

Forschungsdatenmanagement

Notwendige, aber nicht hinreichende Voraussetzung
für den wissenschaftlichen Erkenntnisgewinn

01. Warum Forschungsdatenmanagement? (I) Wesen der Wissenschaft

Dr. habil. Till Biskup

Physikalische Chemie

Universität Rostock

12.04.2024





- 🔑 Wissenschaft ist ein systematisches Unterfangen, das dem Erkenntnisgewinn dient.
- 🔑 Wissenschaft baut immer auf den Vorarbeiten anderer auf und liefert selbst die Grundlage für die künftigen Forschenden.
- 🔑 Forschende haben eine doppelte Verantwortung: gegenüber der Wissenschaft und der Gesellschaft.
- 🔑 Erkenntnis entsteht aus dem Verständnis von Zusammenhängen. Voraussetzung ist die solide Kenntnis der eigenen Disziplin.
- 🔑 Daten sind Grundlage empirischer Wissenschaft. Die Forschenden selbst sind für den richtigen Umgang mit Daten verantwortlich.

🎵 Leitmotiv

Die Qualität eines Großteils veröffentlichter Forschungsergebnisse wird den Ansprüchen der Wissenschaft nicht gerecht.

- ▶ Nur ein **Bewusstsein für die Ansprüche der Wissenschaft** und frühzeitige Vermittlung notwendiger Kompetenzen zum wissenschaftlichen Vorgehen hat Aussicht auf Erfolg.

🎵 Leitmotiv

Forschungsdatenmanagement ist eine notwendige, aber keine hinreichende Voraussetzung für den wissenschaftlichen Erkenntnisgewinn.

Wesen der Wissenschaft:
systematisches Vorgehen und Nachvollziehbarkeit

Die doppelte Verantwortung der Forschenden

Erkenntnisgewinn durch Verständnis von Zusammenhängen

Daten als Grundlage

“ *Eine jede Lehre, wenn sie ein **System**,
d.i. ein nach Prinzipien geordnetes Ganze der Erkenntnis sein soll,
heißt Wissenschaft*

– Immanuel Kant

Wissenschaft als System

- ▶ Voraussetzung: systematisches, diszipliniertes Vorgehen
- ☛ Systematik und Disziplin Frage der persönlichen Einstellung, weniger der Qualität der zur Verfügung stehenden Hilfsmittel
- ☛ Das beste umfassendste Gesamtsystem zum Umgang mit Forschungsdaten ist nur so gut wie seine Anwender.

“ Eine jede Lehre, wenn sie ein **System**,
d.i. ein nach Prinzipien geordnetes Ganze der Erkenntnis sein soll,
heißt Wissenschaft

– Immanuel Kant

Elemente nach Poser

- ▶ Erkenntnis
 - Aussagen einer Wissenschaft müssen *begründet* sein
- ▶ System
 - Wissenschaft ist Resultat eines methodischen Verfahrens
 - Aussagen hängen untereinander zusammen
- ▶ argumentative Struktur
 - „nach Prinzipien geordnetes Ganzes“

“ *Wissenschaft, so verstanden, ist Ausdruck eines Begründungsanspruchs, eines Rationalitätsanspruchs, der begann, als griechische Denker auf den lebenspraktisch gänzlich überflüssigen Gedanken kamen, den aus Feldvermessungen wohlbekannten Satz des Pythagoras zu **beweisen!** Wissenschaft, wo immer und wie immer sie betrieben wird, zielt also ab auf Aussagensysteme oder Theorien, die **begründet** sind.*

– Hans Poser

- ☛ Wissenschaft ist *per se* nicht anwendungsorientiert.
- ☛ Der Erfolg der wissenschaftlichen Herangehensweise sollte nicht über die Grenzen ihrer Anwendbarkeit hinwegtäuschen.

“ *If I have seen further
it is by standing on y^e shoulders of giants.*

– Sir Isaac Newton

Was ist der Kern von Wissenschaft?

- ▶ Erkenntnisgewinn
 - ▶ Unabhängigkeit vom Betrachter/Experimentator
 - ▶ gegründet auf den Erkenntnissen früherer Generationen
 - ▶ überprüfbar, nachvollziehbar, ggf. reproduzierbar
- 👉 Wissenschaftler tragen Verantwortung gegenüber jenen, die auf den gewonnenen Erkenntnissen aufbauen.

