

Forschungsdatenmanagement

Von der Erhebung bis zur Archivierung



FDM Kompetenzzentrum
Heinrich-Heine-Universität Düsseldorf





► FORSCHUNGSDATEN

sind alle Daten, die im Zuge eines wissenschaftlichen Vorhabens z. B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen (DFG 2009).

► FORSCHUNGSDATENMANAGEMENT

FDM ist das systematische Bemühen, diese Daten über den gesamten Datenlebenszyklus hinweg nutzbar zu machen. Dieser reicht von der Erhebung über die Auswertung, (Weiter-)Verarbeitung hin zur Archivierung und ggf. Veröffentlichung der Daten.

► FÖRDERMITTEL

Projekte beruhen auf Zusammenarbeit und der zentrale Zugriff auf Daten vereinfacht die Zusammenarbeit in Kooperationsprojekten.

Zur Bewilligung von Drittmitteln wird der Nachweis eines strukturierten Datenmanagements immer häufiger nachgefragt – haben Sie eine Antwort darauf parat!

► VORTEILE: FDM ...

- verringert die Gefahr von Datenverlust
- macht Daten verfü- und nachnutzbar
- bringt Ordnung in die „Datenflut“
- fördert die Umsetzung ethischer Standards und Grundsätze der Guten Wissenschaftlichen Praxis
- schafft Rechtssicherheit
- verbessert den Datenaustausch innerhalb von Forschungsgruppen (z. B. Datenübergabe zwischen Doktorandengenerationen)
- unterstützt Sie

„Qualitätsgesicherte Forschungsdaten bilden einen Grundpfeiler wissenschaftlicher Erkenntnis und können [...] Grundlage weiterer Forschung sein. Die nachhaltige Sicherung und Bereitstellung von Forschungsdaten dient daher nicht nur der Prüfung früherer Ergebnisse, sondern in hohem Maße auch der Erzielung künftiger Ergebnisse.“

Aus der Präambel der „Grundsätze zum Umgang mit Forschungsdaten“ der Allianz der deutschen Wissenschaftsorganisationen, 2010



SIND IHRE FORSCHUNGSDATEN FAIR?



Findable

Ihre Forschungsdaten sind auffindbar und zitierfähig.



Accessible

Ihre Forschungsdaten sind zugänglich.



Interoperable

Ihre Forschungsdaten sind technisch nachnutzbar (Formate, Software).



Re-usable

Ihre Daten sind verständlich und können wiederverwendet werden.

... weitere Infos auf [forschungsdaten.info](https://www.forschungsdaten.info)

Das FDM Kompetenzzentrum der HHU

... unterstützt Sie in allen Belangen rund ums Forschungsdatenmanagement mit folgenden Services:

BERATUNG ZU

- Antragsstellung
- Datenorganisation
- Datennutzung
- Datenpublikation

TOOLS

- Software für den praktischen Umgang mit Daten
- Rat und Tat im Umgang mit Daten

SCHULUNGEN

- Kurse für Data Management
- FAIR-Prinzipien für Daten
- Datenpublikationen

Datenmanagementpläne

Den Umgang mit Daten planen, strukturieren und abstimmen



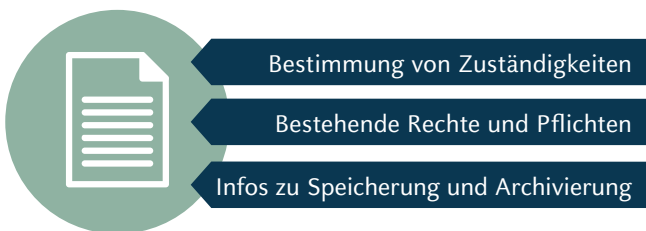
DMP an der HHU

Online ausfüllen – ab in den Antrag!

Was ist ein Datenmanagementplan (DMP)?

Ein DMP unterstützt Sie bei der Strukturierung des Umgangs mit Forschungsdaten, persönlich und für Projekte. DMP dienen als Checklisten und zur laufenden Dokumentation: von der Erhebung bis zur langfristigen Speicherung oder Veröffentlichung der Daten. Einige Förderprogramme der EU oder des BMBF verlangen einen DMP als Teil eines Antrags – haben Sie schon einen DMP erstellt?

Welche Angaben enthält ein DMP?



Legen Sie hier Ihren DMP an:
<https://rdmo.hhu.de>

VORTEILE

- Sie strukturieren Absprachen zum Datenaustausch bei Kooperationen
- Sie dokumentieren für die Berichtspflichten
- Sie planen die Nachnutzung der eigenen Daten
- Sie behalten den Überblick und vermeiden Datenverlust

Speichermedien

Wohin mit den Daten?



Margaret Hamilton neben dem ausgedruckten Quellcode der Software, die sie für die Apollo Mission entwickelt hat, 1969

Speichermedien

Kein Back-up – kein Mitleid

Drehende Festplatten sind nicht für die Ewigkeit gebaut. Manchmal halten sie nicht einmal lange genug, um entsprechend der Guten Wissenschaftlichen Praxis Forschungsdaten für zehn Jahre aufzubewahren.



