

Data Sharing interaktiv vermitteln

Esther Asef, Katarzyna Biernacka, Elisabeth Böker,
Sarah Ann Danker, Juliane Jacob, Janna Neumann, Britta Petersen,
Jessica Rex und Ute Trautwein-Bruns

der *DINI/nestor UAG* „Schulungen/Fortbildungen“

eScience Tage 2021, 05. März 2021



Sofern nicht anders angegeben, stehen die Inhalte dieser Präsentation unter einer unter CC0 1.0 Universell (CC0 1.0) Public Domain Dedication (<https://creativecommons.org/publicdomain/zero/1.0/deed.de>).



Die UAG Schulungen/Fortbildungen

- Unter-AG der DINI/nestor-AG Forschungsdaten
- gegründet am 5. März 2019 in Berlin
- aktuelle Mitgliederzahl: 27
- Ziel:
 - Schulungskonzepte und -Modelle erarbeiten und weiterentwickeln
 - Schulungsmaterialien sammeln und bereitstellen



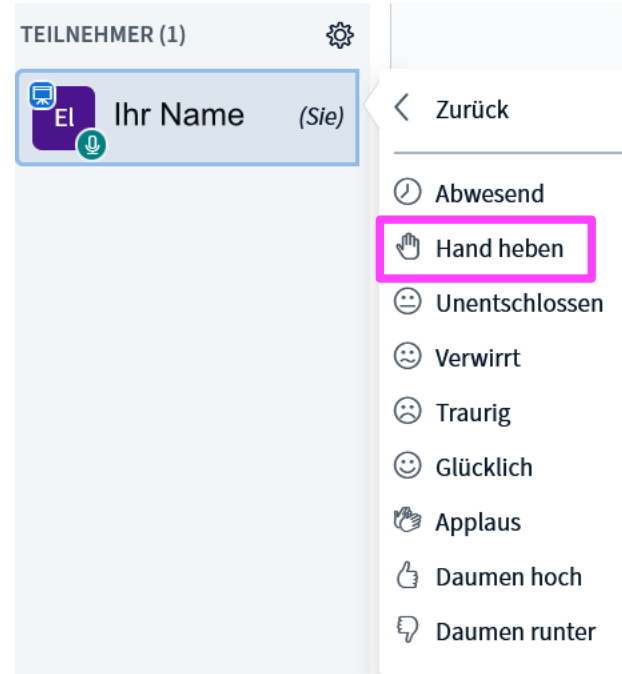
Code of Conduct

- Mikros und Kameras bitte in der Regel aus
- Technische Probleme bitte im Chat melden
- Workshop-Du
- Beteiligt euch rege!
- Fragen bitte über die Funktion "Hand heben"



Funktion "Hand heben"

- Unter „Teilnehmer“





Erwartungsabfrage

Was interessiert dich an diesem Online-Workshop?

Bitte beantworte die BBB-Umfrage





Erwartungsabfrage

Was interessiert dich an diesem Online-Workshop am meisten?

A. Inhalte

B. Didaktische Methoden

C. Beides



Agenda

Orientierung	
<p>inhaltlich</p> <ul style="list-style-type: none">• FAIR Principles und ihre Bedeutung• Data Sharing<ul style="list-style-type: none">• Wie?• Wo?• Was?	<p>methodisch</p> <ul style="list-style-type: none">• Funktionen von BBB<ul style="list-style-type: none">• Breakout-Sessions• Einbindung externer Tools<ul style="list-style-type: none">• PollEv• Mentimeter• Padlet• Unterschiedliche Sozial- und Arbeitsformen
Fragen, Diskussion und Feedback	

- Ziel: "Vermittlung von Data Sharing mit interaktiven Lehrmethoden verbessern"

Die FAIR-Prinzipien

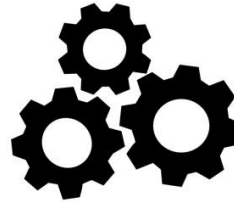
F
indable



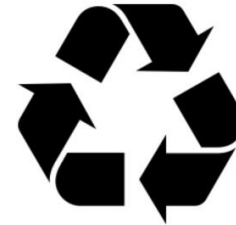
A
ccessible



I
nteroperable



R
eusable



Quelle: Pundir, Sangya. https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg [letzter Zugriff: 10.02.2021], CC-BY-SA-4.0 <https://creativecommons.org/licenses/by-sa/4.0/deed.en>.

