



Forschungsdaten im digitalen Wandel

Aus der Arbeit der Deutschen Forschungsgemeinschaft (DFG)

Digitale Coffee Lecture, Hildesheim, 28.10.2020

Dr. Johannes Fournier

Gedankengang

1. Daten in der Forschung
2. Digitaler Wandel
3. Normative Gestaltung durch die DFG
4. Praktische Gestaltung durch die DFG
5. Internationale und wissenschaftspolitische Gestaltung



1. Warum sind Daten für die Forschung wichtig?

- ▶ Wissenschaftliches Arbeiten ermöglichen
- ▶ Forschungsergebnisse nachvollziehbar machen
- ▶ Anschlussverwendung (auch durch Dritte) stimulieren
- ▶ War das nicht schon immer so?
- ▶ Was hat sich denn geändert?

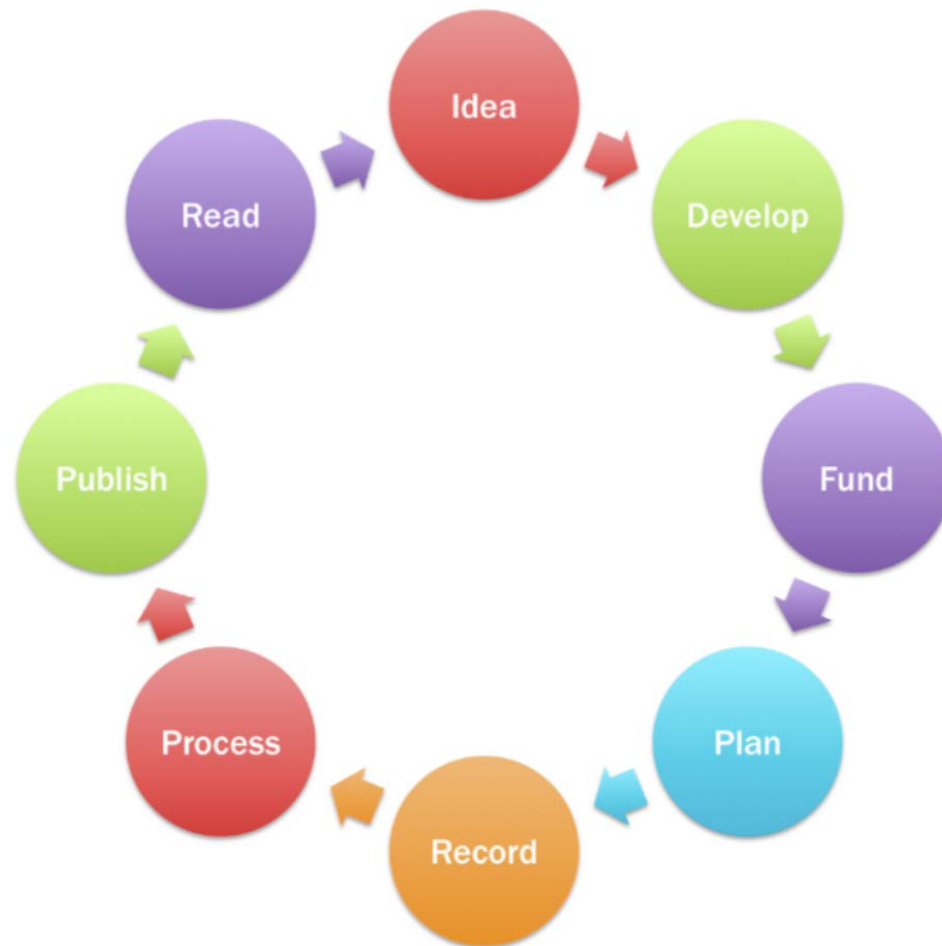


1. Wo zeigt sich der Einfluss digitaler Technologien?

Daten und Software als Schlüsselemente des Wandels

**... an jedem einzelnen Punkt
des Forschungszyklus ...**

Cameron Neylon:
Research Cycle
https://commons.wikimedia.org/wiki/File:Research_cycle.png
(CC BY 2.0)



