

Unverkäufliche Leseprobe



Klaus Mainzer
Die Berechnung der Welt
Von der Weltformel zu Big Data

352 Seiten mit 52 teils farbigen Abbildungen.
Gebunden
ISBN: 978-3-406-66130-3

Weitere Informationen finden Sie hier:
<http://www.chbeck.de/13124349>

Einführung

Epidemien, Marktprodukte und Verbrechen lassen sich, so melden Zeitungen und Medien, immer besser vorausberechnen. Grund sind exponentiell steigende Datenmengen, die immer preiswerter gespeichert und mit gigantisch wachsender Rechenpower immer schneller verarbeitet werden können. Es sind nicht nur die Daten, die viele Millionen von Nutzern in Informations- und Kommunikationsnetzen hinterlassen. Bücher, Musik, Fotografien und Videos gehören ebenso dazu wie Telefonanrufe oder Navigationssysteme unserer Fahrzeuge. Im Internet der Dinge beobachtet die Welt sich selber und produziert über Sensoren Milliarden von Dateneinheiten. Informationskonzerne wie z. B. Google und Facebook vermessen damit die Welt, berechnen Persönlichkeitsprofile und bestimmen unsere Zukunft.

Einige dieser Strategen verkündigen bereits das «Ende der Theorie» – ein radikaler und neuer Paradigmenwechsel, so glaubt man, der die Ursachen und Wirkungen von Krankheiten, Märkten und Verbrechen nicht mehr verstehen muss, sondern durch blitzschnelles Durchforsten von riesigen Datenmengen Muster und Korrelationen erkennt, die Voraussagen in Echtzeit erlauben. Gemeint ist dabei keine wahrscheinlichkeitstheoretisch begründete Hochrechnung aus repräsentativen Stichproben. Gemeint ist die Berechnung von Korrelationen aus nahezu allen Daten eines gesamten Datensatzes. Tatsächlich waren es nur Korrelationen über Anfragen und Kaufverhalten im Internet, die Google 2009 den Ausbruch einer Epidemie

voraussagen ließen, ohne langwierige Datenerhebungen von z. B. Gesundheitsämtern und repräsentative Stichproben abwarten oder sogar medizinisch begründete Modelle des Krankheitsverlaufs kennen zu müssen.

Ebenso lassen sich Markttrends und Profile von Produkten aus scheinbar zufälligen und nicht zusammenhängenden Daten über Personen, ihre Themen und Präferenzen schneller und verlässlicher erschließen als über gezielte Befragungen. Bemerkenswerte Erfolge gelangen in der Prävention von Verbrechen, indem aufgrund von automatischen Datenanalysen die Wahrscheinlichkeit von Diebstahl, Einbrüchen und Tötungsdelikten in bestimmten Straßen berechnet wurde und präventiv Polizei vor Ort die Straftaten verhinderte: Das kommt dem Science-Fiction-Film «Minority Report» erstaunlich nahe, in dem in einer total vernetzten Welt eine Art von Gedankenpolizei Strafdaten vollständig ausschalten wollte.

Technisch bezeichnet man mit dem Schlagwort «Big Data» gigantische Datenmengen, die in konventionellen Datenbanken nicht mehr bearbeitet werden können. Dazu wurden neue Softwarewerkzeuge entwickelt, die nicht mehr wie klassische Computerprogramme mit einem Rechnerprozessor auskommen. Vielmehr werden in Parallelrechnung Hunderte oder Tausende von Prozessoren in Superrechnern gleichzeitig eingesetzt, um so eine Reduktion der Komplexität in der Bearbeitung zu erreichen. In diesem Sinn lassen sich mit «Big Data» konkrete Geschäftsmodelle für Firmen aller Art entwickeln, die möglichst adaptiv, trendsicher und schnell auf Märkte reagieren sollen. Die deutsche Industrie sieht daher durchaus das Potential solcher Art von Big Data-Technologie, gibt sich aber bescheiden und bedeckt gegenüber den weitreichenden Visionen von jenseits des Atlantiks.

Es wäre allerdings leichtfertig und fahrlässig, den Big Data-Hype als typisch amerikanische Marketingstrategie herunterzuspielen, um sich auf derzeit vorhandene Software-Technologien zu beschränken. Tatsächlich wird hier ein Trend sichtbar, der bereits die Dynamik menschlicher Zivilisation maßgeblich bestimmt und auch die Wissenschaften erfasst hat: Was wäre, wenn in Zukunft tatsächlich neue Erkenntnis und die Lösung unserer Probleme nur von der schieren Steigerung von Datenmenge, Datenanalyse und Rechenpower ab-

hängen? Ist die Suche nach Erklärungen, Ursachen und kausalen Zusammenhängen, Gesetzen und Theorien angesichts der steigenden Komplexität der Probleme nicht völlig überholt? Können wir uns angesichts des Tempos zivilisatorischer Entwicklung und der Notwendigkeit schneller Entscheidungen überhaupt noch solche zeitraubende Grundlagenforschung leisten? Sollten wir nicht die «Warum»-Frage vergessen und uns auf das «Was» der Daten beschränken?