LEBENSDAUER VON SPEICHERMEDIEN

Festplatten: 2 bis 10 Jahre

DVD: bis zu 30 Jahre

USB-Sticks: 10 bis 30 Jahre

3 ... 2 ... 1 ... Back-up!

Mit Back-ups beugen Sie diesem Desaster vor, aber wo und wie? Clouds sind praktisch, allerdings gibt es bei der Nutzung auch Unklarheiten: Wo liegen die Daten oder was passiert, wenn der Anbieter gehackt wird oder gar den Dienst einstellt? Sicherheit schafft die Wahl einer zuverlässigen Plattform (z. B. Sciebo) und einer zusätzlichen Speicherethode (z. B. die Server der Hochschule).

3-2-1-REGEL

Mindestens **3** Datenkopien auf **2** verschiedenen Speichermedien und **1** Back-up-Kopie an einem externen Standort.

Und langfristig?

Festplatten können verloren gehen, Repositorien nicht. Wertvolle Daten sollten gut dokumentiert, dauerhaft zugänglich abgelegt und gesichert werden. Hierzu eignen sich Repositorien, wie z. B. HHU ResearchData (<https://researchdata.hhu.de>) oder disziplinär orientierte Repositorien.

Persistent Identifiers

Daten und Autoren wiederfinden –
und gefunden werden!



PID und ORCID

Wie funktioniert verlässliche und dauerhafte Auffindbarkeit im Internet?



> *Persistent Identifier (PID)*

Persistent Identifier (PID) sind flexible Wegweiser, die eine unveränderliche Bezeichnung und die flexible Adresse einer Datei miteinander verbinden. Bei einer technisch notwendigen Adressänderung garantiert die PID eine dauerhafte Erreichbarkeit der gewünschten Datei.

Verlagspublikationen werden mit Hilfe von DOI (Digital Object Identifier) referenziert und auffindbar. An der HHU können DOI durch die Universitäts- und Landesbibliothek vergeben werden.



> *ORCID iD*

Mehrfach vorkommende Namen, unterschiedliche Schreibweisen oder eine Namensänderung können dazu führen, dass Autoren nicht eindeutig zuzuordnen sind. Hier hilft ein PID für Autoren – z.B. die Open Researcher and Contributor ID (ORCID iD). ORCID (ORCID.org) bietet eine dauerhafte digitale Kennung, die Sie von allen anderen Forschenden unterscheidet – wie ein Fingerabdruck. Die ORCID iD lässt sich in wichtige Forschungsabläufe wie die Einreichung von Manuskripten und Anträgen integrieren.

Es gibt weitere PIDs und das FDM Kompetenzzentrum informiert Sie gerne darüber.

Datenpublikation

Daten sprechen für sich



Datenpublikation

Sharing is Caring

Teilen lohnt sich für alle Forschenden: Ihre Daten sind vielleicht vielseitig interessant.

Veröffentlichte Forschungsdaten unterstützen nicht nur „die anderen“, sondern auch einen selbst.

VORTEILE

- die eigenen Publikationen mit Forschungsdaten werden deutlich häufiger zitiert
- Datensätze oder Datenpublikationen werden weiterbenutzt und dann auch zitiert
- Forschungsergebnisse werden nachvollziehbar und nachnutzbar
- Ergebnisse lassen sich einfacher vergleichen
- interdisziplinäre Forschung wird unterstützt
- Metaanalysen werden ermöglicht

Daten publizieren, aber wie?

Es ist sinnvoll, Forschungsdaten in Repositorien abzulegen. Denn hier sind sie sicher aufbewahrt und bleiben abrufbar.

- > **Fachliches Repository:** In vielen Fachcommunitys gibt es bereits anerkannte Dienste. <https://re3data.org>
- > **Generisches Repository:** z. B. Zenodo, Radar
- > **Hochschulrepository:**
HHU ResearchData (<https://researchdata.hhu.de>)
- > **Data Journals:** Fokus auf Beschreibung und Methodik der Datengewinnung
- > **Fachzeitschriften:** ergänzend zu Textpublikationen

Metadaten

Wissen, wonach man sucht



Metadatenstandards

Data about Data

Das Etikett auf einer Dose sagt, was drin ist. Genauso sorgen Metadaten dafür, dass digitale Daten und Objekte gefunden und genutzt werden können. Deshalb ist es wichtig, auch die eigenen Forschungsdaten mit „Etiketten“ zu versehen.

Arten von Metadaten

> *Bibliographische bzw. administrative Daten* enthalten Informationen zur Verwaltung der Daten, zur Entstehung der Gesamtheit der Daten und sind eher allgemein und weniger Community-spezifisch.



> *Inhaltsbeschreibende bzw. fachliche Daten* beschreiben einzelne Aspekte oder Datensätze genauer und bieten zusätzliche Informationen. Je nach Disziplin sind sie sehr verschieden aufgebaut. Viele Fächer haben bereits eigene Metadatenstandards.