Wilkinson, Mark D., Michel Dumontier, IJsbrand J. Aalberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.

Übung: FAIR-Zuordnung

Bitte nimm an der Umfrage ([Poll Everywhere](#)) teil und ordne die Aussagen einer der vier FAIR-Kategorien zu:

- Die Forschungsdaten haben eine klare und zugängliche Nutzungslizenz. (Reusable)
- Den (Meta)Daten wurde ein global eindeutiger und dauerhaft persistenter Identifier zugewiesen. (Findable)
- Die Forschungs- und Metadaten verwenden allgemein übliche, zugängliche und vorzugsweise offene Standards und Formate. (Interoperable)
- Die Metadaten sind und bleiben verfügbar, auch für den Fall, dass die zugehörigen Forschungsdaten nicht mehr vorhanden sind. (Accessible)

[Die Umfrage erfolgt anonym. Es ist nicht nötig einen Namen einzutragen!]

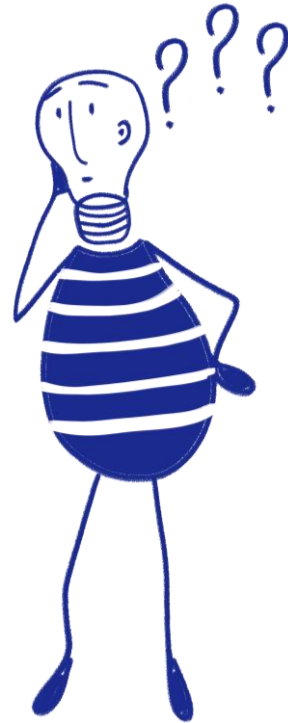
Findable & Accessible

- „F“ und „A“ stehen für eine Kultur des Teilens
- Passend zum Motto der E-Science-Tage 2021 - „Share Your Research Data“
- Passend zum Workshoptitel „Data Sharing interaktiv vermitteln“
- Aber was bedeutet das konkret?

Data Sharing – Wie?

Publikationswege für Daten

1. Ergänzung zu einem begutachteten Artikel („enhanced publication“)
2. Eigenständiges Informationsobjekt in einem Forschungsdaten-Repository
3. Data Journals



Publikationswege für Daten

1. Ergänzung zu einem begutachteten Artikel („enhanced publication“)
 - als sog. Supplement beim dazugehörigen Artikel
 - als Datensatz in einem Repository mit Verlinkung zum dazugehörigen Artikel

Beispiel:

Citation:

Leu, Eva; Brown, Thomas A; Graeve, Martin; Wiktor, Jozef M; Hoppe, Clara Jule Marie; Chierici, Melissa; Fransson, Agneta; Verbiest, Sander; Kvernvik, Ane C; Greenacre, Michael J (2020): Fatty acid composition, ice chemistry and algae biomass during an Arctic sea ice algal bloom on landfast ice near the settlement of Svea, in Van Mijenfjorden, Spitsbergen. *PANGAEA*, <https://doi.pangaea.de/10.1594/PANGAEA.925015> (dataset in review)

Supplement to:

Leu, Eva; Brown, Thomas A; Graeve, Martin; Wiktor, Jozef M; Hoppe, Clara Jule Marie; Chierici, Melissa; Fransson, Agneta; Verbiest, Sander; Kvernvik, Ane C; Greenacre, Michael J (2020): Spatial and Temporal Variability of Ice Algal Trophic Markers - With Recommendations about Their Application. *Journal of Marine Science and Engineering*, **8(9)**, 676, <https://doi.org/10.3390/jmse8090676>

Quelle: <https://doi.pangaea.de/10.1594/PANGAEA.925015>



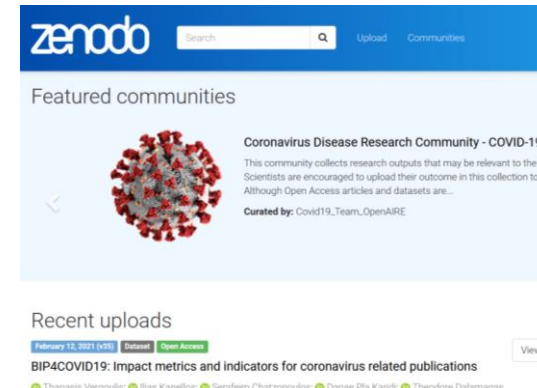
Publikationswege für Daten

2. Eigenständiges Informationsobjekt in einem Forschungsdaten-Repository

- disziplinspezifische Repositorien, z. B. Datorium, Pangaea
- disziplinübergreifende Repositorien, z. B. ZENODO
- institutionelle Repositorien, z. B. Refubium, edoc



