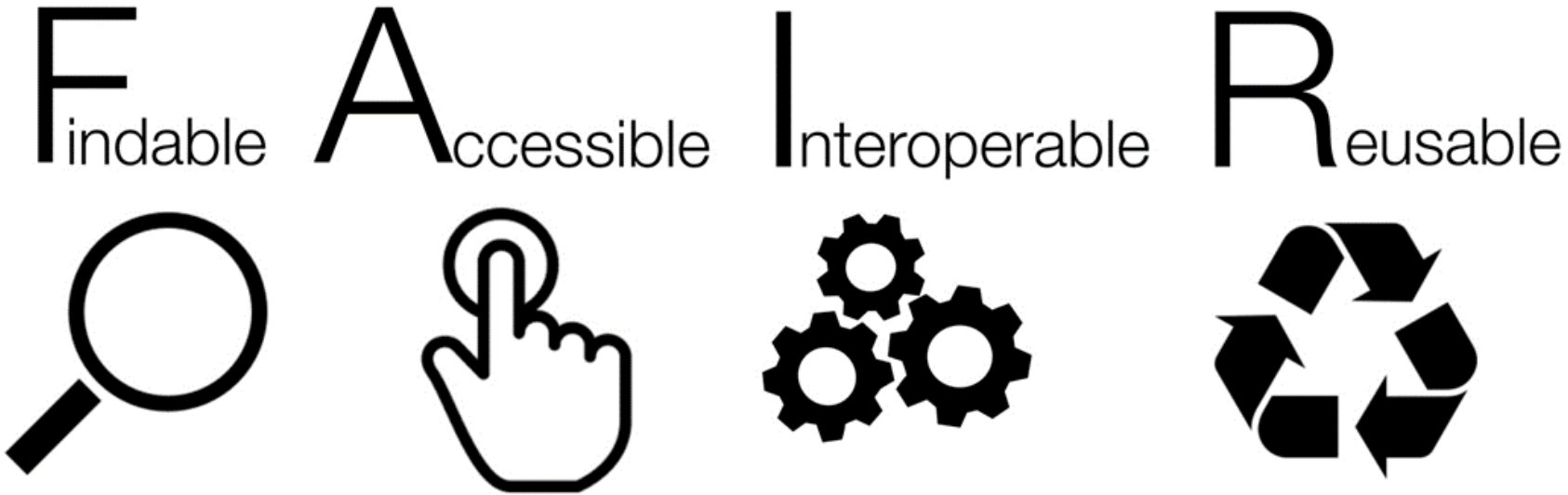
Deutsche Forschungsgemeinschaft (2019) Leitlinien zur Sicherung guter wissenschaftlicher Praxis. https://doi.org/10.5281/zenodo.3923602

17 Archivierung

„[…] werden die zugrunde liegenden Forschungsdate […] in der Regel für einen Zeitraum von zehn Jahren zugänglich und nachvollziehbar in der Einrichtung, wo sie entstanden sind, oder in standortübergreifenden Repositorien aufbewahrt.“ (Leitlinie 17, S. 22)

Deutsche Forschungsgemeinschaft (2019) Leitlinien zur Sicherung guter wissenschaftlicher Praxis. https://doi.org/10.5281/zenodo.3923602

Die FAIR-Prinzipien



Quelle: Pundir, Sangya. *https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg* [letzter Zugriff: 27.09.2018], CC-BY-SA-4.0 https://creativecommons.org/licenses/by-sa/4.0/deed.en

