

FDM Basics

Grundlagen des Forschungsdatenmanagements

Shanna Pilch (HS Worms) & Gery Weyd (TH Bingen)

Netiquette



Mikro
stummschalten



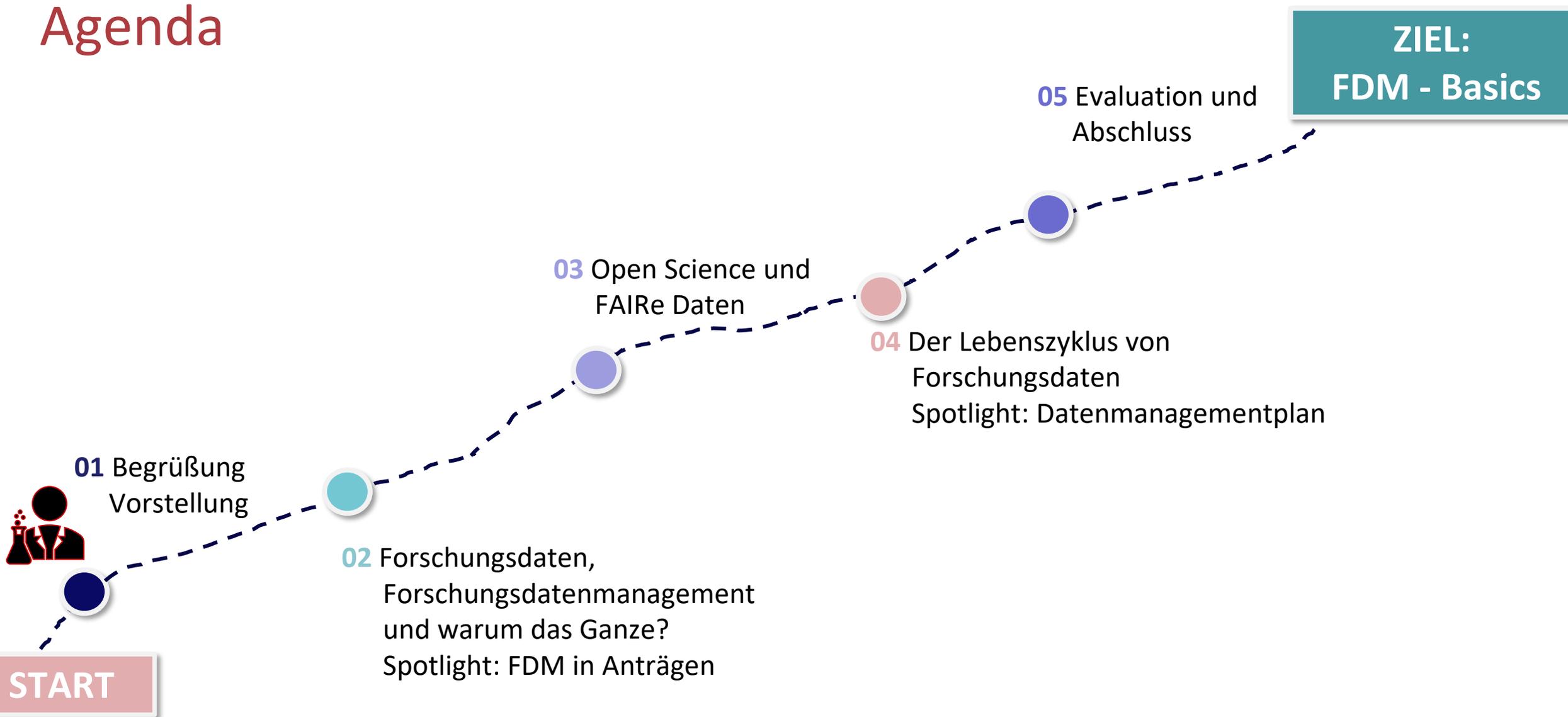
Kamera an ;-)



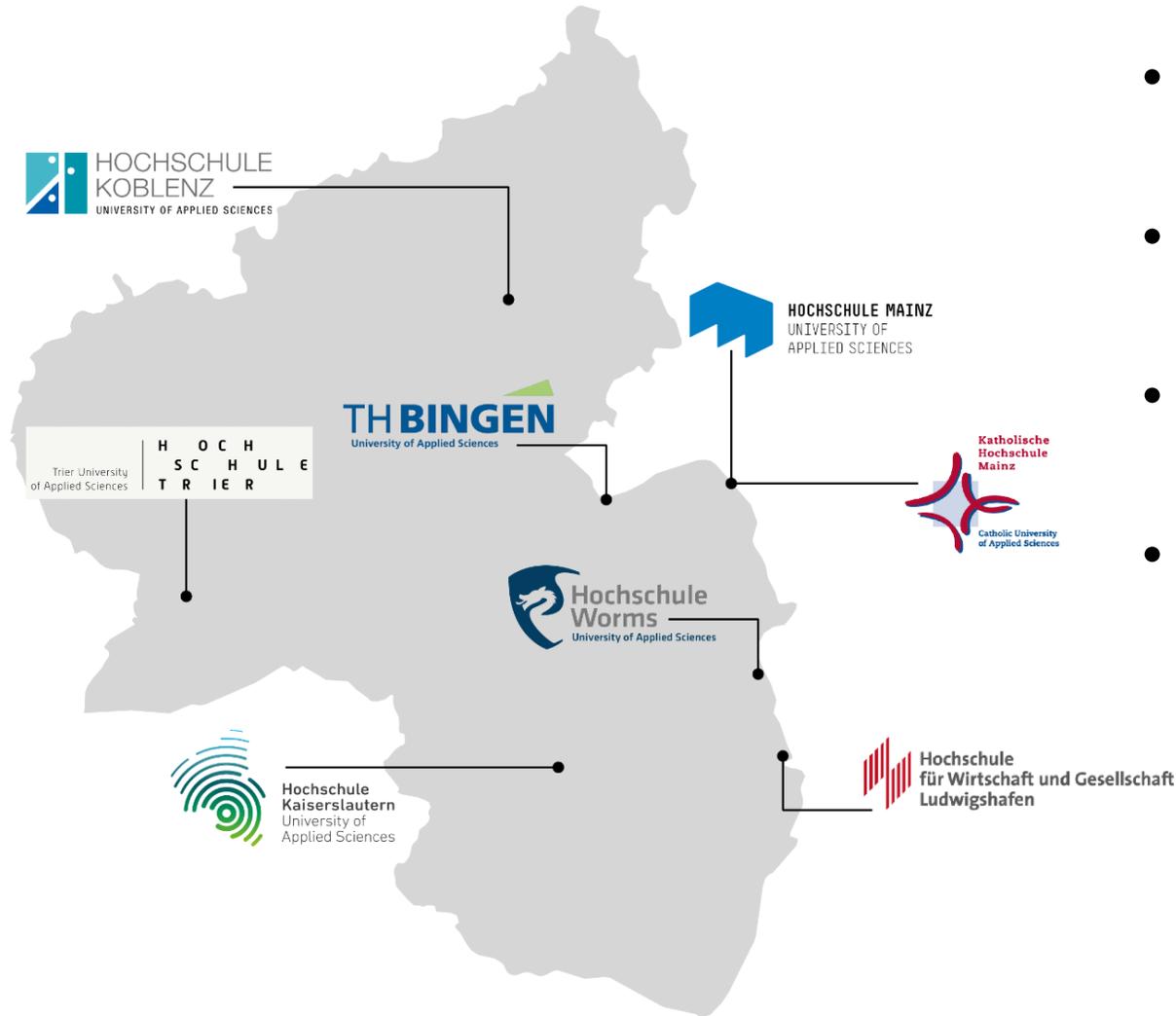
Fragen in den Chat

Teilnahmebescheinigung auf Wunsch!
fdm-haw@hs-mainz.de

Agenda



Projekt FDM@HAW.rlp



- Ziel: Nachhaltige Verankerung von FDM in allen HAWn in RLP
- Verbundprojekt: 7 staatl. Hochschulen für angewandte Wissenschaften + KH Mainz
- Gefördert durch das BMBF und finanziert durch die EU
- Laufzeit 2022 - 2025

<https://www.fdmhawrlp.de/>



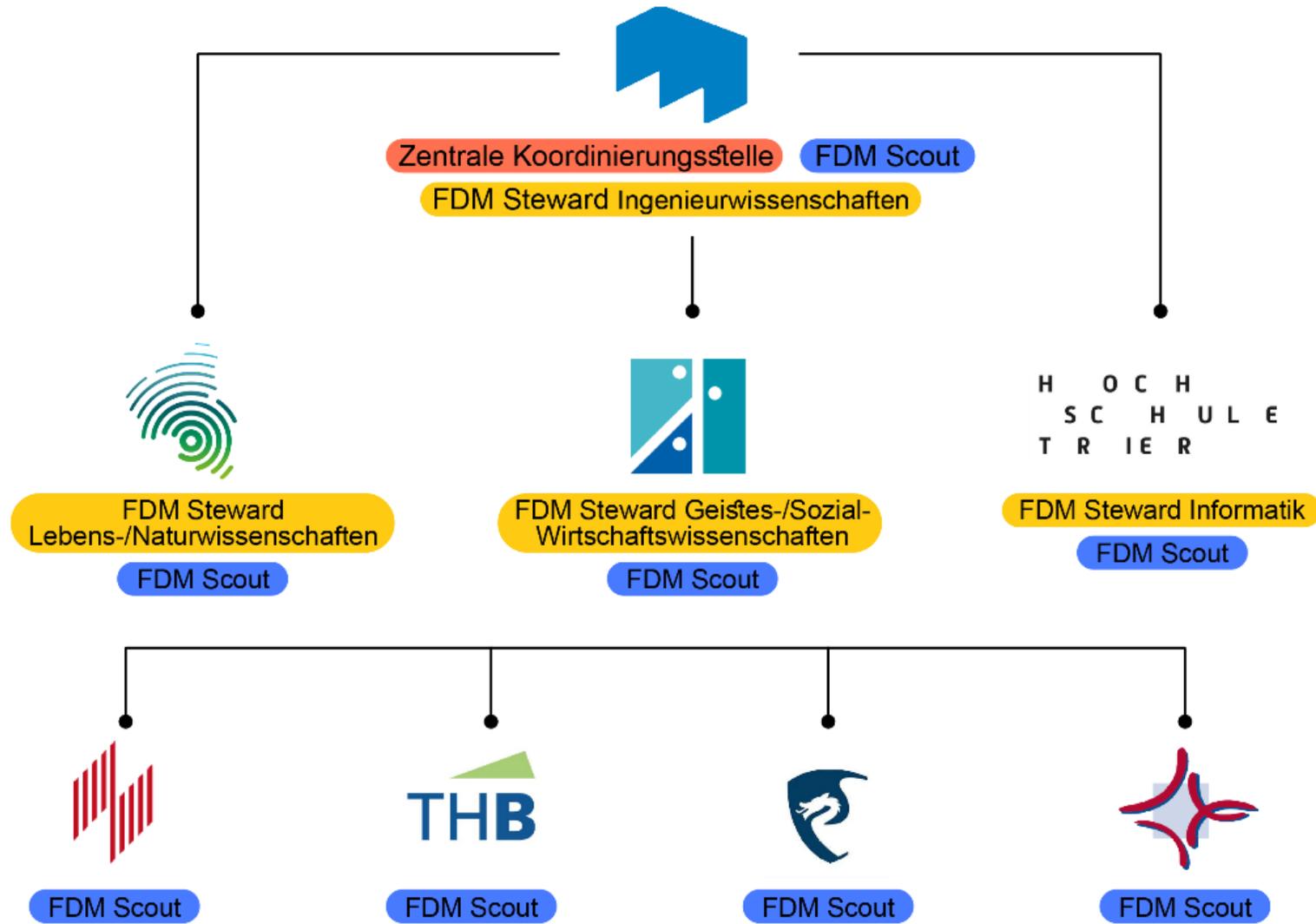
GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung



