

Kompetenzzentrum Forschungsdaten an der Universität Bielefeld

Kolloquium Wissensinfrastruktur, 22.06.2018

Nils Hachmeister, Dirk Pieper, Maik Stührenberg, Johanna Vompras

Übersicht

- Was sind Forschungsdaten?
- FDM als Faktor für Wissenschaftsförderer
- FDM-Historie an der Universität Bielefeld
- Konzeption eines FDM-Kompetenzzentrums
- Vergleichbare FDM-Zentren in Deutschland
- Aktueller Stand und Ausblick

Was sind Forschungsdaten?

Unter Forschungsdaten werden alle Ergebnisse und Bezugsquellen des Forschungsprozesses (u.a. auch Software oder Forschungsobjekte bzw. Proben) verstanden, die im Projekt erhoben, ausgewertet und/oder entwickelt werden.

Merkblatt für Sonderforschungsbereiche der DFG (DFG-Vordruck 50.06 – 10/17)

Was sind Forschungsdaten?

*Unter Forschungsdaten werden alle Ergebnisse und Bezugsquellen des Forschungsprozesses (u.a. auch **Software** oder ~~Forschungsobjekte bzw. Proben~~) verstanden, die im Projekt erhoben, ausgewertet und/oder entwickelt werden.*

Merkblatt für Sonderforschungsbereiche der DFG (DFG-Vordruck 50.06 – 10/17)

Was sind Forschungsdaten?

*Unter Forschungsdaten werden alle Ergebnisse und Bezugsquellen des Forschungsprozesses (u.a. auch **Software** oder ~~Forschungsobjekte bzw. Proben~~) verstanden, die im Projekt erhoben, ausgewertet und/oder entwickelt werden.*

Merkblatt für Sonderforschungsbereiche der DFG (DFG-Vordruck 50.06 – 10/17)

Forschungsdaten sind Daten, die im Zuge der Forschung gesammelt, beobachtet, simuliert, abgeleitet oder generiert werden.

Leitlinien FDM RWTH Aachen

Übersicht

- Was sind Forschungsdaten?
- **FDM als Faktor für Wissenschaftsförderer**
- FDM-Historie an der Universität Bielefeld
- Konzeption eines FDM-Kompetenzzentrums
- Vergleichbare FDM-Zentren in Deutschland
- Aktueller Stand und Ausblick

FDM und Wissenschaftsförderer

Um die Weitergabefähigkeit der Daten an eine geeignete Einrichtung zu gewährleisten, müssen die Antragstellerinnen und Antragsteller Standards des Forschungsdatenmanagements einhalten. [...] Die Umsetzung des Forschungsdatenmanagements ist in der Skizze darzulegen und wird begutachtet [...].

BMBF Richtlinie zur Förderung von Forschung zu "Digitalisierung im Bildungsbereich – Grundsatzfragen und Gelingensbedingungen".

Bundesanzeiger vom 26.09.2017

<https://www.bmbf.de/foerderungen/bekanntmachung-1420.html>

FDM und Wissenschaftsförderer

[...] In einem Sonderforschungsbereich dient ein Teilprojekt Informationsinfrastruktur vor allem dem systematischen Management der im Kontext des Sonderforschungsbereichs relevanten Daten. [...] Grundsätzlich wird ein professionelles Management der Daten, die im Sonderforschungsbereich erhoben, ausgewertet und/oder entwickelt werden, erwartet. [...] Auch nach Ende der Projektförderung soll ein Zugang zu den Daten oder eine Übergabe an bestehende Infrastrukturen gewährleistet sein.

DFG Merkblatt Sonderforschungsbereiche

http://www.dfg.de/formulare/50_06/50_06_de.pdf

FDM und Wissenschaftsförderer

[...] the new research data resulting from the project should be treated permissible according to the FAIR principles, and deposited and shared, according to the national rules of the countries involved. It is strongly advised to make data accessible through through RD-Connect [...]

DFG E-Rare Call for Proposals 2018: Transnational Research Projects on Hypothesis-driven Use of Multiomic Integrated Approaches for Discovery of Disease Causes and/or Functional Validation in the Context of Rare Diseases

http://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_17_87/index.html

FDM und Wissenschaftsförderer

Beispiel Europäische Kommission:

- [Guidelines on FAIR Data Management in Horizon 2020](#)
- [Guideline Open Access to publications and research data](#)

