



Physikalische Chemie, Universität Rostock

**Vorlesung: Forschungsdatenmanagement
im Sommersemester 2024**

Dr. habil. Till Biskup

— Glossar zu Vorlesung 01: „Warum Forschungsdatenmanagement? (I) Wesen der Wissenschaft “ —

Hinweis: Die nachfolgend genannten Begriffe und Definitionen erheben keinen Anspruch auf formale Korrektheit, sondern dienen lediglich dem besseren Verständnis der in der Vorlesung behandelten Themen und sind im jeweiligen Kontext zu sehen. Mehrfache, voneinander abweichende Definitionen in unterschiedlichen Kontexten sind daher möglich. Fremdsprachige Begriffe werden nach Möglichkeit übersetzt, erscheinen aber ggf. unter ihrem ursprünglichen Namen in der Liste. Verweise auf andere Begriffe innerhalb des Glossars sind durch das vorangestellte Symbol ↑ gekennzeichnet.

Abstraktion Nach Edsger Dijkstra [1] das einzige mentale Werkzeug, das es erlaubt, eine große Vielzahl von Fällen abzudecken. Zweck der Abstraktion ist es nicht, vage zu sein, sondern im Gegenteil ein neues Bedeutungsniveau zu schaffen, das präzise Beschreibungen erlaubt.

Axiom nach Aristoteles ein allgemeiner Satz, der aus sich selbst einsichtig ist und deshalb weder bewiesen werden kann noch muss. Nicht beweisbare Grundannahme (Satz) einer Wissenschaft, auf deren Basis sich aber weitere Sätze ableiten lassen. Vier Anforderungen werden an ein Axiomensystem gestellt: (1) Unabhängigkeit der Axiome voneinander, (2) Widerspruchsfreiheit, (3) Vollständigkeit, (4) möglichst einfach, d.h. aus möglichst wenigen Axiomen bestehend.

Berechenbarkeit Formulierbarkeit einer Berechnungsanweisung (Algorithmus) für eine mathematische Funktion

Erkenntnis Aneignung des Sinngehalts von erlebten bzw. erfahrenen Sachverhalten, Zuständen oder Vorgängen, Ergebnis des Vorgangs des Erkennens. Erkenntnis beinhaltet immer eine auf die Erfahrung gestützte Beurteilung und setzt notwendiger Weise ein Subjekt voraus, das erkennt. Neue Erkenntnisse, die von innerer und äußerer Erfahrung unabhängig

sind, sind immer Ergebnis einer schöpferischen Phantasie. Bei der Erkenntnis stehen sich Subjekt und Objekt als Erkennendes und Erkanntes gegenüber. Die Erkenntnis führt zu einem Abbild des Objekts im Subjekt. Die grundsätzliche Unvollständigkeit dieses Abbilds ist die Triebkraft hinter dem Erkenntnisgewinn und letztlich der ↑Wissenschaft. Vgl. [2]

Falsifizierbarkeit Von Karl Popper in seinem Hauptwerk „Logik der Forschung“ [3] eingeführtes Prinzip, nach dem wissenschaftliche Theorien nie bewiesen werden können, aber so formuliert und gestaltet sein müssen, dass sie grundsätzlich widerlegt werden können. In Poppers Worten: „Ein empirisch-wissenschaftliches System muss an der Erfahrung scheitern können.“ [3, S. 17]

Generalisierbarkeit ↑Verallgemeinerbarkeit

Gödelscher Unvollständigkeitssatz zentraler Satz der modernen Logik. Befasst sich mit der Ableitbarkeit von Aussagen in formalen Systemen und zeigt die Grenzen der formalen Systeme ab einer bestimmten Leistungsfähigkeit auf. Nachweis, dass es in hinreichend starken Systemen, wie der Arithmetik, Aussagen geben muss, die man weder formal beweisen noch widerlegen kann. Wurde 1931 von Kurt Gödel veröffentlicht [4] und bedeutete

das Ende des ↑Hilbert-Programms, das die Widerspruchsfreiheit der Mathematik ausgehend von ↑Axiomen zu beweisen suchte.

Halteproblem Fragestellung aus der theoretischen Informatik. Beschreibt die Frage, ob die Ausführung einer Rechenvorschrift (Algorithmus) zu einem Ende gelangt. Alan Turing konnte beweisen, dass es keinen Algorithmus gibt, der diese Frage für alle möglichen Algorithmen und beliebige Eingaben beantwortet [5]. Das Halteproblem ist somit algorithmisch nicht entscheidbar. Dieses Ergebnis ist von grundlegender Bedeutung für die Berechenbarkeitstheorie (↑Berechenbarkeit). Die generelle Unentscheidbarkeit des Halteproblems ist eng verwandt mit ↑Gödels Unvollständigkeitssatz.