Finanziert von der
Europäischen Union
NextGenerationEU

Projektstruktur



Kontakt Daten der Scouts

Technische Hochschule Bingen

Gery Weyd | FDM-Scout

E: [g.weyd\(at\)th-bingen\(dot\)de](mailto:g.weyd(at)th-bingen(dot)de)

Hochschule Kaiserslautern

Maximilian Weyland | FDM-Scout

E: [maximilian.weyland\(at\)hs-kl\(dot\)de](mailto:maximilian.weyland(at)hs-kl(dot)de)

Hochschule Koblenz

Anja Schwarz | FDM-Scout

E: [schwarz1\(at\)hs-koblenz\(dot\)de](mailto:schwarz1(at)hs-koblenz(dot)de)

Hochschule für Wirtschaft und Gesellschaft Ludwigshafen

Andrea Kuhn | FDM-Scout

E: [Andrea.Kuhn\(at\)hwg-lu\(dot\)de](mailto:Andrea.Kuhn(at)hwg-lu(dot)de)

Hochschule Mainz

Jennifer Wettig | FDM-Scout

E: [jennifer.wettig\(at\)hs-mainz\(dot\)de](mailto:jennifer.wettig(at)hs-mainz(dot)de)

Katholische Hochschule Mainz

Darina Schilirò | FDM-Scout

E: [darina.schiliro\(at\)kh-mz\(dot\)de](mailto:darina.schiliro(at)kh-mz(dot)de)

Hochschule Trier

Anja Nied | FDM-Scout

E: [a.nied\(at\)hochschule-trier\(dot\)de](mailto:a.nied(at)hochschule-trier(dot)de)

Hochschule Worms

Shanna Pilch | FDM-Scout

E-Mail: [pilch\(at\)hs-worms\(dot\)de](mailto:pilch(at)hs-worms(dot)de)

Kontakt Daten der Stewards

Ingenieurwissenschaften

Manuela Richter | FDM-Steward
Hochschule Mainz
E: [manuela.richter\(at\)hs-mainz\(dot\)de](mailto:manuela.richter(at)hs-mainz(dot)de)

Geistes-, Sozial- und Wirtschaftswissenschaften

Lucas Krajewski geb. Hamel | FDM-Steward
Hochschule Koblenz
E: [krajewski\(at\)hs-koblenz\(dot\)de](mailto:krajewski(at)hs-koblenz(dot)de)

Informatik

Dr. Johannes Putzke | FDM-Steward
Hochschule Trier
E: [J.Putzke\(at\)umwelt-campus\(dot\)de](mailto:J.Putzke(at)umwelt-campus(dot)de)

Lebens- und Naturwissenschaften

Jörg Panzer | FDM-Steward
Hochschule Kaiserslautern
E: [joerg.panzer\(at\)hs-kl\(dot\)de](mailto:joerg.panzer(at)hs-kl(dot)de)

Lernziele des Seminars

- Sie wissen, was **FDM** ist, und kennen den **Forschungsdatenlebenszyklus**.
- Sie wissen, wie Daten gestaltet sein müssen, damit Sie **FAIR** sind.
- Sie wissen, welche **Entscheidungen** in der **Planungsphase** anstehen können.
- Sie wissen, was ein **Datenmanagementplan** ist.
- Sie kennen Werkzeuge, um FD gut **strukturieren und ordnen** zu können.
- Sie kennen den Nutzen und die Inhalte der **Datendokumentation**.
- Sie kennen die Voraussetzungen für ein sicheres **Backup** und die **Langzeitarchivierung**.
- Sie wissen, was ein **Repository** ist, und kennen verschiedene **Publikationswege** für FD.

Sie erkennen einen Nutzen im Betreiben von Forschungsdatenmanagement für Ihre Arbeit.

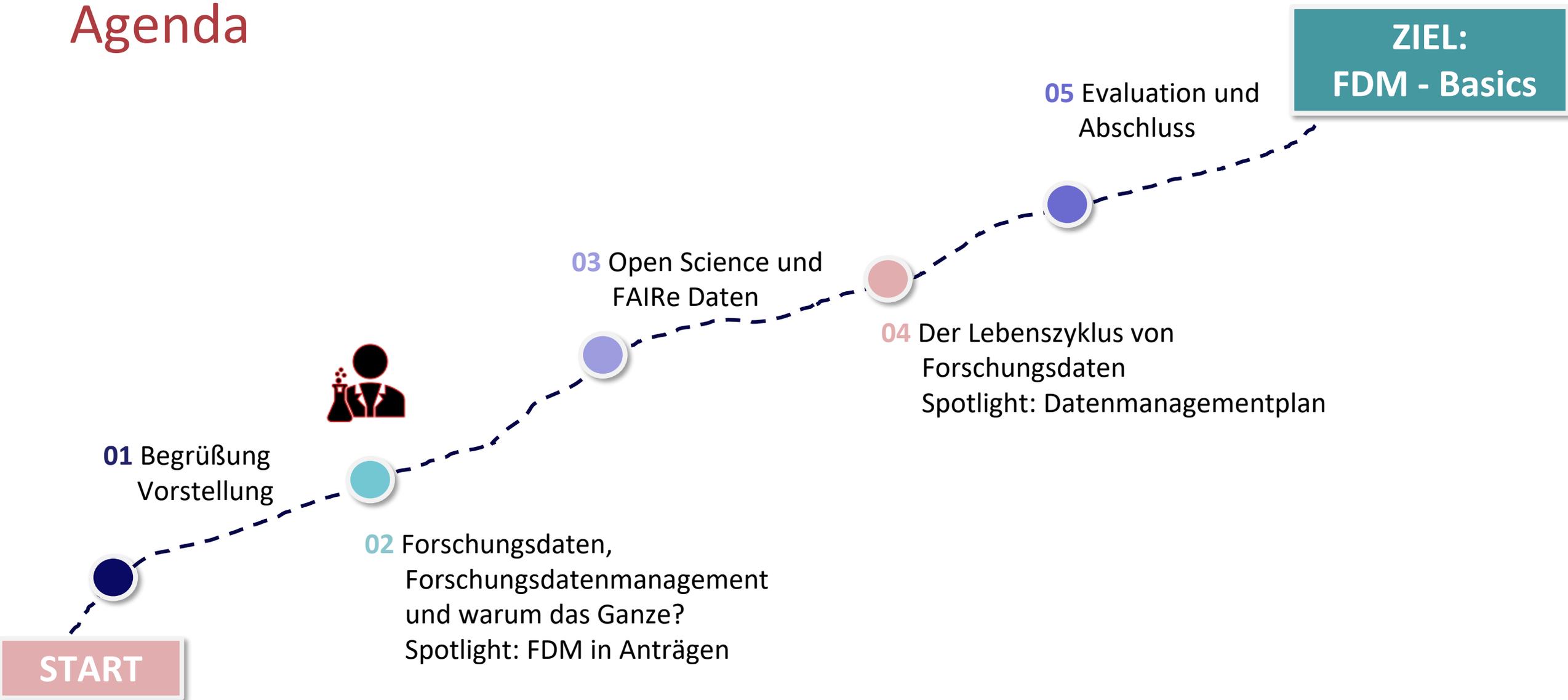
Unsere Fragen an Sie:

Welche Vorkenntnisse bringen Sie heute mit:

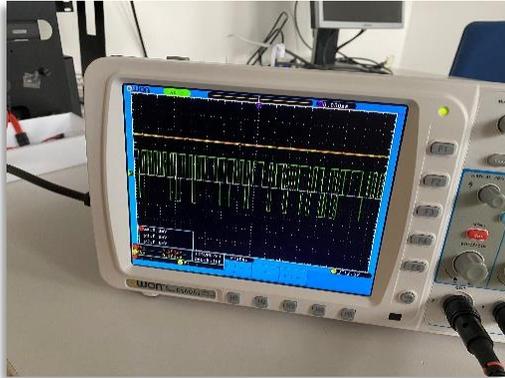
1. Ich habe den Begriff „Forschungsdatenmanagement“ (FDM) bereits gehört.
2. Ich kenne den „Forschungsdatenlebenszyklus“.
3. Ich wurde bei einem Drittmittelantrag dazu aufgefordert, Angaben zum FDM/Umgang mit Forschungsdaten zu machen.
4. Ich habe bereits einen Datenmanagementplan (DMP) erstellt.
5. Ich habe bereits Forschungsdaten veröffentlicht.



Agenda



Was sind Forschungsdaten?



Messdaten/ Messreihen



Umfragen



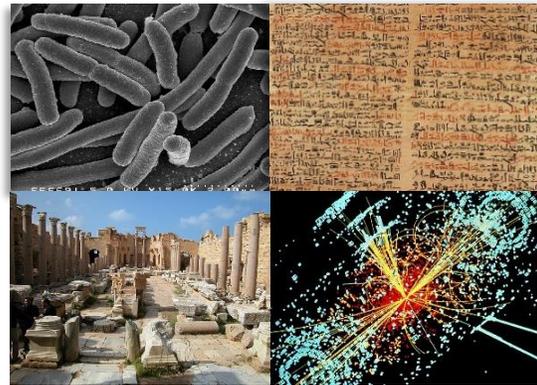
Interview^[1]

Habe nun, ach! Philosophie,
Juristerei und Medizin,
Und leider auch Theologie
Durchaus studiert, mit heißem Bemühn.
Da steh ich nun, ich armer Tor!
Und bin so klug als wie zuvor;
Heiße Magister, heiße Doktor gar
Und ziehe schon an die zehen Jahr
Herauf, herab und quer und krumm
Meine Schüler an der Nase herum-
Und sehe, daß wir nichts wissen können!
Das will mir schier das Herz verbrennen.

Texte



Labordaten



Bilder/Modellierungen^[2]



Video^[3]

...

...und vieles mehr

= alle (digitalen) Daten, die in der Forschung entstehen, entwickelt und ausgewertet werden

Was ist Forschungsdatenmanagement (FDM)?

planen

aufbereiten

verwalten

strukturieren



dokumentieren

archivieren

auswählen

publizieren



FDM: Überlegter und systematisierter Umgang mit Daten von der Planung bis zum Abschluss eines Projektes und w.m. darüber hinaus.

