

# Open Science: Von Daten zu Publikationen

4.1 Modul: Faire Daten

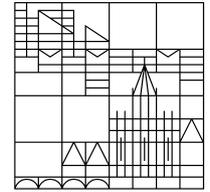
– Vom Märchen zur Wahrheit

<https://www.youtube.com/watch?v=MF07kXV4zJs>

Team Open Science

Kommunikations-, Informations-, Medien-  
Zentrum (KIM), Universität Konstanz



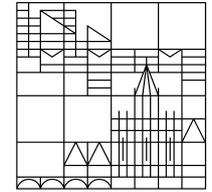


# Maximale Sichtbarkeit für meine Forschung



Konstanz, Mai 2020





# Faire Daten – Vom Märchen zur Wahrheit

Lena Dreher und Dr. Ilona Lang  
Konstanz, Mai 2020



Abbildung: Paulina Halina, CC BY SA 4.0

Alle Inhalte dieser Präsentation stehen, sofern nicht anders angegeben, unter der Lizenz [Creative Commons BY 4.0 International](https://creativecommons.org/licenses/by/4.0/).

# Inhalte

- Was sind **FAIRe Daten** und **Datenmanagement-Pläne**?
- **Wie** erreicht man FAIRe Daten?
- **Tipps** für das eigene (FAIRe) Datenmanagement
- **Angebote des KIM** und weiterführende Links

# Was bedeutet “FAIR” für Daten?

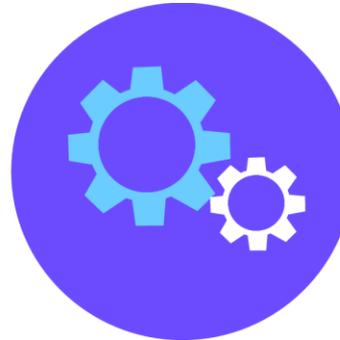
Auffindbar  
(**F**indable)



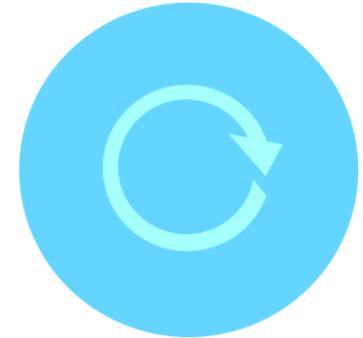
Zugänglich  
(**A**ccessible)



Interoperabel  
(**I**nteroperable)



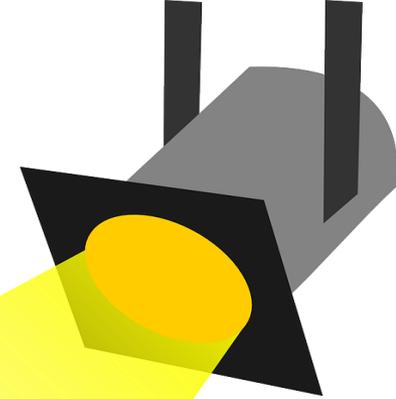
Wiederverwendbar  
(**R**eusable)



Paulina Halina Sieminska. (2019, July 5). A FAIRy tale graphics (Version 1.0.0). Zenodo. <http://doi.org/10.5281/zenodo.3267168> CC BY SA 4.0

# Warum FAIRe Daten nutzen?

- macht Forschung **effizienter** (weniger unbeabsichtigte Dopplungen)
- verbessert **Qualität** durch Transparenz
- erhöht **Sichtbarkeit** (damit höhere Zitationszahlen)
- vereinfacht Entstehung **neuer Forschungsfragen**
- vereinfacht **Kollaborationen**, verbessert Zusammenarbeit in  
Community



# Forderung der EU (Horizon 2020)



([https://ec.europa.eu/research/press/2016/pdf/opendata-infographic\\_072016.pdf#view=fit&pagemode=none](https://ec.europa.eu/research/press/2016/pdf/opendata-infographic_072016.pdf#view=fit&pagemode=none))

## Mein\*e Professor\*in besitzt die Daten – Warum brauche ich FAIR und DMP?

Je nach Förderung (z.B. EU- oder DFG-Gelder) oder Zeitschrift muss sie/er...

- ...die **FAIR-Prinzipien einhalten**
- ...und einen **Datenmanagement-Plan (DMP)** erstellen

# Planung FAIRer Daten: Der Datenmanagementplan (DMP)

## FORSCHUNGSDATEN – OFFEN ALS STANDARD

Projekte brauchen:



Abbildung übersetzt nach: <https://www.openaire.eu/what-is-a-data-management-plan>

# Mein\*e Professor\*in besitzt die Daten

## – Warum brauche ich FAIR und DMP?

Je nach Förderung (z.B. EU- oder DFG-Gelder) oder Zeitschrift muss sie/er...

