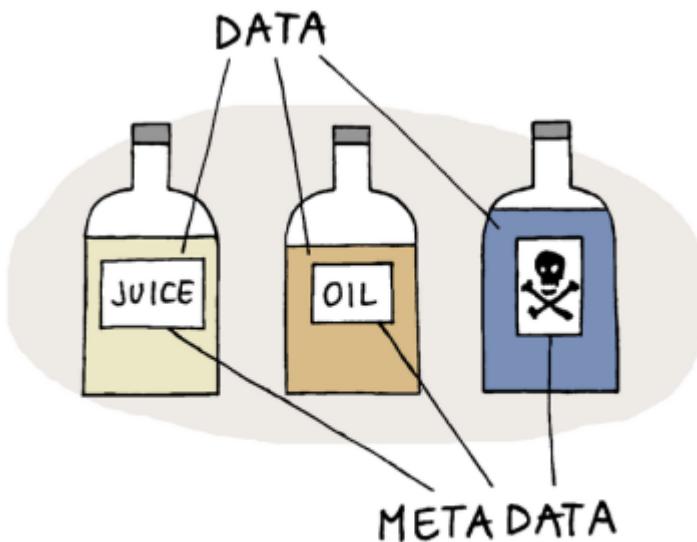


Metadaten

Was sind Metadaten?



 Dataedo /cartoon

Piotr Datnado

Quelle: <https://dataedo.com/cartoon/data-vs-metadata-4>



Bei Metadaten handelt sich um „strukturierte Informationen über andere Daten bzw. Ressourcen und deren Merkmale“ (siehe <https://forschungsdaten.info/praxis-kompakt/glossar/>).

Als Ressourcen sind hier v.a. Forschungsdaten gemeint, aber auch für Publikationen, Poster, Abstracts und weitere wissenschaftliche Veröffentlichungen spielen Metadaten eine bedeutende Rolle. Metadaten erleichtern das Auffinden in Suchmaschinen und Repositorien und das Instituts- und disziplinübergreifende Verständnis vorliegender Forschungsdaten und machen ihre Interpretation und eine spätere Nachnutzung durch andere Personen überhaupt erst möglich. Vor allem in großen mehrjährigen Forschungsprojekten mit verschiedenen Mitarbeitenden erleichtert das einheitliche Verwenden von Metadaten das Verständnis über die erstellten und gespeicherten Forschungsdaten enorm. Sie bieten Kontext und Zusatzinformationen zu vorhandenen Daten und sind damit im weitesten Sinne „**Daten über Daten**“.

Die Vergabe von Metadaten ist ein Baustein zur Umsetzung der **FAIR-Prinzipien** bei der Veröffentlichung von Forschungsdaten. Die FAIR-Prinzipien sollen die Auffindbarkeit, die Zugänglichkeit, die Interoperabilität und die Wiederverwendbarkeit von Forschungsdaten sicherstellen.

Metadaten sollen zumindest bei den folgenden Fragen (analog den **5W1H-Fragen**) helfen, Antworten zu liefern:



- Wer hat die Daten erstellt?
- Was ist der Inhalt der Daten?
- Wann wurden die Daten erstellt?
- Wo wurden die Daten erstellt (Koordinaten)?
- Wie wurden die Daten erstellt?
- Warum wurden die Daten erstellt?

Arten von Metadaten

Es gibt verschiedene Arten von Metadaten. Dabei wird zwischen bibliographischen und administrativen für alle Formen von Forschungsdaten unterschieden, hinzu kommen je nach Fachdisziplin noch Prozessmetadaten sowie inhaltsbeschreibende bzw. deskriptive Metadaten (vgl. siehe

<https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadata-und-metadatastandards/>

