
Kultur der Digitalität

Felix Stalder

edition suhrkamp

SV

Felix Stalder

Kultur der Digitalität

Suhrkamp

edition suhrkamp 2679

Erste Auflage 2016

Originalausgabe

© Suhrkamp Verlag Berlin 2016

Alle Rechte vorbehalten, insbesondere das der Übersetzung,
des öffentlichen Vortrags sowie der Übertragung
durch Rundfunk und Fernsehen, auch einzelner Teile.

Kein Teil des Werkes darf in irgendeiner Form
(durch Fotografie, Mikrofilm oder andere Verfahren)
ohne schriftliche Genehmigung des Verlages reproduziert
oder unter Verwendung elektronischer Systeme
verarbeitet, vervielfältigt oder verbreitet werden.

Satz: Satz-Offizin Hümmer GmbH, Waldbüttelbrun

Druck: Druckhaus Nomos, Sinzheim

Umschlag gestaltet nach einem Konzept

von Willy Fleckhaus: Rolf Staudt

Printed in Germany

ISBN 978-3-518-12679-0

Inhalt

Einleitung: Nach dem Ende der Gutenberg-Galaxis . . .	7
1. Wege in die Digitalität	21
Die Erweiterung der sozialen Basis der Kultur	22
Die Kulturalisierung der Welt	58
Die Technologisierung der Kultur	68
Von den Rändern ins Zentrum der Gesellschaft	92
2. Formen der Digitalität	95
Referentialität	96
Gemeinschaftlichkeit	129
Algorithmizität	164
3. Richtungen des Politischen in der Digitalität	203
Postdemokratie	206
Commons	245
Wider die Alternativlosigkeit	279
Dank	282

position darauf gründet, dass sich eine kritische Masse an teilnehmenden Produzenten aus Eigeninteresse freiwillig unterordnet. Zerbricht der Konsens darüber, wessen Beiträgen besonderes Gewicht zugebilligt werden soll, so droht die Formation ihre innere Struktur zu verlieren und sich zu spalten (»forking« im Jargon der Freien Software).

Algorithmizität

Referenzielle Verfahren in gemeinschaftlichen Formationen schaffen durch persönliche Kommunikation unterschiedlich große und umfassende Zonen der Kultur. Sie breiten sich in den Leerstellen aus, die durch die Erosion etablierter Institutionen und Verfahren zur Produktion von sozialer Bedeutung entstanden sind, und nun, da sich neue Verfahren etablieren, verschärfen sie diese Erosionsprozesse noch zusätzlich. Die Vielzahl der parallel zueinander stattfindenden Austauschprozesse schafft einen Flickenteppich verbundener, konkurrierender oder auch einfach voneinander getrennter Sphären der Bedeutung mit jeweils spezifischen Handlungsperspektiven und Ressourcen, eigenen Voraussetzungen und Potenzialen. Die dadurch generierten Wissens-, Ordnungs- und Handlungsstrukturen sind sowohl ganzheitlich als auch partiell und limitiert. Die Teilnehmer an solchen Strukturen werden auf vielen ehemals funktionell getrennten Ebenen gleichzeitig angesprochen, vormals voneinander unabhängige Sphären, etwa Arbeit und Freizeit, vermischen sich, aber meist nur in Bezug auf Teilbereiche des eigenen Lebens. Die dadurch etablierten Strukturen sind zunächst nur für die aktiven Teilnehmer verbindlich.

Ausweg aus der Bibliothek von Babel

Für sich alleine wären diese neuen Verfahren jedoch nicht in der Lage, aus dem enormen Rauschen der chaotischen Informationssphäre mehr als lokale Inseln der Bedeutung zu erzeugen. 1941 entwarf Jorge Luis Borges in seiner Erzählung »Die Bibliothek von Babel« ein Bild für eine solche Situation. Er schildert die Welt als Bibliothek unvorstellbaren, möglicherweise unendlichen Ausmaßes. Ob es darüber hinaus noch eine Welt außerhalb der Bibliothek gibt, ist den Figuren in der Erzählung nicht bekannt. Es gibt Gründe, die dafür, und andere, die dagegen sprechen. Die Bibliothek beherbergt die vollständige Sammlung aller möglichen Bücher, die auf exakt 4^{10} Seiten geschrieben werden können. In ihr enthalten ist das Versprechen, dass es »kein persönliches, kein Weltproblem [gibt], dessen beredte Lösung nicht existierte«, denn jede mögliche Kombination von Buchstaben, und damit auch jede mögliche Aussage, ist in irgendeinem Buch bereits aufgeschrieben. Nur ist für diese Bibliothek noch kein Katalog gefunden worden (der allerdings irgendwo vorhanden sein muss), und auch sonst ist in der Anordnung der Bücher keinerlei Ordnung zu erkennen. Die »Menschen der Bibliothek«, so Borges, wandern herum auf der Suche nach dem alles erklärenden Buch. Ihre tatsächlichen Entdeckungen sind allerdings viel bescheidener. Nur selten werden Bände gefunden, die mehr als zufällige Zeichenkombinationen enthalten. Schon kleine Regelmäßigkeiten in Textausschnitten sind sensationelle Entdeckungen, um die herum sich konkurrierende Schulen der Interpretation bilden. Trotz Arbeit und Mühen sind die gewonnenen Erkenntnisse aber bescheiden und fragmentarisch, so dass eine düstere Grundstimmung überhandgenommen hat. In der Generation des Erzählers

»glaubt niemand [mehr], irgendetwas entdecken zu können«. ⁷⁵

Obwohl diese Vision in quantitativer Hinsicht erfüllt ist – kein Mensch kann die »Bibliothek« digitaler Informationen überblicken, in praktischer Hinsicht ist sie unendlich groß, und alle Wachstumskurven zeigen immer noch weiter steil nach oben –, ist die kulturelle Wirklichkeit dennoch eine ganz andere als die von Borges beschriebene. Unsere Fähigkeit, mit riesigen Datenmengen umzugehen, hat sich radikal verbessert, und entsprechend ist der Glaube an die Nützlichkeit von Informationen nicht nur ungebrochen, sondern nimmt sogar noch weiter zu. Neu sind gerade die ganz großen Datenmengen (Big Data), die vom Versprechen oder der Warnung begleitet werden, zu neuen Erkenntnissen, zu einem umfassenden Verständnis der Welt, ja, zu »Allwissenheit« zu führen. ⁷⁶ Diese Datengläubigkeit beruht vor allem darauf, dass die beiden bisher beschriebenen Verfahren – Referentialität und Gemeinschaftlichkeit – nicht die einzigen neuen Mechanismen sind, mit denen gefiltert, sortiert, aggregiert und gewichtet wird. Den sozialen Mechanismen der dezentralen, vernetzten Kulturproduktion unterlegt beziehungsweise ihnen vorgeschaltet sind algorithmische Prozesse, welche die unermesslich großen Datenmengen vorsortieren und in ein Format bringen, in dem sie überhaupt durch Einzelne erfasst, in Gemeinschaften beurteilt und mit Bedeutung versehen werden können.

Strenggenommen ist eine kategoriale Unterscheidung zwischen sozialen Prozessen, die sich in und durch technologi-

75 Borges, Jorge Luis (1974), *Die Bibliothek von Babel: Erzählungen*, Stuttgart: Reclam.

76 Geiselberger, Heinrich (Hg.) (2013), *Big Data: das neue Versprechen der Allwissenheit*, Berlin: Suhrkamp.

sche Infrastrukturen vollziehen, und technischen Prozessen, die sozial konstruiert sind, nicht aufrechtzuerhalten. In beiden Fällen versuchen gesellschaftliche Akteure, mit den ihnen zur Verfügung stehenden Ressourcen ihre eigenen Interessen zu realisieren. Die Methoden der (versuchten) Realisierung, die Ressourcen und die Ausformulierung der Interessen beeinflussen sich gegenseitig. Die technologischen Ressourcen schreiben sich in die Formulierung der Ziele ein. Diese öffnen Felder der Imagination und des Begehrens, die wiederum die technische Entwicklung inspirieren.⁷⁷ Obwohl es unmöglich ist, klare theoretische Grenzen zu ziehen, kann der Versuch, eine solche Unterscheidung zu machen, in der Praxis dennoch produktiv sein, da sich auf diese Weise verschiedene Perspektiven hinsichtlich desselben Untersuchungsgegenstands gewinnen lassen.

Der Aufstieg der Algorithmen

Ein Algorithmus ist eine Handlungsanleitung, wie mittels einer endlichen Anzahl von Schritten ein bestehender Input in einen angestrebten Output überführt werden kann: Mithilfe von Algorithmen werden vordefinierte Probleme gelöst. Damit eine Handlungsanleitung zum Algorithmus wird, muss sie in dreifacher Hinsicht determiniert sein. Erstens müssen die Schritte, einzeln und in ihrer Gesamtheit, eindeutig und vollständig beschrieben sein. Dazu ist in der Regel eine forma-

⁷⁷ Dies ist eine der zentralen Erkenntnisse der Science and Technology Studies. Siehe z. B. Bowker, Geoffrey C. und Susan Leigh Star (1999), *Sorting Things Out. Classification and Its Consequences*, Cambridge, MA: MIT Press.

le Sprache notwendig, etwa die Mathematik oder eine Programmiersprache, um die für natürliche Sprachen charakteristischen Unschärfen und Mehrdeutigkeiten zu beseitigen, damit Anweisungen ohne Interpretation angewandt werden können. Zweitens müssen die einzelnen Schritte zusammen praktisch durchführbar sein. Deshalb ist jeder Algorithmus auf den Kontext seiner Realisierung bezogen. Verändert sich dieser, verändert sich auch, welche Handlungsabläufe als Algorithmen formalisiert werden können, und damit, in welcher Weise Algorithmen an der Konstitution der Welt teilhaben. Drittens muss eine Handlungsanweisung mechanisch ausführbar sein, damit sie unter unveränderten Voraussetzungen immer dasselbe Resultat zeitigt.