Open access to research data

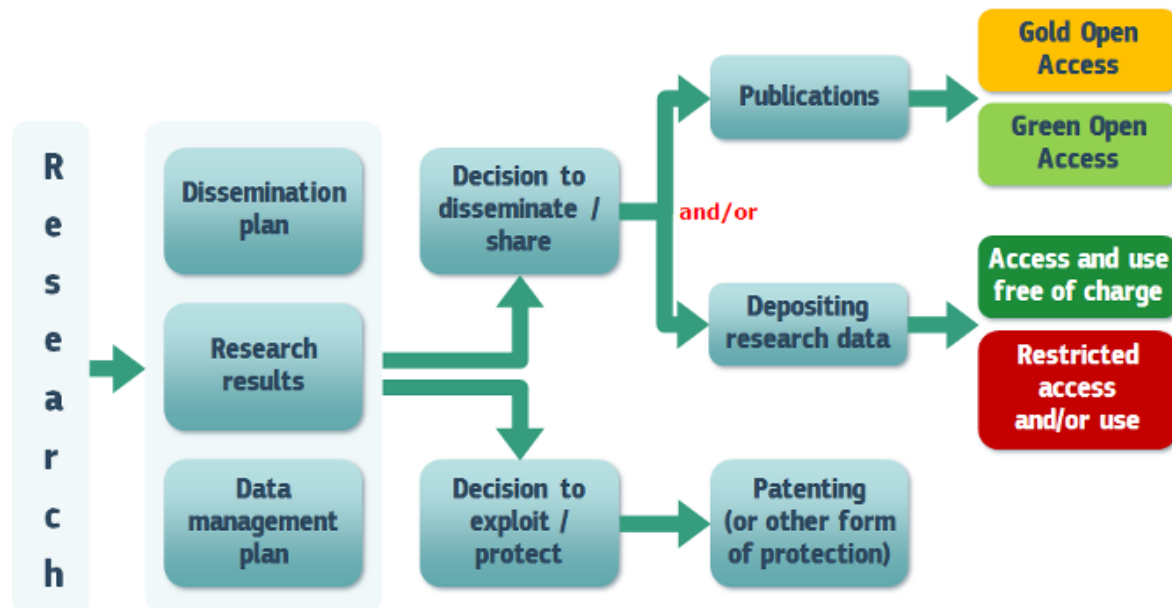
Refers to the right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement.

Research data

Refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.

In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

Users can normally access, mine, exploit, reproduce and disseminate openly accessible research data free of charge.



FDM und Wissenschaftsförderer

Beispiel European Research Council:

The European Research Council supports the basic principle of Open Access to research data. It therefore recommends to all its funded researchers that they follow best practice by retaining files of all the research data they have produced and used during the course of their work, and that they be prepared to share these data with other researchers whenever they are not bound by copyright restrictions, confidentiality requirements, or contractual clauses.

[Guidelines on the Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020](#)

Übersicht

- Was sind Forschungsdaten?
- FDM als Faktor für Wissenschaftsförderer
- **FDM-Historie an der Universität Bielefeld**
- Konzeption eines FDM-Kompetenzzentrums
- Vergleichbare FDM-Zentren in Deutschland
- Aktueller Stand und Ausblick

FDM-Historie an der Universität Bielefeld

2009

Informium

2011

Grundsätze

2013

Resolution

2016

Kontaktstelle

FDM-Historie an der Universität Bielefeld

- Informium (2009): Erste Förderung des Rektorats für den Aufbau hochschulweiter FDM-Services
- SFB 882 (2011-2015): INF Teilprojekt zum Aufbau einer virtuellen Forschungsumgebung inkl. zentraler Servicefunktionen für den FDM-Bereich
- FDM-Aktivitäten der wissenschaftlichen Einrichtungen CEBITEC (de.NBI) und CITEC (Open Science)
- CONQUAIRE
- Weitere FDM-Stelle für die Medizinische Fakultät

Von Informium zur Kontaktstelle FD

- Data Management Planning
- Rechtliche Aspekte des FDM
- Verankerung in Forschung und Lehre, Schulung
- Publikation von Forschungsdaten
- Vernetzung
- Unterstützung bei Drittmittelanträgen (z.B. SFB INF Teilprojekte u.a.)

FDM-Historie an der Universität Bielefeld



Übersicht

- Was sind Forschungsdaten?
- FDM als Faktor für Wissenschaftsförderer
- FDM-Historie an der Universität Bielefeld
- **Konzeption eines FDM-Kompetenzzentrums**
- Vergleichbare FDM-Zentren in Deutschland
- Aktueller Stand und Ausblick

Überlegung zur Einrichtung eines FDM-Zentrums

- Kooperation zwischen UB, BiCDaS und BITS
- Ziel: Verstetigung der bisherigen Aktivitäten
- Vorgehensweise:
 - Darlegung der bisherigen Arbeiten
 - Bedarfserhebung
 - Konzeption eines FDM-Zentrums

Ermittlung der FDM-Bedarfe

- Bedarfserhebung Neuropsychologie-Server
- Umfrage Campus OWL
- Analyse der in PUB veröffentlichten Forschungsdatensätze
- Externe FDM-Bedarfserhebungen

