



Forschungsdaten- management Leitlinien

für die Lebenswissenschaften
an der Universität Tübingen

Präambel

Diese Leitlinien sollen eine Ergänzung zu den bereits bestehenden Regelungen im Kontext Forschungsdatenmanagement (FDM) sein.

Darunter fallen, neben den allgemeinen [Leitlinien Forschungsdatenmanagement](#) der Universität, die [Open Science Policy](#) der Universität, die [Leitlinien zur Sicherung guter wissenschaftlicher Praxis der Universität](#), die [Leitlinien zum Umgang mit geistigem Eigentum \(IPR-Policy\)](#) der Universität sowie die [Leitlinien zum Umgang mit generativen KI-Tools](#) der Universität.

In dieser Leitlinie sollen in erster Linie die fachspezifischen Anforderungen an das FDM in den Lebenswissenschaften adressiert werden.

Biomedizinische Forschungsdaten erfordern durch ihren Personenbezug ein besonderes Maß an Datenschutz sowie Datensicherheit. Zudem stellen größere Datensätze und die Diversität von Datenarten in den Lebenswissenschaften besondere Herausforderungen an die technische Infrastruktur. Zudem ist eine gute Reproduzierbarkeit von Forschungsdaten in den Lebenswissenschaften durch die experimentelle Ausrichtung der Datengenerierung und die damit einhergehenden Hürden bei der Wiederholung von Experimenten essenziell.



Motivation

Ein gutes Datenmanagement
bietet viele Vorteile

- ☑ Mehr Zeit für das eigene Forschungsvorhaben
- ☑ Eine größere Sichtbarkeit der eigenen Forschungstätigkeit
- ☑ Sicheres Datenhandling während des Prozesses und bessere Organisation von Daten
- ☑ Schutz vor Datenverlust und Qualitätssicherung
- ☑ Nachhaltigkeit durch Nachnutzung der Daten für Folgeprojekte und sparsamer Einsatz von Ressourcen
- ☑ Einhaltung der Guten Wissenschaftlichen Praxis
- ☑ Erfüllung der Anforderungen von Drittmittelgebern
- ☑ Rechtssicherheit sowie Absicherung von Urheberrecht, Intellectual Property, Patenten und Erfindungen
- ☑ Absicherung und Nachweis der Urheberschaft sowie erleichterte Nachnutzung durch die Vergabe von Lizenzen



Allgemeine Empfehlungen

Forschungsdatenmanagement basiert auf dem Datenlebenszyklus.
Die folgende Grafik soll nützliche Tipps und Empfehlungen bieten, in
welcher Phase des Lebenszyklus welche Maßnahmen sinnvoll sind.

Individuelle Beratung durch Vor-Ort-Infrastruktur: Core Facilities, NFDI,
Rechenzentrum, Universitätsbibliothek... → fachsspezifische AnsprechpartnerInnen

Frühzeitige Prüfung von Kriterien/Vorgaben des jeweiligen Drittmittelgebers

Verwendung/Nachnutzung von
bereits existierenden Forschungsdaten

Erstellen eines Datenmanagementplans (DMP) → RDMO

Verwendung von zertifizierten
Forschungsdatenrepositorien
→ re3data

Verwendung von offenen und gut dokumentierten Dateiformaten

Nutzung eines Elektronischen Laborjournals (ELN)

Speicherung und Veröffentlichung von annotierten
Forschungsdaten für mind. 10 Jahre → gemäß DFG-
Kriterien

Einbezug von nichtdigitalen Sammlungen und Digitalisierung
von relevanten Daten für die Veröffentlichung

Berücksichtigung der FAIR-Prinzipien

Verwendung von persistenten Identifikatoren, z.B. DOI

Berücksichtigung von ethischen Fragestellungen
→ Nagoya-Protokoll, CARE-Prinzipien, Ethik-Kommissionen, Tierschutz, ...

Publikation von OER in ZOERR

Annotation der Daten nach fachsspezifischen Standards mit repositoryspezifischen Metadaten

Verwendung von ORCID

Datenschutz berücksichtigen → Anonymisierung oder Pseudonymisierung von
personenbezogenen Daten, besonderer Schutz für Patientendaten

Nach Möglichkeit Open Access publizieren



Anlage 1: Forschungsdatenfluss in der grundlegenden **Pflanzenforschung** für LebenswissenschaftlerInnen an der Universität Tübingen

Dieses Beispiel ist von der gelebten Praxis rund um den TRR356 PlantMicrobe inspiriert.