So allgemein definiert könnte man auch die Bauanleitung, die einem typischen Ikea-Möbelstück beiliegt, als Algorithmus verstehen. Sie ist eine Handlungsanleitung, wie in einer endlichen Zahl von Schritten aus einer Kiste voller Einzelteile (Input) ein bestimmtes vordefiniertes Möbelstück (Output) erstellt werden kann. Die Anleitung ist in einer formalen Sprache verfasst, Piktogramme, die jeden Schritt möglichst eindeutig definieren, und sie ist durch einen einzelnen Menschen mit einfachen Werkzeugen ausführbar. Der Ablauf ist wiederholbar, denn am Ende steht immer dasselbe Resultat: Aus einer Billy-Kiste wird immer ein Billy-Regal. Der Mensch übernimmt hier die Rolle der ausführenden Maschine, was trotz all der eindeutig festgelegten Piktogramme zu Problemen führen kann – sei es, dass am fertigen Möbelstück Kratzer und andere Spuren die Einzigartigkeit der (misslungenen) Ausführung bezeugen oder dass, inspiriert vom Mikro-Trend »Ikea Hacking«, bewusst entgegen der offiziellen Anleitung gehandelt wird.

Weil solche Unschärfen vermieden werden sollen, ist die

wichtigste Domäne der Algorithmen in der Praxis die Mathematik und ihre Implementierung im Computer. Der Begriff »Algorithmus« geht auf den im heutigen Usbekistan geborenen Mathematiker, Astronomen und Geografen Muḥammad Ibn-Mūsā al-Ḥwārizmī zurück. Sein um 825 in Bagdad verfasstes Lehrbuch *Über das Rechnen mit indischen Ziffern* wurde in lateinischen Übersetzungen im Mittelalter breit rezipiert und trug wesentlich dazu bei, dass die indisch-arabischen Zahlen und die Ziffer Null ihren Weg nach Europa fanden. Das Buch beginnt mit der Formel *dixit algorizmi ...* (»Algorismi hat gesagt ...«). Im Mittelalter wurde *algorizmi* oder *algorithmi* bald zu einem allgemeinen Begriff für avancierte Rechenoperationen.⁷⁸

Das moderne Streben, Maschinen zu konstruieren, um Anleitungen auch mechanisch auszuführen zu können, erreichte mit Gottfried Wilhelm Leibniz einen ersten Höhepunkt. Ihm wird in Informatiklehrbüchern häufig die Aussage zugeschrieben, dass es »eines ausgezeichneten Mannes nicht würdig [ist], wertvolle Stunden wie ein Sklave im Keller der einfachen Rechnungen zu verbringen. Diese Aufgaben könnten ohne Besorgnis abgegeben werden, wenn wir Maschinen hätten«.⁷⁹ Diese Vision enthält bereits die Unterscheidung zwischen höheren kognitiv-interpretativen Tätigkeiten, die als wahrlich menschlich angesehen werden, und niederen, rein ausführenden, deshalb mechanisierbaren Abläufen. Leibniz selbst entwickelte zu diesem Zweck die erste Rechenmaschine, die alle vier Grund-

78 Krämer, Sybille (1988), *Symbolische Maschinen: Die Idee der Formalisierung in geschichtlichem Abriß*, Darmstadt: Wissenschaftliche Buchgesellschaft, S. 50ff.

79 Etwa in Eigner, Martin, Florian Gerhardt und Torsten Gilz u. a. (2012), *Informationstechnologie für Ingenieure*, Berlin/Heidelberg: Springer.

rechenarten ausführen konnte. Die Motive seines Strebens war nicht die praktische Notwendigkeit von Produktion und Handel – Leibniz' Rechenmaschine blieb, obwohl konzeptionell wegweisend, aufgrund ihrer mechanischen Komplexität ein Einzelstück und wurde nicht angewendet.⁸⁰ »Vielmehr«, so die Einschätzung der Philosophin Sybille Krämer, »sind [die Rechenmaschinen] spekulative Meisterstücke eines Jahrhunderts, welches wie keines zuvor von der Idee der Mechanisierung ›geistiger‹ Abläufe durchdrungen« war.⁸¹ Schon lange bevor Maschinen in großem Umfang eingesetzt wurden, um die materielle Produktion effizienter zu gestalten, dachte Leibniz über ihren Einsatz zur Steigerung geistiger Arbeit nach. Und diese Vision ist nie wieder vollkommen verschwunden. Knapp hundertfünfzig Jahre später formulierte sie der Brite Charles Babbage neu, nun in direkter Verbindung mit der industriellen Mechanisierung und deren Imperativ der zeitsparenden Effizienz.⁸² Doch auch er scheiterte noch am Problem der praktischen Realisierbarkeit einer solchen Maschine.

Den entscheidenden Schritt, um die Vision von kalkulierenden Maschinen Wirklichkeit werden zu lassen, lieferte 1937 Alan Turing. In einem theoretischen Modell wies er nach, dass jeder Algorithmus durch eine Maschine ausführbar sei, sofern

80 Die von Leibniz vorgeschlagene mechanische Konstruktion konnte erst 1820 als praktisch anwendbare (und daher patentierbare) Rechenmaschine unter dem Namen Arithmometer realisiert werden.

81 Krämer, *Symbolische Maschinen*, S. 98.

82 »Die Arbeitsteilung, das dürfte einigen unserer Leser vielleicht paradox erscheinen, [lässt] sich mit gleichem Erfolg und gleicher Zeitersparnis auf geistige Operationen anwenden« (Babbage, Charles [1999 [1832]], *Die Ökonomie der Maschine*, Berlin: Kulturverlag Kadmos, S. 147).

diese in der Lage ist, Schritt für Schritt vorgegebene Zeichen zu lesen, sie nach festgelegten Regeln zu manipulieren und danach wieder auszuschreiben. Für die Gültigkeit dieses Modells spielte es keine Rolle, ob die Maschine analog oder digital, mechanisch oder elektronisch realisiert wurde, denn die Regeln der Manipulation waren zunächst als fester Bestandteil der Maschine gedacht (also in der Hardware implementiert). Den Ausschlag für den elektronisch/digitalen Ansatz gab der Wunsch, dass auch die Instruktionen als Input von der Maschine selbst gelesen werden sollten, so dass die Maschine nicht nur einen, sondern (theoretisch) jeden entsprechend verfassten Algorithmus ausführen können sollte. Umsetzen sollte diese Idee der aus Ungarn stammende Mathematiker John von Neumann. 1945 veröffentlichte er ein Modell, in dem das Programm (der Algorithmus) und die Daten (Input und Output) in einem gemeinsamen Speicher untergebracht waren. So ließ sich beides gleichermaßen manipulieren, ohne dass die Hardware verändert werden musste. Aus der »Turingmaschine« wurde die »universale Turingmaschine«, das heißt der moderne Computer.⁸³

Zwanzig Jahre später prognostizierte Gordon Moore, Mitbegründer des Chipherstellers Intel, dass sich die Komplexität integrierter Schaltkreise und damit die Leistungsfähigkeit von Computerchips in Zukunft alle 18 bis 24 Monate verdoppeln werde. Seit den siebziger Jahren ist seine Voraussage als Moores Gesetz bekannt und ist im Wesentlichen bis heute korrekt, die technische Entwicklung verlief tatsächlich entsprechend exponentiell, nicht zuletzt deshalb, weil sich die Halbleiter-

83 Diese als »Von-Neumann-Architektur« bezeichnete Struktur bildet bis heute die Grundlage fast aller Computer.

industrie an ihr orientierte.⁸⁴ Ein IBM 360/40 Großrechner aus der ersten in bedeutender Stückzahl produzierten Serie konnte etwa 40 000 Kalkulationen pro Sekunde ausführen und kostete bei Markteinführung im Jahre 1965 1,5 Millionen Dollar pro Stück. Knapp vierzig Jahre später konnte ein handelsüblicher Server (zum Beispiel mit einem Vierkern-Intel-Xeon-Prozessor) bereits mehr als vierzig Milliarden Kalkulationen pro Sekunde ausführen, und das zu einem Stückpreis von nur wenig mehr als 1500 Dollar. Das entspricht einer Leistungssteigerung um den Faktor eine Million bei gleichzeitigem Preisrückgang um den Faktor eintausend, also einer Verbesserung des Preis-Leistungs-Verhältnisses um den Faktor eine Milliarde. Würde man die Inflation berücksichtigen, wäre der Faktor noch höher. Ähnlich dramatisch waren die Leistungssteigerungen oder vielmehr der Preisrückgang im Bereich der Datenspeicherung: Ein Gigabyte Daten zu speichern kostete 1980 mehr als 400 000 Dollar, dreißig Jahre später nur noch zehn Cents, ein Preisrückgang um den Faktor vier Millionen. Und die Entwicklung geht auf beiden Gebieten unvermindert weiter.