Heuristik die Kunst, mit begrenztem Wissen und wenig Zeit dennoch zu wahrscheinlichen Aussagen oder praktikablen Lösungen zu kommen; analytisches Vorgehen, bei dem mit begrenztem Wissen über ein System mithilfe von mutmaßenden Schlussfolgerungen Aussagen über das System getroffen werden. Die Aussagen können von der optimalen Lösung abweichen. Ist eine optimale Lösung bekannt, lässt sich durch Vergleich die Güte der Heuristik bestimmen.

Hilbert-Programm auch: Formalismus, nach dem Mathematiker David Hilbert benanntes und von ihm ab 1899 propagiertes Vorhaben, die einzelnen Disziplinen der Mathematik als auf ↑Axiomen basierende formale Systeme zu fassen und so ihre Widerspruchsfreiheit zu beweisen. Erwies sich durch ↑Gödels Unvollständigkeitssatz (1931) endgültig als undurchführbar und damit gescheitert.

hinreichend mathematisches Konzept, das eine Bedingung beschreibt, deren Erfüllung ausreicht, um ein gegebenes Ziel zu erreichen. Vgl. ↑notwendig

intellektuelle Beherrschbarkeit *intellectual manageability*, nach Edsger Dijkstra [1] das Hauptziel der Softwaretechnik (*software engineering*) – und letztlich des Projektmanagements. Unterschiedliche Lösungsansätze für

ein Problem sind unterschiedlich gut intellektuell beherrschbar. Entsprechend ist die intellektuelle Beherrschbarkeit das zentrale Kriterium für die Entscheidung, welche Lösung für ein Problem bevorzugt wird.

Komplexität Eigenschaft der Realität, dass selbst sehr wenige einfache Regeln in ihrer Kombination ein nichttriviales Verhalten erzeugen können, die den Menschen für gewöhnlich überfordern. Eine zentrale Strategie zum Umgang mit Komplexität ist ↑Abstraktion und in der Folge ↑Modularisierung. Nach Fred Brooks [6] lassen sich zwei Arten von Komplexität unterscheiden: ↑vermeidbare Komplexität und ↑unvermeidliche Komplexität.

Metadaten Informationen zu den numerischen Daten, notwendige Voraussetzung für eine sinnvolle Verarbeitung der Daten im Kontext eines ↑Systems zur Datenverarbeitung und für ↑nachvollziehbare Wissenschaft.

Modularisierung Aufteilung der Gesamtaufgabe in kleinere Abschnitte. Die Aufteilung wird so lange fortgesetzt, bis die Lösung für den aktuellen Abschnitt unmittelbar in Form von Quellcode offensichtlich ist. Setzt die Definition von ↑Schnittstellen voraus.

monolithisch aus einem Stück bestehend; zusammenhängend und fugenlos

nachvollziehbare Wissenschaft *reproducible science*, seit der Etablierung rechnergestützter Datenauswertung eigentlich nie mehr erreicht, aber für die Wissenschaft konstituierender Aspekt, dass sich Ergebnisse und Auswertungen unabhängig nachvollziehen lassen, weil alle dazu notwendigen Aspekte vollständig und ausreichend beschrieben wurden (↑Nachvollziehbarkeit). Motivation für die Vorlesung, deren Ziel es ist, die Hörer mit Konzepten vertraut zu machen, die letztlich eine ernstzunehmende nachvollziehbare Wissenschaft ermöglichen. Die ↑Nachvollziehbarkeit geht dabei weit über ↑Replizierbarkeit und ↑Reproduzierbarkeit hinaus.

Nachvollziehbarkeit zentraler Aspekt der ↑Wissenschaft und der wissenschaftlichen Methode, die die Intersubjektivität ihrer Aussagen

ermöglicht. Setzt in der Regel eine ↑hinreichende Beschreibung (und Dokumentation) der einzelnen Schritte voraus, die von einem gegebenen Ausgangspunkt zu einem (neuen) Ergebnis oder auch einer Erkenntnis kommt.

notwendig mathematisches Konzept, das eine Bedingung beschreibt, die zwar erfüllt sein muss, um ein bestimmtes Ergebnis zu bekommen, aber für die Erfüllung nicht ausreicht. Vgl. ↑hinreichend

Replizierbarkeit *replicability*, unabhängige Wiederholbarkeit der (Roh-)Datenerhebung, meist in Form von Experimenten und Beobachtungen, entsprechend nicht in jedem Fall durchführbar. Vgl. ↑Reproduzierbarkeit, ↑Robustheit, ↑Verallgemeinerbarkeit.

Reproduzierbarkeit *reproducibility*, vollständige Wiederholbarkeit einer beschriebenen Datenverarbeitung und -Analyse. Ausgangspunkt sind existierende Daten, entsprechend sollte sie in jedem Fall möglich sein. Vgl. ↑Replizierbarkeit.