Quelle: <https://www.pangaea.de/>, Zugriff 10.02.2021



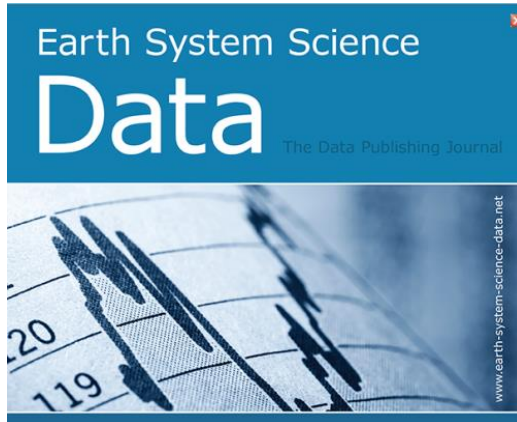
Quelle: <https://zenodo.org/>, Zugriff 10.02.2021



Publikationswege für Daten

3. Data Journals

- publizieren ausführliche Beschreibung der Daten
- z. T. peer-reviewed



Quelle: <https://www.earth-system-science-data.net>, Zugriff 10.02.2021



ISSN: 2352-3409

- Submit Your Paper ▾
- View Articles
- Guide for Authors ▾
- Abstracting/ Indexing

Data in Brief

> Open Access

Editors-in-Chief: Hao-Ran Wang, G

> View Editorial Board

Data in Brief provides a way for researchers to describe their datasets by publishing data articles.

- Thoroughly describe your data, facilitating its reuse.
- Make your data, which is often buried in supplementary materials, easy to find.
- Increase traffic towards associated publications.

Read more

Quelle: <https://www.journals.elsevier.com/data-in-brief>, Zugriff 10.02.2021



Data Sharing – Wo? Repository

„Ein Repository (lateinisch repositorium, Lager‘) ist ein verwalteter Ort zur Aufbewahrung geordneter Dokumente, die öffentlich oder einem beschränkten Nutzerkreis zugänglich sind.“

In einem Archiv (lat. Archivum, Aktenschrank‘) werden dagegen ausschließlich historische Dokumente verwaltet.“

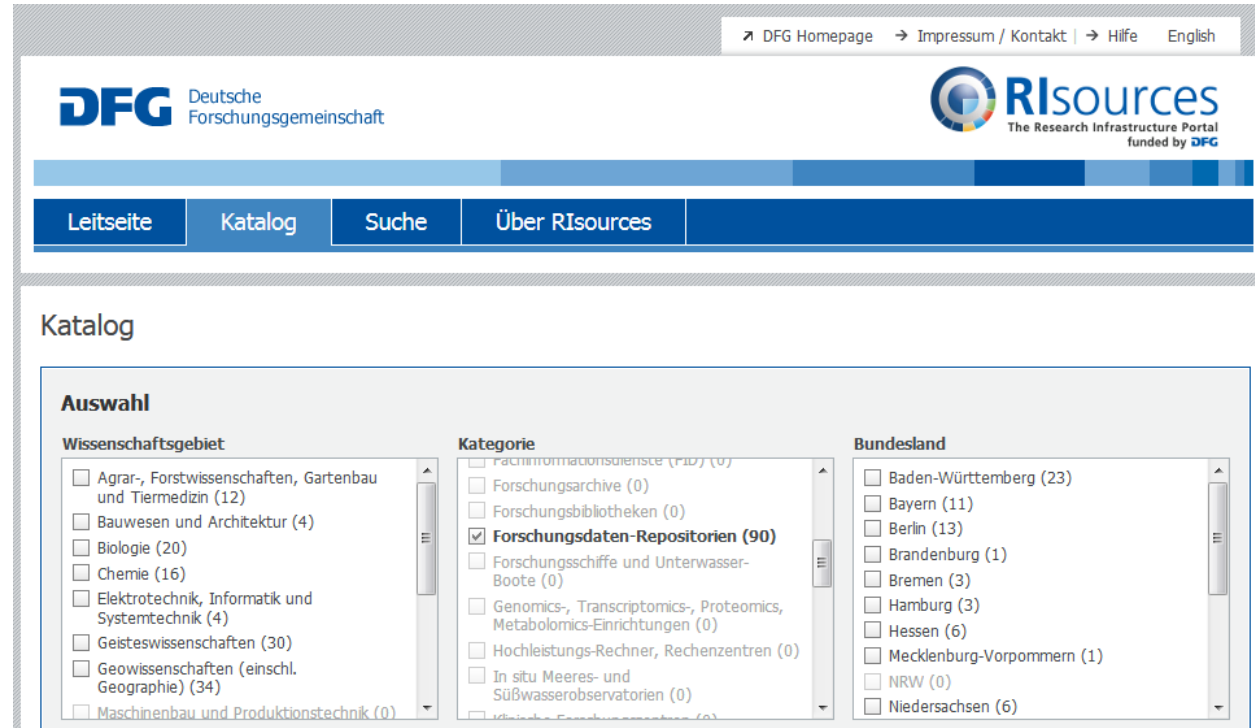
„Digitale Forschungsdaten-Repositoryn sind Informationsinfrastrukturen, die digitale Forschungsdaten möglichst dauerhaft... speichern und organisieren, um die Auffindbarkeit und Zugänglichkeit der Daten zu sichern...“