Es stellt keine Vorgabe dar, sondern soll als Orientierung dienen, um die Abläufe in Forschungsprojekten zu organisieren. Zur einfacheren Lesbarkeit sind eigentlich zyklische Prozesse linear dargestellt.

Datenmanagement-Plan:

- 📁 Verantwortlichkeiten für das FDM im Projekt klären
- 📁 Datenarten bestimmen, die verwendet bzw. generiert werden
- 📁 Geplante Datenvolumina schätzen
- 📁 Festlegung auf einheitlichen Annotierungsstandard, hier ISA-Tab
- 📁 Festlegung der Strukturierung der Daten, hier Annotated Research Context (ARC)
- 📁 (Kollaborative) Speicherumgebung für Datennutzung während Projektlaufzeit klären, hier dedizierte Gitlab-Instanz für TRR356
- 📁 Prozessierungsumgebung klären, hier de.NBI Cloud Tübingen und LRZ-Cloud



Weitere Empfehlungen für den Bereich Pflanzenforschung:

Strukturierung der Daten anhand ISA-Tab beziehungsweise ARC Konzept.

Kontinuierliche Dokumentation aller Experimente, Messungen und Abläufe über ein elektronisches Laborjournal (ELN), hier eLabFTW.

Aktive Nutzung der Speicherumgebung unter konsequenter Verwendung der vorgesehenen Standards für alle Daten, hier TRR356 Gitlab.

Prozessierung der Daten in den vorgesehenen Umgebungen, hier de.NBI Cloud Tübingen und LRZ Compute Cloud, sowie Speicherung der annotierten Ergebnisse in der Speicherumgebung.

Kooperatives Arbeiten im Lab oder mit Partnern unter Nutzung der o.g. Speicher- und Prozessierungsumgebungen.

Ausformulierte Standard-Operating-Procedures (SOP) zu allen relevanten Prozessen inklusive regelmäßiger Überprüfung und Verbesserung.

Publikation der wohlannotierten Forschungsdatensätze in Communityrepositorien und/oder FDAT, üblicherweise in Einklang mit passender Journalpublikation, inklusive Lizenzvergabe (Empfehlung: CC-BY oder CC-BY-SA 4.0) und persistenten Identifikatoren (ORCID und DOI).

**Anlage 2: Forschungsdatenmanagement für
LebenswissenschaftlerInnen an der Medizinischen Fakultät**



**Universitätsklinikum
Tübingen**

Geniale Idee für ein datengestütztes klinisches Forschungsprojekt aber offene Fragen zu Datenmanagement, Datenschutz, IT, Governance....?



Wie komme ich jetzt an die
Daten, die ich brauche?



Damit
konnte doch
keiner
rechnen!



Hätten wir das
nicht auch
anders planen
können?

Wir haben da eine Lösung: das **Center for Digital Health (CDH)** des UKT und der MFT



The screenshot shows the website of the Center for Digital Health (CDH). The header features the title "Center for Digital Health" in white text against a dark background with a glowing network diagram. Below the header is a breadcrumb trail: "Medizinische Fakultät > Forschung > CDH - Center for Digital Health". The main content area is titled "Unsere Mission" and contains a paragraph describing the center's mission. At the bottom right, there is a yellow button with the text "Beratung und Support anfragen".

Center for Digital Health

Medizinische Fakultät > Forschung > CDH - Center for Digital Health

Unsere Mission

Das Center for Digital Health (CDH) an der Medizinischen Fakultät der Universität Tübingen und dem Universitätsklinikum bündelt die Expertise in der Medizin- und Bioinformatik, mit den Schwerpunktthemen Data Science und Maschinellem Lernen. Es fördert die datengetriebene interdisziplinäre Forschung und schafft die Voraussetzungen für eine effiziente Nutzung und Analyse großer Gesundheitsdatenmengen mit dem Ziel der Integration von innovativen digitalen Ansätzen aus der Grundlagenforschung in die Versorgung (Translation).