Diese Steigerungen der Leistungsfähigkeit bilden die materielle Grundlage für die rasant angewachsene Zahl algorithmisch ausführbarer Handlungen. Heute ist der Punkt erreicht, an dem die von Leibniz getroffene Unterscheidung zwischen schöpferisch-kreativen mentalen Funktionen und »einfachen Rechnungen« zunehmend verwischt wird. Die Diskussion der letzten Jahre um die angeblich drohende »Herrschaft der Computer« wurde weniger durch den vermehrten Einsatz von Algorithmen als solchen befeuert als durch das langsame

84 The Inquirer Staff (2005), »Gordon Moore says aloha to Moore's Law«, in: *The Inquirer*, 13.04. (online).

Verwischen dieser Unterscheidung und den neuen Möglichkeiten, immer weitere Bereiche des vermeintlich kreativen Denkens zu formalisieren und zu mechanisieren.⁸⁵ Tätigkeiten, die noch vor Kurzem unzweifelhaft der menschlichen Intelligenz vorbehalten schienen, beispielsweise das Verfassen von Texten oder die Inhaltsanalyse von Bildern, übernehmen inzwischen immer häufiger Maschinen. Bereits 2010 wurde die Anwendung Stats Monkey vorgestellt, die kurze Berichte zu Baseballspielen anfertigen kann. Alles, was das Programm dafür benötigt, sind umfassende Daten zu den Spielen, die mittlerweile routinemäßig erhoben und dank verbesserter algorithmischer Bilderkennung und neuer Sensoren immer detaillierter werden. Das Programm extrahiert aus den Daten die entscheidenden Momente und Akteure eines Spiels, erkennt charakteristische Muster im Spielverlauf (etwa »frühe Führung stetig ausgebaut«, »dramatische Aufholjagd« oder Ähnliches) und generiert darauf aufbauend eigene Berichte. Dabei kann eine Vielzahl von Variablen bestimmt werden, etwa ob ein Artikel aus der Perspektive eines neutralen Beobachters oder vom Standpunkt eines der beiden Teams geschrieben werden soll. Oder man kann, weil Eltern nicht gerne über die Fehler ihrer Kinder lesen, festlegen, dass vor allem die gelungenen Spielzüge hervorgehoben werden sollen, wenn es um Partien in der Schülerliga geht. Der Algorithmus wurde rasch patentiert, und aus dem ursprünglich interdisziplinären Forschungsprojekt wurde ein Start-up-Unternehmen: Narrative Science. Dieses bietet heute Texte aller Art an, neben Sport- vor allem

85 Meckel, Miriam (2011), *Next. Erinnerungen an eine Zukunft ohne uns*, Reinbek bei Hamburg: Rowohlt. Man könnte auch sagen, dass die Angst dadurch ausgelöst wird, dass die Mechanisierung der Arbeit nun auch die Arbeitsfelder der Mittelklasse zu erfassen beginnt.

auch Finanzberichterstattung, ebenfalls ein Feld, für das statistische Informationen in Fülle vorliegen. Diese Texte werden in renommierten Medien wie dem Wirtschaftsmagazin *Forbes* veröffentlicht und dort mit dem Autorenvermerk »narrative science« gekennzeichnet. Noch beschränken sich diese Beiträge auf relativ einfache Themen. Dabei soll es aber nicht bleiben. Auf die Frage, wie groß der Anteil algorithmisch verfasster Nachrichten in fünfzehn Jahren sein werde, antwortete Kristian Hammond, Technologiechef von Narrative Science und einer der Gründer des Unternehmens, 2012 selbstbewusst mit »neunzig Prozent« – und fügte gleich noch hinzu, dass ein Algorithmus schon in fünf Jahren zum ersten Mal den Pulitzer-Preis gewinnen werde.⁸⁶ Da wird einiges an Hype und Selbstmarketing mitschwingen, aber als generelle Einschätzung ist seine Aussage nicht unglaubwürdig. Ob die Algorithmen den klassischen Journalismus ersetzen oder ihn ergänzen werden, ist noch offen. Da die Medienhäuser inzwischen aber unter starkem finanziellen Druck stehen, ist es durchaus realistisch, dass viele journalistische Texte in Zukunft automatisiert erstellt werden. Es lassen sich aber auch ganz andere Anwendungen denken. So sieht der Publizist Alexander Pschera ein neues Zeitalter im Verhältnis zwischen Mensch und Natur anbrechen, da mit Sendern und Sensoren bestückte Tiere uns als Individuen und nicht mehr nur als Gattungswesen entgegentreten, wenn sie erst mithilfe entsprechender Software ihre eigene Geschichte erzählen können.⁸⁷

Noch ist es allerdings nicht so weit. Aber dass sich auch die

86 Levy, Steven (2012), »Can an algorithm write a better news story than a human reporter?«, *wired.com*, 24.04. (online).

87 Pschera, Alexander (2014), *Das Internet der Tiere. Der neue Dialog zwischen Mensch und Natur*, Berlin: Matthes & Seitz.

CIA für Narrative Science interessiert und über ihre Beteiligungsgesellschaft In-Q-Tel in das Unternehmen investiert hat, ist ein Hinweis darauf, dass bereits heute Anwendungen jenseits des Journalismus entwickelt werden. Zu Propagandazwecken lassen sich mit Algorithmen beispielsweise problemlos massenhaft Einträge in Onlineforen und in den sozialen Massenmedien erstellen.⁸⁸ Narrative Science ist nur eine von vielen Firmen, die automatisierte Textanalyse und -produktion anbieten. Sogenannte E-Discovery-Software, wie sie unter anderem von IBM vertrieben wird, verspricht, den zeitlichen Aufwand dramatisch zu reduzieren, der nötig ist, um die stetig wachsenden Aktenmengen bei komplexen Rechtsfällen zu analysieren. Ohne solche Software wäre die juristische Aufarbeitung ab einer gewissen Dokumentenmenge praktisch unmöglich. Auch bei Wikipedia sind zahlreiche Bots (automatisierte Editionsprogramme) aktiv. In der deutschen Ausgabe ist es ihnen noch verboten, eigene Artikel anzulegen, doch in der schwedischen Variante ist das anders. Gemessen an der Zahl der Einträge ist sie mittlerweile die weltweit zweitgrößte Ausgabe der Onlineenzyklopädie, denn im Sommer 2013 legte ein einzelner Bot mehr als 200 000 Artikel an.⁸⁹ Darüber hinaus bietet die Firma Epagogix seit 2003 Software an, die basierend auf historischen Daten Filmskripts auf ihr Marktpotenzial prüft. Mindestens eines der großen Holly-

88 Die amerikanischen Geheimdienste sind hier keineswegs einzigartig. *Der Spiegel* berichtete martialisch, dass auf russischer Seite ganze »Bot-Armeen in die Propaganda-Schlacht« geschickt würden (Bilder, Benjamin [2015], »Nemzow-Mord: Die Propaganda der russischen Hardliner«, in: *Spiegel Online*, 28.02. ([online])).

89 Guldbrandsson, Lennart (2013), »Swedish Wikipedia surpasses 1 million articles with aid of article creation bot«, [blog.wikimedia.org](http://blog.wikimedia.org/2013/06/17/06), 17.06. (online).

woodstudios nutzt diese Software, ohne Drehbuchautoren und Regisseure einzubeziehen, denn diese würden »nervös«, so der CEO von Epagogix, wenn sie davon wüssten, wie ihre kreativen Werke analysiert werden.⁹⁰ Oder man denke an die typische Ansage zu Beginn eines Anrufs bei einer Telefonhotline – »Dieser Anruf kann zu Schulungszwecken aufgezeichnet werden«. Geschult werden sollen aber immer öfter nicht die Angestellten im Call-Center, sondern die Algorithmen. Diese sollen lernen, den Persönlichkeitstypus des Anrufers zu erkennen und darauf basierend das Skript für die Interaktion mit den Kunden anzupassen, das die schlecht ausgebildeten und befristet angestellten Mitarbeiter anleitet.⁹¹ Ein anderes Beispiel sind Algorithmen, die Studentenessays benoten können,⁹² oder ... Es ist nicht notwendig, diese Liste mit Beispielen noch weiter zu ergänzen. Auch ohne weitere Hinweise auf vergleichbare Entwicklungen im Bereich der Bild-, Ton-, Sprach- und Filmanalyse ist schon jetzt deutlich, dass an vielen Fronten die Grenzen zwischen dem, was als kreativ, und dem, was als mechanisch verstanden wird, verschoben werden.⁹³

90 Bunnell, Thomas (2007), »The mathematics of film«, in: *Boom Magazine*, November, S. 51 (online).

91 Steiner, Christopher (2012), »Automatons get creative«, in: *Wall Street Journal*, 17.08. (online).

92 Dies war die Aufgabe bei einem 2012 veranstalteten und mit 100 000 Dollar dotierten Preisausschreiben. 159 Teilnehmer lieferten je einen eigenen Algorithmus ab. Die Besten erzielten Resultate vergleichbar mit jenen menschlicher Tutoren und Assistenten. Siehe »The Hewlett Foundation: Automated essay scoring«, kaggle.com, 10.02.2012 (online).

93 Ayres, Ian (2007), *Super Crunchers: How Anything Can Be Predicted*, London: Bookpoint.

Dynamische Algorithmen

Die Algorithmen, die für solche Aufgaben eingesetzt werden, sind jedoch nicht mehr einfache Abfolgen von statischen Instruktionen. Sie werden nicht immer wieder unverändert ausgeführt, sondern sind hochgradig dynamisch und adaptiv. Die heute zur Verfügung stehende Rechenleistung wird genutzt, um Programme zu schreiben, die sich halbautomatisch und auf der Grundlage von Feedback selbst verändern und bessern können.

Was das bedeutet, lässt sich exemplarisch anhand von evolutionären und selbstlernenden Algorithmen erläutern. Ein evolutionärer Algorithmus wird in einem iterativen Verfahren entwickelt, das so lange durchlaufen wird, bis das gewünschte Resultat erzielt wird. Die Werte der Variablen der ersten Generation von Algorithmen werden dabei meist zufällig ausgewählt, um den Einfluss von Vorannahmen der Programmierer auf die Ergebnisse zu verringern. Ganz lässt sich dieser aber nicht vermeiden, denn die Art der Variablen (unabhängig von ihrem Wert) muss ja zunächst überhaupt erst einmal bestimmt werden. Auf dieses Problem komme ich noch zu sprechen. Danach erfolgt eine Phase der Evaluation: Der Output jedes der getesteten Algorithmen wird dahingehend bewertet, wie nahe er der gewünschten Lösung kommt. Anschließend werden die besten ausgewählt und miteinander kombiniert. Zudem werden Mutationen, also zufällige Veränderungen, eingeführt. Danach werden diese Schritte so oft wiederholt, bis der Algorithmus den Vorgaben entsprechend »gut genug« ist beziehungsweise sich keine weiteren Verbesserungen einstellen. Dabei werden Algorithmen in extrem rechenaufwendigen Verfahren quasi »gezüchtet«, das heißt in einer hohen Zahl erprobt, anstatt einen einzigen analytisch zu entwerfen

und dann zu implementieren. Im Zentrum der Suche steht eine funktionsfähige Lösung, die sich experimentell und in der Praxis bewährt, von der man aber hinterher möglicherweise nicht mehr weiß, warum sie funktioniert, oder ob sie wirklich die bestmögliche Lösung ist. Die grundlegenden Verfahren für dieses Vorgehen stammen meist aus den siebziger Jahren, der ersten Phase der künstlichen Intelligenz, nur sind sie heute um einiges leistungsfähiger. Eines der bekanntesten Beispiele für einen evolutionär entwickelten Algorithmus ist sicherlich »Google Flu Trends«. Um vorherzusagen, welche Regionen besonders von der alljährlichen Grippewelle betroffen sein werden, wertet er die geografische Verteilung von Suchanfragen für bestimmte Begriffe aus (zum Beispiel »Erkältungsmittel«). Für die Entwicklung des Programms testete Google 450 Millionen verschiedene Modelle – bis sich eines herauschälte, das lokale Grippeepidemien relativ zuverlässig ein bis zwei Wochen vor den nationalen Gesundheitsbehörden feststellen konnte.⁹⁴

