Love Data Week 2025 / University of Cologne

Aber ich habe doch gar keine Daten!

Felix Rau
Data Center for the Humanities

DCH

Universität zu Köln

Forschungsdatenmanagement für Menschen ohne Daten

DCH

Universität zu Köln

Keine Daten.
Was heißt das?

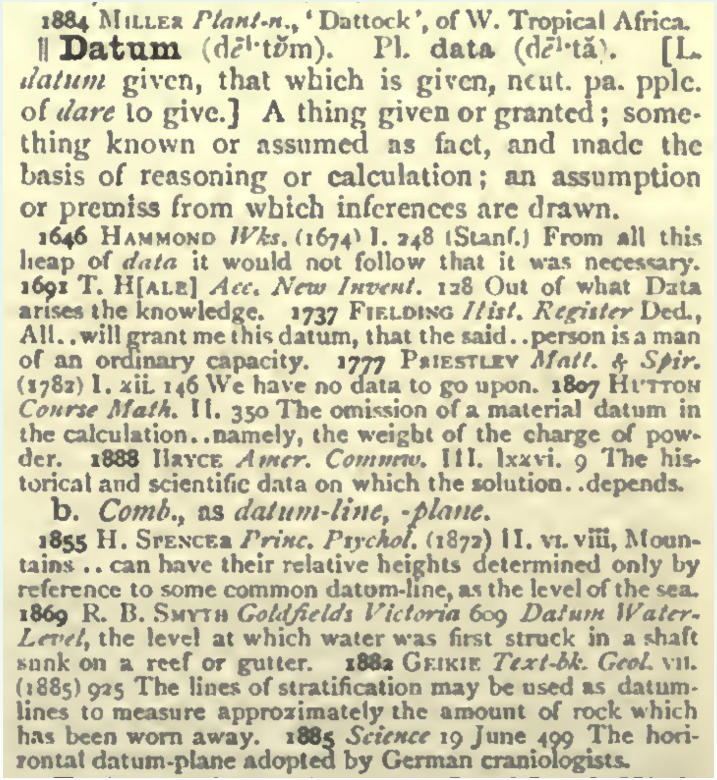
DCH

Universität zu Köln

Daten

Data / Daten

Pluralform des
Partizip Perfekt Passiv von
lat. *dare* »geben«
(sg. *datum*, pl. *data*)



1884 MILLER *Plant-n.*, 'Dattock', of W. Tropical Africa.

‖ **Datum** (dē'·tŭm). Pl. **data** (dē'·tă). [L. *datum* given, that which is given, neut. pa. pple. of *dare* to give.] A thing given or granted; something known or assumed as fact, and made the basis of reasoning or calculation; an assumption or premiss from which inferences are drawn.

1646 HAMMOND *Wks.* (1674) I. 248 (Stanf.) From all this heap of *data* it would not follow that it was necessary. 1691 T. H[ALE] *Acc. New Invent.* 128 Out of what Data arises the knowledge. 1737 FIELDING *Hist. Register* Ded., All..will grant me this datum, that the said..person is a man of an ordinary capacity. 1777 PRIESTLEY *Matt. & Spir.* (1782) I. xii. 146 We have no data to go upon. 1807 HUTTON *Course Math.* II. 350 The omission of a material datum in the calculation..namely, the weight of the charge of powder. 1888 HAYCE *Amer. Commw.* III. lxxvi. 9 The historical and scientific data on which the solution..depends.

b. *Comb.*, as *datum-line, -plane.*

1855 H. SPENCER *Princ. Psychol.* (1872) II. vi. viii, Mountains..can have their relative heights determined only by reference to some common datum-line, as the level of the sea. 1869 R. B. SMYTH *Goldfields Victoria* 609 *Datum Water-Level*, the level at which water was first struck in a shaft sunk on a reef or gutter. 1882 GEIKIE *Text-bk. Geol.* VII. (1885) 925 The lines of stratification may be used as datum-lines to measure approximately the amount of rock which has been worn away. 1885 *Science* 19 June 499 The horizontal datum-plane adopted by German craniologists.

Quelle: Oxford English Dictionary

DCH

Universität zu Köln

Eine hilfreiche Definition

Leonelli 2016 defines data

as any product of research activities, ranging from artifacts such as photographs to symbols such as letters or numbers, that is collected, stored, and disseminated *in order to be used as evidence for knowledge claims*.

Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago ; London: The University of Chicago Press.

Konsequenzen

Warum sagen wir dann, dass wir keine Daten haben?

Beispiele

1. Im Fach heißen sie anders.
2. Sie haben gar keinen Namen.
3. Es gibt keinen Daten-Output.
4. Der Gegenstand ist nicht digital.

Daten = Evidenz für Wissensaussagen

- Ich werde die Evidenz für Wissensaussagen methoden- und fachunabhängig *Daten* nennen.

- Was Daten sind und welche Rolle sie im Forschungsprozess spielen, ist methoden- und fachabhängig.

- Entsprechend ist Datenmanagement in weiten Teilen abhängig von Methode und Fach.

FDM – ein paar Thesen

1. Datenmanagement findet in jedem Forschungsprozess statt, egal ob man es als solches konzeptualisiert.
2. Datenmanagement wird in jedem Schritt des Forschungsprozesses gemacht und ist nicht beschränkt auf einen bestimmten Schritt.
3. Datenmanagement ist in weiten Teilen trivial, muss aber trotzdem – und um so mehr – gut umgesetzt werden.
4. Gute Datenmanagementpraktiken lassen sich auf angrenzende Bereiche wie Bibliographiemanagement, Notizenmanagement und Personal Knowledge Management übertragen.

DCH

Universität zu Köln

FDM im Forschungsprozess

DCH

Universität zu Köln

FDM im Forschungsprozess

1. erwerben
2. lagern
3. handhaben
4. ablegen

Rechtliches
Ethisches

DCH

Universität zu Köln

Grundlegendes FDM

DCH

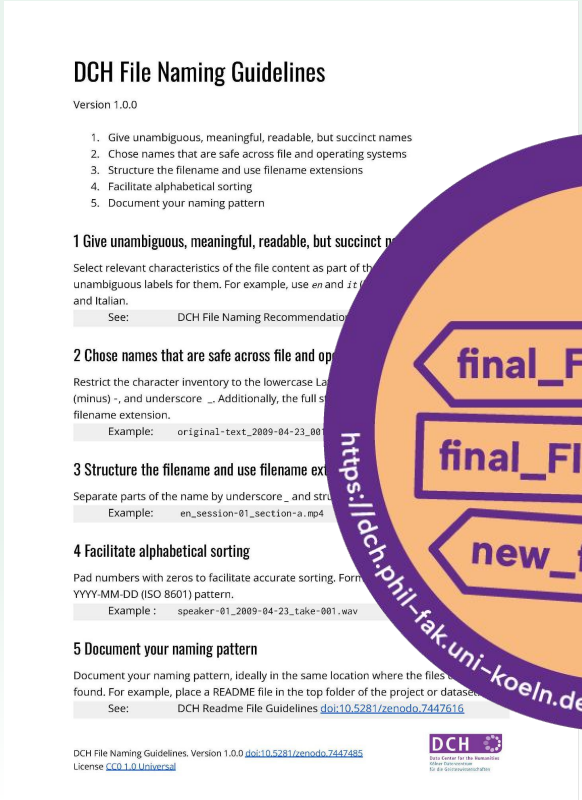
Universität zu Köln

Dateibenennung

Dateinamen

```
Paper_final_final(2).docx
Untitled.pdf
Einfl?sse.xslx

Original-text_2009-04-23_001.mp4
```