- ...die FAIR-Prinzipien einhalten
  - ...und einen Datenmanagement-Plan (DMP) erstellen
- der Antrag für Fördermittel und der DMP ist eine sehr gute Lektüre für die **Einarbeitung** in ein neues (Bachelor-, Master-, PhD-) Projekt
- bei (eiligen) Projektberichten können **Nachfragen bei Ihnen** bzgl. Ihrer **Daten-Planung, Erhebung, Benennung** und **Speicherung** entstehen
- wenn Sie die Prinzipien kennen, vermeiden Sie doppelte Arbeit und unterstützen Ihre\*n Professor\*in

# Wie erreiche ich “FAIRe” Daten?

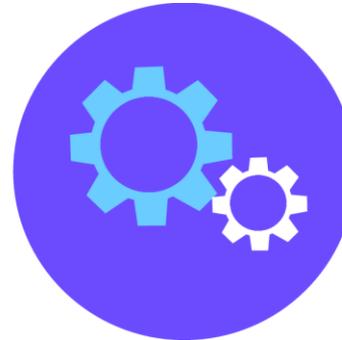
Auffindbar  
(**F**indable)



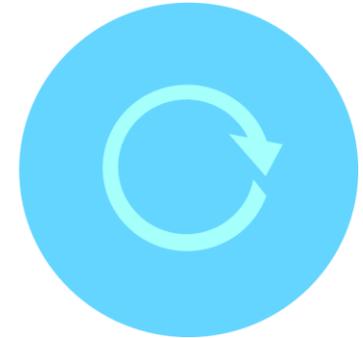
Zugänglich  
(**A**ccessible)



Interoperabel  
(**I**nteroperable)



Wiederverwendbar  
(**R**eusable)



Paulina Halina Sieminska. (2019, July 5). A FAIRy tale graphics (Version 1.0.0). Zenodo. <http://doi.org/10.5281/zenodo.3267168> CC BY SA 4.0

# Meine Daten sind auffindbar!



Dafür benötigt man...

- umfangreiche **Metadaten** (Autor\*in, Projekt...) zu Daten, explizit **verknüpft**
- eindeutige und dauerhafte **Identifikatoren** für Daten (z.B. DOI)
- die Registrierung und Indexierung von Daten in einer **durchsuchbaren Ressource** (z.B. ein Repository)

# Meine Daten sind zugänglich!



Dafür benötigt man...

- am besten ein **Repositorium**, welches folgendes gewährleistet:
  - zugängliche Meta(Daten) über einen Identifier mit einem **standardisierten Kommunikationsprotokoll** (z.B. HTTPS)
  - offenes, freies und universell implementierbares Protokoll
  - Protokoll erlaubt **Authentifizierung** und **Autorisierung**, wenn nötig
- **Metadaten bleiben zugänglich**, auch wenn die eigentlichen Daten nicht mehr zur Verfügung stehen

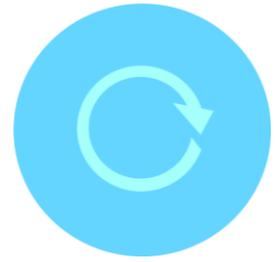
# Meine Daten sind interoperabel!



Dafür benötigt man...

- (Meta)Daten mit einer formalen, zugänglichen, gemeinsam genutzten und allgemein anwendbaren **Sprache zur Wissensrepräsentation**
- **Vokabulare** für (Meta)Daten, die den FAIR-Prinzipien (auffindbar...) folgen
- in (Meta)Daten qualifizierte **Verweise auf andere wichtige (Meta)Daten**
  - alle nötigen Dateien zum Verständnis Ihrer Dateien sind verlinkt

# Meine Daten sind wiederverwendbar!



Dafür benötigt man...

- Meta(Daten), die reichhaltig mit einer Vielzahl von **genauen und relevanten Attributen** beschrieben sind
  - klare und zugängliche **Datennutzungslizenz**
  - detaillierte **Herkunftsangabe**
  - domänenrelevante **Gemeinschaftsstandards**

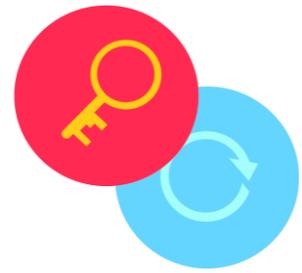
# Meine Dateien sind FAIR, auch für mich!