Historisch steht die «Warum»-Frage am Anfang menschlichen Denkens in Wissenschaft und Philosophie. Warum bewegen sich Sterne und Planeten in regelmäßigen Bahnen? Ist die Vielfalt der Stoffe aus einfachen Grundbausteinen aufgebaut? In griechischer Tradition entstand eine faszinierende Idee, die den weiteren Entwicklungsgang von Forschung grundlegend beeinflusste. Der scheinbar chaotischen Vielfalt der Sinneseindrücke liegen einfache Gesetze der Symmetrie, Regelmäßigkeit und Harmonie zugrunde, die mathematisch beschreibbar sind. Diese Entwicklung begann mit den regulären Körpern Platons und führte zur Suche nach Symmetrien in den Grundgesetzen der modernen Elementarteilchenphysik. Die Suche nach der Weltformel ist der Versuch, die Komplexität der Welt durch einfache Symmetriegesetze zu bewältigen (*Kapitel 1*). Das ist der Trend einer theoriegeleiteten («hypothesen-driven») Forschung. Dahinter steht die Überzeugung: Erst wenn wir eine gute Theorie haben, können wir wissen, wonach wir suchen, um die Vielfalt der Welt zu verstehen und zu bewältigen.

Aber auch die datengetriebene («data-driven») Forschungsperspektive ist keineswegs neu, wie leider historisch wenig gebildete Trendmacher von Big Data glauben machen wollen. Vielmehr ist dieser Trend tief in den Anfängen der Wissenschaftsentwicklung verwurzelt. In einer technisch-wissenschaftlich geprägten Zivilisation bekommen die dahinterstehenden wissenschaftlichen und philosophischen Traditionen eine dramatische Bedeutung. Die Forderung, auf Fakten und Messdaten zu setzen und Hypothesen zu misstrauen, findet sich von Francis Bacon bis Isaac Newton. Der schottische Aufklärungsphilosoph David Hume kritisiert kausale Verknüpfungen von Ereignissen als Hirngespinnste und führt sie auf Korrelationen von Sinneseindrücken zurück. Mit Auguste Comtes Positivismus zieht der Glaube an Fakten und Daten auch in die Sozialwissenschaften ein.

Daten werden Zahlen zugeordnet und damit berechenbar. Gesetze werden zu Rechenregeln, um mathematische Gleichungen zu lösen. Ende des 18. Jahrhunderts ist für den Mathematiker und Astronomen Pierre Simon Laplace die Welt durch Anfangsdaten und Bewegungsgleichungen vollständig bestimmt. Daher kommt es nur auf die Berechnung von Gleichungslösungen an, um zu präzisen Voraussagen zu gelangen. Wenn also, so argumentiert Laplace, einer «Intelligenz» alle diese Daten und Gleichungen gegeben wären, müsste für sie die Welt total berechenbar sein. Diese von Laplace unterstellte «Intelligenz» geht als Laplacescher Geist in die Geschichte ein (*Kapitel 2*). Naheliegend ist es heute, sich darunter einen Superrechner vorzustellen. Berechnungen sind nach Laplace keineswegs auf deterministische Abläufe beschränkt. Die Wahrscheinlichkeit von abhängigen und unabhängigen Ereignissen und ihre Erwartungswerte sind Thema seines zweibändigen Werks über Wahrscheinlichkeitsrechnung von 1812.

Sind Gesetze aber tatsächlich überflüssig, ein Relikt aus einer Zeit, als Naturgesetze noch wie bei Galilei und Newton als «Gedanken Gottes» in der Sprache der Mathematik aufgefasst wurden? Von Nietzsches «Gott ist tot» zum «Tod der Gesetze» als unumkehrbarer Trend der modernen Welt? Massen von Daten und Zahlen alleine sind für uns aber ebenso sinnlos wie die Milliarden von Sinneseindrücken, die unsere Sinnesorgane tagtäglich bombardieren. Seit frühesten Jugend haben wir gelernt, uns an Mustern und Regelmäßigkeiten dieser Daten zu orientieren. Stellen wir uns ein Gerät vor, das eine Folge von Werten aus den Ziffern 0 und 1 (Bits) generiert. In der Bitfolge 010101010101010101 erkennen wir die periodische Abfolge des Paares 01. Es ist daher kürzer, die Regel «10 mal 01» zu notieren und mit dieser Regel die nächsten Schritte dieser Abfolge voraussagen. In der Datenfolge 01100010111001011110 ist kein Muster zu erkennen und damit auch keine Möglichkeit der Voraussage. Um diese Abfolge zu beschreiben, gibt es keine kürzere Darstellung als die Folge selber. Regeln und Gesetze sind also zunächst Datenkompressionen, die ein Muster zum Ausdruck bringen.

Unser Gehirn wurde während seiner Evolution auf Datenkompression und Reduktion von Komplexität trainiert. Blitzschnelle Entscheidungen hängen von dieser Fähigkeit ab. Das traf nicht nur

im Überlebenskampf während der Steinzeit zu. Auch im heutigen Geschäftsleben und in der Politik stehen wir unter dem Druck häufig reflexartiger Entscheidungen. Superrechner und Big Data scheinen diesen Trend zur schnellen Entscheidung zu bedienen. Gelegentlich bilden wir uns aber auch Zusammenhänge und Muster ein, denen nur scheinbare Korrelationen von Ereignissen zugrunde liegen. Wetterregeln unserer Vorfahren waren häufig nicht besser begründet als das Zockerverhalten von Börsenspekulanten. Aber die Muster und Korrelationen von Big Data bleiben zufällig, wenn wir die zugrunde liegenden Zusammenhänge nicht verstehen. Natürlich greift ein Krebspatient in seiner äußersten Not nach dem Strohalm einer statistischen Korrelation zwischen einer unverstandenen Medikamentenwirkung und einer beobachteten Lebensverlängerung. Die langjährige Forschung nach den biochemischen Gesetzen, die dieser Korrelation zugrunde liegen oder auch nicht, mag für ihn persönlich zu spät kommen. Endgültig bieten aber nur diese Gesetze eine verlässliche und reproduzierbare Therapie.

