



Love Data Week 2025 / University of Cologne

# Aber ich habe doch gar keine Daten!

Felix Rau  
Data Center for the Humanities



Universität zu Köln

# Forschungsdatenmanagement für Menschen ohne Daten



Keine Daten.  
Was heißt das?

# 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'tōm). Pl. *data* (dē'tā). [*L. datum* given, that which is given, neut. pa. pp. 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 *Math. & 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. 1883 HAYCE *Amer. Commw.* III. lxxvi. 9 The historical and scientific *data* on which the solution. .depends.  
b. *Comb.*, as *datum-line*, *-plane*.  
1855 H. SPENCEA *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](#)

# 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.

# FDM im Forschungsprozess

# FDM im Forschungsprozess

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

Rechtliches  
Ethisches

# Grundlegendes FDM

# Dateibenennung

## Dateinamen

Paper\_final\_final(2).docx

Untitled.pdf

Einflussse.xlsx

Original-text\_2009-04-23\_001.mp4

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

## DCH File Naming Guidelines

Version 1.0.0

1. Give unambiguous, meaningful, readable, but succinct names
2. Chose names that are safe across file and operating systems
3. Structure the filename and use filename extensions
4. Facilitate alphabetical sorting
5. Document your naming pattern

### 1 Give unambiguous, meaningful, readable, but succinct names

Select relevant characteristics of the file content as part of the unambiguous labels for them. For example, use *en* and *it* for English and Italian.

See: [DCH File Naming Recommendation](#)

### 2 Chose names that are safe across file and operating systems

Restrict the character inventory to the lowercase Latin alphabet (minus *-*, and underscore *\_*). Additionally, the full system filename extension.

Example: `original-text_2009-04-23_001.mp4`

### 3 Structure the filename and use filename extensions

Separate parts of the name by underscore *\_*, and structure the name.

Example: `en_session-01_section-a.mp4`

### 4 Facilitate alphabetical sorting

Pad numbers with zeros to facilitate accurate sorting. Form YYYY-MM-DD (ISO 8601) pattern.

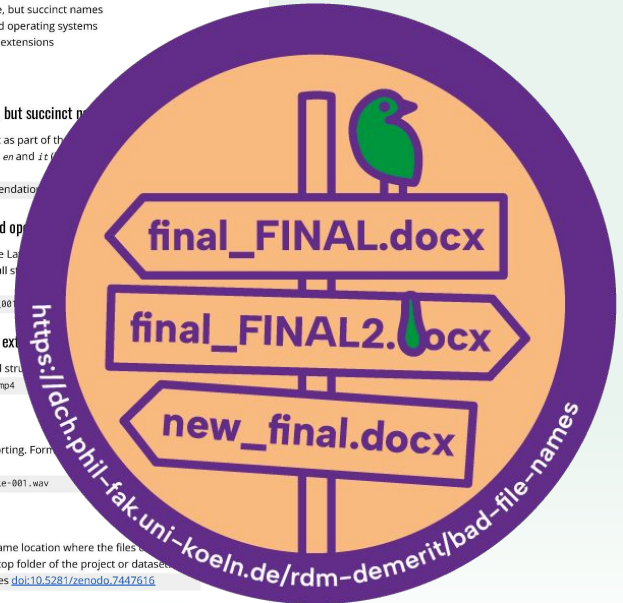
Example: `speaker-01_2009-04-23_take-001.wav`

### 5 Document your naming pattern

Document your naming pattern, ideally in the same location where the files are found. For example, place a README file in the top folder of the project or dataset.

See: [DCH Readme File Guidelines doi:10.5281/zenodo.7447616](#)

DCH File Naming Guidelines, Version 1.0.0 [doi:10.5281/zenodo.7447485](https://doi.org/10.5281/zenodo.7447485)  
License [CC 0 1.0 Universal](#)



# 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>



## RDM Demerit Badges



Here you find a number of badges representing some common *demerits* in day-to-day research data management (RDM).

If you think you have earned a badge, come and talk to us – we will not judge. We have been there, too.

For more information and support visit us on our website:



When working with research data, and in fact with any digital data, things inevitably go wrong. From a minor whoopsie to a major catastrophe, things happen ... even to the most experienced of us.

Since we have grown from our mistakes, we are not ashamed of them. In fact, we carry them as badges of honour on our devices.