NΨ-Server

- Anforderungsprofil für einen FDM-Server für neurophysiologische Daten (EEG, fMRT) 2016/17
→ beispielhafte fachspezifische FDM-Anforderungen
- Genannte formulierte Anforderungen:
 - Rechte-Verwaltung (Nutzer/Gruppen – Rollen/Rechte) mit Anbindung an zentrale IDM-Systeme
 - Standardisierte Metadaten zur Beschreibung von FD
 - Aggregation und Versionierung von Datensätzen
 - Vorverarbeitung von Datensätzen
 - Konvertierung in verschiedene Dateiformate

Umfrage Campus OWL

- Umfrage unter den OWL-Hochschulen 2017 zur Nutzung von Forschungsdaten
- Beteiligte Universitäten
 - Hochschule für Musik (Detmold)
 - Hochschule OWL
 - FH Bielefeld
 - Universität Bielefeld
 - Universität Paderborn

Umfrage Campus OWL

Ergebnisse:

- Heterogenität der verwendeten Datenformate – von Standard-Dateitypen bis hin zu gerätespezifischen Rohdaten (z. B. Sensorik)
- Heterogenität der Daten bzgl. Art – von Text über audio-visuelle Daten (z. B. Interviews), Skizzen, Simulationen, Datenbanken
- Varianz des Datenvolumens bzgl. Kapazität und Handling (kein kollaboratives Arbeiten)
- Große Spanne von freien Daten bis hin zu personenbezogenen Daten mit hohem Schutzbedarf

Analyse der PUB-Daten

Erhebung der in PUB gespeicherten
Forschungsdatensätze 12/2017

- 205 Einträge, 1393 Dateien → insgesamt 1598 Objekte
- PUB-JSON-Export wurde nach XML konvertiert und anschließend per XSLT analysiert
- Insgesamt sehr viel disziplinbezogene FD → flexible Infrastruktur bzgl. Datentyp, Dateigröße, Metadaten

Analyse der PUB-Daten

Ergebnisse im Detail

- 1370 Einträge mit Open Access-Zugang, 23 nur auf Nachfrage zugänglich
- 161 Einträge mit Nutzungslizenzen
- 35 unterschiedliche MIME-Types (bei 1125 Datentyp-Angaben insgesamt) unterstreichen die Heterogenität der Datentypen – vom Office-Dokument bis hin zur Virtual Appliance (VBox-Instanz)
- Dateigrößen von 89 Byte bis 14,06 GB

Externe FDM-Bedarfserhebungen

Aufgrund zeitlicher Constraints wurde auf vorhandene Untersuchungen rückgegriffen

- Umfragen der HU Berlin und Uni Kiel
- Umfrage TIB

FDM-Bedarfserhebungen – HU Berlin und Uni Kiel

Umfrage im Oktober 2013 an der HU Berlin sowie 2014 an der Uni Kiel

- 24 Fragen mit den Themenbereichen
 - Datentypen
 - Zugang
 - Infrastruktur und Services
- HU Berlin: 123 ProfessorInnen und 267 Wiss. Mitarbeitern
- Uni Kiel: 218 Forschende (ohne Aufschlüsselung nach Statusgruppen)

FDM-Bedarfserhebungen – HU Berlin und Uni Kiel

Ergebnisse

- Während in Berlin größtenteils Texte verarbeitet werden (gefolgt von Datenbanken und Tabellen, Bildern), bilden in Kiel strukturierte Daten (Messwert-Reihen) einen Schwerpunkt
- Dateigrößen größtenteils < 1 TB – Kiel mit 2,5% der Befragten mit Bedarf > 100 TB
- Speicherung der Daten erfolgt oft auf privaten oder dienstlichen Rechnern, seltener auf zentralen Servern
- Forschungsdaten werden auch im Rahmen von Publikationen eingereicht

FDM-Bedarfserhebungen – HU Berlin und Uni Kiel

Ergebnisse

- Bereitschaft zur Ablage von Forschungsdaten in einem Archiv unterschiedlich: 60% wären dazu bereit, allerdings wird auch der Aufwand (Anonymisierung, Datensicherheit, Zugang) kritisch gesehen
→ hoher Bedarf bzgl. Beratung und Infrastruktur
- Unterschiede zwischen der Bevorzugung lokaler Datenarchive vs. disziplinärer internationaler Repositorien
- Backups selten organisiert
- Teilweise unklare Verantwortlichkeit für Daten

FDM-Bedarfserhebungen – TIB Hannover

Quantitative Online-Umfrage im Auftrag der TIB Hannover Ende 2016/Anfang 2017

- Über 1400 Teilnehmer aus natur- und ingenieurwissenschaftlichen Disziplinen sowie Mathematik
- Schwerpunkt auf Nutzung und Entstehung nicht-textueller wissenschaftlicher Materialien