FORSCHUNGSDATEN-MANAGEMENT IN DER PRAXIS

Aspekte von Forschungsdatenmanagement

PLANEN

ORGANISIEREN UND DOKUMENTIEREN

ETHIK, SICHERHEIT UND RECHT

SPEICHERN

ARCHIVIEREN

TEILEN UND PUBLIZIEREN

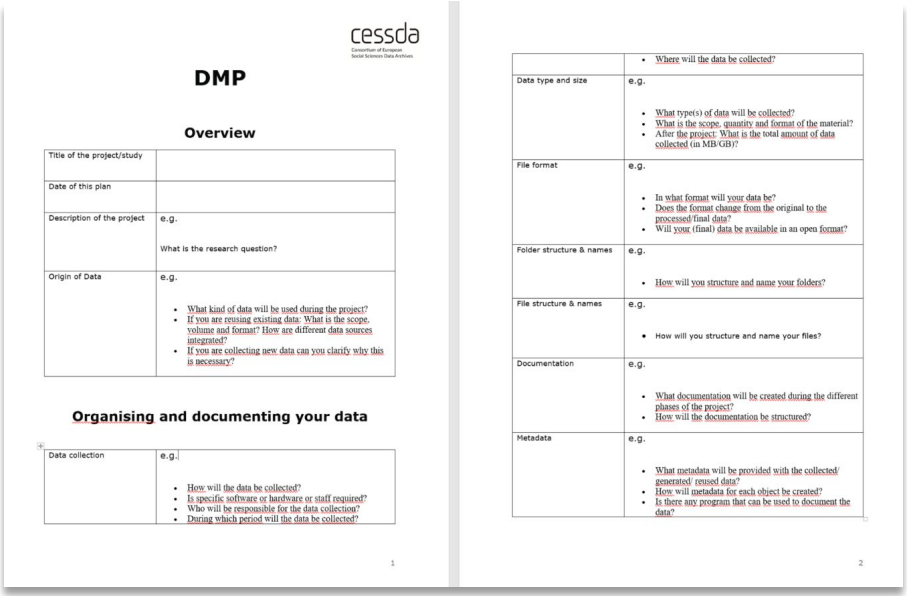
ZITIEREN UND NUTZEN

Datenmanagementplan

- Strukturierte Planung und Dokumentation aller Fragen rund um das FDM für ein Projekt
- über den gesamten Projektverlauf sowie darüber hinaus
- Inhalt: Informationen über Datenbestand, Datenerstellung, Workflows, Verarbeitung, Archivierung und Veröffentlichung
- Schlüsselinstrument für nachnutzbare Daten, z.B. für Anschluss- und /oder Folgeforschung
- „living document“: regelmäßige Ergänzung
- Wird zunehmend im Kontext von Förderanträgen verlangt
- Tools: RDMO, DMPonline

Bestandteile eines DMP

- Administrative Informationen (Projektname, Datenurheber*in, weitere Mitwirkende, Kontakt, Förderprogramm usw.)
- Projekt-und Datensatzbeschreibung
- Datentypen, -formate, -umfang
- Methoden
- Angaben zu Metadaten und Standards
- Archivierung und Sicherung der Daten
- Veröffentlichung der Daten
- Verantwortlichkeiten
- Kosten



cessda
Consortium of European Social Sciences 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. • 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. • 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?

1

	• Where will the data be collected?
Data type and size	e.g. • 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. • 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. • How will you structure and name your folders?
File structure & names	e.g. • How will you structure and name your files?
Documentation	e.g. • What documentation will be created during the different phases of the project? • How will the documentation be structured?
Metadata	e.g. • What metadata will be provided with the collected/ generated/ reused data? • How will metadata for each object be created? • Is there any program that can be used to document the data?

2

[Muster und Templates von CESSDA](#) (Screenshot)
[Muster und Templates der HU](#)

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

Backup und Datensicherheit

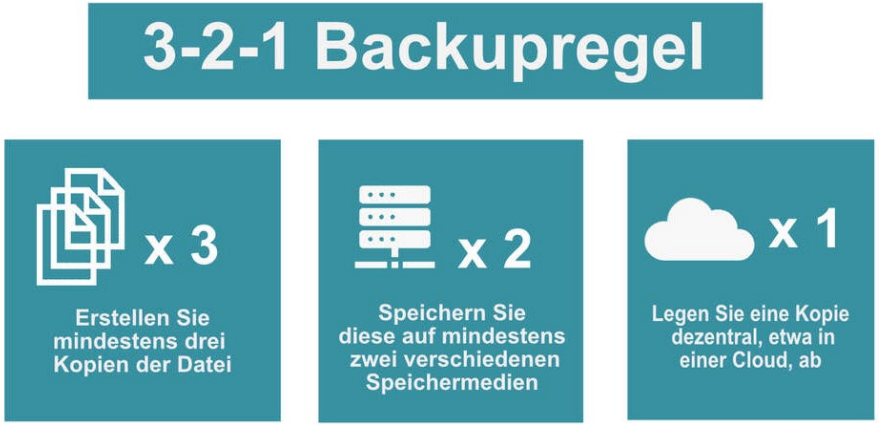


Abb: 3-2-1 Backupregel - I. Lang/Bearbeitung E. Böker / CC BY 4.0

- Regelmäßige Backups zu einem festgelegten Zeitpunkt durchführen
- Sicherheitskopien an einem anderen Ort als die Originale aufbewahren
- Systeme der Universität verwenden
- Vor unbefugten Zugriff schützen, z.B. mit Passwortschutz