https://doi.org/10.5281/zenodo.7447485

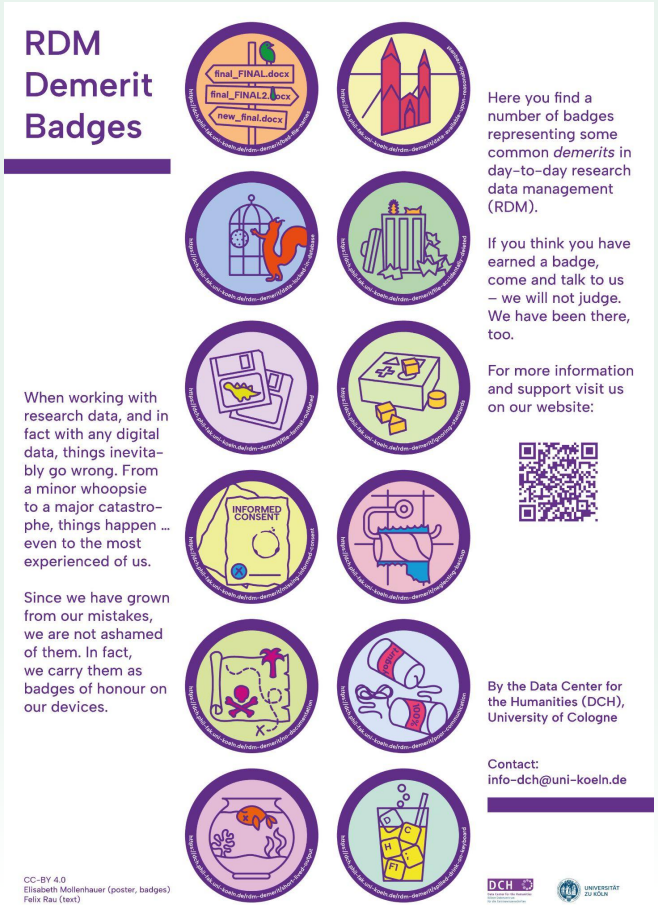
DCH

Universität zu Köln

Werbung

Aufkleber z.B. für den Laptop mit gängigen FDM Fehlern.

Ziel:
Aus Fehlern lernen,
zu Fehlern stehen,
über Fehler reden
https://dch.phil-fak.uni-koeln.de/rdm-demerit



DCH

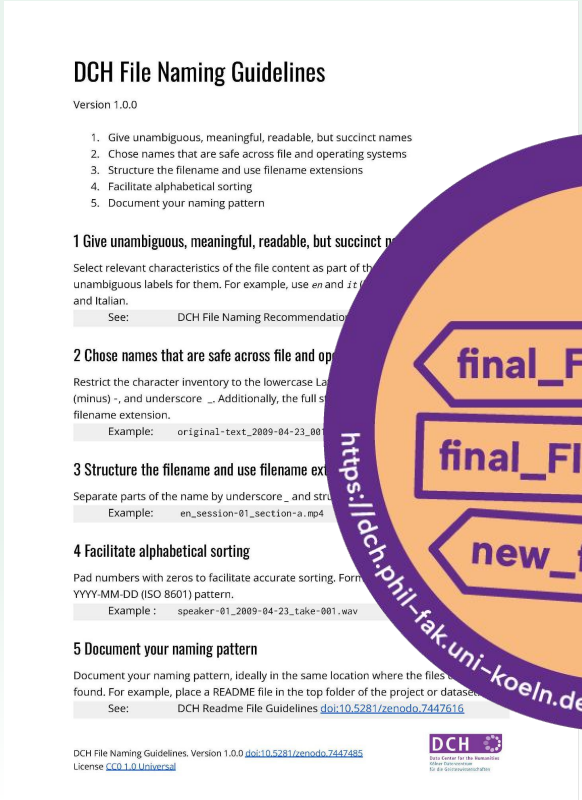
Universität zu Köln

Dateibenennung

Dateinamen

```
Paper_final_final(2).docx
Untitled.pdf
Einfl?sse.xslx

Original-text_2009-04-23_001.mp4
```



