

**Open Science  
Forschungs-  
datenmanagement**

**Matthias Landwehr**

KIM, 13.02.2023

**LOVE  
DATA  
Week  
2023**

**„Open science is the idea that scientific knowledge of all kinds should be openly shared as early as is practical in the discovery process.“**

Michael Nielsen,  
Quantum physicist

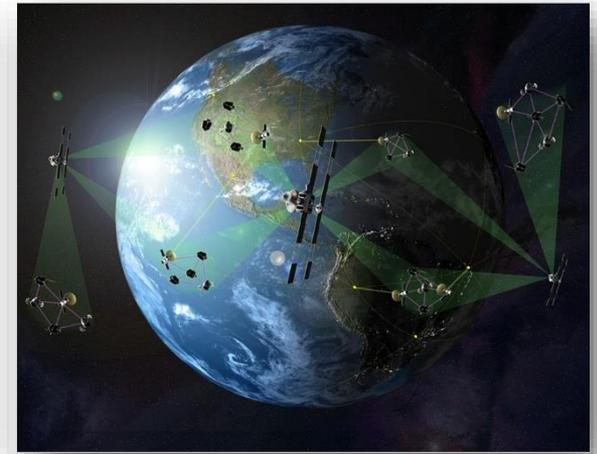
# Definition und Abgrenzung

*„Unter Forschungsdaten sind [...]*

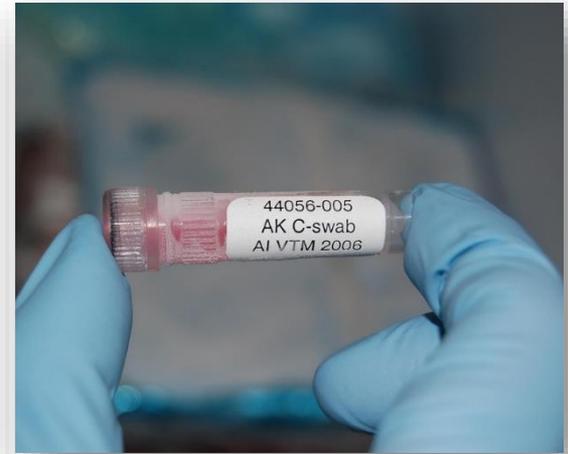
Daten zu verstehen, die im Zuge eines wissenschaftlichen Vorhabens z. B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“

(DFG 2009)