Arten von Metadaten	Beispiele	Ziele
Bibliografisch	Titel, Autor_innen/Mitwirkende, Abstract, Keywords, Sprache, Identifier [Link zum Wiki-Artikel]	Ermöglichen die Zitation der Daten, Dienen der Auffindbarkeit und thematischen Eingrenzung
Administrativ	Datum/Zeitpunkt, Ort/Koordinaten, Nutzungsrechte/Lizenzen, Zugriffsrecht, Dateiformat, Dateigröße	Helfen bei der Verwaltung der Daten und deren langfristiger Erhaltung
Prozessmetadaten	Projekt, Methodik, Arbeitsschritte, Hilfsmittel, Geräte/Instrumente, Anwendungsprogramme	Aufzeigen der verwendeten Methoden und Hilfsmittel bei der Entstehung und Verarbeitung der Daten
Inhaltsbeschreibende bzw. deskriptive	Fachbegriffe, domänenspezifisches Wissen, Kontextinformationen	Disziplinspezifische Auskunft über zusätzliche Informationen zu Inhalt und Entstehung der Daten



Tipp: Vergeben Sie Metadaten am besten direkt beim Entstehen der Forschungsdaten. Je größer das Forschungsvorhaben ist und je komplexer und variantenreicher die anfallenden Daten sind, desto schwieriger wird es, dies im Nachhinein zu rekonstruieren.

Hilfsmittel bei der Vergabe von Metadaten

Es gibt verschiedene Möglichkeiten Metadaten zu erfassen und im Zusammenhang mit den Daten zu speichern:

- Metadaten in den Daten (z.B. im Header von Dateien)
- Metadaten an den Daten (z.B. im Dateinamen)
- Metadaten bei den Daten (z.B. als eigene ReadMe-Datei)
- Metadaten in einer Datenbank oder einem Wiki
- Verknüpfung zwischen Metadaten mittels Identifier (z.B. DOI) und den Daten

Da die meisten der genannten Möglichkeiten nur begrenzte Informationen liefern können, ist ein sinnvolles Zusammenspiel zu empfehlen. Redundante Angaben können dabei nicht vermieden werden, dienen aber dem Abgleich und dürfen sich nicht widersprechen.

Werden Metadaten nicht nur menschen-, sondern auch **maschinenlesbar** erstellt, können diese Informationen mittels Schnittstellen über verschiedene Systeme hinweg verknüpft und ausgetauscht werden und erfahren dadurch eine größere Verbreitung. Dabei können Metadaten unabhängig oder zusammen mit den eigentlichen Forschungsdaten gespeichert werden. Zum Beispiel ist das Erstellen einer **ReadMe-Datei** [Link zum Wiki-Artikel] sinnvoll.

Durch die Nutzung von **Tools** zur (automatischen) Vergabe von Metadaten kann die Arbeit erheblich erleichtert und die Qualität der Dateneinreichung verbessert werden. Für generische Metadaten eignen sich folgende Services:

- **DataCite Metadatengenerator**: Hier werden Metadaten aus dem Metadatenstandard [Link nach unten] von **DataCite** erstellt. Das Ergebnis ist eine XML-Datei mit interoperablen Metadaten. Diese Datei kann zusammen mit den Forschungsdaten und evtl. weiteren Dokumenten gespeichert und veröffentlicht werden.

Dieses Videotutorium von FDM Bayern (2020) zeigt die Beschreibung von Forschungsdaten durch die Nutzung des DataCite-Metadatengenerators und des **DataCite-Best-Practice-Guides**:
<https://youtu.be/y7Xullpa6gk>

- **CodeMeta generator** der **CodeMeta Initiative**
- weitere Tools für die Metadatenerstellung listet die Research Data Alliance auf:
<https://rdamsc.bath.ac.uk/tool-index>

Wozu dienen Metadaten?

- Verwaltung und Organisation von Forschungsdaten sowie Identifizierung, Gruppierung und Kategorisierung von Daten
- Auffindbarkeit: effiziente Recherche nach Forschungsdaten sowie innerhalb größerer Datenmengen
- Nachvollziehbarkeit/Reproduzierbarkeit: aufgrund der Kenntnisse über Prozess der Datenerhebung, -verarbeitung und -analyse
- Nachnutzbarkeit: längerfristige Weiter- und Wiederverwendung von Forschungsdaten aufgrund von Kontext- und Zugriffsmöglichkeiten (inkl. rechtlicher Bedingungen)
- Langzeitarchivierung
- Referenzierung der eigenen Forschungsdaten
- Einheitliche Metadaten erleichtern die (interdisziplinäre und internationale) Kooperation
- Vergleichbar- und Verknüpfbarkeit von Forschungsdaten aus unterschiedlichen Quellen