FDM-Bedarfserhebungen – TIB Hannover

Ergebnisse

- Rund 50% der Befragten gaben an, dass Forschungsdaten im Rahmen ihrer Tätigkeit entstehen, bei 30% entsteht (auch) Software
- 45% der FD-Erzeuger nutzen Datenrepositorien oder planen es zumindest
- 70% der Wissenschaftler, die aktuell noch keine FD erzeugen, würden Repositorien nutzen, sofern sie an der Einrichtung verortet sind

Zusammenfassung der Bedarfe

Benötigt werden:

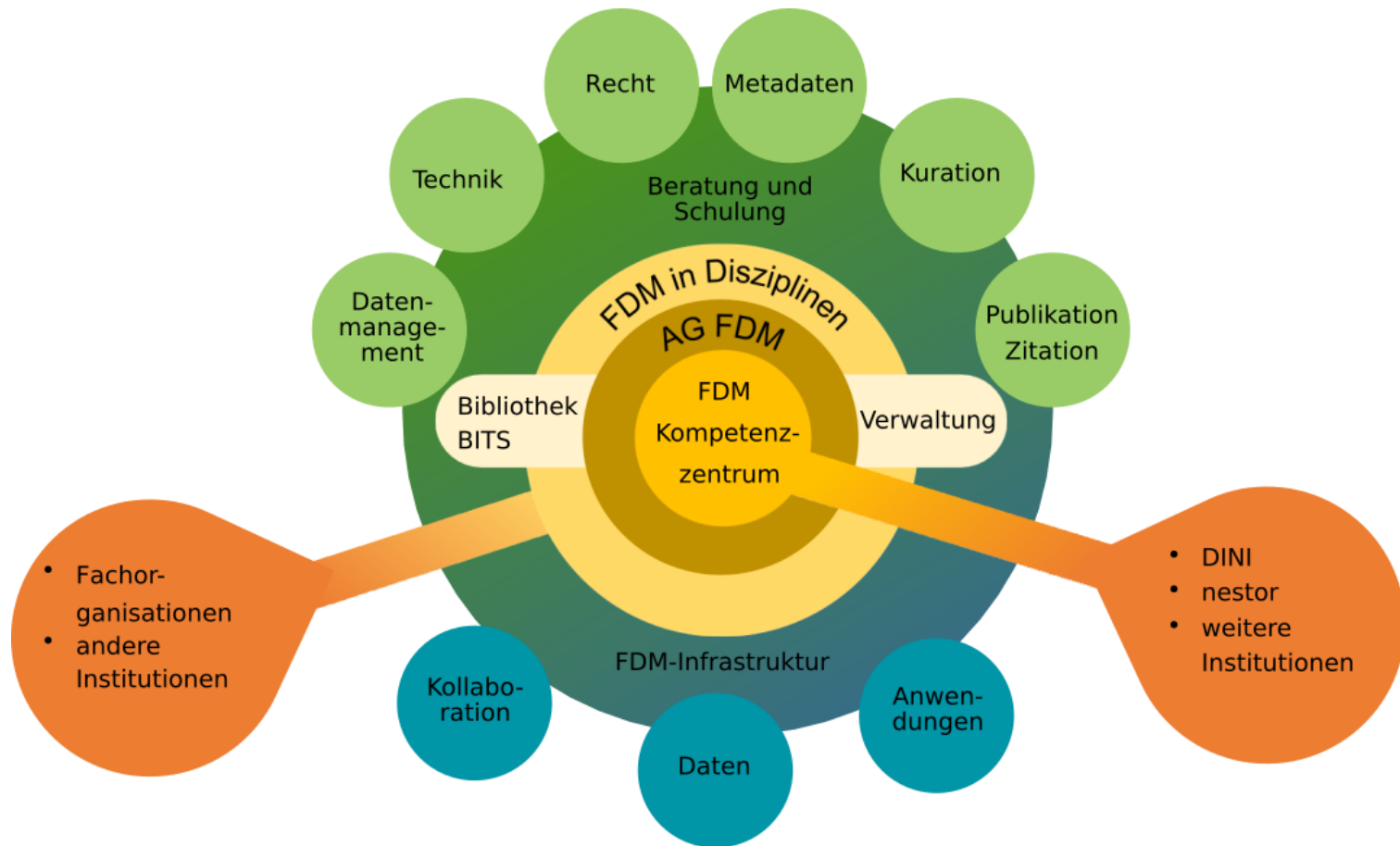
- Infrastruktur/Werkzeuge
 - zur Kollaboration (z. B. GitLab, Sciebo, Wikis) mit Rechte-Management (intern und extern)
 - zum Authoring (wie kollaborativ nutzbare Datenbanken und Arbeitsplattformen)
 - zur Bereitstellung der Daten übers Web
 - zur strukturierten Metadatenerfassung
- Services
 - Schulungen zu diversen Aspekten (technisch, rechtlich, etc.)
 - Langzeitsicherung

