

Forschungsdatenmanagement

Notwendige, aber nicht hinreichende Voraussetzung
für den wissenschaftlichen Erkenntnisgewinn

17. Bewährte Verfahren (aus eigener Anschauung)

Dr. habil. Till Biskup

Physikalische Chemie

Universität Rostock

05.07.2024





- Die Entwicklung von Werkzeugen, Strategien und Lösungen setzt die eigene aktive Forschungstätigkeit voraus.
- Schlüssel zum Erfolg ist die langjährige praktische Erfahrung und der reflektierte Umgang mit dem eigenen Scheitern.
- Die Entwicklung von Lösungen setzt Kompetenzen in der Fachwissenschaft und der Werkzeugimplementierung voraus.
- Lösungen sind immer kontextspezifisch, beruhen aber meist auf Werkzeugen und bewährten Verfahren (Mustern).
- Lösungen müssen dezentral, modular, flexibel erweiterbar, robust, resilient, attraktiv, bewährt und langfristig verfügbar sein.

- 1 Ausgangspunkt und Motivation
 - Wissenschaft; Datenmenge und Digitalität
- 2 Aspekte eines forschungsnahen Forschungsdatenmanagements
 - Forschungsdatenlebenszyklus
- 3 Bausteine eines funktionierenden, individuellen Forschungsdatenmanagements
 - Eigenschaften; Prinzipien; notwendige Kompetenzen; Werkzeuge
- 4 Hindernisse und Probleme
 - Hindernisse und mögliche Abhilfe; Antimuster
- 5 **Funktionierende Lösungen**
 - Bewährte Verfahren (aus eigener Anschauung)

These

Es gibt bewährte Verfahren (Muster),
die für konkrete Lösungen herangezogen werden können.

- ▶ Viele bewährte Verfahren existieren in anderen Fachbereichen (z.B. Softwareentwicklung) und scheinen innerhalb des Forschungsdatenmanagements unbekannt zu sein.
- ▶ Bewährte Verfahren liefern nur Vorlagen, die immer an die realen Gegebenheiten angepasst werden müssen. Das setzt ein entsprechendes Verständnis und Kompetenzen im Umgang mit den zugrundeliegenden Werkzeugen voraus.
- ▶ Der Einsatz der Verfahren erfordert eine Kosten–Nutzen-Analyse, da sie immer sowohl Vor- als auch Nachteile haben.

Schlüssel zum Erfolg: langjährige praktische Erfahrung

Muster, keine schlüsselfertigen Lösungen

Kurze Wiederholung: Anforderungen

Übersicht: nachfolgend vorgestellte Lösungen

Leitmotiv

Forschungsdatenmanagement ist primär die Verantwortung der individuellen Forschenden.

These

Den meisten, die sich offiziell mit Forschungsdatenmanagement beschäftigen, fehlt die praktische Kompetenz und Erfahrung für die Entwicklung und Implementierung funktionierender Lösungen. Nur wer selbst hinreichend an seinen Daten gescheitert ist und darüber reflektiert und die Gründe für das Scheitern verstanden hat, wird in der Lage sein, tragfähige Lösungen zu erarbeiten.

Schlüssel zum Erfolg: langjährige praktische Erfahrung

- ▶ Bedeutung für individuelle Forschende
 - Persönliche Erfahrung lässt sich nur machen, nicht einkaufen.
 - (Neue) Lösungen können nur mit entsprechender Erfahrung entwickelt, implementiert und angepasst werden.
 - Gewissenhafte und reflektierte Nutzung ist schon Studierenden ab dem ersten Semester (in den jeweiligen Grenzen) zuzutrauen.
- ▶ Grundlage der vorgestellten Lösungen
 - persönliche langjährige praktische Erfahrung
 - intellektuelle Durchdringung von Abläufen und Zusammenhängen
 - kritische Reflexion des wiederholten persönlichen Scheiterns an den eigenen Ansprüchen (und jenen der Wissenschaft)
 - Erfahrung sowohl des Dozenten als auch von Personen, mit denen er im engen Austausch stand und steht

These

Man kann aus den Fehlern anderer lernen,
vor allem aber auf der Erfahrung anderer (positiv) aufbauen.