In diesen Größenordnungen lassen sich die nötigen Prozesse nur noch mithilfe von Computerprogrammen bewältigen. Die Testreihen werden nicht länger von Programmierern durchgeführt, sondern von Algorithmen. Kurz: Man setzt Algorithmen ein, um neue Algorithmen zu schreiben beziehungsweise deren Variablen zu bestimmen. Wird nun dieser reflexive Prozess wiederum in einen Algorithmus eingebaut, wird die-

94 Getestet wurden diese Modelle jeweils auf Basis der 50 Millionen häufigsten Suchbegriffe aus den Jahren 2003-2008, geordnet nach Zeit und Ort der Abfrage. Die Resultate wurden gegen die Daten der Gesundheitsbehörden abgeglichen. Siehe Ginsberg, Jeremy u. a. (2009), »Detecting influenza epidemics using search engine query data«, in: *Nature* 457/7232, S. 1012-1014 (online).

ser »selbstlernend«: Die Programmierer legen nicht die Regeln für seine Ausführung fest, sondern Regeln, nach denen der Algorithmus lernen soll, ein bestimmtes Ziel zu erreichen. Seine Lösungsstrategien sind in vielen Fällen so komplex, dass sie sich nicht einmal im Nachhinein nachvollziehen lassen. Sie können nur noch experimentell, aber nicht mehr logisch überprüft werden. Solche Algorithmen sind im Grunde Black Boxes, Objekte, die nur über ihr äußeres Verhalten verstanden werden können, deren innere Struktur sich aber der Erkenntnis entzieht.

Automatische Gesichtserkennung, wie sie etwa in Überwachungstechnologien und bei der Authentifizierung von Zugangsberechtigungen zum Einsatz kommt, beruht darauf, dass Computer große Mengen an Bildern von Gesichtern auswerten, um dann im ersten Schritt ein allgemeines Modell eines Gesichts zu erstellen und danach jene Variablen zu identifizieren, die ein Gesicht einmalig und damit wiedererkennbar machen. Mit sogenannten *Unsupervised-* oder *Deep-learning-*Algorithmen gehen einige Entwickler und Unternehmen sogar noch einen Schritt weiter: Computer sollen lernen, Gesichter aus unstrukturierten Bildern, also aus Bildmengen, die sowohl Bilder mit als auch ohne Gesichter enthalten, zu extrahieren – und das ohne vorher ein Modell des Gesichts zu besitzen. Aus unstrukturiertem Material unbekannte Muster extrahieren und auswerten, das gelang bisher nur mit sehr einfachen Mustern – etwa Kanten oder Flächen in Bildern –, denn solche Lernprozesse zu programmieren ist äußerst komplex und rechenintensiv. Aber in den letzten Jahren hat die verfügbare Rechenleistung enorme Sprünge gemacht, und sowohl die Dateninputs als auch die Komplexität der Lernmodelle haben exponentiell zugenommen. Ausgehend von einfachen Mustern können Algorithmen heute Schritt für Schritt

komplexe Bildinhalte immer besser erkennen. Sie verfeinern sich sozusagen selbstständig. Mit dem Begriff *deep learning* soll eben diese Vielschichtigkeit angezeigt werden. 2012 konnte Google die Leistungsfähigkeit der neuen Programme auf beeindruckende Weise demonstrieren: Aus rund zehn Millionen zufällig ausgewählten Einzelbildern aus YouTube-Videos, analysiert in einem Cluster von 1000 Rechnern mit 16000 Prozessoren, konnte in nur drei Tagen ein Modell erstellt werden, das die Gesichtserkennung in unstrukturierten Bildern um siebzig Prozent verbesserte.⁹⁵ Der Algorithmus »weiß« natürlich nicht, was ein Gesicht ist, aber er erkennt zuverlässig eine Klasse von Formen, die Menschen als Gesicht bezeichnen. Ein Vorteil eines Modells, das nicht auf Basis vorgegebener Parameter entsteht, ist etwa, dass es Gesichter auch in nichtstandardisierten Situationen (zum Beispiel wenn sich die Person im Hintergrund befindet, das Gesicht halb verdeckt ist oder aus einem spitzen Winkel aufgenommen wurde) identifizieren kann. Dank dieser Technik kann man direkt nach Bildinhalten suchen und nicht wie bislang primär nach Bildbezeichnungen. Solche Algorithmen werden auch eingesetzt, um Menschen auf Bildern zu identifizieren und sie etwa in sozialen Netzwerken mit dem Profil der Person, die auf dem Bild zu sehen ist, zu verbinden, ganz ohne Zutun der Nutzer selbst. Solche Algorithmen sollen auch dazu dienen, direkt Handlungen in der »unstrukturierten« Realität zu steuern, beispielsweise in selbstfahrenden Autos oder anderen auto-

95 Absolut gesehen war die Rate der Treffer zwar immer noch relativ niedrig, nur 15,8% der Treffer waren korrekt. Zufälliges Raten jedoch hätte nur eine Trefferquote von 0,005% gebracht. Siehe Quoc, V. Le et al. (2012), »Building high-level features using large scale unsupervised learning«, research.google.com (online).

nomen mobilen Anwendungen, die besonders für das Militär von großem Interesse sind.

Solche Algorithmen können direkt auf Veränderungen der Umwelt reagieren und sich anpassen. Mit dieser Rückkopplung verkleinert sich aber auch der Zeithorizont, innerhalb dessen sie repetitive und damit vorhersehbare Ergebnisse generieren. Algorithmen, die Kräfte der Berechenbarkeit, können so selbst unberechenbar werden. An der Börse sind sogenannte *subsecond extreme events*, das heißt Kursschwankungen, die in weniger als einer Sekunde geschehen, bereits häufig zu beobachten.⁹⁶ Dramatische *flash crashes*, wie jener vom 6. Mai 2010, als der Dow-Jones-Index in wenigen Minuten (und damit für Menschen wahrnehmbar) knapp 1000 Punkte verlor, sind bisher jedoch noch recht selten.⁹⁷ Mit der Einführung der Sprachsteuerung von Mobiltelefonen (etwa Apples Siri im Jahre 2011) erreichten die ersten Programme, die auf selbstlernenden Algorithmen aufbauen, die breite Öffentlichkeit und dringen nun in immer weitere Bereiche des Alltags vor.

96 Johnson, Neil u. a. (2013), »Abrupt rise of new machine ecology beyond human response time«, in: *Nature. Scientific Reports* 3 (online). Die Autoren beziffern die Zahl dieser Events zwischen Januar 2006 und Februar 2011 auf 18520, also im Durchschnitt etwa 15 pro Handelstag.

97 Nestler, Gerald (2014), »Mayhem in Mahwah: The case of the flash crash; or, forensic re-performance in deep time«, in: Forensic Architecture (Hg.), *Forensis: The Architecture of Public Truth*, Berlin/London: Sternberg Press, S. 125-146.

Durch Algorithmen generierte Ordnungen sind ein konstitutiver Bestandteil der Kultur der Digitalität. Zum einen ist die maschinelle Vorsortierung der (informationellen) Welt Voraussetzung dafür, überhaupt die riesigen unstrukturierten Datenmengen bewältigen zu können, zum anderen liefern diese großen Datenmengen und die Rechenzentren, in denen sie gespeichert und prozessiert werden, die materiellen Voraussetzungen, um immer komplexere Algorithmen zu entwickeln. Notwendigkeiten und Möglichkeiten treiben sich gegenseitig voran.⁹⁸

Die wohl bekanntesten Algorithmen, welche die digitale Infosphäre sortieren und in der heutigen Form überhaupt erst nutzbar machen, sind die der Suchmaschinen, allen voran Googles PageRank. Dank ihnen finden wir uns in einer Welt unstrukturierter Informationen zurecht und können immer größere Teile der (informationellen) Welt in die Ordnung der Unstrukturiertheit überführen, ohne dass dadurch die Bibliothek von Babel entsteht. »Unstrukturiert« heißt hier, dass es keine vorgegebene Ordnung gibt, wie etwa, um noch einmal das Bild von der Bibliothek zu bemühen, ein Katalogisierungs-

98 Einen Eindruck von der Geschwindigkeit des Fortschritts liefert ein anderer Google-Bilderkennungsalgorithmus. Dieser konnte bereits 2011 mit 80%er Wahrscheinlichkeit Hunde in Bildern korrekt identifizieren. In gerade einmal drei Jahren wurde diese Rate nicht nur auf 93,5% gesteigert (was in etwa den Fähigkeiten eines Menschen entspricht). Der Algorithmus konnte nun auch noch über 200 verschiedene Hunderassen identifizieren, was kaum ein Mensch vermag (McMillan, Robert [2015], »This guy beat Google's super-smart AI – but it wasn't easy«, wired.com, 15.01. [online]).

system, das jedem Buch einen spezifischen Platz in einem Regal zuweist. Vielmehr liegen alle Bücher kreuz und quer im Raum verteilt und werden je nach Anfrage dynamisch angeordnet, so dass die passenden Bücher für jeden Besucher immer gleich am Eingang bereitstehen. Doch die Metapher mit den kreuz und quer verteilten Büchern ist problematisch, denn »Unstrukturiiertheit« bedeutet nicht einfach die Abwesenheit jeglicher Struktur, sondern die Präsenz einer anderen Art von Ordnung, einer Metastruktur, einer Ordnungspotenz, aus der sich zahllose konkrete Ordnungen ad hoc generieren lassen. Diese Metastruktur wird von Algorithmen erstellt. Anschließend leiten sie aus ihr eine tatsächliche Ordnung ab, die dem Nutzer begegnet, wenn er etwa die Liste mit den Treffern durchscrollt, die ihm eine Suchmaschine liefert. Was der Nutzer nicht sieht, sind die komplexen Voraussetzungen für die Zusammenstellung seiner Suchergebnisse: Allein der Google-Index umfasste beispielsweise Mitte 2014 laut Firmenangaben mehr als hundert Millionen Gigabytes an Daten.