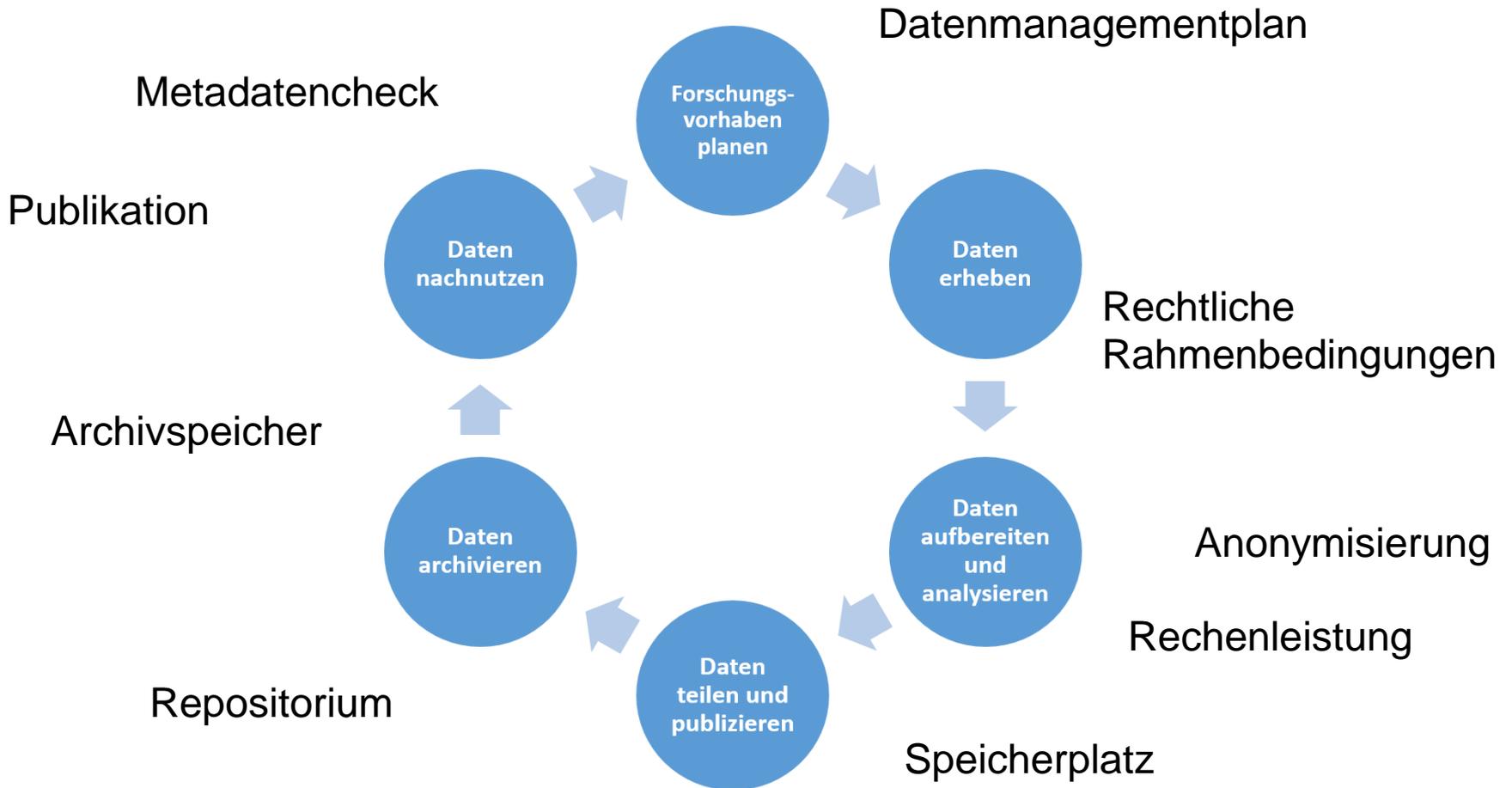
# Beispiele



# Beispiele

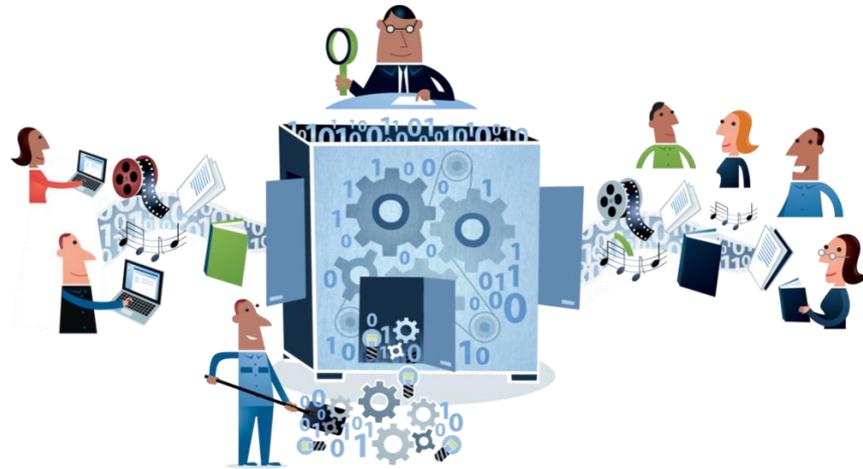


# Forschungsdatenlebenszyklus



# Warum Forschungsdatenmanagement?

- Schnellere Auffindbarkeit von Daten
- Wissenserhalt
- Datenverlust vorbeugen
- leichtere Zusammenarbeit
- Weitergabe und Nachnutzung der Daten
- Erhöhung der eigenen Sichtbarkeit
- Nachvollziehbarkeit von (eigenen) Forschungsergebnissen
- Referenzierbarkeit (Verwendung von persistenten Identifikatoren empfohlen)
- Gesamtgesellschaftlicher Nutzen offener Daten
- Vorgaben in Forschungsdaten-Policies (Universität, Drittmittelgebern, Verlagen)