Zusammenfassung der Bedarfe

Kein One-fits-all-Ansatz

- Heterogenität der Datenformate und -typen
- Varianz des Speicherbedarfs
- Generische vs. disziplin-bezogene Infrastrukturen

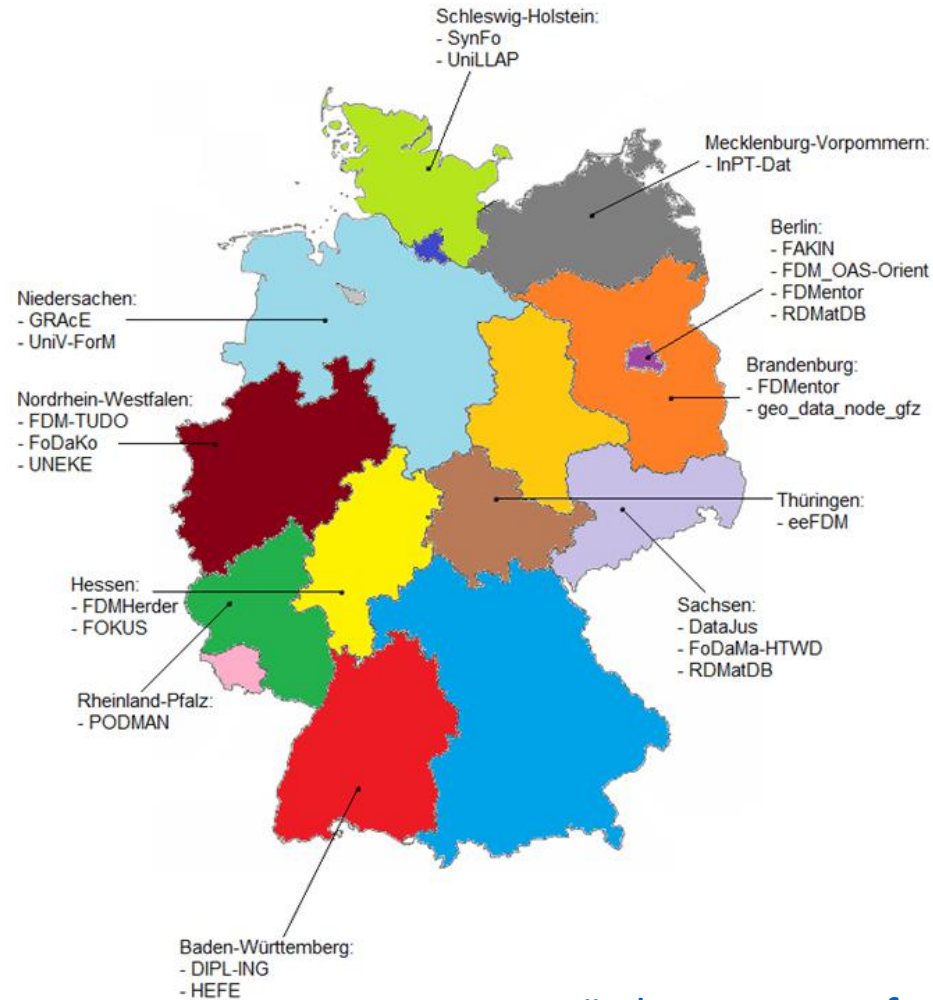
Zusammenfassung der Bedarfe



Übersicht

- Was sind Forschungsdaten?
- FDM als Faktor für Wissenschaftsförderer
- FDM-Historie an der Universität Bielefeld
- Konzeption eines FDM-Kompetenzzentrums
- **Vergleichbare FDM-Zentren in Deutschland**
- Aktueller Stand und Ausblick

FDM-Zentren in Deutschland



[BMBF Förderung zur Erforschung des FDM 2016](#)

FDM-Zentren in Deutschland – Göttingen

Göttingen/ **eResearch Alliance**


[Data Management](#)
[Tools & Services](#)
[Projects & Events](#)
[News](#)
[Blog](#)
[About](#)

Your research project! | Your data! | Our services!

We understand eResearch as *enhanced* research, which to us means an optimized **usage of digital technologies and methods for innovative research**. We offer information, personal advice and support for key issues related to digital research through all phases of the research life cycle:



Ideas

- Project proposal support
- Data management planning
- Expert network

[more](#)


Research

- Workshops & Trainings
- ICT services
- Visualisation & Exploration
- Data strategy implementation

[more](#)


Results

- Persistent Identification
- Data publication
- Long term archiving

[more](#)

News

- 02.08.-16.08.2018: Summer School Data Science
- Registration for the Pre-RDA-Symposium is open now
- 19.03./20.03.2018: Pre-RDA Symposium in Göttingen: The critical role of university RDM infrastructure in transforming data to knowledge
- DARIAH-DE Repository released
- 24.11.2017: Info event on research data management and research services at UMG