Elektronische Laborbücher

Das Ende des Kladdenchaos

1. 76 - 77.

8. 46

9. 46

10. 46

11. 46

12. 46

13. 46

14. 46

15. 46

16. 46

17. 46

18. 46

19. 46

20. 46

21. 46

22. 46

23. 46

24. 46

25. 46

26. 46

27. 46

28. 46

Yndkalm:

Li Ba Br

sonstige

Reihe - Serie.

2. 67. 10. 6

R

Reihe	Serie	Yndkalm	Li	Ba	Br	sonstige
17.16.	11 ¹⁸	0 71398				
		2 71562				
		6 71885				
		10 72170				
		16 72610				
		11.06 28 73511				
		12.00 0 74000				
		16.10 10 75317				
		12.47 0 76854				
		15.55 6 77300				
		12.59 12 77576				
		15.15 0 78202				
		6 78557				
		13.25 12 78875				
		16.15 0 82117				
		6 82455				
		16.25 12 82822				
		18.0 0 85455				
		18.28 20 86612				
		19.21 0 88127				
		19.33 12 88862				
		11.41 20 91372				
		19.15 13 92111				
		13.30 15 93115				
		17.20 11 94201				
		11.59 14 97735				
		23.11 11 97454				
		11.51 21 96484				
		21.30 15 980 18980				
		13.52 50 23171				
			104/0,82			
			323/4,81			
			185/4,71			
			440/6,73			
			901/12,75,1			
			658/10,765,8			
			346/6,57,7			
			576/6,62,7			
			355/6,59,2			
			318/6,53,0			
			388/6,56,3			
			367/6,61,2			
			1157/20,57,9			
			735/12,61,3			
			1245/20,62,0			
			1004/13,77,1			
			1534/18,25,2			
			1600/20,81,5			
			419,50,83,8			

Elektronische Laborbücher

Den Workflow im Labor digital vereinfachen

Laborbücher gehören in den Naturwissenschaften zum Forschungsalltag: Hier landen Messergebnisse, Skizzen von Versuchsaufbauten oder Auswertungen. Aber was, wenn diese Daten digital sind? Alles ausdrucken und in die Kladde kleben?

Besser nicht. Elektronische Laborbücher (ELB) schaffen Abhilfe.



WOZU EIN ELB?

- Laborbuch führen mit PC, Handy oder Tablet
- Volltextsuche in sämtlichen Inhalten
- kollaborativ arbeiten
- Datensicherheit, Zugriffskontrolle
- sichert Datenintegrität, schafft Belegbarkeit durch Zeitstempel
- beliebige Dateien importieren oder verknüpfen
- andere Systeme anbinden (API)
- in PDF (und andere Formate) exportieren



Überblick

<https://fdm.hhu.de>

Archivierung von Bild- und audiovisuellen Dateien

Eine Frage des Formats



Archivierung von Bild- und audiovisuellen Dateien

... auch „born digital“

Ohne Abspielgerät sind Kassetten Plastikmüll. Auch digitale Dateiformate entwickeln sich mit der Zeit weiter: Forschungsdaten lassen sich dann im schlimmsten Fall nicht mehr öffnen oder bearbeiten.

Wer zur Archivierung seiner Daten folgende Formate nutzt, ist aber auf der sicheren Seite.

FIT FÜR'S ARCHIV: FORMATE

► Rastergrafiken

- Tagged Image File Format (TIFF) – unkomprimiert > *.tif
- Portable Network Graphics (PNG) > *.png
- JPEG2000 > *.jpg, *.jpeg

► Vektorgrafiken

- Scalable Vector Graphics (SVG) > *.svg, *.svgz

► Computer-Aided Design (CAD)

- AutoCAD Drawing > *.dwg
- Drawing Interchange Format, AutoCAD > *.dxf
- Extensible 3D, X3D > *.x3d, *.x3dv, *.x3db

► Ton, Audio

- Waveform Audio File Format (WAV) – unkomprimiert > *.wav

► Video

- FFV1 Codec in Matroska Container > *.mkv
- Motion JPEG 2000 (ISO / IEC 15444-4) > *.mj2
- AVI - unkomprimiert > *.avi

Personenbezogene Daten

Undercover unterwegs



(Forschungs-)Datenschutz

Alles rechtens?

Ob DSGVO, Ethikleitlinien, Bundes- oder Landesdatenschutzgesetz, schriftliche Einwilligung, Zweckbindung oder „Broad Consent“, Pseudonymisierung oder Anonymisierung: Wer verantwortungsvoll und rechtskonform mit personenbezogenen Forschungsdaten umgehen will, hat viele Vorschriften zu beachten.



WARUM DATENSCHUTZ?

- Einhaltung ethischer Standards
- Vertrauen in die Forschung stärken
- Vorgaben von Fördergebern einhalten
- langfristige Speicherung, Nutzung und/oder Weitergabe auf ein rechtlich solides Fundament stellen



Weitere Informationen erhalten Sie unter <https://hhu.de/datenschutz>