Dieses Buch ist ein Plädoyer für Theorie und Gesetze, gegen Big Data-Verheißungen und für die ungebrochene Bedeutung von Grundlagenforschung und philosophischer Reflexion. Das wird zunächst am Beispiel mathematischen Denkens und algorithmischer Computerverfahren gezeigt, mit denen die technisch-wissenschaftliche Welt erst möglich wurde. Bereits in den historischen Anfängen der Mathematik finden sich die Vertreter von Big Data. Es sind die Babylonier, die für damalige Verhältnisse große Massen von Daten über astronomische Beobachtungen, Ernteergebnisse, Handel, Gewerbe und Verwaltungsabläufe auf unzähligen Tontafeln in Keilschrift festhalten. Aus den Regelmäßigkeiten in den astronomischen Daten werden erstaunliche Voraussagen über Planetenkonstellationen abgeleitet, ohne sie allerdings erklären zu können und zu wollen.

Ganz anders die griechische Mathematik: Im Zentrum ihres Interesses stehen geometrische Modelle und Beweise. In der Astronomie werden Planeten mit idealen Kugelschalen verbunden, die sich zentrisch um die Erde, aber in anderen Modellen auch bereits um die Sonne drehen. Beobachtete Schleifenbewegungen von Planeten werden mathematisch exakt auf kombinierte Umläufe von Kugelschalen

zurückgeführt. Das sind noch keine Naturgesetze im Sinn neuzeitlicher Physik, aber die Reduktion komplexer Erscheinungen der Natur auf einfache mathematische Gesetze. In Euklids «Elementen» entwirft die griechische Mathematik erstmals das Idealbild einer mathematischen Theorie, die aus wenigen als wahr vorausgesetzten Grundgesetzen (Axiomen und Postulaten) logisch exakt weitere Sätze (Theoreme) ableitet, um so geometrische Konstruktionen mit Zirkel und Lineal zu rechtfertigen. Der Satz des Pythagoras gilt nicht, weil die Menschheit bei der Feldvermessung immer wieder diese Datenverhältnisse angetroffen hat, sondern weil sie logisch aus den Euklidischen Axiomen folgen. Kurz: Es gibt Wahrheiten, die nicht von Big Data abhängen, mit denen wir aber Big Data beherrschen können: die Gesetze der Mathematik!

Big Data wird mit Zahlen dargestellt. Big Data mag heute und zukünftig unvorstellbar große Zahlen verarbeiten. Diese Anwendungen sind aber nur ein Bruchteil der unendlichen mathematischen Welt der Zahlen, die in der Zahlentheorie axiomatisch beschrieben wird. Hier sehen wir par excellence, was Gesetzeserkenntnis durch Datenkompression und Reduktion von Komplexität bedeuten. Schon das einfachste Gesetz der Arithmetik, das dem Zählen zugrunde liegt, kondensiert die unendliche Folge der natürlichen Zahlen 1, 2, 3, ... in eine einfache Regel aus endlich vielen Symbolen: (a) Beginne mit 1. (b) Wenn bis zur Zahl n gezählt ist, addiere 1 zu n hinzu. Durch sukzessive Anwendung dieser Regel entsteht die unendliche Folge 1, 1+1, 1+1+1, ..., die den bekannten Symbolen der natürlichen Zahlen 1, 2, 3, ... entspricht.

So einfach die Zahlenwelt in ihren Anfängen zu sein scheint, so dass sie jedem Schulkind zugänglich ist, so ist man schon nach wenigen Schritten in den tiefsten Geheimnissen der Mathematik wie z. B. der Fermatschen Vermutung, deren Beweis über Jahrhunderte verschlossen blieb und die erst durch höchst abstrakte Theorien und Beweise entschieden werden konnte. Anfang des 20. Jahrhunderts hatte der große Mathematiker David Hilbert eine Liste von offenen Beweisen und Problemen aufgestellt, die noch zu lösen waren. Wenn alle mathematischen Theorien nach dem Vorbild der Arithmetik in axiomatischen Regeln mit endlich vielen Symbolen formalisiert werden, dann müssten, so war seine feste Überzeugung, alle mathematischen

Wahrheiten formal bewiesen werden können: «Wir müssen wissen – wir werden wissen,» ließ Hilbert sich auf seinen Grabstein meißeln (*Kapitel 3*).

Hilberts Programm schien den Laplaceschen Geist in einer entscheidbaren und berechenbaren Welt der Mathematik zu vollenden. Bald schon bewies aber Kurt Gödel, dass bereits die Arithmetik keine vollständige Formalisierung zulässt. Er konnte in der Formalisierung der Arithmetik eine nicht entscheidbare Aussage nachweisen. Selbst wenn das Axiomensystem um diese nicht entscheidbare Aussage erweitert würde, gäbe es wieder eine nicht entscheidbare Aussage im erweiterten System. Es bleiben also prinzipiell immer blinde Flecken, die von den jeweiligen Axiomatisierungen nicht erfasst werden. Auch die Widerspruchsfreiheit der Arithmetik lässt sich mit den Methoden der Arithmetik selber nicht beweisen (*Kapitel 4*).