Digitalbevaring.dk

# Open Science Policy

- **Im 11/2021 entstanden aus OA-Policy und Policy zum Forschungsdatenmanagement**
  - [...] Die Universität Konstanz erkennt die Veröffentlichung von Forschungsdaten als Teil des wissenschaftlichen Prozesses an.
  - Wissenschaftler\*innen der Universität Konstanz sind verpflichtet, Forschungsdaten nach geltenden fachspezifischen Standards aufzubereiten, zu sichern, zu dokumentieren und aufzubewahren. Dies umfasst auch die Dokumentation verwendeter Methoden und Werkzeuge. Die Orientierung an den FAIR-Prinzipien wird empfohlen.
  - Die Universität Konstanz fordert ihre Forschenden nachdrücklich auf, soweit zulässig und nicht unzumutbar, Forschungsdaten zum frühestmöglichen Zeitpunkt in Form einer zitierbaren Datenpublikation öffentlich zugänglich zu machen [...] Es werden unterstützende Infrastruktur- und Beratungsangebote bereitgestellt.

# Was bedeutet “FAIR” für Daten?

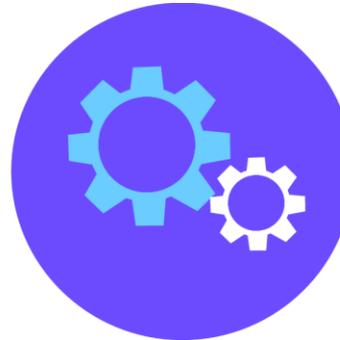
Auffindbar  
(**F**indable)



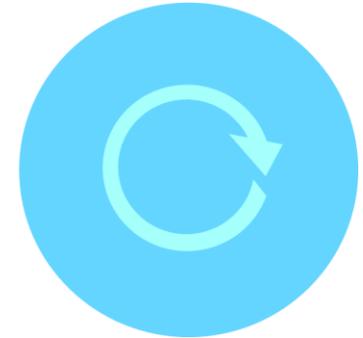
Zugänglich  
(**A**ccessible)



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

# 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>

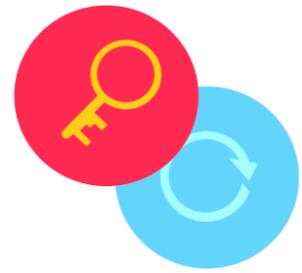
# Meine Dateien sind FAIR, auch für mich!

Ein paar Fragen, die Sie für sich beantworten sollten:

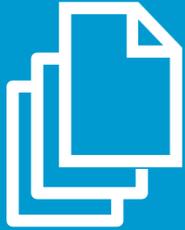
- **Wo** will ich meine Daten aufheben?
  - *private Festplatte, USB-Stick, Cloud...*
- **Von wo** brauche ich Zugriff auf die Daten?
  - USB-Stick in der Schublade oder Cloud?
- **Wer** braucht noch Zugriff auf die Daten?
  - ich, meine AG, internationale Kooperationspartner?
- **Wie viel Rechenleistung** 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?

