



D@tenflut

Erlanger Universitätstage 2016

Herausgegeben von Rudolf Freiburg

FAU Forschungen, Reihe A, Geisteswissenschaften 9



D@tenflut

FAU Forschungen, Reihe A
Geisteswissenschaften

Band 9

Herausgeber der Reihe:
Wissenschaftlicher Beirat der FAU University Press

D@tenflut

Fünf Vorträge

herausgegeben von
Rudolf Freiburg

Erlangen
FAU University Press
2017

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im
Internet über <http://dnb.d-nb.de> abrufbar.

Das Werk, einschließlich seiner Teile, ist urheberrechtlich geschützt.
Die Rechte an allen Inhalten liegen bei ihren jeweiligen Autoren.
Sie sind nutzbar unter der Creative Commons Lizenz BY-NC-ND.

Der vollständige Inhalt des Buchs ist als PDF über den OPUS Server der
Friedrich-Alexander-Universität Erlangen-Nürnberg abrufbar:
<https://opus4.kobv.de/opus4-fau/home>

Verlag und Auslieferung:
FAU University Press, Universitätsstraße 4, 91054 Erlangen
<http://www.university-press.fau.de>
university-press@fau.de

Druck: Verlagsdruckerei SCHMIDT, Neustadt/Aisch]

ISBN: 978-3-96147-046-4 (Druckausgabe)
eISBN: 978-3-96147-047-1 (Online-Ausgabe)
ISSN: 2199-014X

Inhaltsverzeichnis

Vorwort	3
Einleitung: D@tenflut – Eine essayistische Einschätzung der digitalen Revolution.....	5
SEBASTIAN M. BÜTTNER Algorithmen, Datenkraken und gläserne Bürger: <i>Big Data</i> als Herausforderung für die Gesellschaftsanalyse	39
GEORG GLASZE Digitale Geographien	61
SVENJA HAGENHOFF „Außer Kontrolle“: Alte und neue Informationsfluten im Publikationswesen	77
KLAUS MEYER-WEGENER Erstmal einfach alles speichern: <i>Big Data</i> als Aufgabe für die Informatik	99
HANS-ULRICH PROKOSCH Der vermessene Mensch: Wissen wir, was wir alles über uns preisgeben?	117
Autoren- und Herausgeberverzeichnis.....	141
Index.....	145
Namensregister	153

Vorwort

Der vorliegende Band präsentiert die Druckfassungen der Vorträge zum Thema ‚D@tenflut‘, die anlässlich der Erlanger Universitätstage vom 16. Februar bis zum 15. März 2016 in Amberg und vom 20. Oktober bis zum 24. November 2016 in Ansbach gehalten wurden. Die fünf Erlanger Wissenschaftlerinnen und Wissenschaftler, Professor Dr. Svenja Hagenhoff, Professor Dr. Georg Glasze, Professor Dr. Klaus Meyer-Wegener, Professor Dr. Hans-Ulrich Prokosch und Dr. Sebastian Büttner – aus den Disziplinen der Buchwissenschaft, der Geographie, der Informatik, der Medizinischen Informatik und der Soziologie – führten in beeindruckender Weise die Chancen und Gefahren vor Augen, die das sich immer klarer abzeichnende ‚Zeitalter der großen Datenmengen‘ mit sich bringt. Mit ihren geschickt didaktisch aufbereiteten Beiträgen gelang es ihnen ausnahmslos, ein erfreulich interessiertes Publikum zu begeistern. Ihnen sei ganz besonders für ihr Engagement, aber auch für ihre Bereitschaft gedankt, die Vorträge für den Druck freizugeben.

Die Erlanger Universitätstage können nur deshalb jedes Jahr aufs Neue durchgeführt werden, weil sie durch zahlreiche Institutionen und Personen wertvolle Unterstützung erfahren. Wie in den Jahren zuvor gilt unser Dank an erster Stelle der Universitätsleitung der Friedrich-Alexander-Universität, insbesondere ihrem Präsidenten Professor Dr. Joachim Hornegger, der sich nicht nur intensiv dafür einsetzt, die schöne Tradition der Universitätstage fortzusetzen und für die notwendige finanzielle Unterstützung zu sorgen, sondern es sich auch nicht nehmen lässt, die Vortragsreihen mit persönlichen Grußworten zu eröffnen. Ein herzliches Dankeschön ergeht auch an die Adresse der MitorganisatorInnen in Amberg, besonders an Herrn Oberbürgermeister Michael Cerny, Herrn Bürgermeister Martin Preuß und Frau Bürgermeisterin Brigitte Netta, den Kulturreferenten Herrn Wolfgang Dersch, den Kulturamtsleiter Herrn Thomas Boss und Frau Barbara Cosima Frey. Für die Begleitung der Universitätstage in Ansbach ist vor allem Frau Oberbürgermeisterin Carda Seidel und Herrn Bürgermeister Martin Porzner herzlich zu danken. Herr Dr. Thomas Bauer, Regierungspräsident von Mittelfranken, nahm sich trotz seines dichten Terminkalenders auch im Jahre 2016 die Zeit, uns zu einem Pressegespräch bei Kaffee und Gebäck in seinen Amtsräumen willkommen zu heißen; in seiner Begrüßungsrede im Rahmen der Auftaktveranstaltung der Universitätstage in Ansbach, die im Jahre 2016 nach längerer Pause endlich wieder in der ‚Alten Bibliothek‘ im Ansbacher

Schloss stattfanden, wählte er zutreffende Einführungsworte für das aktuelle Thema, wenn er darauf hinwies, dass sich der ‚digitale Wandel nie wieder so langsam vollziehen werde, wie zur Zeit‘. Auch ihm sei herzlich für sein Engagement und sein Interesse an den Veranstaltungen gedankt.

Wie in den Jahren zuvor wird die Finanzierung des Bandes durch namhafte Beiträge der gastgebenden Städte gewährleistet, aber auch durch den Universitätsbund Erlangen-Nürnberg e.V., in Amberg durch Herrn Professor Dr. Volker Groß und in Ansbach durch Herrn Gerhard Neumeyer vertreten, denen allen aufrichtiger Dank gebührt. Zu großem Dank verpflichtet sind wir auch dem 1. Vorsitzendem des Universitätsbundes, Herrn Dr. Siegfried Balleis, der das Konzept der Erlanger Universitätstage tatkräftig unterstützt.

Für die Drucklegung des Bandes konnte ich mich wieder auf ein bewährtes Team verlassen; hier gilt mein Dank besonders meinen Sekretärinnen Barbara Cunningham und Evelin Werner sowie meiner Mitarbeiterin Susanne Wagner, die intensiv dazu beigetragen haben, die eingereichten Manuskripte der Beiträge in druckfertige Texte zu verwandeln. Für seine zuverlässigen Recherchearbeiten möchte ich mich bei Lukas Pracht bedanken. Ein herzliches Dankeschön ergeht schließlich an Herrn Markus Putnings von der Universitätsbibliothek Erlangen-Nürnberg, der uns die Aufnahme des Bandes in die Schriftenreihe der ‚FAU University Press‘ ermöglichte, und an Frau Eva Volleth, die uns während der Endphase der Drucklegung mit Rat und Tat zur Seite stand.

Erlangen, im Februar 2017

Rudolf Freiburg

Einleitung

D@tenflut:

Eine essayistische Einschätzung der digitalen Revolution

Rund um die Uhr erreichbar. Nirgendwo außer acht sein. Per Mausklick erfaßt. Rückläufig bis hin zum Babypuder verdatet. Nichts geht verloren. Alltägliche Gänge in Billigläden, ins Kino, aufs Klo sind unvergänglich. Auch unserer Liebe langatmiger Weg speichert ein Chip von Fingernagelgröße. Kein Versteck mehr. Immer im Blickfeld. Sogar der Schlaf behütet. Nie mehr allein.¹

Im November des Jahres 2016 wurden in zahlreichen renommierten deutschen Zeitungen – darunter die *ZEIT*, die *Süddeutsche Zeitung*, die *Welt*, und der *Tagesspiegel* – ganzseitige Anzeigen veröffentlicht, in denen die Forderung nach einer „Europäischen Charta der digitalen Grundrechte“ erhoben wird. Zu den Unterstützern der Charta zählen neben dem *ZEIT*-Chefredakteur Giovanni Di Lorenzo, dem ehemaligen Präsidenten des Europäischen Parlamentes Martin Schulz und dem Philosophen und Soziologen Jürgen Habermas auch der Ratsvorsitzende der Evangelischen Kirche in Deutschland Heinrich Bedford-Strohm sowie die Schriftstellerinnen Juli Zeh und Eva Menasse.² Die Forderung nach einem ‚digitalen Grundgesetz‘ für Europa war lange überfällig und reagiert auf die existenzielle Gefährdung des Einzelnen durch eine Revolution, die sich bislang weitgehend im Stillen – noch unbemerkt von der Mehrheit der Menschen – im Bereich der Datensammlung und -verarbeitung abspielt und bereits in naher Zukunft das Leben der gesamten Menschheit in entscheidender Weise verändern wird, eine Revolution, die ähnlich wie ihr französisches Vorbild im Jahr 1789 immer mehr in den Verdacht gerät, ihre eigenen

¹ Günter Grass, „Ohnmacht“, in: *Vonne Endlichkeit [sic]* (Göttingen: Steidl Verlag, 2015), 13; vgl. auch „Allabendliches Gebet“, ebd., 14.

² Vgl. Anonymous, „Für digitale Grundrechte! – Ein Bürgerrecht: Die *ZEIT*-Stiftung unterstützt eine Digitalcharta für Europa“. (*online*). (Für diese und alle folgenden Angaben von Internetquellen in den Fußnoten wird auf die vollständige Nennung im Literaturverzeichnis verwiesen.)

Kinder zu fressen'.³ Der euphemistisch wirkende Name ‚Datenflut‘ und die verharmlosenden Anglizismen ‚Big Data‘ oder ‚Deep Learning‘, die an die ideologischen Intentionen von George Orwells ‚Newspeak‘ erinnern, dürfen nicht darüber hinwegtäuschen, dass die digitale Revolution einen Einschnitt im Leben aller Menschen darstellen wird. Der Katalog der existenziellen Kränkungen der Menschheit, von denen Sigmund Freud zu Beginn des 20. Jahrhunderts berichtet hatte, wird durch einen folgenreichen Eintrag erweitert. Freud hatte behauptet, dass der Mensch durch den Verlust des geozentrischen Weltbildes einen Teil seines Selbstbewusstseins eingebüßt habe. Durch die Erkenntnisse des Darwinismus habe er dann in der Folge auf extrem demütigende Weise zu der Einsicht gelangen müssen, dass er vom Affen abstamme; und die Tiefenpsychologie schließlich habe bei ihm einen Erkenntnisschock ausgelöst, von dem er sich lange nicht erholen werde: Angesichts der komplexen Interdependenz von *id*, *ego* und *super-ego* sei er nicht länger ‚Herr im Hause seines eigenen Bewusstseins‘.⁴ Neben den von Freud diagnostizierten kosmischen, biologischen und psychologischen Traumata umfasst der erweiterte Katalog der Kränkungen seit Beginn des 21. Jahrhunderts zweifellos nunmehr auch die digitale Erschütterung des menschlichen Selbstbewusstseins. Die sich in der Gegenwart eindeutig abzeichnende Datenflut offenbart bereits jetzt ein erschreckend klares destruktives Potenzial,⁵ das sich zu apokalyptisch wirkenden Ausmaßen steigert, wenn man die aktuellen Entwicklungen in die Zukunft extrapoliert. Der Schock, der der Menschheit bevorsteht, falls die aktuelle Datenflut nicht durch moralische, ethische und juristische Regelungen eingedämmt wird, ist derjenige einer radikalen Überwachung und Marginalisierung des Einzelnen, der von jetzt ab dazu verurteilt sein könnte, im Reich der maschinengesteuerten Lebenswelt ein Nischendasein zu führen oder gar völlig vor der Übermacht der künstlichen Intelligenz und ihrer Auswirkungen zu kapitulieren. Es mutet an, als könne sich die ungesteuerte Energie der Datenfluten zu einem Tsunami ausweiten, in

³ Das ‚geflügelte Wort‘ wird Pierre Vergniaud zugesprochen; vgl. *The Oxford Dictionary of Quotations*, ed. Angela Partington (London et al.: Oxford University Press, 1992 [1941]), 710.

⁴ Vgl. Sigmund Freud, „Eine Schwierigkeit der Psychoanalyse“, in: ders., *Abriß der Psychoanalyse: Einführende Darstellungen* (Frankfurt a.M.: Fischer Taschenbuch Verlag, 1998 [1917]), 187-194.

⁵ Zu diesen Gedanken vgl. auch Joseph Schumpeters Konzept der ‚kreativen Zerstörung‘ in Yvonne Hofstetters Buch *Das Ende der Demokratie: Wie die künstliche Intelligenz die Politik übernimmt und uns entmündigt* (München: C. Bertelsmann, 2016), 60-61.

dem das alttestamentarische Bild der Sintflut als digital-atheistische *Nemesis Divina* fröhliche Urständ feiert. Autonom steuernde Autos, Industrie 4.0, das ‚Internet der Dinge‘ und Superrechner, die in der Lage sind, selbstständig weitreichende Entscheidungen zu treffen, reduzieren die Aufgabe des Menschen bestenfalls auf diejenige eines passiven Nutznießers von Produkten, die er kaum noch verstehen, geschweige denn völlig kontrollieren kann:

Wir befinden uns mitten in einem riesigen Infrastrukturprojekt, das in mancher Hinsicht denen der Vergangenheit ebenbürtig ist, von den römischen Aquädukten bis zur *Encyclopédie* der französischen Aufklärung. Aber wir haben die Bedeutsamkeit dieses Projektes noch nicht vollständig erkannt, weil es noch so neu ist, weil wir noch mittendrin stecken, und weil das Ergebnis, anders als das Wasser, das über Aquädukte fließt, sich nicht einfach fassen lässt. Dieses Projekt ist die Datafizierung. So wie die früheren Fortschritte unserer Infrastruktur wird auch die Datafizierung die Gesellschaft grundlegend verändern.⁶

Erneut offenbart die europäische Aufklärung, die auf den Prinzipien der Rationalität beruht, ihr dialektisches Potenzial, auf das Theodor W. Adorno und Max Horkheimer in ihren Arbeiten hingewiesen haben;⁷ der Siegeszug der Vernunft, der in der bahnbrechenden Schrift Sir Francis Bacons im 17. Jahrhundert noch mit dem Ziel, dem ‚Wohle der Menschheit‘ zu dienen, assoziiert worden war,⁸ kennt nun keinerlei Grenzen mehr; hatte Bacon mit Recht dazu aufgefordert, das bekannte Terrain des Wissens, das emblematisch durch die legendären Säulen des Herakles begrenzt worden war, hinter sich zu lassen und zu neuen Welten aufzubrechen, so entpuppt sich jene *terra incognita*, an deren Ufer uns die Datenfluten spülen, immer mehr als menschenfeindliches Land, als Terrain des Posthumanen. Auch Bacons Vorstellungen vom menschlichen Vermögen, seine einflussreiche Fakultätenanalyse,⁹ erfahren eine drastische Neuauslegung: Die neue *ratio* könnte nicht mehr Teil eines menschlichen, sondern eines künstlichen Gehirns sein; die neue *memoria*

⁶ Viktor Mayer-Schönberger und Kenneth Cukier, *Big Data: Die Revolution, die unser Leben verändern wird*, übers. Dagmar Mallett (München: Redline Verlag, 2013), 123.

⁷ Vgl. Max Horkheimer und Theodor W. Adorno, *Dialektik der Aufklärung: Philosophische Fragmente* (Frankfurt: Fischer Verlag, 2012 [1944]).

⁸ Sir Francis Bacon, „The Advancement of Learning“ (1605), *The Advancement of Learning and New Atlantis*, ed. Arthur Johnston (Oxford: Clarendon Press, 1974), 3-200; 36.

⁹ Vgl. ebd., 116-118.

versteht sich nicht länger als labiler Speicherort für ephemere ‚Eindrücke‘, wie die traditionsreiche Metaphorik von der ‚Wachstafel‘ (*tabula rasa*) nahelegt, sondern offenbart sich als Hypermnese gigantischer Datenbanken, die kein Vergessen mehr dulden, und selbst die *imaginatio*, auf die der Mensch so lange stolz war,¹⁰ wird der sterilen Welt sich selbst generierender Algorithmen übereignet.¹¹ Während zu Beginn der Aufklärung eine philanthropische Emanzipation stand, die Immanuel Kant mit dem Appell verband, man möge sich von den Fesseln der Tradition befreien, indem man sich seines eigenen Verstandes bediene, um aus der ‚selbstverschuldeten Unmündigkeit auszutreten‘,¹² zeichnet sich im 21. Jahrhundert eine Entwicklung ab, in der das wohl grandioseste Instrument rationaler Lebensbewältigung, der Computer,¹³ sich anschickt, jene emanzipatorische Entwicklung, als deren spätes Produkt er ja begriffen werden kann und muss, zu stornieren.¹⁴ Wer sich heute eines Netzwerk-Rechners bedient, schwebt permanent in der Gefahr, in eine neue ‚Unmündigkeit‘ zu geraten, die er in gewisser Weise selbst verschuldet hat, etwa wenn er gedankenlos nicht gelesene Geschäftsbedingungen von Großkonzernen unterschreibt.¹⁵ Das Thema der Datenfluten kann nicht länger das Sujet von Informatikern, Mathematikern und Statistikern sein;

¹⁰ Auch der Bereich der Kunst kann sich nicht gegen die Invasion digitaler Techniken wehren. Als Beispiel möge die Vorstellung des Klavier-Roboters Teo Tronico dienen, der am 26. August 2012 zusammen mit den Berliner Philharmonikern musizierte; vgl. ‚Robot-Pianist Teo Tronico ‚live‘ at Berlin Philharmonie“. (*online*).

¹¹ Selbst der menschliche ‚Traum‘ ist Gegenstand von Algorithmen geworden; vgl. den Hinweis auf ‚träumende Netze‘ in Christoph Drösser, *Total Berechenbar: Wenn Algorithmen für uns entscheiden* (München: Hanser, 2016), 210.

¹² Vgl. Immanuel Kant, *Beantwortung der Frage: Was ist Aufklärung?* (Frankfurt a.M.: Neuer Frankfurter Verlag, 1900).

¹³ Der Computer stellt den vorläufigen Höhepunkt des Bemühens, die Welt zu vermessen, dar; zur spannenden Geschichte des Versuchs, die ‚Wirklichkeit‘ mit den Augen der Mathematik zu sehen, vgl. Klaus Mainzer, *Die Berechnung der Welt: Von der Weltformel zu Big Data* (München: Beck, 2014); beginnend mit Platons Vorstellungen von einer Welt geometrischer Formen verfolgt Mainzer Ähnlichkeiten und Unterschiede analytischer Ordnungssysteme bei Pierre-Simon Laplace, Gottfried Wilhelm Leibniz, Kurt Gödel, David Hilbert, Konrad Zuse und Richard Feynman.

¹⁴ In seiner Analyse der unterschiedlichen Angriffe auf die Freiheit in der Gegenwart bezeichnet Harald Welzer dieses Verhalten der „Selbstentmündigung“ als „Signatur der Epoche“; vgl. Harald Welzer, *Die smarte Diktatur: Der Angriff auf unsere Freiheit* (Frankfurt a.M.: Fischer Verlag, 2016), 116-127.

¹⁵ Vgl. auch den Werbetext zu Hofstetters Buch *Das Ende der Demokratie*: „Yvonne Hofstetter warnt: Die Rückkehr in eine selbst verschuldete Unmündigkeit hat begonnen; dass sie in smartem selbst-optimierendem Gewand daherkommt, mag man als Ironie der Menschheitsgeschichte betrachten“.

der einzelne Bürger ist genötigt, um die Kontrolle über seine Daten zu kämpfen.¹⁶ Die riesigen Datenmengen infiltrieren alle Lebensbereiche und verändern sie in fundamentaler Weise. Selbst derjenige, der sich entscheidet, kein Smartphone zu benutzen und weitgehend auf die Verwendung von Computern zu verzichten, kann der digitalen Magie nicht enttrinnen. In diesem Essay soll untersucht werden, wie sich die Welt der *Big Data* auf Epistemologie und Hermeneutik, die Psychologie, die Soziologie, die Ontologie und – nicht zuletzt – die Ethik auswirkt.

Erkenntnis des Bekannten: Epistemologie und Hermeneutik im Zeichen von *Big Data*

Die menschliche Neugierde,¹⁷ als deren Produkt auch die Computer und Großrechner der Gegenwart bezeichnet werden können, folgt dem Drang, die Welt zu erkennen und zwischen Sein und Schein zu unterscheiden; Wahrnehmung und Wissen stehen dabei in einem interdependenten, nicht immer fruchtbar zu nennenden Zusammenhang. Bereits in der antiken Philosophie galt das Interesse Platos und Aristoteles' dem Versuch, die Bedingungen des Erkennens zu identifizieren und zu optimieren. Platos Anliegen, die Welt der Ideen als eigentlich Seiendes zu definieren und der erfahrbaren ‚Realität‘ sowie der menschlichen Existenz den Status von Schattenhaftigkeit zuzuweisen, lässt sich nicht zuletzt auf den Wunsch zurückführen, man möge nicht länger durch irriige Annahmen getäuscht werden. Aristoteles' Analysetechniken im *Organon* (4. Jh. v. Chr.) stehen mit ihrer Tendenz zur Kategorienbildung im Dienste des Bemühens, die ‚Instrumente‘ der menschlichen Erkenntnis zu justieren und Fehlannahmen bei der Urteilsbildung zu minimieren. In der Zeit der Aufklärung wurde die Erkenntnistheorie zu einem wichtigen Instrument, das den Siegeszug der Vernunft vorbereitete. Sie widmete sich in dezidiertester Weise der Aufgabe, Irrtümer zu eliminieren. In seiner Schrift *Novum Organum Scientiarum* (1620) entwickelte Bacon seine Idolenlehre, in der er erstmals auf das Verfälschungspotenzial aufmerksam machte, das in der Sprache (*idola foris*) und der (auf aristotelischen Grundannahmen

¹⁶ Vgl. Frank Schirrmacher, *Payback: Warum wir im Informationszeitalter gezwungen sind zu tun, was wir nicht tun wollen, und wie wir die Kontrolle über unser Denken zurückgewinnen* (München: Blessing, 2009).

¹⁷ Vgl. Hans Blumenberg, *Der Prozess der theoretischen Neugierde* (Frankfurt a.M.: Suhrkamp, 1973).

beruhenden) philosophischen Tradition (*idola theatri*) verborgen war;¹⁸ mehr als anderthalb Jahrhunderte später widmete sich Kant in seiner *Kritik der reinen Vernunft* (1781) dem Versuch, das menschliche Denken von den Störprozessen negativer Urteile zu bereinigen. Der Wunsch nach vorurteilslosem Erkennen und Verstehen wurde zu einer Grundbedingung für die Erfolgsgeschichte der Aufklärung, da nur eine von Fehlurteilen gereinigte Wahrnehmung der Welt dazu dienen konnte, diese auch im technischen und pragmatischen Sinne für die Zwecke des Menschen zu nutzen, wie dies bereits Bacon gefordert hatte.¹⁹ Die Weiterentwicklung der antiken Hermeneutik und mittelalterlichen Exegese, die vor allem durch Zielsetzungen theologischer Provenienz im 18. Jahrhundert intensive Impulse zur Vervollkommnung erfahren hatte, nahm diese Erkenntnisse der Epistemologie in ihr Programm auf und entwickelte Strategien zum ‚angemessenen‘, also *idealiter*, vorurteilsfreien Verstehen vornehmlich von Texten aus der Theologie, aber auch aus der Philosophie, der Kunst und der Literatur. Unter dem Einfluss von Wilhelm Dilthey, Friedrich Schleiermacher, später im 20. Jahrhundert dann unter der Ägide von Hans-Georg Gadamer, Eric D. Hirsch und Wolfgang Iser, entwickelte sich eine hochkomplexe Hermeneutik, die die Lehre vom Verstehen präziserte und dynamisierte.²⁰ Der bekannte ‚hermeneutische Zirkel‘ – oder besser die ‚hermeneutische Spirale‘ – beschrieb den abstrakten Zusammenhang von allgemeinen Vorurteilsstrukturen und ihrer Korrektur durch Erkenntnisprozesse (Erfahrung oder Lektüre), die im Idealfall zu einem immer angemesseneren Verstehen des komplexen Erkenntnisgegenstandes führen sollten. Sowohl die Epistemologie als auch die Hermeneutik stellten dabei bestimmte Grundbedingungen des Verstehensprozesses nicht in Frage; vereinfacht könnte man sagen, dass einem Erkenntnisprozess in der Regel eine Hypothese oder eine Theorie vorausgehen, die zu verifizieren oder zu falsifizieren seien, dass sich der Erkenntnisprozess in einem linearen zeitlichen Raum vollziehe und die Grundgesetze der Kausalität zu beachten habe. Auch die Gesetze des Syllogismus seien weiterhin gültig – allerdings nicht, wenn sie bloß auf unbegründeten, nur von der Tradition, nicht aber der Erfahrung oder ‚wissenschaftlichen Überprüfung‘, als ‚wahr‘ bezeichneten Axiomen

¹⁸ Vgl. Francis Bacon, „Novum Organum“ (1620), in: *The Works of Francis Bacon*, eds. James Spedding, Robert Leslie Ellis and Douglas Denon Heath; 14 vols.; vol. 1 (London: Longman et al., 1858), 149-365.

¹⁹ Vgl. Bacon, *The Advancement of Learning*, 36.

²⁰ Vgl. hierzu auch Rudolf Freiburg, *Autoren und Leser: Studien zur Intentionalität literarischer Texte* (Frankfurt a.M. et al.: Lang Verlag, 1985), 13-44.

beruhen sollten. Der amerikanische Theoretiker Hirsch betonte die erkenntnisprägende Macht der generischen Vorerwartungen, wenn er das Wesen und die Funktion des ‚wahren Genres‘, eines Konstruktes, das den Konturen des Typbegriffes von Edmund Husserl nachempfunden war, für den Verständnisprozess herausstrich.²¹ Die von der Konstanzer Schule eingeleitete allgemeine Psychologisierung des Verstehensprozesses, die den Rezipienten als substanziellen Koproduzenten der Sinnhaftigkeit von Texten und Erkenntnisobjekten in den Mittelpunkt ihrer Ausführung stellte, akzentuierte auch in ihrer weiteren Entwicklung zur Rezeptionsästhetik historischer und systematischer Provenienz den dynamischen Charakter aller Verstehensakte. Erkennen und Verstehen, so wurde deutlich, waren in vielfältiger Weise aufeinander bezogen, aber nicht frei von Irrtümern: *Errare humanum est*.

Im Kontext der Datenfluten, im Machtbereich von *Big Data* und *Deep Learning*, verlieren viele dieser Grundannahmen an Überzeugungskraft. Einige Axiome der Epistemologie und Hermeneutik sind von nun an – wenn nicht gar in Frage zu stellen – zumindest zu modifizieren und zu revidieren. Wenn Sir Karl Popper in seiner Wissenschaftstheorie auf die Unmöglichkeit der Verifikation von Erkenntnissen aufmerksam machte, so verwies er unter anderem auch auf das Unvermögen des Menschen, alle möglichen auftretenden Fälle und Eventualitäten eines Problemkomplexes zu berücksichtigen.²² Die Begrenztheit des menschlichen Verstandes war natürlich schon in der Antike erkannt worden,²³ und schon immer hatte man versucht, diese natürlichen Grenzen durch Instrumente zu überschreiten. Die durch die eingeschränkte menschliche Sinneswahrnehmung gesetzten Grenzen konnten durch logische Überlegungen und abstrakte Denkvorgänge transzendiert werden, sodass man etwa die Mathematik als ‚Teleskop des menschlichen Geistes‘ bezeichnen dürfte.²⁴ Der Computer stellt in dieser Hinsicht ein noch weitaus mächtigeres Instrument zur Verbesserung der menschlichen Erkenntnisfähigkeit dar, aber erst im Zusammenhang mit den schier unergründlichen Möglichkeiten der informationstechnisch ‚neuronalen‘ Vernetzung

²¹ Vgl. hierzu Eric D. Hirsch, „The Concept of Genre“, in: *Validity in Interpretation* (New Haven: Yale University Press, 1971), 68-126.

