

Forschungsdatenmanagement

Notwendige, aber nicht hinreichende Voraussetzung
für den wissenschaftlichen Erkenntnisgewinn

05. Planen

Dr. habil. Till Biskup

Physikalische Chemie

Universität Rostock

03.05.2024





- 🔑 Planbarkeit der Forschung steht immer in einem gewissen Spannungsverhältnis zur Unvorhersehbarkeit der Ergebnisse.
- 🔑 Planung beginnt mit der Präzisierung der Forschungsfrage unter Einbeziehung der verfügbaren Literatur.
- 🔑 Notwendige Ressourcen sollten im Vorfeld abgeschätzt und die relevanten Akteure rechtzeitig involviert werden.
- 🔑 Rechtliche Fragen und deren Relevanz für das Projekt sollten im Vorfeld mit den relevanten Akteuren geklärt werden.
- 🔑 Die Verantwortlichkeit für die Planung liegt grundsätzlich bei der wissenschaftlichen Projektleitung.

Der Forschungsdatenlebenszyklus

Modell der wissenschaftlichen Methode aus Sicht der Forschungsdaten



Wissenschaftliche Aspekte

Organisatorische Aspekte

Verantwortung

Anforderungen und notwendige Werkzeuge

- ▶ Forschungsfrage präzisieren
 - unter Einbeziehung der verfügbaren Literatur
 - setzt Kompetenzen bzgl. Literaturrecherche voraus
 - Stand der Forschung kennen und berücksichtigen
- ▶ Überblick verschaffen über mögliche Werkzeuge zur Datenerhebung und -Auswertung
 - Datenerhebung: Experiment, Umfrage, Literaturrecherche, ...
 - Was kann/habe ich selbst? Wo brauche ich Kooperationen?
 - Was kann/will ich selbst entwickeln/etablieren?
- ▶ Gibt es existierende Daten, die genutzt werden können/sollen?
 - Sind die Daten verfügbar?
 - Sind die Daten (prinzipiell) geeignet?
 - Ist die Qualität der Daten hinreichend sichergestellt?

- ▶ Welche Daten sollen warum erhoben werden?
 - Daten zu erheben ist vergleichsweise einfach geworden.
 - Typisches Problem: zu viele bzw. ungeeignete Daten
 - Das Feld „Zweck“ in den Metadaten zur Datenerhebung sollte immer bewusst ausfüllbar sein – und diese Perspektive bereits in der Planung vor der Datenerhebung eingenommen werden.
- ▶ Wie sollen die Ergebnisse veröffentlicht werden?
 - Wissenschaft lebt vom Austausch (vgl. das Newton-Zitat).
 - Können/sollen/dürfen Daten veröffentlicht werden?
- ▶ Vorgehen planen
 - Untersuchungsmethoden, Probenherstellung, Kooperationen
 - Abschätzung der anfallenden Datenmengen und -Formate
 - Auswertung und dazu notwendige Infrastruktur und Software
- ☞ Aspekte werden typischerweise in einem Projektantrag behandelt.

- ❓ Ist Planung mit dem Charakter der Wissenschaft vereinbar und nicht eine unzulässige Einschränkung der Freiheit der Forschung?
- ▶ Wissenschaft ist ein systematisches Unterfangen.
 - Wissenschaft transzendiert Raum, Zeit und Forschende.
 - Forschende tragen Verantwortung gegenüber anderen.
- ▶ Wissenschaft ist letztlich nicht vorhersehbar.
 - Planung ist immer bis zu einem gewissen Grad vorläufig.
 - Details ergeben sich aus Erfahrung und mit zunehmender Durchdringung der Fragestellung.
- ▶ Vorausplanung ist notwendig – aber immer vorläufig.
 - Unvorhersehbarkeit liefert Kriterien für Art der Planung und eingesetzte Werkzeuge: Flexibilität, Modularität, Erweiterbarkeit.
 - Viele allgemeine Aspekte sind immer anwendbar und notwendig, egal wie sich Fragestellung und Situation (weiter) entwickeln.

These

Freiheit der Forschung und Wissenschaft ist nicht die Freiheit, nicht zu planen bzw. unstrukturiert vorzugehen – weil das weder der Verantwortung gegenüber der Wissenschaft noch gegenüber der Gesellschaft gerecht wird –, sondern die Freiheit, die Forschungsfrage selbst zu definieren und immer wieder zu ändern – sowohl aufgrund der Unvorhersehbarkeit als auch des zunehmenden Verständnisses der Fragestellung.

- 👉 Freiheit geht immer mit Verantwortung einher.
 - Je mehr Freiheit, desto mehr Verantwortung.
- 👉 Mangelnde Planung ist unprofessionell und meist eine Ausrede.