Philosophisch wurde Gödels Ergebnis häufig als Scheitern des Hilbertschen Programms gedeutet. Tatsächlich handelt es sich nur um eine Relativierung. Ein Schüler Hilberts, Gerhard Gentzen, bewies die Widerspruchsfreiheit der Zahlentheorie, wenn stärkere Methoden vorausgesetzt werden, als sie in der Arithmetik zur Verfügung stehen. Es ist also wie bei dem Baron Münchhausen, der sich nicht selber aus dem Sumpf (hier der Widersprüche) herausziehen kann, aber mit einem Werkzeug von außen. Es geht also weiterhin um logisch exakte Beweise von Gesetzen und nicht um quasi-empirisches Probieren. Gerhard Gentzen leitete die Entwicklung einer neuen mathematischen Theorie der Beweise ein. Diese mathematische Beweistheorie verschafft uns erst neue Erkenntnis über die unterschiedlichen Stärken von Beweismethoden. Gödel war also nur der Anfang. Ferner sind Gentzens Methoden von Bedeutung, um konsistentes und verlässliches Arbeiten von komplexen Computerprogrammen in der praktischen Anwendung zu garantieren (*Kapitel 5*).

Big Data ist ohne immer stärker werdende Superrechner nicht möglich. Aber auch der Computertechnik liegt eine mathematische Theorie zugrunde, die erst die Rechenleistungen der Vergangenheit und Zukunft möglich macht. Es war der britische Logiker und Mathematiker Alan Turing, der als Zeitgenosse von Gödel und Gentzen maßgeblich die Theorie der Berechenbarkeit begründete. Mit der

logisch-mathematischen Definition der nach ihm benannten Turingmaschine schuf er den Prototyp, auf den wir mathematische Theoreme über Berechenbarkeit unabhängig von technischen Standards ihrer Realisation beziehen können. In Kapitel 3 werden wir sehen, wie er damit eine prinzipiell nicht berechenbare Zahl definieren konnte. Im nächsten Schritt kann bewiesen werden, dass es prinzipiell keine Rechenmaschine geben kann, die für jedes Rechenverfahren (Algorithmus) und beliebige Eingaben entscheiden kann, ob die Maschine nach endlich vielen Schritten stoppt und ein Ergebnis liefert oder unbegrenzt weiter suchen muss. Es lässt sich beweisen, dass Gödels Unvollständigkeit formaler Theorien daraus unmittelbar folgt.

Analog zu Gödels Satz ist Turings Ergebnis keine absolute Grenze der Berechenbarkeit, sondern eröffnet eine mathematische Theorie, in der sich unterschiedliche Grade und Stärken von Berechenbarkeit unterscheiden lassen. Berechenbarkeit und Beweisbarkeit werden also genauer unterschieden und besser begründet. Das hat unmittelbare Folgen auch für die Praxis: Wenn wir von einem komplexen Computerprogramm einer industriellen Produktionsstraße vorher mathematisch beweisen können, dass es widerspruchsfrei und vollständig arbeitet, können später auftretende Unfälle und teure Produktionsfehler vermieden werden. Probieren und mehr oder weniger zufällige Daten helfen nur wenig. Umso komplexer die moderne Lebenswelt wird und umso abhängiger wir von Softwareprogrammen werden, umso größer sind die Herausforderungen an ein sicheres Software-Engineering. Verlässlichkeit und Nachprüfbarkeit kommt an Beweisen und Gesetzen nicht vorbei.

Andererseits machen die Sirenenklänge von schnellen Erfolgen mit Big Data und Superrechnern auch vor der Mathematik nicht halt. Der amerikanische Logiker Gregory Chaitin propagiert ein quasi-empirisches Vorgehen in der Mathematik. Gödels und Turings Ergebnisse werden als Grenzen der klassischen Auffassung von Beweisbarkeit und Berechenbarkeit aufgefasst. Weitere Axiome sind bestenfalls Hypothesen wie in den Naturwissenschaften, die als plausibel gelten, sich bisher bewährt haben und deren Annahmen für neue Problemlösungen dienen. Das ist der Maßstab ihrer Akzeptanz. Ob sie beweisbar sind oder sich widerspruchsfrei in Theorien einfügen, spielt keine Rolle mehr. Ähnlich hatte schon der Popper-Schüler

Imre Lakatos argumentiert, der mathematisches Denken als quasi-empirische Problemlösungssuche mit Hypothesen statt Axiomen und Beweisen beschrieb. Big Data und Superrechner versprechen eine neue Auflage dieser Problemlösungssuche unter den Bedingungen moderner Technik.

So verkündete der amerikanische Informatiker und Software-Unternehmer Stephen Wolfram 2002 eine «neue Art der Wissenschaft» («A New Kind of Science»), in der Computerexperimente anstelle mathematischer Beweise und Theorien treten werden (*Kapitel 6*). Wolfram hatte umfangreiche Musterentwicklungen von zellulären Automaten in einem bis dahin nicht gekannten Umfang durchgeführt und bemerkenswerte Zusammenhänge zwischen vielfältigen Strukturbildungen beobachtet. Zelluläre Automaten bestehen aus schachbrettartigen Gittern, deren Zellen nach ausgewählten Regeln ihre Zustände (z. B. die Farben Schwarz oder Weiß) wechseln und dabei von der Farbverteilung der jeweiligen Zellumgebung abhängen. Schnelle Computerleistungen erlaubten Musterentwicklungen in vielen nachfolgenden Generationen, die vorher nicht möglich waren. Wie heute bei Big Data konnte man nun feststellen, dass sich bestimmte komplexe Muster aus scheinbar zufälligen Regeln gebildet hatten. Die Frage «Warum» blieb unbeantwortet. Stattdessen wurden, wie heute bei Big Data, Klassifikationen und Korrelationen von beobachteten Gemeinsamkeiten vorgenommen.