# Auffindbar und Wiederverwendbar 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

# Auffindbar und Wiederverwendbar durch gute Datei-Benennung

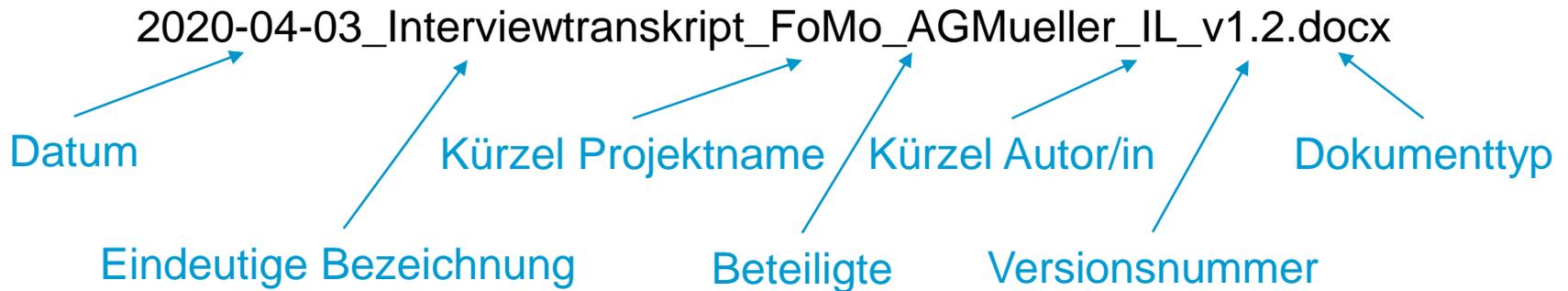


- **Kurze Dateinamen** (unter 225 Zeichen!)
  - auf Abkürzungen einigen und diese Metadaten z.B. in einem **README.txt** festhalten
- **Einheitliche Versionierung**
  - machen Sie Angaben zur Versionierung; z.B. V1 oder v1.1
- **Benennen Sie das Dateiformat**
  - z.B. .docx, .csv...

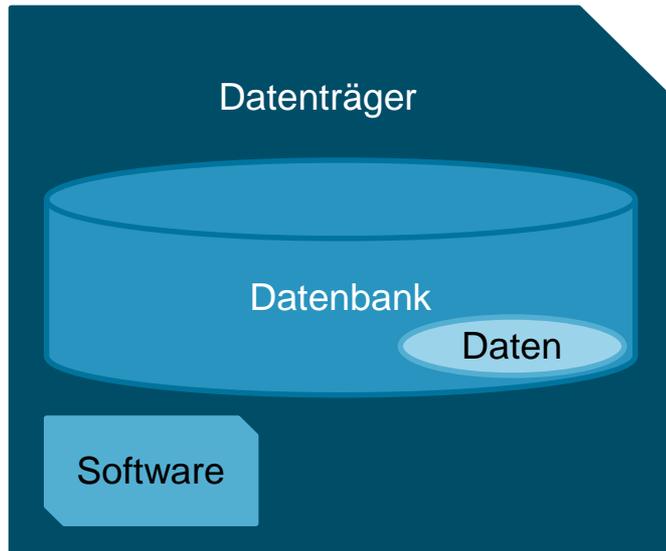
# Auffindbar und Wiederverwendbar durch gute Datei-Benennung



## Beispiel für eine gute Datei-Benennung:



# Rechtspositionen bei Forschungsdaten



## Einzelne Daten / Datensätze

- **Werkschutz** von Daten
- **Schutzlücken**, wenn Daten keine Schöpfungshöhe aufweisen. Umstritten
- **Wissenschaftler-Persönlichkeitsrecht?**
- **Datenschutzrechte** der Betroffenen

## Datenbank

- Rechte am Datenbankwerk
- Datenbankherstellerrecht

## Software

Werkschutz

## Datenträger

Eigentum

# Große Datenmengen

- 17 TB Daten

- Wo speichern?

- Cloud, Fileserver, Dropbox, Google Drive?

- Wie übertragen?

- USB: 82 Stunden
- 10G-LAN: 4 Stunden 5G
- Experim.: 0,44 Sekunden 319 Terabit/Sekunde

# KonDATA

Login

DE

Kommunikations-, Informations-, Medienzentrum (KIM)

## KonDATA

Universität  
Konstanz



KonDATA ist das Forschungsdatenrepositorium der Universität Konstanz. Dieser Service des KIM bietet Ihnen die Möglichkeit, Ihre Forschungsdaten langfristig sichtbar und zitierfähig zu publizieren.