Wissenschaftliche Aspekte

Organisatorische Aspekte

Verantwortung

Anforderungen und notwendige Werkzeuge

- ▶ Planung umfasst *alle* Stationen des Forschungsdatenlebenszyklus.
 - Wissenschaft setzt ein systematisches (geplantes) Vorgehen voraus.
 - „Vom Ende her denken“: Weiterverwendung von Erkenntnissen (nicht notwendigerweise Weiterverwendung von Daten)
- ▶ Forschungsdatenlebenszyklus ist ein Planungsinstrument
 - Das Modell ist nicht die Umsetzung, es beschreibt sie höchstens.
 - Spannungsfeld: Was wird wo thematisiert?

Unterschiedliche Arten von Planung auf unterschiedlichen Ebenen

- ▶ Gesamtstruktur: im Vorfeld festlegen
 - Werkzeuge und Methoden der Forschungsfrage angemessen
 - Datenablage: organisiert, nachvollziehbar, langfristigkeit(er) sicher
 - Sicherstellung der Vollständigkeit erhobener Metadaten
 - rechtliche Aspekte
- ▶ Konkrete Umsetzung und Einsatz von Werkzeugen
 - überwiegend auf der Sachebene in der Praxis entscheiden

- ▶ wissenschaftliche Aspekte schriftlich festhalten
 - Projektantrag
 - alternativ Projektskizze für „interne“ Projekte
- ▶ „klassische“ Projektplanung
 - Ziele und ihre Überprüfbarkeit festlegen
 - Strukturen etablieren zur Projektverwaltung:
regelmäßige Treffen, Überblick über Projektfortschritt behalten
 - viel wird (implizit) gemacht, mitunter fehlt es aber auch
 - ein Unterschied zwischen guter und schlechter Betreuung
- ▶ andere Stationen des Forschungsdatenlebenszyklus mitdenken
 - auf noch eher abstraktem Niveau
 - für alle Arten von Forschungsdaten
(auch Werkzeuge/Software mitdenken/berücksichtigen)
 - Hilfsmittel: Daten- bzw. Softwaremanagementplan (DMP, SMP)

- ▶ Modularisierung
 - Verteilen der Aufgaben auf Verantwortliche
 - Aufteilung in Unterprojekte und Arbeitspakete
- ▶ notwendige Infrastruktur
 - Sind die Voraussetzungen für die Datenerhebung erfüllt? (Geräte, Archive, Sammlungen, Bibliotheken, ...)
 - Umfang zu erwartender Forschungsdaten grob abschätzen und notwendige Infrastruktur zur Speicherung rechtzeitig beschaffen
- ▶ rechtliche Aspekte abklären
 - Urheberrecht, Datenschutz, Lizenzen, weitere Schutzrechte (Datenbanken, Geschäftsgeheimnisse, Patentrecht, ...)
 - Verantwortliche der Institution kennen und involvieren
- ▶ Beachtung von Leitlinien zum Umgang mit Forschungsdaten
 - Quellen: lokale Institution, Fachgesellschaften, Fördermittelgeber
 - lokale Institution: verantwortliche Stelle kennen, ggf. kontaktieren

Wissenschaftliche Aspekte

Organisatorische Aspekte

Verantwortung

Anforderungen und notwendige Werkzeuge

- ▶ individuelle Verantwortlichkeiten
 - Gruppen-/Projektleitung
 - individuelle Forschende
- ▶ institutionelle Verantwortlichkeiten
 - (Heimat-)Institution
 - Forschungsförderer

Leitmotiv

Forschungsdatenmanagement ist primär die Verantwortung der individuellen Forschenden.

- ☞ Erst wenn wir unsere Aufgaben erfolgreich erledigt haben, können wir mit dem Finger auf andere zeigen...

Gruppen-/Projektleitung

- ▶ Formulieren der (übergreifenden) Forschungsfrage
- ▶ Überblick über den Stand der Forschung (Forschungsgebiet)
- ▶ Überblick über den aktuellen Stand des (Gesamt-)Projektes
- ▶ relevante organisatorische Aspekte möglichst rechtzeitig klären
- ▶ Ermöglichen und Einfordern der Planung durch die Durchführenden
- ▶ Sicherstellen, dass die notwendigen Kompetenzen bei den Durchführenden vorhanden sind
 - ggf. selbst aktiv werden (Lehre)
 - alternativ auf existierende Angebote verweisen
 - notwendigen Kompetenzerwerb ermöglichen und einfordern

individuelle Forschende

- ▶ Formulieren der konkreten Forschungsfrage
- ▶ Überblick über den Stand der Forschung (Forschungsfrage)
 - Tipp: Theorieteil vielleicht nicht erst anfangen zu schreiben, wenn man alle Experimente gemacht hat
 - Erpart mitunter die Enttäuschung, wenn man feststellt, dass andere vor Jahrzehnten die Frage längst beantwortet haben.
- ▶ systematisches Vorgehen: setzt Plan (im Kleinen) voraus
- ▶ (eigenständiger) Erwerb der notwendigen Kompetenzen
- ▶ Bewusstsein für die Notwendigkeit der Planung
- ▶ ständiger Realitätsabgleich:
wieviel Planung muss sein, passen die Planungswerkzeuge?