²² Vgl. Karl Popper, *Logik der Forschung* (Berlin: Akademie Verlag, 2004).

²³ Vgl. den sokratischen Ausspruch „*scio me nihil scire*“; das griechische Original des Ausspruchs findet sich in *The Oxford Dictionary of Quotations*, ed. Partington, 654.

²⁴ Vgl. etwa Thomas Weth, „Faszination Unendlichkeit: Über die Struktur des Unendlichen“, in: *Unendlichkeit: Fünf Vorträge*, hg. Rudolf Freiburg (Erlangen: FAU Press, 2016), 105-138; 105.

bietet sich ein Potenzial der epistemologischen Grenzüberschreitung, das alles bislang Dagewesene in den Schatten stellt. Die den heutigen parallel arbeitenden Superrechnern zur Verfügung stehenden riesigen Datenmengen aus Exa- und Zettabytes machen eine Neudefinition des Begriffs ‚Erkenntnisgegenstand‘ nötig.²⁵ Zwar beruhen die im Kontext von *Big Data* gewonnenen Erkenntnisse nach wie vor auf empirischen Verfahren, aber auch diese haben sich in ihrem Wesen geändert. Während in traditionellen Erkenntnisprozessen darum gerungen wurde, Fehler, Unschärfen und Ungenauigkeiten in der Datenerhebung zu eliminieren, nehmen die neuen – auf dem Einsatz von Algorithmen beruhenden – Verfahren diese Defizite bewusst in Kauf, da die hohe Datenmenge aus statistischen Gründen einzelne Fehler als ‚vernachlässigenswert‘ erscheinen lässt.²⁶ Schon Bacon hatte den erkenntnishemmenden Faktor der wissenschaftlichen Tradition erkannt, der durch gedankenloses Beharren auf Kontinuität Neues verhindert. Jeder Wissenschaftler, der versucht, Hypothesen für Forschungen aufzustellen, weiß, wie schwer es ist, den Spagat zwischen dem Bekannten und dem Unbekannten zu leisten; die Hypothese kann sich nur schwer aus dem Gravitationsfeld des Bekannten lösen, das Traditionelle färbt selbst innovative Thesen ein. *Big-Data*-Informationssysteme hingegen ermöglichen in naher Zukunft offenbar eine *nicht menschliche* Hypothesenbildung, die durch künstliche Intelligenz gesteuert wird.²⁷ Mit der zunehmenden Entwicklung der künstlichen Intelligenz wird diese vom Bereich des Humanen distanzierte Theoriebildung an Bedeutung gewinnen, menschliche Vorurteilsstrukturen einen immer geringeren Raum einnehmen und somit Erkenntnisprozesse nicht mehr in negativer Weise beeinträchtigen. Die nur noch von Maschinen erfassbaren – für ein menschliches Gehirn unvorstellbar großen – Datenmengen führen aber auch zu einer Revision von Grundvoraussetzungen des bisherigen Denkens, das sich bislang in der Regel ‚linear‘ vollzog und auf den Prinzipien der ‚Kausalität‘ beruhte. Zukünftige Quantencomputer könnten eine derartige Linearität obsolet erscheinen lassen, und schon jetzt wird das ‚Kausalitätsdenken‘ von den *Big-Data*-Statistikern zwar nicht völlig ersetzt, aber doch ergänzt. Mit Blick auf das immense Datenvolumen kommt dem Prinzip der ‚Korrelation‘ eine immer größere

²⁵ Vgl. hierzu auch Klaus Meyer-Wegener, „Erstmal einfach alles speichern: *Big Data* als Aufgabe für die Informatik“, 99-115 in diesem Band.

²⁶ Vgl. Mayer-Schönberger und Cukier, *Big Data*, 45-46.

²⁷ Vgl. Viktor Mayer-Schönberger, „Neue Erkenntnisse über die Wirklichkeit: Wie wir die Diktatur der Daten verhindern können“, *Forschung und Lehre* 9 (2014): 706-707; 706.

Bedeutung zu.²⁸ Erkenntnisse werden auch dann fruchtbar und nützlich sein,²⁹ wenn sie nicht in jedem Fall ‚erklärt‘ werden können:

Im Zeitalter von Big Data werden diese neuen Analysemethoden irgendwann eine ganze Welle neuer Erkenntnisse und nützlicher Voraussagen auslösen. Wir werden bisher völlig unbekannte Zusammenhänge erkennen. Wir werden auf komplexe technische und soziale Dynamiken aufmerksam werden, die uns trotz aller Bemühungen bisher entgangen sind. Aber am wichtigsten ist, dass solche nichtkausalen Analysen unser Verständnis der Welt verbessern werden, indem sie nach dem *Was* statt nach dem *Warum* fragen.³⁰

Zwar darf mit Recht bezweifelt werden, dass mit dem Einzug von *Big Data* in das wissenschaftliche Denken das „Ende der Theorie“ eingeleitet werde,³¹ aber der Stellenwert des theoretischen Begründungszusammenhangs könnte in der weiteren Entwicklung epistemologischer Strategien wohl an Bedeutsamkeit verlieren.³² Das hermeneutische Grundgesetz, das das Verstehen als ein Produkt der Interdependenz erklärt, bei dem das ‚Einzelne‘ aus dem Blickwinkel des ‚Ganzen‘ und das ‚Ganze‘ aus dem Blickwinkel des ‚Einzelnen‘ aufzufassen sind, erfährt eine tiefgreifende Revision, da sich der Begriff des ‚Ganzen‘ im Angesicht des *Big-Data*-Kosmos in frappierender Weise verändert hat. Während ein menschlicher Leser nur eine begrenzte Zahl von literarischen Texten lesen kann, um zu einem ‚hermeneutischen Urteil‘ zu gelangen, erfassen zeitgemäße Scanner unter Umständen alle zu einem definierten Textkorpus zu zählenden Quellen. Statt traditioneller hermeneutischer Prozesse regelt nun das intelligente Auswerten von ‚Mustern‘ den Verständnisprozess. Weitert man den ‚Text‘-Begriff im Sinne kulturwissenschaftlicher Praktiken aus, bezieht also ‚Geschichte‘, ‚Lebenswelt‘, ‚Politik‘ und ‚Kultur‘ in ihrer ‚Textualität‘ mit ein, so lassen diese Muster sogar Schlüsse auf die Zukunft,

²⁸ Mayer-Schönberger, „Neue Erkenntnisse über die Wirklichkeit“, 707.

²⁹ Vgl. auch Mayer-Schönberger und Cukier, *Big Data*, 70: „Korrelationen sind in einer Welt der wenigen Daten nützlich, aber im Kontext von Big Data kommen sie richtig groß heraus. Mit ihrer Hilfe gelangen wir schneller, leichter und besser an Erkenntnisse als zuvor.“

³⁰ Mayer-Schönberger und Cukier, *Big Data*, 82-83.

³¹ Vgl. auch Chris Anderson, „Das Ende der Theorie: Die Datenschwemme macht wissenschaftliche Methoden obsolet“, in: *Big Data: Das neue Versprechen der Allwissenheit*, hg. Heinrich Geiselberger und Tobias Moorstedt (Frankfurt a.M.: Suhrkamp, 2013), 124-130.

³² Vgl. auch Klaus Mainzer, „Die Berechnung der Welt: Können Big Data-Ergebnisse Theorie und Beweis ersetzen?“, *Forschung und Lehre* 9 (2014): 696-699.

also Prognosen und Vorhersagen zu.³³ Die Welt der *Big Data* erscheint wie in einer digitalen Version behavioristischer Lehren als unerklärliche ‚schwarze Box‘ (*black box*), auf einen digitalen *stimulus* folgt ein *response*, das Ergebnis zählt, die Theorie ist sekundär.

Im Universum der *Big Data* kristallisiert sich eine Welt heraus, die ahuman und posthuman zugleich erscheint und zahlreiche gravierende Probleme zeitigt. Die Daten selbst sind nämlich nun nicht mehr wie in der traditionellen Forschung frei verfügbar, sondern werden in einem Großteil der Fälle von kommerziell agierenden Großkonzernen wie *Google*, *Facebook* oder *Twitter* generiert, sodass die Freiheit der Forschung an die Grenzen kommerzieller Interessen stößt. Prinzipien wie Inklusion und Exklusion (wer hat Zugriff auf die Daten, wer kann und darf sie auswerten?) werden wohl Gerichte und Ethikräte beschäftigen.³⁴ Allmählich werden die Universitäten ihren Status als „Torwächter zur Bildung“ verlieren.³⁵

Während die Wende zum Ahumanen im Bereich der Erkenntnis einerseits aufgrund der Eliminierung von Vorurteilen durchaus Chancen zur Verbesserung bietet, kann die Welt der *Big Data* in hermeneutischen Prozessen andererseits aber auch zu einer sehr großen Gefahr werden. Die Datafizierung der Welt und die Zusammenführung der Daten zu hochkomplexen mathematischen Systemen ist als ‚digitales‘ Erkenntnisinstrument keinesfalls einem simplen ‚Teleskop‘ vergleichbar, sondern dieses dynamische Instrument interagiert auf vielfältige Weise mit dem erkennenden Subjekt. Während die klassische Hermeneutik den ‚hermeneutischen Kreis‘ zugunsten der Metaphorik von der ‚hermeneutischen Spirale‘

³³ In der Sekundärliteratur sind viele Belege für Zukunftsvorhersagen zu finden; so sagte *Google Flu* aufgrund der Suchanfragen im Jahre 2009 korrekt eine Grippeepidemie in den USA voraus; vgl. Mayer-Schönberger und Cukier, *Big Data*, 7-8. Auch das Beispiel von dem ahnungslosen Vater, der durch eine Firma, die das Einkaufsverhalten ihrer Kunden mit Algorithmen auswertete, erfahren musste, dass seine noch minderjährige Tochter schwanger sei, wird immer wieder als Beispiel für *Big-Data*-Prognosen angeführt; vgl. ebd., 76-77; Drösser, *Total Berechenbar*, 117-118, und Markus Morgenroth, *Sie kennen dich! Sie haben dich! Sie steuern dich! Die wahre Macht der Datensammler* (München: Knauer, 2016), 16-17. Gefährlich ist es auch, wenn *Big-Data*-Systeme politischen Aufruhr vorhersagen können; zum „*riot forecasting*“ vgl. Ramón Reichert, „Einführung“, in: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*, hg. Ramón Reichert (Bielefeld: Transcript Verlag, 2014), 10-31. (Für diesen Sammelband wird im Folgenden die Sigle *BD* benutzt.)

³⁴ Vgl. Lev Manovich, „Trending: Verheißungen und Herausforderungen der Big Social Data“, in: *BD*, 65-83; 70-80.

³⁵ Vgl. David M. Berry, „Die Computerwende: Gedanken zu den Digital Humanities“, in: *BD*, 47-64; 53.

aufgab, da jeder Verstehensprozess *idealiter* auf eine neue und höhere Ebene des Verstehens führen sollte, mutet der Umgang mit *Big Data* in gewisser Weise wie eine Rückkehr in die Zirkularität des Denkens und des Verstehens an. Das in die *Big-Data*-Welt eingebundene Erkenntnisobjekt ist selbst nun wieder auch Erkenntnisobjekt der *Big-Data*-Erkenntnisstrategien, mithin ihr Opfer. Durch jede digital präsentierte Anfrage in einem Suchportal gibt das Subjekt Daten von sich selbst preis, stillt durch sein Erkenntnisinteresse unentwegt den Wissensdurst der intelligenten Maschinen, die ihrerseits im Hintergrund stetig an der Vervollständigung des Benutzerprofils arbeiten, um jedermann dann exakt jenes Wissen zu bieten, nach dem es ihm verlangt. Menschliche Neugierde kollidiert mit dem technokratischen Wissensdurst eines gigantischen digitalen Nimmersatts, der nie schläft. Dieser digital-hermeneutische Zirkel entpuppt sich als ein veritabler Teufelskreis, da er im Extremfall gar keinen Erkenntnisgewinn mehr präsentiert, sondern den Erkenntnisuchenden mit einem digitalen Spiegelbild seiner Hauptinteressen abspeist; der Erkenntnisuchende wird zu einem hermeneutischen Narziss. Die Theorie kennt für diesen Vorgang den Ausdruck der ‚Internet-Blase‘: Eingesponnen in ein Kokon der eigenen virtuellen Welt, ist der *user* dazu verurteilt, permanent im engen Sinnbezirk der eigenen Vorlieben zu verweilen.³⁶ Jede *Google*-Anfrage lässt den Kokon undurchdringlicher für die ‚Außenwelt‘ werden, die Isolation führt zu einer neuen Form eines digital initiierten Solipsismus.

Vernetzte Einsamkeit: Der Einfluss der *Big-Data*-Welt auf die menschliche Psyche

Schon unterhalb der Schwelle des als *Big Data* zu bezeichnenden Datenaufkommens ist die Reizüberflutung durch die tägliche Datenflut eine starke Belastung für den menschlichen Geist. In der Geschichte der Telekommunikation wurde immer wieder die Metaphorik des Nervensystems bemüht,³⁷ um die Verbindungen zwischen einzelnen – auf elektronische Weise miteinander kommunizierenden – Teilnehmern zu verdeutlichen. Im Zeitalter der Datenfluten bekommt der skeptische Beobachter den Eindruck, dass das digitale Nervensystem der Menschheit an einer schweren Form der Neurasthenie erkrankt ist, die sich auf den menschlichen

³⁶ Vgl. hierzu Eli Pariser, *The Filter Bubble: What the Internet is Hiding from You* (New York: Penguin, 2011).

³⁷ Vgl. hierzu Regine Buschauer, „(Very) Nervous Systems: Big Mobile Data“, in: *BD*, 405-436, 407.

Benutzer auswirkt: Bei *Facebook* und *Twitter* werden pro Tag derartig viele Nachrichten gesendet,³⁸ dass kein einzelner Mensch sie mehr in seinem Leben zu lesen vermag, auf *Youtube* werden pro Stunde gigantische Mengen von Videos heraufgeladen, sodass ein Einzelner mit deren Konsum völlig überfordert wäre. Die neu eingerichteten *Clouds* der Internet-Betreiber lassen den digitalen Starkregen von Bildern und Musik auf den Konsumenten herunterprasseln, die Mitglieder von sozialen Netzwerken überbieten sich in der Bereitstellung von Trivialitäten, so sehr sogar, dass die menschliche Psyche keine Ruhe mehr findet. Georg Simmel prägte zu Beginn des 20. Jahrhunderts den Begriff der ‚Blasiertheit‘, mit dem er die geistige Paralyse beschreiben wollte, die von jenen als psychischer Verteidigungsmechanismus entwickelt wird, die Tag für Tag den unaufhörlichen Reizen des Lebens in modernen Metropolen ausgesetzt sind.³⁹ Eine derartige Verteidigungsstrategie gegen Reizüberflutung ist auch im Kontext des Digitalen unausweichlich: Es gibt keine ‚Sendepausen‘ und daher keine Schutzräume mehr.⁴⁰ Die neuen Medien, die die Datenflut ins Unermessliche steigern, überwinden mühelos die Grenzen von Ort und Zeit; jedermann kann jederzeit von fast jedem Ort der Erde mit fast jedem anderen, der eine entsprechende digitale Ausrüstung besitzt, kommunizieren. Doch paradoxerweise entsteht gerade durch die enge Vernetzung der kommunizierenden Subjekte eine soziale Vereinsamung, die überall im Alltagsleben beobachtet werden kann. Der *homo connexus* ist auf fürchterliche Weise mit sich allein.⁴¹ Wer heute an einer großen Universität die Studierenden vor dem *Auditorium Maximum* beobachtet, während diese auf den Beginn einer Vorlesung warten, wird kaum Zeuge einer aktiv im Hier und Jetzt stattfindenden Konversationskultur werden; stattdessen sitzt jeder einzelne isoliert von allen anderen, vertieft in sein Smartphone oder Tablet, um auf *Facebook* zu posten, dass niemand mit

³⁸ Das Verhalten der *Facebook-User* wird unter anderem ausgewertet, um deren Kreditwürdigkeit zu überprüfen; vgl. Morgenroth, *Sie kennen dich!*, 192-193.

³⁹ Vgl. Georg Simmel, „Die Großstädte und das Geistesleben“, in: *Die Großstadt: Vorträge und Aufsätze zur Städteausstellung*, hg. Thomas Petermann (Dresden: Zahn & Jaensch, 1903), 185-206.

⁴⁰ Dieser beklagenswerte Mangel an ‚Ungestörtsein‘ bedingt die Unfähigkeit vieler, mit der ‚Einsamkeit‘ und auch der ‚Langeweile‘ zurechtzukommen. Die beste Studie zum Einfluss mobiler Telekommunikation auf die Psyche und das Sozialverhalten von Individuen findet sich in Sherry Turkle, *Reclaiming Conversation: The Power of Talk in a Digital Age* (New York: Penguin, 2015).

⁴¹ Zum Begriff *homo connexus* vgl. auch Joachim M. Buhmann, „Eine Revolution unseres Weltbildes: Über die informationstechnologische und erkenntnistheoretische Herausforderung von *Big Data*“, *Forschung und Lehre* 9 (2014): 700-702.

ihm sprechen möchte. Anstatt kurz an die Tür des Mitbewohners im Studentenheim zu klopfen, wird eher eine Sprachnotiz, besser noch ein Kurztext bei *WhatsApp* lanciert, der menschliche Kontakt wird als unangemessene ‚Nähe‘ – vielleicht sogar als ‚Bedrohung‘ – empfunden. Dieser paradoxe Zustand des „gemeinsamen Alleinseins“ ist längst zum Gegenstand psycho-sozialer Studien geworden;⁴² im Angesicht des Zusammenwachsens sozialer Netzwerke driftet der Einzelne in die totale Vereinsamung ab, die einen fruchtbaren Nährboden für psychische Erkrankungen darstellt:

After an evening of avatar-to-avatar talk in a networked game, we feel, at one moment, in possession of a full social life and, in the next, curiously isolated, in tenuous complicity with strangers. We build a following on Facebook or MySpace and wonder to what degree our followers are friends. We recreate ourselves as online personae and give ourselves new bodies, homes, jobs, and romances. Yet, suddenly, in the half-light of virtual community, we may feel utterly alone. As we distribute ourselves, we may abandon ourselves.⁴³

In der zweiten Dekade des 21. Jahrhunderts scheint das Subjekt in den Datenfluten, die es lebenslang begleiten werden, unterzugehen. Wenn Virginia Woolf zu Beginn des zurückliegenden Jahrhunderts bedauernd anmerkte, das moderne Subjekt habe sich etwa im Jahre 1910 merklich verändert,⁴⁴ so gilt diese Beobachtung etwas mehr als hundert Jahre später in noch viel beängstigenderer Weise. Formen des zwischenmenschlichen Agierens werden aufgegeben; menschliche Nähe wird durch maschinelle ‚Streicheleinheiten‘ vorgetäuscht: In Altenheimen etwa werden Demenz-Kranke mit künstlichen Robben ausgestattet, damit sie das Gefühl bekommen, ein lebendiges Wesen habe sich ihnen zugewandt.⁴⁵ Nach dem Verebben der Tamagotchi-Mode schafften sich zahlreiche Haushalte in Amerika künstliche Haustiere an, kinderlose Ehen trösteten sich mit artifiziellen Babys, und selbst das Liebesleben ist nicht gegen die digitale Invasion immun, wenn ‚Roxxy‘, der ‚erste Sex-Roboter in der Geschichte der

⁴² Vgl. Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other* (New York: Basic Books, 2011).

⁴³ Ebd., 11-12.

⁴⁴ Vgl. hierzu Pericles Lewis, „On or about December, 1910“, in: *The Cambridge Introduction to Modernism* (Cambridge: Cambridge University Press, 2010 [2007]), 86-94.

⁴⁵ Vgl. Turkle, *Alone Together*, 103-109.

Menschheit' im Schlafzimmer Regie führt.⁴⁶ Virtuelle Welten wie *Second Life* bieten digitale Kompensation für frustrierte Zeitgenossen, die sich im Netz eine erfreulichere Existenz mit erfolgreicherer Karriere aufbauen; Psychologen ermuntern Ratsuchende durchaus dazu, im Internet unter falscher Identität nach Bedürfnisbefriedigung zu suchen,⁴⁷ und Online-Spiele verschaffen den ultimativen ‚Kick‘ durch Rollenidentifikationen, die von einer vorbildlichen Laufbahn als Banker oder Börsianer bis zur kriminellen Karriere eines apokalyptisch agierenden ‚Terminators‘ reichen können. In der Gegenwart spalten sich die Persönlichkeiten in Portfolios selbststilisierter Wunschbilder auf, wenn in Netzwerken Eigenschaften ge-*postet* werden, die man im wahren Leben wohl kaum vorweisen kann. Es bildet sich eine Form der ‚digitalen Schizophrenie‘ heraus, die auf Dauer keiner Gemeinschaft bekömmlich sein kann. Die Unmöglichkeit, dem eigenen Arbeitsplatz durch räumliche Distanz zu entkommen, die Verfügbarkeit für Firma und Chef über die Grenzen der Arbeitszeit hinweg, lassen zudem bei vielen das Gefühl einer ‚digitalen Paranoia‘ aufkommen, die noch durch den Wegfall der gewohnten Orientierungsstrukturen von Raum und Zeit intensiviert wird.⁴⁸ Während man mit dem Strom der Datenflut schwimmt, erfährt man eine Beschleunigung, aber die Orientierung geht verloren; im Angesicht der unübersehbaren Informationen reduziert sich der Eigenanteil am Ganzen zu einer *quantité négligeable*. Die Minderwertigkeitskomplexe, die sich früher im Bewusstsein herausgebildet hatten, man sei auf dem Bereich der menschlichen Kreativität zum Epigontum verdammt, erfahren aufgrund der rasanten digitalen Beschleunigung eine paradoxe Revision: In der Gegenwart ist der Zeitgenosse ein Epigone des Gegenwärtigen, er kommt auch dann zu spät, wenn er pünktlich ist. Wenn ein Wissenschaftler heute noch glaubt, originell und kreativ zu sein, muss er häufig schmerzhaft feststellen, dass er einfach nicht lange genug recherchiert hat.

Zwischen Exhibitionismus und erzwungener Entblößung: Gesellschaftlicher Wandel in der Welt der Daten

Die psychologischen Veränderungen, die durch *Big Data* ausgelöst werden, stehen in engem Zusammenhang mit den soziologischen Wandlun-

⁴⁶ Vgl. ebd., 53-66, 67-73 und 65-66.

⁴⁷ Vgl. Jan Kalbitzer, *Digitale Paranoia: Online bleiben, ohne den Verstand zu verlieren* (München: Beck, 2016), 91-94.

⁴⁸ Vgl. ebd., 55-61 (Störungen der Zeitstruktur) und 68-78 (Störungen in Bezug auf das Ortsempfinden).

gen in der Gegenwart.⁴⁹ Bereits als das *Time Magazine* im Jahre 2006 in der Folge des Siegeszugs des Internet-Video-Portals *Youtube* den anonymen ‚user‘, der längst zum ‚producer‘ mutiert war, also den Internet-Konsumenten (‚You‘), zur ‚Person des Jahres‘ wählte, zeichnete sich der tiefgreifende Wandel in der Gesellschaft ab.⁵⁰ Die hohe Anzahl von *Twitterern*, *Bloggern*, Produzenten von Videos, Literatur und Bildern zeugt davon, dass bislang passive Konsumentenschichten Zugang zu Formen der Veröffentlichung eigenen Materials bekamen und diesen auch intensiv nutzten. Inzwischen ist eine weitere Dekade vergangen, und das Material, das auf soziale Plattformen hochgeladen wird, erlebt ein exponentielles Wachstum. Neben dieser intentionalen Preisgabe persönlicher Daten,⁵¹ die in manchen Fällen gar einem digitalen Exhibitionismus gleichkommen mag,⁵² finden sich in der Datenwelt aber auch unzählige Spuren von Aktivitäten, die nicht wissentlich hinterlassen werden:

Kameras, die in Einkaufspassagen bisher Diebstähle aufklärten, analysieren neuerdings Kleider und Schmuck ihrer Kunden und informieren die Ladenbesitzer über deren sozioökonomischen Status. Spielekonsolen zählen, wie viele Kinder vor dem Bildschirm sitzen, und erkennen, ob sie Spaß haben. *E-Book-Reader* zählen, welche Stellen häufig gelesen, welche überblättert werden und wo sich die Pupillen weiten.⁵³

Jede Anfrage bei *Google*, jeder Klick auf einen ‚like‘-Button der sozialen Netzwerke, jede Bestellung bei *Amazon*, jeder elektronische Bezahlvorgang mit EC- oder Kreditkarte hinterlässt eine Myriade von Daten und Metadaten, die mit den ausgeklügelten Algorithmen der *Big-Data*-Welt zu Protokollen zusammengefasst und zur Kontrolle der Kunden genutzt

⁴⁹ Vgl. hierzu auch Sebastian Büttner, „Algorithmen, Datenkraken und gläserner Bürger: *Big Data* als Herausforderung für die Gesellschaftsanalyse“, 39-60 in diesem Band.

⁵⁰ Vgl. *Time Magazine* (25. Dezember 2006, 1. Januar 2007), Titelseite. (online.)

⁵¹ Auf diese besondere Gefahr weist auch Welzer hin; vgl. Welzer, *Die smarte Diktatur*, 200: „Die Auflösung der Demokratie geschieht im Rahmen der Demokratie.“ [Hervorhebung im Original].

⁵² *Big-Data*-Welten kommen dem Bedürfnis des Menschen nach, „Geständnisse“ abzulegen; in jedem Fall erinnern sie aber an die Einrichtung eines *Panopticum*s; zu diesen auf den Ideen Foucaults beruhenden Vorstellungen vgl. auch Tom Boellstorff, „Die Konstruktion von Big Data in der Theorie“, in: *BD*, 105-131; 117-118.

⁵³ Morgenroth, *Sie kennen dich!*, 13; zum Einfluss von *Big Data* auf den Buchmarkt vgl. auch Svenja Hagenhoff, „Außer Kontrolle und auch noch aus der Form: Zur Digitalisierung des Publikationswesens“, 77-98 in diesem Band.

werden können.⁵⁴ Die Welt der persönlichen Daten erscheint wie ein nie versiegender Fluss, die Datenflut trägt also Informationen mit sich, die von geschickten Datensuchern und -händlern abgeschöpft werden können.⁵⁵ Supercomputer errechnen aus den Abermillionen von Angaben Muster, die ihrerseits wiederum von hohem Wert für Unternehmen sind, die ihre Werbestrategien auf diesem Wege personalisieren können. Informationen sind also mit Geld gleichzusetzen, und der „Informationskapitalismus“ blüht derzeit wie nie zuvor: „Wenn intelligente Maschinen der Motor des Informationskapitalismus sind, ist Big Data ihr Treibstoff und das Internet ihr Chassis.“⁵⁶

Wie man in einschlägigen Monographien leicht nachlesen kann, lässt sich aus dem Musikgeschmack von Menschen, die ihre Trennung bei Facebook öffentlich gemacht haben,⁵⁷ ein Trend herauslesen;⁵⁸ diese Information kann für den Verkauf von CDs und das Angebot von MP3-Musik-Plattformen höchst rentabel sein. Die im Netz veröffentlichten Fotos bringen Entwicklungslabors, Software-Anbieter und Rahmenhändler nahe an ihr Zielpublikum, wie etwa auch die Aktivitäten der in Deutschland tätigen Firma *Fotocommunity* beweisen. Zu diesen Datenfluten gesellen sich die unvorstellbar großen Mengen an Informationen, die rund um die Uhr von den Smartphones,⁵⁹ Tablets, ‚wearables‘, Navigationssystemen, Kunden-, Bank- und Kreditkarten an die Server von Großkonzernen gesendet werden.⁶⁰

Vereinzelt sind diese Daten weder besonders wertvoll, noch besonders gefährlich; nach der algorithmischen Zusammenführung und statistischen

⁵⁴ Vgl. auch Alexander Galloway und Eugene Thacker, „Protokoll, Kontrolle und Netzwerke“, in: *BD*, 289-311; 296: „Wir wollen noch einmal betonen: Das Protokoll ist eine materialisierte Emanation von dezentraler Kontrolle.“

⁵⁵ Vgl. Mayer-Schönberger und Cukier, *Big Data*, 168-175.