Quelle: Heinz Pampel, Vortrag in einer Vorlesung im Modul „Ausgewählte Aspekte digitaler Informationsversorgung“ an der Humboldt-Universität zu Berlin, Institut für Bibliotheks- und Informationswissenschaft (IBI), 12.06.2014

Repositorien finden: resources.dfg.de

- Angebot der DFG
- Informationsportal
- Deutschlandweit
- Forschungsinfrastrukturen
- Für Forschende



The screenshot shows the RIsources website interface. At the top, there are navigation links for 'DFG Homepage', 'Impressum / Kontakt', 'Hilfe', and 'English'. The DFG logo and 'Deutsche Forschungsgemeinschaft' are on the left, and the RIsources logo 'The Research Infrastructure Portal funded by DFG' is on the right. Below the logos is a navigation bar with 'Leitseite', 'Katalog', 'Suche', and 'Über RIsources'. The main content area is titled 'Katalog' and contains three filter sections: 'Auswahl', 'Wissenschaftsgebiet', 'Kategorie', and 'Bundesland'. The 'Auswahl' section is currently expanded to show 'Wissenschaftsgebiet' and 'Kategorie' filters. In the 'Kategorie' filter, 'Forschungsdaten-Repositorien (90)' is selected with a checkmark. The 'Bundesland' filter shows a list of German states with their respective counts.