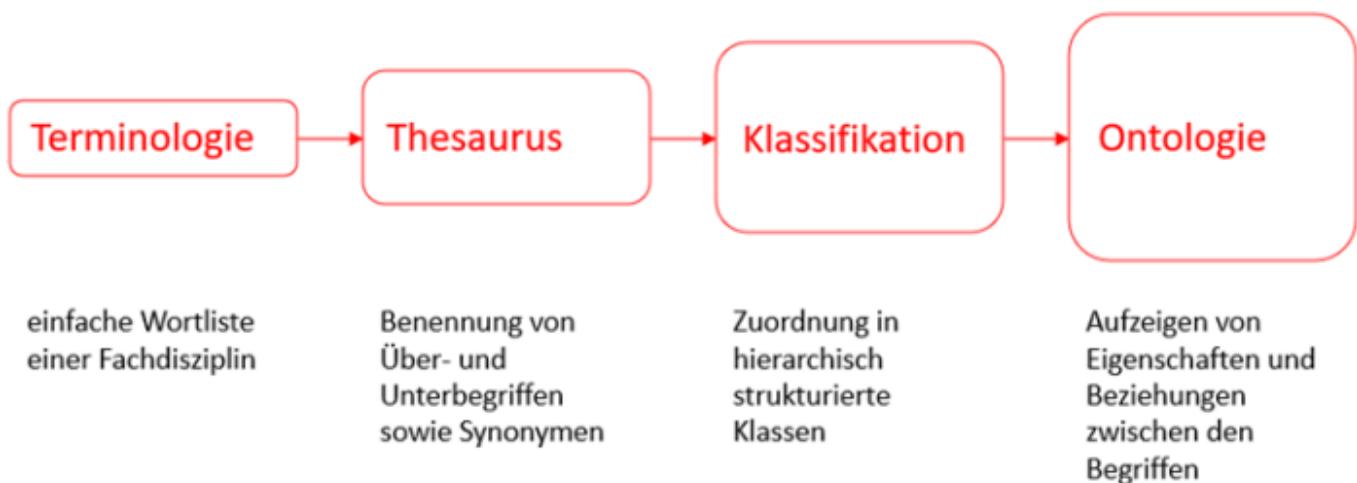




- Metadatenvergabe ist Teil guter wissenschaftlicher Praxis

Was sind Terminologien / Ontologien?

Metadatenschemata geben eine Struktur für Metadaten vor. Sie legen fest, welche Angaben in welchem Format erfasst werden sollen. Terminologien standardisieren den Inhalt bestimmter Metadatenfelder, d.h. sie bilden Wortlisten, aus denen man wählen kann. So können zum Beispiel unterschiedliche (oder auch fehlerhafte) Schreibweisen vereinheitlicht werden. Sie eignen sich gut, um innerhalb eines Forschungsprojektes die Verwendung von Begriffen festzulegen und damit einheitlich zu verwenden. Dabei kann die inhaltliche Breite und Tiefe solcher kontrollierten Vokabulare sehr unterschiedlich sein:



Ontologien sind v.a. relevant bei der maschinellen Verarbeitung großer Datenmengen mittels KI.

Einen Überblick über bestehende Terminologien gibt das Basic Register of Thesauri, Ontologies and Classifications [BARTOC](#).

Weitere Beispiele für kontrollierte Vokabulare:

- [Linked Open Vocabularies \(LOV\)](#) vor allem für den ingenieurwissenschaftlichen Bereich
- [Umweltklassifikation](#) und der [Umweltthesaurus](#) vom Umweltbundesamt
- Thesaurus für die Sozialwissenschaften von GESIS: <https://lod.gesis.org/thesoz/de/>
- Physics Subject Headings (PhySH): <https://physh.aps.org/>
- Mathematics Subject Classification (MSC): <https://mathscinet.ams.org/msc/msc2010.html>
- Art and Architecture Thesaurus (AAT): <http://www.aat-deutsch.de/>
- Medical Subject Headings (MeSH): <https://www.nlm.nih.gov/mesh/>
- Standard Thesaurus Wirtschaft (STW): <https://zbw.eu/stw/version/latest/about.de.html>
- [Getty Thesaurus of Geographic Names](#) Online
- [Biodiversity Thesaurus](#)

Darüber hinaus ist auch die Verwendung von **Normdaten** zur Beschreibung von Forschungsdaten sinnvoll. Im Unterschied zu den oben genannten Vokabularen, haben Normdaten jeweils eigene IDs, die diese eindeutig referenzierbar machen. Vor allem in der maschinellen Verarbeitung und dem

Austausch von Daten zwischen verschiedenen Systemen können Normdaten eindeutig interpretiert werden.