⁵⁶ Yvonne Hofstetter, *Sie wissen alles: Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen* (München: Bertelsmann, 2014), 12.

⁵⁷ Facebook ist offenbar in der Lage, sich anbahnende Partnerschaften zwischen seinen *usern* vorherzusagen, bevor diese selbst von ihren Sympathien wissen; vgl. Morgenroth, *Sie kennen dich!*, 188; zur Rolle von Facebook im *Big-Data*-Kontext vgl. ebd., 185-193.

⁵⁸ Es gibt sogar einen ‚*Facebook-Happiness-Index*‘; vgl. hierzu Ramón Reichert, „Facebooks Big Data: Die Medien- und Wissenstechniken kollektiver Verdattung“, in: *BD*, 437-452; 443.

⁵⁹ Zum problematischen Wesen von Smartphones, die auch die Ökologie gefährden, vgl. auch Welzer, *Die smarte Diktatur*, 74-82.

⁶⁰ Vgl. auch die Beispiele in Nikolas Forgó, „Wer bin ich schon?: Die Datafizierung des Lebens und das Datenschutzrecht“, *Forschung und Lehre* 9 (2014): 704-705.

Auswertung stellen sie aber eine veritable Fundgrube für die verschiedensten Interessensgemeinschaften dar.⁶¹

Aber auch für seriös arbeitende Soziologen zeichnet sich somit ein recht neues Forschungsgebiet ab,⁶² in dem die Daten nicht erst durch Befragungen mühsam erhoben werden müssen; die Nutzer des Internets stellen sie mehr oder minder freiwillig zur Verfügung, sodass Trends leicht mit statistischen Mitteln erkannt werden können:

Das Aufkommen sozialer Medien im ersten Jahrzehnt des 21. Jahrhunderts schuf neue Möglichkeiten, dynamische soziale und kulturelle Prozesse zu studieren. Zum ersten Mal können wir die Fantasien, Meinungen, Ideen und Gefühle Hunderter Millionen von Menschen verfolgen. Wir können die Bilder und Videos, die sie produzieren, betrachten und kommentieren, die Unterhaltungen, die sie gerade führen, belauschen, ihre *Blogs* und Twitternachrichten lesen, nach ihren Karten navigieren, uns ihre Musik anhören und ihre Wege im physischen Raum verfolgen. Und dafür müssen wir sie nicht einmal um Erlaubnis bitten, da sie uns dazu regelrecht ermutigen, indem sie all diese Daten veröffentlichen.⁶³

Inzwischen hat sich das Gebiet der ‚*culturomics*‘ etabliert, mit deren Hilfe man Aussagen über den Zustand bestimmter Kulturen und die Befindlichkeit ihrer Vertreter treffen kann.⁶⁴ Googles Instrument *nGram* bietet jedem Nutzer die Möglichkeit,⁶⁵ die Verwendung von Begriffen (etwa ‚*Big Data*‘) für einen bestimmten Zeitraum zu überprüfen.⁶⁶ Die sozialen Foren geben Auskunft darüber, wie sich die Gesprächskultur in der Gegenwart

⁶¹ Die einschlägigen Monographien zum Thema *Big Data* übertrumpfen sich gegenseitig mit Beispielen der totalen Überwachung; vgl. etwa den informationsreichen Band von Bruce Schneier, *Data und Goliath: Die Schlacht um die Kontrolle der Welt*, übers. Sigrid Schmid und Claudia Van Den Block (München: Redline Verlag, 2015).

⁶² Vgl. auch Jean Burgess und Axel Bruns, „Twitter-Archive und die Herausforderungen von ‚Big Social Data‘ für die Medien- und Kommunikationswissenschaft“, in: *BD*, 191-202; 192.

⁶³ Lev Manovich, „Trending: Verheißungen und Herausforderungen der Big Social Data“, in: *BD*, 65-83; 67.

⁶⁴ Vgl. hierzu Richard Rogers, „Nach dem Cyberspace: Big Data, Small Data“, in: *BD*, 173-87; 178.

⁶⁵ Vgl. auch Daniel Rosenberg, „Daten vor Fakten“, in: *BD*, 133-156; 142.

⁶⁶ Analysen dieser Art sind durchaus geeignet, bisherige Grundannahmen der Geistes- und Kulturgeschichte in Frage zu stellen; vgl. auch Thomas Weitin, „Thinking Slowly: Literatur lesen unter dem Eindruck von Big Data“, *LitLingLab Pamphlet 1* (März 2015), 1-18; Weitin berichtet etwa davon, dass die Leerstellentheorie Wolfgang Iersers durch *eyetracker*-Methoden (Fixation, Sakkaden, Regressionen) auf ihre Richtigkeit hin überprüft werden könnte (14).

entwickelt, meist mit dem bedenkenswerten Ergebnis, dass die Aggression zunimmt, die Hemmungen in Bezug auf xenophobe und rassistische Aussagen schwinden und das Netz zum Ventil für angestaute Frustration enttäuschter Zeitgenossen mutiert. Biometrische und soziometrische Analysen stellen einerseits die Instrumente dar, mit denen das *panta rhei* der Datenfluten eingedämmt werden soll, andererseits bedingen sie ihrerseits selbst ein Anschwellen der Fluten. Nicht erst seit den Enthüllungen, die der als ‚Whistleblower‘ diffamierte Edward Snowden im Jahre 2013 über die weltweiten Abhör- und Spionagemethoden der Geheimdienste veröffentlichte, weiß die Allgemeinheit, wie transparent digitale Kommunikation geworden ist. Mit den Mitteln der neuen Technik können Regierungen, Konzerne, Firmen, aber auch Individuen bequem aus der Distanz überwacht werden. Die aus den Smartphones und anderen digitalen Geräten generierten Metadaten lassen auf unkomplizierte Weise Persönlichkeits- und Bewegungsprofile entstehen,⁶⁷ die zur Vorhersage von Verhaltensweisen und zukünftigen Aktivitäten werden.⁶⁸ Im Angesicht der neueren algorithmischen Informationsverarbeitungssysteme wirkt George Orwells ‚*Big Brother*‘ wie der kleine Bruder von ‚*Big Data*‘:

Big Data läutet einen neuen Bund ein: Es ist der Bund zwischen Kapitalismus und Diktatur mit der Verheißung neuer, profitabler Geschäftsmodelle der totalen Überwachung. Als *Informationskapitalismus* etabliert Big Data die Diktatur von Informationseliten, weil sie über unsere Daten und über Schlüsseltechnologien zu deren Analyse verfügen.⁶⁹

Man braucht nicht viel Phantasie, um vorherzusagen, dass die privaten medizinischen Daten,⁷⁰ die die immer populärer werdenden *wearables* zur Verfügung stellen, zwar einerseits von hohem Wert für die Prognose von Epidemien und Krankheiten sind,⁷¹ wohl auch für die Therapie von

⁶⁷ Vgl. hierzu auch Georg Glasze, „Wie *Big (Spatial) Data* und *Crowdsourcing* uns alle zu Kartographen machen – ob wir wollen oder nicht“, 61-75 in diesem Band.

⁶⁸ Morgenroth berichtet, dass die Firma ‚Sociometric Solutions‘ ein Gerät entwickelt hat, mit dem Arbeitgeber ihre Belegschaft rund um die Uhr in Bezug auf Bewegungsprofile und biometrische Daten überwacht; vgl. Morgenroth, *Sie kennen dich!*, 135.

⁶⁹ Hofstetter, *Sie wissen alles*, 10.

⁷⁰ Vgl. hierzu auch Hans-Ulrich Prokosch, „Der vermessene Mensch: Wissen wir, was wir alles über uns preisgeben?“, 117-139 in diesem Band.

⁷¹ Zur anfälligen Vorhersage von Epidemien durch *Google Flu* vgl. auch Annika Richterich, „Infodemiologie – von ‚Supply‘ zu ‚Demand‘: Google Flu Trends und transaktionale Big Data in der epidemiologischen Surveillance“, in: *BD*, 333-363; 351.

körperlichen Defiziten genutzt werden können,⁷² dass sie andererseits aber auch dazu führen mögen, in Zukunft den Abschluss einer Kranken- oder Lebensversicherung für den Einzelnen zu erschweren. Die digitalen Fahrtenschreiber in den neuen autonom steuernden Fahrzeugen werden unter Umständen eine Erhöhung von Beiträgen für die Kfz-Versicherung des Kraftfahrers bedingen. Die Rekonstruktion der Waren in einem Warenkorb, die durch den Einsatz von Kundenkarten ermöglicht wird, ließe auf Suchtverhalten eines Individuums schließen, das in Zukunft – anders als bislang – für seine Erkrankung verantwortlich gemacht werden könnte. Längst führen Unternehmen bei Bewerbungen Hintergrundrecherchen zum Leben und Verhalten des Bewerbers durch (*background check*), die zur Ablehnung des Arbeitssuchenden führen, obwohl dieser alle für die Tätigkeit erforderlichen ‚berufsbezogenen‘ Voraussetzungen zu erfüllen verspricht.⁷³ Es wäre nicht auszudenken, was passieren würde, wenn diese Daten konsequent in den Dienst düsterer politischer Absichten gestellt und gegen Andersdenkende Verwendung fänden. Im Zusammenspiel mit totalitären Formen der Machtausübung würden die vorhandenen Datenbanken zur zynischen Legitimationsbasis für sozialen Ostracismus, der in die Vernichtung nicht erwünschter Institutionen und Individuen einmünden könnte. Bereits jetzt werden in den USA und auch in zunehmendem Maße in Europa auf „prädiktiven Algorithmen“ basierende Sozialprognosen getätigt,⁷⁴ mit denen die Entwicklung der Verbrechensrate vorhergesagt werden kann. Die in Steven Spielbergs Film *Minority Report* (2002) beschriebene dystopische Vision einer Gesellschaft, in der jemand für ein Verbrechen verurteilt werden kann, das er noch gar nicht begangen hat, aber mit großer Wahrscheinlichkeit noch begehen wird, offenbart längst nicht mehr die Konturen einer unscharfen Phantasie.⁷⁵ Unter dem Stichwort der ‚*crime preview*‘ werden Praktiken angewandt, die zur Vorverurteilung von Individuen und marginalisierten Gruppen führen. Zum Lernbereich der Sozial-Prognose-Algorithmen scheint die Lektion ‚*political correctness*‘ noch nicht zu zählen.⁷⁶

⁷² Vgl. Iris Pigeot und Svenja Schauer, „Bereits allgegenwärtig: Große Datensammlungen im Gesundheitswesen“, *Forschung und Lehre* 9 (2014): 708-709.

⁷³ Vgl. hierzu Morgenroth, *Sie kennen dich!*, 96-151.

⁷⁴ Vgl. hierzu Drösser, *Total Berechenbar*, 120-121; mit diesen prädiktiven Algorithmen werden ‚heiße Listen‘ (*heat lists*) zusammengestellt, auf denen sich die Namen verdächtiger Individuen finden (vgl. ebd., 118).

⁷⁵ Vgl. Steven Spielberg, *Minority Report* (USA, 2002).

⁷⁶ Vgl. auch das Kapitel „Algorithmen können diskriminieren“, in: Drösser, *Total Berechenbar*, 220-223.

Das Oxymoron ‚wahrer Scheinwelten‘: Ontologischer Wandel im Kosmos der Daten

In gewisser Weise offenbart die *Big-Data*-Welt die Konturen eines Oxymorons, das man als ‚phantastische Faktizität‘ oder in der Inversion auch als ‚faktische Phantasie‘ bezeichnen könnte, und vielleicht ist es kein Zufall, dass das Wort „postfaktisch“ von der ‚Gesellschaft für deutsche Sprache‘ zum ‚Wort des Jahres 2016‘ gewählt wurde,⁷⁷ ein Jahr, in dem Donald Trump die Präsidentschaftswahlen in den USA für sich entschied, England mit dem Brexit den Ausstieg aus der Europäischen Union beschloss und viele Demokratien das Phänomen erlebten, dass zweifelhaftes gesellschaftliche Kräfte mit konstruierten Daten politisch avancieren und Volksentscheide manipulieren.⁷⁸ Statistiken können gefälscht werden, Daten sind niemals evaluativ neutral, und die Zusammenstellung von statistischem Material mithilfe von Algorithmen generiert dezidiert virtuelle Welten.⁷⁹ Diese konstruierten Welten sind keinesfalls immun gegen Angriffe, können durch gezielte digitale Strategien in ihrer Funktionalität beeinträchtigt oder gar dekonstruiert werden:

This year will be remembered for the bad. Hardly a week passed without news of some kind of digital breach, somewhere in the world, often establishing some kind of record – for sheer scope, for novel tactics or for setting an ominous new precedent. Hackers broke into the U.S. Department of Justice, the Internal Revenue Service and likely the national Security Agency. They stole or tried to sell data from private companies including Adult FriendFinder, LinkedIn, Mail.ru and Yahoo. They leaked the confidential medical records of elite athletes Simone Biles and Serena and Venus Williams, the private photos of celebrities including comedian

⁷⁷ Vgl. Anonymous (dpa), „Postfaktisch‘ ist Wort des Jahres“, *Erlanger Nachrichten*, 10. und 11. Dezember 2016, 1.

⁷⁸ Vgl. hierzu auch Katrin Weller, „Twitter und Wahlen: Zwischen 140 Zeichen und Milliarden von Tweets“, in: *BD*, 239-257; 255. Die Einflussnahme auf Wahlen kann aber auch einer veritablen Manipulation gleichkommen, wie die Präsidentschaftswahlen im November 2016 nahelegen. Neben Donald Trump, Recep Tayyip Erdoğan und Hillary Clinton waren auch „The Hackers“ als ‚Person of the Year 2016‘ nominiert; vgl. Matt Vella, „The Hackers: They Made Vulnerability the New Normal and Took Aim at Democracy Itself“ (*Time Magazine*, 19. Dezember 2016): 62-65.

⁷⁹ Vgl. hierzu die Kapitel „Algorithmen sind nicht objektiv“ und „Algorithmen entscheiden, wer und was sichtbar ist“, in: Drösser, *Total Berechenbar*, 215-217 und 224-226.

Leslie Jones and, along the way the embarrassing password choices of a billionaire hacker named Zuckerberg.⁸⁰

An dieser Stelle konvergieren die Welten von *Big-Data*-Theorie und Postmoderne. Als vermeintlich präzise mathematische Abbildungen einer Wirklichkeit suggerieren Daten Authentizität und Verlässlichkeit; doch bereits ein Blick in die Geschichte der Berechnung der Welt beweist,⁸¹ dass man bei dieser Auffassung einem Trugschluss unterliegt:

Fakten sind ontologisch, Evidenz ist epistemologisch. Daten sind rhetorisch. Ein Datum kann auch ein Faktum sein, so wie ein Faktum eine Evidenz sein kann. Aber seit seiner ersten sprachlichen Formulierung ist die Existenz eines Datums unabhängig von jeder Berücksichtigung einer entsprechenden ontologischen Wahrheit. Wenn sich ein Faktum als falsch erweist, hört es auf, ein Faktum zu sein. Falsche Daten hingegen bleiben dennoch Daten.⁸²

Es ist ein Wesensmerkmal von Daten, dass sie zumindest eine ‚numerische‘ Distanz zur dargestellten Welt schaffen. Die auf Datenbasis entstehenden Modelle, Graphen, Illustrationen, Koordinatensysteme und Tabellen rücken den Betrachter in einen durchaus als ‚ontologisch‘ zu bezeichnenden Abstand, der die Vor- und Nachteile einer virtuellen Welt zu erkennen gibt. Wie fatal diese Erscheinung ist, lässt sich sicherlich am besten am Beispiel des militärischen Einsatzes von Datentechnik erkennen; ein militärischer Akt, der nicht mehr mit der Waffe in der Hand ausgeführt wird, sondern sich auf dem Bildschirm als Simulation durch die Verbindung von zwei Punkten darstellt, wird in seiner existenziellen Dimension vom Ausführenden vielleicht gar nicht mehr begriffen oder verarbeitet. Der tatsächliche Einsatz einer Drohne zum Zweck der Eliminierung von Terroristen erscheint im Gewand eines vermeintlich folgenlosen Computerspiels, bei dem der Spieler sogar noch essen, trinken oder Musik hören könnte. An Jean Baudrillards provokativ-zynische These, dass der Golfkrieg ein Videospiel gewesen sei, wäre hier zu erinnern.⁸³ Die Konturen von Raum und Zeit, die Faktizität des Physischen und wohl auch die Unerträglichkeit menschlichen Leids

⁸⁰ Vella, „The Hackers“, 62.

⁸¹ Vgl. zu dieser Thematik auch das informative Buch von Mainzer, *Die Berechnung der Welt*.

⁸² Rosenberg, „Daten vor Fakten“, in: *BD*, 136.

⁸³ Vgl. Hanjo Berressem, „Jean Baudrillard“, in: *Metzler Lexikon Literatur- und Kulturtheorie*, hg. Ansgar Nünning (Stuttgart und Weimar: Metzler, 1998), 39-40.

scheinen sich irrtümlicherweise hinter dem undurchdringlichen Vorhang der Daten zu verbergen. Aber virtuelle Welten erleichtern nicht nur die Anwendung von Gewalt in militärischen Kontexten, sie führen auch dazu, dass die Entscheidungsträger im Finanzbereich nicht nur aufgrund ihrer Büros in den Führungsetagen im obersten Stockwerk von Bankgebäuden Lebenswirklichkeit und Erdschwere aus dem Blick verloren zu haben scheinen. Zwischen ihrer Existenz und den Konsequenzen ihrer Entscheidungen verläuft die obskure Grenze eines undurchdringlichen Datennebels, hinter dem moralische und ethische Bedenken unsichtbar bleiben. Wenn Finanzentscheidungen auf der Grundlage von *Big-Data*-Statistiken vorgenommen werden, rücken die menschlichen Rechte, Bedürfnisse und Nöte von Millionen von Menschen nicht selten zugunsten von Profitinteressen selbsternannter Eliten in den Hintergrund.⁸⁴ Diese Tendenz zur Emergenz einer Welt des Ahumanen und Posthumanen kennzeichnet die Gegenwart in vielerlei Bereichen,⁸⁵ und es sind nicht nur die Banken und Großkonzerne, die sich in die vermeintlichen Schutzräume des Virtuellen zurückziehen:

Es mag sein, dass die Daten, die wir sammeln und übertragen, keinerlei Beziehung zu irgendeiner Wahrheit oder Wirklichkeit jenseits der Wirklichkeit haben, die wir mit ihrer Hilfe konstruieren.⁸⁶

Die Welt des Internet-Benutzers mutiert immer mehr zu einer virtuellen Existenz; ein *simulacrum* entsteht,⁸⁷ aus dessen Perspektive die Realität nun ihrerseits nur noch als schlechte ‚Kopie‘ erlebbar wird. Die im Netz vorgefundenen Informationen verkommen zu Varianten des postmodernen Konzeptes ‚großer Erzählungen‘ (*grands récits*),⁸⁸ ansatzweise in sich konsistent, aber ohne Wahrheitswert und mit hohem Anteil von Ideologie belastet. Wer im Juli des Jahres 2016 den Amoklauf in München am Bildschirm verfolgte, konnte erleben, dass die Berichtersteller der

⁸⁴ Vgl. hierzu Marc Bauders Dokumentarfilm *Master of the Universe: Der Banker* (Deutschland, 2013), in dem der ehemalige Investment-Banker Rainer Voss schildert, in welcher Distanz zur Lebenswirklichkeit finanzielle Entscheidungen getroffen werden; auch die Bankiers scheinen in einer ‚Blase‘ zu agieren.

⁸⁵ Ein gutes Beispiel für diesen Trend ist der „Hochfrequenzhandel“ an den Börsen; vgl. hierzu das Kapitel „Wie Algorithmen die Märkte beherrschen“, in: Drösser, *Total Berechenbar*, 137-150; 148.

⁸⁶ Rosenberg, „Daten vor Fakten“, in: *BD*, 155.

⁸⁷ Vgl. Berressem, „Baudrillard“, 39.

⁸⁸ Vgl. Walter Reese-Schäfer, „Jean-François Lyotard“, in: *Metzler Lexikon Literatur- und Kulturtheorie*, hg. Nünning, 334-335.

öffentlichen Medien in starkem Maße auf Informationsmaterial zurückgriffen, das aus ungesicherten Internetquellen zusammengestellt worden war. Obwohl die Datenflut zu diesem Ereignis stündlich anstieg, waren der Wahrheitswert und die Übereinstimmung mit dem ‚tatsächlichen Geschehen‘ als eher gering einzustufen. Mit Hilfe der Medien wurde eine Scheinrealität plausibel gemacht, die erst geraume Zeit später revidiert werden konnte. Immer wieder sorgen kolportierte Nachrichten im Netz für Verwirrung in den sozialen Medien. Das Verschwinden der Realität hinter der virtuellen Welt eines digitalen *simulacrum* kann aber auch im Alltagsleben auf jeder Straße bemerkt werden: Passanten nehmen weder den Straßenverkehr noch die Natur wahr, sondern halten sich beim Gehen, Autofahren oder Warten in selbstgewählten virtuellen Welten auf. Die von vielen Kritikern beschriebene ‚hermeneutische Blase‘ weitet sich naturgemäß zu einer ‚ontologischen Blase‘ aus, die sich als Isolation und Solipsismus zu erkennen gibt. Wie einzelne Forscher berichten, ziehen Kinder es vor, im Museum virtuelle Illustrationen von Tieren zu beobachten, anstatt diese *in natura* zu erleben.⁸⁹ 3D-Brillen, die sich immer mehr zu Verkaufsschlagnern entwickeln, ermöglichen eine Simulation der ‚Realität‘, die sich in zunehmendem Maße Anwendungen im Bereich Kommunikation, Sport, aber wohl auch Sexualität und Pornographie öffnen werden. Es ist zu erwarten, dass diese Erscheinungsformen des Digitalen auf lange Sicht zu einer ‚Umstrukturierung des menschlichen Gehirns‘ führen;⁹⁰ schon jetzt kann man in Schulen und Universitäten erleben, dass Wissen nicht mehr präsent ist,⁹¹ sondern dass sich die Lernenden auf die digitalen Quellen des Wissens verlassen und in der Gefahr schweben, in „digitaler Demenz“ zu versinken:⁹²

Was mit der digitalen Weiterentwicklung des Gehirns droht, ist eine Art Vergiftung unserer vom lesenden Gehirn ererbten Fähigkeiten durch einen

⁸⁹ Turkle, *Alone Together*, 3-4.

⁹⁰ Vgl. Bernard Stiegler, „Licht und Schatten im digitalen Zeitalter: Programmatische Vorlesung auf dem Digital Inquiry Symposium am Berkeley Center for New Media“, in: *BD*, 35-46; 41.

⁹¹ Die Datafizierung der Welt unterstreicht somit Platos These, dass jede Technisierung der Welt und jede Verschriftlichung von Wissen mit einer Einbuße der menschlichen Erinnerungsfähigkeit einhergehe; vgl. hierzu auch Frederica Frabetti, „Eine neue Betrachtung der Digital Humanities im Kontext originärer Technizität“, in: Reichert, *BD*, 85-102; 89.

⁹² Vgl. hierzu auch Manfred Spitzer, *Digitale Demenz: Wie wir unsere Kinder um den Verstand bringen* (München: Droemer, 2014).

Verlust des Wissens, das im Laufe von über zwei Jahrtausenden Kultur produziert wurde.⁹³

Lebensweltliche Ereignisse wie Geburt, Hochzeit, Schul- und Berufsabschlüsse scheinen nur dann Anspruch auf ‚reale Gültigkeit‘ zu haben, wenn sie – in paradoxer Verkehrung der Wirklichkeit – die höheren Weihen der ontischen Existenz durch ihre Präsenz in sozialen Netzwerken erhalten und ihre ‚Authentizität‘ durch die allgegenwärtigen ‚like‘-Buttons bestätigt worden ist. Die Flucht ins Virtuelle ist zum täglichen Habitus geworden, das Leben findet in einer Vielzahl von Fällen nur noch vor den Bildschirmen der Computer statt. Traditionelle als ‚human‘ zu bezeichnende Lebensformen wie Kommunikation von Angesicht zu Angesicht, Begrüßung, Abschied, Kranken- und Höflichkeitsbesuche, Anrufe und Benachrichtigungen degenerieren in dem Maße, in dem die digitalen Austauschmöglichkeiten und Kontaktaufnahmen an Bedeutung gewinnen. Leben ohne elektronische Hilfsmittel findet längst nur noch in Nischen statt, als Prinzip weicht das Humane vor dem Posthumanen zurück. Mit der für die nächsten Jahre prognostizierten radikal durchgeführten digitalen Vernetzung aller Gegenstände des alltäglichen Lebens im Kontext von „*smart house*“- , „*smart phone*“- und „*smart car*“-Konzepten wird die Verdrängung des Menschen aus dem Zentrum der Lebenswirklichkeit immer drastischere Formen annehmen.⁹⁴ Philosophisch betrachtet wird der ontische Status der Objekte mit Einzug in die Welt des „Internets der Dinge“ aufgewertet, oder – wie Joseph Weizenbaum und andere Kritiker bemerken – die Unschärfe zwischen Dingen/Maschinen und dem Menschen wird immer mehr an Intensität zunehmen.⁹⁵ Die ‚schöne neue Welt‘ des Posthumanen zeichnet sich durch die intendierte Suspension von Wesensmerkmalen aus, die bislang zum Beschreibungskatalog des ‚typisch‘ Menschlichen gezählt haben: Rationale Prozesse können in zunehmenden Maße von intelligenten Maschinen geleistet werden, Erinnerungen werden auf künstliche Speicherorte transferiert, Entscheidungen

⁹³ Stiegler, „Licht und Schatten im digitalen Zeitalter“, 41-42.

⁹⁴ Im Laufe des Jahres 2016 wurde das „*Dash-System*“ zum Bestseller bei *Amazon*; durch Drücken von Knöpfen, die man an Kühlschränken oder Waschmaschinen anbringt, werden automatisch Lebensmittel oder Waschpulver nachgeliefert. Aufgrund der *Big-Data*-Systeme weiß die die Firma vor dem Verbraucher, was dieser benötigen wird; vgl. Anonymous (dpa), „Einkauf auf Knopfdruck: Amazon bringt Dash-Bestellsystem nach Deutschland“, *Erlanger Nachrichten* (1. September 2016), 21.

⁹⁵ Die Entwicklung führt von der Turingmaschine über Joseph Weizenbaums „*ELIZA*“-Programm zu den ‚bots‘ in der aktuellen Internetszene; vgl. auch Turkle, *Alone Together*, 23-24.

an Supercomputer übereignet, die besser als jedes menschliche Wesen mit dem hohen Komplexitätsgrad unterschiedlichster Informationen umzugehen ‚verstehen‘. In der sich abzeichnenden posthumanen Welt erscheint der *homo rationalis* als Fremd- und Störkörper, dessen Einfluss auf das Funktionieren des Gesamtsystems minimalisiert werden muss, um den algorithmischen Ablauf der Prozesse nicht zu gefährden. Hatte sich das individuelle Subjekt im Laufe der europäischen Aufklärung zum verantwortungsvoll handelnden Wesen emanzipiert, so leistet es im Angesicht der *Big-Data*-Welt einen philosophischen Offenbarungseid, indem es – bedingt durch die überlegene Macht der Rechner – nun nicht länger mehr wagt, selbst zu denken. Es wird zum Opfer jener Maschinen, die es – zumindest bevor diese eine autonome algorithmische Evolution etablierten – selbst erschuf:

Die Entmündigung des immer bequemer werdenden Bürgers, der seine Gedankenarbeit herunterfährt und den Kühlschrank die Einkäufe erledigen lässt, mündet in das totale Geführtwerden und damit in die Unfreiheit, prophezeit der ehemalige Landesdatenschutzbeauftragte Joachim Wahlbrink. „Die Industrie scheint Gefallen an dieser Vision zu haben, weil man davon ausgeht, dass der Bürger seine Helfer, die langsam, aber sicher seine Herrscher werden, immer wieder in Anspruch nehmen möchte. Wir werden eine Umwertung und Entwertung all dessen erleben, was wir bisher an Selbstbestimmung, Steuerung und Eigenverantwortung hatten.“⁹⁶

Algorithmische Entmündigungen: Ethik im Zeitalter der Datenflut

Die Welt der *Big-Data*-Maschinen stellt somit eine „sozialethische Herausforderung“ ersten Ranges dar.⁹⁷ Das wohl wichtigste ethische Anliegen der näheren Zukunft wird sein, ob der Mensch länger als Instanz von folgenreichen Entscheidungen in den Bereichen von Wirtschaft, Politik, Medizin und Gesellschaft, um nur einige der betroffenen Felder zu nennen, bleiben kann.⁹⁸

⁹⁶ Morgenroth, *Sie kennen dich!*, 206-207.