Warum das Ganze? Mehrwert für Forschende



Überblick behalten



Zusammenarbeit erleichtern



Zeit und Nerven sparen



Qualität wissenschaftlicher Arbeit gewährleisten



formale Anforderungen erfüllen

Warum das Ganze? (Wissenschafts)Politische Dimension

„Daten sind das neue Öl“.
(Clive Humby, Mathematiker) [1]



Open Science Policy der EU

The screenshot shows the European Commission website page for Open Science. The URL is https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en#the-eus-open-science-policy. The page features the European Commission logo, a search bar, and a navigation menu. The main heading is "Open Science", followed by a sub-heading: "An approach to the scientific process that focuses on spreading digital and collaborative technology. Expert groups, ...". Below this, there is a "PAGE CONTENTS" section with links to "The EU's open science policy", "8 ambitions of the EU's open science policy", "Future of open science under Horizon Europe", and "Tracking open research trends - Open Science Monitor". A "Documents" link is also present. On the right side, there is a "European Open Science Cloud - EU Node" section with a navigation menu (Home, About, Services, Resource Hub, Support, Contributors, News & Events) and a sub-heading "EOSC EU Node" with the text "A European platform and information gateway to explore, engage, and enrich your research collaborations."

Screenshots: Thomas Schimmer 04/2024

Spotlight Anträge: Welche Fragen werden in der Regel gestellt?

Beschreibung der Daten

Datentypen, Datenvolumen, Speicherung, Sicherung, Austausch

Rechtliche Aspekte

Datenschutz, Anonymisierung, Pseudonymisierung, Lizenzen

Leitlinien und Standards

z.B. DFG-Leitlinien, FAIR-Prinzipien, disziplinspez. Standards & Umsetzung

Dokumentation

Was wird wie dokumentiert? (z.B. ReadMe-Datei, Codebook, Metadaten)

Archivierung

Was, wo und wie lange? (z.B. 10 Jahre nach DFG-Leitlinie)

Publikation und Nachnutzung

Repositorien, Lizenzen, Embargofrist

Verantwortlichkeiten & Ressourcen

Während und nach Projektlaufzeit

Spotlight Anträge: FDM bei EU – Bund – DFG – Stiftungen

Förderinstitution	FDM-Angaben	Inhalte
Europäische Kommission (Horizon Europe)	verpflichtend	DMP (öffentlich), PID, Lizenzierung, Repositorium, Dokumentation, FAIR
Bund	uneinheitlich	ggf. FDMP / Verwertungsplan, FAIR
DFG	verpflichtend	FDM im Antrag, ggf. zusätzlich FDM-Verbund- Konzept, fachspez. Standards, Repositorium, FAIR
VolkswagenStiftung	verpflichtend	DMP, PID, Lizenzierung, fachspez. Standards, Repositorium, FAIR
Carl-Zeiss-Stiftung	uneinheitlich	Ggf. Forschungsdatenkonzept

Ausschlaggebend ist das Förderprogramm und der jeweilige Ausschreibungstext.

[1]

Europäische Kommission (Horizon Europe)

FDM-Antragsunterstützung

- [Informationen und Inhalte für die Antragsstellung](#) (OpenAIRE)
- [Informationen zu FDM allgemein für Horizon Europe](#) (OpenAIRE)
- [Horizon Europe Programm Guide](#) (PDF)
- [FDM-Template \(RDM-Template\)](#) (PDF)

FDM-Antragsanforderungen

- Projektantrag: **Stellungnahme zum FDM** (max. 1 Seite)
- Im Projektverlauf: **DMP** entsprechend der DMP-Vorlage verpflichtend:
 - 1. Version innerhalb von 6 Monaten nach Projektstart
 - 2. Version - Aktualisierung: Projektmitte
 - 3. Version: Abschluss

ALLGEMEIN GILT:

- FAIR-Prinzipien beachten
- Forschungsdaten in Repositorien zugänglich machen
- Nachnutzung ermöglichen:
 - Lizenzierung: mgl. CC-BY oder "Public Domain" CC0
 - Ausführliche Dokumentation (Daten, Erhebungsmethoden, Werkzeuge bzw. Instrumente, inkl. Softwarecode, Algorithmen oder Analyseprotokolle)

[1]

Bundesministerien (u. a. BMBF, BMWi, BMLE)

FDM-Antragsunterstützung

Bundesministerium für Bildung und Forschung (BMBF):

EBF-Rahmenprogramm (empirische Bildungsforschung) =>
Maßstab und Orientierung

Küsten-, Meeres- und Polarforschung

- FDM im Verwertungsplan
- Qualität des DMs für Bewertung Projektskizze relevant

Bundesanstalt für Landwirtschaft und Ernährung (BLE)

- [FDMP Skizzenphase \(Stand 01-2024\)](#)
- [FDMP Projektphase \(Stand 01-2024\)](#)
- [Merkblatt FDM \(Stand 08-2023\)](#)

FDM-Antragsanforderungen

- Keine einheitlichen Vorgaben

Abhängig vom Förderprogramm (Ausschreibungstext beachten!):

- Verwertungsplan mit FDM-Angaben
- DMP / FDMP (Forschungsdatenmanagementplan)

ALLGEMEIN GILT:

- FDM soll betrieben werden, aber ohne Vorschrift wie
- Man soll mgl. einem fachspezifischen Standard folgen
- Daten sollen mgl. FAIR-nachnutzbar gemacht werden
- Auffindbarkeit & Nachnutzbarkeit wichtig :
DOI, Repository, Dokumentation (s. BLE)

ENTWICKLUNG DER VORGABEN ORIENTIERT SICH AN NFDI

s. [Aktionsplan BMBF](#)

Deutsche Forschungsgemeinschaft (DFG)

FDM-Antragsunterstützung

- [Fragenkatalog \(Checkliste PDF\)](#) zum Umgang mit Forschungsdaten (FD)
- [Informationen zu beantragbaren Mitteln](#)
„Kosten, die für Datenmanagement anfallen und der Nutzbarkeit diese Daten dienen, können beantragt werden“ (keine Grundausstattung!)
- [Förderprogramm](#) - spezifische Informationen
- [Fachspezifische Empfehlungen](#)

FDM-Antragsanforderungen

Grundsätzliche Förderkriterien:

- Schilderung Umgang mit FD im (Teil-)Projekt (Abschnitt 2.4 des Antrags, Hilfsmittel: Fragenkatalog der DFG)
 - Konzept zum Erhalt der FD nach Förderende
 - Empfehlung der DFG: frühzeitige Suche nach Repositorium
 - Anschlussfähigkeit an bestehende Strukturen (etwa bei Schaffung von Dateninfrastrukturen und Methoden)
 - Standards der Fachkollegien beachten

Verbundanträge:

In Skizze & Antrag (Abschnitt 2.3):

Umgang mit FD aufnehmen

=> Entwicklung eines gemeinschaftlichen FDM-Verbund-Konzepts!

Volkswagenstiftung

FDM-Antragsunterstützung

- [Basis-DMP](#) mit 7 Fragen
- [FAQs für Antragsstellende](#)

Förderprogramm: [Data-Reuse](#)

Mittel für FDM bis 6 Monate nach Projektende

FDM-Antragsanforderungen

- Datenmanagementplan (DMP) bei datenintensiven Forschungsanträgen (Basis-DMP inkl. ggf. Vorgaben des Zielrepositoriums)
- **Skizze des Datenmanagements im Antrag sinnvoll, auch wenn nicht im Förderprogramm erwähnt!**

Hinweise zum DMP – erwünscht ist:

- Beachtung FAIR-Prinzipien & community-spezifische Standards
- Vergabe von persistenten Identifikatoren (z. B. DOI)
- Lizenzvergabe (während und nach Projekt)
- Repositoriumsüberlegung/-auswahl

Carl-Zeiss-Stiftung

FDM-Antragsunterstützung

- [Übersicht der Fördermöglichkeiten für HAW](#)