https://doi.org/10.5281/zenodo.7447485

DCH

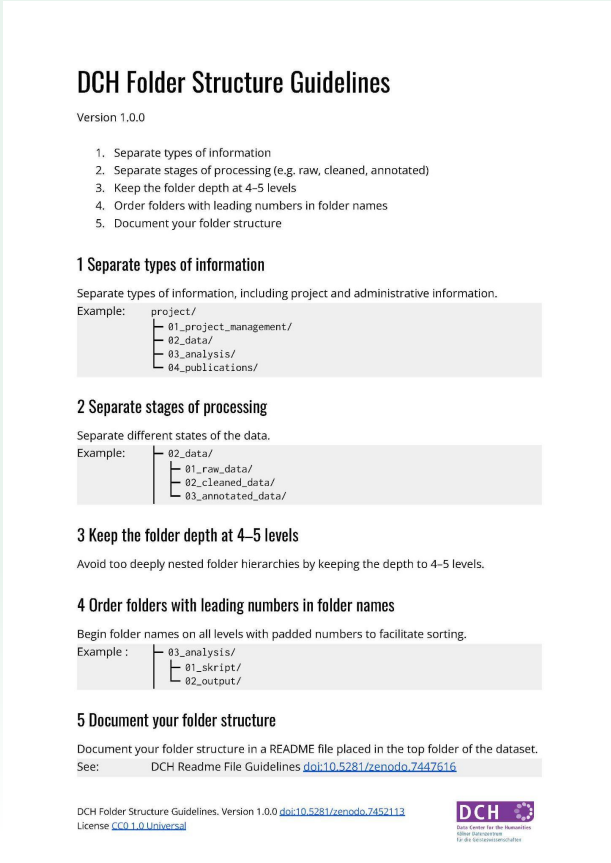
Universität zu Köln

Ordnerstruktur

Wiederkehrende simple Ordnerstrukturen schaffen Überblick

```
project/
    ├─ 01_project_management/
    ├─ 02_data/
    ├─ 03_analysis/
    └─ 04_publications/
```

https://doi.org/10.5281/zenodo.7452113



DCH Folder Structure Guidelines

Version 1.0.0

1. Separate types of information
2. Separate stages of processing (e.g. raw, cleaned, annotated)
3. Keep the folder depth at 4–5 levels
4. Order folders with leading numbers in folder names
5. Document your folder structure

1 Separate types of information

Separate types of information, including project and administrative information.

Example:
```
project/
├─ 01_project_management/
├─ 02_data/
├─ 03_analysis/
└─ 04_publications/
```

2 Separate stages of processing

Separate different states of the data.

Example:
```
├─ 02_data/
│  ├─ 01_raw_data/
│  ├─ 02_cleaned_data/
│  └─ 03_annotated_data/
```

3 Keep the folder depth at 4–5 levels

Avoid too deeply nested folder hierarchies by keeping the depth to 4–5 levels.

4 Order folders with leading numbers in folder names

Begin folder names on all levels with padded numbers to facilitate sorting.

Example :
```
├─ 03_analysis/
│  ├─ 01_skript/
│  └─ 02_output/
```

5 Document your folder structure

Document your folder structure in a README file placed in the top folder of the dataset.

See: DCH Readme File Guidelines doi:10.5281/zenodo.7447616

DCH Folder Structure Guidelines. Version 1.0.0 doi:10.5281/zenodo.7452113
License CC0 1.0 Universal

DCH

DCH

Universität zu Köln

Backup

Three things are certain:

Death, taxes, and lost data.

Guess which has occurred.

David Dixon

Backup bedeutet:

- Automatisch
- Bewusster Zyklus und Auswahl
- Duplizieren (3-2-1)
  - 3 Instanzen
  - 2 verschiedene Medien
  - 1 Off-site Kopie
- Wiederherstellung testen

https://doi.org/10.5281/zenodo.7760967

DCH

Universität zu Köln

[Backup-Workshop für Studierende (Mac&PC)]

10.02.2025, 12-13:30 Uhr
11.02.2025, 10-11:30 Uhr
13.02.2025, 12-13:30 Uhr

Universitäts- und Stadtbibliothek, Raum 4006

DCH

Universität zu Köln

Mangelnde Dokumentation

Mangelnde Dokumentation führt zu

- Unbeabsichtigter Varianz im Arbeitsablauf
- Unklarer Semantik in Daten und Strukturen
- Misverständnissen und Misattributionen

und macht

- Mitarbeiter:innen
- Dem zukünftigen ich

das Leben schwer.

Hier hilft schon eine kleine Textdatei (README) im Ordner. Eine ausgiebige Dokumentation mit wissenschaftlichen Metadaten kann notwendig sein.

DCH

Universität zu Köln

Eingeschlossene Daten

Daten, Dokumente und Notizen sollten wo immer möglich unabhängig von einer spezifischen Software lesbar und nutzbar sein:

- Notiz-Apps wie Notion oder Evernote
- Filemaker-Datenbank
- Webseite mit Datenbank zu Projektende

DCH

Universität zu Köln

Webseite mit Datenbank zur Ergebnispräsentation

Digitale Präsentation muss gut geplant und in den Projektablauf eingebettet sein. Die Kompetenz muss nachhaltig vorhanden sein.