Für Wolfram war das ein neues Forschungsparadigma, wie zukünftig auch Mathematik und theoretische Physik sich entwickeln werden: Mit gewaltigen Rechenleistungen wird man probieren und experimentieren, um Problemlösungen zu finden. Theorien, Beweise und Erklärungen werden überflüssig, da sie zu aufwendig seien und bestenfalls im Nachhinein nur das bestätigen, was man sowieso schon gesehen und beobachtet hat. Man sollte die Ressourcen stattdessen lieber nutzen, um weiter Neues zu entdecken und zu erzeugen. Zehn Jahre später entwickelt Wolfram mit seiner Firma die Such- und Wissensmaschine WolframAlpha, mit der er nach dem Vorbild von Big Data gewaltige Datenmengen in Facebook mit Mustern, Clustern und Korrelationen durchforstet. Wieder lautet die Devise: Computerexperiment, Überraschung und Entdeckung, statt Begründung, Erklärung und Beweis!

2011 schrieb ich mit Leon Chua, Informatiker und Elektroingenieur der Universität Berkeley, das Gegenbuch «The Universe as Automaton». Am Beispiel der zellulären Automaten bewiesen wir systematisch, dass die Musterbildungen sich keineswegs zufällig aus den zugrunde gelegten Regeln ergaben. Jeder zelluläre Automat konnte durch ein Gesetz in Form einer Gleichung charakterisiert werden, mit dem präzise Erklärungen und Voraussagen aller Musterbildungen mathematisch gefolgert werden konnten. Einige Annahmen Wolframs über Musterbildungen erwiesen sich sogar als fehlerhaft oder ungenau. Neue Entdeckungen wurden durch Beweise erst möglich. Schließlich konnten wir systematisch Theoreme über Musterbildung aus einer grundlegenden Theorie beweisen, die durch fundamentale Symmetrien ausgezeichnet ist. Wie in der Physik erklären also Gesetze und Theorien die beobachteten Phänomene, die bei zellulären Automaten in den Musterbildungen von Computern auftreten. Wenigstens in der Welt der zellulären Automaten hatten wir die «Weltformel» zur Erklärung aller beobachteten Zusammenhänge gefunden. Big Data alleine reicht nicht. Klasse statt Masse, heißt die Devise. Wenn auch noch nicht die «Weltformel», aber Formeln, Gesetze und Beweise werden durch Big Data nicht überflüssig!

Überall in der Natur finden wir komplexe Muster- und Strukturbildungen, die aus der Wechselwirkung vieler Elemente mit wenigen Regeln entstehen. Zelluläre Automaten sind nur ein einfaches Beispiel für Musterbildungen, bei dem der Computer die Rolle der Natur übernimmt. In der Natur kann es sich um Musterbildungen in chemischen Reaktionen, um die Bildung von Zellstrukturen, von Mustern auf Fell oder Gefieder von Tieren oder Verschaltungsmuster von Neuronen im Gehirn handeln. Man kann die ungeheure Vielfalt dieser Muster wieder nur beobachten und klassifizieren, ohne ihre Ursachen erklären und ihren Entstehungsprozess gesetzmäßig voraussagen zu können.

Das wäre erneut der Ansatz von Big Data, nämlich Sammeln und Generieren von Daten und Berechnung von Korrelationen. Tatsächlich lassen sich wenigstens für Physik, Chemie, Biologie und Gehirnforschung die gesetzmäßigen Grundlagen in einer gemeinsamen mathematischen Theorie angeben. Mit den Gesetzen dieser Theorie in Form von Gleichungen lassen sich nicht nur genaue Er-

klärungen, sondern auch die Prognose von neuen Phänomenen präzise berechnen.

Dies nenne ich das Newtonsche Forschungsparadigma, das die Erfolge der modernen Physik und Technik durch Theorie- und Gesetzesbildung erst möglich machte, auch wenn die abschließende «Weltformel» noch nicht gefunden ist. Newton hatte die axiomatische Methode Euklids aus der Mathematik in die Physik übertragen. Seiner Mechanik liegen fundamentale Bewegungsgesetze als Axiome zugrunde, die er durch die Erfahrung als begründet annimmt. Beobachtete Wirkungen in der Natur sollen durch Kräfte als Ursachen erklärt werden. Bei Newton stand die Gravitation im Zentrum. Die beobachteten Planetengesetze Keplers oder Wurfgesetze Galileis, aber auch die gesammelten Daten von Ebbe und Flut konnten aus den Grundaxiomen der Mechanik und dem Gravitationsgesetz mathematisch gefolgert und als Erklärungen vergangener Ereignisse bzw. Voraussagen zukünftiger Ereignisse verwendet werden.

Die Gesetze von Kepler, Galilei und Newton sind zwar geniale Datenkompressionen und Reduktionen von unbegrenzten Datenmassen in einfachen Gesetzen. Andererseits wurden sie aber nicht durch Big Data möglich. Die historisch nachgewiesenen Versuche von Galilei mit der schiefen Ebene halten sich deutlich in Grenzen. Und Newtons Idee des Schwerkraftgesetzes kam ihm nicht, weil er unentwegt Äpfel von Bäumen fallen ließ. Entscheidend war eine mathematische Theorie mit Gesetzen und Beweisen, mit denen er überschaubar viele Daten seiner Messungen erklären konnte.