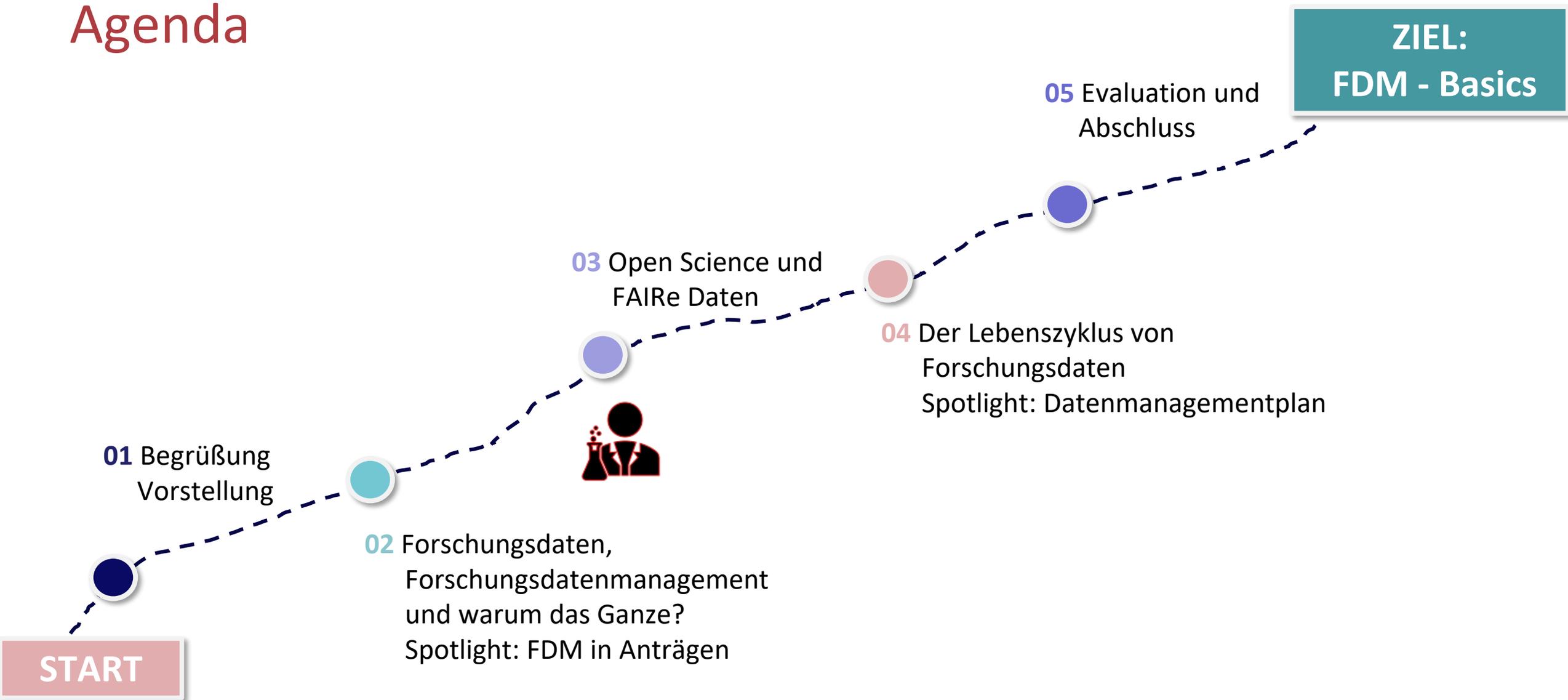
FDM-Antragsanforderungen

Abhängig vom Förderprogramm

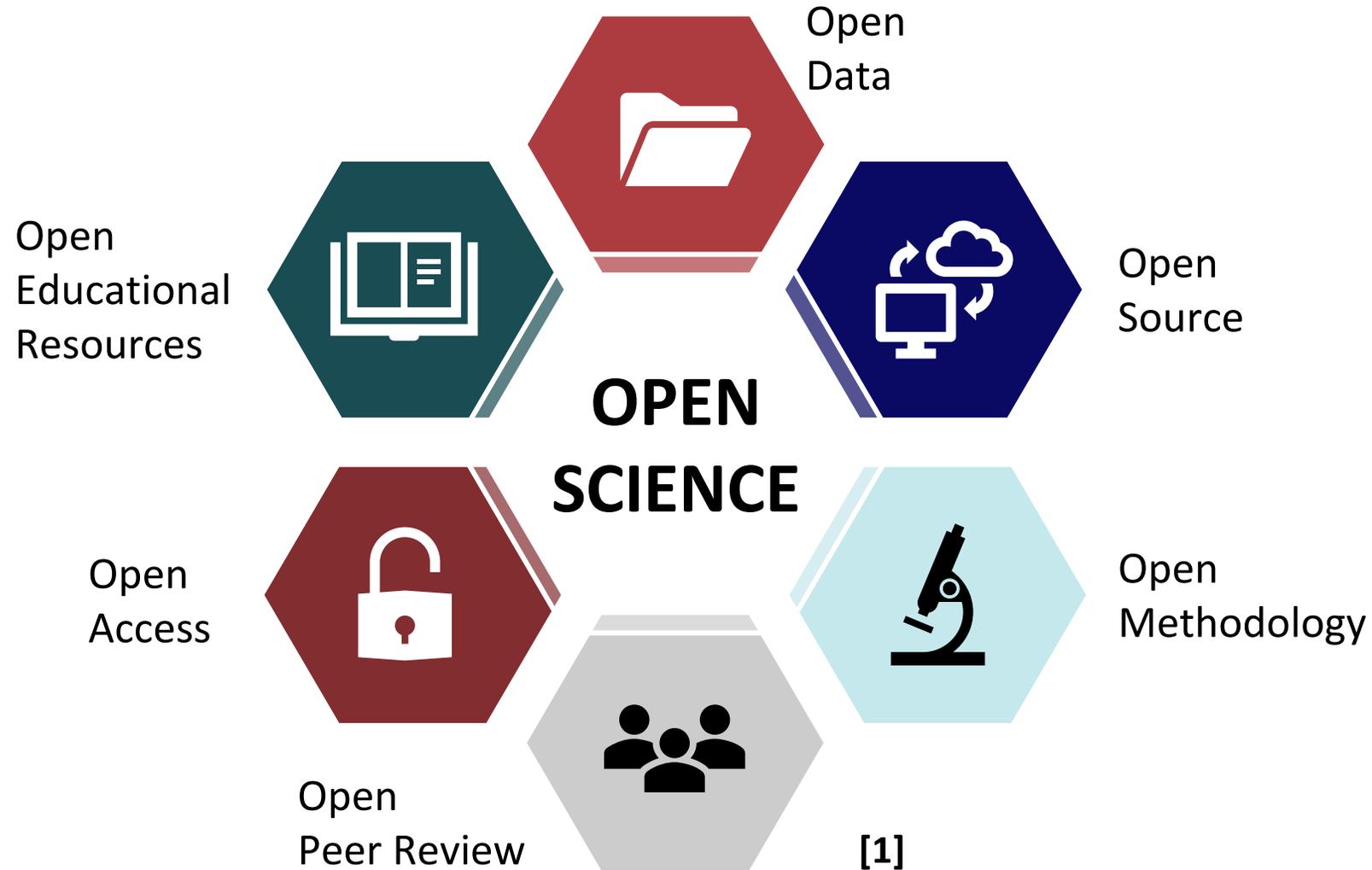
Ggf. Forschungsdatenkonzept (keine Aktualisierung):

- Datenschutz
- Open Source
- Open Access
- gemeinsame Datenbanken
- Reproduzierbarkeit

Agenda



Prinzipien der Open Science



FAIR-Prinzipien

Findable



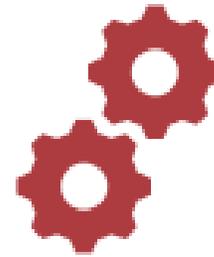
Auffindbar

Accessible



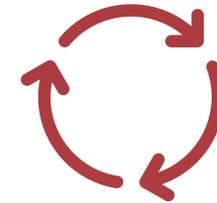
Zugänglich

Interoperable



Interoperabel

Reusable



Wiederverwendbar

[1]

FAIR-Prinzipien gemäß Wilkinson et al. 2016

Findable:

- F1. (Meta)data are assigned a globally unique and persistent identifier (URI. . .)
- F2. Data are described with rich metadata (defined by R below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

Accessible:

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
 - A1.1 The protocol is open, free, and universally implementable
 - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

FAIR-Prinzipien gemäß Wilkinson et al. 2016

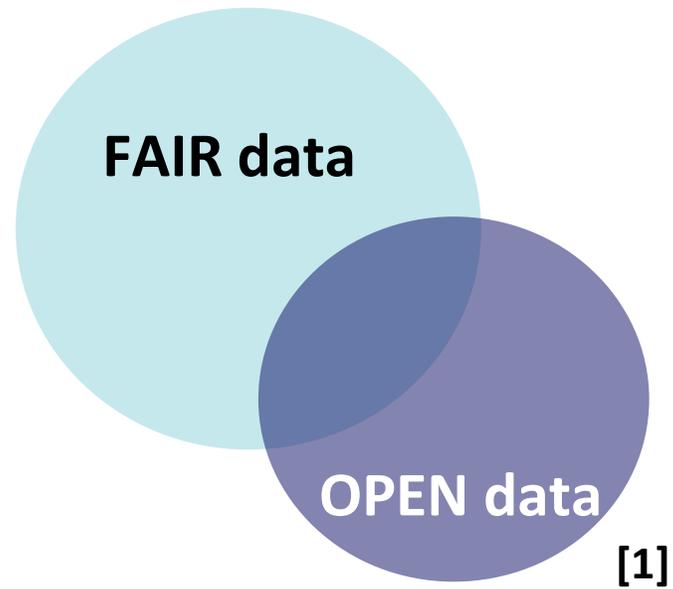
Interoperable

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

Reusable

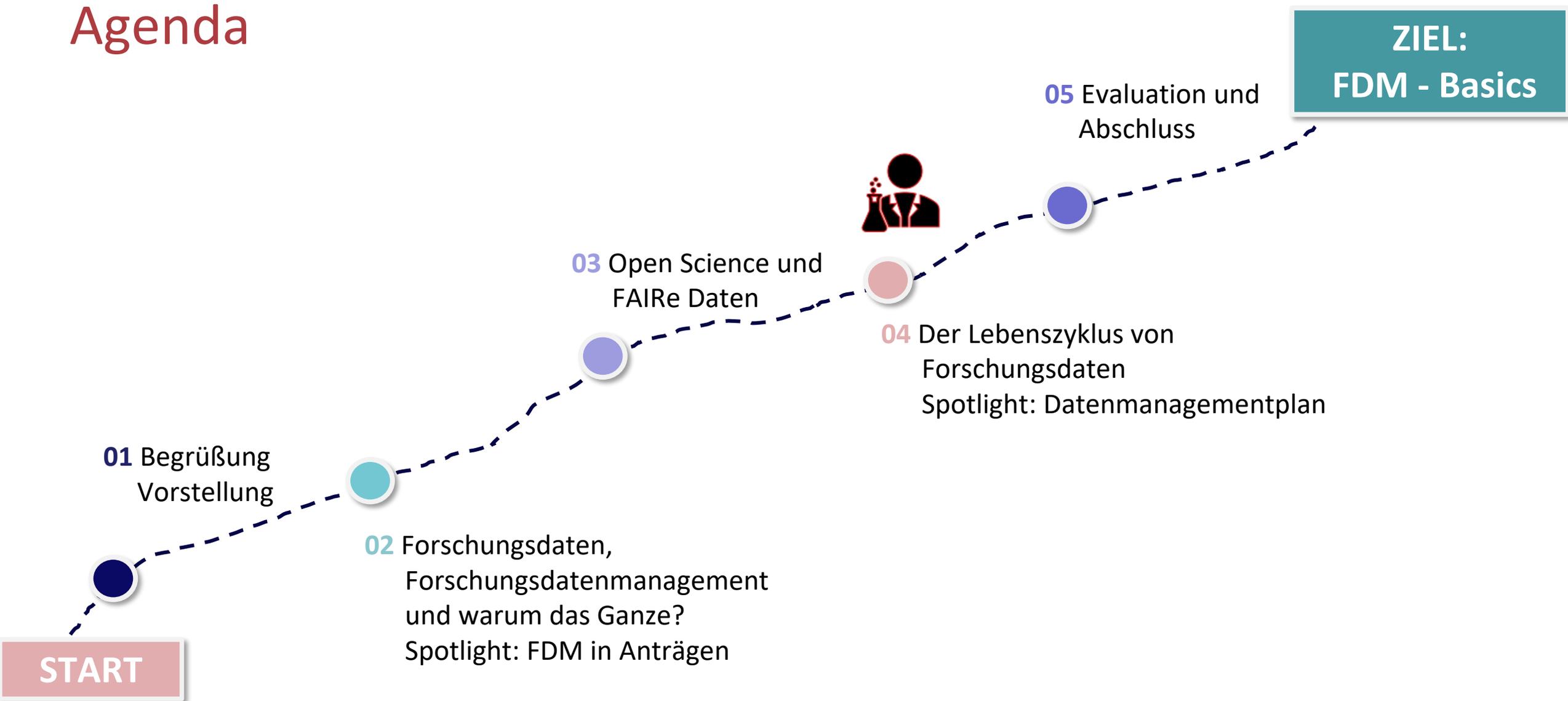
- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
 - R1.1. (Meta)data are released with a clear and accessible data usage license
 - R1.2. (Meta)data are associated with detailed provenance
 - R1.3. (Meta)data meet domain-relevant community standards