Ursprünglich, also in der zweiten Hälfte der neunziger Jahre, funktionierte PageRank so, dass der Algorithmus die Struktur der Links im World Wide Web analysierte und dabei erstens auf die Zahl der Links achtete, die auf ein bestimmtes Dokument verwiesen, und zweitens die »Relevanz« der Seite bewertete, von der aus ein Link zum fraglichen Dokument führte. Die Relevanz einer Seite wurde wiederum durch die Anzahl der Links bestimmt, die auf diese Seite führten. Aus diesen beiden Variablen wurde jedem von der Suchmaschine erfassten Dokument ein Wert, der PageRank, zugewiesen. Dieser diente dazu, die zu einem Suchbegriff gefundenen Dokumente als hierarchische Liste (Suchresultate) zu präsentieren, wobei das Dokument mit dem höchsten Wert zuerst ange-

zeigt wurde.⁹⁹ Dieser Algorithmus war überaus erfolgreich, weil er das unüberblickbare Chaos des World Wide Web auf eine Aufgabe reduzierte, die ein einzelner Nutzer problemlos meistern konnte: einen Suchbegriff eingeben und aus den angebotenen »Treffern« einen auswählen. Mit der Verbindung aus Qualität der algorithmischen Vorauswahl und Einfachheit der menschlichen Endauswahl setzte sich Google rasch gegen die Konkurrenz durch.

Dem Verfahren liegt die Annahme zugrunde, dass jeder Link ein Hinweis auf Relevanz ist und dass Links von häufig verlinkten, das heißt populären Quellen wichtiger sind als solche von weniger verlinkten, also unpopulären. Der Vorteil dieser Annahme ist, dass sie sich in rein quantitativen Variablen erfassen lässt und keinerlei direktes Verständnis für den Inhalt der Dokumente oder den Kontext, in dem sie sich befinden, notwendig ist.

Das Problem, wie sich die Relevanz von Dokumenten beurteilen lässt, deren Inhalt bestenfalls ansatzweise erfasst werden kann, war Mitte der neunziger Jahre, als die erste Version des PageRank-Algorithmus entwickelt wurde, nicht mehr neu. Wissenschaftsadministratoren an Universitäten und Förderinstitutionen sahen sich schon ab den fünfziger Jahren mit dieser Schwierigkeit konfrontiert. Im Zuge der Entstehung der Wissensökonomie war die Zahl wissenschaftlicher Publikationen rasant angestiegen. Auch die wissenschaftlichen Felder, Perspektiven und Methoden hatten sich vervielfacht und ausdifferenziert, so dass selbst Experten ihr eigenes Gebiet

99 Brin, Sergey und Lawrence Page (1998): »The anatomy of a large-scale hypertextual Web search engine«, in: *Computer Networks and ISDN Systems* 30/1-7, S. 107-117 (online).

nicht mehr überblicken konnten.¹⁰⁰ Statt also die unzähligen neuen Publikationen zu lesen und inhaltlich zu beurteilen, verlagerten sie ihre Analyse auf eine höhere Abstraktionsebene. Sie begannen zu zählen, wie häufig ein Aufsatz oder ein Buch zitiert wurde, und verwendeten diese Information, um zu beurteilen, welchen Stellenwert dem Autor oder der Publikation zukam.¹⁰¹ Die zugrunde liegende Annahme war (und ist es bis heute), dass nur auf wichtige Dinge verwiesen wird und deshalb jedes Zitat und jede Referenz als indirektes Votum für Relevanz gelten könne.

In beiden Fällen – der Klassifikation einer chaotischen Informationssphäre und der Administration des expandierenden Wissenschaftsbetriebs – geht es darum, dynamische Ordnungen für sich rasch wandelnde Felder zu entwickeln, anhand denen man Aussagen darüber treffen kann, wie wichtig einzelne Dokumente sind, auch wenn man über deren Inhalte nichts weiß. Weil Zitat- beziehungsweise Linkanalysen rein quantitativ operieren, lassen sich mit ihnen große Datenmengen schnell strukturieren und besonders relevante Positionen bestimmen. Der zweite Vorteil ist, dass dieser Ansatz keine Annahmen über die Konturen von Feldern oder deren Verhältnis zueinander machen muss. Das ermöglicht es, ungeordnete oder dynamische Inhalte zu ordnen. In beiden Fällen werden die von den Akteuren selbst geschaffenen Verweise – Zitate in einem wissenschaftlichen Text, Links auf Webseiten – genutzt. Deren ordnungsstiftender Wert für das Feld als Ganzes wird aber

100 Garfield, Eugene (1955), »Citation indexes for science. A new dimension in documentation through association of ideas«, in: *Science* 122, S. 108-111.

101 Die dafür notwendigen Daten werden seit 1964 als Science Citation Index (SCI) publiziert.

nur in der Aggregation, etwa als Häufigkeit, mit der ein Artikel zitiert wird, sichtbar.¹⁰² In beiden Fällen ist der Wechsel von einer Analyse der »Daten« (Inhalte der Dokumente im herkömmlichen Sinn) zur Analyse der »Metadaten« (Beschreibung der Dokumente im Hinblick auf ihre Beziehungen untereinander) Voraussetzung dafür, die wachsenden Informationsmengen überhaupt nutzen zu können.¹⁰³ Mit dieser Verschiebung wird eine neue Abstraktionsebene eingeführt. Informationen werden nicht länger als Repräsentationen einer externen Realität verstanden, ihre Bedeutung nicht im Hinblick auf die Beziehung zwischen »Information« und »Welt« beurteilt, etwa mit einem qualitativen Kriterium wie »wahr«/»falsch«. Vielmehr wird die Informationssphäre als eine selbstreferentielle, geschlossene Welt aufgefasst, und Dokumente werden dementsprechend nur noch in Bezug auf ihre Position innerhalb dieser Welt bewertet, und zwar anhand quantitativer Kriterien wie »zentral«/»peripher«.

Auch wenn der PageRank-Algorithmus sehr leistungsfähig

102 Die Annahme, dass die Subjekte diese Strukturen indirekt und ohne strategische Absicht produzieren, erwies sich in beiden Kontexten als problematisch. In der Wissenschaft bildeten sich sogenannte Zitationskartelle, Gruppen von Wissenschaftlern, die sich gegenseitig häufig referenzieren, mit dem Ziel, ihre Position im SCI zu verbessern. Mit den Suchmaschinen entstand auch der Industriezweig der Suchmaschinenoptimierer (*search engine optimizers*), die mit allerlei Mitteln versuchen, eine Website für die Bewertung durch den Suchalgorithmus zu optimieren.

103 Zur Geschichte des SCI und seinem Einfluss auf die frühe Fassung des Google PageRank siehe Mayer, Katja (2009), »Zur Soziometrik der Suchmaschinen. Ein historischer Überblick der Methodik«, in: Becker, Konrad und Felix Stalder (Hg.), *Deep Search. Die Politik des Suchens jenseits von Google*, Innsbruck: Studienverlag, S. 64-83.

war und Google schnell zu einer marktbeherrschenden Stellung verhalf, so war er doch zu Beginn relativ einfach und in seiner Wirkungsweise ansatzweise transparent. Er folgte dem klassischen statischen Modell eines Algorithmus. Ein Dokument bzw. eine Seite, auf die viele Links verwiesen, wurde als wichtiger erachtet als eines, auf das wenige Links verwiesen.¹⁰⁴ Der Algorithmus analysierte die gegebene Struktur der Informationsordnung und bestimmte die Position jedes Dokuments darin, weitgehend unabhängig vom Kontext der Suchanfrage, über die keine Annahmen gemacht wurden. Dieser Ansatz funktionierte so lange relativ gut, wie die Informationsmenge eine gewisse Größe nicht überschritt und die Nutzer sowie ihre Anfragen einander einigermaßen ähnlich waren. Beides ist heute nicht länger der Fall. Die Informationsmenge, die es vorzusortieren gilt, wird immer größer, und die Nutzer suchen in allen möglichen Situationen und an allen möglichen Orten nach allem nur Erdenklichen. Zu der Zeit, als Google gegründet wurde, wäre es niemanden in den Sinn gekommen, unterwegs schnell das Tagesmenü des Restaurants um die Ecke im Internet zu suchen, heute ist das dank der Smartphones eine Selbstverständlichkeit.