- Daten- und Projektergebnisse sollten ohne spezielle Präsentation verständlich und nutzbar sein
- Software und Datenbanken brauchen Pflege
- Pflege muss nach Projektende gewährleistet sein

DCH

Universität zu Köln

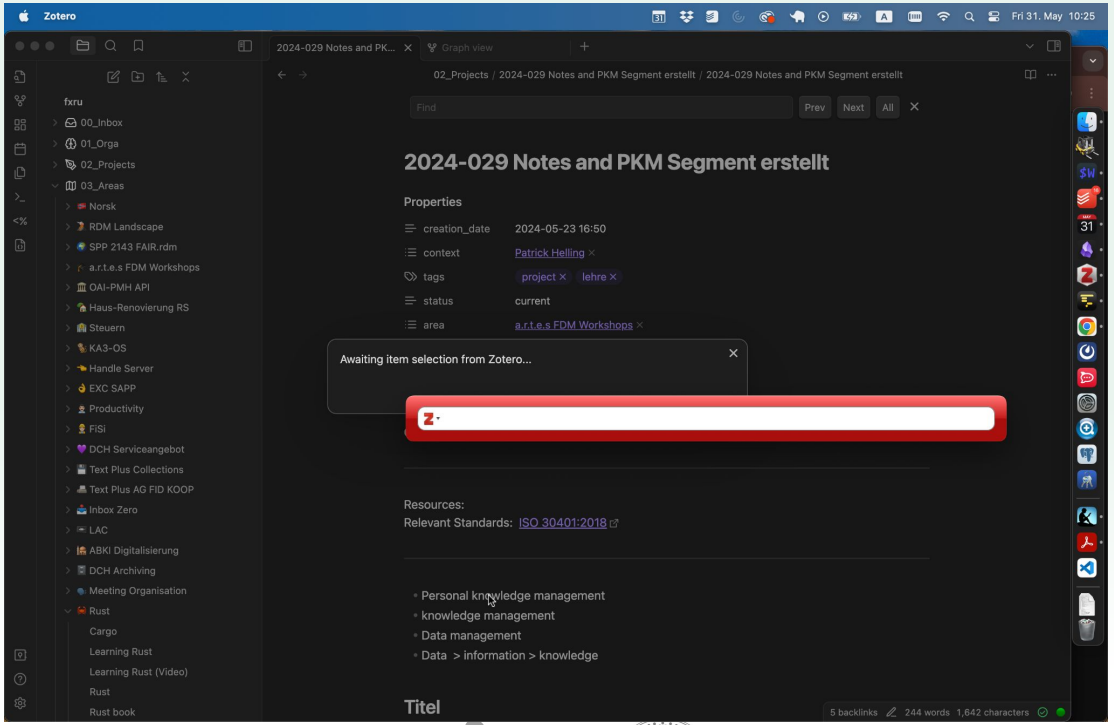
Jenseits des FDM

DCH

Universität zu Köln

Notizen und Personal Knowledge Management

- Workflows
- Digitale Objekte
- Dateibenennung
- Ordnerstruktur
- Metadaten
- READMEs