⁹⁷ Alexander Filipović, „Algorithmen und Roboter – Die Datafizierung der Welt: Eine ethische Vermessung des digitalen Wandels“, *Communicatio Socialis* 48.1 (2015): 6-15; 9.

⁹⁸ Zu den wichtigsten Fragestellungen des ethischen Denkens vgl. Michael Quante, *Einführung in die Allgemeine Ethik* (Darmstadt: Wissenschaftliche Buchgesellschaft, 2011 [2003]), 24-39.

Vor dem Siegeszug atheistischer und agnostischer Ansichten im späten 19. und im 20. Jahrhundert war es in den Zeitaltern zuvor durchaus üblich, bei Entscheidungen höchster Dringlichkeit eine ‚höhere Instanz‘ um Rat zu fragen. Gebet und Meditation sollten das entscheidende Subjekt darin unterstützen, das ‚Richtige‘ zu tun, da der gläubige Mensch auf die Überlegenheit des göttlichen Rats vertraute. Die divinen Eigenschaften Omnipräsenz, Omnipotenz und Omniszienz ließen jedoch die Entscheidung, sich einer höheren Macht anzuvertrauen, bei genauerem Hinsehen als eine Schein-Entscheidung erkennen, da sie – vor allem im Kontext bekannter Prädestinationslehren – bereits im göttlichen Heilsplan vorgesehen waren. In gewisser Weise waren also ‚freie‘ Entscheidungen nur Realisierungen von Entschlüssen, die längst auf höherer metaphysischer Ebene beschlossen worden waren. Diese Hinwendung zur Metaphysik ging aber häufig mit einer Entlastung des Einzelnen einher, der auf das Wirken einer höheren Macht verweisen konnte.

Zwar gehört es sicherlich zu den Akten der Blasphemie, die Welt der *Big Data* mit göttlichen Attributen auszustatten, aber der Gedanke liegt doch nahe, dass die Supercomputer eine Simulation von Allmacht, Allwissenheit und Allgegenwart vortäuschen, wenn sie sie nicht gar schon tatsächlich erreicht haben. Die pseudo-divine Omnipräsenz zeitgenössischer Datenwelt lässt nicht nur Katholiken an die frühkindlichen Ermahnungen denken, Gott sehe alles, wisse alles, könne alles und werde sich alles merken, was man tue. Auch jedes Haar des zukünftigen Sünders habe er ja bereits gezählt.

Der neue ‚Baum der Erkenntnis‘ steht nun im digitalen Garten Eden, betreut von ‚Apple‘ und anderen; jetzt lautet das Gebot, man möge von ihm essen, wann immer man es wünsche, und so viel man wolle, der Gott des Digitalen verlange es. Aber wohl auch hier endet die Geschichte mit der Vertreibung aus dem Paradies, weil dieser Baum mit jedem Ernten seiner Früchte wächst, geradezu bis ins Unendliche Wurzeln und Äste sprießen lässt, bis für das menschliche Wesen kein Platz mehr in diesem Garten Eden bleibt.

In seiner analytischen Schärfe könnte der kategorische Imperativ Kants sicherlich in die Algorithmen von maschinellen Entscheidungsträgern überführt werden; selbst hochkomplexe Bewertungsszenarien ethischer Dilemmata aus der Perspektive der deontologischen oder utilitaristischen Ethik wären in Sekundenbruchteilen durchführbar,⁹⁹ aber die

⁹⁹ Vgl. ebd., *Einführung in die Allgemeine Ethik*, 130-142.

entscheidende Grundfrage bleibt: Werden ethische Beurteilungen ganz der kalten Logik von Maschinen übereignet werden?¹⁰⁰ Sollen Operationen auf der Grundlage von *Big-Data*-Berechnungen vorgenommen werden? Reicht die datenanalytische Prognose von wahrscheinlichen Straftaten aus, um einen ‚Delinquenten‘ in Haft zu nehmen? Gilt vor Gericht in Zukunft noch das Prinzip „Aussage gegen Aussage“ oder muss es dem Slogan „Datenevidenz vor Aussage“ weichen? Gibt der Computer den statistisch begründeten Befehl zum ‚finalen Rettungsschuss‘, wenn er die Bewegungsdaten der Täter anlässlich einer Geiselnahme analysiert hat? Wird der Innenminister entmündigt, wenn die europäische ‚Union der künstlichen Intelligenz‘ beschließt, die europäischen Grenzen gegen Asylsuchende zu schließen?¹⁰¹ Selbst die alltägliche Straßenszene bietet im Angesicht ‚selbstfahrender Autos‘ ethische Probleme in großer Fülle:

Es gibt gravierende ethische Fragestellungen, wie diese Autos programmiert werden sollen. Es kann immer zu fatalen Situationen kommen, und die Algorithmen, die Rechenprozesse müssen mit Schadensszenarien umgehen. Die Frage ist: Wie dürfen sie das? Elementare Standards des demokratischen Verfassungsstaates dürfen dabei nicht verletzt werden. Kameras können identifizieren, ob eine ältere oder eine jüngere Person auf der Straße steht. Eine Programmierung könnte dann lauten: immer den älteren Menschen töten. Das darf natürlich nicht sein. Ein Grundsatz könnte sein: immer den Schaden möglichst gering halten. Leichter gesagt als getan: da kommen Ihnen zwei Motorradfahrer entgegen, einer mit Helm, einer ohne. Der ohne Helm stirbt beim Unfall wahrscheinlich eher. Aber soll er jetzt dafür belohnt werden, dass er sich verkehrswidrig verhält, weil der andere Fahrer gerammt wird, der wahrscheinlicher, aber schwerverletzt, überlebt? Da bekommen wir Riesenprobleme; wir müssen uns deshalb erst kulturell an diese autonomen Systeme gewöhnen.¹⁰²

¹⁰⁰ Vor allem aus der Sicht des Christentums wären derartige Entscheidungen ein Affront gegen das „Personenprinzip“, das als „Kern der christlichen Ethik“ bezeichnet werden darf; vgl. Filipović, „Algorithmen und Roboter“, 13.

¹⁰¹ Vgl. hierzu auch die Fiktion des artifiziellen Politikers ‚Ai‘ in Hofstetter, *Das Ende der Demokratie*, 19-33.

¹⁰² Alexander Jungkunz, Michael Husarek, Dieter Schwab und Michael Kasperowitsch, „Wenn Autos selbst fahren – was tun sie bei Unfällen?: Interview mit Peter Dabrock“, *Erlanger Nachrichten* (12. November 2016): 6.

Algorithmen verändern die Welt; vielleicht wird eine Maschine die Entscheidung auf der Grundlage einer Myriade von vergleichbaren Einzelfällen treffen, aber ob eine solche Entscheidungsfindung ethisch vertretbar ist, muss – zumindest noch zum jetzigen Zeitpunkt – von Menschen verantwortet werden.¹⁰³ All diese Einzelentscheidungen bieten somit genügend Anlass zu ethischer Reflexion. Noch gravierender wiegen aber die Fragen, die sich auf die Machtentfaltung der *Big-Data*-Welt beziehen. Wie jedes Instrument oszilliert auch der allgegenwärtige Algorithmus zwischen ‚gut‘ und ‚böse‘ und schwebt in der Gefahr, missbraucht zu werden. Die zunehmende Symbiose von Datenmacht und Informationskapitalismus ist wohl kaum noch zu stoppen. Aber wie ist es zu verhindern, dass diese Macht in die Hände totalitärer Regierungsformen übergeht? Unter der Herrschaft eines diktatorischen Regimes werden die Datenfluten zu Instrumenten der totalen Auslöschung, zu Werkzeugen von individueller und nationaler Erpressung und zur existenziellen Bedrohung des Lebens von Millionen.

Noch gibt es Hoffnung, dass es sich der Mensch nicht gefallen lassen wird, umfassend berechnet und vollkommen berechenbar zu werden. Nicht nur der Schlaf der Vernunft, auch die Vernunft selbst, scheint Ungeheuer zu gebären, zumindest wenn man sie mit der „instrumentellen Vernunft“ gleichsetzt. Diese brachiale Erscheinungsform des ‚Vernünftigen‘ gilt es zu bekämpfen. Zwischen dem Schlaf der Vernunft und dem Größenwahn des Rationalen muss es auch weiterhin ein Drittes geben: Dort sollte das Prinzip des Humanen auch förderhin Heimat finden.

Schluss

Die Welt der Datenfluten darf also nicht länger ausschließlich die Domäne von Informationstheoretikern, Mathematikern und Statistikern bleiben. Sie geht jeden einzelnen an, selbst dann, wenn er keinen Computer benutzt. Auf dem Plan steht nichts Geringeres als die menschliche Existenz, ihre Freiheit und ihre Würde. Es ist das Anliegen dieses Bandes, zum Verständnis der ethischen Herausforderung der Menschheit durch die Welt der *Big-Data*-Vernetzungen beizutragen. Niemand möge sich in diesem Kontext mit der Voltaireschen Formel „Il faut cultiver notre

¹⁰³ Wahrscheinlich kann im Moment noch niemand vorhersagen, ob auch ‚metaethische‘ Fragen zum Gegenstand algorithmischer Problemlösungsstrategien zählen werden.

jardin“ zufriedengeben,¹⁰⁴ denn die unkontrollierte Datenflut wird sich zu einem Tsunami ausweiten, vor dem kein noch so kultivierter Garten sicher ist.¹⁰⁵

Rudolf Freiburg

¹⁰⁴ Voltaire, *Candide*, in: *Romans et Contes*, ed. René Pomeau (Paris: Garnier-Flammarion, 1966), 259.

¹⁰⁵ Für die Hilfe bei Recherchen und der Erstellung des Manuskriptes gilt mein Dank Barbara Gabel-Cunningham, Evelin Werner, Lukas Pracht und Susanne Wagner.

Literaturverzeichnis

- Anderson, Chris. „Das Ende der Theorie: Die Datenschwemme macht wissenschaftliche Methoden obsolet“. In: *Big Data: Das neue Versprechen der Allwissenheit*. Hg. Heinrich Geiselberger und Tobias Moorstedt. Frankfurt a.M.: Suhrkamp, 2013. 124-130.
- Anonymous (dpa). „Einkauf auf Knopfdruck: Amazon bringt Dash-Bestellsystem nach Deutschland“. *Erlanger Nachrichten* (1. September 2016): 21.
- Anonymous (dpa). „‚Postfaktisch‘ ist Wort des Jahres“. *Erlanger Nachrichten* (10. und 11. Dezember 2016): 1.
- Bacon, Francis. „The Advancement of Learning“ (1605). In: *The Advancement of Learning and New Atlantis*. Ed. Arthur Johnston. Oxford: Clarendon Press, 1974. 3-200.
- Bacon, Francis. „Novum Organum“ (1620). In: *The Works of Francis Bacon*. Ed. James Spedding, Robert Leslie Ellis and Douglas Denon Heath. 14 vols.; vol. 1. London: Longman et al., 1858. 149-365.
- Berressem, Hanjo. „Jean Baudrillard“. In: *Metzler Lexikon Literatur- und Kulturtheorie*. Hg. Ansgar Nünning. Stuttgart und Weimar: Metzler, 1998. 39-40.
- Berry, David M. „Die Computerwende: Gedanken zu den Digital Humanities“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 47-64.
- Blumenberg, Hans. *Der Prozess der theoretischen Neugierde*. Frankfurt a.M.: Suhrkamp, 1973.
- Boellstorff, Tom. „Die Konstruktion von Big Data in der Theorie“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 105-131.
- Buhmann, Joachim M. „Eine Revolution unseres Weltbildes: Über die informationstechnologische und erkenntnistheoretische Herausforderung von *Big Data*“. *Forschung und Lehre* 9 (2014): 700-702.
- Burgess, Jean und Axel Bruns. „Twitter-Archive und die Herausforderungen von ‚Big Social Data‘ für die Medien- und Kommunikationswissenschaft“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 191-202.
- Buschauer, Regine. „(Very) Nervous Systems: Big Mobile Data“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 405-436.

- Drösser, Christoph. *Total Berechenbar: Wenn Algorithmen für uns entscheiden*. München: Hanser, 2016.
- Filipović, Alexander. „Algorithmen und Roboter – Die Datafizierung der Welt: Eine ethische Vermessung des digitalen Wandels“. *Communicatio Socialis* 48.1 (2015): 6-15.
- Forgó, Nikolas. „Wer bin ich schon?: Die Datafizierung des Lebens und das Datenschutzrecht“. *Forschung und Lehre* 9 (2014): 704-705.
- Frabetti, Frederica. „Eine neue Betrachtung der Digital Humanities im Kontext originärer Technizität“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 85-102.
- Freiburg, Rudolf. *Autoren und Leser: Studien zu Intentionalität literarischer Texte*. Frankfurt a.M. et al.: Lang Verlag, 1985.
- Freud, Sigmund. „Eine Schwierigkeit der Psychoanalyse“. In: *Abriß der Psychoanalyse: Einführende Darstellungen*. Frankfurt a.M.: Fischer Taschenbuch Verlag, 1998 (1917). 187-194.
- Galloway Alexander und Eugene Thacker, „Protokoll, Kontrolle und Netzwerke“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 289-311.
- Grass, Günter. „Ohnmacht“. In: *Vonne Endlichkait [sic]*. Göttingen: Steidl Verlag, 2015. 13.
- . „Allabendliches Gebet“. In: *Vonne Endlichkait [sic]*. Göttingen: Steidl Verlag, 2015. 14.
- Hirsch Eric. D. „The Concept of Genre“. In: *Validity in Interpretation*. New Haven: Yale University Press, 1971. 68-126.
- Hofstetter, Yvonne. *Das Ende der Demokratie: Wie die künstliche Intelligenz die Politik übernimmt und uns entmündigt*. München: C. Bertelsmann, 2016.
- . *Sie wissen alles: Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen*. München: Bertelsmann, 2014.
- Horkheimer, Max und Theodor W. Adorno. *Dialektik der Aufklärung: Philosophische Fragmente*. Frankfurt: Fischer Verlag, 2012 (1944).
- Jungkunz, Alexander, Michael Husarek, Dieter Schwab und Michael Kasperowitsch. „Wenn Autos selbst fahren – was tun sie bei Unfällen?‘: Interview mit Peter Dabrock“. *Erlanger Nachrichten* (12. November 2016): 6.
- Kalbitzer, Jan. *Digitale Paranoia: Online bleiben, ohne den Verstand zu verlieren*. München: Beck, 2016.

- Kant, Immanuel. *Beantwortung der Frage: Was ist Aufklärung?* Frankfurt a.M.: Neuer Frankfurter Verlag, 1900.
- Lewis, Pericles. „On or about December, 1910“. In: *The Cambridge Introduction to Modernism*. Cambridge: Cambridge University Press, 2010 (2007). 86-94.
- Mainzer, Klaus. *Die Berechnung der Welt: Von der Weltformel zu Big Data*. München: Beck, 2014.
- . „Die Berechnung der Welt: Können Big Data-Ergebnisse Theorie und Beweis ersetzen?“. *Forschung und Lehre* 9 (2014): 696-699.
- Manovich, Lev. „Trending: Verheißungen und Herausforderungen der Big Social Data“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 65-83.
- Mayer-Schönberger, Viktor. „Neue Erkenntnisse über die Wirklichkeit: Wie wir die Diktatur der Daten verhindern können“. *Forschung und Lehre* 9 (2014): 706-707.
- und Kenneth Cukier. *Big Data: Die Revolution, die unser Leben verändern wird*. Übers. Dagmar Mallett. München: Redline Verlag, 2013.
- Morgenroth, Markus. *Sie kennen dich! Sie haben dich! Sie steuern dich! Die wahre Macht der Datensammler*. München: Knaur, 2016.
- Pariser, Eli. *The Filter Bubble: What the Internet is Hiding from You*. New York: Penguin, 2011.
- Partington, Angela. Ed. *The Oxford Dictionary of Quotations*. London et al.: Oxford University Press, 1992 [1941].
- Pigeot, Iris und Svenja Schauer. „Bereits allgegenwärtig: Große Datensammlungen im Gesundheitswesen“. *Forschung und Lehre* 9 (2014): 708-709.
- Popper, Karl. *Logik der Forschung*. Berlin: Akademie Verlag, 2004.
- Quante, Michael. *Einführung in die Allgemeine Ethik*. Darmstadt: Wissenschaftliche Buchgesellschaft, 2011 [2003].
- Reese-Schäfer, Walter. „Jean-François Lyotard“. In: *Metzler Lexikon Literatur- und Kulturtheorie*. Hg. Ansgar Nünning. Stuttgart und Weimar: Metzler, 1998. 334-335.
- Reichert, Ramón. „Einführung“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 10-31.
- . „Facebooks Big Data: Die Medien- und Wissenstechniken kollektiver Verdatung“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 437-452.

- Richterich, Annika. „Infodemiologie – von ‚Supply‘ zu ‚Demand‘: Google Flu Trends und transaktionale Big Data in der epidemiologischen Surveillance“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 333-363.
- Rogers, Richard. „Nach dem Cyberspace: Big Data, Small Data“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 173-187.
- Rosenberg, Daniel. „Daten vor Fakten“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 133-156.
- Schirmacher, Frank. *Payback: Warum wir im Informationszeitalter gezwungen sind zu tun, was wir nicht tun wollen, und wie wir die Kontrolle über unser Denken zurückgewinnen*. München: Blessing, 2009.
- Schneier, Bruce. *Data und Goliath: Die Schlacht um die Kontrolle der Welt*. Übers. Sigrid Schmid und Claudia Van Den Block. München: Redline Verlag, 2015.
- Simmel, Georg. „Die Großstädte und das Geistesleben“. In: *Die Großstadt: Vorträge und Aufsätze zur Städteausstellung*. Hg. Thomas Petermann. Dresden: Zahn & Jaensch, 1903. 185-206.
- Spitzer, Manfred. *Digitale Demenz: Wie wir unsere Kinder um den Verstand bringen*. München: Droemer, 2014.
- Stiegler, Bernard. „Licht und Schatten im digitalen Zeitalter: Programmatische Vorlesung auf dem Digital Inquiry Symposium am Berkeley Center for New Media“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 35-46.
- Turkle, Sherry. *Reclaiming Conversation: The Power of Talk in a Digital Age*. New York: Penguin, 2015.
- . *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books, 2011.
- Vella, Matt. „The Hackers: They Made Vulnerability the New Normal and Took Aim at Democracy Itself“. *Time Magazine* (19. Dezember 2016): 62-65.
- Voltaire. *Romans et Contes: Candide*. Ed. René Pomeau. Paris: Garnier-Flammarion, 1966.
- Weitin, Thomas. „Thinking Slowly: Literatur lesen unter dem Eindruck von Big Data“. *LitLingLab Pamphlet #1* (März 2015): 1-18.

- Weller, Katrin. „Twitter und Wahlen: Zwischen 140 Zeichen und Milliarden von Tweets“. In: *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Hg. Ramón Reichert. Bielefeld: Transcript, 2014. 239-257.
- Welzer, Harald. *Die smarte Diktatur: Der Angriff auf unsere Freiheit*. Frankfurt a.M.: Fischer Verlag, 2016.
- Weth, Thomas. „Faszination Unendlichkeit: Über die Struktur des Unendlichen“. *Unendlichkeit: Fünf Vorträge*. Hg. Rudolf Freiburg. Erlangen: FAU Press, 2016. 105-138.

Film

- Bauders, Marc. *Master of the Universe: Der Banker*. Deutschland, 2013.
- Spielberg, Steven. *Minority Report*. USA, 2002.

Internetquellen

- Anonymous, „Für digitale Grundrechte! – Ein Bürgerrecht: Die ZEIT-Stiftung unterstützt eine Digitalcharta für Europa“
<http://www.zeit.de/digital/internet/2016-11/europaeische-charta-digitale-grundrechte-hintergrund>
- „Person of the Year: You“. *Time Magazine* (25 December 2006; 1 January 2007) <http://content.time.com/time/covers/0,16641,20061225,00.html>
- Robot-Pianist Teo Tronico ‚live‘ at Berlin Philharmonie
<https://www.youtube.com/watch?v=U5rKjrQ3JoE>

Algorithmen, Datenkraken und gläserne Bürger: *Big Data* als Herausforderung für die Gesellschaftsanalyse

SEBASTIAN M. BÜTTNER

Abstract: Die Digitalisierung der technischen Infrastrukturen und der rasante Ausbau von Technologien zur umfassenden Erhebung und Verarbeitung von Daten aller Art stellen nicht nur die Gesellschaft, sondern auch die Gesellschaftsanalyse vor ganz neue Herausforderungen. Unternehmen wie *Google* und *Facebook*, Markt- und Meinungsforschungsinstitute, Datenhändler und auch Geheimdienste dringen immer mehr auf das Gebiet der Gesellschaftsanalyse vor und streben danach, die Verhaltensweisen von sozialen Gruppen und auch von Einzelpersonen in ihrer Gesamtheit zu erfassen und zu analysieren. Bemerkenswert ist daran nicht nur der grenzenlose Datenbeschaffungsseifer, sondern vor allem auch die zunehmende Automatisierung der Datenanalyse und der damit verbundene Anspruch an allumfassende Information und Transparenz. Ausgehend von einer kurzen Darstellung der geschichtlichen Entwicklung der Bevölkerungsanalyse geht dieser Beitrag aktuellen Trends der technisch hochgerüsteten Datenanalyse nach, die derzeit unter dem Schlagwort ‚*Big Data*‘ eine breite öffentliche Aufmerksamkeit erfahren. Der Beitrag schließt mit der Diskussion einiger problematischer Aspekte der gegenwärtigen digitalen Datenflut.

Die Digitalisierung der Datenanalyse als Herausforderung für die Soziologie

Die Digitalisierung und die Expansion von Technologien zur Analyse von immer größeren Datenmengen – heute weithin bekannt unter dem Schlagwort ‚*Big Data*‘ – fordern die Soziologie in ihrer Rolle als Gesellschaftswissenschaft in vielfältiger Weise heraus. Denn für die Soziologie stellt sich zum einen ganz grundsätzlich die Frage, wie mit einem Thema umzugehen ist, das in der Öffentlichkeit bereits intensiv diskutiert wird und normativ stark aufgeladen ist. Die Digitalisierung und die damit verbundenen neuen technischen Entwicklungen werden einerseits von Vielen emphatisch begrüßt und mit großem Entwicklungseifer und Fortschrittsoptimismus gefeiert. Andererseits – dies verdeutlichen die im Titel dieses Beitrags genannten Begriffe „Algorithmen“, „Datenkraken“ und „gläserne Bürger“ – rufen die Enthüllungen der neuen Spionagetechniken von Geheimdiensten wie *PRISM* oder *TEMPORA* sowie die Berichte über den Datensammlungseifer von Konzernen wie *Facebook* oder von Banken und Versicherungen auch ein großes Unbehagen bei den Bürgern hervor

und ein Gefühl der Rat- und Machtlosigkeit angesichts der neuen technologischen Entwicklungen.¹ Für die Soziologie beziehungsweise für die wissenschaftliche Gesellschaftsanalyse heißt es hier, eine eigene Position im allgemeinen Getöse der öffentlichen Aufregung zu finden.

Eine zweite Herausforderung, die sich für die Soziologie angesichts fortschreitender Digitalisierung stellt, ist die adäquate Einschätzung der Art, der Geschwindigkeit und der Radikalität der aktuellen Veränderungen. Hier stellt sich vor allem die Frage, inwieweit die Gesellschaftsanalyse selbst Schritt halten kann mit dem technologischen Wandel und ob die zur Verfügung stehenden theoretischen Modelle und Kategorien ausreichen, um die aktuellen Wandlungsprozesse zu verstehen: Bedeutet die Digitalisierung tatsächlich einen ähnlich einschneidenden Wandel wie die Industrialisierung im 19. Jahrhundert, wie es von manchen Kommentatoren behauptet wird?² Haben wir es heute tatsächlich mit einem radikalen Umbruch und einem epochalen Wandel zu tun? Oder fügt sich der aktuelle Trend zur Digitalisierung von Arbeit und Gesellschaft in die mittlerweile über hundertjährige Geschichte der Elektrifizierung von technischen Infrastrukturen und Kommunikationsmedien ein? – Es ist schwierig, Trendaussagen über gesellschaftliche Wandlungsprozesse zu treffen, wenn die Veränderungen selbst gerade in vollem Gange sind. Fakt ist aber auch: Technologische Entwicklungen sind Teil der menschlichen Kultur und haben insofern einen erheblichen Einfluss auf die Art der Vergesellschaftung und auf die Formen sozialer Interaktion.³ Die Frage ist, wie der Wandel in einem größeren Rahmen gesellschaftlicher Entwicklung einzuordnen ist und inwieweit auch hier Vorsicht geboten ist vor allzu starker Übertreibung.

Die dritte große Herausforderung ist schließlich technischer und methodologischer Art. Denn die bisherigen Routinen der empirischen Sozialforschung und der wissenschaftlichen Gesellschaftsanalyse werden durch neue Technologien,⁴ Verfahren und Möglichkeiten der Datenana-

¹ Vgl. Frank Schirmacher, hg., *Technologischer Totalitarismus: Eine Debatte* (Berlin: Suhrkamp, 2015).

² Vgl. Jeremy Rifkin, *Die dritte industrielle Revolution: Die Zukunft der Wirtschaft nach dem Atomzeitalter* (Frankfurt: Campus, 2011) und ders., *Die Null-Grenzkosten-Gesellschaft: Das Internet der Dinge, kollaboratives Gemeingut und der Rückzug des Kapitalismus* (Frankfurt: Campus, 2014).

³ Vgl. Werner Rammert, hg., *Technik und Sozialtheorie* (Frankfurt a.M.: Campus, 1998).

⁴ Die empirische Sozialforschung ist ein Kerngeschäft der Soziologie. Hier steht der Soziologie ein ganzer Strauß an unterschiedlichen Forschungsmethoden zur Verfügung, die üblicherweise in qualitative und quantitative Verfahren unterschieden

lyse stark herausgefordert.⁵ Dies hat zum einen mit den gesteigerten Objektivitäts- und Erklärungsversprechen der neuen *Big-Data*-Forschung zu tun. Es hat zum anderen auch damit zu tun, dass die technischen Neuerungen größtenteils außerhalb der Soziologie und auch außerhalb der wissenschaftlichen Sphäre im engeren Sinne stattfinden, nämlich überwiegend in der Informatik und in der Mathematik sowie in privaten Unternehmen und Konzernen. Die wissenschaftliche Welt droht somit den Anschluss an die neuesten Entwicklungen im Bereich der Datenanalyse zu verpassen. Damit geht jedoch auch die politisch brisante Frage einher, was es bedeutet, wenn ein immer größerer Teil der empirischen Sozialforschung und der Gesellschaftsanalyse von kommerziellen Akteuren und aus ökonomischen Interessen vorangetrieben wird.

Dies sind zentrale Herausforderungen, die sich für die Soziologie im Zuge der fortschreitenden Digitalisierung der Gesellschaft und der Gesellschaftsanalyse stellen. Im Mittelpunkt der nachfolgenden Ausführungen steht jedoch nicht die ausführliche Diskussion jeder dieser Herausforderungen oder eine Diskussion von möglichen Lösungswegen. Der vorliegende Beitrag soll vielmehr einen Eindruck von den Veränderungen vermitteln, die sich gegenwärtig auf dem Gebiet der Gesellschaftsanalyse durch *Big Data* vollziehen. Dabei werden zunächst einige Grundlagen der Gesellschaftsanalyse und wichtige Entwicklungsschritte in der Geschichte der ‚Vermessung der Gesellschaft‘ nachvollzogen. Denn historisch betrachtet fügt sich der Wandel hin zu *Big Data* nahtlos in die Geschichte einer immer weitergehenden ‚Vermessung‘ und rationalen Durchdringung der Welt ein.⁶ Mit Blick auf einige neue Entwicklungen auf dem Gebiet der Gesellschaftsanalyse, die mit der Digitalisierung einhergehen, wird jedoch im hinteren Teil des Beitrags deutlich gemacht, dass wir derzeit durchaus einen Umbruch mit sehr weitreichenden gesellschaftlichen

werden. Vgl. Rainer Schnell, Paul B. Hill und Elke Esser, *Methoden der empirischen Sozialforschung* (München: Oldenbourg, 2011).