2. Wie macht sich der digitale Wandel bemerkbar?

Idealtypische Ausprägungen

► **Transformativer Wandel**

- Überführung analoger Informationen und Praktiken in digitale Formate
- Digitalisierung von Texten, Bildern, Objekten usw., Suche und Erhebung

► **Ermöglicher Wandel**

- Nutzung datenintensiver Technologien zur Bearbeitung von Forschungsfragen, die sonst überhaupt nicht bearbeitet werden könnten (z.B. OMICS oder Kernphysik)

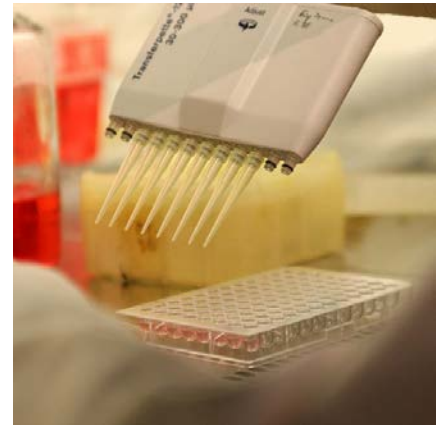
► **Substituierender Wandel**

- Nutzung digitaler Technologien, um konzeptionelle Teile des Forschungsprozesses zu unterstützen oder gar abzulösen
 - Ersetzen von Experimenten durch Modellierungen und Simulationen
 - Anwendung neuronaler Netze anstelle einer vorausgehenden Hypothesenbildung
- *Vollständige Neuorientierung des Forschungsansatzes*

2. Wie zeigt sich der Daten-bezogene Wandel noch?

► **Veränderte Forschungspraktiken durch ...**

- Skalierungseffekte (Volumen, Geschwindigkeit, Komplexität), die Erkenntnismöglichkeiten erweitern
- Kombination höchst umfangreicher Datensätze, die ein Prüfen von Entstehungskontexten und Erhebungsmethoden erheblich erschweren



► **Erhöhte Aufmerksamkeit erforderlich für ...**

- eine Qualitätssicherung als zwingende Voraussetzung, valide und nachvollziehbare Ergebnisse zu erlangen
- die notwendige Entwicklung wissenschaftsimmanenter Gütekriterien

Nähere Ausführungen zum digitalen Wandel finden sich in DFG (2020): Digitaler Wandel in den Wissenschaften. Impulspapier. – Das Papier wird voraussichtlich noch im Lauf dieser Woche zugänglich werden unter https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/digitaler_wandel/index.html

3. Förderregularien – Anforderungen und Normen

Leitlinien zum Umgang mit Forschungsdaten (2015)

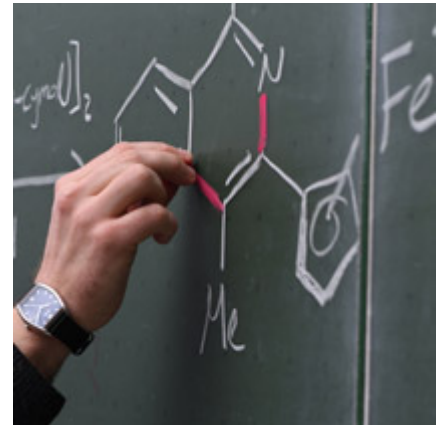
- Beitrag zu Qualitätssicherung und Nachvollziehbarkeit
- Möglichkeit der Anschlussverwendung – auch durch Dritte

► „Ausführungsbestimmungen für Anträge“

- Könnten „meine“ Daten für andere Forschende relevant sein?
- Welche Daten fallen an? Wie wird deren Qualität gesichert?
- Wo sind diese Daten – ggf. auch für Dritte – verfügbar?