Die Veröffentlichung von Forschungsdaten wird durch den Mentalitätswandel hin zu Open Science immer wichtiger. Mit einer Veröffentlichung auf KonDATA können Sie die Auflagen zur Datenpublikation diverser Forschungsförderer erfüllen. Die in KonDATA veröffentlichten Daten werden mit einem Digital Object Identifier (DOI) versehen. Dadurch können sie global eindeutig und dauerhaft identifiziert und zitiert werden. Informationen zu Ihren Daten und Metadaten werden über standardisierte Schnittstellen an RADAR, einem Anbieter für wissenschaftliche Repositorien, und andere Meta-Suchmaschinen übertragen. Dadurch sind sie auch außerhalb von KonDATA zu finden. Bei der Veröffentlichung von Daten führt das KIM eine formale Kuration durch. Dadurch wird sichergestellt, dass Ihre Daten nach gängigen Standards beschrieben werden und somit verständlich für Nachnutzer\*innen sind.

Betreiber ist das KIM der Universität Konstanz. KonDATA setzt auf dem von FIZ Karlsruhe angebotenen Dienst RADAR auf. Die Speicherung der Daten findet ausschließlich auf IT-Infrastruktur des KIM an der Universität Konstanz statt.

### Daten publizieren

Melden Sie sich in KonDATA an und starten Sie Ihre Datenpublikation

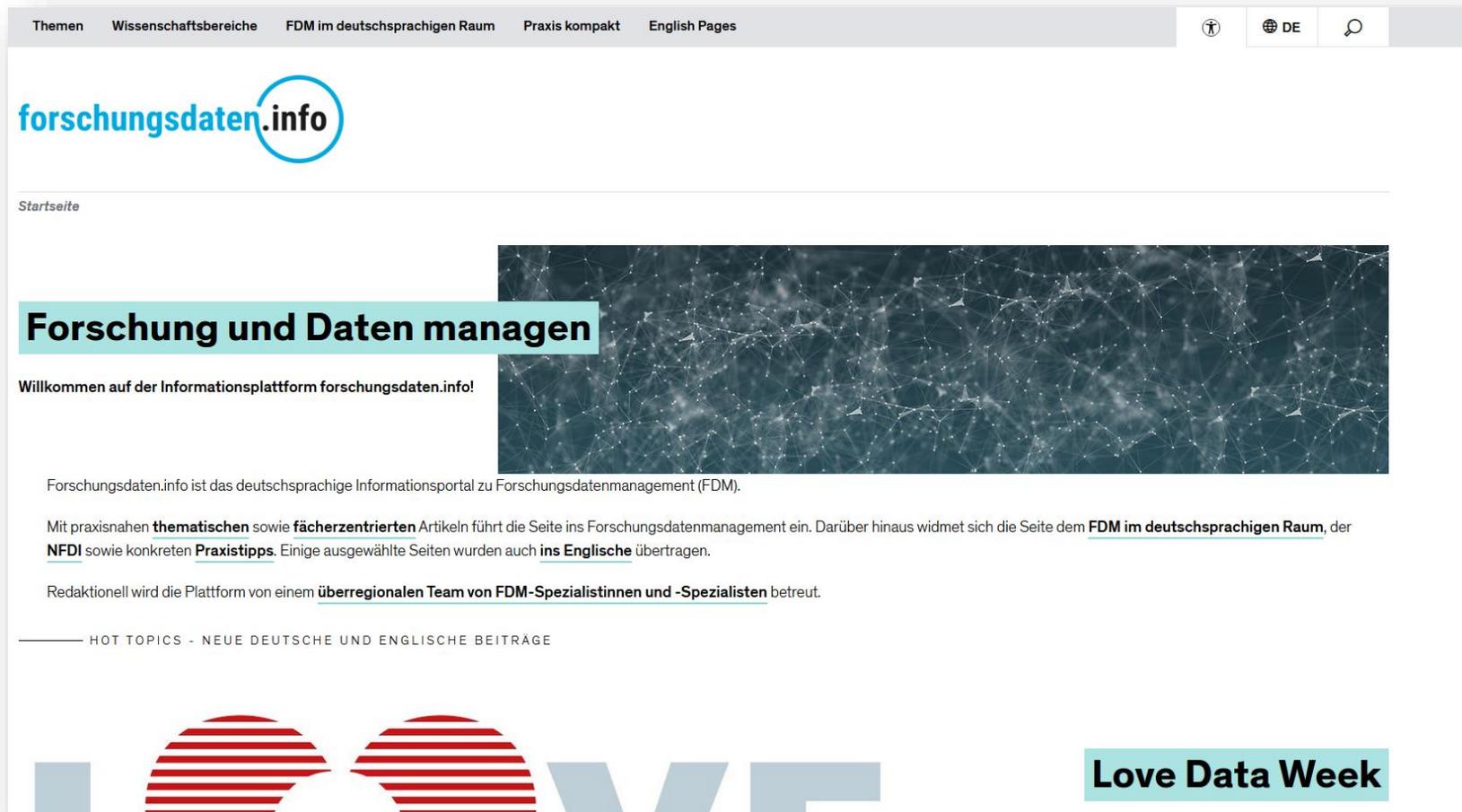
### Daten durchsuchen

Finden Sie Datensätze in KonDATA, filtern Sie die Ergebnisse.

### Nutzungsbedingungen

### Hilfe & Unterstützung

# www.forschungsdaten.info



The screenshot shows the homepage of the website 'forschungsdaten.info'. At the top, there is a navigation bar with links for 'Themen', 'Wissenschaftsbereiche', 'FDM im deutschsprachigen Raum', 'Praxis kompakt', and 'English Pages'. On the right side of the navigation bar, there are icons for accessibility, a globe with 'DE', and a search icon. The main header features the logo 'forschungsdaten.info' in blue. Below the logo, the text 'Startseite' is visible. A large banner image with a network pattern is partially obscured by a teal box containing the text 'Forschung und Daten managen'. Below the banner, a teal box contains the text 'Willkommen auf der Informationsplattform forschungsdaten.info!'