Wesen der Wissenschaft:
systematisches Vorgehen und Nachvollziehbarkeit

Die doppelte Verantwortung der Forschenden

Erkenntnisgewinn durch Verständnis von Zusammenhängen

Daten als Grundlage

Verantwortung gegenüber der Wissenschaft

- ▶ Professionalität: saubere, nachvollziehbare Arbeit nach bestem Wissen und Gewissen
- ▶ Kommunikation der eigenen Ergebnisse und Erkenntnisse zeitnah und für ein Fachpublikum nachvollziehbar
- ▶ Kenntnis des Standes der Forschung auf dem eigenen Gebiet und explizite Bezugnahme darauf in Fachpublikationen

Verantwortung gegenüber der Gesellschaft

- ▶ Verantwortung gegenüber dem Steuerzahler: ökonomischer Umgang mit den verfügbaren Ressourcen
- ▶ Wissenschaftskommunikation: verständliche Vermittlung gesellschaftlich relevanter Erkenntnisse

Empirische Wissenschaften

- ▶ Interpretationen ändern sich, Daten sollten Bestand haben.
- ▶ Voraussetzung: Daten nach bestem Wissen und Gewissen akkurat aufgenommen (und dokumentiert)
- 👉 Nachvollziehbarkeit und saubere Dokumentation sind wesentliche Aspekte von Wissenschaftlichkeit.

These

Datenauswertung, die nicht vollständig dokumentiert und nachvollziehbar ist, ist letztlich unwissenschaftlich.

“ *Scientists regard the free interchange of knowledge and insights as essential, and, in consequence, being non-secretive is one of their rules of professional conduct.*

– Edsger Dijkstra

- ▶ Kommunikation der Ergebnisse/Erkenntnisse, damit andere Forschende darauf aufbauen und ihrerseits neue Erkenntnisse gewinnen können (Newton-Zitat)
- ▶ Ordnen, sortieren, auswählen und geeignet präsentieren der Ergebnisse und der daraus gezogenen Schlussfolgerungen
- ▶ Aufbereitung der Erkenntnisse für eine breitere Allgemeinheit
- ▶ Angemessene Kommunikation in Wort und Schrift ist eine wesentliche Qualifikation für wissenschaftliches Arbeiten.

E. W. Dijkstra, „On the Role of Scientific Thought“
in: Selected Writings on Computing: A Personal Perspective. Springer, New York 1982

Kenntnis des Standes der Forschung auf dem eigenen Gebiet

- ▶ Verantwortung der individuellen Forschenden
(*Nicht* auf die Betreuung verlassen!)
- ▶ Voraussetzung: Vertrautheit mit den grundlegenden Konzepten
- ☛ Literaturrecherche inkl. Vorwärts- und Rückwärtssuche ausgehend von einer Publikation ist eine notwendige Grundqualifikation.

Explizite Bezugnahme darauf in Fachpublikationen

- ▶ großes Bild/Kontext in der Einleitung
(Hier finden sich i.d.R. die meisten Referenzen.)
- ▶ direkte Bezugnahme auf relevante Arbeiten in der Diskussion
- ☛ gilt gleichermaßen für Abschlussarbeiten und Fachpublikationen

“ *Kunst und Wissenschaft, Forschung und Lehre sind frei.
Die Freiheit der Lehre entbindet nicht von der
Treue zur Verfassung.*

– Art. 5, Abs. 3 GG

- ▶ Deutsche Besonderheit, erklärbar aus der Geschichte.
- ▶ Art. 5 GG gilt nur für *unabhängig* Forschende.
- ▶ Freiheit geht immer mit Verantwortung einher.
- ▶ Verantwortungsvoller Umgang mit den bereitgestellten Ressourcen lässt sich nicht überprüfen oder einklagen.

Verantwortung gegenüber dem Steuerzahler

- ▶ Großteil der Wissenschaft ist öffentlich finanziert
 - soll die Unabhängigkeit sichern
 - ist außerhalb der Naturwissenschaften vermutlich noch wichtiger
- ▶ Zumindest in den Naturwissenschaften ist schon das Studium für den Steuerzahler ziemlich teuer...
- ▶ Verantwortung bedeutet nicht, dass alle Forschung absehbar gesellschaftlich relevante Ergebnisse zeitigen muss.