- ▶ Vorarbeiten anderer
 - oftmals langjährige technische Mitarbeitende
- ▶ Voraussetzungen
 - fruchtbarer Austausch über Anforderungen und Lösungsansätze
 - auf der Basis je eigener persönlicher Erfahrung
 - intellektuelle Durchdringung der Problemstellung
 - Reflexion über die eigenen Lösungen

These

Die Entwicklung von Werkzeugen, Strategien und Lösungen für das Forschungsdatenmanagement setzt die praktische Einbindung in die Forschung und das eigene Erleben der Notwendigkeiten und Bedarfe sowie der tatsächlichen Rahmenbedingungen voraus.

zwei Personengruppen

- ▶ **Strategen**
 - fachwissenschaftliche Kompetenz und Erfahrung
 - Kenntnis und Erfahrung bzgl. Implementierung der Werkzeuge
- ▶ **Implementierer**
 - notwendige Fachexpertise für die Umsetzung
 - selbstständige Implementierung in einem vorgegebenen Rahmen

Schlüssel zum Erfolg: langjährige praktische Erfahrung

Muster, keine schlüsselfertigen Lösungen

Kurze Wiederholung: Anforderungen

Übersicht: nachfolgend vorgestellte Lösungen

These

Es gibt (meist) keine schlüsselfertigen Lösungen – nur Werkzeuge. Werkzeuge müssen immer an die individuellen Gegebenheiten angepasst und der kompetente Umgang mit ihnen erlernt werden.

- ▶ Werkzeuge mit Lösungen zu verwechseln, führt meist zu Frust.
 - Fehlende Anpassung an die individuellen Gegebenheiten führt bestenfalls für Mehrarbeit, schlimmenenfalls zu Zielverfehlung.
 - mangelndes Bewusstsein für die notwendige Einarbeitung (sowohl seitens der Nutzenden als auch der jeweiligen Leitung)
- ▶ bewährte Verfahren: Muster
 - Beschreibung eines Problems und einer abstrakten Lösung
 - Kontextualisierung des Einsatzes geeigneter Werkzeuge

Muster (nicht nur) in der Softwareentwicklung

- ▶ Muster beschreiben sowohl ein wiederkehrendes Problem als auch eine abstrakte (mögliche) Lösung.
- ▶ Muster müssen immer in einem konkreten Kontext implementiert werden (Muster selbst sind keine Lösung).
- ▶ Muster können zu wenig und zu viel verwendet werden.
- ▶ Muster beschreiben in der Praxis *bewährte* Verfahren.
- ▶ Muster beinhalten notwendigerweise eine Kosten–Nutzen–Abwägung.

Ursprung von Mustern

1977 Christopher Alexander (Architekturtheoretiker)

1995 Gamma *et al.*: Entwurfsmuster (Software)

Schlüssel zum Erfolg: langjährige praktische Erfahrung

Muster, keine schlüsselfertigen Lösungen

Kurze Wiederholung: Anforderungen

Übersicht: nachfolgend vorgestellte Lösungen

These

Funktionierende Werkzeuge und Lösungen müssen dezentral, modular, flexibel und erweiterbar sein.

- ▶ Dezentral: Nur Forschende selbst können im realen Forschungsalltag die Komplexität der relevanten Abläufe erfassen.
- ▶ Modular: Die Aufgabe ist zu groß, als dass sie von einem Team auf einmal angegangen werden kann. Modularität ermöglicht Priorisierung und überhaupt erst die (schrittweise) Umsetzung.
- ▶ Flexibel: Ansprüche ändern sich ständig, auch aufgrund der fortschreitenden intellektuellen Durchdringung der Fragestellung.
- ▶ Erweiterbar: Anforderungen entwickeln sich ebenso mit der fortschreitenden intellektuellen Durchdringung der Fragestellung.

► dezentral

- unter Kontrolle der individuellen Forschenden
- unabhängig von zentraler Infrastruktur
- geringere Angriffsfläche
- tendenziell geringere Komplexität
- trägt zur Robustheit bei
- dezentral bedeutet *nicht* (ausschließlich) lokal

► modular

- klar definierte Schnittstellen
- Links statt ein Gesamtsystem
- einzelne Komponenten austauschbar
- einzige Chance für intellektuelle Beherrschbarkeit der Bausteine
- mit aktuellen Ressourcen umsetzbar
- Voraussetzung für Flexibilität und Erweiterbarkeit

- ▶ flexibel
 - intrinsische Anforderung der Wissenschaft
 - Aufgaben sind ggf. ähnlich, aber immer individuell
 - Voraussetzung: Modularität und damit Schnittstellen
 - Balance zwischen Freiheit und Komplexität
- ▶ erweiterbar
 - intrinsische Anforderung der Wissenschaft
 - Freiheit, nur das gerade Notwendige zu implementieren
 - Voraussetzung: Modularität und damit Schnittstellen
 - Herausforderung: Systemarchitektur
 - Balance zwischen Generalisierbarkeit und Beherrschbarkeit

► robust

- Toleranz gegenüber Fehlbedienung/falschen Nutzereingaben
- Langlebigkeit/Langfristigkeit (s.u.)
- (numerische) Stabilität von Lösungen
- Idempotenz (gleicher Systemzustand bei mehrfacher Ausführung)
als Voraussetzung für Fehlertoleranz komplexer Abläufe
- Gegenteil von Zerbrechlichkeit bei (Software-)Systemen
und damit Voraussetzung für wartbare und erweiterbare Systeme