(Heimat-)Institution

- ▶ Ansprechpartner für juristische Fragen
 - Datenschutz und weitere Schutzbedarfe
 - Recht an Daten (Urheberrecht)
- ▶ ggf. Ansprechpartner für technische/organisatorische Fragen
 - Aufstellung von Geräten
 - Nutzung institutioneller Infrastruktur
 - Bereitstellung von Arbeitsmaterial
 - vorhandene Lösungen für die Speicherung/Ablage von Daten jeglicher Art (digital, analog, physische Objekte, ...)
- ▶ Leitlinien zum Umgang mit Forschungsdaten

Forschungsförderer

- ▶ Leitlinien zum Umgang mit Forschungsdaten
- ▶ Vorlagen für Datenmanagementpläne (DMP)

Wissenschaftliche Aspekte

Organisatorische Aspekte

Verantwortung

Anforderungen und notwendige Werkzeuge

- ▶ Anforderungen zu formulieren ist notwendiger Ausgangspunkt
 - Anforderungsanalyse liefert Kriterien für die Wahl der Werkzeuge
 - das Richtige machen, nicht das Falsche/Unwichtige richtig machen
 - Anforderungen zu formulieren setzt eine intellektuelle Durchdringung der Fragestellung voraus und ist nicht trivial.
- ▶ Werkzeuge werden hier (noch) nicht im Detail vorgestellt
 - Beispiele teilweise in den weiteren Abschnitten der Vorlesung
 - Werkzeuge sind immer (auch) kontextspezifisch.
- ▶ Werkzeuge sind nicht nur bzw. oft nicht digital.
 - Digitalität ist kein Qualitätskriterium.
- ▶ Werkzeuge sind keine Lösungen
 - Werkzeuge können durch kompetenten Einsatz und spezifische Anpassung an die jeweiligen Gegebenheiten zu Lösungen werden.

These

Nur Systeme, die hinreichend einfach nutzbar sind und deren Verwendung offensichtliche Vorteile bietet, werden genutzt werden.

- ▶ Generell: dem Charakter der Wissenschaft entsprechend
 - Kriterien: Flexibilität, Modularität, Erweiterbarkeit
 - möglichst minimaler Verwaltungsaufwand
- ▶ Formulierung der Forschungsfrage etc.
 - ausdrucksstarke Formate: Referenzen, mathematische Formeln, ...
- ▶ Daten-/Softwaremanagementplan
 - Aktualisierung muss für alle Beteiligten einfach möglich sein
 - setzt entsprechende Dokumentformate und Ablageorte voraus
 - Versionierung?

These

Lebenszeit skaliert nicht, ein geschriebenes Dokument tendenziell schon.

- ▶ Dokumentations- und Ablagesysteme
 - modulare, hierarchische Verzeichnisstrukturen
 - flexible Dateiformate
- ▶ Strategien, um auf dem Stand der Forschung zu bleiben
 - Literaturrecherche für die Präzisierung der Forschungsfrage
 - Netzwerke, Newsletter, Email-Alerts, ...
- ▶ administrative Aspekte
 - Überblick über die rechtliche Situation
 - (institutionelle) Ansprechpartner für entsprechende Rückfragen
- ▶ Vorlagen für einen Daten-/Softwaremanagementplan

- ▶ Datenmanagementplan (DMP)
 - Welche Daten werden erzeugt (Art, Umfang)?
 - Wie werden die Daten dokumentiert und organisiert?
 - Wie werden Daten gespeichert und ggf. gesichert?
 - Was passiert mit den Daten nach Projektende?
 - Wie werden die Daten zur Weiterverwendung/Nachnutzung bereitgestellt, soweit möglich und notwendig?

- ▶ Softwaremanagementplan (SMP)
 - Wie wird die Software versioniert?
 - Wie wird die Software veröffentlicht, lizenziert und verteilt?
 - Wie ist die Software für alle relevanten Gruppen dokumentiert?
 - Wie wird die Qualität der Software sichergestellt?
 - Wie wird die langfristige Wartung/Wartbarkeit gewährleistet?
 - Wer hängt von der Verfügbarkeit der Software ab?

DMP: K. Briney: Data Management for Researchers. Pelagic, Exeter 2015, S. 34

SMP: Martinez-Ortiz *et al.*: Practical Guide to Software Management Plans. doi:10.5281/zenodo.7589725



- 🔑 Planbarkeit der Forschung steht immer in einem gewissen Spannungsverhältnis zur Unvorhersehbarkeit der Ergebnisse.
- 🔑 Planung beginnt mit der Präzisierung der Forschungsfrage unter Einbeziehung der verfügbaren Literatur.
- 🔑 Notwendige Ressourcen sollten im Vorfeld abgeschätzt und die relevanten Akteure rechtzeitig involviert werden.
- 🔑 Rechtliche Fragen und deren Relevanz für das Projekt sollten im Vorfeld mit den relevanten Akteuren geklärt werden.
- 🔑 Die Verantwortlichkeit für die Planung liegt grundsätzlich bei der wissenschaftlichen Projektleitung.