By the Data Center for the Humanities (DCH), University of Cologne

Contact:  
info-dch@uni-koeln.de

CC-BY 4.0  
Elisabeth Mollenhauer (poster, badges)  
Felix Riss (text)



DCH



Universität zu Köln



# Dateibenennung

## Dateinamen

Paper\_final\_final(2).docx

Untitled.pdf

Einflussse.xlsx

Original-text\_2009-04-23\_001.mp4

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

## DCH File Naming Guidelines

Version 1.0.0

1. Give unambiguous, meaningful, readable, but succinct names
2. Chose names that are safe across file and operating systems
3. Structure the filename and use filename extensions
4. Facilitate alphabetical sorting
5. Document your naming pattern

### 1 Give unambiguous, meaningful, readable, but succinct names

Select relevant characteristics of the file content as part of the unambiguous labels for them. For example, use *en* and *it* for English and Italian.

See: [DCH File Naming Recommendation](#)

### 2 Chose names that are safe across file and operating systems

Restrict the character inventory to the lowercase Latin alphabet (minus *-*, and underscore *\_*). Additionally, the full system filename extension.

Example: `original-text_2009-04-23_001.mp4`

### 3 Structure the filename and use filename extensions

Separate parts of the name by underscore *\_*, and structure the filename extension.

Example: `en_session-01_section-a.mp4`

### 4 Facilitate alphabetical sorting

Pad numbers with zeros to facilitate accurate sorting. Form YYYY-MM-DD (ISO 8601) pattern.

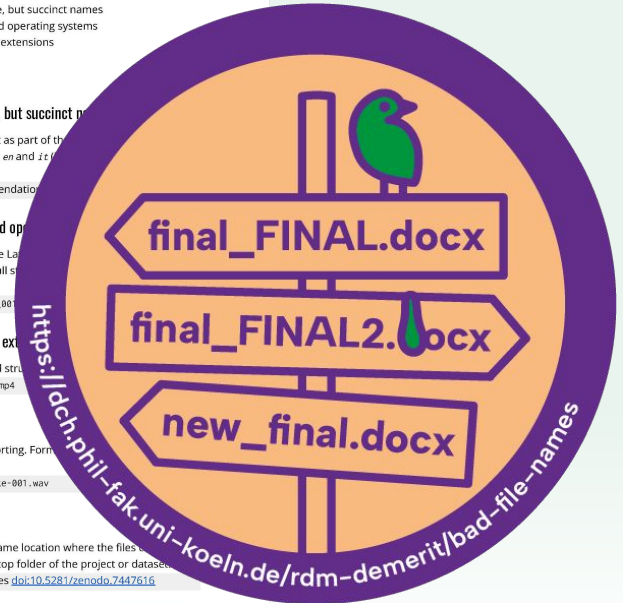
Example: `speaker-01_2009-04-23_take-001.wav`

### 5 Document your naming pattern

Document your naming pattern, ideally in the same location where the files are found. For example, place a README file in the top folder of the project or dataset.

See: [DCH Readme File Guidelines doi:10.5281/zenodo.7447616](#)

DCH File Naming Guidelines, Version 1.0.0 [doi:10.5281/zenodo.7447485](https://doi.org/10.5281/zenodo.7447485)  
License [CC 0.1.0 Universal](#)



# 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](https://doi.org/10.5281/zenodo.7447616)

DCH Folder Structure Guidelines, Version 1.0.0 [doi:10.5281/zenodo.7452113](https://doi.org/10.5281/zenodo.7452113)  
License: [CC0 1.0 Universal](https://creativecommons.org/licenses/by/4.0/)



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 Project Data Storage Guidelines

Version 1.0.0

1. Keep three instances
2. Keep two local instances on different media
3. Keep one off-site copy
4. Automate data replication and backup
5. Document procedure and responsibilities

### 1 Three instances

Keep three instances of the data. Best practice is to keep two local instances and one off-site copy of your project data.

### 2 Two local instances on different media

Keep two local instances of the data on different media. One should be the working instance on the SSD or hard drive, the other on a different medium, a local NAS, or a network drive. Example at the University of Colorado: [https://www.colorado.edu/it/infrastructure/backup](#)

### 3 One off-site copy

At least one instance of the data should be stored off-site. This should not affect all copies of the data. Example at the University of Colorado: [https://www.colorado.edu/it/infrastructure/backup](#)

### 4 Automate data replication and backup

The process of regular local replication and backup should be automated. The frequency of replication and backup must be sufficient to minimize the amount of data loss in case of an incident. Example solutions: [https://www.ubuntu.com/server/docs/backup](#), [https://www.backintime.org/](#)

### 5 Document procedure and responsibilities

Document the data storage procedure – including location of data, frequency of replication and backup – as well as the responsibilities for the procedure e.g. in your data management plan.