. The main content area contains several paragraphs of text: 'Forschungsdaten.info ist das deutschsprachige Informationsportal zu Forschungsdatenmanagement (FDM).', 'Mit praxisnahen **thematischen** sowie **fächerzentrierten** Artikeln führt die Seite ins Forschungsdatenmanagement ein. Darüber hinaus widmet sich die Seite dem **FDM im deutschsprachigen Raum**, der **NFDI** sowie konkreten **Praxistipps**. Einige ausgewählte Seiten wurden auch **ins Englische** übertragen.', and 'Redaktionell wird die Plattform von einem **überregionalen Team von FDM-Spezialistinnen und -Spezialisten** betreut.' Below this text, there is a horizontal line followed by the text 'HOT TOPICS - NEUE DEUTSCHE UND ENGLISCHE BEITRÄGE'. At the bottom of the page, there is a decorative graphic with red and white striped circles and a teal box containing the text 'Love Data Week'.

Themen Wissenschaftsbereiche FDM im deutschsprachigen Raum Praxis kompakt English Pages

forschungsdaten.info

Startseite

## Forschung und Daten managen

Willkommen auf der Informationsplattform forschungsdaten.info!

Forschungsdaten.info ist das deutschsprachige Informationsportal zu Forschungsdatenmanagement (FDM).

Mit praxisnahen **thematischen** sowie **fächerzentrierten** Artikeln führt die Seite ins Forschungsdatenmanagement ein. Darüber hinaus widmet sich die Seite dem **FDM im deutschsprachigen Raum**, der **NFDI** sowie konkreten **Praxistipps**. Einige ausgewählte Seiten wurden auch **ins Englische** übertragen.

Redaktionell wird die Plattform von einem **überregionalen Team von FDM-Spezialistinnen und -Spezialisten** betreut.

HOT TOPICS - NEUE DEUTSCHE UND ENGLISCHE BEITRÄGE

Love Data Week

# Forschungsdatenmanagement

**Matthias Landwehr**

openscience@uni-konstanz.de  
kim.uni.kn/openscience

## Mailingliste

- Community bilden
- Für Wissenschaftler\*innen und Angehörige der Universität

openscience@mailman.uni-konstanz.de

## Weitere Informationen

- open-access.network
- forschungsdaten.info

**Herzlichen Dank!**

**Be open!**