Beispiel für Normdaten:

- GND – Gemeinsame Normdatei der Deutschen Nationalbibliothek <https://explore.gnd.network/>
- [Geonames](#) für geographische Begriffe
- [ORCID](#) für Personendaten
- [Open Funder Registry](#) für Forschungsförderer und deren geförderte Projekte

Auch die **NFDI** [Link zum Wiki-Artikel] (Nationale Forschungsdateninfrastruktur für Deutschland) hat den Wert von Ontologien und Terminologien erkannt und viele Konsortien bieten eigene Lösungen an bzw. arbeiten derzeit an eigenen Lösungen:

NFDI4Biodiversity: Konsortium für Biodiversität und Umweltdaten: Terminologie-Service (teilweise noch Work in Progress) <https://terminologies.gfbio.org/>

NFDI4Cat: NFDI für Wissenschaften rund um Katalyse: Ontologie Sammlung <https://www.nfdi4cat.org/services/ontologie-sammlung/> + (Meta-)Datenstandards basierend auf Use Cases in Planung

NFDI4Chem: Chemistry Consortium in the NFDI: Terminology Service <https://terminology.nfdi4chem.de/ts/>

NFDI4Culture: Konsortium für Forschungsdaten zu materiellem und immateriellem kulturellem Erbe: Antilope zur Terminologie-Suche <https://service.tib.eu/annotation/>

NFDI4DataScience: NFDI for Data Science and Artificial Intelligence: Nutzung von ORKG <https://orkg.org/>

NFDI4Energy: NFDI für die interdisziplinäre Energiesystemforschung: Metadatenregister für Digitale Objekte in Planung

NFDI4Ing: NFDI für Ingenieurwissenschaften: Ontologie-Sammlung <https://terminology.nfdi4ing.de/ts/>

NFDI4Mat: NFDI für Materialwissenschaft und Werkstoffkunde: Ontologie-Service in Planung

NFDI4Microbiota: Konsortium für Microbiota-Forschung: (Meta-)Datenstandards und Terminologie-Service in Planung

BERD@NFDI: NFDI für Betriebswirtschaftslehre, Volkswirtschaftslehre und verwandte Daten: Ontologie-Dienst für Daten deutscher Unternehmen (Work in Progress)

FAIRagro: FAIRe Dateninfrastruktur für die Agrosystemforschung: Terminologie-Service in Planung

Text+: Konsortium für text- und sprachbasierte Forschungsdaten: föderierte Metadaten- und Dateninfrastruktur in Planung

Quelle: NFDI-N(HAW)igator, Stand: 03/2025

Empfehlungen

Folgende Grundsätze sind beim Thema Metadaten zu empfehlen:

1. Vergeben Sie ausreichend Metadaten, v.a. Identifier [Link zum Wiki-Artikel].
2. Nutzen Sie vor allem einheitliche Metadaten aus einem Metadatenstandard.
3. Nutzen Sie automatische Tools zur Metadatenvergabe.
4. Nutzen Sie möglichst einen fachspezifischen Metadatenstandard.
5. Vergeben Sie so viele deutsche und englische Metadaten wie möglich, um eine gute Auffindbarkeit zu gewährleisten – auch über die notwendigen Angaben in Repositorien oder Elektronischen Laborbüchern hinaus.
6. Vergeben Sie Metadaten an allen zur Verfügung stehenden Orten (unter Umständen redundant, aber nicht widersprüchlich); z.B. in den Forschungsdaten selbst, in einer ReadMe-Datei [Link zum Wiki-Artikel] oder auf der Landing Page der Datensätze.
7. Erstellen Sie maschinenlesbare Metadaten, z.B. durch eine ReadMe-Datei [Link zum Wiki-Artikel].
8. Achten Sie auf die verschiedenen Ebenen bei der Erfassung von Metadaten, also die zutreffenden Angaben für die einzelne Datei oder für einen größeren Datensatz bzw. Gruppe von Dateien oder für ein Forschungsprojekt im Ganzen.
9. Verwenden Sie möglichst auch kontrollierte Vokabulare und Normdaten zur Referenzierung Ihrer Forschungsdaten.