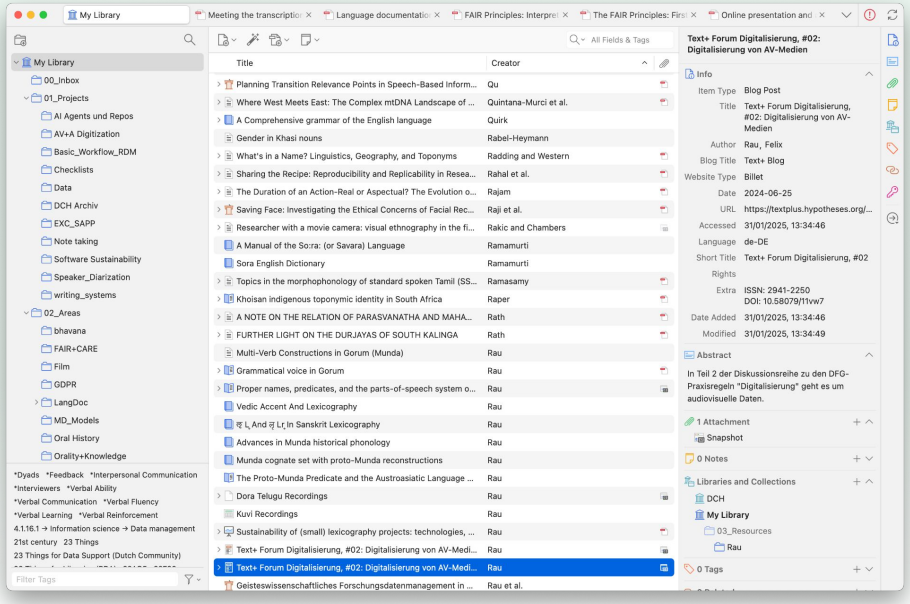


DCH

Universität zu Köln

Bibliographiemanagement

- Standardisierte, offene Formate
- Metadaten (z.B. Tags)
- Organisationsprinzipien



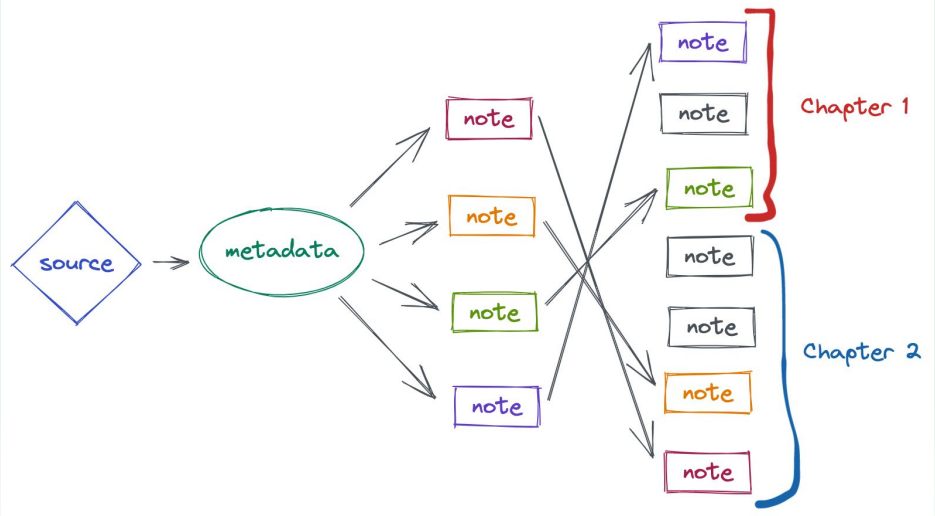
Zotero- und/oder Citavi-Kurse an der USB Köln

DCH

Universität zu Köln

Workflows und Integration

Beispiel-Workflow für die Arbeit mit wissenschaftlichen Artikeln, Literatur oder ähnlichen Dokumenten



PDF > Zotero > Zotero Notes > Obsidian

Razlogova, Elena. n.d. “01 Notetaking for Historians - Doing History with Zotero and Obsidian - Obsidian Publish.” Doing History with Zotero and Obsidian. Accessed October 26, 2023. https://publish.obsidian.md/history-notes/01+Notetaking+for+Historians.

DCH Universität zu Köln

Materialien (Lokal in Köln):

Guidelines:
https://dch.phil-fak.uni-koeln.de/guidelines-fuer-forschende

Badges:
https://dch.phil-fak.uni-koeln.de/rdm-demerit

C3RDM FDM Wissen:
https://fdm.uni-koeln.de/fdm-wissen

Kurse an der USB:
https://ub.uni-koeln.de/kurse-beratung/kurse-der-usb

DCH

Universität zu Köln

Vielen Dank!

Data Center for the Humanities
Kölner Datenzentrum
für die Geisteswissenschaften

Philosophische Fakultät
Universität zu Köln

https://dch.phil-fak.uni-koeln.de/

Bei Fragen zu FDM: info-dch@uni-koeln.de

Guidelines:
https://dch.phil-fak.uni-koeln.de/guidelines-fuer-forschende

Badges:
https://dch.phil-fak.uni-koeln.de/rdm-demerit

DCH

Universität zu Köln