Im 18. Jahrhundert wurden Newtons geometrische Gesetze in Gleichungen übertragen. Kraftursachen wurden nun als Lösungen von Bewegungsgleichungen aufgefasst, die durch den neuen Kalkül der Differential- und Integralrechnung berechenbar waren. Die Laplacesche Vision der totalen Berechenbarkeit der Natur ging davon aus. Nach der Mechanik traten weitere physikalische Theorien wie Elektrodynamik und Quantenphysik für die neu entdeckten Kernkräfte und die schwache Wechselwirkung hinzu. Auch diese Theorien sind durch fundamentale Grundgleichungen bestimmt. Die Physik verfolgt das Forschungsprogramm, diese Kräfte unter enormen Energiebedingungen schrittweise zu vereinigen und die so vereinigten Kräfte (z. B. elektromagnetische und schwache Wechsel-

wirkung) durch neue fundamentale Gleichungen zu erklären. Die bisher gefundenen Gesetze einer vereinigten Theorie von elektromagnetischer, schwacher und starker Wechselwirkung unter den Bedingungen der Quantenphysik unterstützt die Hoffnung auf eine finale Theorie der Vereinigung aller Grundkräfte, mit der alle physikalischen Ereignisse erklärbar wären. Das wäre dann die «Weltformel».

Die Verfolgung dieses Forschungsziels ist allerdings nur mit Big Data möglich. Die Teilchenkollisionen im CERN, mit der die ungeheuren Energien von Vereinigungszuständen erreicht werden sollen, produzieren Datenmassen, die nur teilweise durch die größten Superrechner verarbeitet werden können. Auch in der Astrophysik und Kosmologie produzieren unsere Satelliten, Raumsonden und Teleskope Datenmengen, die ohne Superrechner nicht zu erfassen sind. Big Data ist überall, aber auch eine hochentwickelte Mathematik von Gleichungen, in denen Muster und Korrelationen erfasst werden. Sie müssen allerdings auf die fundamentalen Naturkonstanten und Grundgleichungen der Physik zurückführbar sein, um die Ursachen dieser Korrelationen zu verstehen und um sie als zufällige Zusammenhänge auszuschließen.

Die verwendeten Gleichungen sind aber in der Regel nicht mehr «von Hand» lösbar, wie man das von Schulaufgaben gewohnt ist. Aufwendige Computerprogramme berechnen numerische Approximationen mit digitalen Daten. Die Vorgänge der physikalischen Welt werden in dem Umfang berechnet, der in digitalen Modellen erfasst ist. Big Data führt zu einer digitalen Physik der Superrechner, in der Gleichungen in Computerprogramme überführt werden. Physiker wie Richard Feynman und Computerpioniere wie Konrad Zuse haben sich daher die Frage gestellt, ob die Natur nicht selber als ein Computer aufgefasst werden kann, in dem Veränderungen und Wechselwirkungen von Atomen und Molekülen Rechenschritten und Rechenprozessen entsprechen. Diese Idee führte zum Konzept der Quantencomputer mit neuartigen Möglichkeiten, aber auch Grenzen der Berechenbarkeit (*Kapitel 7*).

Grenzen der Berechenbarkeit zeigten sich schon in der Mechanik, als der Mathematiker und Philosoph Henri Poincaré knapp hundert Jahre nach Laplace Instabilität und Chaos in komplexen Rückkopp-

lungsschleifen der Himmelsmechanik entdeckte. Aber auch hier handelt es sich nicht um absolute Grenzen der Berechenbarkeit. Mit der Theorie komplexer dynamischer Systeme eröffnet sich vielmehr eine genaue Analyse und Bestimmung nichtlinearer Dynamik mit ihren überraschenden Nebeneffekten. Genau das benötigen wir aber für Frühwarnsysteme unserer Klimamodelle, Turbulenzen und extremen Ereignisse, die unsere Erde erschüttern (*Kapitel 8*).

In den Lebenswissenschaften und der Gehirnforschung bricht Big Data weiter in die Naturwissenschaften ein. Gensequenzen und Proteinfaltungen produzieren Datenmengen, deren Korrelationen nur durch aufwendige Algorithmen entdeckt werden können. Wie verlässlich sind aber solche Korrelationen, wenn keine gesetzmäßigen Erklärungen vorliegen? In dem Buch «Local Activity Principle» vertrete ich mit Leon Chua die Auffassung, dass solche Gesetze identifiziert werden können und müssen. Tatsächlich gelingt es für bestimmte Klassen der Struktur- und Musterbildung, die kausalen Ursachen gesetzmäßig zu erklären und mathematisch zu berechnen. Wie können scheinbar leblose Zellen «am Rand des Chaos» zum Leben erweckt werden und durch ihre Wechselwirkungen neue Ordnung und Struktur erzeugen?

Diese Vorgänge finden wir auch im Gehirn als einem der komplexesten Systeme der Natur. Muster- und Verschaltungsnetze von Neuronen, die mentalen und kognitiven Zuständen des Wahrnehmens, Denkens, Fühlens und des Bewusstseins zugrunde liegen, lassen sich mit den Hodgkin-Huxley-Gleichungen genau prognostizieren und berechnen. Damit verstehen wir zwar die «Maschinensprache» des Gehirns auf der molekularen und zellulären Ebene. Die Brücke zur Kognition und damit zur Berechnung der Dynamik unserer Gefühle, Gedanken und Vorstellungen bleibt aber nach wie vor eine große Herausforderung. Gilt auch hier Hilberts Devise «Wir müssen wissen – wir werden wissen»? Jedenfalls können wir nur so die Gründe für die Entstehung menschlicher Kreativität, aber auch von mentalen Krankheiten verstehen (*Kapitel 9*).