Ökonomischer Umgang mit den verfügbaren Ressourcen

- ▶ Verantwortung, Geräte auch zeitlich möglichst auszulasten
- ▶ Groß-/Größtgeräte: Schichtbetrieb bzw. Automatisierung

Forschende gehören zur intellektuellen Elite

- ▶ Verpflichtung zu Redlichkeit gehört zum Berufsethos.
 - eigene Grenzen kennen und ggf. klar benennen
- ▶ Verstecken in der eigenen Nische ist meist unredlich.

Vermittlung gesellschaftlich relevanter Erkenntnisse

- ▶ Forschende haben eine gewisse Bringschuld.
- ▶ Entscheidungen müssen von der Politik getroffen werden.
- 👉 Wissenschaftskommunikation für eine breite Öffentlichkeit ist eine ganz eigene Qualifikation, die nur wenige besitzen.
- 👉 Der Charakter der Wissenschaft (Vorläufigkeit der Erkenntnisse, Widerstreit der Ideen) muss immer wieder neu vermittelt werden.

Wesen der Wissenschaft:
systematisches Vorgehen und Nachvollziehbarkeit

Die doppelte Verantwortung der Forschenden

Erkenntnisgewinn durch Verständnis von Zusammenhängen

Daten als Grundlage

Erkenntnisgewinn (Kant)

- ▶ setzt eine zur Erkenntnis fähige Person voraus
 - „künstliche Intelligenz“ (ML) ist nicht zur Erkenntnis fähig
- ▶ beruht auf dem Verständnis von Zusammenhängen
- ▶ nur Kommunikation der Erkenntnisse trägt zur Wissenschaft bei

Verständnis von Zusammenhängen (Newton)

- ▶ lässt sich nicht lernen oder vermitteln
- ▶ Wissen und Werkzeuge als Voraussetzung zum Verständnis lassen sich vermitteln und erlernen
- ▶ setzt solide Kenntnis der Grundlagen der eigenen Disziplin voraus
- 👉 praktische Aspekte des wissenschaftlichen Vorgehens

“ *the adjective „scientific“ when used in the expression „scientific thought“ refers more to a way of thinking than to what the thoughts are about. [...]*

Let me try to explain to you what to my taste is the characteristic for all intelligent thinking. It is that one is willing to study in depth an aspect of one's subject matter in isolation for the sake of its own consistency, all the time knowing that one is occupying oneself only with one of the aspects. [...]

It is what I sometimes have called „the separation of concerns“, which, even if not perfectly possible, is yet the only available technique for effectively ordering one's thoughts that I know of. [...] *It is being one- and multiple-track minded simultaneously.*

– Edsger Dijkstra

- ▶ Abstraktion
 - setzt tiefes Verständnis der Zusammenhänge voraus
 - Abstraktion ermöglicht Präzision und Verallgemeinerung
- ▶ systematisches und strukturiertes Vorgehen
 - Fragestellungen sind viel zu komplex, um sie anzugehen
 - iterative Zerlegung in einfachere Teilprobleme
- ▶ Trennung der Belange (*separation of concerns*)
 - Fokussierung auf einzelne Teilprobleme
 - ohne den Gesamtzusammenhang aus den Augen zu verlieren
- ▶ intellektuelle Beherrschbarkeit (*intellectual manageability*)
 - Ziel aller Abstraktion und Trennung der Belange
 - Grund: unsere intellektuellen Fähigkeiten sind begrenzt

“ *Part of becoming a scholar in any field is learning how to evaluate data, make decisions about reliability and validity, and adapt to conditions of the laboratory, field site, or archive. Publications that report findings set them in the context of the domain, grounding them in the expertise of the audience. Information necessary to understand the argument, methods, and conclusions are presented. Details necessary to replicate the study are often omitted because the audience is assumed to be familiar with the methods of the field. Replication and reproducibility, although a common argument for releasing data, are relevant only in selected fields and difficult to accomplish even in those. Determining which scholarly products are worth preserving is the harder problem.*

– Christine L. Borgman

C. L. Borgman: *Big Data, Little Data, No Data: Scholarship in the Networked World.*
MIT Press, Cambridge, MA, 2015. S. xviii

Reproduzierbarkeit und Replizierbarkeit

Nachvollziehbarkeit und Generalisierbarkeit sind wichtigere Konzepte.



		Daten	
		gleich	unterschiedlich
Analyse	gleich	reproduzierbar	replizierbar
	unterschiedlich	robust	verallgemeinerbar

- ☛ Wichtiger als Reproduzierbar- und Replizierbarkeit ist die unabhängige Bestätigung wissenschaftlicher Thesen.

verändert nach: The Turing Way Community

“ *Everything should be made as simple as possible, but not simpler.*
– Albert Einstein

Sparsamkeitsprinzip

Von allen hinreichenden Erklärungen ist die zu bevorzugen,
die mit den wenigsten Annahmen auskommt.