⁵ Vgl. Jochen Mayerl, „Bedeutet ‚Big Data‘ das Ende der sozialwissenschaftlichen Methodenforschung?“, *Soziopolis: Gesellschaft beobachten*. (online).

⁶ Das Motiv der ‚Vermessung‘ ist hier ganz bewusst an den Titel des Romans *Die Vermessung der Welt* von Daniel Kehlmann angelehnt. Kehlmann stellt in diesem Buch in fiktiver Form das Leben des Naturforschers Alexander von Humboldt und des Mathematikers Carl Friedrich Gauß dar, zweier Pioniere der Aufklärung und der modernen Naturwissenschaften. Vgl. Daniel Kehlmann, *Die Vermessung der Welt* (Reinbek: Rowohlt, 2005). In der Soziologie wurden der besondere Geist der Moderne und die Tendenz zur unablässigen Quantifizierung und Berechnung der Welt von Max Weber einst auf den Begriff der ‚Rationalisierung‘ gebracht. Von ihm stammt auch das Motiv der ‚Entzauberung‘ der Welt durch fortschreitende Rationalisierung.

Konsequenzen erleben. Die neuen Technologien erlauben zumindest ein Ausmaß an Datensammlung und -analyse, das in der Geschichte bisher seinesgleichen sucht.

Die Vermessung der Gesellschaft: Geschichte der Gesellschaftsanalyse

Wenn wir das Phänomen *Big Data* und insbesondere den Wandel, der durch *Big Data* auf dem Gebiet der Gesellschaftsanalyse angestoßen wird, begreifen möchten, ist es zunächst lohnenswert, einen Blick in die Vergangenheit zu werfen und sich die Geschichte der Gesellschaftsanalyse genauer vor Augen zu führen. Dabei stellen wir fest, dass es immer schon Versuche gegeben hat, genauere Informationen über ökonomische und gesellschaftliche Zusammenhänge und vor allem auch über die Bevölkerung in einem bestimmten Gebiet zu sammeln. Von den uns überlieferten Aufzeichnungen aus dem Alten Ägypten oder aus dem Alten Mesopotamien wissen wir etwa,⁷ dass es weit vor unserer Zeit bereits relativ systematische und zum Teil auch sehr ausgeklügelte Versuche der Sammlung und Aufzeichnung von wirtschaftlichen und bevölkerungsspezifischen Informationen („Daten“) gab.⁸ Auch finden sich in historischen Quellen immer wieder Hinweise auf größere Bevölkerungsschätzungen und Volkszählungen. Eine der frühesten überlieferten Volkszählungen ist den meisten von uns bestens bekannt. Denn sie ist in der Bibel im *Neuen Testament* an prominenter Stelle im „Lukas-Evangelium“ als Auftakt zur berühmten Erzählung von der Geburt Jesu Christi niedergeschrieben:

Es begab sich aber zu der Zeit, dass ein Gebot von dem Kaiser Augustus ausging, dass alle Welt geschätzt würde. Und diese Schätzung war die allererste und geschah zu der Zeit, da Cyrenius Landpfleger von Syrien

⁷ Vgl. Jan Assmann, *Das kulturelle Gedächtnis: Schrift, Erinnerung und politische Identität in frühen Hochkulturen* (München: Beck, 2013).

⁸ Unter ‚Daten‘ lassen sich prinzipiell viele verschiedene Arten von Informationen über die soziale und die natürliche Umwelt verstehen. Über ‚Daten‘ erschließen wir uns die Welt. Sie geben uns Hinweise über die Beschaffenheit der Welt. Sehr häufig treten Daten in Ziffernform auf. Daten können aber viele weitere symbolische Formen annehmen oder als Bilder, Geräusche, elektromagnetische Signale verarbeitet und gespeichert werden. Vgl. Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences* (London: Sage, 2014), 1-12.

war. Und jedermann ging, dass er sich schätzen ließe, ein jeglicher in seine Stadt.⁹

Dieses Zitat ist nicht nur aufgrund seiner besonderen historischen Bedeutung bemerkenswert. Im Hinblick auf unsere Diskussion über *Big Data* in der Gesellschaftsanalyse belegt es eindrucksvoll, dass es den Anspruch einer möglichst umfassenden Erhebung von Bevölkerungsdaten anscheinend schon sehr lange gibt – zumindest seit Beginn unserer Zeitrechnung. Das Zitat ist auch Beleg dafür, dass die Bemühungen um eine möglichst genaue Erhebung und Erfassung der Bevölkerungszahlen auch damals bereits mit einem ähnlichen Allumfassendheitsanspruch formuliert wurden („dass alle Welt geschätzt würde“), wie es heute in den Diskursen über *Big Data* häufig zu vernehmen ist. Darüber hinaus zeigt das Beispiel der Volkszählung von Kaiser Augustus, dass die Erhebung von Bevölkerungsdaten unterschiedlicher Art anscheinend schon sehr lange zum Standardrepertoire des Regierens gehört und eine wichtige Herrschaftstechnologie darstellt.¹⁰ Die Formel „Wissen ist Macht“,¹¹ die eigentlich erst zu Beginn der Neuzeit und im Zuge des Aufstiegs der modernen Wissenschaften und der modernen bürokratischen Herrschaft eine besondere Bedeutung erlangt hat, galt in gewissem Sinne wohl auch bereits für die Herrschaftsordnungen der Antike. Denn je mehr Wissen und Information über die Zusammensetzung der Bevölkerung in einem Herrschaftsgebiet vorliegen, desto besser lässt sich eine Herrschaftsordnung stabilisieren und verwaltungsmäßig strukturieren. Dies diente nicht nur als Grundlage für die Erhebung von Steuern und Abgaben, sondern auch für die Berechnung der wehrfähigen Bevölkerung und des Reservoirs an potenziellen Soldaten.¹²

Die Methoden der ökonomischen Buchführung und der Bevölkerungszählung waren in der Antike zwar schon recht weit entwickelt, die Geburt der modernen Bevölkerungsstatistik und der systematischen Analyse von gesellschaftlichen und sozio-ökonomischen Entwicklungen wird jedoch erst in der frühen Neuzeit verortet. Als eine der ersten Publikationen der

⁹ „Lukas-Evangelium“, Kapitel 2, Vers 1-3 in der Übersetzung von Martin Luther von 1912. (*online*).

¹⁰ Vgl. Michel Foucault, *Geschichte der Gouvernementalität* (Frankfurt: Suhrkamp, 2006).

¹¹ Vgl. Sir Francis Bacon, „Nam et ipsa scientia potestas est“, in: *Meditationes Sacrae*, zitiert nach *The Oxford Dictionary of Quotations*, ed. Angela Partington (London et al.: Oxford University Press, 1992 [1941]), 45.

¹² Peter Miller, „Governing by Numbers: Why Calculative Practices Matter“, *Social Research* 68.2 (2000): 379-396.

modernen Bevölkerungsanalyse gilt die 1662 erschienene Schrift *Natural and Political Observations Made upon the Bills of Mortality* von John Graunt.¹³ In dieser Arbeit analysierte Graunt Sterbeverzeichnisse in London, um ein Frühwarnsystem zu erstellen, das genauere Informationen über das Auftreten und die Verbreitung der Pest liefern sollte. Etwas mehr als ein Jahrzehnt später, im Jahr 1676, veröffentlichte William Petty, ein Freund und Kollege Graunts, sein Werk *Political Arithmetick*, das als das Grundlagenwerk dieser neuen wissenschaftlichen Bewegung gilt. Die Verfechter der ‚politischen Arithmetik‘ verfolgten von Anfang nicht allein wissenschaftliche Ziele, sondern sie waren vor allem an der Lösung von sozialen und politischen Problemen durch möglichst umfassende wissenschaftliche Datenanalysen interessiert.¹⁴

Obwohl die ‚politische Arithmetik‘ nicht unumstritten war, bildete sie bis zum Ende des 18. Jahrhunderts die zentrale Grundlage für die systematische Erfassung und Analyse von Bevölkerungsdaten und sozio-ökonomischen Entwicklungen. Die Sammlung und Aufbereitung der Bevölkerungsdaten verblieb jedoch im Wesentlichen im Bereich der Administration und diente nach wie vor überwiegend politischen Zwecken. Eine eigenständige wissenschaftliche Disziplin zur systematischen empirischen Analyse von sozialen Entwicklungen und gesellschaftlichen Prozessen wie die ‚Soziologie‘, existierte zu diesem Zeitpunkt noch nicht. Das Nachdenken über die Gesellschaft und gesellschaftliche Bedingungen war hauptsächlich Gegenstand der Philosophie, genauer gesagt: der politischen Philosophie sowie der Sozial- und Moralphilosophie. Die einzige wissenschaftliche Disziplin, die sich im 18. Jahrhundert mit der systematischen empirischen Analyse von sozio-ökonomischen Entwicklungen und unter anderem auch mit der Analyse von Bevölkerungsdaten auseinandersetzte, war die sich in dieser Zeit allmählich formierende Schule der ‚Nationalökonomie‘.¹⁵

¹³ Vgl. Philipp Lepenies, *Die Macht der einen Zahl: Eine politische Geschichte des Bruttoinlandsprodukts* (Berlin: Suhrkamp, 2013), 23-36.

¹⁴ Vgl. ebd., 31-36.

¹⁵ Hier ist zum einen natürlich der besondere Beitrag von Adam Smith (1723-1790) hervorzuheben, insbesondere sein 1776 erstmals veröffentlichtes Hauptwerk *An Inquiry into the Nature and Causes of the Wealth of Nations*. Eigens erinnert sei hier jedoch an Thomas Robert Malthus (1766-1834), der mit seinem 1798 erschienen *Essay on the Principle of Population* und seinem 1820 veröffentlichten Hauptwerk *Principles of Economics* eine für das politische Denken des 19. Jahrhunderts sehr einflussreiche ökonomische Theorie der Bevölkerungsentwicklung formulierte, vgl. Lepenies, *Die Macht der einen Zahl*, 31-49.

An der Schwelle vom 18. zum 19. Jahrhundert beginnt sich das statistische Denken in Wissenschaft, Politik und Verwaltung allmählich immer mehr durchzusetzen. In diesen Zeitraum fallen auch die Gründung der ersten staatlichen Statistikbüros und damit der Beginn der amtlichen Statistik und einer systematischen Erfassung von Bevölkerungsdaten.¹⁶ Die der Bevölkerungsstatistik zugrunde liegenden statistischen Methoden und Verfahren blieben im Verlauf des 19. Jahrhunderts jedoch noch weitgehend rudimentär. Die amtliche Statistik beschränkte sich im Wesentlichen auf die deskriptive Statistik, also auf eine Beschreibung von Häufigkeiten und Häufigkeitsverteilungen. Schätzungen und Stichprobenverfahren galten im 19. Jahrhundert noch als äußerst fehleranfällig. Dies änderte sich erst zu Beginn des 20. Jahrhunderts mit der Entstehung und massiven Expansion der sogenannten ‚quantitativen Sozialforschung‘ – insbesondere im Bereich der angewandten und politiknahen Sozialforschung in Großbritannien und in den Vereinigten Staaten – und mit dem Siegeszug der ‚mathematischen Ökonomie‘ in den Wirtschaftswissenschaften.¹⁷

Aufschwung der empirischen Sozialforschung im 20. Jahrhundert

Der Beginn des 20. Jahrhunderts markiert zweifellos den Anbruch eines neuen Zeitalters der umfassenden Vermessung und wissenschaftlichen Durchdringung der Gesellschaft. In den ersten Jahrzehnten des 20. Jahrhunderts – und in immer intensiverem und rasanterem Ausmaß ab Mitte des 20. Jahrhunderts – wurden die Bemühungen um eine systematische empirische Erfassung und Analyse möglichst aller menschlichen Lebensbereiche und gesellschaftlichen Entwicklungen gewaltig ausgeweitet. Dies ist einerseits der Weiterentwicklung und immer stärkeren Ausdifferenzierung der Human- und Sozialwissenschaften in der ersten Hälfte des 20. Jahrhunderts geschuldet. Neue Disziplinen traten auf den Plan, wie etwa die Psychologie, die Soziologie oder die Politikwissenschaften. Es ist andererseits auch in einem engen Zusammenhang zu sehen mit der

¹⁶ Es ist hier allerdings wichtig zu betonen, dass es regional sehr unterschiedliche Entwicklungen der Verwaltung und der Statistik gab. Für eine ausführliche und genaue Darstellung der Entwicklung der politischen Arithmetik und der Statistik im Spannungsfeld von Wissenschaft, Politik und Verwaltung siehe insbesondere das Buch von Alain Desrosières, *Die Politik der großen Zahlen: Eine Geschichte der statistischen Denkweise* (Berlin: Springer, 2005).

¹⁷ Vgl. ebd., 182-238; 311-357.

Entwicklung und Weiterentwicklung der Methoden der empirischen Sozialforschung und mit einer stärkeren Integration von Verfahren der induktiven (oder ‚schließenden‘) sowie der explorativen Statistik in die Praxis der Sozialforschung. Sehr einflussreich ist zudem die immer stärkere wissenschaftliche Unterfütterung des Regierungshandelns und der bürokratischen Organisation der im 19. Jahrhundert entstandenen ‚modernen‘ Nationalstaaten. Viele Aspekte der staatlichen Planung und der politischen Regulierung würde es ohne die Berücksichtigung immer neuer empirischer Daten und statistischer Informationen überhaupt nicht geben. Die Expansion der wissenschaftlichen Forschung und damit auch die Produktion immer neuer Daten und Informationen über sozio-ökonomische Zusammenhänge und Entwicklungen wurden von den staatlichen Bürokratien stark befördert und vorangetrieben.¹⁸

Schließlich ist hervorzuheben, dass die Expansion der nicht-staatlichen kommerziellen Sozialforschung, insbesondere die Entstehung der Markt- und Meinungsforschung in den ersten Jahrzehnten des 20. Jahrhunderts, auch in erheblichem Maße zur Entwicklung von neuen Forschungsmethoden und zur fortschreitenden Vermessung des gesellschaftlichen Alltagslebens beitrug. Als eines der ersten kommerziellen Forschungsinstitute gilt das sogenannte *National Analysts Worldwide*, ein Marktforschungsinstitut, das 1911 von Charles Coolidge Parlin in Philadelphia (USA) gegründet wurde.¹⁹ In den folgenden Jahrzehnten weitete sich die kommerzielle Marktforschung immer weiter aus und damit eine immer umfassendere Erforschung der Bedürfnisstrukturen und des Konsumverhaltens der Bevölkerung und einzelner Käuferschichten. Legendär ist darüber hinaus auch der Aufstieg der kommerziellen Wahl- und Meinungsforschung, der vor allem mit dem Namen George Gallup, dem Leiter des *American Institute of Public Opinion (AIPO)*, in Verbindung gebracht wird.²⁰ Denn Gallup hat zusammen mit seinen Kollegen, so die offizielle Narration, den Sieg von Franklin D. Roosevelt im Präsidentschaftswahlkampf von 1936 entgegen anders lautender Prognosen bis auf

¹⁸ Hierzu insbesondere Lutz Raphael, „Die Verwissenschaftlichung des Sozialen als methodische und konzeptionelle Herausforderung für eine Sozialgeschichte des 20. Jahrhunderts“, *Geschichte und Gesellschaft* 22.2 (1996): 165-193; sowie Peter Wagner, *A History and Theory of the Social Sciences* (London: Sage, 2001).

¹⁹ Vgl. Harm G. Schröter, „Zur Geschichte der Marktforschung in Europa im 20. Jahrhundert“, in: *Geschichte des Konsums: Erträge der 20. Arbeitstagung der Gesellschaft für Sozial- und Wirtschaftsgeschichte*, hg. Rolf Walter (Stuttgart: Steiner, 2004), 319-336.

²⁰ Vgl. ebd., 321.

wenige Prozentpunkte genau vorhergesagt. Das Erfolgsrezept seiner Prognose war eine repräsentative Stichprobe und nicht – wie damals noch weitgehend üblich – eine breit angelegte Massenbefragung. Laut offiziellen Darstellungen markierte dies einen Meilenstein in der Geschichte der quantitativen Sozialforschung. Das Stichproben-Prinzip beziehungsweise die Durchführung von Befragungen nach dem Prinzip der Zufallsauswahl entwickelte sich in der Folge zur Standardmethode der quantitativen Sozialforschung und verhalf der empirischen Sozialforschung in der zweiten Hälfte des 20. Jahrhunderts zu einem großen Aufschwung.²¹

Ab Mitte des 20. Jahrhunderts lässt sich folglich ein weltweiter Boom der Sozialforschung und der Gesellschaftsanalyse feststellen. Dieser Boom ist sowohl von einer starken Expansion der empirischen Forschung in den Human-, Sozial- und Wirtschaftswissenschaften als auch von einer immensen Ausweitung und Profilierung der kommerziellen Markt- und Meinungsforschung geprägt. Heute gibt es kaum einen Bereich des menschlichen Lebens und der modernen Alltagswelt, der nicht durch empirische Forschung erschlossen ist und mithilfe von empirischen ‚Daten‘ und ‚Indikatoren‘ beschrieben wird. Für eine Vielzahl von Lebensbereichen gibt es heute Daten, die gesellschaftliche und wirtschaftliche Entwicklungen möglichst umfassend und detailliert beschreiben sollen. Wir haben uns bereits so stark an die statistische Erfassung unseres Alltagslebens gewöhnt, dass es uns in keiner Weise als ungewöhnlich erscheint, wenn wir in den täglichen Nachrichten oder in Zeitungen permanent mit statistischen Daten und Indikatoren konfrontiert werden. Seien es zum Beispiel die täglichen Meldungen über die Entwicklung der wichtigsten Aktien-Indizes, des Wirtschaftswachstums oder des Konsumklimas, aber auch Meldungen über die Entwicklung der Arbeitslosigkeit, der Lebenszufriedenheit oder der Ungleichheit. Wir leben also bereits seit geraumer Zeit in einer ziemlich stark vermessenen und durch vielfältige Datengenerierung und -analyse geprägten Welt.²² Dennoch scheint dieser Prozess der totalen Vermessung unseres Alltagslebens mit der Digitalisierung eine neue Dynamik erhalten zu haben, die bisherige Standards, Gewohnheiten und Gewissheiten der Gesellschaftsanalyse ins Wanken bringt.

²¹ Zu Stichprobenverfahren in der empirischen Sozialforschung und als Einführung in Methoden der empirischen Sozialforschung siehe Rainer Schnell *et al.*, *Methoden der empirischen Sozialforschung* (München: Oldenbourg, 2011).

²² Vgl. Steffen Mau, *Das metrische Wir: Über die Quantifizierung des Sozialen* (Berlin: Suhrkamp, 2017) und Uwe Vormbusch, *Die Herrschaft der Zahlen: Zur Kalkulation des Sozialen in der kapitalistischen Moderne* (Frankfurt a.M.: Campus, 2012).

Big Data: Ein neues Zeitalter der Totalvermessung?

In den einschlägigen Schriften zu *Big Data* ist immer wieder davon die Rede, dass sich im Zuge der Digitalisierung von Technologien aller Art vor allem die Möglichkeiten der Datensammlung, des Datentransfers und der technisch gestützten Datenanalyse bereits heute schon stark ausgeweitet haben.²³ In der Folge erleben wir einen großen Zuwachs des ‚Volumens‘ und der ‚Varietät‘ von Daten sowie der ‚Geschwindigkeit‘ der Datengenerierung und der Datenanalyse.²⁴ Dieser Umstand ist zum einen ganz schlicht der Tatsache geschuldet, dass die Digitalisierung von Daten aller Art eine leichtere und schnellere Verarbeitung dieser Daten ermöglicht. Zum anderen lassen sich durch Digitalisierung Daten erheblich einfacher komprimieren und immer größere Datenmengen auf immer kleineren Datenträgern speichern. Und mit der Entwicklung von immer schnelleren Computerprozessoren und immer kleineren Speichermedien wachsen die Möglichkeiten der Datenverarbeitung weiter an. Dies lässt sich am Beispiel der Veränderung von Tonträgern oder am Beispiel von USB-Sticks sehr gut nachvollziehen: Während es bis in die 1990er Jahre noch üblich war, Musik auf unzähligen Platten und anderen Tonträgern wie Kassetten und CDs zu lagern, hat die Erfindung des Speicherformats MP3 die Lagerung von Tonträgern geradezu überflüssig gemacht. Heute haben viele Menschen ganze Musikarchive auf einer einzigen Computerfestplatte gespeichert, oder sie ‚streamen‘ Musik direkt über eine Internet-Verbindung. Ähnliches lässt sich über die Entwicklung von Speichermedien sagen: Während wir in den 1980er Jahren auf einer 5,25-Zoll-Diskette noch relativ überschaubare Datenmengen speichern konnten (zwischen 80 und 1200 Kilobyte), hat heute jeder USB-Stick so viel Speicherkapazität wie eine mittelgroße Festplatte. Hier liegen wir derzeit bei Kapazitäten von bis zu 512 Gigabyte.

Diese Beispiele veranschaulichen sehr deutlich, wie stark sich die technischen Möglichkeiten zur Datenverarbeitung in den vergangenen zwei

²³ Siehe grundlegend hierzu: Kitchin, *The Data Revolution*; vgl. auch Ramón Reichert, hg., *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie* (Bielefeld: Transcript, 2014), Yvonne Hofstetter, *Sie wissen alles: Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen* (München: Bertelsmann, 2014) und Heinrich Geiselberger und Tobias Moorstedt, hg., *Big Data: Das neue Versprechen der Allwissenheit* (Berlin: Suhrkamp, 2013).

²⁴ Im englischsprachigen Kontext wird dieser Wandel daher gerne mit drei V-Wörtern beschrieben: ‚volume‘, ‚velocity‘ und ‚variety‘; vgl. Kitchin, *The Data Revolution*, 68.

bis drei Jahrzehnten verändert haben – und die Entwicklung ist längst nicht abgeschlossen. Neben der Ausweitung von Möglichkeiten zur Speicherung und Verarbeitung von immer größeren Datenmengen ermöglicht die Digitalisierung auch eine sehr leichte Verknüpfung („Relationierung“) von verschiedenen Datensätzen und eine relativ hohe Flexibilität im Umgang mit unterschiedlichen Datenarten. Die Daten sind nahezu beliebig wandelbar und skalierbar und lassen sich mühelos in eine gewünschte Form übertragen und weiterverarbeiten.²⁵

Eine der größten und einschneidendsten Neuerungen stellt jedoch die Zunahme der automatisierten Datensammlung und Datenanalyse mithilfe von Computern beziehungsweise mit spezieller Software dar, die bestimmte schematische Anweisungen („Algorithmen“) zur Verarbeitung und Analyse von Daten enthält. Algorithmen sind automatisierte Filter, nach denen Maschinen Informationen („Daten“) aus ihrer Umwelt selektieren und gemäß eines genau festgelegten Handlungsschemas („Skript“) verarbeiten.²⁶ Die algorithmische Verarbeitung von Daten ermöglicht die Analyse von riesigen Datenmengen, die ein einzelner Mensch oder ein einzelnes Forschungsteam längst nicht mehr erfassen und verarbeiten könnten. Algorithmen und algorithmisch gesteuerte Datenanalysen sind heute längst zum normalen Bestandteil unseres Alltags und unseres Arbeitslebens geworden. Bei jeder *Google*-Anfrage werden Algorithmen aktiviert, die uns eine Auswahl an Treffern anzeigen. Die Banner-Werbung im Internet oder weitere Kaufempfehlungen beim Online-Shopping werden algorithmisch erzeugt. Selbst der Börsenhandel wird heute in hohem Maße über Algorithmen gesteuert und abgewickelt. Die eingestellte Software erkennt in Bruchteilen von Sekunden Schwankungen im Marktgeschehen, die von menschlichen Analysten in dieser Geschwindigkeit nicht zu erkennen wären. Dies ermöglicht es Anlegern, bereits auf kleinste Kursentwicklungen zu reagieren.

Wenn heute also von *Big Data* die Rede ist, dann sind in der Regel die soeben beschriebenen Prozesse gemeint, allen voran die Tendenz zur automatisierten Verarbeitung und Analyse von immer größeren Datenmengen. Es liegt auf der Hand, dass die neuen technischen Möglichkeiten zur Sammlung und Verarbeitung von Daten und die neuen Verfahren der Datenanalyse auch die bisherige Praxis der Gesellschaftsanalyse und der empirischen Sozialforschung herausfordern. Denn die neuen Technologien ermöglichen ein Ausmaß an Datenverarbeitung, das bisher nicht für

²⁵ Vgl. ebd., 80-99.

²⁶ Hofstetter, *Sie wissen alles*, 87-148.

möglich gehalten wurde. Die quantitative Sozialforschung arbeitet zwar schon sehr lange mit Massendaten und bearbeitet diese mithilfe von Computerprogrammen. Der Fokus der Analyse war und ist hierbei jedoch zumeist begrenzt auf eine relativ überschaubare Gruppe von Personen und auf eine eigens für bestimmte Forschungszwecke ausgewählte Stichprobe. Die Grundlage der bisherigen Praxis der Sozialforschung bildet jedoch gerade nicht ‚Big‘ Data, sondern Angaben einer überschaubaren Auswahl von Individuen, die üblicherweise in einem Interview mithilfe eines Fragebogens ermittelt und anonymisiert weiterverarbeitet werden. Dies ist bis heute die übliche Praxis bei nahezu allen standardisierten Befragungen in der wissenschaftlichen Forschung wie in der Markt- und Meinungsforschung.