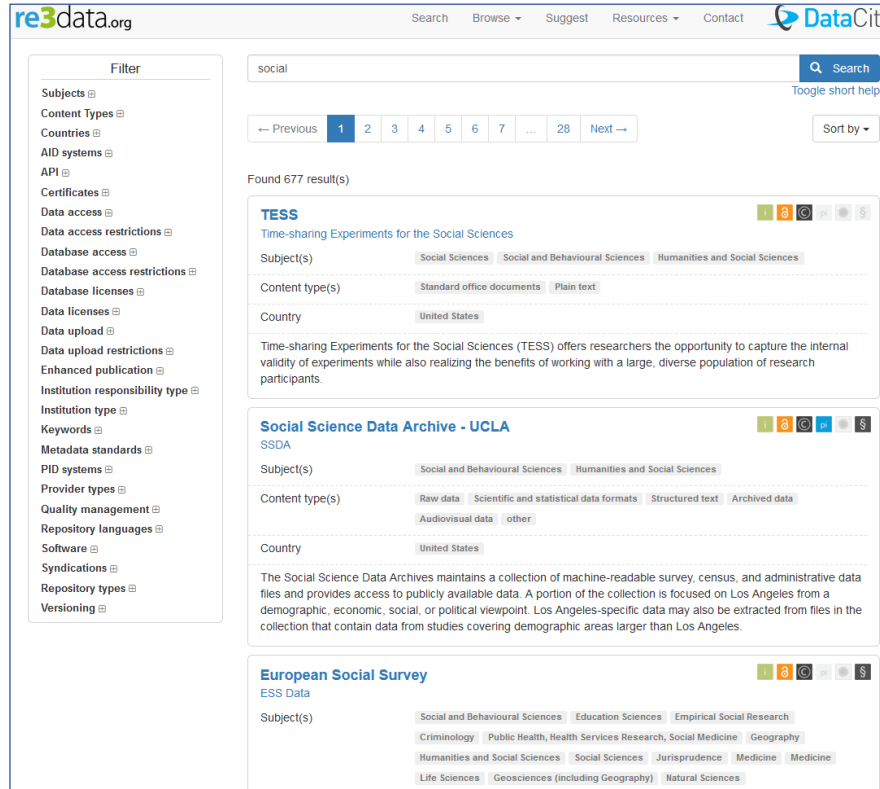
Wissenschaftsgebiet	Kategorie	Bundesland
<input type="checkbox"/> Agrar-, Forstwissenschaften, Gartenbau und Tiermedizin (12)	<input type="checkbox"/> Fachinformationswissenschaft (0)	<input type="checkbox"/> Baden-Württemberg (23)
<input type="checkbox"/> Bauwesen und Architektur (4)	<input type="checkbox"/> Forschungsarchive (0)	<input type="checkbox"/> Bayern (11)
<input type="checkbox"/> Biologie (20)	<input type="checkbox"/> Forschungsbibliotheken (0)	<input type="checkbox"/> Berlin (13)
<input type="checkbox"/> Chemie (16)	<input checked="" type="checkbox"/> Forschungsdaten-Repositorien (90)	<input type="checkbox"/> Brandenburg (1)
<input type="checkbox"/> Elektrotechnik, Informatik und Systemtechnik (4)	<input type="checkbox"/> Forschungsschiffe und Unterwasser-Boote (0)	<input type="checkbox"/> Bremen (3)
<input type="checkbox"/> Geisteswissenschaften (30)	<input type="checkbox"/> Genomics-, Transcriptomics-, Proteomics, Metabolomics-Einrichtungen (0)	<input type="checkbox"/> Hamburg (3)
<input type="checkbox"/> Geowissenschaften (einschl. Geographie) (34)	<input type="checkbox"/> Hochleistungs-Rechner, Rechenzentren (0)	<input type="checkbox"/> Hessen (6)
<input type="checkbox"/> Maschinenbau und Produktionstechnik (0)	<input type="checkbox"/> In situ Meeres- und Süßwasserobservatorien (0)	<input type="checkbox"/> Mecklenburg-Vorpommern (1)
	<input type="checkbox"/> Klassische Experimentelle (0)	<input type="checkbox"/> NRW (0)
		<input type="checkbox"/> Niedersachsen (6)

Quelle: http://resources.dfg.de/index.html#q=*%&sort=RI_SORT_DE%20asc&rows=10&RI_EXT=Y. Zugriff 10.02.2021



Repositorien finden: re3data.org

- Sammlung von Repositorien
- Weltweit
- Verschiedene Disziplinen
- Forschende, Förderer, Verlage und Institutionen



The screenshot shows the re3data.org search interface. The search term 'social' is entered in the search bar, and the results are filtered to show 677 results. The first three results are:

- TESS** (Time-sharing Experiments for the Social Sciences):
 - Subject(s): Social Sciences, Social and Behavioural Sciences, Humanities and Social Sciences
 - Content type(s): Standard office documents, Plain text
 - Country: United States
 - Description: Time-sharing Experiments for the Social Sciences (TESS) offers researchers the opportunity to capture the internal validity of experiments while also realizing the benefits of working with a large, diverse population of research participants.
- Social Science Data Archive - UCLA** (SSDA):
 - Subject(s): Social and Behavioural Sciences, Humanities and Social Sciences
 - Content type(s): Raw data, Scientific and statistical data formats, Structured text, Archived data, Audiovisual data, other
 - Country: United States
 - Description: The Social Science Data Archives maintains a collection of machine-readable survey, census, and administrative data files and provides access to publicly available data. A portion of the collection is focused on Los Angeles from a demographic, economic, social, or political viewpoint. Los Angeles-specific data may also be extracted from files in the collection that contain data from studies covering demographic areas larger than Los Angeles.
- European Social Survey** (ESS Data):
 - Subject(s): Social and Behavioural Sciences, Education Sciences, Empirical Social Research, Criminology, Public Health, Health Services Research, Social Medicine, Geography, Humanities and Social Sciences, Social Sciences, Jurisprudence, Medicine, Medicine, Life Sciences, Geosciences (including Geography), Natural Sciences

Quelle: re3data About. <http://service.re3data.org/about>. Zugriff 10.02.2021



Übung: Repositorien finden auf re3data.org

Suche auf re3data.org nach einem Repository, das für deine Daten passend ist.
Diskutiert in der Gruppe (Breakout-Sessions) eure Suche und Auswahl (Padlet).