► Unterstützung durch die DFG

- Beantragung projektspezifischer Kosten für die Aufbereitung von Forschungsdaten (insbes. zur Nachnutzung durch Dritte)
- Infrastrukturförderung der DFG
- Information (Website zu Forschungsdaten) und Verzeichnisse



3. Förderregularien – Anforderungen und Normen

Leitlinien zum Umgang mit Forschungsdaten (2015)

► **Appell an wissenschaftliche Fachgesellschaften**

- Reflexion der Praxis und Entwicklung von Regularien
- Würdigen des Engagements im Datenmanagement

► **Fachspezifische Empfehlungen** (laufend ergänzt)

- Initiiert durch Fachkollegien (u.a. Alte Kulturen, Erziehungswissenschaften, Linguistik, MatWerk)
- Initiiert durch Fachverbände (u.a. DGP, DGSKA, RatSWD)

► **DFG-interne AG Forschungsdaten** (seit Juni 2020)

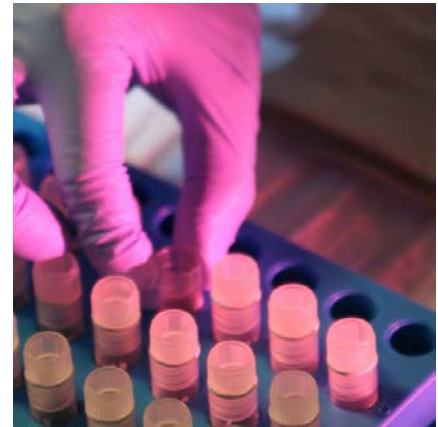
- Anknüpfen an Ergebnisse des Projekts zum digitalen Wandel
- Einführen einer Checkliste zum Umgang mit Daten in Anträgen und Berichten (wird aktuell diskutiert)



3. Förderregularien – Anforderungen und Normen

Kodex zur guten wissenschaftlichen Praxis

- ▶ Ermutigung zum redlichen wissenschaftlichen Handeln
- ▶ **Nr. 7 Phasenübergreifende Qualitätssicherung**
 - Erheben und Analysieren von Forschungsdaten
 - Sorgfältige Dokumentation der Daten
- ▶ **Nr. 10 Nutzungsrechte**
 - Vertragliches Vereinbaren der Verfügbarkeit von Daten
 - Nutzungsberechtigte entscheiden über Zugang für Dritte
- ▶ **Nr. 11 Methoden und Standards**
 - Wesentliche Voraussetzung für die Vergleichbarkeit von Forschungsergebnissen



3. Förderregularien – Anforderungen und Normen

Kodex zur guten wissenschaftlichen Praxis

► Nr. 12 Dokumentation

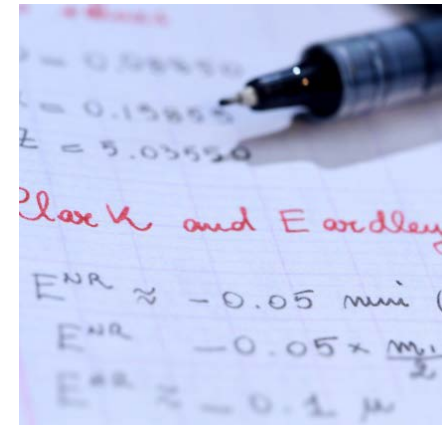
- Information über verwendete Forschungsdaten bieten
- Zugang zu Forschungsdaten gewähren, falls möglich

► Nr. 13 Herstellung von öffentlichem Zugang

- Archivierung nach FAIR-Prinzipien

► Nr. 17 Archivierung

- Speicherung am Ort des Entstehens
- Speicherung in standortübergreifenden Repositorien
- Aufbewahrung – mit fachlichen Unterschieden – für 10 Jahre