FAIR vs. OPEN data

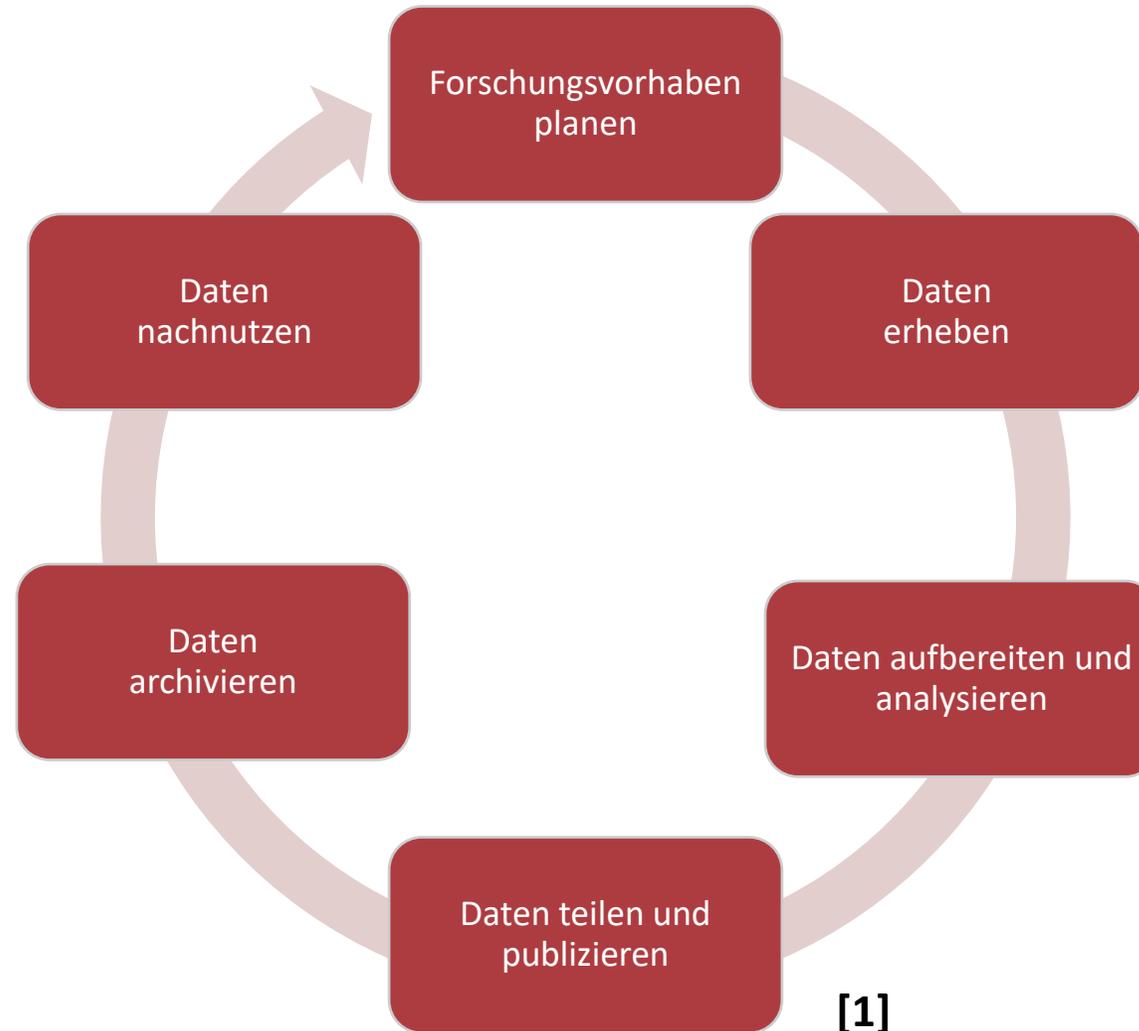


„As open as possible, as closed as necessary“
(EU Open Science-Policy)

Agenda

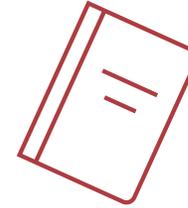


Forschungsdatenlebenszyklus



4. Der Lebenszyklus von Forschungsdaten – Inhalte

4.1. Planung des Forschungsvorhabens



4.2. Datenorganisation und -dokumentation

4.3. Datensicherung und -archivierung

4.4. Datenpublikation

1. Planung des Forschungsvorhabens (I)

Welche **Daten** sollen erhoben werden?

Wie organisiere ich das **Backup** meiner Daten?

Wie sollen die Daten im Projekt **geteilt** werden?

Welche **Software und Datenformate** möchte ich verwenden?

Wo sollen die Daten **gespeichert** werden?

Gibt es bereits Daten, die ich **nachnutzen** kann?

1. Planung des Forschungsvorhabens (II)

Welche **Vorgaben**
seitens der
Drittmittelgeber
gibt es?

Muss/soll ein **DMP**
erstellt werden?

Welche Kosten entstehen ggf. für
Nachnutzung, Speicherung oder
Archivierung? Können diese
erstattet werden?

Welche **rechtlichen**
Rahmenbedingungen
sind relevant?

Müssen **rechtl.**
Vereinbarungen
getroffen werden?

Spotlight: Datenmanagementplan (DMP)



Ein Datenmanagementplan (DMP) ...

- ... ist ein strukturiertes Dokument, enthält alle Infos, wie mit den Daten während und nach Projektende verfahren wird;
- ... erleichtert die Abstimmung zwischen Projektpartnern;
- ... hilft bei der frühzeitigen Erkennung von Problemen und Lösungsfindung;
- ... beugt Sicherheitslücken vor;
- ... ggfls. Anforderung von Drittmittelgebern;
- ... kann mittels Checklisten und online-Tools (z.B. RDMO) erstellt werden.

Spotlight: Datenmanagementplan (II)



Datenmanagementpläne nach der DMP-Vorlage von Science Europe

Inhalte:

Basisinformation

Datenbeschreibung

Dokumentation & Datenqualität

Speicher und Backup

Recht und Ethik

Publikation und Nachnutzung

Verantwortlichkeiten & Ressourcen

[1]

4. Der Lebenszyklus von Forschungsdaten

4.1. Planung des Forschungsvorhabens

4.2. Datenorganisation und -dokumentation

4.3. Datensicherung und -archivierung

4.4. Datenpublikation

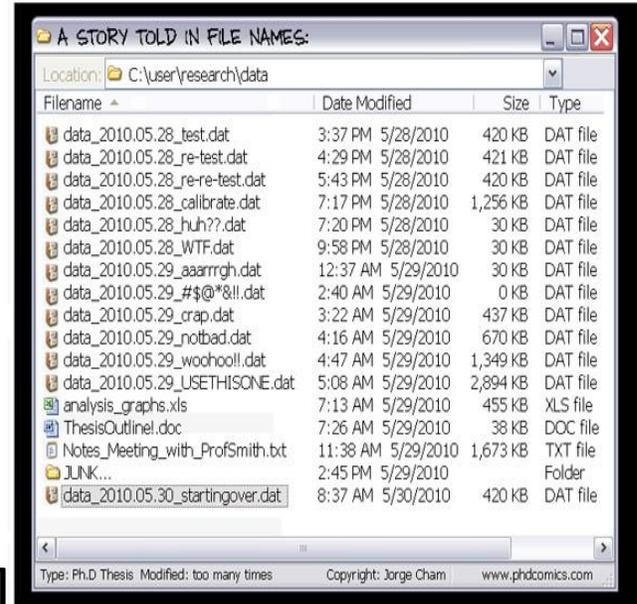


4.2.1. Datenorganisation

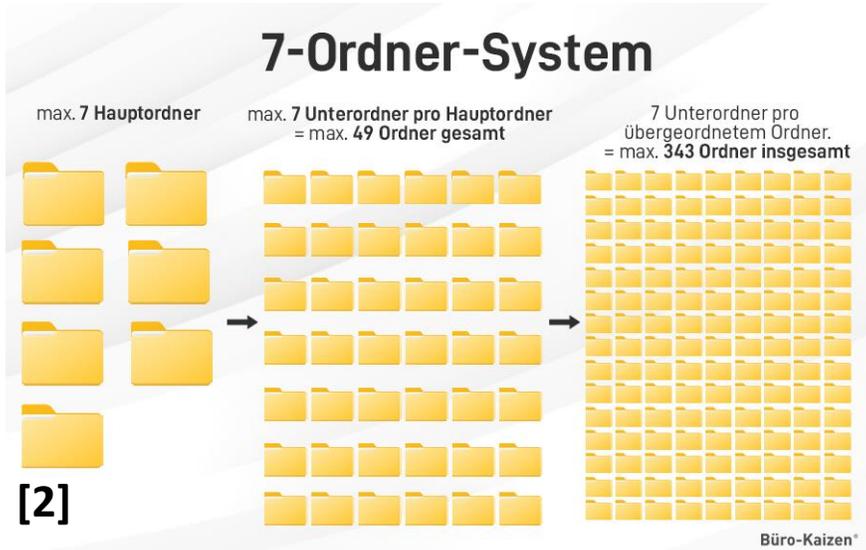
Ordnerstruktur

Dateibenennung

Versionierung



[1]



[2]

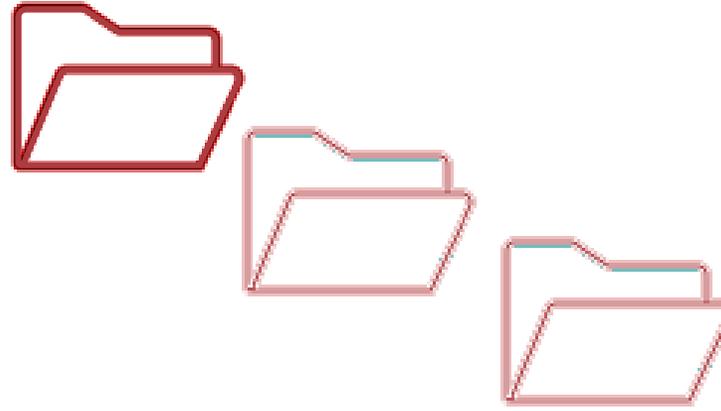


[3]

Ordnerstruktur

Eine gute Ordnerstruktur...

... ist hierarchisch gegliedert;
... baut inhaltlich aufeinander auf.