Quelle: Wikipedia / Daniel Mietchen / CC BY 4.0

Data Sharing – Was? Daten auswählen

- Meilenstein-Version
- Daten, die einer Publikation zugrunde liegen
- Datenqualität
- Rechte
- Einzigartigkeit der Daten
- Zeit/Kosten/Nutzen
- Technischer Aufwand



Quelle: Pixabay

Metadaten – Daten über Daten

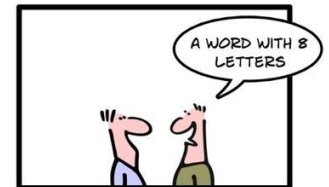
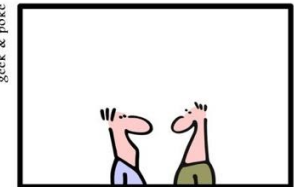
Inhaltliche Metadaten:

- Titel
- Beschreibung
- Autor*in
- Urheberrechts-Inhaber*in
- Kontaktdaten
- Lizenzangaben
- Schlüsselwörter

Technische Metadaten:

- Aufnahmezeitpunkt
- Brennweite
- Blende
- Belichtungsdauer
- Geographische Koordinaten
- und viele weitere

SIMPLY EXPLAINED:
METADATA



geek & poke

by Geek and Poke, CC-BY-3.0

Beispiele für Metadatenstandards

Eine Übersicht zu disziplinspezifischen und fachübergreifenden Metadatenstandards unter:

<http://rd-alliance.github.io/metadata-directory/subjects/>

Beispiele:

- Fachübergreifende Metadatenstandards:
 - Dublin Core
 - MARC21
- Geowissenschaften:
 - ISO 19115
- Biodiversität:
 - Darwin Core
- Geisteswissenschaften:
 - Text Encoding Initiative (TEI)
- Naturwissenschaften:
 - ICAT Schema
 - Cristallographic Information Framework
- Sozialwissenschaften:
 - Data Documentation Initiative (DDI)

Normdaten

Dienen der eindeutigen Identifikation von Personen, Institutionen, Forschungsförderer...

- Gemeinsame Normdatei (GND)
- International Standard Name Identifier (ISNI)
- Virtual International Authority File (VIAF)
- Open Funder Registry



Kontrolliertes Vokabular

Klassifikation

dient der Zuordnung von Objekten in (meist hierarchisch strukturierten) Klassen. Die Klassen sind durch bestimmte Merkmale charakterisiert.

Thesaurus

Ist eine natürlich-sprachliche geordnete Sammlung von Begriffen und deren Beziehungen zueinander.



Persistente Identifier (PID)

- dauerhafter, digitaler Identifikator aus Ziffern und/oder alphanumerischen Zeichen
- PIDs sind digitalem Objekt zugeordnet und verweist direkt darauf
- Beispiele für PIDs:
 - DOI (Digital Object Identifier)
 - URN (Uniform Resource Name)
 - Handle
 - ORCID (Personen)
- PID bleibt unverändert, auch wenn sich Ablageort ändert
- wird das Objekt gelöscht, bleibt Landing-Page mit Metadaten erhalten
- Durch PIDs sind digitale Objekte dauerhaft auffindbar, abrufbar und zitierbar.


ZITATIONSVORSCHLAG


Biernacka, Katarzyna, Petra Buchholz, Sarah Ann Danker, Dominika Dolzycka, Claudia Engelhardt, Kerstin Helbig, Juliane Jacob, Janna Neumann, Carolin Odebrecht, Ute Trautwein-Bruns, Cord Wiljes, und Ulrike Wuttke: *Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement*. Version 3.1. Berlin, 2020. <https://doi.org/10.5281/zenodo.4322849>

Was sind Lizenzen?