DCH Project Data Storage Guidelines. Version 1.0.0 [doi:10.5281/zenodo.7760967](https://doi.org/10.5281/zenodo.7760967)  
License [CC0 1.0 Universal](#)





## [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

# Mangelnde Dokumentation

Mangelnde Dokumentation führt zu

- Unbeabsichtigter Varianz im Arbeitsablauf
- Unklarer Semantik in Daten und Strukturen
- Missverständnissen 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 README File Guidelines

Version 1.0.0

1. Keep it simple
2. Keep the README file with the dataset
3. Describe the dataset and its context
4. Document the internal structure of the dataset
5. Document file naming structure (including used

### 1 Keep it simple

The README file should be written as plain text and concise structure (e.g. this [template](#)). To separate the profile into the README file, use a separate

### 2 Keep the README file with the dataset

The README file should be stored next to the dataset

### 3 Describe the dataset and its context

The README file should contain a description of the dataset was collected. The description may include any other information relevant for

### 4 Document the internal structure of the dataset

The README file should contain a description of the internal structure of the dataset should be clear which information belongs to which file

See: [DCH Folder Structure Guidelines](#)

### 5 Document file naming pattern

Document the file naming pattern; describe the file naming pattern used for different categories.

See: [DCH File Naming Guidelines](#) [doi:10.5281/zenodo.7447616](https://doi.org/10.5281/zenodo.7447616)

See also: [DCH README File Template 1.0.0](#) [doi:10.5281/zenodo.7452055](https://doi.org/10.5281/zenodo.7452055)

DCH README File Guidelines, Version 1.0.0 [doi:10.5281/zenodo.7447616](https://doi.org/10.5281/zenodo.7447616)  
License [CC0 1.0 Universal](#)