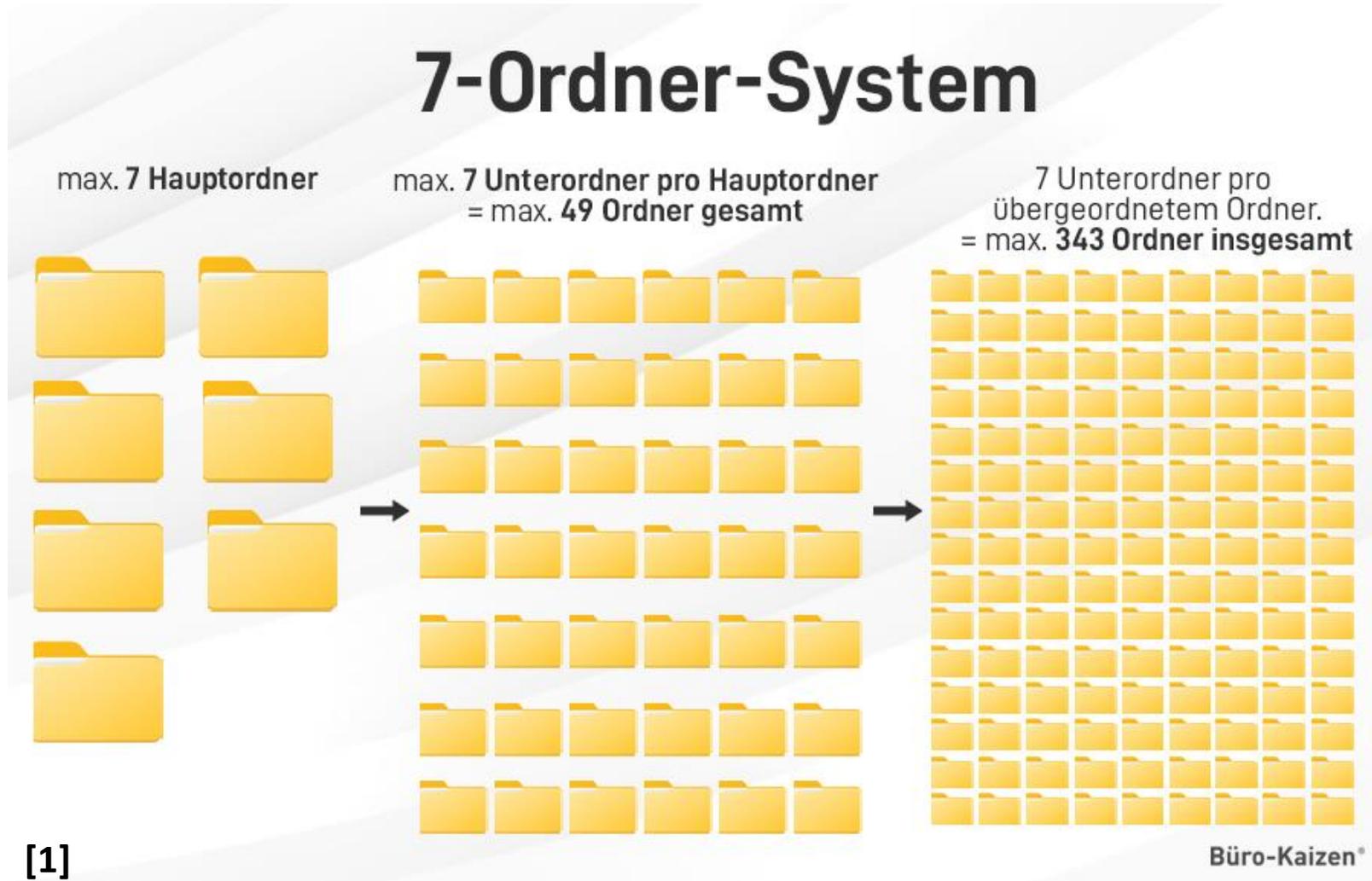


Ordnerstrukturen auf Servern sollten möglichst identisch mit denen auf lokalen Rechnern sein.



Abgestimmte Benennungskonventionen etc. schriftlich festhalten (z.B. Readme-Datei oder DMP)!

Beispiel: 7-Ordner-3-Ebenen-System



Beispiele Verzeichnisstruktur:

01_Organisatorisches

02_Literatur

03_Software

04_Messdaten

05_Analyse_Auswertung

06_Bildmaterial

07_Paper

01_Verwaltung

02_Kommunikation

03_Forschungsdaten

04_Auswertung

05_Dokumente

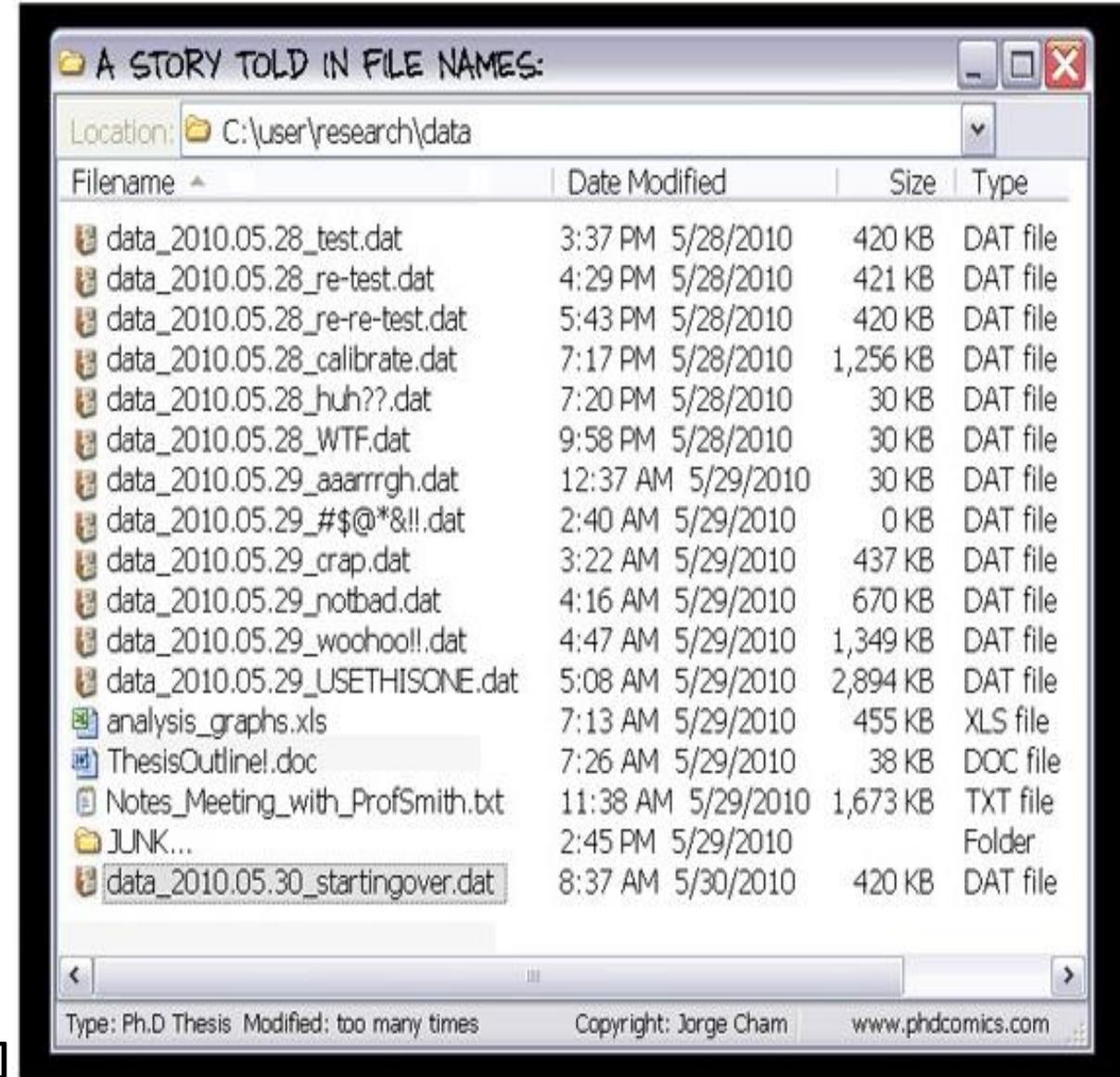
06_Abschluss

07_Sonstiges

Dateibenennung

Dateien und Ordner systematisch benennen und ordnen, damit...

- ...sie jetzt und in Zukunft leicht auffindbar und zugänglich sind.
- ...längeres Suchen oder das Vergleichen verschiedener Versionen vermieden wird.
- ...Änderungen nachvollziehbar sind.
- ...die Dateien nicht versehentlich gelöscht oder überschrieben werden.



[1]

Eine klare Dateibenennung ...

- ... ist objektiv, intuitiv und personenunabhängig nachvollziehbar
- ... ist einheitlich & konsistent
- ... ist eindeutig und lässt Rückschlüsse über den Inhalt zu
- ... ist so lang wie nötig und so kurz wie möglich
(zulässige Pfadlänge Windows 10: 256 Zeichen)
- ... vermeidet Leerzeichen und Sonderzeichen

Mögliche Namensbestandteile bei der Dateibenennung

- Inhaltsstichwort oder Kurztitel
- Name oder Kürzel des Erstellers
- Erstellungsdatum (z.B. JJMMTT oder JJJ-MM-TT)
- Bearbeitungsdatum
- Bezeichnung der Arbeitsgruppe / des Projekts
- Publikationsdatum
- Projektnummer
- Versionsnummer (z.B. V01 oder V_01)
- Anhängsel für Software, mit der die Datei erstellt wurde (z.B. .xls/.xlsx für Excel; .txt für Textdateien etc.)

Beispiele für einheitliche Namensgebung

Für chronologische Sortierung mit Datumsangabe beginnen:

z.B.: JJJJMMTT_Name

- 20160512_Klimamessung1_original.jpg
- 20160522_Klimamessung1_MHU_Ausschnitt.jpg
- 20160523_Klimamessung1_MHU_Ausschnitt_bearbeitet_Farbe.jpg