Ein paar Fragen, die Sie für sich beantworten sollten:

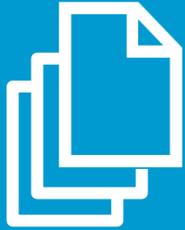
- **Wo** will ich meine Daten aufheben?
  - *Private Festplatte, USB-Stick, Cloud...*
- **Von wo** brauche ich Zugriff auf die Daten?
  - *USB-Stick in der Schublade oder Cloud?*
- **Wer** braucht noch Zugriff auf die Daten?
  - *Ich, meine AG, internationale Kooperationspartner?*
- **Wie viel Rechenleistung** brauche ich?
  - *Reicht mein Laptop bei Big Data?*



Das KIM hilft Ihnen gerne bei diesen und weiteren Fragen!



## 3-2-1 Backup-Regel



x 3

Erstellen Sie 3  
Kopien der Datei



x 2

Speichern sie die  
Kopien auf mind. 2  
verschiedenen  
Speichermedien



x 1

Speichern Sie  
mind. 1 Kopie in  
der NextCloud

# Meine Dateien sind FAIR, auch für mich!

Ich habe meine Masterarbeit vor einem Jahr abgegeben und will nun die Dateien für meine Promotion wieder verwenden...

Daten Erhebung 1 - Kopie	24.04.2020 14:44	Microsoft Excel-Arbeits...
Daten Erhebung 1 - neu	24.04.2020 14:44	Microsoft Excel-Arbeits...
Daten Erhebung 1	24.04.2020 14:44	Microsoft Excel-Arbeits...
Daten Erhebung bearbeitet	24.04.2020 14:44	Microsoft Excel-Arbeits...
Daten Erhebung final (2)	24.04.2020 14:44	Microsoft Excel-Arbeits...
Dok 1	24.04.2020 14:36	Microsoft Word-Dokum...
final 1		
final neu	24.04.2020 14:36	Microsoft Word-Dokum...
Thesis 1	24.04.2020 14:36	Microsoft Word-Dokum...
Thesis 2	24.04.2020 14:36	Microsoft Word-Dokum...
Thesis final 2	24.04.2020 14:36	Microsoft Word-Dokum...
Diagramm 1 neu	24.04.2020 14:39	PNG-Datei
Diagramm 1	24.04.2020 14:39	PNG-Datei
Diagramm 2	24.04.2020 14:39	PNG-Datei
SPSS Tabelle 4 - Kopie.sav	26.08.2019 07:24	SAV-Datei
SPSS Tabelle 4.sav	26.08.2019 07:24	SAV-Datei
R script Bachelor - k	24.04.2020 14:46	Textdokument
R script Bachelor	24.04.2020 14:46	Textdokument
R script korrigiert	24.04.2020 14:46	Textdokument
R script neu	24.04.2020 14:46	Textdokument

Wo ist "Erhebung 2"?

Welches ist jeweils die finale Version?

Warum hat alles das gleiche Datum?

Ich habe kein SPSS mehr, womit öffne ich diese Datei?

# Auffindbarkeit und Wiederverwendbarkeit durch gute Datei-Benennung



- **Datum** der Erstellung angeben
  - Amerikanisches Format „JJMMDD“ ermöglicht sinnvolle Sortierung
- **Einheitlich, konsistente und eindeutige** Benennung
  - nicht einmal „Thesis 1“ dann „final 1“
- **Keine Umlaute, Leer- und Sonderzeichen** (z.B. & \* % \$ £ ] { ! @)
  - viele Programme haben damit Probleme (z.B. „R“)
  - zur Abgrenzung Punkte, Unterstrich oder Bindestrich verwenden