Algorithmenwolken

Um auf solche Veränderungen des Nutzerverhaltens zu reagieren – und sie dadurch gleichzeitig weiter voranzutreiben –, wird der Google-Suchalgorithmus laufend verändert. Er wird

¹⁰⁴ Eines, auf das keine Links verwiesen, konnte vom Algorithmus gar nicht erfasst werden, denn die Suchmaschinen indizierten das Web, indem ihre »Crawler« selbst den Links folgten.

immer komplexer und nimmt immer mehr kontextuelle Informationen auf, die zusätzlich zum Wert des PageRanks die Ordnung der Suchresultate beeinflussen. Der Algorithmus ist kein feststehendes Objekt mehr, kein unveränderliches Rezept, sondern wandelt sich mehr und mehr zu einem dynamischen Prozess, einer undurchsichtigen Wolke aus vielen interagierenden Algorithmen, die kontinuierlich verfeinert werden, Schätzungen zufolge rund 500-600 Mal pro Jahr. Diese Weiterentwicklungen sind so umfassend, dass seit 2003 jedes Jahr mehrere neue Versionen der Algorithmenwolke mit eigenem Namen versehen werden. Allein 2014 führte Google dreizehn große Updates durch, so viele wie noch nie.¹⁰⁵

Im Zuge dieser Änderungen entstehen immer neue Abstraktionsebenen, so dass der Algorithmus zusätzliche Variablen wie Ort und Zeit der Abfrage ebenso wie das bisherige erfasste Verhalten einer Person berücksichtigt, aber auch ihre Einbindung in ein soziales Umfeld und vieles mehr. 2005 nahm Google diese Personalisierung und Kontextualisierung in seinen Suchalgorithmus auf. Anfänglich konnten die Nutzer selbst wählen, ob sie diese verwenden wollten. Seit 2009 sind sie fester Bestandteil und für alle verbindlich, die über Google eine Suchanfrage starten.¹⁰⁶ Bis Mitte 2013 war der Suchalgorithmus auf mindestens zweihundert Variablen angewachsen.¹⁰⁷ Was relevant ist, bestimmt der Algorithmus also nicht länger primär über die Position eines Dokuments in einer dynamischen, aber für alle existenten äußeren informationellen Welt.

105 »Google algorithm change history«, moz.com (online).

106 Feuz, Martin, Matthew Fuller und Felix Stalder (2011), »Personal web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation«, in: *First Monday* 17/2 (online).

107 Dean, Brian (2013), »Google's 200 ranking factors«, in: *Search Engine Journal*, 31.05. (online).

Stattdessen weist er Inhalten nun innerhalb eines dynamischen und singulären, das heißt auf jeden einzelnen Nutzer eigens zugeschnittenen Informationskosmos einen Platz zu. Für jede Person wird eine andere Ordnung erstellt und nicht mehr nur ein Ausschnitt einer vorgängig bestehenden Ordnung angezeigt. Die Welt wird nicht mehr repräsentiert; sie wird für jeden User eigens generiert und anschließend präsentiert. Google ist nicht das einzige Unternehmen, das diesen Weg eingeschlagen hat. Von Algorithmen erstellte Ordnungen werden immer stärker darauf ausgerichtet, dem individuellen Nutzer seine eigene, singuläre Welt zu schaffen. Facebook, Partnerbörsen und andere soziale Massenmedien verfolgen diesen Ansatz noch viel radikaler als Google.

Vom Datenschatten zum synthetischen Profil

Für diese Form der Welterzeugung sind nicht nur detaillierte Informationen über die externe Welt, also die von allen geteilte Realität, nötig, sondern auch über die je eigene Beziehung der Einzelnen zu eben dieser Welt.¹⁰⁸ Zu diesem Zwecke werden von jedem Nutzer Profile angelegt, je ausführlicher sie sind, desto besser für die Algorithmen. Ein Profil, wie es etwa (aber nicht nur) Google erstellt, erfasst den Nutzer auf drei Ebenen: als »Wissensperson«, die sich über die Welt informiert (dazu

¹⁰⁸ Es sind also nicht nur die Erfordernisse der Werbung, die die Erfassung persönlicher Daten vorantreiben, sondern auch die Entwicklung personalisierender Algorithmen zur Ordnung der Datenflut. Dies lässt vermuten, dass die Erfassung persönlicher Daten nicht zurückgehen wird, nicht einmal dann, wenn die kommerziellen Erfordernisse sich ändern sollten, etwa durch neue, nicht auf Werbung beruhende Geschäftsmodelle.

gehört zum Beispiel das Aufzeichnen der Suchanfragen, des Surfverhaltens etc.), als »physische Person«, die sich in der Welt befindet und sich in ihr bewegt (dazu gehört zum Beispiel die Ortung über das Smartphone, Sensoren im Smart Home oder die Erfassung von Körpersignalen) und als »soziale Person«, die mit anderen Menschen interagiert (dazu gehört zum Beispiel das Verfolgen der Aktivitäten in den sozialen Massenmedien).¹⁰⁹

Anders als in den neunziger Jahren sind diese Profile aber keine einfachen Abbildungen der Einzelnen, keine *digital personas* oder *data shadows* mehr. Sie repräsentieren nicht mehr das, was konventionellerweise als »Individuum«, im Sinn einer in räumlicher und zeitlicher Hinsicht einheitlichen Identität, bezeichnet wird. Die Profile bestehen vielmehr einerseits aus subindividuellen Elementen, jenen Fragmenten des erfassten Verhaltens, die sich für eine bestimmte Anfrage auswerten lassen, ohne dabei den Anspruch zu erheben, eine Person als Ganze zu repräsentieren, und andererseits aus dynamischen Clustern von mehreren Personen, so dass die zu modellierende Person gleichzeitig verschiedene Positionen in der Zeit einnehmen kann. Aufgrund dieser zeitlichen Spreizung werden Vorhersagen wie diese möglich: Eine Person, die bisher x getan hat, wird mit einer Wahrscheinlichkeit y die Handlung z ausführen. Amazon stellt auf diese Weise seine Buchempfehlungen zusammen, denn der Onlinehändler weiß, dass innerhalb des personalen Clusters, der einen Teil jedes persönlichen Pro-

109 Für eine detaillierte Darstellung, wie diese drei Ebenen erfasst werden, siehe Stalder, Felix und Christine Mayer (2009), »Der zweite Index. Suchmaschinen, Personalisierung und Überwachung«, in: Becker, Konrad und Felix Stalder (Hg.), *Deep Search. Die Politik des Suchens jenseits von Google*, Innsbruck: Studienverlag, S. 112-131.

files ausmacht, ein bestimmter Prozentsatz der Personen diese Sequenz von Handlungen bereits durchschritten hat. Oder wie es die *Data-mining*-Firma Science Rockstars (!) auf ihrer Website markig ausdrückt: »Ihre nächste Handlung ist eine Funktion des Verhaltens von anderen und Ihrer eigenen Vergangenheit.«

Google und andere Anbieter von Algorithmen erstellter Ordnungen setzen verstärkt auf diese Prognosefähigkeit ihrer Programme, um den verwirrenden und potenziell zeitaufwendigen Schritt der Suche obsolet zu machen. Das Ziel besteht darin, den im Akt des Suchens zum Ausdruck kommenden Bruch zwischen der Welt, wie sie jede einzelne Person erlebt – von Unsicherheit geprägt, denn Suchen impliziert, etwas »nicht zu wissen« –, und der von Algorithmen erstellten Ordnung – in der Gewissheit herrscht, weil schon alles wohlgeordnet ausgelegt ist – zu minimieren. Idealerweise sollen Fragen beantwortet werden, bevor sie gestellt werden. Ein erster Versuch aus dem Hause Google, diesen Bruch zum Verschwinden zu bringen, heißt Google Now und firmiert unter dem Slogan »Die richtigen Informationen zur richtigen Zeit«. Die ursprünglich als App entwickelte Anwendung, die inzwischen auch als Erweiterung für Googles eigenen Browser Chrome verfügbar ist, versucht, auf der Grundlage vorhandener Daten die nächsten Schritte eines Nutzers zu antizipieren und ihm die notwendigen Informationen bereitzustellen, bevor er sie sucht, um diese Schritte möglichst effizient erledigen zu können. So werden beispielsweise Daten aus dem Kalender herangezogen, um herauszubekommen, wohin sich ein User als Nächstes bewegen wird. Auf Basis von Echtzeitverkehrsdaten wird ihm dann der optimale Weg dorthin vorgeschlagen. Für Autofahrer wird die Verkehrsdichte in die Kalkulation aufgenommen. Sie wird mithilfe von Bewegungsprofilen an-

derer Autofahrer erhoben, die erkennen lassen, ob der Verkehr fließt oder ob mit Stau zu rechnen ist. Auf der Grundlage historischer Verkehrsdaten, so die Vision, sollen Autofahrer so umgelenkt werden können, dass Staus gar nicht erst entstehen.¹¹⁰ Für die Nutzer öffentlicher Verkehrsmittel wertet Google Now die Echtzeitdaten der örtlichen Verkehrsbetriebe aus. Davon ausgehend wird die optimale Route vorgeschlagen, und abhängig von der errechneten Wegzeit wird mal etwas früher, mal etwas später eine Erinnerung angezeigt, dass es nun wirklich Zeit sei, aufzubrechen. Was Google noch experimentell und auf einen eindeutig erkennbaren Kontext beschränkt testet, ist bei Facebook schon Alltag. Hier arbeiten die Ordnungsoperationen unter dem Namen EdgeRank bereits ohne jede explizite Nutzerinteraktion vollständig im Hintergrund. Auf der Grundlage der drei Variablen Affinität (bisherige Interaktionen zwischen zwei Nutzern), Gewichtung (Interaktionsrate aller Nutzer in Bezug auf einen einzelnen Beitrag) und Aktualität (Alter des Postings) selektiert der Algorithmus die Inhalte aus dem Strom der Statusupdates des eigenen Freundeskreises, die den Nutzern jeweils angezeigt werden.¹¹¹ Auf diese Weise will Facebook den Update-Strom auf der Nachrichtenseite der Mitglieder angenehm übersichtlich gestalten und, so darf man vermuten, genug Platz für Werbung schaffen. Auf das darin enthaltene Potenzial zur Manipulation durch im Hintergrund wirkende Algorithmen komme ich im nächsten Kapitel zu sprechen.

110 Das eröffnet die Frage, welche Autofahrer auf einen Umweg geschickt werden, damit kein Stau entsteht, und welche doch den direkten Weg – nun staufrei – angezeigt bekommen.