Automatisch generierte Namen (z.B. von Digitalkameras) vermeiden



Skalierbarkeit berücksichtigen: 01 oder 001

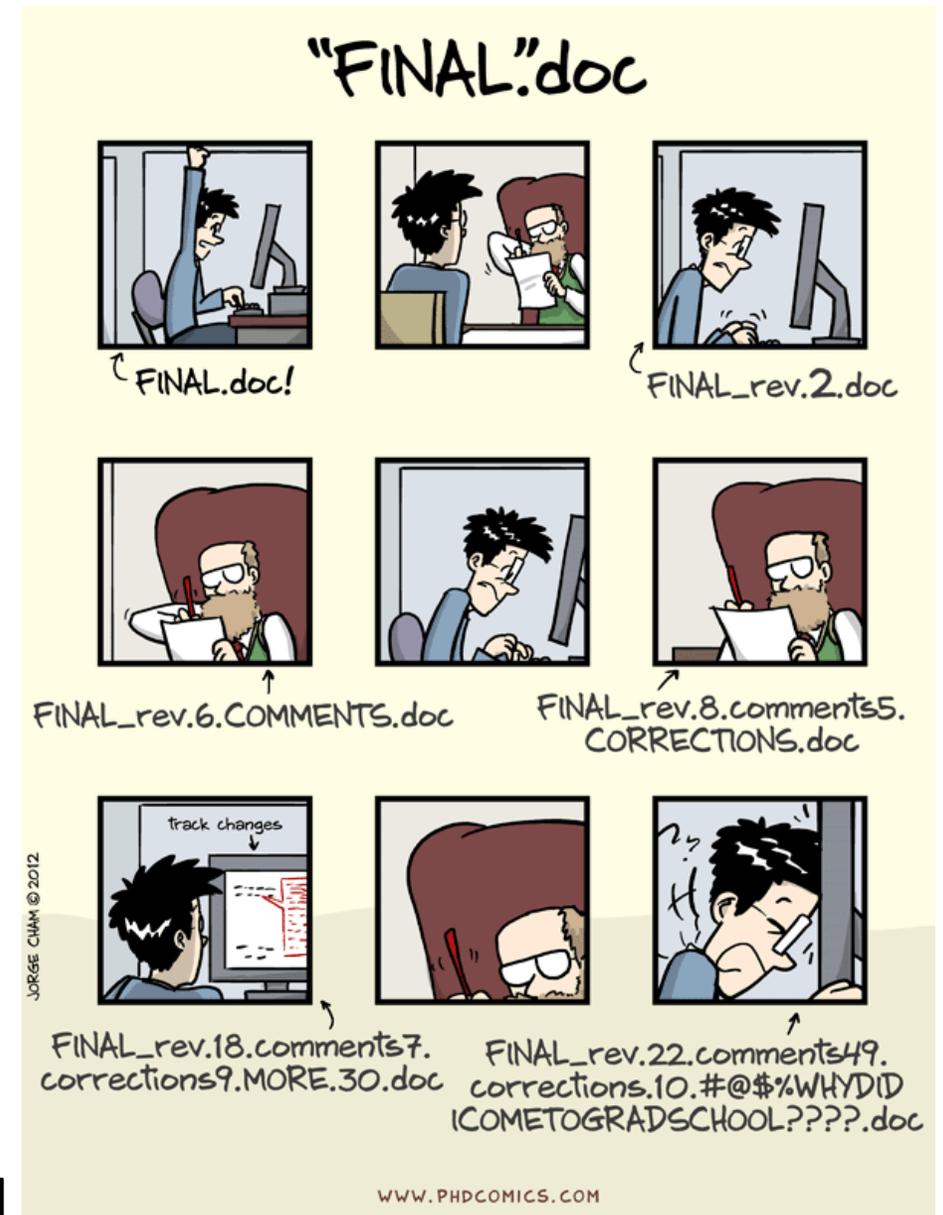
Versionierung im Dateinamen

- anhand natürlicher Zahlen (v1, v2, v3)
- Unterscheidung zwischen größeren und kleineren Änderungen

Beispieldatei_v1 -> Beispieldatei_v2
Beispieldatei_v1_1 -> Beispieldatei_v1_2

- verwirrende Zusatzbezeichnungen vermeiden;
- Dateinamen konstant beibehalten

Vgl.: <https://www.forschungsdaten-bildung.de/dateien-benennen>



[1]

Weitere Hinweise zur Versionierung

- jede Änderung vermerken (insbesondere in größeren Teams)
- veraltete Versionen getrennt speichern, verwerfen oder löschen
- die Originaldatei sollte stets erhalten bleiben
- bei großen Projekten bzw. umfangreichen Änderungen: Versionskontroll-Software zur Verwaltung von Versionen nutzen (z.B. Git, Subversion)
- automatische Versionierung und Änderungsverfolgung bei kollaborativen Dokumenten und Speicherorten (z.B. Wiki, Google Docs, Cloud-Dienste)

4.2.2. Datendokumentation

Die Rolle von Metadaten in der Datendokumentation

- Was sind Metadaten?
- Wozu dienen Metadaten?
- Metadaten-
Standards (generisch/fachspezifisch)

Weitere Formate der Datendokumentation

- Beispiele

Was muss alles dokumentiert werden?

Alles, was Dritte zum Verständnis des Datensatzes benötigen!



Was sind Metadaten?

- Beschreiben oder definieren andere Daten näher (z.B. Datensätze od. Code).
- Enthalten technische, rechtliche, administrative, fachliche und deskriptive Informationen (Datenvolumen, Datenformate, Lizenzen, Autor:in, Titel, Thema, Kurzbeschreibung, Schlagworte etc.).
- Werden unabhängig von oder zusammen mit den Daten, die sie beschreiben, gespeichert oder verknüpft.

rich metadata



poor metadata



[1]

Wozu dienen Metadaten?

- Sorgen für Auffindbarkeit und Nachvollziehbarkeit der Daten und Medien (z.B. in Datenbanken/Suchmaschinen)
- Erleichtern die maschinelle und automatisierte Verarbeitung von Daten
- Erleichtern die Nachnutzung (auch die eigene) und ermöglichen die Reproduzierbarkeit

Gut dokumentierte Daten mit ausführlichen Metadaten werden häufiger gefunden, zitiert und genutzt!



Metadaten-Standards

- Gewährleisten die Interoperabilität von Metadaten
- Hinweis: Repositorien haben meist eigene Metadatenstandards

Übersicht über eine Auswahl verschiedener Metadatenstandards:

- [Metadata Standards Catalog](#) (Research Data Alliance [RDA])
- [RDA Metadata Standards Directory Working Group](#)
- [FairSharing.org](#)
- [DDC \(Digital Curation Centre\)](#)

Tools zur Erstellung von Metadaten unterschiedlicher Disziplinen:

- [Metadata Standards Directory \(RDA\)](#)

[1]

Weitere Formate für die Datendokumentation:

Die Dokumentation kann z.B. erfolgen in

- einer begleitenden ReadMe-Datei
- einem projektinternen Wiki
- einem (elektronischen) Laborbuch
- einem Datenmanagementplan (DMP)
- systematisierten Ordnerstrukturen und Dateibenennungen (z.B. Autor, Titel, Thema, Schlagworte)

4. Forschungsdaten im Lebenszyklus

4.1. Planung des Forschungsvorhabens

4.2. Datenorganisation und -dokumentation

4.3. Datensicherung und -archivierung

4.4. Datenpublikation



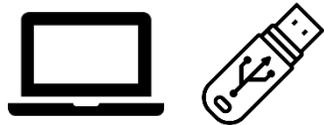
Datensicherung

- Regelmäßige **Backups** (Nutzung von Hochschulservern inkl. Backup-Systeme; festen Zeitpunkt festlegen)
- Sicherungskopien **getrennt von Originalen** aufbewahren
- Softwareanwendungen in die Sicherungsmaßnahmen einbeziehen
- Mehrere **Speichermedien** verwenden
- **Datenschutz** berücksichtigen (Pseudonymisierung und Anonymisierung personenbezogener Daten) => ggf. Datenschutzbeauftragten hinzuziehen
- **Zugriffsrechte** klären bzw. Zugriff sichern (z.B. durch Passwortschutz)

Die 3-2-1-Regel



3 Kopien der Datei



2 unterschiedliche Medien



1 externe Speicherung

4.2. Backup, Publikation, Archivierung: Ein Vergleich

	Backup	Publikation	Archivierung
Ziel	Aktive Nutzung	Teilen der Daten	Sichern finaler Daten
Zeitraum	kurzzeitig (z.B. Analysezeitraum, Projektlaufzeit)	Offen	langfristig bzw. dauerhaft (10 Jahre DFG-Vorgabe)
Metadaten	Hauptsächlich Zugriffssteuerungsliste (ACL) und Dateieigenschaften, Versionierung	FAIR und nach Vorgaben des Repositoriums	datenspezifisch, standardisiert

Besonderheiten der Langzeitarchivierung

Speicherung „für immer“

- Gut überlegen, was wirklich für die Nachwelt zur Verfügung stehen sollte
- Im Falle von personenbezogenen Daten auch den Datenschutz beachten (vgl. Art. 5 DSGVO)
- Verwendung von Formaten möglichst auf Basis offener, dokumentierter und nicht proprietärer Standards
- Vermeidung von Kompression
- Migration der Daten in aktuelle Datenformate
- Nutzung professioneller Archive

LZA – Nachhaltige Dateiformate

Datentyp	empfohlene Formate	weniger geeignete/ ungeeignete Formate
Audio	*.wav *.flac	*.mp3
Computer-aided Design (CAD)	*.dwg *.dxf *.x3d *.x3db *.x3dv	-
Datenbanken	*.sql *.xml	*.accdb *.mdb
Rastergrafiken & Bilder	*.tif (unkomprimiert) *.jp2 *.jpg2 *.png	*.gif *.jpeg *.jpg *.psd
statistische Daten	*.por	*.sav (IBM®SPSS)
Tabellen	*.csv *.tsv *.tab	*.xls *.xlsx *.xlx
Texte	*.odf *.rtf *.txt PDF/A	*.docx *.doc PDF
Vektorgrafiken	*.svg *.svgz	*.cdr
Video	*.mp4 *.mkv *.mj2 *.avi (unkomprimiert)	*.mov *.wmv

4. Forschungsdaten im Lebenszyklus

4.1. Planung des Forschungsvorhabens

4.2. Datenorganisation und -dokumentation

4.3. Datensicherung und Langzeitarchivierung

4.4. Datenpublikation



4.4. Datenpublikation - Inhalte

4.4.1. **Publikationswege** für Forschungsdaten

4.4.2. **Persistierende Identifikatoren** (DOI; ORCID)

4.4.3. **Vergabe von Nutzungsrechten** am Beispiel von Creative Commons-Lizenzen

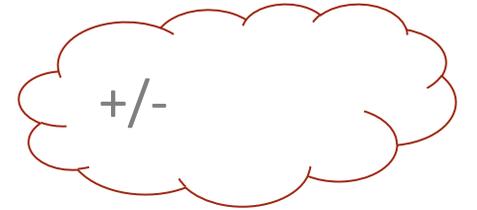
4.4.1. Publikationswege für Forschungsdaten



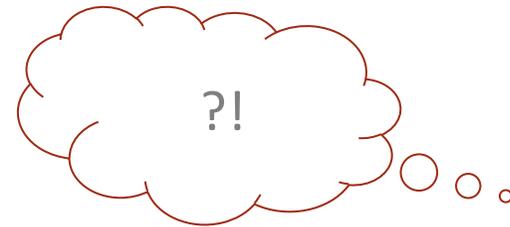
Publikation von Forschungsdaten **als eigenständiges Informationsobjekt** in einem Repository



Publikation der Datenbeschreibung in einem **Data Journal**



Publikation von Forschungsdaten **als Supplement** zur Veröffentlichung der Ergebnisse in Fachartikeln bei einem Verlag

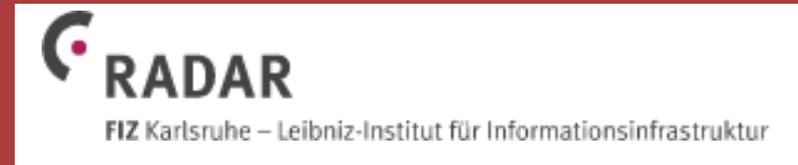


1010 Publikation in einem Repository

Repositorien = Speicherorte für digitale Objekte

Arten von Repositorien:

- Institutionell (z.B. einige Universitäten)
- Fachspezifisch (z.B. SowiDataNet, gesis, DARIAH-DE etc.)
- Disziplinübergreifend bzw. generisch (ZENODO, RADAR, GitHub, Figshare)



Wie finde ich ein Repository?