Robustheit *robustness*, im Kontext der Datenverarbeitung die Tatsache, dass unterschiedliche, unabhängige Analysen derselben Daten zum gleichen Ergebnis führen. Vgl. ↑Reproduzierbarkeit, ↑Replizierbarkeit, ↑Verallgemeinerbarkeit

Schnittstelle in ihrer allgemeinsten Form ein (impliziter) „Vertrag“ zwischen einer Partei, die einen Dienst bereitstellt, und den Nutzern dieses Dienstes. Das Ziel von Schnittstellen ist die Entkopplung, sie sind eine wesentliche Voraussetzung für die ↑Modularisierung. Ein alltägliches Beispiel für eine Schnittstelle wären elektrischer Stecker und Steckdose.

Sparsamkeitsprinzip ↑Heuristik in der wissenschaftlichen Theorienbildung, die diejenige Erklärung bevorzugt, die mit der geringsten Anzahl an Annahmen auskommt, um ein Phänomen zu beschreiben. Verwandte Begriffe sind „Denkökonomie“ und „Ockhams Skalpell“ (nach Wilhelm von Ockham).

System zur Datenverarbeitung hier: Gesamtsystem für wissenschaftliche Datenverarbeitung

von der Datenaufnahme bis zur fertigen Publikation, das alle Aspekte umfasst und das ↑nachvollziehbare Wissenschaft möglich macht und gewährleistet. Definitiv ein größeres Projekt, das nicht nur eine ↑monolithische Anwendung umfasst, sondern viele Aspekte darüber hinaus. Setzt entsprechende ↑Infrastruktur und in der Umsetzung der einzelnen Komponenten sauberen Code und eine solide Softwarearchitektur voraus.

transzendieren die Grenzen eines Bereichs überschreiten

Trennung der Belange *separation of concerns*, grundlegendes Prinzip für ↑Modularisierung, nach Edsger Dijkstra [7] die einzig effektive Möglichkeit, seine Gedanken zu ordnen, indem man sich auf einen Aspekt eines ↑komplexen Problems fokussiert, ohne dabei zu vergessen, dass es lediglich ein Teilaspekt ist.

unvermeidliche Komplexität *essential complexity*, nach Fred Brooks [6] jener Teil der ↑Komplexität eines Systems, der in der Komplexität der Fragestellung begründet ist und der sich nicht verkleinern lässt. Eine gute Systemarchitektur zielt auf die Beherrschung dieser unvermeidlichen Komplexität u.a. durch Einsatz von ↑Abstraktion und ↑Modularisierung. Vgl. ↑vermeidbare Komplexität.

Verallgemeinerbarkeit auch: Generalisierbarkeit, *generalisability*, im Kontext der Datenverarbeitung die Tatsache, dass sowohl unabhängig erhobene Daten als auch voneinander unabhängige Analysemethoden zum gleichen Ergebnis führen. Baustein zur unabhängigen Bestätigung wissenschaftlicher Hypothesen. Vgl. ↑Reproduzierbarkeit, ↑Replizierbarkeit, ↑Robustheit

vermeidbare Komplexität *accidental complexity*, nach Fred Brooks [6] jener Teil der ↑Komplexität eines Systems, der *nicht* in der Komplexität der Fragestellung begründet ist und der sich durch geschickten Einsatz von (etablierten) Strategien beheben lässt. Ein wesentlicher Baustein zur Verringerung dieser vermeidbaren Komplexität ist die Verwendung

guter ↑Abstraktionen. Vgl. ↑unvermeidbare Komplexität.

Wissenschaft Auf den Erkenntnisgewinn ausgerichteter, systematisches menschliches Unterfangen, das in der Regel eine Reihe von Kri-

terien erfüllt bzw. erfüllen sollte: Unabhängigkeit vom Beobachtenden bzw. Durchführenden, gegründet auf den Erkenntnissen früherer Generationen, sowie überprüfbar, nachvollziehbar und ggf. reproduzierbar. Für Einführungen vgl. u.a. [8, 9].

Literatur

- [1] Edsger W. Dijkstra. The humble programmer. *Communications of the ACM* 15 (1972), S. 859–865.
- [2] Heinrich Schmidt. *Philosophisches Wörterbuch*. 22. Aufl. Neu bearbeitet von Georgi Schischkoff. Stuttgart: Kröner, 1991.
- [3] Karl Popper. *Logik der Forschung*. 11. Auflage. Tübingen: Mohr Siebeck, 2005.
- [4] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik* 38 (1931), S. 173–198.
- [5] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42 (1936), S. 230–265.
- [6] Frederick P. Brooks. *The Mythical Man Month*. Anniversary edition with four new chapters. Boston: Addison Wesley Longman, 1995.
- [7] Edsger W. Dijkstra. „On the Role of Scientific Thought (EWD447)“. In: *Selected Writings on Computing: A Personal Perspective*. New York: Springer-Verlag, 1982, S. 60–66.
- [8] Alan F. Chalmers. *What is this thing called Science?* Third edition. Berkshire, UK: Open University Press, 1999.
- [9] Hans Poser. *Wissenschaftstheorie*. Stuttgart: Reclam, 2001.