Dieses etablierte Stichproben-Prinzip steht nun mit dem Wachstum von *Big Data* verstärkt zur Disposition. Denn die neuen Technologien und Analyseverfahren ermöglichen eine Erhebung und Weiterverarbeitung von riesigen Datenmengen in Bruchteilen von Sekunden und häufig sogar in Echtzeit. ‚Warum nur eine Stichprobe analysieren, die tendenziell verzerrt sein kann und falsche Ergebnisse liefert, wenn man auch größere Mengen von Daten erheben kann und prinzipiell eine komplette Grundgesamtheit erfassen könnte?‘ So ließe sich der Geist umschreiben, der in der aktuellen *Big-Data*-Bewegung zuweilen vorherrscht und von Apologeten der neuen Analyseverfahren zum Teil sehr offensiv propagiert wird. Folglich lässt sich im Wandel hin zu *Big Data* eine Reihe von Entwicklungen ausmachen, die die bisherige Praxis der Sozialforschung stark in Frage stellen:

Wir können erstens einen *Trend zur Erfassung von riesigen Datenmengen* feststellen, der zum Teil bereits bis zur Gesamterfassung reicht. Das heißt, es werden heute relativ wahllos die verschiedensten Daten gesammelt und gespeichert. Nicht unbedingt, weil man alle erhobenen Daten am Ende letztlich braucht und für eine Analyse verwendet, sondern einfach weil es technisch möglich ist und kaum Kosten verursacht. Mit dem Trend zur endemischen Erfassung von Daten geht im Kontext der *Big-Data*-Bewegung zweitens ein *neues Objektivitätsversprechen* einher, das bisherige Forschungsstrategien entwertet und *Big Data* nach dem Motto ‚*big is beautiful*‘ zur überlegenen Forschungsstrategie erhebt. Dieser neue Datenfetisch und dieser neue Glaube an die Überlegenheit von *Big-Data*-Analysen gegenüber bisherigen Forschungsstrategien werden deutlich in folgendem Zitat von Chris Anderson, dem ehemaligen Chef-Redakteur des Technologie-Magazin *Wired*. Er sieht im neuen Zeitalter von *Big Data* nicht nur ein Ende etablierter Forschungswege und etablierter

Annahmen über menschliches Verhalten nahen, sondern insgesamt ein ‚Ende der Theorie‘ – eine Provokation, mit der Anderson in den vergangenen Jahren vor allem auch in wissenschaftlichen Kreisen viel Aufmerksamkeit erregt hat:

Wir leben in einer Welt, in der riesige Mengen von Daten und angewandte Mathematik alle anderen Werkzeuge ersetzen, die man sonst noch so anwenden könnte. Ob in der Linguistik oder in der Soziologie: Raus mit all den Theorien des menschlichen Verhaltens! Vergessen Sie Taxonomien, die Ontologie und die Psychologie! Wer weiß schon, warum Menschen sich so verhalten, wie sie sich gerade verhalten? Der springende Punkt ist, dass sie sich so verhalten und dass wir ihr Verhalten mit einer nie gekannten Genauigkeit nachverfolgen können. Hat man erst einmal genug Daten, sprechen die Zahlen für sich selbst.²⁷

Neben den beiden bisher erwähnten Tendenzen lässt sich drittens ein *Trend zur Individualisierung und Personalisierung der Datenanalyse* ausmachen. Während die bisherige Praxis der Sozialforschung nicht unbedingt an den Daten von Einzelnen interessiert war – dies gilt nicht nur für die wissenschaftliche Forschung, sondern auch für die Markt- und Meinungsforschung – ist gerade im Bereich der Marktforschung eine starke Tendenz zur Personalisierung der Datenerhebung und der Datenanalyse erkennbar. Personalisierte Informationen ermöglichen differenziertere und damit effektivere Werbestrategien. Deshalb werden Daten, insbesondere personenbezogene Daten, heute bisweilen auch als das „neue Gold“ des Internetzeitalters bezeichnet.²⁸

Der Wandel hin zu *Big Data* und zur automatisierten Datenanalyse, dies ist der vierte und letzte Aspekt, der hier hervorgehoben werden soll, wird überwiegend von Privatunternehmen und Akteuren mit kommerziellen Interessen vorangetrieben. Im Kontext von *Big Data* und in der neuen Forschung, die mithilfe von *Big-Data*-Analysen betrieben wird, finden wir derzeit eindeutig *eine Dominanz von kommerziellen Akteuren* vor. *Big Data* wird sicherlich auch in der Wissenschaft weiter zunehmen. Es gibt immer mehr wissenschaftliche Projekte, die mit *Big-Data*-Analysen arbeiten. Aber das Gros der Datensammlung und der Datenanalyse findet heute noch überwiegend außerhalb der Wissenschaft in

²⁷ Chris Anderson, „Das Ende der Theorie: Die Datenschwemme macht wissenschaftliche Methoden obsolet“, in: *Big Data: Das neue Versprechen der Allwissenheit*, hg. Heinrich Geiselberger und Tobias Moorstedt (Berlin: Suhrkamp, 2013), 124-130.

²⁸ Vgl. Carsten Dierig, „Daten sind das neue Gold“, *Die Welt*, 27.04.2014. (online).

privaten Unternehmen statt. Diese Unternehmen verfügen über die Ressourcen und Technologien zur groß angelegten Datenerhebung. Und sie haben aufgrund ihrer Dienstleistungen einen direkten Zugang zu einer Masse an Kundendaten und zu Kunden, die ihnen – mal bewusst, doch häufig auch unbewusst – bereitwillig diese Daten liefern.

In diesem Trend zur Privatisierung der gesellschaftsbezogenen Datenanalyse liegt derzeit eine der größten Herausforderungen für die wissenschaftliche Sozialforschung. Denn die großen Datenkonzerne wie *Facebook*, *Google*, *Amazon* und Co. haben Zugang zu Daten, die der wissenschaftlichen Sozialforschung nicht zur Verfügung stehen. Sie arbeiten mit Technologien und Analyseverfahren, die der wissenschaftlichen Entwicklung weit voraus sind. Es ist unklar, was mit diesen Daten geschieht und welche Erkenntnisse in Zukunft aus diesen Daten abgeleitet werden. Es besteht jedoch die Gefahr, dass die kommerzielle *Big-Data*-Forschung Erkenntnisse über die Einstellungen, Verhaltensdispositionen und die Lebenssituation von Gruppen und Individuen generiert, die weit über das Maß hinaus gehen, das uns durch die bisherige Praxis der wissenschaftlichen Analyse von Gesellschaft bekannt ist. Hier stellt sich ganz grundsätzlich die Frage, ob es normativ wünschenswert ist, dass kommerzielle Akteure immer mehr Daten von uns und über uns sammeln und daraus entsprechende Erkenntnisse ableiten. Und es stellt sich darüber hinaus die Frage, wie sich auch der Blick auf die Gesellschaft verändert, wenn die bisherige Sozialforschung immer mehr von Forschung überlagert wird, die überwiegend auf Datensammlungen und -analysen von kommerziellen Akteuren basiert. Ich möchte im Folgenden kurz einige zentrale Entwicklungen im Kontext der kommerziellen Datensammlung skizzieren, die zeigen sollen, wie massiv und umfassend heute Daten erhoben und ausgewertet werden.

Daten, Daten, Daten ...: Die Totalvermessung von Individuen und Haushalten

Es ist heute weithin bekannt, dass wir im Internet viele Spuren hinterlassen und dass diese Spuren von Unternehmen genutzt und weiterverarbeitet werden, um ihre Dienste zu verbessern, mehr über uns als potenzielle Kunden zu erfahren oder um Werbung besser platzieren zu können und damit am Ende mehr Einnahmen zu erzielen. Nur deshalb sind viele Internet-Anwendungen wie Suchmaschinen, *Apps*, E-Mail- und *Messenger*-Dienste, *Social-Media*-Portale wie *Facebook* oder Nachrichtenportale kostenlos nutzbar. Ohne den Mehrwert, den Internet-Dienst-

anbieter durch die Datensammlung und durch zusätzliche Werbeeinnahmen erzielen, wären viele Anwendungen und Dienstleistungen heute längst nicht mehr frei und unentgeltlich verfügbar. Manche dieser Anwendungen gäbe es wohl überhaupt nicht. Doch das Geschäft mit den Daten und die unablässige Sammlung von immer mehr Nutzerinformationen ist für die meisten Unternehmen so lukrativ und so wichtig, dass sie mehr auf eine möglichst weite Verbreitung und Nutzung ihrer Dienste aus sind als auf die Erzielung von Einnahmen aus Abonnements und Nutzerentgelten. So weit, so gut.

Was Vielen jedoch noch nicht klar ist, sind die Reichweite und das Ausmaß, mit der die Datensammlung heute bereits betrieben wird, und der Umfang an individuellen Daten, die wir unablässig durch Nutzung bestimmter Dienste produzieren und die – ob wir das wollen oder nicht – unablässig von uns gespeichert werden. Die Journalistin Barbara Junge hat sich vor einigen Jahren einmal die Mühe gemacht und sich bei diversen Unternehmen über die Art und den Umfang der über sie gespeicherten Daten informiert.²⁹ Von einigen Unternehmen hat sie keine Antwort bekommen. Andere Unternehmen haben bereitwillig, die Informationen zur Verfügung gestellt, die sie von ihr in ihren Datenbanken gespeichert haben. Allein von *Amazon* erhielt sie eine 44 Seiten umfassende Auflistung aller ihrer Aktivitäten mitsamt einer Analyse ihres Kaufverhaltens, ihren zentralen Interessen und Vorlieben, ihrer sozialen Situation und ihrer Kreditwürdigkeit. Junge schreibt in ihrem Beitrag dazu:

Der Einblick, den allein unser Warenbestellverhalten liefert, ist größer, als wir denken. Ohne Scheu hinterlegen wir Körpermaße, Bildungsgrad, Familienverhältnisse, Freizeitinteressen, gerne die Automarke oder die bevorzugte Ernährungsweise. Auch unsere Kaufbereitschaft, die Zahl der benutzten Kreditkarten oder Hinweise auf unsere Gebührentreue haben ihren ökonomischen Wert.³⁰

Auch hier ließe sich vielleicht noch einwenden, dass dies im Prinzip nichts Neues ist. Auskunftsteien und Informationsdienste gibt es schon lange, die Verbraucherinformationen speichern und bei Bedarf an Unternehmen weiterverkaufen, etwa wenn sie gesicherte und unabhängige Informationen über unsere Kreditwürdigkeit benötigen. Man denke hier

²⁹ Barbara Junge, „Wer hat meine Daten?“, in: *Big Data: Das neue Versprechen der Allwissenheit*, hg. Heinrich Geiselberger und Tobias Moorstedt (Berlin: Suhrkamp, 2013), 23-34.

³⁰ Ebd., 31.

etwa an traditionsreiche Unternehmen wie die *Schufa Holding AG*, die aus der *SCHUFA e.V.* hervorgegangen ist, die Schutzgemeinschaft für allgemeine Kreditsicherung. Auch ist prinzipiell nichts dagegen einzuwenden, wenn Verkaufsportale unseren Geschmack und unsere Vorlieben besser einschätzen können, damit sie uns bei weiteren Besuchen nur passende Werbung und Kaufangebote anzeigen. Was wäre jedoch, wenn unsere persönlichen Daten für Zwecke benutzt werden, mit denen wir nicht einverstanden sind? Oder wenn sie an Personen oder Unternehmen gelangen würden, denen wir nicht explizit die Erlaubnis zur Verwendung unserer Daten gegeben haben?

Dass dies mittlerweile geschieht, darauf hat der Journalist Dirk von Gehlen unlängst in einem Beitrag in der *Süddeutschen Zeitung* hingewiesen.³¹ Er sei selbst ziemlich schockiert gewesen, so schreibt er, als er von einer Kollegin einen Datensatz zugespielt bekommen habe, der seine gesamte Internet-Nutzungshistorie aus dem August 2016 enthielt. Die Daten stammten aus einer ‚Produktprobe‘, die ein Reporterteam des *NDR* bei einer Datenfirma zum Schein gekauft hatte. Der Inhalt seiner Datei ließ von Gehlen erschauern:

In diesen verkauften Daten kann ich nun nachlesen, welche Webseiten ich ansurfte, welche Begriffe ich suchte und sogar minutengenau, welche Reiseverbindungen mich interessierten (dank einer Nachlässigkeit der Deutschen Bahn). Man sieht darin, was mich beschäftigte, wann ich zur Arbeit ging und wohin ich reiste. Diese Daten stammen aus dem Webbrowser, den ich nutze, sie fügen sich zu einem Bild, das mich digital nackt dastehen lässt – und all das kann man kaufen.

Dieses Beispiel offenbart in eklatanter Weise ein Grundproblem der ungebremsten Datensammlung im heutigen digitalen Zeitalter: Je mehr die Menschen ihre alltäglichen Verrichtungen über Internetdienste und über internetfähige elektronische Geräte abwickeln, desto mehr persönliche Informationen geben sie von sich preis, und desto mehr Daten sind über sie und über ihr Leben verfügbar. Da Daten heute jedoch leichter transferierbar und kombinierbar sind denn je und prinzipiell überall speicherbar, ist allerdings eine genaue Kontrolle der Weiterverwendung der Daten kaum mehr möglich. Ein Schutz der privaten Daten und ein Schutz der Privatsphäre sind unter diesen Bedingungen folglich nicht mehr zu gewährleisten. Denn die individuellen Nutzerdaten werden in

³¹ Dirk von Gehlen, „Plötzlich nackt im Netz“, *Süddeutsche Zeitung* (2. November 2016). (online).

der Regel auf Servern gespeichert, die fernab vom Einflussbereich des Einzelnen oder der geltenden Datenschutzbestimmungen liegen. Häufig willigen die Kunden, zum Teil ohne es zu wollen oder es genau zu wissen, bei der Nutzung einer *App*, einer Suchmaschine oder eines Web-Browsers anstandslos der Speicherung und Weiterverarbeitung ihrer Daten oder gar der Weitergabe an Dritte ein. Was mit ihren Daten geschieht, welche Daten möglicherweise an Dritte weiterverkauft werden, und wer diese ‚Dritten‘ sein könnten, wird in den entsprechenden Allgemeinen Geschäftsbedingungen und Datenschutzbestimmungen nicht weiter spezifiziert. Den Nutzern bleibt in den meisten Fällen ohnehin keine andere Wahl, als entsprechenden Vorgaben in den Allgemeinen Geschäftsbedingungen in vollem Umfang zuzustimmen, wenn sie einen bestimmten Service oder eine bestimmte *App* nutzen möchten.

Welche Folgen und welche bedenklichen Konstellationen sich daraus ergeben können, hat Eric Schmidt, der ‚*Executive Chairman*‘ der aus *Google Inc.* hervorgegangenen *Alphabet Holding*, im Jahr 2010 in einem Vortrag einmal relativ offen und unverblümt zum Ausdruck gebracht. Das folgende Zitat ist auch ein Beleg für den Eifer, mit dem Internetkonzerne wie *Google* ihre Datensammlungen betreiben, und welche Allmachtsfantasien daran anscheinend geknüpft sind:

The more information we have about you the more we can improve the quality of our searches. [...] We don't need you to type at all, 'cause we know where you are – with your permission. We know where you've been – with your permission. We can more or less guess what you're thinking about.³²

Die größten Datenlieferanten sind gegenwärtig Computer und mobile Endgeräte, wie Tablets oder Smartphones, die viele Menschen die meiste Zeit ganz selbstverständlich bei sich tragen. Diese mobilen Endgeräte können nicht nur das komplette Bewegungsprofil ihrer Benutzer abbilden, sondern auch viele weitere Aspekte unseres Alltags und unseres Soziallebens. Es sind jedoch nicht nur die großen Internet-Konzerne wie *IBM*, *Google*, *Facebook* oder *Microsoft* sowie Hersteller von mobilen Endgeräten und *Apps*, die heute massenhaft Daten sammeln und weiterverarbeiten. Durch die zunehmende Verbreitung von netz- und softwarebasierten, sogenannten ‚smarten‘ Technologien in nahezu allen Lebens-

³² Auszug aus einem Vortrag von Eric Schmidt auf dem *Washington Ideas Forum* 2010. Zitiert nach Jan-Felix Schrape, „Big Data als Erwartungsraum“, *Soziopolis – Gesellschaft beobachten*. (online).

bereichen, werden auch Haushalts- und TV-Geräte, Küchenmaschinen, Heizungs- und Lichtanlagen oder Spielekonsolen zu Datensammlern. Immer mehr dieser Geräte zeichnen unentwegt Daten über ihre Nutzung, über gewählte Programme, über technische Störungen und so fort auf und übertragen diese per Datenverbindung an die jeweiligen Vertriebs- oder Herstellerfirmen. Auf diese Weise werden immer mehr Firmen zu Datensammlern und -analysten. Die Datensammlung dient nach offiziellen Angaben zumeist ausschließlich der Verbesserung und Weiterentwicklung der jeweiligen Produkte. Es ist jedoch auch nicht ausgeschlossen, dass Daten an Dritte weiterverkauft werden, die ihrerseits wieder ein Interesse der Sammlung und Analyse von Kunden- und Haushaltsdaten haben.

Die Sammlung von personenbezogenen Informationen und von detaillierten Informationen über Haushalte sind heute jedenfalls zu einem lukrativen Geschäft geworden. Es gibt Unternehmen wie *PayBack*, das sicherlich Vielen bekannt ist, die im großen Stil das Kaufverhalten von Millionen von Konsumenten aufzeichnen und bei Bedarf Konsumentendaten an interessierte Kunden weiterverkaufen. Darüber hinaus finden sich im Bereich der Marktforschung auch neue Datenunternehmen wie *Acxiom* oder *Schober*, die in der breiten Öffentlichkeit weniger präsent und deshalb auch weniger bekannt sind. Das Geschäftsmodell dieser Firmen besteht darin, Privatadressen und Kundenprofile zu sammeln und mit so vielen Zusatzinformationen wie möglich anzureichern (etwa Angaben zum Bildungshintergrund, zum Konsumverhalten, zur Arbeits-, Wohn- und Lebenssituation und vieles mehr). Diese Informationen werden an Interessenten weiterverkauft und je nach Kundenwunsch zielgruppenspezifisch aufbereitet. Die Sammlung und Aufbereitung von Sozialdaten und von personenbezogenen Daten birgt gegenwärtig, so scheint es jedenfalls, ein großes Marktpotenzial in sich. Es ist derzeit nicht absehbar, dass und wie sich dieser Trend aufhalten ließe.

Auf dem Weg in die Transparenzgesellschaft?

In diesem Beitrag wurde ein Bogen gespannt von den Anfängen der Gesellschaftsanalyse bis zur heutigen Diskussion über ein neues Zeitalter der Totalvermessung der Gesellschaft mithilfe von digitalen Technologien und *Big-Data*-Analysen. Es wurde deutlich gemacht, dass es immer schon Versuche gegeben hat, die Gesellschaft und ihre einzelnen Mitglieder möglichst vollständig zu erfassen und zu analysieren, oder wenigstens mithilfe von Daten zu ordnen und zu kategorisieren. Ein Großteil der

Forschung diene stets konkreten politischen, wirtschaftlichen, verwaltungstechnischen oder wissenschaftlichen Zielen. Der Anspruch einer möglichst umfassenden Vermessung der Gesellschaft ist somit nicht neu. Dennoch stellt die *Big-Data*-Forschung mit ihren neuen Möglichkeiten der automatisierten Datensammlung und -verarbeitung, wie wir sehen konnten, durchaus eine Zäsur in der Geschichte der Gesellschaftsanalyse dar. Noch nie war es so einfach und so schnell möglich, eine schier unfassbare Masse an personenbezogenen Daten zu generieren, zu verarbeiten und zu analysieren. Die Technologien der Datenerfassung und Datenerhebung sind derzeit sicherlich noch nicht so weit entwickelt und ausgereift, wie manche es vielleicht gerne hätten. Allerdings ist klar, dass ein weiterer Fortschritt hier nicht lange auf sich warten lassen wird und dass die Technologien und Verfahren bereits in ein paar Jahren noch mehr Daten verarbeiten und noch detailliertere Aussagen über menschliches Verhalten und gesellschaftliche Entwicklung liefern werden. Hier stellt sich die Frage, welche gesellschaftlichen Implikationen und auch welche Implikationen für die Weiterentwicklung der Gesellschaftsanalyse sich aus den heutigen Entwicklungen ablesen lassen?

Eine zentrale Entwicklung, die sich bereits heute abzeichnet, ist eine fundamentale Veränderung unseres bisherigen Verständnisses von Privatheit und eine Verschiebung der bestehenden Trennlinie zwischen der öffentlichen und der privaten Sphäre. Hier ist die entscheidende Frage, in welche Richtung sich die gesellschaftlichen Normen und Wertvorstellungen verschieben werden: hin zu einem stärkeren Schutz der Privatsphäre und der Persönlichkeitsrechte unter den neuen digitalen und umfassend vernetzten Vorzeichen? Oder bewegen wir uns auf ein neues Zeitalter allumfassender Transparenz und Zurschaustellung individueller Lebenswege und sozialer Lebenswelten zu, wie es sich im Bauplan der neuen Technologien, in den Geschäftsstrategien ihrer Hersteller und Betreiber sowie im Gebrauch dieser Technologien durch ihre Nutzer bereits abzeichnet? Aus heutiger Sicht wirkt die Vorstellung einer Welt allumfassender Sichtbarkeit und Transparenz noch wie eine dystopische Zukunftsvision oder gar wie ein Schreckensszenario. Es ist allerdings auch gut möglich, dass sich die nachfolgenden Generationen Stück für Stück in dieser neuen transparenten Welt der Daten einrichten und ganz eigene Umgangsformen mit den neuen Technologien entwickeln werden.³³

³³ Wer genauer nachvollziehen möchte, wie sich dieser Wandel in der heutigen Arbeitswelt bereits vollzieht und welche Absurditäten und Gefahren gerade in der subtilen Veränderung im Alltag und im alltäglichen Miteinander liegen können, dem

Für die Soziologie und die angrenzenden Kultur- und Medienwissenschaften bedeutet dies, den Wandel hin zur Transparenzgesellschaft genauer zu betrachten und die sozialen und gesellschaftlichen Implikationen, die mit *Big Data* einhergehen, nüchtern zu analysieren. Dies bedeutet, dass man weder in allzu große Panik und Panikmache verfallen sollte, noch dass man sich den naiven Fortschrittsvorstellungen und Allmachtsfantasien der Protagonisten der neuen *Big-Data*-Bewegung unkritisch anschließen sollte. *Big Data* stellt ohne Zweifel eine große Herausforderung für die Soziologie und für die wissenschaftliche Gesellschaftsanalyse dar. Das Ende der Soziologie und der soziologischen Theorien bedeutet sie sicher nicht. Denn die Soziologie erschöpft sich gerade nicht in einer reinen datengetriebenen quantitativen Sozialforschung. Ihre Stärke liegt auch und vor allem darin, dass sie mithilfe von Theorien und hermeneutischen Verfahren in der Lage ist, sich auch ohne *Big Data* in die Logik und in die innere Struktur von sozialen Phänomenen und Prozessen hineinzusetzen. Im Kontext der aktuellen *Big-Data*-Debatte liegt die besondere Aufgabe und Herausforderung der Soziologie wohl vor allem darin, zum aktuellen *Hype* um *Big Data* ein Stück weit kritisch Distanz zu halten, ohne sich den neuen Möglichkeiten und Potenzialen der digitalen Datenanalyse gänzlich zu verschließen.

empfehle ich die Lektüre von Dave Eggers, *The Circle* (New York: Random House, 2011).

Literaturverzeichnis

- Anderson, Chris. „Das Ende der Theorie: Die Datenschwemme macht wissenschaftliche Methoden obsolet“. In: *Big Data: Das neue Versprechen der Allwissenheit*. Hg. Heinrich Geiselberger und Tobias Moorstedt. Berlin: Suhrkamp, 2013. 124-130.
- Assmann, Jan. *Das kulturelle Gedächtnis: Schrift, Erinnerung und politische Identität in frühen Hochkulturen*. München: Beck, 2013.
- Desrosières, Alain. *Die Politik der großen Zahlen: Eine Geschichte der statistischen Denkweise*. Berlin: Springer, 2005.
- Eggers, Dave. *The Circle*. New York: Random House, 2011.
- Foucault, Michel. *Geschichte der Gouvernementalität*. 2 Bde. Frankfurt: Suhrkamp, 2006.
- Geiselberger, Heinrich und Tobias Moorstedt. Hg. *Big Data: Das neue Versprechen der Allwissenheit*. Berlin: Suhrkamp, 2013.
- Hofstetter, Yvonne. *Sie wissen alles: Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen*. München: Bertelsmann, 2014.
- Junge, Barbara. „Wer hat meine Daten?“. In: *Big Data: Das neue Versprechen der Allwissenheit*. Hg. Heinrich Geiselberger und Tobias Moorstedt. Berlin: Suhrkamp, 2013. 23-34.
- Kehlmann, Daniel. *Die Vermessung der Welt*. Reinbek: Rowohlt, 2005.
- Kitchin, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: Sage, 2014.
- Lepenies, Philipp. *Die Macht der einen Zahl: Eine politische Geschichte des Bruttoinlandsprodukts*. Berlin: Suhrkamp, 2013.
- Mau, Steffen. *Das metrische Wir: Über die Quantifizierung des Sozialen*. Berlin: Suhrkamp, 2017.
- Miller, Peter. „Governing by Numbers: Why Calculative Practices Matter“. *Social Research* 68.2 (2000): 379-396.
- Partington, Angela, Ed. *The Oxford Dictionary of Quotations*. London et al.: Oxford University Press, 1992 [1941].
- Rammert, Werner. Hg. *Technik und Sozialtheorie*. Frankfurt a.M.: Campus, 1998.
- Raphael, Lutz. „Die Verwissenschaftlichung des Sozialen als methodische und konzeptionelle Herausforderung für eine Sozialgeschichte des 20. Jahrhunderts“. *Geschichte und Gesellschaft* 22.2 (1996): 165-193.
- Reichert, Ramón. Hg. *Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Bielefeld: Transcript, 2014.

- Rifkin, Jeremy. *Die dritte industrielle Revolution: Die Zukunft der Wirtschaft nach dem Atomzeitalter*. Frankfurt: Campus, 2011.
- . *Die Null-Grenzkosten-Gesellschaft: Das Internet der Dinge, kollaboratives Gemeingut und der Rückzug des Kapitalismus*. Frankfurt: Campus, 2014.
- Schirrmacher, Frank. Hg. *Technologischer Totalitarismus: Eine Debatte*. Berlin: Suhrkamp, 2015.
- Schnell, Rainer, Paul B. Hill und Elke Esser. *Methoden der empirischen Sozialforschung*. München: Oldenbourg, 2011.
- Schröter, Harm G. „Zur Geschichte der Marktforschung in Europa im 20. Jahrhundert“. In: *Geschichte des Konsums: Erträge der 20. Arbeitstagung der Gesellschaft für Sozial- und Wirtschaftsgeschichte*. Hg. Rolf Walter. Stuttgart: Steiner, 2004. 319-336.
- Vormbusch, Uwe. *Die Herrschaft der Zahlen: Zur Kalkulation des Sozialen in der kapitalistischen Moderne*. Frankfurt a.M.: Campus, 2012.
- Wagner, Peter. *A History and Theory of the Social Sciences*. London: Sage, 2001.

Internetquellen

- Die Bibel*. „Lukas-Evangelium“, Kapitel 2, Vers 1-3. Übers. Martin Luther. 1912. http://www.bibel-online.net/buch/luther_1912/lukas/2/#1 [zuletzt besucht 20.03.2017].
- Dierig, Carsten. „Daten sind das neue Gold“. In: *Die Welt* (27. April 2014). <https://www.welt.de/print/wams/finanzen/article127344442/Daten-sind-das-neue-Gold.html> [zuletzt besucht 15.01.2017].
- Gehlen, Dirk von. „Plötzlich nackt im Netz“. *Süddeutsche Zeitung* (2. November 2016). <http://www.sueddeutsche.de/digital/datenhandel-ploetzlich-nackt-im-netz-1.3231155> [zuletzt besucht 15.01.2017].
- Mayerl, Jochen. „Bedeutet ‚Big Data‘ das Ende der sozialwissenschaftlichen Methodenforschung?“. In: *Soziopolis – Gesellschaft beobachten*. <http://www.soziopolis.de/beobachten/wissenschaft/artikel/bedeutet-big-data-das-ende-der-sozialwissenschaftlichen-methodenforschung/> [zuletzt besucht 15.01.2017].
- Schrape, Jan-Felix. „Big Data als Erwartungsraum“. *Soziopolis – Gesellschaft beobachten*. <http://www.soziopolis.de/beobachten/wissenschaft/artikel/big-data-als-erwartungsraum/> [zuletzt besucht 15.01.2017].

Digitale Geographien

GEORG GLASZE

Abstract: Ein großer Teil der vielfach als *Big Data* bezeichneten ‚Datenflut‘ verweist in der einen oder anderen Form auf eine räumliche Situation auf der Erde, das heißt diese (Geo-)Daten haben eine geographische Referenz – sind also *big geospatial data*. Teilweise wurde und wird mit *Big Data* die Erwartung verknüpft, dass die ungeheuren Mengen an Daten soziale und räumliche Wirklichkeiten abbilden und unmittelbar für die Analyse zugänglich machen. In diesem Beitrag wird eine solche Perspektive als naiv zurückgewiesen: Ansätze der *critical data studies* weisen darauf hin, dass Daten immer in spezifischen sozio-technischen Kontexten produziert werden und daher immer selektiv sind. Bereits am Beispiel der traditionellen Kartographie haben die Arbeiten der Kritischen Kartographie gezeigt, dass geographische Informationen immer von den sozio-technischen Kontexten geprägt sind, in denen sie hergestellt werden. Im Mittelpunkt des Beitrags steht dann die Frage, wie die mit dem Boom an Geodaten einhergehenden sozio-technischen Transformationen es beeinflussen, welche Bilder wir uns von der Welt machen (digitale Kartographien), und letztlich auch, wie wir in der Welt leben (digitale Geographien).