Guidelines

- Policies on Research Data and Open Access as "Amtliche Mitteilung" (PDF, German only)
- Research data policy of the Georg-August-University Göttingen (incl. UMG) - English version
- Open Access policy of the Georg-August-University Göttingen (incl. UMG) - English version

FDM-Zentren in Deutschland – Aachen

🏠 ▶ [Forschung](#) ▶ [Forschungsdaten-Management](#) ▶ [Forschungsdaten-management an der RWTH Aachen](#)



Forschungsdatenmanagement
allgemein

▼ Forschungsdaten- management an der RWTH Aachen

Leitlinie zum
Forschungsdaten-
management an der RWTH
Aachen

Best Practices

Unterstützungsangebote für
den Forschungsprozess

Beratung

Weiterbildungsangebote

Archivieren und Publizieren

Forschungsdatenmanagement an der RWTH Aachen

Die RWTH Aachen ist bestrebt, ihre Forschenden auf der Basis verbindlicher Standards und Verfahren praxisnah und individuell beim Forschungsdatenmanagement zu unterstützen.

Dazu gibt es Beratungs-, Schulungs- und Infrastrukturangebote, die ein effizientes, unkompliziertes und rechtlich abgesichertes Arbeiten ermöglichen.

KONTAKT



ServiceDesk Forschungsdaten- management

☎ +49 241 80 24680

✉ [E-Mail schreiben](#)

FDM-Zentren in Deutschland – Hamburg



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

ZENTRUM
FÜR NACHHALTIGES
FORSCHUNGSDATENMANAGEMENT

ÜBER UNS

FORSCHUNGSDATEN

PUBLIKATIONEN

E-RESEARCH

SERVICE

KONTAKT



Zentrum

für nachhaltiges Forschungsdatenmanagement



Foto: K. Helmholz

FDM-Zentren in Deutschland – Hamburg



ZENTRUM
FÜR NACHHALTIGES
FORSCHUNGSDATENMANAGEMENT

ÜBER UNS FORSCHUNGSDATEN PUBLIKATIONEN E-RESEARCH SERVICE KONTAKT



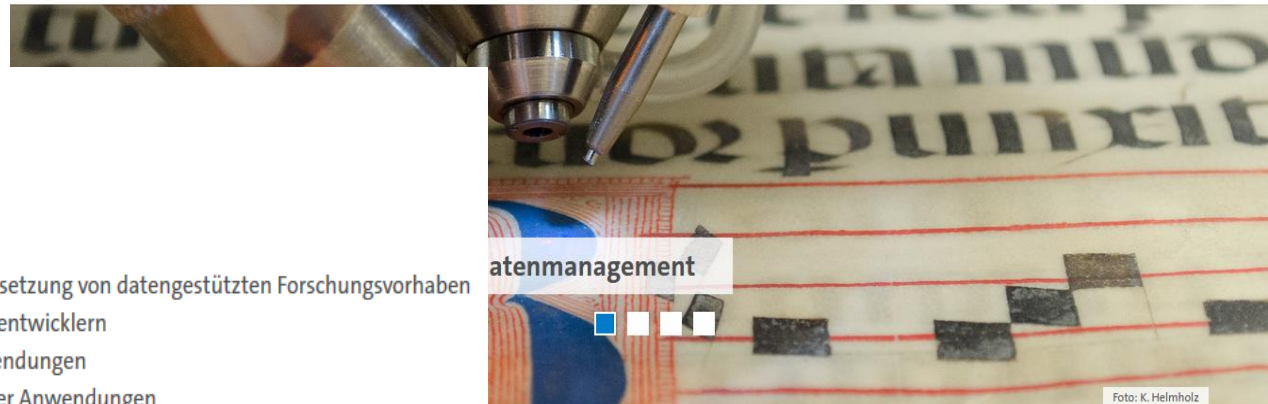
SERVICE

Was wir Ihnen anbieten können:

- Beratung bei der Beantragung und Umsetzung von datengestützten Forschungsvorhaben
- Hilfe bei der Vermittlung von Softwareentwicklern
- Kuration bestehender Daten und Anwendungen
- Beratung bei der Migration vorhandener Anwendungen
- Beratung bei der Erstellung von Metadaten und Datenbeschreibungen
- Kooperation mit dem Regionalen Rechenzentrum (RRZ)
- Kurzpräsentation Ihrer Datenbank, Ihres Labors auf unserer Webseite
- Vermittlung von Kontakten und Kooperationen
- Beratung bei Digitalisierungsfragen

Was wir leider nicht leisten können:

- Übernahme umfangreicher Softwareentwicklungen für Projekte
- Erstellung umfangreicher Datenbeschreibungen bzw. Metadatensätze
- Digitalisierung von Material
- Migration komplexer Altanwendungen



FDM-Zentren in Deutschland – Heidelberg

Universität Heidelberg > Einrichtungen > Kompetenzzentrum Forschungsdaten >

Kompetenzzentrum Forschungsdaten

Das Kompetenzzentrum Forschungsdaten (KFD) bietet Wissenschaftlern der Universität Heidelberg gemäß der [Research Data Policy der Universität](#) zentrale Dienste zur Archivierung und Veröffentlichung von Forschungsdaten an. Das KFD ist eine gemeinsame Serviceeinrichtung des [Universitätsrechenzentrums](#) und der [Universitätsbibliothek](#) Heidelberg.