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



# 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

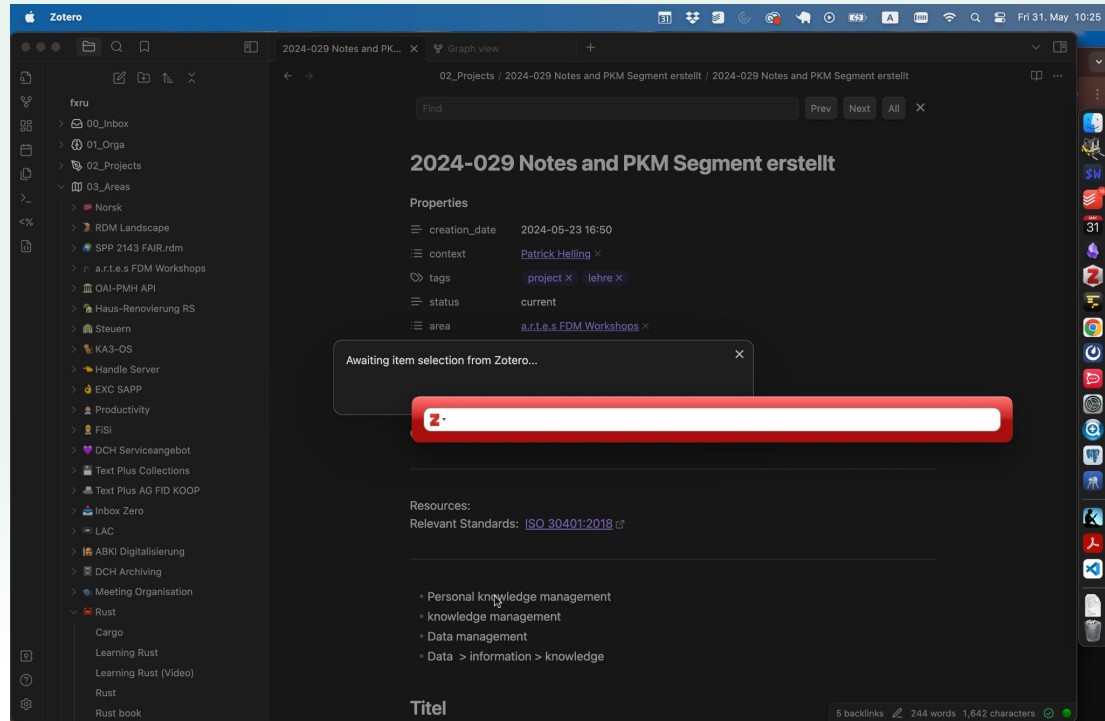


# Jenseits des FDM



# Notizen und Personal Knowledge Management

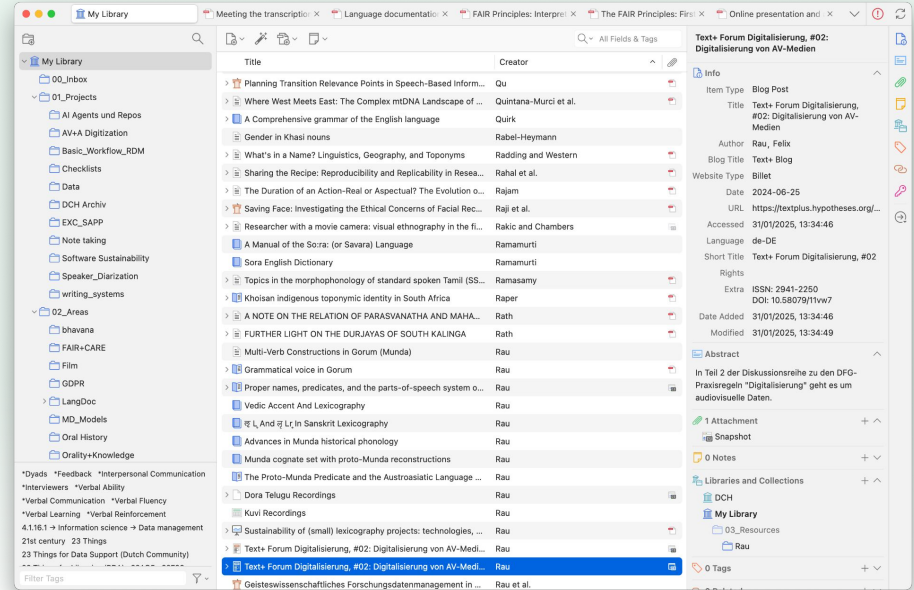
- Workflows
- Digitale Objekte
- Dateibenennung
- Ordnerstruktur
- Metadaten
- READMEs



# Bibliographiemanagement

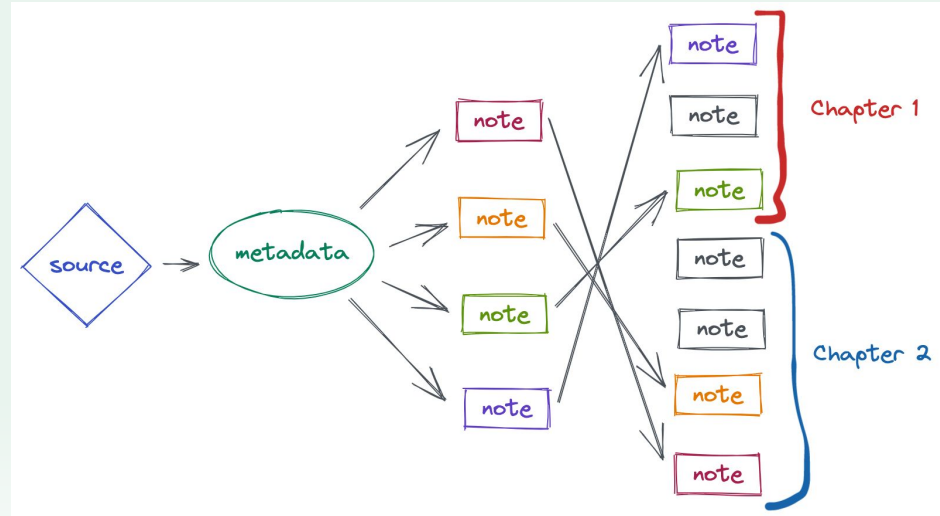
- Standardisierte, offene Formate
- Metadaten (z.B. Tags)
- Organisationsprinzipien

Zotero- und/oder  
Citavi-Kurse an der USB  
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>.

# 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>



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](mailto: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>



Universität zu Köln