Beratung und Support anfragen

<https://www.mezizin.uni-tuebingen.de/de/medizinische-fakultaet/forschung/cdh-center-for-digital-health>

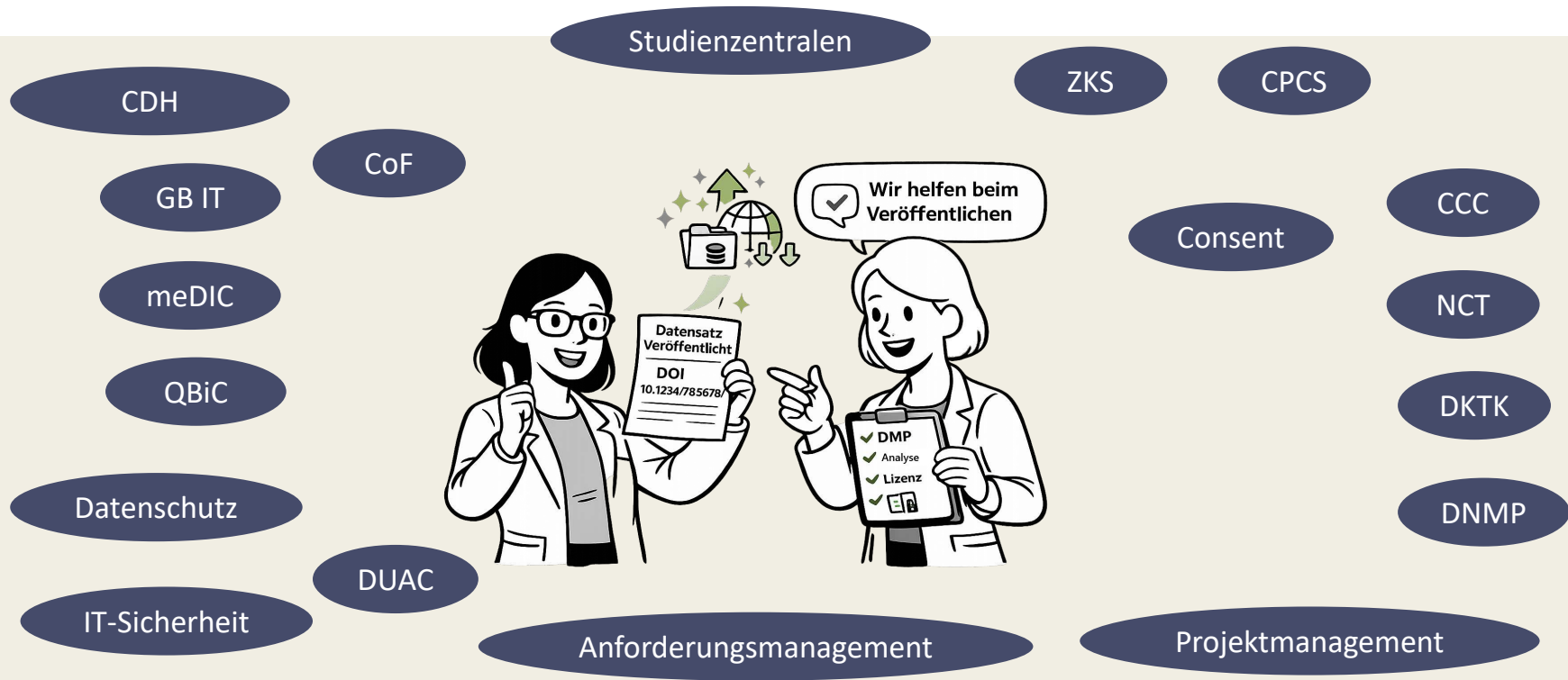


Beratung, Hilfestellungen und Vermittlung an zuständige Stellen bzgl. ...

- Anforderungsanalyse
- Machbarkeit
- Datenmanagement
- Projektmanagement
- IT-Infrastruktur
- Datenschutz
- IT-Sicherheit
- Analyseplanung
- ...



Erfolgreich forschen und publizieren an der Medizinischen Fakultät



Beispiel: Retrospektive Datenauswertung, lokal / monozentrisch

Zu klären:

- Studiendesign und Dokumentation
 - Ethikantrag?
- Eigenforschung?
 - Rechtsgrundlage?
- Daten bereits zugänglich?
 - Nein? -> Daten am UKT verfügbar?
- Wie/wo/durch wen wird ausgewertet?
- Müssen Verträge geprüft werden?

Freiwillige oder verpflichtende
Beratungsangebote

Studienzentrale ZKS

Ethikkommission

Datenschutzbüro

GBIT

meDIC

CCC

Rechtsabteilung

Ggfls.
Freigabeprozess

DUAC

Daten