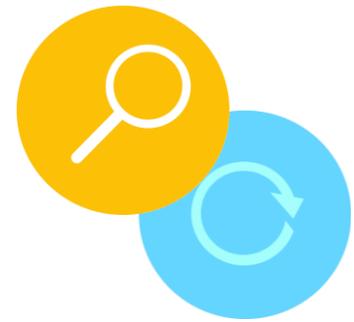
*z.B. 200518\_Masterarbeit\_IL\_v2*

# Auffindbarkeit und Wiederverwendbarkeit durch gute Datei-Benennung

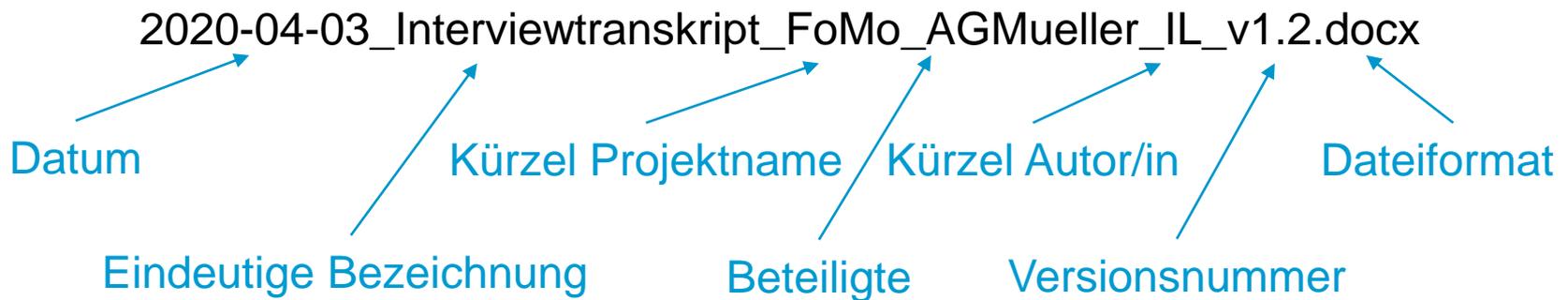


- **Kurze Dateinamen** (unter 225 Zeichen!)
  - auf Abkürzungen einigen und diese Metadaten z.B. in einem **README.txt** festhalten
- **Einheitliche Versionierung**
  - machen Sie Angaben zur Versionierung; z.B. V1 oder v1.1
- **Benennen Sie das Dateiformat**
  - z.B. .docx, .csv...

# Auffindbarkeit und Wiederverwendbarkeit durch gute Datei-Benennung



## Beispiel für eine gute Datei-Benennung:



## Weitere Tipps...

- Erstellen Sie bzw. fragen Sie nach einer **Anleitung** für die Benennung und Speicherung ihrer Dateien
  - Orientieren Sie sich möglichst an Standards ihres Fachbereichs
- Achten Sie möglichst auf die Verwendung von **freien Dateiformaten**
  - z.B. lassen sich .sav Datentabellen nur mit SPSS öffnen, lieber zusätzlich .csv wählen (auch mit R zu verwenden)
- **Schulen** Sie ggf. ihre Kolleg\*innen und neue Mitarbeiter\*innen
  - Hiwis, Technische Angestellte, andere Studierende mit Abschlussarbeit...
- **Aktualisieren** und **kontrollieren** Sie regelmäßig & zeitnah Ihre Dateien und ggf. Ihren Datenmanagementplan

# Angebote des KIM und weiterführende Links

Open Science



**Individuelle Beratung auf Anfrage**

[kim.uni.kn/openscience](http://kim.uni.kn/openscience)



**Informationsplattform**

[forschungsdaten.info](http://forschungsdaten.info)



**OpenAIRE-Portal zu FAIRen Daten**

[openaire.eu/how-to-make-your-data-fair](http://openaire.eu/how-to-make-your-data-fair)



**Research Data Management Organizer**

<https://rdmo.forschungsdaten.info/>

# Übersicht über alle Kursmodule

Alle Module auf einen Blick

1 Was bietet mir Open Science?

2 Warum Forschungsdatenmanagement?

3 Open Access – Es ist nicht alles Gold was glänzt

4 **Maximale Sichtbarkeit für meine Forschung**

4.1 **FAIR**

4.2 Repositorien

4.3 Persistente Identifikatoren

5 Arbeitserleichterung durch freie Bildungsmaterialien

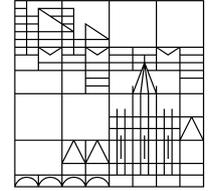
6 Rechtsfragen beim Veröffentlichen

# Quellenangaben

- **Internationale GO FAIR Initiative:** <https://www.go-fair.org/> [20.04.2020].
- **OpenAIRE** European Open Science Infrastructure, for open scholarly and scientific communication [www.openaire.eu](http://www.openaire.eu) [20.04.2020].
- **European Commission** Open Science Strategy  
<https://ec.europa.eu/research/openscience/index.cfm> [20.04.2020].
- Informationsplattform **forschungsdaten.info**  
<https://www.forschungsdaten.info/> [20.04.2020].

# Abbildungsnachweise

- 1. Titelfolie, Folie 5: OpenClipart-Vectors/pixabay - Bearbeitung I. Lang
- 1. Titelfolie, Folie 25: mcmurryjulie/pixabay - Bearbeitung I. Lang
- 2. Titelfolie, Folie 4,10-16, 18-20,: A FAIRy tale graphics. Paulina Halina Sieminska. (2019, July 5) (Version 1.0.0). Zenodo. <http://doi.org/10.5281/zenodo.3267168> *CC BY SA 4.0*  
Bearbeitung I. Lang
- Folie 16: Ilona Lang (3-2-1 Backup-Regel)



**Herzlichen  
Dank!**

**Lena Dreher  
Dr. Ilona Lang**

Team Open Science  
[openscience@uni-konstanz.de](mailto:openscience@uni-konstanz.de)