- ▶ keine Aussage über den Wahrheitsgehalt der Erklärung
- ▶ Heuristik für die Wahl wissenschaftlicher Hypothesen

“ *Jede Beobachtung ist eine Beobachtung im Lichte einer Theorie.*

– Hans Poser

“ *Bei der Entwicklung der Wissenschaft spielen Beobachtungen und Experimente nur die Rolle von kritischen Argumenten.*

– Karl Popper

“ *Ein empirisch-wissenschaftliches System muß an der Erfahrung scheitern können.*

– Karl Popper

“ *All models are wrong but some are useful.*

– George E. P. Box

H. Poser: *Wissenschaftstheorie*. Reclam, Stuttgart 2001, S. 126

K. Popper, in: D. Miller (Hg.): *Karl Popper Lesebuch*. Mohr Siebeck, Tübingen 1997, S. 9

K. Popper: *Logik der Forschung*. 11. Aufl., Mohr Siebeck, Tübingen 2005, S. 17

G. E. P. Box, in: Launer und Wilkinson (Hg.): *Robustness in Statistics*, Academic Press, New York 1979

- 1922 David Hilbert: Hilbert-Programm
„Wir müssen wissen, wir werden wissen.“
Versuch, die Widerspruchsfreiheit der Mathematik ausgehend von Axiomen zu beweisen.
- 1931 Kurt Gödel: Unvollständigkeitssätze
Eine geschlossene, widerspruchsfreie und vollständige Gründung der Mathematik auf Axiome ist unmöglich.
- 1935 Karl Popper: Logik der Forschung
„Ein empirisch-wissenschaftliches System muss an der Erfahrung scheitern können.“
- 1936 Alan Turing: Halteproblem
Es gibt Algorithmen, für die nicht beweisbar ist, ob ihre Ausführung je zu einem Ende kommt.

Wesen der Wissenschaft:
systematisches Vorgehen und Nachvollziehbarkeit

Die doppelte Verantwortung der Forschenden

Erkenntnisgewinn durch Verständnis von Zusammenhängen

Daten als Grundlage

- ▶ Daten sind Grundlage der *empirischen* Wissenschaften
 - Einschränkung erlaubt in einer Vorlesung zu FDM
 - keine Daten – keine Notwendigkeit für FDM
- ▶ Daten sind Mittel zum Zweck
 - Ziel der Wissenschaft ist Erkenntnisgewinn
 - Daten sind nicht *per se* langfristig erhaltenswert.
- ▶ Daten allein genügen nicht
 - Entscheidend ist die sorgfältige Planung der Datenerhebung
 - instruktives historisches Beispiel: Gregor Mendel
- ▶ Daten ohne Kontext sind unbrauchbar
 - Kontext: Metadaten (Informationen über die Daten)
 - Verantwortung: Metadaten mit erheben

Blick über den Tellerrand: Archäologie

“ *Der Ausgräber trägt eine große Verantwortung, da jede Bodenuntersuchung die Zerstörung eines niemals wiederherstellbaren Zustandes, d.h. eines geschichtlichen Dokumentes bedeutet. Seine Pflicht besteht also darin, die Ausgrabungen so sorgfältig wie nur irgend möglich durchzuführen, genau zu dokumentieren und die Grabungsergebnisse auch zu veröffentlichen.*

– Tonio Hölscher

☛ Archäologen wissen, dass Daten ohne Kontext wertlos sind – und warum Dokumentation entsprechend wichtig ist.



- 🔑 Wissenschaft ist ein systematisches Unterfangen, das dem Erkenntnisgewinn dient.
- 🔑 Wissenschaft baut immer auf den Vorarbeiten anderer auf und liefert selbst die Grundlage für die künftigen Forschenden.
- 🔑 Forschende haben eine doppelte Verantwortung: gegenüber der Wissenschaft und der Gesellschaft.
- 🔑 Erkenntnis entsteht aus dem Verständnis von Zusammenhängen. Voraussetzung ist die solide Kenntnis der eigenen Disziplin.
- 🔑 Daten sind Grundlage empirischer Wissenschaft. Die Forschenden selbst sind für den richtigen Umgang mit Daten verantwortlich.