111 Vaughan, Pamela (2013), »Demystifying how Facebook's EdgeRank algorithm works«, hubspot.com, 23. 04. (online).

Variablen und Korrelationen

Jeder komplexe Algorithmus enthält eine Vielzahl von Variablen und eine meist noch größere Zahl von Verknüpfungsmöglichkeiten. Jede Variable und jede Relation, auch wenn sie noch so technisch oder mathematisch ausgedrückt ist, kodifiziert Annahmen, die eine spezifische Position in der Welt ausdrücken. Es kann keine rein beschreibenden Variablen geben, ähnlich wie es keine »Rohdaten« geben kann.¹¹² Beides, Daten wie Variablen, sind immer schon »gekocht«, das heißt, sie wurden durch kulturelle Operationen erzeugt und in kulturellen Kategorien geformt.¹¹³ Mit jeder Nutzung der produzierten Daten, mit jeder Ausführung eines Algorithmus werden die darin eingebetteten Annahmen aktiviert, und die in ihnen enthaltenen Positionen wirken mit an der Welt, die der Algorithmus generiert und präsentiert.

Die frühe Version des PageRank-Algorithmus beruhte, wie bereits erwähnt, auf der im Grunde recht simplen Annahme, dass häufig verlinkte Inhalte relevanter sind als nur selten verlinkte und dass Links auf Seiten, die selbst häufig verlinkt sind, mehr Gewicht beigemessen werden sollte als solchen, die auf Seiten gefunden wurden, auf die selbst nur wenige Links verweisen. Die Substitution des qualitativen Kriteriums »Relevanz« durch das quantitative »Popularität« erwies sich nicht nur als ungemein praktikabel, sondern auch als äußerst folgen-

112 Gitelman, Lisa (Hg.) (2013), *»Raw data« is an Oxymoron*, Cambridge, MA/London: MIT Press.

113 Das Begriffspaar »roh« im Sinne von unbearbeitet und »gekocht« im Sinne von bearbeitet geht zurück auf den Anthropologen Claude Lévi-Strauss, der es einführte, um den Unterschied zwischen Natur und Kultur zu verdeutlichen. Siehe Lévi-Strauss, Claude (1976), *Das Roh und das Gekochte*, Frankfurt a. M.: Suhrkamp.

reich, denn Suchmaschinen beschreiben die Welt nicht nur, sie bringen sie auch hervor. Was auf den obersten Plätzen der Suchergebnisse gelistet ist, ist nicht einfach schon populär, sondern wird es beziehungsweise bleibt es. Ein Drittel der Nutzer klickt auf den ersten Suchtreffer, und rund 95 Prozent schauen sich nicht mehr als die ersten zehn Resultate an.¹¹⁴ Sogar die früheste Version des PageRank-Algorithmus lieferte also kein Abbild der bestehenden Realität, sondern konstituierte diese mit.

Popularität ist aber nicht das einzige Element, mit dem Algorithmen die Welt der User aktiv gestalten. Eine Suchmaschine kann immer nur jenen Teil der Informationen sortieren, gewichten und zugänglich machen, den sie zuvor in ihren Index aufgenommen hat. Alles andere bleibt unsichtbar. Das Verhältnis zwischen dem erfassten Teil des Internets (*surface web*) und dem nichterfassten (*deep web*) ist kaum zu bestimmen. Schätzungen schwanken zwischen 1:5 und 1:500.¹¹⁵ Es gibt zahlreiche Gründe, warum Inhalte für Suchmaschinen unerreichbar sind. Vielleicht sind die Informationen in Formaten gespeichert, die Suchmaschinen nicht oder nur schlecht lesen können, oder sie werden bewusst hinter proprietären Schranken, etwa Bezahlssystemen, versteckt. Um den Bereich des von

114 Lee, Jessica (2013), »No. 1 position in Google gets 33% of search traffic«, searchenginewatch.com, 20.06. (online).

115 Eine immer noch oft zitierte Schätzung ist bereits veraltet: Bergman, Michael K. (2001), »White paper: The deep web: surfacing hidden value«, in: *The Journal of Electronic Publishing* 7/1 (online). Sie kommt auf ein Verhältnis von 1:500. Je mehr Inhalte dynamisch aus Datenbanken generiert werden, desto fragwürdiger werden solche Schätzungen. Unbestritten ist hingegen, dass nur ein kleiner Teil der online verfügbaren Information von Suchmaschinen erfasst wird.

Algorithmen Verwertbaren zu erweitern, geben die Betreiber von Suchmaschinen umfangreiche Hinweise, wie Anbieter ihre Seite gestalten sollten, damit die Suchwerkzeuge sie optimal finden können. Diesen Hinweisen muss man nicht folgen, aber wenn man die zentrale Rolle der Suchmaschinen als Sortierungs- und Filterungsinstanzen berücksichtigt, dann ist ersichtlich, dass hier Macht durch Standardsetzung ausgeübt wird.¹¹⁶

Dass der Einzelne sich »freiwillig« dieser Macht unterordnet, ist typisch für Netzwerkmacht, die keine Anweisungen gibt, sondern Voraussetzungen konstituiert. Aber es ist im Eigeninteresse (fast) jedes Informationsproduzenten, im Index der Suchmaschinen optimal verzeichnet zu sein, und entsprechend besteht ein starker Anreiz, die Voraussetzungen zu akzeptieren. Berücksichtigt man noch zusätzlich den Quasimonopolcharakter vieler Anbieter von Algorithmen erstellter Ordnungen und den hohen Preis, den man entrichten muss, wenn die eigene Seite für andere kaum oder gar nicht sichtbar ist, bekommt der Ausdruck »Freiwilligkeit« einen fahlen Beigeschmack. Hier wird also auf mehr oder weniger subtile Weise die Welt so vorformatiert, dass sie von Algorithmen optimal erfasst werden kann.¹¹⁷

116 Röhle, Theo (2009), »Die Demontage der Gatekeeper. Relationale Perspektiven zur Macht der Suchmaschinen«, in: Becker, Konrad und Felix Stalder (Hg.), *Deep Search. Die Politik des Suchens jenseits von Google*, Innsbruck: Studienverlag, S. 133-148.

117 Das Phänomen, dass die Welt zur Erfassung durch Algorithmen hergerichtet wird, ist nicht nur auf die digitalen Netze beschränkt. So wurde in Deutschland bereits 1994 eine neue Schriftart (die fälschungerschwerende Schrift, kurz: FE-Schrift) für Autokennzeichen eingeführt, die auch deren Maschinenlesbarkeit erhöht und die automatische Verkehrskontrolle erleichtert. Dem menschlichen

Die Anbieter von Suchmaschinen rechtfertigen solche Verfahren meist mit dem Hinweis, dadurch »effizientere« Dienste und »relevantere« Resultate anbieten zu können. Die vermeintlich technisch-neutralen Begriffe wie »Effizienz« und »Relevanz« können den politischen Charakter der Definition von Variablen aber nur notdürftig verdecken. Effizient im Hinblick worauf? Relevant für wen? Dies sind Fragen, die meist ohne viel Diskussion von den Entwicklern und Institutionen, welche die Algorithmen als ihr Eigentum reklamieren, entschieden werden. Hin und wieder entzündeten sich jedoch öffentliche Auseinandersetzungen um sie, meist dann, wenn die Interessen eines Anbieters mit denen der Konkurrenz kollidieren. So argumentierte etwa die Initiative FairSearch,¹¹⁸ Google missbrauche seine Marktmacht als Suchmaschine dazu, seine eigenen Inhalte privilegiert zu bewerten und in den Suchresultaten entsprechend prominent zu platzieren. Als Beispiel führten die Vertreter von FairSearch unter anderem an, dass Google bei Adresssuchen vornehmlich den eigenen Kartendienst und bei Produktsuchen das eigene Preisvergleichsangebot anzeige. Das Argument zeigte Wirkung: Im November 2010 leitete die Europäische Kommission ein kartellrechtliches Prüfverfahren gegen Google ein. 2014 wäre es fast zu einer Einigung gekommen, die dem US-Internetgiganten einige Zugeständnisse abgetrotzt hätte, doch Mitglieder der Kommission, das EU-Parlament und Verbraucherschützer sprachen sich gegen eine solche Einigung aus. Im April

Auge erscheint sie hingegen, so der hilfreiche Wikipedia-Eintrag zur FE-Schrift, »seltsam proportioniert und unförmig«.

118 Fairsearch.org wird offiziell von vielen Konkurrenten von Google unterstützt, darunter Microsoft, Tripadvisor und Oracle.

2015 wurde das Kartellverfahren durch die inzwischen neugewählte Kommission wiederaufgenommen. Zur Begründung hieß es, dass Google »das Sanktionssystem, das es auf der Grundlage bestimmter Parameter auf andere Preisvergleichsdienste anwendet, nicht auf seinen eigenen Preisvergleichsdienst an[wendet], was dazu führen kann, dass sie auf den allgemeinen Suchergebnisseiten von Google auf einem niedrigeren Rang erscheinen«. ¹¹⁹ Mit anderen Worten, die Kommission wirft dem Unternehmen vor, die Suchergebnisse zum eigenen Vorteil und zum Nachteil der Nutzer zu manipulieren.

Das ist aber nicht der einzige Fall, durch den die politische Seite der Suchalgorithmen an die Öffentlichkeit kam. Im Sommer 2012 kündigte Google an, dass Webangebote, gegen die besonders häufig Löschanträge wegen Urheberrechtsverletzungen eingereicht werden, in den Ergebnissen zukünftig weiter hinten eingereiht würden. ¹²⁰ Das Unternehmen führte also explizit politisch-ökonomische Kriterien ein, um zu beeinflussen, was die Nutzer nach Maßgabe einiger mächtiger Akteure (zum Beispiel aus der Filmbranche) sehen. ¹²¹ Auch hier könnte man von einer Personalisierung der Suche sprechen, nur dass nicht mehr die natürliche Person des Nutzers, sondern die juristische Person des Rechteinhabers im Zentrum steht. Nach deren Interessen und Vorlieben wird die Suche

119 Europäische Kommission (2015), »Kartellrecht: Kommission übermittelt Google Mitteilung der Beschwerdepunkte zu seinem Preisvergleichsdienst«, Brüssel, 15.04. (online).