Big Data eröffnet erstaunliche Möglichkeiten, um aus unüberschaubar vielen Daten und Spuren im Netz persönliche Profile über Vorlieben, Veranlagungen und Neigungen zu erschließen. Personalisierte Ernährung und Gesundheitsfürsorge werden so möglich. Die

Rede ist bereits von einer personalisierten Medizin. Das wäre tatsächlich die passende Antwort auf den hochkomplexen menschlichen Organismus, der keine starren Standards wie ein mechanischer Automatismus kennt. Von der Zytologie der Zelle über die Physiologie der Organe bis zur Immunologie und ihren psychosomatischen Wirkungen können individuell sehr unterschiedliche Krankheitsverläufe entstehen. Kurz: Was der eine wegsteckt, kann für den anderen gefährlich werden. Big Data wird den Trend zur personalisierten Medizin unterstützen. Aber Heilen und Helfen setzt ein gründliches Verständnis von Ursachen voraus.

Der Rohstoff «Daten» wird zunehmend nicht in den vielen einzelnen Computern, Datenbanken und Datenträgern gelagert, die wir im Alltag benutzen. Das wäre so wie früher die Daten und Informationen in den Büchern und Zeitschriften zu Hause, am Arbeitsplatz oder in einigen öffentlichen Bibliotheken. Sie sind in unsichtbaren Informations- und Kommunikationsnetzen global verteilt und werden daher als Daten-«Wolke» (englisch «cloud») bezeichnet. Das Internet gilt als Nervensystem der globalisierten Welt. Die Metapher vom Nervensystem passt insofern gut, da Nerven und Gehirne nur mit einem Organismus funktionieren. In der Tat wächst die digitale Welt mit den sozialen Infrastrukturen zusammen (*Kapitel 10*). Energiesysteme, logistische Versorgungssysteme und mobile Infrastrukturen werden durch «intelligente» Netze gesteuert, die zunehmend autonom und lernfähig Entscheidungen treffen. Die Rede ist von soziotechnischen Systemen und Cyberphysical Systems, in denen Technik und Mensch über Elektronik, Sensorik und Software zusammenwachsen. Das Schlagwort von Industrie 4.0 ist eine Umsetzung des Internets der Dinge am Arbeitsplatz: Nicht mehr der starre Ablauf der Produktion am Fließband, sondern Werkstücke und Maschinen sind es, die sich über Sensoren wahrnehmen, mit Menschen kommunizieren und ihren Arbeitsprozess bedarfsgerecht selber organisieren. Damit zeichnet sich eine neue Industriewelt ab, in der nicht länger nach gleichen Standards für alle produziert wird. Von Automobilen bis zur Kleidung könnte in dieser Infrastruktur nach persönlichen Wünschen und Profilen on-demand hergestellt und geliefert werden. Der maßgeschneiderte Anzug, der individuell zu einem gewünschten Zeitpunkt entsteht, wird für alle preiswert

möglich. Big Data liefert Informationen über Trends dieser soziotechnischen Systeme.

In soziotechnischen Systemen wachsen Infrastrukturen und IT-Welt zusammen. Daher sind Risiken nicht alleine technisch bedingt, sondern ebenso ökonomischer, ökologischer und sozialer Natur. Die mathematische Berechnung von Risiken ist mit der Entstehung der Wahrscheinlichkeitstheorie eng verbunden (*Kapitel 11*).

In der mathematischen Naturwissenschaft werden Gesetze und Modelle angenommen, deren Erklärungen und Voraussagen in Experimenten und Voraussagen überprüfbar sind. In den Sozial- und Wirtschaftswissenschaften liegen zwar auch mathematische Modelle vor. Ihr Voraussagepotential ist aber bekanntermaßen schwach, wie die vergangenen Finanz- und Wirtschaftskrisen zeigen. Von Frühwarnsystemen konnte in diesen Situationen keine Rede sein. Vielmehr führten falsch verstandene Modelle noch zusätzlich tiefer in die Krise. Nassib Taleb, Finanzspekulant, Hobby-Philosoph und Erfolgsschriftsteller, machte aus der Not eine Tugend, wetterte gegen den «Platonismus» mathematischer Modelle, pries den Zufall und vertraute seiner Intuition. Big Data scheint diese Haltung zu unterstützen: Statt Intuition wird nun auf das Erfassen aller erreichbaren Daten gesetzt, die nach hilfreichen Korrelationen zu durchforsten sind, ohne die kausalen Ursachen und Wirkungen zu verstehen. Sind hier grundsätzliche Grenzen der Berechenbarkeit erreicht, oder fehlt einfach noch der «Newton der Wirtschaftswissenschaften»? Wir zeigen, dass man sich bereits in der Finanzwirtschaft nicht auf ein einziges Modell verlassen darf. Vielmehr muss eine Klasse von möglichen Modellen Stresstests unterzogen werden, um ihre Konsequenzen unter extremer Belastung (worst case) zu bestimmen. Dazu wurden robuste Risikomaße entwickelt. Ohne genaue Analyse des «Faktors Mensch» hängen mathematische Modelle aber in der Luft. Verhaltensökonomie und experimentelle Ökonomik liefern neue Daten über kognitive und emotionale Einflüsse auf Entscheidungen unter Unsicherheit, die gleichwohl in mathematischen Modellen berücksichtigt werden können.

Wie berechenbar ist aber die soziale Welt? Voraussagen über soziales Verhalten gelten als praktisch ausgeschlossen, da für einzelne Menschen keine Bewegungs- und Entwicklungsgleichungen wie bei

Molekülen und Zellen bekannt sind. Tatsächlich zeigt Big Data in ersten Ansätzen, dass überraschend genaue Voraussagen möglich sind – nur durch große Datenanalyse und Berechnung von Korrelationen (*Kapitel 12*). So könnten in Zukunft Konflikte in bestimmten Wohnvierteln vorausgesagt und Schadensbegrenzung durch präventive Maßnahmen der Polizei erreicht werden. Damit sind aber die sozialen Ursachen wie z. B. Jugendarbeitslosigkeit oder Ghettobildung von Migranten noch nicht verstanden.