3. Förderregularien – Anforderungen und Normen

Kodex zur guten wissenschaftlichen Praxis

► „Dritte Ebene“ als dynamisches Webportal

- Fachspezifische Ausführungen
- Fallbeispiele
- FAQs

► Workshops mit Forschenden

- Fächerübergreifend und dezidiert fachbezogen
- Geistes- und Sozial-, Lebens-, Natur- und Ingenieurwissenschaften

► Veröffentlichung geplant gegen Ende 2020

- Pflege, Aktualisierung und Ergänzung von Inhalten



4. Förderprogramme – Daten in der Forschungspraxis

Datenmanagement in ausgewählten Bereichen

- ▶ Antragstellung für die Sachbeihilfe
- ▶ Förderung der wissenschaftlichen Infrastruktur
- ▶ Exkurs mit Beispielen zur Gestaltung des digitalen Wandels über die DFG-Forschungsförderung
- ▶ NFDI



4. Förderprogramme – Daten in der Forschungspraxis

Leitfaden für die Antragstellung

► „Umgang mit im Projekt erzielten Forschungsdaten“

- Art, Umfang, Dokumentation und Archivierung erläutern
- Eventuelle Nachnutzung durch Dritte ausführen
- Fach-übliche Standards und Repositorien nutzen
- Mittel für die Aufbereitung der Daten beantragen



► Rückbindung an die „Leitlinien“

- Unterstützung, Beratung, Förderung
 - [Re3data](#) (Registry of Research Data Repositories)
 - [RDMO](#) (Research Data Management Organisier)

4. Förderprogramme – Daten in der Forschungspraxis

Grundgedanken zur DFG-finanzierten Datenaufbereitung

► Trennung von Grundaufgaben und zusätzlichen Aufwänden

- Von wissenschaftlichen Einrichtungen wird erwartet, dass sie für ihre Forschenden eine grundlegende Architektur zur sicheren Archivierung von Daten bereitstellen.
- Es wird nicht erwartet, dass jede Einrichtung sämtliche technisch-organisatorische Voraussetzungen erfüllen kann, um Daten interoperabel / nachnutzbar zu machen.
- Diese Aufwände können oft nur von standortübergreifenden Infrastrukturen gestemmt werden.

► Voraussetzungen der Kostenübernahme für Inanspruchnahme standort-übergreifender Infrastrukturen

- Die Infrastrukturen haben einen Kosten-Leistungs-Katalog, der die Zusammensetzung der Kosten und die Abgrenzung von Zusatzaufwänden und Grundbetrieb erläutert.
- Durchschnittswerte für die Aufwände sind aktuell nicht zu benennen, da Aufwände für die Aufbereitung unterschiedlicher Datensätze verschieden sind und es vielfältige Kostenmodelle (z.B. Nutzungsgebühren, Mitgliedsbeiträge) gibt.

4. Förderprogramme – Daten in der Forschungspraxis

Infrastrukturförderung

► Informationsinfrastrukturen für Forschungsdaten

- Daten nachhaltig sichern und nachnutzbar machen
- Förderung nach ggf. unterschiedlichem Entwicklungsstand der Communities akzentuieren
 - Erarbeiten und Umsetzen von Standards für FDM unterstützen
 - Aus- und Aufbau der Daten-Infrastruktur (z.B. Repositorien) fördern
- Zusammenarbeit von Forschenden mit Infrastruktur-einrichtungen gezielt stärken



► **Ausbau auf Grundlage eines DFG-Positionspapiers**

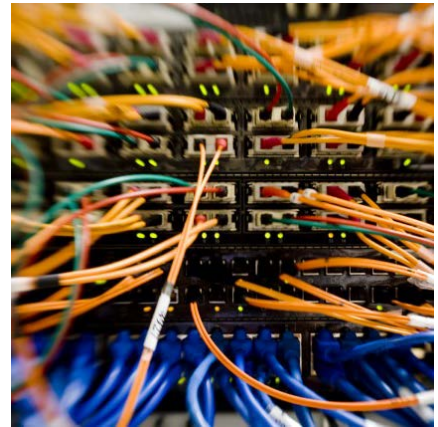
- Policies und Regelwerke weiter entwickeln
- Kompetenz im FDM systematisch stärken
- Bestehende Infrastrukturen anschlussfähig vernetzen