120 Singhal, Amit (2012), »An update to our search algorithms«, inside-search.blogspot.co.uk, 10.08. (online). Mitte 2014 erhielt Google laut eigenen Angaben rund 20 Millionen Anfragen, Links wegen angeblicher Urheberrechtsverletzungen aus dem Index zu entfernen.

121 Wragge, Alexander (2012), »Google-Ranking: Herabstufung ist ›Zensur light‹«, irights.info, 23.08. (online).

ausgerichtet. Amazon bediente sich ähnlicher Methoden. 2014 veränderte der Onlinehändler seinen vielgepriesenen Empfehlungsalgorithmus gezielt dahingehend, dass Bücher missliebiger Verlage, mit denen er sich gerade in Preisverhandlungen befand, seltener angezeigt wurden.¹²²

Kontroversen um das Vorgehen von Amazon oder Google sind jedoch eher die Ausnahme als die Regel. Fast immer finden die unabdingbaren – aber niemals neutralen – Setzungen in Bezug auf Erfassung und Auswertung der Daten durch die Algorithmen abseits jeder Diskussion statt. Die Logik des ursprünglichen PageRank-Algorithmus wurde schon im Jahr 2000 kritisiert, weil sie im Grunde einfach die kommerzielle Logik der Massenmedien abbilde und systematisch wenig populäre, aber vielleicht anderweitig relevante Informationen benachteilige und so die »Vision des Internets als inklusiven, demokratischen Ort untergrabe«.¹²³ Die seither vorgenommenen Veränderungen des Suchalgorithmus haben diese Tendenz zwar modifiziert, aber nicht abgeschwächt. Zusätzlich zur Konzentration auf das, was populär ist, privilegieren die neuen Variablen kürzlich hochgeladene und sich ständig aktualisierende Inhalte. Die Auswahl der Suchergebnisse orientiert sich nun am Ort des Nutzers und berücksichtigt seine soziale Vernetzung. Sie wird auf den Durchschnitt einer dynamisch modellierten Gruppe ausgerichtet. Mit anderen Worten: Der neue Google-Algorithmus favorisiert das, was im sozialen

122 Manjoo, Farhad (2014), »Amazon's tactics confirm its critics' worst suspicions«, in: *The New York Times*, Bits Blog, 23.05. (online).

123 Introna, Lucas D. und Helen Nissenbaum (2000), »Shaping the web: Why the politics of search engines matters«, in: *The Information Society* 16, S. 169-185.

Netzwerk eines Nutzers gerade an Popularität gewinnt. Die globalen Dörfer werden so immer dörflicher.¹²⁴

Daten-Behaviorismus

Damit wiederholen und verstärken Algorithmen wie der von Google eine Tendenz, die bereits auf der Ebene des einzelnen Nutzers wie auch auf der der gemeinschaftlichen Formationen sichtbar wurde: Um mit der Masse und Komplexität der Informationen umzugehen, wird der Blick nach innen gerichtet, womit allerdings auch hier nicht das Innere einzelner Personen gemeint ist. Der einzelne Mensch mit einer Innenwelt, mit Ideen, Träumen und Wünschen ist als Referenzgröße nicht relevant. Für die Algorithmen sind Menschen Black Boxes, die rein über ihre messbaren Reiz-Reaktions-Beziehungen erfasst werden. Bewusstsein, Wahrnehmung oder Intention spielen für sie keine Rolle. Die Rechtsphilosophin Antoinette Rouvroy spricht in diesem Zusammenhang von »Daten-Behaviorismus«.¹²⁵ Sie weist damit auf die schleichende Rückkehr eines lange Zeit als diskreditiert geltenden Ansatzes in der Verhaltenspsychologie hin, der postulierte, dass man das menschliche Verhalten rein über von außen beobachtbares und messbares Verhalten erklären, vorhersagen und steu-

124 Pariser, Eli (2012), *Filter Bubble. Wie wir im Internet entmündigt werden*, München.

125 Rouvroy, Antoinette (2013), »The end(s) of critique: data-behaviourism vs. due-process«, in: Vries, Katja de und Mireille Hildebrandt (Hg.), *Privacy, Due Process and the Computational Turn: The Philosophy of Law Meets the Philosophy of Technology*, Abingdon/Oxon/New York: Routledge, S. 143-165.

ern könne.¹²⁶ Psychologische Dimensionen wurden (und werden auch in der Neuauflage des Behaviorismus) ignoriert, weil sie schlecht empirisch beobachtbar sind. Entsprechend entfällt auch die Notwendigkeit, Personen direkt zu befragen oder deren subjektives Erleben, Denken und Empfinden in Betracht zu ziehen. Menschen galten damals (und gelten heute wieder) als unzuverlässig, schlecht in ihrer Selbstwahrnehmung und nur bedingt wahrheitsgetreu in ihrer Auskunft. Deshalb fordert das Gebot einer streng empirischen Wissenschaft, alles, was zu keiner physischen, beobachtbaren Aktion führt, als irrelevant zu erachten. Aus dieser Perspektive lässt sich sogar komplexes Verhalten in Reiz-Reaktions-Einheiten zerlegen. Das führt zu der Einschätzung, dass ein Beobachter über einen Handelnden immer mehr weiß als dieser über sich selbst, denn anders als der Handelnde, dessen Eindrücke unzutreffend sein können, verfügt der Beobachter über objektive und vollständige Informationen. Schon früh wurde Kritik an diesem Ansatz laut. Er sei mechanistisch, reduktionistisch und autoritär, weil er den beobachtenden Wissenschaftler gegenüber dem Subjekt privilegiere. In der Praxis stieß er schnell an seine Grenzen: Es war schlicht und ergreifend zu aufwendig und zu kompliziert, detaillierte Daten über menschliches Verhalten zu erheben.

Doch das hat sich in den letzten Jahren radikal geändert. Mittlerweile lassen sich immer mehr Handlungen, Zustände

126 »Wir möchten das Verhalten des einzelnen Organismus vorhersagen und kontrollieren. Das ist unsere ›abhängige Variable‹ – die Wirkung, für die wir eine Ursache finden müssen. Unsere ›unabhängige Variable‹ – die Ursachen des Verhaltens – sind die äußeren Bedingungen, von denen Verhalten eine Funktion ist« (Skinner, Burrhus Frederic [1973 {1953}], *Wissenschaft und menschliches Verhalten*, München: Kindler, S. 42).

und Zusammenhänge empirisch messen. Algorithmen wie die von Google oder Amazon bilden den technischen Hintergrund für das Revival eines nach wie vor mechanistischen, reduktionistischen und autoritären Ansatzes, mit dem der alte, lange für überwunden gehaltene Traum eines objektiven Blicks, des Blicks von nirgendwo, wiederbelebt wird.¹²⁷ Jede Kritik an dieser positivistischen Perspektive – etwa dass jedes Messresultat nicht nur das Gemessene, sondern auch den Messenden widerspiegelt – wird mit dem Verweis auf die schiere Menge an Daten beiseitegewischt.¹²⁸ Diese Haltung untermauert den Anspruch derjenigen, die über die neuen und umfassenden Beobachtungskapazitäten verfügen – dazu gehören neben Google und Facebook auch die Geheimdienste westlicher Länder –, einzelne Personen besser zu kennen als diese sich selbst und so ihre Fragen beantworten zu können, bevor sie diese stellen. Ein Ziel, das Google immer wieder anzustreben betont.

Bei der »Wendung nach innen« geht es also um den Raum der »gemeinschaftlichen Formation«, konstituiert durch die Summe aller Handlungen der interagierenden Akteure. Allerdings wird eine gemeinschaftliche Formation nicht bewusst in einem horizontalen Prozess geschaffen und erhalten, sondern synthetisch als eine rechnerische Funktion konstruiert, zu der Einzelne, je nach Kontext und Notwendigkeit, zugeordnet oder aus der sie entfernt werden können. All dies geschieht hinter dem Rücken der Nutzer, angereichert mit den Zielen und Po-

127 Jurgenson, Nathan (2014), »View from nowhere. On the ideology of big data«, *thenewinquiry.com*, 09.10. (online).

128 boyd, danah und Kate Crawford (2012), »Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon«, in: *Information, Communication & Society* 15/5, S. 662-679.

sitionen, die für die Entwickler des jeweiligen Algorithmus relevant sind, sei es Profit- oder Überwachungsoptimierung, soziale Normierung, Serviceverbesserung oder anderes. Die so generierten Resultate werden den Usern als personalisierter Service, als effiziente Dienstleistung, die quasimagische Resultate produziert, verkauft. Aus dem gigantischen Heuhaufen durchsuchbarer Informationen werden Ergebnisse generiert, die dabei gleich zu der gesuchten Nadel erklärt werden. Wie diese Resultate zustande gekommen sind, welche Positionen in der Welt damit gestärkt beziehungsweise geschwächt werden, ist im besten Fall nur ansatzweise nachvollziehbar. Aber solange die Nadel einigermaßen funktionstüchtig ist, sind die meisten Nutzer zufrieden, und der Algorithmus registriert diese Zufriedenheit, um sich selbst zu validieren. In der dynamischen Welt der Unübersichtlichkeit sind Nutzer geleitet von einem radikalen, kurzfristigen Pragmatismus. Sie lassen sich die Welt gerne vorsortieren, um besser in ihr handeln zu können. Ein angemessenes Urteil im Hinblick auf die Frage, ob die gelieferten Informationen die Welt richtig oder falsch repräsentieren, können sie sich ohnehin nicht bilden, denn diese ist ohne Hilfsmittel letztlich unfassbar. Außer in rasch schrumpfenden Domänen des Spezial- oder Alltagswissens kann die Welt immer weniger ohne vorsortierende Mechanismen überblickt werden. Die User können die Suchergebnisse einzig pragmatisch daraufhin beurteilen, ob sie helfen, ein konkretes Problem zu lösen. Dabei steht nicht die beste Lösung oder die richtige Antwort im Vordergrund, sondern eine, die verfügbar und gut genug ist. Das verleiht den Institutionen und Verfahren, die die Lösungen und Antworten liefern, einen enormen Einfluss.