Ähnliches gilt für die Prävention von Terror und Krieg: Was ist die Zukunft von Krieg und Terror in einer Welt von Big Data und soziotechnischen Infrastrukturen? Auch hier könnte Big Data als Frühwarnsystem für Attentate und kriegerische Auseinandersetzungen dienen. Jedenfalls sind verbesserte Vorwarnzeiten denkbar, um sich besser vorbereiten zu können. Damit sind aber wiederum die Konfliktursachen nicht verstanden, um sie nachhaltig zu verändern. Die persönliche Killerdrohne maßgeschneidert für einen Terroristen löst das Problem nicht und wäre auch mit dem deutschen Grundgesetz nicht vereinbar. Intelligente Computerviren, die Infrastrukturen lahmlegen und zerstören, sind der Albtraum soziotechnischer Systeme. Sie wären der Zusammenbruch global vernetzter Zivilisation.

Big Data bricht nicht nur in die Sozial- und Wirtschaftswissenschaften ein, sondern macht auch vor den Geisteswissenschaften nicht halt. Die Rede ist von «Digital Humanities». Es geht nicht nur um Digitalisierung von Texten, sondern ihre statistische Auswertung und Vermessung. Ein Leser kann nur begrenzt viele Texte zu einem Thema lesen. Hier kommt Big Data ins Spiel, in dem aus der Masse der maschinell erfassten Dokumente Korrelationen und Zusammenhänge bestimmt werden. Diese Kontexte, so hofft man, lassen auch auf Textinhalte von Literatur schließen. Andere befürchten die Ausreibung des Geistes aus den Geisteswissenschaften. Die Linguistik meldet bereits Erfolge mit Textverarbeitungssystemen, die selbstständig im Stil eines Autors Texte verfassen. Auch hier hält digitale Personalisierung und Profilbildung Einzug. Was ist von Filmen und Fernsehsendungen, Musik, Kunst und Theater zu halten, die aufgrund von Big Data Korrelationen den Publikumsgeschmack genau vorausberechnen, um Quote und Zustimmung zu erreichen? Spätestens

dann stellt sich die Frage nach Ursache und Gründen von Kreativität, die immer wieder für Überraschungen sorgte, auf Widerstand stieß und schließlich überzeugte. So manche Biographie von großen Schriftstellern, Malern und Komponisten führt uns das eindrucklich vor Augen.

Schließlich stellt sich die Frage nach der Berechenbarkeit unserer Zukunft (*Kapitel 13*). Formeln erweisen sich als Schlüssel zur Macht – Macht über die Natur, Macht über die Gesellschaft und Macht über die Zukunft. Ein Blick in die Geschichte zeigt, wie Theoretiker die Grundlagen der Zukunft gelegt haben. Wie weit reichen aber ihre Modelle und Gesetze? Werden tatsächlich Roboter unsere Alltagsarbeit übernehmen, intelligente Infrastrukturen sich selber steuern, Nanosensoren unseren Körper täglich auf entstehende Krankheiten scannen, Medizintechnologien verbrauchte Organe durch gezüchtetes Gewebe ersetzen, Drohnen entstehende Konflikte frühzeitig erkennen und Raumschiffe zunächst unser Planetensystem erforschen und dann zu interstellaren Reisen aufbrechen? Das sind Zukunftsvisionen bestehender Technologien und Forschungstrends, die im Rahmen physikalischer Gesetze hochgerechnet werden.

Zukunft bedeutet aber auch Lebensstil und Lebensqualität, die sich Menschen wünschen. Big Data kann hier Trends und unbewusste Neigungen entdecken. Globale Kommunikationssysteme, die Menschen zeitnah mit Daten und Informationen versorgen, bestärken den Wunsch, Zukunft mitgestalten zu wollen und sich nicht vom Diktat der Märkte, Produkte und Technologien überrollen zu lassen. Wir sprechen bereits von partizipativer Demokratie, in der klassische repräsentative Politikformen sich auflösen und Teilhabe und Mitgestaltung in den Vordergrund rücken. Auch diese politische Veränderung wird letztlich durch die Entwicklung von Informationstechnologie und Big Data möglich. Wie weit können wir dabei gehen, ohne das Gesamtsystem nachhaltig zu destabilisieren? Jedenfalls wird deutlich, dass nur von Zukünften und nicht von «der» Zukunft die Rede sein kann.

Am Ende kommen wir noch einmal zum Ausgangspunkt: Warum passen die Mathematik und ihre Formeln so gut auf die Welt? (*Kapitel 14*) Mathematisches Denken ist eine der erstaunlichsten Fähigkeiten des Menschen. Ursprünglich aus unserer evolutionären Fähigkeit

zur Mustererkennung in der Datenvielfalt entstanden, konnten wir uns abstrakte Welten mit mathematischen Strukturen schaffen, von denen nur ein Bruchteil Anwendung in der physischen Welt findet. Die Mathematik lässt uns die Grundlagen und Berechenbarkeit der Welt verstehen, in der wir leben. Die Philosophie lässt uns fragen, warum wir uns auf welche ihrer möglichen Zukünfte einlassen sollten. Datenmassen ohne theoretische Grundlagen bleiben blind. Theorien und Formeln ohne Daten sind aber leer. Am Ende benötigen wir Urteilskraft, die Big Data mit theoretischer Fundierung verbindet, um über unsere Zukunft zu entscheiden.

[...]

Mehr Informationen zu diesem und vielen weiteren Büchern aus dem Verlag C.H.Beck finden Sie unter: www.chbeck.de