4. Förderprogramme – Daten in der Forschungspraxis

Infrastrukturförderung durch e-Research-Technologien

► Verbesserung von Forschung, die auf digitale Unterstützungsprozesse angewiesen ist

- Infrastrukturen für die Beschaffung, Zugänglich- und Nutzbarmachung, Bearbeitung, Auswertung und Sicherung wissenschaftlich relevanter Informationen
- Werkzeuge, Verfahren, Organisation, Finanzierungsmodelle
 - Infrastrukturen für Forschungssoftware, um an/mit Daten zu arbeiten



► Antragstellung nach Entwicklungsphasen

- Forschung und Entwicklung (Thesen prüfen)
- Konsolidieren (Betriebsstrukturen aufbauen)
- Optimieren (Nachhaltigkeit sichern; Komplexität reduzieren)

4. Förderprogramme – Daten in der Forschungspraxis

Der digitale Wandel im Spiegel von Ausschreibungen

► **Forschungssoftware**

- Nutzbarmachung und [Nachhaltigkeit](#) (2016)
- [Qualitätssicherung](#) von Forschungssoftware (2019)
 - Generierung, Verarbeitung, Analyse, Visualisierung von FD

► **Hochdurchsatzsequenzierung (2017/2018)**

- Untersuchung von Variationen in DNA und RNA
- Förderung von NGS-Kompetenzzentren in Deutschland
- Förderung von Sequenzierkosten

► [Förderinitiative](#) **Künstliche Intelligenz (2020)**

- Verzahnung von KI-Methodenforschung mit Fachforschung
- Stärken von Karrierewegen und wettbewerblicher Position



4. Förderprogramme – Daten in der Forschungspraxis

Bund-Länder-Programm zur [NFDI](#)

► Ziel, Vorgehen, Verantwortlichkeiten

- Daten aus Wissenschaft und Forschung systematisch erschließen, sichern und zugänglich machen
- Bis zu 30 eigeninitiativ agierende Konsortien über drei Auswahlrunden fördern
- Als Bund und Länder pro Jahr bis zu 85 Mio. € für die Finanzierung der Konsortien bereitstellen
- Als DFG die Begutachtung und Bewertung von Anträgen durchführen und die Mittelbewirtschaftung administrieren



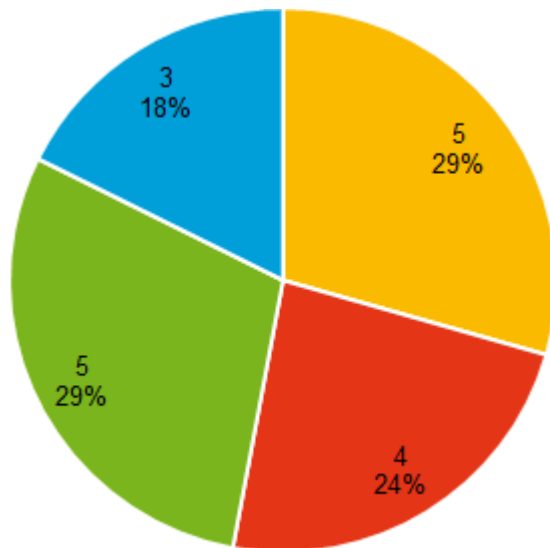
► Aktueller Sachstand

- Förderung von 9 Konsortien aus der ersten Runde
- Begutachtung von 17 Anträgen aus der zweiten Runde