► resilient

- Forschungsdatenmanagement ist komplex und daher fehleranfällig
- Redundanz durch Vorhalten mehrerer gleicher Systeme
- Backup, um im Fehlerfall Informationen wiederherzustellen
- technisch komplett unabhängige Verfahren (Bsp.: Papier/digital)
- automatisierte Konfiguration von (Ersatz-)Systemen

- ▶ attraktiv
 - Nur hinreichend einfach bedienbare Systeme mit öffentlichen Vorteilen ihrer Nutzung werden genutzt werden.
 - intuitiv: Anordnung von Elementen, offensichtliche Bedienung
 - automatisiert: Konzentration auf relevante Aspekte
 - robust: tolerant gegenüber Fehlbedienung
 - ermöglichend: hilft, Ansprüchen gerecht zu werden
 - wirkmächtige Abstraktionen bietend
- ▶ plattformunabhängig
 - digitale Werkzeuge: Unterstützung aller gängigen Betriebssysteme
 - sichert Kompatibilität und Austauschbarkeit
- ▶ keine Bindung an einen Anbieter (*vendor lockin*)
 - Wissenschaft lebt von Offenheit und Unabhängigkeit
 - setzt offengelegte und hinreichend dokumentierte Schnittstellen und Austauschformate voraus

- ▶ langfristig wartbar
 - Umfeld ändert sich, Werkzeuge müssen anpassbar sein
 - komplexe, teure Werkzeuge: verantwortungsvoller Ressourceneinsatz
 - Voraussetzung: konzeptionelle/technische Dokumentation
- ▶ aktiv gewartet
 - ohne aktive Wartung ist Funktionalität nicht sicherzustellen
 - Wartung i.d.R. nicht lokal umsetzbar (mangelnde Ressourcen)
- ▶ bewährt
 - Werkzeuge sind meist Mittel zum Zweck der Zielerfüllung.
 - Hilfe von anderen Nutzenden verfügbar
 - Voraussetzung: hinreichende Verbreitung und Nutzerbasis
- ▶ langfristige Perspektive
 - Wissenschaft denkt in Generationen . . .
 - kompetente Nutzung setzt Einarbeitung voraus
 - Zeit für die Einarbeitung muss gut investiert sein

Schlüssel zum Erfolg: langjährige praktische Erfahrung

Muster, keine schlüsselfertigen Lösungen

Kurze Wiederholung: Anforderungen

Übersicht: nachfolgend vorgestellte Lösungen

Vorgestellte Lösungen

- ▶ Metadaten während der Datenerhebung
 - Schema(ta) und Dateiformat
 - Fokus: menschen-schreibbar und maschinen-verarbeitbar
- ▶ Lückenloses Protokoll der Datenverarbeitung
 - Gesamtsystem zur wissenschaftlichen Datenauswertung
 - Fokus: Automatisierung, Nachvollziehbarkeit
- ▶ Elektronisches Laborbuch
 - Dokumentation der Abläufe im Labor
 - Fokus: Struktur, Nachvollziehbarkeit
- ▶ Modulare digitale (dezentrale) Infrastruktur
 - Wissens- und Projektmanagement, lokale PIDs, Datenablage, ...
 - Fokus: Organisation, Dokumentation, Nachvollziehbarkeit

Gliederung der jeweiligen Präsentationen

- ▶ Problemstellung, Herausforderungen und Anforderungen
 - Welche Fragestellung wird adressiert?
 - Was sind die Herausforderungen und Anforderungen?
- ▶ Konzept und Umsetzung
 - Wie sieht die Lösung aus?
 - Wie wird sie konkret umgesetzt?
- ▶ Anwendung auf die eigene Praxis
 - Wie lässt sich das Konzept auf die eigene Praxis anwenden?
 - Was gibt es dabei zu beachten?
- ▶ Vorteile gegenüber Alternativen
 - Welche Vorteile bietet die vorgestellte Lösung?
 - Wie wird sie den aufgestellten Kriterien gerecht?



- Die Entwicklung von Werkzeugen, Strategien und Lösungen setzt die eigene aktive Forschungstätigkeit voraus.
- Schlüssel zum Erfolg ist die langjährige praktische Erfahrung und der reflektierte Umgang mit dem eigenen Scheitern.
- Die Entwicklung von Lösungen setzt Kompetenzen in der Fachwissenschaft und der Werkzeugimplementierung voraus.
- Lösungen sind immer kontextspezifisch, beruhen aber meist auf Werkzeugen und bewährten Verfahren (Mustern).
- Lösungen müssen dezentral, modular, flexibel erweiterbar, robust, resilient, attraktiv, bewährt und langfristig verfügbar sein.