Siehe auch den Entscheidungsbaum zu Metadaten vom Projekt [FDM-ndsHAW](#): „Schwere Entscheidung – Wie gehe ich mit Metadaten um?“ <https://zenodo.org/records/15396322>

Best Practice

Best-Practice-Beispiel aus dem Forschungsalltag zur Verwaltung von Daten und Metadaten: [Interview mit Dipl.-Inf. Colin Fischer, Datenmanager im Graduiertenkolleg „Integrität und Kollaboration in dynamischen Sensornetzen“ \(i.c.sens\) im Rahmen der Reihe Data Champions an der Gottfried Wilhelm Leibniz Universität Hannover \(LUH\)](#)

Weitere Links

- Einführung zum Thema Metadaten von FDM Bayern: <https://youtu.be/4HJENeUY4Uc>
- spezielle Suchmaschine für Metadaten [Open Archive Initiative](#): findet standardisierte Metadaten in Metadatensammlungen und damit die dazugehörigen Forschungsdaten
- Top 10 FAIR Data & Software Things – Sammlung eigenständiger Leitfäden für FAIRe Daten und Software in verschiedenen Fachdisziplinen. <https://librarycarpentry.org/Top-10-FAIR/>
- Übersicht verschiedener Rollen von Mitwirkenden an einer Publikation: [CRediT – Contributor Role Taxonomy](#)
- Handbuch zur Erstellung diskriminierungsfreier Metadaten: <https://maehr.github.io/diskriminierungsfreie-metadaten/#sec-Schritt-f%C3%BCr-Schritt-Anleitu>

ng (inkl. einer Checkliste zur Genauigkeit, Vollständigkeit, Konsistenz, Interoperabilität, Inklusivität und ethnischen Überlegungen)

Quellen

- Biernacka, K., Dockhorn, R., Engelhardt, C., Helbig, K., Jacob, J., Kalová, T., Karsten, A., Meier, K., Mühlichen, A., Neumann, J., Petersen, B., Slowig, B., Trautwein-Bruns, U., Wilbrandt, J., & Wiljes, C. (2023). Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement (Version 5). Zenodo. <https://doi.org/10.5281/zenodo.10122153>
- Blümm, M., Fritsch, K., Bock, S., Hackenbuchner, J., Arning, U., & Förstner, K. U. (2024). 06_LE_Metadaten. FDM@Studium.nrw Blended-Learning-Basiskurs „Forschungsdatenmanagement“ (Version 1.0). https://landesinitiativefdmnrw.github.io/FDMatStudium/thk/texte/06_LE_Metadaten.html
- Malzer, Christian: Metadaten. Warum und wie beschreibe ich meine Forschungsdaten? Online-Vortrag zur Love Data Week 2024 am 14.02.2024
- Webseite eRA Göttingen: <https://www.eresearch.uni-goettingen.de/de/knowledge-base/explain-data/explain-metadata/>
- Webseite FDM-ndsHAW: <https://fdm-nds-haw.hawk.de/de/de/fdm-ressourcen>
- Webseite Forschungsdaten.info: <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadata-standards/> und <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadata-im-forschungsalltag/>
- Webseite Friedrich-Schiller-Universität Jena: <https://www.zedif.uni-jena.de/2596/metadata>
- Webseite Leibniz Universität Hannover <https://www.fdm.uni-hannover.de/de/tools>
- Webseite RDM-Compas: <https://rdm-compas.org/articles/was-sind-metadaten>

From:
<https://forschung-wiki.jade-hs.de/> - **Forschung-Wiki**



Permanent link:
<https://forschung-wiki.jade-hs.de/de/fdm/metadata/start?rev=1751454957>

Last update: **2025/07/02 11:15**