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

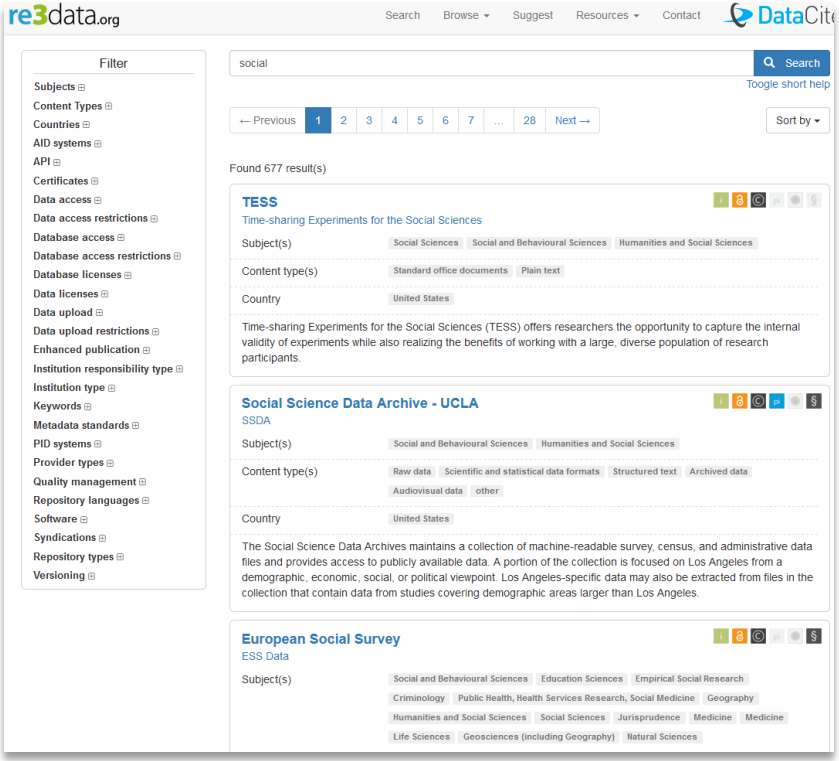
Forschungsdaten publizieren

Als eigenständige Publikation in Forschungsdaten-Repositorien:

- fachspezifische Repositorien, z.B. PANGAEA
- generische Repositorien, z. B. ZENODO
- institutionelle Repositorien, 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



Checkliste Veröffentlichung

- Möglichst passender Ort für die Daten (Datentyp und Fachgebiet)
- Vertrauenswürdige Anbieter, Nachhaltigkeit prüfen
- Vergabe eines persistenten Identifikators: langfristige Identifikation und Referenzierung
- Klare Nutzungsbedingungen für Datenautor*innen sowie Nutzer*innen
  - z. B. Lizenzen, Entgelte, Embargofristen
- Möglichst offene Formate und Lizenzen
- Daten ausreichend beschreiben, damit sie für Nachnutzende auffindbar und verständlich sind
  - Benennung, Metadaten und Dokumentation

Forschungsdaten finden

- direkt in Fachrepositorien und fachübergreifenden Repositorien (via re3data)
- Datenjournale (Liste)
- mittels Metasuchmaschinen z. B.
  - B2FIND http://b2find.eudat.eu
  - BASE https://www.base-search.net/Search/Advanced
  - CrossRef https://search.crossref.org/
  - DataCite Metadata https://search.datacite.org/
  - OpenAire https://www.openaire.eu/search/find

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

Weiterführende Informationen

- Forschungsdaten.info https://www.forschungsdaten.info/

- MANTRA Research Data Management Training
  https://mantra.edina.ac.uk/

Danke für Ihre Aufmerksamkeit!

Esther Asef
Referentin für Forschungsdatenmanagement
Freie Universität Berlin
Universitätsbibliothek

E-Mail: forschungsdaten@fu-berlin.de
Web: https://www.fu-berlin.de/forschungsdatenmanagement

Dieses Werk ist unter einer Creative Commons Namensnennung 4.0 International Lizenz 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 oder Lizenz https://support.microsoft.com/de-de/office/einfügen-von-piktogrammen-in-microsoft-office-e2459f17-3996-4795-996e-b9a13486fa79