Ein kritischer Blick auf Daten und Karten

Vor dem Hintergrund rasch wachsender Datenmengen wurde teilweise argumentiert, dass Daten soziale (und natürliche) Prozesse und Strukturen umfassend abbilden und diese damit beobachtbar werden. In einem viel diskutierten Aufsatz stellte beispielsweise der Herausgeber des US-amerikanischen Magazins *Wired* 2008 in Aussicht, dass bald „Computer mit einem Knopfdruck Wissen aus Daten enthüllen“.¹ Eine Auseinandersetzung mit Theorien, Hypothesen oder statistischen Verfahren würde damit überflüssig. Arbeiten der *critical data studies* kritisieren diese Perspektive auf *Big Data* als naiv und letztlich gefährlich.² So legen Rob Kitchin and Tracey P. Lauriault 2014 dar, dass immer bestimmte Daten in bestimmten gesellschaftlichen und technischen Zusammenhängen erzeugt werden. Daten können daher immer nur selektiv und partiell

¹ Chris Anderson, „The End of Theory: The Data Deluge Makes the Scientific Method Obsolete“, *Wired Magazine* 16.8 (2008). (online). (Für diese und alle folgenden Angaben von Internetquellen in den Fußnoten wird auf die vollständige Nennung im Literaturverzeichnis verwiesen.)

² Vgl. Danah Boyd and Kate Crawford, „Critical Questions for Big Data“, *Information, Communication & Society* 15.5 (2012): 662-679.

sein – auch *big (geospatial) data*.³ Ebenso ist die Interpretation von Daten immer komplex und wiederum eingebunden in spezifische sozio-technische Kontexte.⁴ Der kanadische Geograph Trevor Barnes weist 2013 zudem darauf hin, dass die Versprechen, die mit *Big Data* verknüpft werden, nicht so neu sind, wie das manche darstellen. Ähnliche Versprechen wurden schon in den *social physics* seit dem 19. Jahrhundert und insbesondere in der sogenannten ‚quantitativen Revolution‘ in der Sozial- und Wirtschaftsgeographie in den 1960er Jahren formuliert. Der bislang vielfach geschichtslose ‚Hype‘ um *Big Data* sollte daher mit den Erfahrungen sowie den kritischen Auseinandersetzungen mit den Ursprüngen und Vorläufern von *Big Data* konfrontiert und kontextualisiert werden.⁵

Auch Karten wurde und wird vielfach zugesprochen, dass sie Wirklichkeit ‚abbilden‘. So definieren viele Lehrbücher Karten beziehungsweise Kartographie über die Abbild-Metapher:

Kartographie: als Lehre von der maßstabsgerecht verkleinerten Abbildung der Erdoberfläche [...].⁶

Eine Karte ist ein verkleinertes, vereinfachtes und verebnetes Abbild der Erdoberfläche [...] und die Kartographie ist das Fachgebiet, welches sich mit der Herstellung derartiger Abbilder befasst [...].⁷

Im Kontext der poststrukturalistisch orientierten Kritik an den Grundlagen und Wahrheitsansprüchen moderner Wissenschaft, wie sie insbesondere seit den 1980er Jahren laut wurde, entwickelte sich allerdings eine wissenschaftliche Auseinandersetzung mit den Wahrheitsansprüchen der modernen Kartographie – ein Diskussionszusammenhang, der vielfach als

³ Vgl. Rob Kitchin and Tracey P. Lauriault, „Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work“, *The Programmable City Working Paper* 2. (online).

⁴ So ist auch die *Software*, mit der Daten verarbeitet und analysiert werden, niemals neutral. Die sozialwissenschaftlichen *software studies* wollen untersuchen, wie *Software* zum einen eingebettet ist in spezifische soziale und technische Entstehungsprozesse und zum anderen, wie *software* die Welt mit gestaltet, in der wir leben. Einführend dazu beispielsweise: Martin Dodge, Rob Kitchin and Mathew Zook, „How Does Software Make Space? Exploring Some Geographical Dimensions of Pervasive Computing and Software Studies“, *Environment and Planning A* 41.6 (2009): 1283-1293.

⁵ Vgl. Trevor Barnes, „Big Data, Little History“, *Dialogues in Human Geography* 3.3 (2013): 297-302.

⁶ Hartmut Leser, *Diercke Wörterbuch Allgemeine Geographie* (München: dtv, 2005), 382.

⁷ Peter Kohlstock, *Kartographie: Eine Einführung* (Paderborn: UTB, 2004), 15.

Kritische Kartographie bezeichnet wird.⁸ Die Autoren und Autorinnen der Kritischen Kartographie entlarvten den Anspruch der Kartographie, eine neutrale Abbildung der Erdoberfläche zu schaffen, als modernistischen Mythos. Ein Mythos, der letztlich verschleiert hat, dass Karten immer nur ein spezifisches Bild sein können, das bestimmte Dinge in einer bestimmten Weise darstellt und unweigerlich immer eine Vielzahl anderer möglicher Darstellungen ‚verschweigt‘. Damit verbunden war und ist vielfach eine Kritik an der Kartographie als einer elitären Praxis. Wenige privilegierte Akteure prägen die Weltbilder, die die Kartographie vermittelt.⁹

Big Geospatial Data:

Sozio-technische Hintergründe des Booms von Geodaten

Eine Vielzahl von Entwicklungen trägt dazu bei, dass viele der Daten, die als *Big Data* bezeichnet werden, in der einen oder anderen Form auf eine räumliche Situation auf der Erde verweisen:

- Mit der Öffnung des genauen Signals der Satelliten des US-amerikanischen *Navstar Global Positional System (GPS)* steht seit 2000 weltweit eine präzise Lokalisationstechnik zur Verfügung.
- Im Zusammenspiel von *GPS* mit dem mobilen Internet nutzen Millionen von Smartphone-Usern Navigationsdienste und weitere ortsspezifische Dienstleistungen. Dabei produzieren sie vielfach gleichzeitig Daten über ihre Aufenthaltsorte und damit Geodaten.
- *GPS* war eine entscheidende Grundlage für den Boom der ‚*crowdsourced geodata*‘.
- Ein großer Teil der Kommunikation in sozialen Medien produziert auch georeferenzierbare Daten.
- Viele der automatisch erhobenen Sensordaten, die eine wesentliche Quelle von *Big Data* darstellen, enthalten auch Angaben über den Ort des Sensors und sind damit Geodaten.
- Die Sensoren der Fernerkundungssatelliten produzieren täglich enorme Mengen georeferenzierter Daten.
- Im Zuge der *OpenData*-Bewegung machen vor allem öffentliche Organisationen bislang weitgehend unzugängliche Daten publik.

⁸ Vgl. Georg Glasze, „Kritische Kartographie“, *Geographische Zeitschrift* 97.4 (2009): 181-191.

⁹ Vgl. John Brian Harley, „Deconstructing the Map“, *Cartographica* 26.2 (1989): 1-20.

- Mit der Digitalisierung können vormals analog organisierte Informationen, wie beispielsweise historischer Reiseberichte oder Presseartikel, semantisch georeferenziert und damit in (neue) Geodaten übersetzt werden.

Satellitengestützte Navigationssysteme wurden seit den 1960er Jahren zunächst vom Militär entwickelt. Das bislang leistungsfähigste und bekannteste System, das US-amerikanische *Navstar Global Positional System* (*GPS*), wurde seit 1970 als sogenannte *Dual-Use-Technologie* für militärische und zivile Zwecke entwickelt und ist seit 1995 weltweit funktionsfähig.¹⁰ Es ermöglicht elektronischen Empfängern, ihre Position im Koordinatensystem mittels Funksignalen von Satelliten zu bestimmen. Bis 2000 wurde dabei zwischen einem Signal mit hoher Genauigkeit für das US-Militär und einem öffentlichen Signal mit limitierter Genauigkeit differenziert.¹¹

Die Öffnung des präzisen Signals des *GPS*-Systems war Startschuss für einen Boom neuer Navigationsdienste. Im Zusammenspiel mit dem ‚mobilen Internet‘ ermöglichen die neuen Lokalisationstechniken zudem die Entwicklung weiterer sogenannter *location based services*.¹² Dabei handelt es sich um verschiedene Dienstleistungen, die spezifisch für bestimmte Orte angeboten werden, wie etwa ortsbezogene Werbung über *GPS*- und internetfähige Smartphones. Gleichzeitig können auf Basis dieser Techniken die Aufenthaltsorte von Objekten und Personen bestimmt beziehungsweise deren Mobilität verfolgt werden. Dabei werden permanent neue Geodaten produziert – häufig unbewusst und manchmal unfreiwillig (beispielsweise beim sogenannten ‚*Tracking*‘ von *Android*-Smartphones durch *Google* als eine Grundlage für Echtzeit-Staumeldungen). Daneben war die satellitengestützte Ortsbestimmung allerdings auch eine Voraussetzung für die Entwicklung der *Web*

¹⁰ Das Militär der Sowjetunion und später Russlands hat seit den 1970er Jahren ein eigenes System entwickelt (*GLONASS*). Die Europäische Union (*GALILEO*) und China (*BEIDOU*) bauen seit einigen Jahren ebenfalls eigene Systeme auf. Vielfach verwenden neue Lokalisationssysteme in Smartphones oder Navigationssystemen die Signale mehrerer Systeme – teilweise ergänzt durch weitere terrestrische Lokalisationstechniken, wie die Zuordnung zu Funkzellen des Mobilfunks oder die Orientierung durch *WLAN*-Signale.

¹¹ Vgl. Saamer Kumar and Kevin Moore, „The Evolution of Global Positioning System (*GPS*) Technology“, *Journal of Science Education and Technology* 11.1 (2002): 59-80.

¹² Rick Sturdevant, „The Navstar Global Positioning System: From Military Tool to Global Utility“, in: *Down to Earth: Satellite Technologies, Industries, and Cultures*, eds. Lisa Parks and James Schwoch (New Brunswick, N.J.: Rutgers University Press, 2012), 99-121.

2.0-Kartographie, bei der Tausende von Freiwilligen geographische Informationen erheben – man spricht hier auch von ‚*crowdsourced geodata*‘.

Neben den mobil erhobenen Geodaten liefern unzählige ortsfeste Sensoren große Mengen an Daten, die auf den Standort des Sensors referenzieren und damit auch Geodaten produzieren – so beispielsweise Sensoren zur Verkehrszählung. Eine wichtige Quelle digitaler Geodaten sind zudem die Sensoren der Fernerkundungssatelliten. Die Entwicklung der satellitengestützten Fernerkundung ab Mitte des 20. Jahrhunderts wurde zunächst vor allem von den Regierungen der USA und der Sowjetunion vorangetrieben. Der Start des ersten Satelliten ‚*Sputnik 1*‘ 1957 durch die Sowjetunion markiert den Beginn einer neuen Phase der Fernerkundung. Die Akzeptanz dieses Satelliten beziehungsweise das Ausbleiben eines Protests seitens der US-Regierung sowie anderer Regierungen schuf die Grundlage für das Recht auf freien Überflug im Weltraum: die sogenannte *Open Sky Doctrine* im Gegensatz zur Doktrin der territorialen Souveränität, das sich für die luftfahrtgestützte Fernerkundung durchgesetzt hatte. Nur zwei Jahre später, 1959, schickte die CIA den ersten Spionagesatelliten ins Weltall.¹³ Nach dem Ende des Kalten Krieges beschleunigte sich die Verfügbarkeit von Satellitenbilddaten: Die Regierungen der USA und der Sowjetunion beziehungsweise Russlands gaben riesige Mengen von Daten zur zivilen und kommerziellen Nutzung frei. Gleichzeitig hoben die USA Beschränkungen für die kommerzielle Satellitenindustrie nach und nach auf. Es entwickelte sich eine Situation der zunehmenden wirtschaftlichen Konkurrenz der US-amerikanischen Anbieter mit Satellitenbilddatenanbietern beispielsweise in Frankreich und Kanada.¹⁴

Im Zuge der *OpenData*-Bewegung machen insbesondere öffentliche Einrichtungen immer mehr Daten frei zugänglich – vielfach georeferenzierte Daten wie beispielsweise Messdaten von Luftgütestationen, Planungsdaten, Verkehrsmessungsdaten und so weiter.¹⁵ Nicht zuletzt ermöglichen digitale Ortsverzeichnisse (*Gazeteers*), dass geographische Informationen, die beispielsweise in historischen Reiseberichten oder Zeitungsartikeln in unstrukturierter Form vorliegen, heutzutage rasch in

¹³ Vgl. Walter A McDougall, ... *the Heavens and the Earth: A Political History of the Space Age* (New York: Basic Books, 1985).

¹⁴ Vgl. John C. Baker, Ray A. Williamson and Kevin M. O’Connell, „Introduction“, in: *Commercial Observation Satellites: At the Leading Edge of Global Transparency*, eds. John C. Baker, Ray A. Williamson and Kevin M. O’Connell (Santa Monica, CA: Rand, 2001), 1-11.

¹⁵ Siehe beispielsweise die *Open Government Data* der Stadt Wien. (online).

strukturierte Geodaten transformiert werden können. Die geographische Referenzierung wird zu einem wichtigen Ordnungsschema digitaler Informationen.

Von der analogen zur digitalen Kartographie

Menschen machten sich schon lange vor der Entwicklung der modernen Kartographie und auf sehr unterschiedliche Arten Bilder der Welt. Die moderne, vermessende und nach geometrisch exakter Abbildung strebende Kartographie ist jedoch ein Phänomen der (europäischen) Neuzeit. Sie ist eng verwoben mit der Herausbildung der modernen Wissenschaften und der neuzeitlichen Territorialstaaten. Die jungen Territorialstaaten, insbesondere in Europa und seinen Kolonien, waren es auch, die im 18. und 19. Jahrhundert Organisationen etablierten, die lange Zeit die Kartographie dominierten.

Seit den 1960er Jahren wurde die analoge Print-Kartographie rasch und umfassend von Geographischen Informationssystemen (*GIS*) verdrängt. Diese Systeme ermöglichen es, Geodaten computergestützt zu erfassen, zu speichern, zu analysieren und (kartographisch) zu präsentieren. Die Karte steht damit nicht länger im Mittelpunkt, sie wird zu einer Präsentationsform digitaler Geoinformationen.¹⁶ *GIS* wurden und werden für verschiedene Zwecke verwendet und dies zu großen Teilen außerhalb der etablierten Organisationen der staatlichen und wissenschaftlichen Kartographie: Marktforschungsunternehmen nutzen *GIS* für Standortanalysen, die Polizei verwendet *GIS* für die räumliche Analyse von Kriminalitätsdaten, Städte und Gemeinden setzen *GIS* für die Flächennutzungsplanung ein und so fort. Nicht zuletzt war und ist das Militär, und dabei insbesondere das US-Militär, ein wichtiger Anwender und Förderer der technischen Fortentwicklung von *GIS*.¹⁷ *GIS* haben die alten topographischen Karten als Grundlage militärischer Operationen abgelöst, virtuelle 3D-Landschaftsmodelle in *GIS* ermöglichen die präzise Navigation militärischen Geräts.

Die Techniken und Praktiken von *GIS*, die wachsende Verfügbarkeit digitaler Geodaten infolge der *GPS*-Technik und der Entwicklung einer kommerziellen Satellitenbildindustrie sowie die zunehmend einfache

¹⁶ Vgl. Daniel Sui and Richard Morrill, „Computers and Geography: From Automated Geography to Digital Earth“, in: *Geography and Technology*, eds. Stanley D. Brunn, Susan L. Cutter and James W. Harrington (Dordrecht: Kluwer Academic, 2004), 81-108.

¹⁷ Vgl. Neil Smith, „History and Philosophy of Geography: Real Wars, Theory Wars“, *Progress in Human Geography* 16.2 (1992): 257-271.

Nutzung des Internets über Desktop- und mobile Computer (Smartphones) sind die wichtigsten Bausteine für die Entwicklung des *Geoweb* –¹⁸ und somit für die grundlegende Transformation von Geoinformation und kartographischer (Re-)Präsentation im digitalen Zeitalter.¹⁹ Der Begriff ‚*Geoweb*‘ wird genutzt, um die wachsende Bedeutung von Geodaten für das Internet zu beschreiben sowie den Boom neuer webbasierter Technologien, die Geodaten nutzen und vielfach produzieren. Neben virtuellen Globen wie insbesondere *Google Earth* und virtuellen Kartendiensten wie *Google Maps*, *OpenStreetMap* oder *Here* sind damit auch georeferenzierte Artikel in *Wikipedia*, georeferenzierte Bilder in *Flickr* oder georeferenzierte Meldungen in einem der *social media*-Dienste Teil des *Geoweb*s. In zunehmender Weise wird im *Geoweb* geformt, was wir über Orte und Räume der Erde wissen und letztlich vielfach auch, wie wir in der Welt agieren. Die Entwicklung des *Geoweb* wurde und wird in hohem Maße von Unternehmen bestimmt, die bis vor wenigen Jahren keinen Bezug zu Geoinformation und Kartographie hatten oder noch überhaupt nicht existierten, wie zum Beispiel *Google*, *Navteq/Here* oder *TomTom*. Gleichzeitig ermöglicht der Kontext des *Web 2.0* die Entwicklung von nichtkommerziellen, offenen Projekten wie *OpenStreetMap* und *Wikimapia*, in denen Tausende Freiwillige geographische Informationen erheben, organisieren und präsentieren – sogenannte ‚*volunteered geographic information*‘ (*VGI*).²⁰

Google, der sicherlich wichtigste Akteur des *Geoweb*s, kaufte im Jahr 2004 das *Start-up Where2Technologies*, das eine benutzerfreundliche Web-Oberfläche zur Präsentation geographischer Informationen entwickelt hatte. *Google* entwickelte die Software zum Online-Kartendienst *Google Maps* weiter, der nach dem Start 2005 rasch zur meist genutzten digitalen Kartenplattform wurde. Der Dienst bietet inzwischen neben Online-Karten beispielsweise auch fotorealistische Panoramen von Straßenzügen sowie Routenplanung und Verkehrsinformationen in Echtzeit. Ebenfalls 2004 kaufte *Google* das Unternehmen *Keyhole Inc.*, das auf der Basis einer Videospiele-Software einen virtuellen Globus aus Satellitendaten und Luftbildern entwickelt hatte und zeitweise von der *CIA*

¹⁸ Mark Graham and Matthew Zook, „Augmented Realities and Uneven Geographies: Exploring the Geolinguistic Contours of the Web“, *Environment and Planning A* 45.1 (2013): 77-99.

¹⁹ Georg Glasze, „Sozialwissenschaftliche Kartographie-, GIS- und Geoweb-Forschung“, *Kartographische Nachrichten* 64.3 (2014): 123-129.

²⁰ Michael Goodchild, „Citizens as Sensors: The World of Volunteered Geography“, *GeoJournal* 69.4 (2007): 211-221.

gefördert worden war. Basierend auf der *Keyhole*-Software entwickelte Google den virtuellen Globus *Google Earth*, der 2005 online ging und einer breiten Öffentlichkeit fotorealistische Bilder der Erde in einer bis dahin unbekanntem Qualität, Quantität und Abdeckung bietet. Bereits wenige Monate nach dem Start von *Google Maps* wurde das Programm von einem kalifornischen Informatiker gehackt und genutzt, um Immobilienangebote in Kalifornien räumlich differenziert zu präsentieren. Google erkannte, dass die Zusammenführung der *Google*-Basiskarte mit allen möglichen Arten weiterer georeferenzierter Daten neue Dienstleistungen ermöglicht und viele neue Nutzer zu *Google* führt. Rasch schuf *Google* eine Schnittstelle (*application programming interface, API*), die solche Zusammenführungen erleichtert und es auch Menschen ohne Programmier- oder Kartographieausbildung ermöglicht, sogenannte *map mashups* zu schaffen. Auch wenn die Kartendienste von *Google* derzeit zumindest für die nicht-kommerzielle Nutzung kostenfrei verfügbar sind, die zugrunde liegenden Geodaten sind allerdings nicht zugänglich und bleiben im Besitz des Unternehmens.

Bei offenen *Geoweb*-Projekten wie dem besonders erfolgreichen *Web 2.0* Kartographieprojekt *OpenStreetMap (OSM)* sind diese Daten hingegen frei verfügbar. *OSM* präsentiert sich auf der eigenen Webseite als „Projekt mit dem Ziel, eine freie Weltkarte zu erschaffen“ –²¹ vielfach wird *OSM* auch als „*Wikipedia* der Kartographie“ bezeichnet.²² Gestartet wurde das Projekt ebenfalls 2004 durch einen britischen Informatikstudenten, der von der restriktiven Lizenzpolitik des staatlichen britischen Kartographie-dienstleisters *Ordnance Survey* frustriert war. Gemeinsam mit weiteren Freiwilligen der *OpenData*-Bewegung in London schuf er die notwendige Infrastruktur zum Start des Projekts. *Mapping Parties* in immer mehr Regionen brachten neue Freiwillige zu dem Projekt, die auf der Basis von selbst erhobenen *GPS-Tracks* und Beobachtungen im Gelände Geodaten beitrugen. Die gesamten *OSM*-Geodaten sind frei nutzbar und bilden die Grundlage für zahlreiche Kartendienste und andere raumbezogene Dienstleistungen. Zehn Jahre nach dem Start von *OSM* übertrifft die Datendichte und Aktualität der *OSM*-Geodaten in vielen Regionen das Angebot staatlicher und kommerzieller Anbieter.

²¹ Vgl. *OpenStreetMap*. (online).

²² Frederik Ramm und Jochen Topf, *OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten* (Berlin: Lehmanns Media, 2010), 4.

Neue und nicht so neue Weltbilder: Ein kritischer Blick auf das Geoweb

Die sozialwissenschaftliche Auseinandersetzung mit dem neuen Boom an Geodaten und der Entwicklung des *Geoweb* baut auf den Perspektiven der Kritischen Kartographie auf. Vor dem Hintergrund der Transformation von Kartographie und Geoinformation im digitalen Zeitalter und der Bedeutung von Daten, Software und dem Internet weist die *Geoweb*-Forschung darüber hinaus Beziehungen zur sozialwissenschaftlichen Daten-, Software- und Internetforschung auf.

Unmittelbar anschließen kann die *Geoweb*-Forschung an die Ansätze einer stärker prozessualen Kartographieforschung sowie einer sozialwissenschaftlich-orientierten *GIS*-Forschung (*Critical GIS*). Die prozessuale Kartographieforschung geht in stärkerem Maße als die frühen Arbeiten der Kritischen Kartographie über Fragen der (Re-)Präsentation und damit des Kartenbildes hinaus und rückt auch die Praktiken, Konventionen und Techniken der Herstellung und Verwendung von Karten ins Blickfeld. Sie fragt danach, wie diese Praktiken an der Herstellung der sozialen Wirklichkeiten, in denen wir leben, beteiligt sind.²³ Neuere Arbeiten einer solchen ‚prozessualen Kartographie‘ beziehen ihre Anregungen vor allem aus der sozialwissenschaftlichen Wissenschafts- und Technikforschung (den ‚*science and technology studies*‘).

Parallel und vielfach in Austausch mit dieser prozessualen Kartographie entspannte sich in den 1990er Jahren auch eine Debatte über die sozio-technischen Rahmenbedingungen und sozialen Konsequenzen von *GIS* – teilweise als *GIS & Society* oder *Critical GIS*-Debatte gekennzeichnet.²⁴ Kritisiert wurde (1) der große Einfluss wirtschaftlicher und militärischer Interessen auf die Entwicklung von *GIS*,²⁵ (2) die Ungleichheiten in der Gewinnung und im Zugang zu geographischen Informationen, welche die lange Zeit teuren und komplexen Geoinformationssysteme mit sich brachten, sowie nicht zuletzt (3) die Fokussierung von *GIS*-Analysen auf quantifizierbare und metrische Informationen, auf ‚Zählen‘ und ‚Messen‘ und damit die Gefahr einer Marginalisierung von verstehender Sozialforschung.

²³ Vgl. John Pickles, *A History of Spaces: Cartographic Reason, Mapping, and the Geo-Coded World* (London: Routledge, 2004), 5.

²⁴ Für einen Überblick siehe Nadine Schuurman, „Trouble in the Heartland: GIS and its Critics in the 1990s“, *Progress in Human Geography* 24.4 (2000): 569-590.

²⁵ Zur Kritik an der Nähe von GIS zu militärischen Operationen siehe beispielsweise Smith, „History and Philosophy of Geography“.

Die sich konstituierende *Geoweb*-Forschung untersucht die sozio-technischen Prozesse, in denen geographische Informationen im digitalen Zeitalter hergestellt und genutzt werden. Sie fragt danach, welche neuen (und teilweise auch nicht so neuen) Weltbilder im *Geoweb* geformt und vermittelt werden.²⁶ Dabei lassen sich (mindestens) zwei Spannungsfelder identifizieren:²⁷

- Universalisierung *versus* Personalisierung/Kommerzialisierung und fortdauernde Fragmentierung: Zumindest die Staaten des globalen Nordens konnten bis weit ins 20. Jahrhundert hinein weitgehend die Produktion von Karten und allen möglichen Formen geographischer Informationen innerhalb ihrer Grenzen kontrollieren. Mit der wachsenden Verfügbarkeit von Satellitenbilddaten und dem Erfolg kommerzieller Akteure im *Geoweb* sowie den offenen Initiativen einer *volunteered geographic information* scheinen Geoinformationen heute jedoch immer weniger durch die Nationalstaaten kontrolliert werden zu können. Neue Akteure wie *Google* oder *OSM* versprechen hingegen universell-globale Geoinformationen. Allerdings zeigen sich auch Gegenteiligkeiten. So gab *Google* 2013 mit der Einführung einer neuen Version von *Google Maps* die Idee einer universellen Weltkarte auf: Je nach Suchanfrage, den besuchten Orten, dem jeweiligen individuellen Verlauf bisheriger Suchanfragen und besuchter Orte, den Spracheinstellungen und der Lokalisierung des abrufenden Computers personalisieren die Algorithmen von *Google* die Inhalte der Karte. Der Grund liegt im Geschäftsmodell von *Google*: gezielte, also möglichst personalisierte Werbung. So führen die kommerziellen Interessen privatwirtschaftlicher *Geoweb*-Dienstleister dazu, dass ‚geographische Filterblasen‘ entstehen. Daneben werden aber auch nationalstaatliche Fragmentierungen reproduziert. Einige *Geoweb*-Anbieter passen inzwischen die präsentierten Geoinformationen dem jeweiligen nationalen Kontext an: So unterscheidet *Google* seit 2014 beispielsweise drei kartographische Präsentationen der Halbinsel Krim. Für Computer mit IP-Adressen aus der Ukraine wird die Krim als Teil der Ukraine dargestellt, für IP-Adressen aus Russland ist die Ukraine durch eine nationale Grenze von der Ukraine abgetrennt und Teil Russlands,

²⁶ Vgl. Glasze, „Sozialwissenschaftliche Kartographie“.

²⁷ Vgl. Georg Glasze, „Neue Kartografien, neue Geografien: Weltbilder im digitalen Zeitalter“, *APUZ* 65.41-42 (2015): 29-37.

für alle anderen Internetnutzer zeigt *Google* eine gestrichelte Linie im Norden der Krim als umstrittene Grenze.

- Öffnung und Demokratisierung *versus* alte und neue Exklusionen: Insbesondere die Projekte der *volunteered geographic information* sind vielfach als ‚Öffnung‘ beziehungsweise ‚Demokratisierung‘ der Kartographie sowie der gesamten Geoinformation begrüßt worden:²⁸ Neue Akteure bekommen Zugang zu Geoinformationen, und es eröffnen sich Chancen, bislang ‚verschwiegene‘ Informationen zu vermitteln. Inzwischen konnten jedoch mehrere Studien belegen, dass sich auch in den *VGI*-Projekten Fragen von Zugang und Exklusion stellen. Vielfach werden diese Projekte von soziodemografisch sehr homogenen Gruppen geprägt: Die Teilnehmenden sind überwiegend männlich, technik-affin, jung, europäisch oder nord-amerikanisch und weiß.²⁹ Diese Ungleichheit prägt, welche Daten und wo sie erhoben werden, sowie die Art und Weise, wie diese Daten verarbeitet und präsentiert werden (siehe Abbildung 1).³⁰

Nicht zuletzt wächst mit dem *Geoweb* die Bedeutung von Code und Software für die Organisation und Vermittlung von Geoinformation und damit ein Bereich, dessen Funktionsweise und Entwicklung für die allermeisten Nutzer kaum einsichtig und verständlich ist.

²⁸ Vgl. Chris Perkins and Martin Dodge, „The Potential of User-Generated Cartography: A Case-Study of the OpenStreetMap Project and Manchester Mapping Party“, *North West Geography* 8.1 (2008): 19-32.