- Lizenzen regeln Bedingungen der Nachnutzung von veröffentlichten Daten
- Freie Lizenzen erlauben die Nutzung, Weiterverbreitung und Änderung urheberrechtlich geschützter Werke
 - stehen i.d.R zum freien Einsatz zur Verfügung und müssen nur verlinkt werden
 - Voraussetzung ist, dass du der/die Rechteinhaber*in bist
- Auswahl der Lizenz ist von der Art der Daten abhängig:
 - z.B. Creative Commons (CC) - Lizenzen für Artikel, Monografien, Bilder etc.
 - Open-Database-License (ODbL) für DB oder CC ab Version 4
 - General Public Licence (GNU) für Software
- wird keine Lizenz vergeben, gilt ggf. das strengere Urheberrecht

CC-Lizenzen


 CC0 (Public Domain)


 CC BY (Namensnennung)

 CC BY-SA (Namensnennung - Weitergabe unter gleichen Bedingungen)

 CC BY-ND (Namensnennung - Keine Bearbeitung)

 CC BY-NC (Namensnennung - Nicht kommerziell)

 CC BY-NC-SA (Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen)

 CC BY-NC-ND (Namensnennung - Nicht-kommerziell - Keine Bearbeitung)



Vorsicht!

nestor 

Lizenzfrei
≠
freie Lizenz





Übung: Stimmungslage Data Sharing

Gebt eine Einschätzung zu den folgenden Aussagen bzgl. Data Sharing ab
(Mentimeter-Umfrage)



Checkliste für Data Sharing

- Policies: fordern i.d.R. „FAIRe“ Daten
 - „So offen wie möglich, so geschlossen wie nötig“
- Access: so niederschwellig wie möglich
- Lizenzen: so frei wie möglich
- PIDs: stabile Identifikation und Verlinkung
- Zertifikate und Standards: Vertrauen und Interoperabilität
- Informationen für Nachnutzer verständlich?
 - Benennung
 - Formate
 - Dokumentation



Quelle: Pixabay

- Ziel von Data Sharing ist Nachnutzung
- Wird Inhalt eines nächsten Workshops



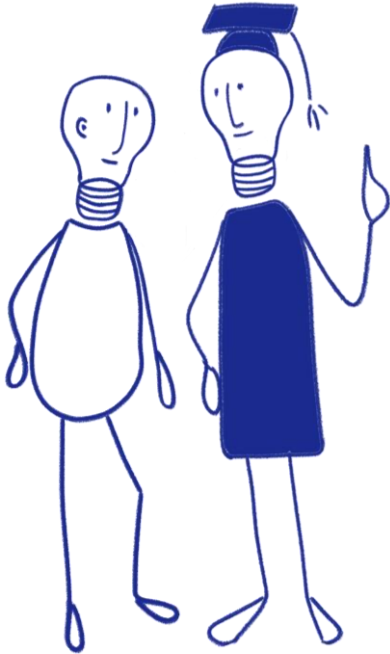
Q&A

Gibt es inhaltliche Fragen?

- aus dem Chat
- Wortmeldungen



Dokumentation des Workshops



Alle Materialien unter:

<https://www.forschungsdaten.org/index.php/Online-Workshops>

- Lehdrehbuch
- Präsentationsfolien
- Ergebnisse:
 - PollEv-Umfrage
 - Mentimeter-Umfrage
 - Break-Out-Diskussion im Padlet



Feedback

Dieser virtuelle Workshop war...

- A. grandios - so sollten Workshops immer sein.
- B. hilfreich, ich habe viel gelernt.
- C. so lala, könnte man besser machen.
- D. nicht hilfreich - ich habe mehr erwartet.



Feedback: Methodik

Wie bewertest du den Einsatz der Umfragen in diesem Online-Workshop?

A. Überflüssig

B. Angemessen

C. Ich habe dazu keine Meinung



Feedback: Methodik

Wie bewertest du den Einsatz der Breakout-Sessions in diesem Online-Workshop?

- A. Überflüssig
- B. Angemessen
- C. Ich habe dazu keine Meinung





Feedback: Methodik

Wie bewertest du den Einsatz von Padlet in diesem Online-Workshop?

A. Überflüssig

B. Angemessen

C. Ich habe dazu keine Meinung





Esther Asef (FU Berlin)

Katarzyna Biernacka (HU Berlin)

Elisabeth Böker (KIM Konstanz)

Sarah Ann Danker (Uni Duisburg-Essen)

Juliane Jacob (Uni Hamburg)

Janna Neumann (TIB)

Britta Petersen (CAU Kiel)

Jessica Rex (TU Ilmenau)

Ute Trautwein-Bruns (RWTH Aachen University)

https://www.forschungsdaten.org/index.php/UAG_Schulungen/Fortbildungen