Unsere Dienste:
Unterstützung an jedem Punkt des
Forschungsprozesses

Projektplanung und
Datenmanagement

Datenbearbeitung
im Projekt

Archivierung
und Veröffentlichung

forschungsdaten.info

FDM-Zentren in Deutschland – Berlin u.a.

Humboldt-Universität Berlin, Freie Universität Berlin,
Technische Universität Berlin, Universität Potsdam,
Europa-Universität Viadrina Frankfurt (Oder)

- Handreichung entwickeln
- Forschungsdaten Policy entwickeln
- Aufbau von Schulungen (Multiplikatoren)
- Projektkoordination



Infrastrukturen – Geistes- und Kulturwissenschaften

Common Language Resources and Technology Infrastructure

- DMPs
- Lehrmaterialien
- Persistente Ids
- Langzeitarchivierung
- Webbasierte Analysetools

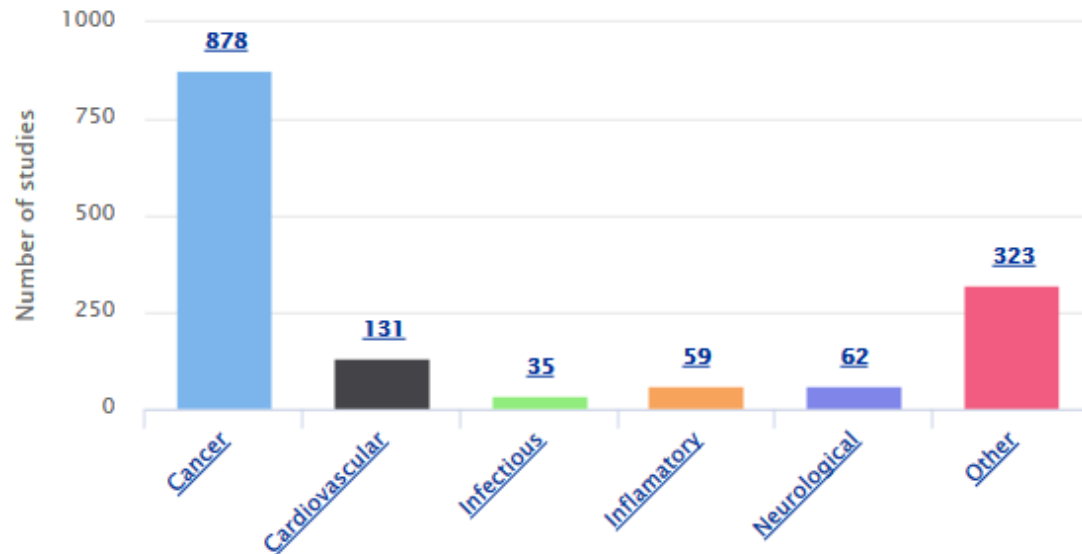


Infrastrukturen – Lebenswissenschaften

What is in the EGA?