- Weltweit: www.re3data.org
- In Deutschland: risources.dfg.de,
www.nfdi.de (über Konsortium),
fairsharing.org
- Open Access Repositorien:
<https://v2.sherpa.ac.uk/opensoar/>

Auswahlkriterien:

- Archivierung mind. 10 Jahre
- Automatische Vergabe von persistierenden Identifikatoren (z.B. DOI)
- Zertifizierung (z.B. Core Trust Seal)
- Klare Nutzungsbedingungen (Entgelte, Embargofristen etc.)
- Metadatenstandards

Link:

[Checkliste. So erkennen Sie ein FAIRes Datenrepository](#)



Publikation in Data Journals

- Datenbeschreibung und Dokumentation der in einem Repository veröffentlichten Daten
- Qualitätssicherung durch Peer-Review-Verfahren
- Open-Access (z.B. Data in Brief) und Closed-Access-Zeitschriften
- Meist gebührenpflichtig

Link:

[Eine Liste mit Data Journals](#)

Publikation als Supplement

- Unterstützung und Verdeutlichung der Forschungsergebnisse
- Meist aggregierte Daten (Bilder, Tabellen etc.)

Link:

[Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B \(2020\) The citation advantage of linking publications to research data. PLoS ONE 15\(4\): e0230416. <https://doi.org/10.1371/journal.pone.0230416>](https://doi.org/10.1371/journal.pone.0230416)

Abstract

Introduction

Materials and methods

Results

Discussion

Conclusion

Data and code availability

Acknowledgments

References

Reader Comments

Figures

Citation: Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. <https://doi.org/10.1371/journal.pone.0230416>

Editor: Jelte M. Wicherts, Tilburg University, NETHERLANDS

Received: July 5, 2019; **Accepted:** February 28, 2020; **Published:** April 22, 2020

Copyright: © 2020 Colavizza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: Code and data can be found at: <https://doi.org/10.5281/zenodo.3470062>.

Funding: This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1 and by Macmillan Education Ltd, part of Springer Nature, through grant RG92108 "Effect of data sharing policies on articles' citation counts" granted to BM. Springer Nature provided support in the form of salaries for author IH, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section.

Competing interests: One of the authors (IH) is at the time of publication in the journal, employed by PLOS, publisher of PLOS ONE. IH was employed by Springer Nature, publisher of the BMC journals, at the time of planning and conducting the research and writing of the original manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials. There are no patents, products in development or marketed products associated with this research to declare. All other authors have declared that no other competing interests exist.

Publisher's Note: The article involves the independent analysis of data from

4.4.2. Persistente Identifikatoren

Digital Object Identifier (DOI)

- dienen der dauerhaften und eindeutigen Referenzierbarkeit von Forschungsdaten
- machen Daten zitierfähig
- werden für Artikel/Texte von Registrierungsagenturen (z.B. DataCite) über Verlage oder Bibliotheken vergeben
- werden für Forschungsdaten von Repositorien (z.B. Zenodo) vergeben

Open Researcher and Contributors ID (ORCID)

- legen eine dauerhafte, 16-stellige Identifizierungsnummer für Autor:innen / Forschende fest
- verknüpfen Publikationen eindeutig mit einer Person (auch bei Namensgleichheit)

Registrierung: <https://orcid.org/register> (für Autor:innen kostenfrei)

4.4.3. Vergabe von Nutzungsrechten – CC-Lizenzen



Namensnennung



Namensnennung, Weitergabe unter gleichen Bedingungen

Empfohlen im Sinne von
Open Science



Namensnennung, keine Bearbeitungen



Namensnennung, nicht kommerziell



Namensnennung, nicht kommerziell, Weitergabe unter gleichen Bedingungen



Namensnennung, nicht kommerziell, keine Bearbeitungen

[1]

Vorteile einer Publikation von Forschungsdaten

- ✓ Ermöglicht die Nachnutzung Ihrer Daten durch andere Wissenschaftler:innen
- ✓ Ermöglicht die Reproduzierbarkeit Ihrer Forschungsergebnisse
- ✓ Stärkt den wissenschaftlichen Austausch
- ✓ Reduziert Mehrfacherhebung von Daten
- ✓ Erhöht die Transparenz der Forschungsergebnisse durch Verfügbarkeit der Primärdaten

Agenda

START

01 Begrüßung
Vorstellung

02 Forschungsdaten,
Forschungsdatenmanagement
und warum das Ganze?
Spotlight: FDM in Anträgen

03 Open Science und
FAIRe Daten

04 Der Lebenszyklus von
Forschungsdaten
Spotlight: Datenmanagementplan

05 Evaluation und
Abschluss



ZIEL:
FDM - Basics

6. Veranstaltungsevaluation

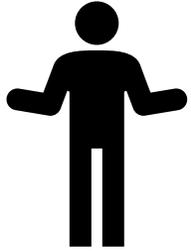
Direkter Link zur Umfrage:

<https://evasys.hochschulevaluierungsverbund.de/evasys/online.php?pswd=FDMB09>

Alternativ <https://evasys.hochschulevaluierungsverbund.de/evasys/online/> & folgende Losung eingeben: **FDMB09**

Denken sie an Ihre Teilnahmebestätigung!
fdm-haw@hs-mainz.de





Fragen und Antworten



Vertiefende Schulungen:

Ethik und Recht im
Forschungsdatenmanagement (FDM)

10.09.2024, 9-11 Uhr, Zoom

Dateimanagement –
FAIRes Dateimanagement

24.09.2024, 9-11 Uhr, Zoom

Links zum Thema

Umfassende Info-Plattform Forschungsdaten.info

<https://www.forschungsdaten.info>

Infonetzwirk: Forschungsdaten.org

<https://www.forschungsdaten.org/index.php/Hauptseite>

Services für die Bildungsforschung von kooperierenden Bildungsdatenzentren

<https://www.forschungsdaten-bildung.de/>

Kurse und Info-Tutorials:

<https://www.forschungsdaten.info/praxis-kompakt/tutorialskurse-und-anleitungen/>

FDM-FAQ

https://www.fdmhawrlp.de/?page_id=46

Sekundärliteratur zum Thema

Handbuch Forschungsdatenmanagement / hrsg. von Stephan Büttner ... – Bad Honnef : Bock + Herchen, 2011. - 223 S. E-book (lizenzfrei)

<http://opus.kobv.de/fhpotsdam/volltexte/2011/241/pdf/HandbuchForschungsdatenmanagement.pdf>

Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten : Grundlagen und praktische Lösungen für den Umgang mit quantitativen Forschungsdaten / Uwe Jensen, Sebastian Netscher, Katrin Weller (Hrsg.) - Opladen : Verlag Barbara Budrich, 2019. (233 Seiten). E-book (lizenzfrei)

http://www.contentselect.com/index.php?id=bib_view&ean=9783847412601

Nützliche Tools

- Wie FAIR sind Ihre Daten: <https://satisfyd.dans.knaw.nl/>
- Interaktiver Assistent, wie ich Datenschutz in meinem Projekt beachten sollte: https://wiki.bib.uni-mannheim.de/xerte/play.php?template_id=217#page1

Quellennachweise

Folie 14: [1] [Arthur, Charles \(23 August 2013\): "Tech giants may be huge, but nothing matches big data". The Guardian. Retrieved 30 April 2019.](#)

Folien 16-21: [1] Tabellen nach Boelter, Sarah, Schimmer, Thomas, Richter, Manuela, Schmidt, Diana, Schneidenbach, Esther (Februar 2024): [Und jetzt noch FDM?! Ein Erfahrungsaustausch zur Antragsberatung an HAW/FH](#), Love Data Week 2024.

Folie 25-26: [1] Wilkinson, Mark D., Dumontier, Michel, Aalbersberg, Ijsbrand J., Appleton, Gabrielle, Axton, Myles, Baak, Arie, Blomberg, Niklas, Boiten, Jan-Willem, da Silva Santos, Luiz Bonino, Bourne, Philip E., Bouwman, Jildau, Brookes, Anthony J., Clark, Tim, Crosas, Mercè, Dille, Ingrid, Dumon, Olivier, Edmunds, Scott, Evelo, Chris T., Finkers, Richard, Gonzlaez-Beltran, Alejandra, Gray, Alsdair J.G., Growth, Paul, Goble, Carole, Grethe, Jeffrey S., Heringa, Jaap, Hoen, Peter A.C't, Hooft, Rob, Kuhn, Tobias, Kok, Ruben, Kok, Joost, Lusher, Scott J., Martone, Maryann E., Mons, Albert, Packer, Abel L., Persson, Bengt, Rocca-Serra, Philippe, Roos, Marco, van Schaik, Rene, Sansone, Susanna-Assunta, Schultes, Erik, Sengstag, Thierry, Slater, Ted, Strawn, George, Swertz, Morris A., Thompson, Mark, van der Lei, Johan, van Mulligen, Erik, Velterop, Jan, Waagmeester, Andra, Wittenburg, Peter, Wolstencroft, Katherine, Zhao, Jun, Mons, Barend: The FAIR Guiding Principles for scientific data management and stewardship. In: *Sci Data* 3, 160018 (2016).

DOI: <https://doi.org/10.1038/sdata.2016.18>.

Folie 34: [1] vgl. Science Europe (2021): Practical Guide to the International Alignment of Research Data Management - Extended Edition. DOI: [10.5281/zenodo.4915862](https://doi.org/10.5281/zenodo.4915862).

Folie 37: [1] vgl.: Biernacka, Katarzyna, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020, S. 59 ff. DOI: <https://doi.org/10.5281/zenodo.3938533>.

Folie 42-43: [1] vgl.: Biernacka, Katarzyna, Buchholz, Petra, Dolzycka, Dominika, Helbig, Kerstin, Neumann, Janna, Odebrecht, Carolin, Wiljes Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020, S. 59 ff. DOI: <https://doi.org/10.5281/zenodo.3938533>.

Folie 48: [1] Vgl. Piwowar, Heather A., Vision, Todd J. : Data Reuse and the Open Data Citation Advantage, in: *PeerJ* 1 (1. Oktober 2013): e175. DOI: <https://doi.org/10.7717/peerj.175>.

Folie 49: [1] vgl. <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards/>.

Folie 50: [1] Vgl.: <https://forschungsdaten.info/themen/beschreiben-und-dokumentieren/datendokumentation>.

Folie 56: [1] in Anlehnung an: Biernacka, Katarzyna, Buchholz, Petra, Danker, Sarah Ann, Dolzycka, Dominika, Engelhardt, Claudia, Helbig, Kerstin, Jacob, Juliane, Neumann, Janna, Odebrecht, Carolin, Petersen, Britta, Slowig, Benjamin, Trautwein-Bruns, Ute, Wiljes, Cord, Wuttke, Ulrike: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 3.0. Berlin, 2020. DOI: <https://doi.org/10.5281/zenodo.3938533>.

FDM@HAW.rlp



Vielen Dank für Ihre Teilnahme!