4. Förderprogramme – Daten in der Forschungspraxis

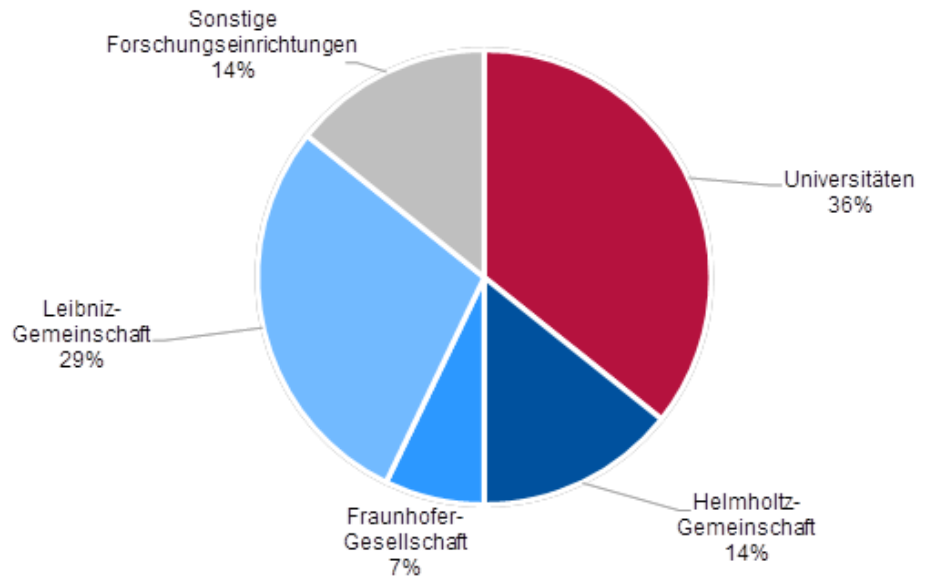
Statische Auswertungen zur zweiten NFDI-Antragsrunde

- Anzahl der 17 Anträge nach primärem Wissenschaftsbereich



- Geistes- und Sozialwissenschaften
- Lebenswissenschaften
- Naturwissenschaften
- Ingenieurwissenschaften

- Antragstellende Einrichtung nach Institutionentyp



5. Ausblick

Internationales und politisches Umfeld

► Internationale Aktivitäten mit DFG-Beteiligung

- **Science Europe / [Priority Research Data](#)**

- Vorschlag disziplin-spezifischer Datenmanagementpläne
- Überlegungen zum internationalen Abgleich von Daten-Policies

- **Knowledge Exchange / Open Scholarship**

- [Task Force](#) „Publishing Reproducible Research Output“
- [Ausschreibung](#) zu „Praxis und Hürden der Reproduzierbarkeit“

- **EOSC**

- Stellungnahmen der Allianz der Wissenschaftsorganisationen
- [Interessensbekundung](#) für eine Mitgliedschaft in der EOSC Association



5. Ausblick

Internationales und politisches Umfeld

► Datenstrategie der Bundesregierung

- [Eckpunkte](#) einer Datenstrategie und Online Konsultation
- Schlaglichter aus der DFG-Beteiligung an der Umfrage
 - Standards sollten von den Forschenden entwickelt, nicht von der Politik gesetzt werden
 - Ethikkommissionen sollten mit Datenschutzkompetenzen ausgestattet werden
- Auswertung und [Ergebnisse](#) der Konsultation



► EU KOM

- [European Data Strategy](#)
- Idee der „Common European Data Spaces“ in unterschiedlichen Sektoren (u.a. Health, Energy, Mobility, Finance, Agriculture)

5. Ausblick

Wie wollen wir – und wie können wir – die wechselseitige Nutzung von Daten aus Forschung, Industrie und Wirtschaft sinnvoll gestalten?



Vielen Dank für Ihre Aufmerksamkeit

Weitere Informationen

- ▶ zur DFG: <http://www.dfg.de>
- ▶ zum Förderatlas: <http://www.dfg.de/foerderatlas>
- ▶ zu allen geförderten Projekten: <http://www.dfg.de/gepris>
- ▶ zu den deutschen Forschungseinrichtungen: <http://gerit.org>