Studies in the EGA by disease

Click on a column to view category subgroups



If applicable, a study may be included in more than one category



- Kein freier Zugriff
- Keine sensiblen Daten

Infrastrukturen – Astronomie



LOFAR

Projects of db.lofar.target.rug.nl

Total number of projects 434

Total number of users 645

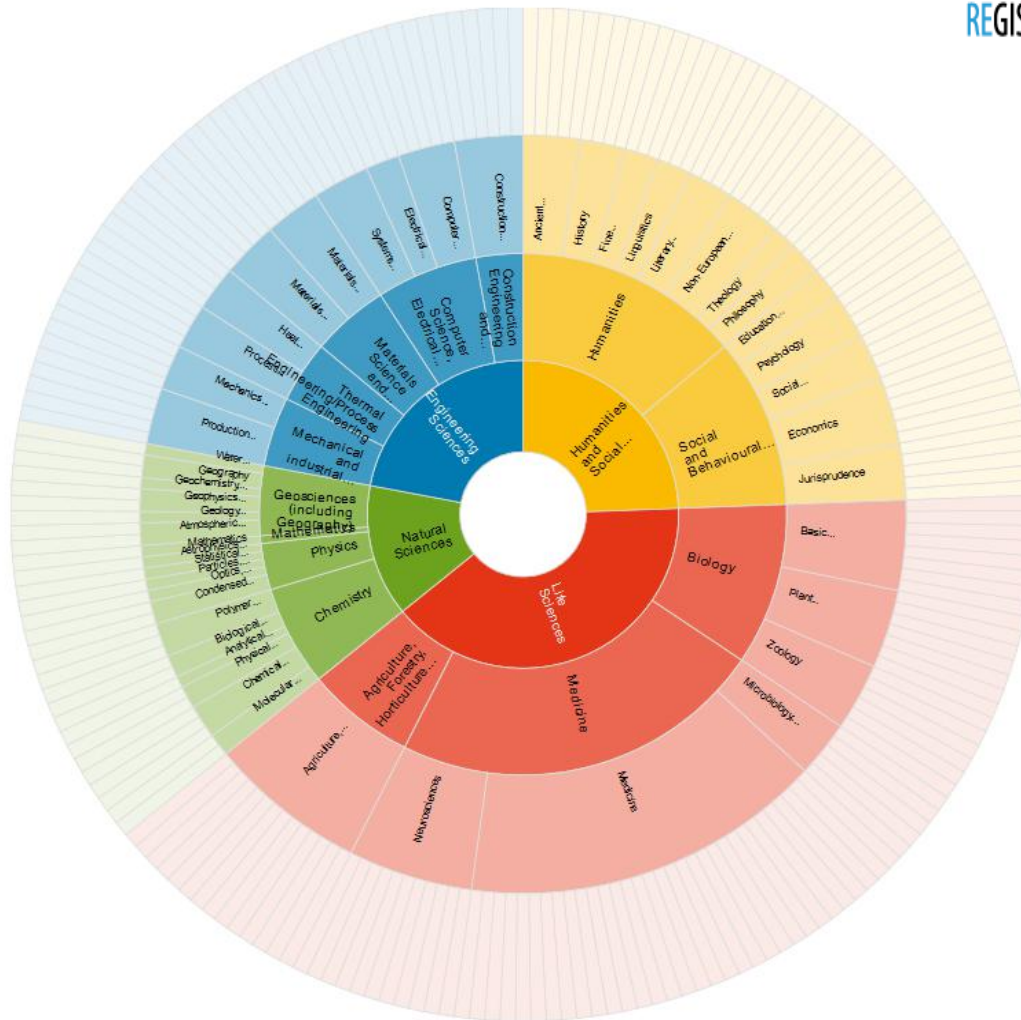
Click on a project name to set the project. Only projects you are a member of are selectable.

The following project groupings are available:

- **All projects**
- All public data
- Commissioning
- Cycle 0
- Cycle 1
- Cycle 10
- Cycle 2
- Cycle 3
- Cycle 4
- Cycle 5
- Cycle 6
- Cycle 7
- Cycle 8
- Cycle 9
- Other projects

Each DataProduct type has two columns. The first column gives the number of DataProducts that can be staged. The second column gives the number of public DataProducts.

Infrastrukturen – re3data



Übersicht

- Was sind Forschungsdaten?
- FDM als Faktor für Wissenschaftsförderer
- FDM-Historie an der Universität Bielefeld
- Konzeption eines FDM-Kompetenzzentrums
- Vergleichbare FDM-Zentren in Deutschland
- **Aktueller Stand und Ausblick**

Aktueller Stand

- Erarbeitung des Konzeptpapiers und Vorstellung im CIO-Gremium 02/2018
- Die Einrichtung eines FDM-Kompetenzzentrums wurde vom Rektorat bewilligt
- Stellenanteil wurde auf 50% reduziert, d.h. 1,5 Stellen
→ Stelleneinsparung soll nicht die Breite des Angebots beeinflussen, allerdings wird aktive Mitarbeit in SFB nicht ohne Kompensation möglich sein
- Stellenausschreibung in naher Zukunft

Geplante Arbeiten des Kompetenzzentrums

- Organisations- und Prozesssteuerung
 - Projektsteuerung
 - Öffentlichkeitsarbeit
 - Berichtswesen
- Dienstleistungs-Portfolio
- Technische Infrastruktur
 - Aufbau einer technischen Infrastruktur
 - Integration vorhandener Ressource
 - Pflege und Weiterentwicklung

Geplante Arbeiten des Kompetenzzentrums

- Beratung
 - Antragsberatung
 - Unterstützung bei Auswahl technischer Infrastruktur
 - Datenpublikation
 - Rechtliche Aspekte
- Kooperation
 - Technisch
 - Disziplinar

Ausblick

- Mit dem Kompetenzzentrum wird FDM an der Universität Bielefeld als zentrale Aufgabe wahrgenommen
- Verstetigung und Ausbau des Angebots nach erster Förderphase wird angestrebt
- FDM ist ein Querschnittsthema
→ Kooperation mit Fachdisziplinen erforderlich

**Vielen Dank für die
Aufmerksamkeit!**
Fragen? Anmerkungen?