²⁹ Vgl. Mordechai Haklay, „Neogeography and the Delusion of Democratisation“, *Environment and Planning A* 45.1 (2013): 55-69.

³⁰ Für ein lokales Beispiel vgl. Christian Bittner, „Reproduktion sozialräumlicher Differenzierungen in OpenStreetMap: Das Beispiel Jerusalem“, *Kartographische Nachrichten* 64.3 (2014): 136-144.

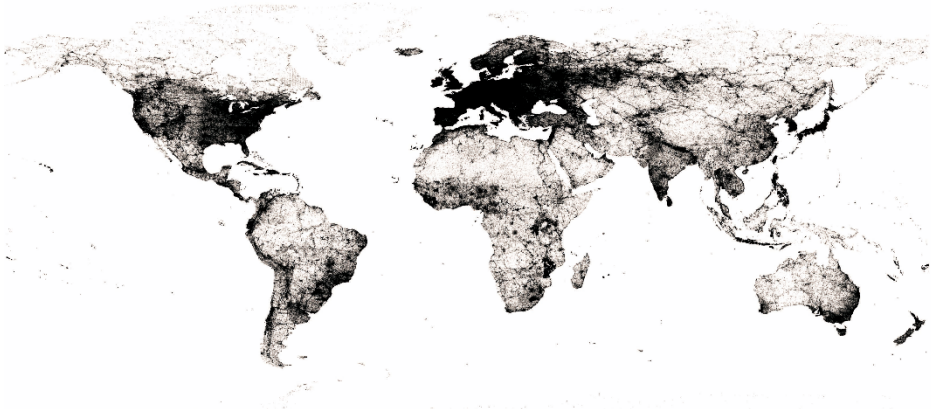


Abbildung 1: Ungleiche Geographien von *OpenStreetMap* im weltweiten Vergleich (je dunkler, desto mehr Bearbeitungen wurden vorgenommen).³¹

Ausblick: Digitale Geographien

Vor dem Hintergrund der stetig wachsenden Menge georeferenzierter Daten sowie deren Einbettung in viele neue Praktiken können digitale Geoinformationen nicht länger nur als (Re-)Präsentation von Räumen betrachtet werden. Sie werden vielmehr zu einem wichtigen Teil gesellschaftlicher Räumlichkeit. Mit dem Schlagwort der ‚*software sorted geographies*‘ hat Steven Graham schon 2005 darauf hingewiesen, wie im Zusammenspiel von (Geo-)Daten und Software über die Zugänglichkeit von Räumen ‚entschieden‘ wird – beispielsweise in elektronischen Zugangssystemen an Flughäfen.³² Noch umfassender ist das Konzept der ‚augmentierten Geographien‘, wie es Graham und Zook vorgeschlagen haben.³³ Sie bezeichnen damit die Erweiterung von Raumbeschreibungen/Geographien um digitale Informationen: Beschreibungen von Orten in sozialen Medien, Bewertungen von Hotels oder Restaurants in Bewertungsplattformen, Projekte der *Web 2.0*-Kartographie und so weiter.

³¹ Vgl. Finn Dammann, *Sprachgeographie in OpenStreetMap: Eine geolinguistische Metadaten-Analyse* (Unveröffentlichte Masterarbeit am Institut für Geographie der FAU Erlangen-Nürnberg, 2016).

³² Vgl. Steven Graham, „Software-Sorted Geographies“, *Progress in Human Geography* 29.5 (2005): 562-580.

³³ Vgl. Graham and Zook, „Augmented Realities and Uneven Geographies“.



Abbildung 2: Eine Smartphone-App veranschaulicht ‚augmentierte Geographien‘ (flickr.com/photos/lge/5708231997; 10.01.2017).

Die immer spezifischeren sozio-technischen Entstehungshintergründe von Daten – beispielsweise die Dominanz kommerzieller Interessen im *Geoweb*, aber auch die neuen und alten Ungleichheiten, welche die offenen Projekte der *volunteered geographic information* prägen, führen dazu, dass (auch) diese digitalen Geographien höchst ungleich sind. Da sie aber in zunehmendem Maße prägen, was wir über die Welt wissen, und wie wir in der Welt leben, wird es zur Aufgabe einer wissenschaftlichen Sub-Disziplin ‚Digitale Geographie‘, die sozio-technischen Hintergründe dieser Geographien zu erforschen und sie für öffentliche Debatten zugänglich zu machen.

Literaturverzeichnis

- Baker, John C., Ray A. Williamson and Kevin M. O'Connell. „Introduction“. In: *Commercial Observation Satellites: At the Leading Edge of Global Transparency*. Eds. John C. Baker, Ray A. Williamson and Kevin M. O'Connell. Santa Monica, CA: Rand, 2001. 1-11.
- Barnes, Trevor. „Big Data, Little History“. *Dialogues in Human Geography* 3.3 (2013): 297-302.
- Bittner, Christian. „Reproduktion sozialräumlicher Differenzierungen in OpenStreetMap: Das Beispiel Jerusalem“. *Kartographische Nachrichten* 64.3 (2014): 136-144
- Boyd, Danah and Kate Crawford. „Critical Questions for Big Data“. *Information, Communication & Society* 15.5 (2012): 662-679.
- Dammann, Finn. *Sprachgeographie in OpenStreetMap: Eine geolinguistische Metadaten-Analyse*. Unveröffentlichte Masterarbeit am Institut für Geographie der FAU Erlangen-Nürnberg, 2016.
- Dodge, Martin, Rob Kitchin and Mathew Zook. „How Does Software Make Space? Exploring Some Geographical Dimensions of Pervasive Computing and Software Studies“. *Environment and Planning A* 41.6 (2009): 1283-1293.
- Glasze, Georg. „Kritische Kartographie“. *Geographische Zeitschrift* 97.4 (2009): 181-191.
- . „Sozialwissenschaftliche Kartographie-, GIS- und Geoweb-Forschung“. *Kartographische Nachrichten* 64.3 (2014): 123-129.
- . „Neue Kartografien, neue Geografien: Weltbilder im digitalen Zeitalter“. *APUZ* 65.41-42 (2015): 29-37.
- Goodchild, Michael. „Citizens as Sensors: The World of Volunteered Geography“. *GeoJournal* 69.4 (2007): 211-221.
- Graham, Mark and Matthew Zook. „Augmented Realities and Uneven Geographies: Exploring the Geolinguistic Contours of the Web“. *Environment and Planning A* 45.1 (2013): 77-99.
- Graham, Steven. „Software-Sorted Geographies“. *Progress Human Geography* 29.5 (2005): 562-580.
- Haklay, Mordechai. „Neogeography and the Delusion of Democratisation“. *Environment and Planning A* 45.1 (2013): 55-69.
- Harley, John Brian. „Deconstructing the Map“. *Cartographica* 26.2 (1989): 1-20.
- Kohlstock, Peter. *Kartographie: Eine Einführung*. Paderborn: UTB, 2004.
- Kumar, Saamer and Kevin Moore. „The Evolution of Global Positioning System (GPS) Technology“. *Journal of Science Education and Technology* 11.1 (2002): 59-80.

- Leser, Hartmut. *Diercke Wörterbuch Allgemeine Geographie*. München: dtv, 2005.
- McDougall, Walter. ... *the Heavens and the Earth: A Political History of the Space Age*. New York: Basic Books, 1985.
- Perkins, Chris and Martin Dodge. „The Potential of User-Generated Cartography: A Case Study of the OpenStreetMap Project and Manchester Mapping Party“. *North West Geography* 8.1 (2008): 19-32.
- Pickles, John. *A History of Spaces: Cartographic Reason, Mapping, and the Geo-Coded World*. London: Routledge, 2004.
- Ramm, Frederik und Jochen Topf. *OpenStreetMap: Die freie Weltkarte nutzen und mitgestalten*. Berlin: Lehmanns Media, 2010.
- Schuurman, Nadine. „Trouble in the Heartland: GIS and its Critics in the 1990s“. *Progress in Human Geography* 24.4 (2000): 569-590.
- Smith, Neil. „History and Philosophy of Geography: Real Wars, Theory Wars“. *Progress in Human Geography* 16.2 (1992): 257-271.
- Sturdevant, Rick. „The Navstar Global Positioning System: From Military Tool to Global Utility“. In: *Down to Earth: Satellite Technologies, Industries, and Cultures*. Eds. Lisa Parks and James Schwoch. New Brunswick N.J: Rutgers University Press, 2012. 99-121.
- Sui, Daniel and Richard Morrill. „Computers and Geography: From Automated Geography to Digital Earth“. *Geography and Technology*. Eds. Stanley D. Brunn, Susan L. Cutter and James W. Harrington. Dordrecht: Kluwer Academic, 2004. 81-108.

Internetquellen

- Anderson, Chris. „The End of Theory: The Data Deluge Makes the Scientific Method Obsolete“. *Wired Magazine* 16.8 (2008).
<https://www.wired.com/2008/06/pb-theory/> [zuletzt besucht 07.02.2017].
- Kitchin, Rob und Tracey P. Lauriault. „Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work“. *The Programmable City Working Paper* 2.
http://eprints.maynoothuniversity.ie/5683/1/KitchinLauriault_CriticalDataStudies_ProgrammableCity_WorkingPaper2_SSRN-id2474112.pdf [zuletzt besucht 10.12.2016].
- OpenStreetMap*. <https://www.openstreetmap.de>
[zuletzt besucht 13.03.2017].

„Außer Kontrolle“: Alte und neue Informationsfluten im Publikationswesen

SVENJA HAGENHOFF

Abstract: Technologische, organisatorische und institutionelle Entwicklungen haben im Laufe der Zeit den Kreis derjenigen vergrößert, die an Kommunikation teilhaben und insbesondere auch selber als Sender Informationen und Meinungen öffentlich kommunizieren können. Bereits der Druck mit beweglichen Lettern, aber auch Fortschritte in der Papierproduktion, bei physischen und organisatorischen Distributionsinfrastrukturen, bei Schreib- und Lesefähigkeiten und jüngst die Digitalisierung als Totalphänomen haben kontinuierlich zu Entgrenzungen in der Art geführt, dass die Menge und die Vielfalt an Inhalten im Laufe der Zeit immer umfänglicher geworden sind. Lange Zeit waren die Prozesse des Publizierens von Inhalten in der Hand professioneller Akteure, die als Verleger oder Drucker *know-how* und Technologien innerhalb von Organisationen akkumuliert haben und sowohl Publikationsprozesse als auch die Beschaffenheit von Medien vollständig kontrollieren konnten. Im Zuge vor allem technologischer Veränderungen in jüngerer Zeit verwischen die Leistungsrollen der Erzeuger und der Rezipienten von Inhalten: Vormals passive Rezipienten werden aktive Kommunikatoren und Miterzeuger von Inhalten, die sich in Menge, Vielfalt und Beschaffenheit der Kontrolle etablierter *gatekeeper* entziehen. Das Verhältnis von Nutzsignal zu Rauschen wird tendenziell immer ungünstiger, sodass Filter benötigt werden, die den Rezipienten das Relevante zugänglich machen.

Informationsflut, Kommunikationssystem und Publikationswesen

Der Begriff Informationsflut wurde von Alvin Toffler geprägt (engl. *information overload*).¹ Er charakterisiert die Situation, dass der Mensch über ein ‚Zuviel‘ an Informationen zur Meinungsbildung und Entscheidungsfindung verfügt und darin förmlich ertrinkt. Der Abgleich des Informationsbedarfs eines Rezipienten mit dem Angebot an Information gestaltet sich mit zunehmender Informationsmenge immer schwieriger, da das Verhältnis von Nutzsignal zu Rauschen immer ungünstiger wird. Als eine Ursache für die anwachsenden Informationsfluten werden die kontinuierlichen Fortschritte bei der Gestaltung von Kommunikationssystemen genannt, dargelegt zum Beispiel bei Ben H. Bagdikian anhand

¹ Alvin Toffler, *Future Shock* (New York: Random House, 1970), 350. In der deutschen Übersetzung des Buchs wird der Begriff *Informationsüberfütterung* verwendet. Alvin Toffler, *Der Zukunftsschock* (Bern: Scherz, 1971), 280.

der Beispiele von Buchdruck, Rotationsmaschine, Eisenbahn, Ozeandampfer und elektronischem Telegraph.² Die große Bedeutung, die man Kommunikationssystemen und den Effekten der Veränderungen daran für Gesellschaften beimisst, verdichtet sich im Diskurs dann auch häufig im Begriff der ‚Kommunikationsrevolutionen‘.³

Ein Kommunikationssystem in diesem Sinne ist ein komplexes Gewebe aus *Akteuren*, die Informationen erzeugen (Kommunikator, Autor), distribuieren sowie rezipieren (Rezipient); *Medien* als Artefakte zum Transport und gegebenenfalls zur Speicherung von Informationen; *Technologien* zur *Produktion* dieser Medien sowie *Infrastrukturen* und *Organisationsleistungen* zur physischen wie akquisitorischen *Distribution* dieser Medien hin zu den potenziellen Rezipienten.⁴ Auf das Kommunikationssystem wirken *Institutionen*, verstanden als Regelwerke, die die Verhaltensweisen der handelnden Akteure fördern oder limitieren und damit maßgeblich die Funktionalität eines konkreten Kommunikationssystems beeinflussen: Positiv wirken auf die Diffusion von Informationen beispielsweise gesetzlich gesicherte Informations- und Meinungsfreiheiten, ein Mindestmaß an Schulbildung oder Regelungen zur Preisgestaltung von Medien, wie sie beispielsweise in Deutschland mit der Festlegung des verminderten Mehrwertsteuersatzes, der gesetzlich fixierten Buchpreisbindung und der vom Gesetz gegen Wettbewerbsbeschränkungen tolerierten, branchenintern festgelegten Preisbindung von Zeitungen und Zeitschriften existieren.

Kommunikation wird verstanden als ein Prozess aus absichtsvollen Mitteilungen von Informationen zwischen Akteuren. Der Kommunikator verknüpft Zeichen aus einem gegebenen Repertoire – zum Beispiel Buchstaben des lateinischen Alphabets – nach bestimmten Regeln oder Mustern der Zeichenverwendung, von denen er annimmt, dass potenzielle

² Vgl. Ben H. Bagdikian, *The Information Machines: Their Impact on Men and the Media* (New York: Harper & Row, 1971), 3ff.

³ Vgl. beispielsweise Michael North, hg., *Kommunikationsrevolutionen: Die neuen Medien des 16. und 19. Jahrhunderts* (Köln: Böhlau, 1995); Wolfgang Behringer, *Im Zeichen des Merkur: Reichspost und Kommunikationsrevolution in der Frühen Neuzeit* (Göttingen: Vandenhoeck & Ruprecht, 2003); Bill Kovarik, *Revolutions in Communication: Media History from Gutenberg to the Digital Age* (New York: Continuum, 2011).

⁴ Der Begriff ‚physisch‘ bezieht sich auf die ‚Überbrückung von Raum und Zeit‘; unter ‚akquisitorischer Distribution‘ versteht man die ‚Schaffung von Markttransparenz‘, damit Angebot und Nachfrage adäquat zusammenfinden. Vgl. auch Svenja Hagenhoff, „Distribution“, in: *Reclams Sachlexikon des Buches*, hg. Ursula Rautenberg (Stuttgart: Reclam, 2015), 124.

Rezipienten diese Regeln kennen. Wenn Rezipienten über diese semiotische Kompetenz verfügen,⁵ können sie die Mitteilung entschlüsseln und ihre Schlussfolgerungen aus dem Mitgeteilten ziehen. Seit Jahrhunderten kommuniziert der Mensch erfolgreich mittels Schrift- und Lesemedien: Bücher, Flugblätter und -schriften, Zeitungen, Magazine oder *Blogs* als Publikationen dienen dem Menschen der raum- und zeitunabhängigen Massenkommunikation. Informationen und Gedanken werden unabhängig von Personen transportiert und archiviert, die Inhalte der Publikationen lassen sich an größere, dem Kommunikator persönlich nicht bekannte, disperse Publika (Massenpublika) vermitteln.⁶

Im Laufe der Entwicklungsgeschichte des zeit- und raumunabhängigen Kommunikationswesens haben technische wie auch organisatorische und institutionelle Veränderungen dazu geführt, dass schriftbasierte Publikationen in immer größerer Menge hergestellt und verteilt werden konnten. Die Produktions- und Distributionskosten wurden tendenziell immer geringer. Auch hat sich der Kreis derjenigen, die an schriftbasierter Kommunikation als Kommunikatoren oder Rezipienten teilhaben konnten, verändert und vor allem vergrößert.⁷ Die quantitativen und qualitativen Veränderungen stellen ein Phänomen dar, das von jeher von eben jenen Klagen über ‚Fluten‘, ‚Vermassung‘, ‚Entgrenzung‘, ‚Kontrollverlust‘ oder ‚Sinnverlust‘ begleitet wurde und wird.

Ziel des Beitrags ist es aufzuzeigen, dass entstehende Fluten an Informationen oder Inhalten kein Phänomen der Digitalisierung sind, sondern ein kontinuierlicher Begleiter mediengestützter, raum- und zeitunabhängiger Kommunikation. Schon immer waren deswegen Filtermechanismen erforderlich, die dabei helfen, das Angebot an Informationen zu sichten, zu bewerten, zu sortieren und zu selektieren. Der Beitrag gliedert sich wie folgt: Im nächsten Kapitel werden zunächst Veränderungen des Publikationswesens als spezielle Ausprägung eines Kommunikationssystems in groben Zügen erläutert. Anschließend wird anhand einiger Beispiele zu konstatierten Fluten verschiedenster Ausprägungen das Dauerphänomen Kulturkritik dargelegt. Das folgende Kapitel verdeutlicht die Flut in Mengen und Gewicht exemplarisch

⁵ Vgl. Rudi Keller, *Zeichentheorie: Zu einer Theorie semiotischen Wissens* (Tübingen: Francke, 1995).

⁶ Vgl. Gerhard Maletzke, *Psychologie der Massenkommunikation: Theorie und Systematik* (Hamburg: Hans Bredow-Institut, 1963), 33-35.

⁷ Vgl. Michael Giesecke, *Von den Mythen der Buchkultur zu den Visionen der Informationsgesellschaft: Trendforschungen zur kulturellen Medienökologie* (Frankfurt a.M.: Suhrkamp, 2002), 81.

anhand des aktuellen deutschen Buchmarkts. Das letzte Kapitel widmet sich der Notwendigkeit der Informations- und Medienselektion und formuliert ein Forschungsdesiderat.

Veränderung des Publikationswesens im Laufe der Zeit

Das Publikationswesen kann anhand von technologischen Veränderungen in Bezug auf die Produktion von Schrift- und Lesemedien in vier große Abschnitte unterteilt werden:⁸ in das Zeitalter der handwerklichen Produktion, das Zeitalter der mechanischen Produktion, das Zeitalter der industriellen Produktion und das Zeitalter der Digitalisierung. Die Abschnitte sind nicht trennscharf zueinander. Es gab und gibt Phasen der Überlappungen und des Nebeneinanders verschiedener Arbeitsweisen und daraus resultierender Medienarten.

Vor Erfindung des Drucks mit beweglichen Lettern waren Schriftmedien typischerweise keine Druckerzeugnisse, sondern Handschriften. Schriftmedien wurden dadurch erzeugt, dass Schreiber Texte handschriftlich niedergeschrieben beziehungsweise abgeschrieben haben (*ars artificialiter scribendi*).⁹ Die Produktion war handwerklich geprägt, da Schriftwerke in der Regel Auftragsarbeiten waren, die nur angefertigt wurden, wenn ein individuelles Lese- oder Besitzinteresse bestand.¹⁰ Der Produzent des gefertigten Schriftmediums kannte den Empfänger meistens, sodass das unternehmerische Absatzrisiko gering blieb. Handschriften waren eine relativ teure und damit auch knappe Ware, die nicht Jedermann zu Kommunikations- wie auch Rezeptionszwecken zur Verfügung stand. Aufgrund geringer Alphabetisierungsquoten hatten sie eine vergleichsweise eingeschränkte Wirkmächtigkeit. Darüber hinaus war das gesamte Kommunikationssystem von wenig Freiheitsgraden für die Akteure gekennzeichnet: Die Schreiber befanden sich in der Regel in monastischen oder universitären und damit hoch regulierten und kontrollierten Arbeitsumgebungen.¹¹ Eine Öffentlichkeit als Adressat von

⁸ Die folgenden Ausführungen stammen in Teilen aus Svenja Hagenhoff, „Buch und Buchsachgruppen“, in: *Handbuch Medienökonomie*, hg. Jan Krone und Tassilo Pellegrini (Berlin: Springer, 2016). (online).

⁹ Vgl. Michael Giesecke, *Der Buchdruck in der frühen Neuzeit: Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien* (Frankfurt a.M.: Suhrkamp, 2006), 67ff.

¹⁰ Vgl. Ute Schneider, „Wozu lesen? Persistente Funktionen des Lesens im sozialen Kontext“, *IASL* 39.1 (2014): 268-283.

¹¹ Vgl. hierzu z.B. das Pecien-System der Vervielfältigung der in Universitäten benötigten Literatur.

Schriftkommunikation im heutigen Sinne existierte nicht, im Gegenteil, die Vergemeinschaftung privater Gedanken war legitimierungsbedürftig.¹²

Das Zeitalter des mechanischen Drucks beginnt mit der Erfindung der beweglichen Lettern durch Gutenberg um 1450. Aus dem Handwerk wurde ein maschinengestützter Produktionsprozess (*ars nova ingeniosa*).¹³ Werkstoffe in Form flexibel einsetzbarer, wiederverwendbarer Typen sowie ein größeres Produktionsmittel in Form der Druckerpresse erforderten größere Investitionen zur Herstellung der Produktionsbereitschaft, damit verbunden war die Notwendigkeit einer hohen Auslastung der Produktionsanlage. Schriftmedien wurden nun in höherer Auflage auf Vorrat für einen anonymen Massenmarkt produziert. Diese Vorratsproduktion hat dazu geführt, dass in der Regel die produzierte Menge und die nachgefragte Menge nicht mehr zusammenpassten. Es entstand das bis heute anhaltende Phänomen des ‚Vergriffen-Seins‘, das Kaufwillige unversorgt lässt, aber sehr viel häufiger auch das des Makulierens, also der Notwendigkeit, überschüssige Druckexemplare zu entsorgen beziehungsweise den wertvollen, da knappen Beschreibstoff anderweitig zu verwerten. Die Veränderung der Produktion führte zwangsläufig auch zu einer Veränderung der Organisationsleistung auf der Distributionsstufe, denn der anonyme Massenmarkt musste zunächst über das Vorhandensein neuer Buchtitel informiert werden (akquisitorische Distribution), und die produzierten Exemplare mussten zu den potenziellen Käufern transportiert werden (physische Distribution). Wolfgang Behringer verdeutlicht, dass die viel zitierte ‚Gutenberg-Galaxis‘ eigentlich eine ‚Taxis-Galaxis‘ ist, denn ohne gut ausgebaute und standardisierte Transportsysteme kann das Produzierte räumlich nicht verteilt werden, und schriftbasierte Informationen können nicht zur Zirkulation gelangen.¹⁴ Gleichermäßen stellt Uwe Spiekermann fest, dass die nicht-handwerkliche Massenproduktion ohne güterverbreitende Infrastrukturen wertlos ist.¹⁵ Im Laufe des langen Zeitalters des mechanischen Drucks veränderten sich aber nicht nur die Produktionsweise und daraus resultierend die distributiven Infrastrukturen. Verändert haben sich auch die Arten von Autoren und die Arten von Lesern: Schriftlich kommuniziert haben nun auch Nicht-Approbierte und

¹² Vgl. Giesecke, *Buchdruck*, 189.

¹³ Vgl. ebd., 67ff.

¹⁴ Vgl. Wolfgang Behringer, „Von der Gutenberg-Galaxis zur Taxis-Galaxis: Die Kommunikationsrevolution: Ein Konzept zum besseren Verständnis der Frühen Neuzeit“, *Historische Zeitschrift: Beihefte* 41 (2005): 39-54.

¹⁵ Vgl. Uwe Spiekermann, *Basis der Konsumgesellschaft: Entstehung und Entwicklung des modernen Kleinhandels in Deutschland 1850-1914* (München: Beck, 1999), 138-218.

damit Jedermann. Die Leserschaft beschränkte sich nicht mehr auf institutionell Privilegierte, sondern der ‚Laie‘ oder der ‚gemein man‘ wurde zum Leser.¹⁶

Das Zeitalter der industriellen Produktion ab Anfang des 19. Jahrhunderts hat im Vergleich zum mechanischen Druck vor allem Geschwindigkeits-, Mengen- und Kostenveränderungen gebracht. Zwei ganz wesentliche Veränderungen können identifiziert werden: Zum einen hat es mit der Entwicklung der Papiermaschine Ende des 18. Jahrhunderts sowie der Erfindung des Holzschliffpapiers 1843 große Fortschritte in der Papierproduktion gegeben,¹⁷ die dazu geführt haben, dass sich ausgeprägte Knappheitsphänomene und damit auch die Kostensituation für diesen zentralen Rohstoff deutlich entspannt haben. Zum anderen hat die Erfindung der Schnellpresse zu Beginn des Jahrhunderts zusammen mit der Entwicklung von Möglichkeiten der industriellen Massenproduktion der benötigten Bleilettern ab den 1860er Jahren – statt handgegossen nun aus einer Gießmaschine stammend – den Ausstoß von Druckwerken pro Zeiteinheit massiv verändert:¹⁸ Während um 1800 die *Stanhope*-Presse noch immer Arbeitsschritte erforderte, die im Wesentlichen denen in der Gutenberg'schen Offizin (Handsatz, Handpresse, Bedrucken von einzelnen Blättern) entsprachen und einen Ausstoß von circa 250 Druckbogen pro Stunde erzeugte, hat die erste Schnellpresse 1814 bereits 1 100 Bogen Papier pro Stunde bedruckt (Vervierfachung). Die *Walter*-Rotationsmaschine ab circa 1862 kam auf bereits 10 500 Druckbogen pro Stunde (Faktor 40), und 1895 hat der Ausstoß 36 000 Bogen betragen (Faktor 144 zum Jahr 1800). Ein moderner Rollenoffset schafft heute 40 000 Bogen à 16 Seiten pro Stunde, das ist eine Steigerung des Ausstoßes um den Faktor 60 im Vergleich zur Mitte des 19. Jahrhunderts und eine Steigerung um den Faktor 2 500 im Vergleich zum Jahr 1800. Helen Müller legt dar, wie der technologische Fortschritt im Zusammenspiel mit massiv fortschreitender Alphabetisierung, verbesserten Transportinfrastrukturen und erhöhtem Urbanisierungsgrad ab der Mitte des 19. Jahrhunderts zu

¹⁶ Vgl. Helen Müller, „Verlagswesen und europäische Massenkommunikationsgesellschaft um 1900“, *IASL* 27 (2002) 1: 170-197, hier 175-177; vgl. auch Michael Giesecke, *Buchdruck*, 329.

¹⁷ Vgl. Alexander Friebel, „Ohne Papier keine Zeitung: Die Erfindung des Holzschliffpapiers als Meilenstein in der Entwicklung des deutschen Pressewesens“, *Jahrbuch für Kommunikationsgeschichte* 3 (2001): 132-156.

¹⁸ Folgende Leistungsdaten nach Martyn Lyons, *Das Buch: Eine illustrierte Geschichte* (Hildesheim: Gerstenberg Verlag, 2012), 132-134; sowie Alexander Friebel, „Ohne Papier keine Zeitung“, hier 135.

einem ausdifferenzierten Verlags- und Pressewesen führte und die Gesellschaften der Länder Deutschland, Frankreich und Großbritannien zu hochentwickelten Massenkommunikationsgesellschaften transformierte.¹⁹

Das Zeitalter der Digitalisierung ist durch eine umfassende Entmaterialisierung gekennzeichnet: Die Digitalisierung in Kombination mit der Elektronifizierung erlaubt für Schrift- und Lesemedien Produktions- und auch Distributionsprozesse, die von spezifischen Materialitäten und Übertragungstechnologien, wie Papier oder mechanischen Teilen, vollständig unabhängig sind. Knappheitsphänomene, Limitationen und ökonomisch bedingte Bündelungsnotwendigkeiten (zum Beispiel mehrere Artikel zu einer Zeitung) schwinden in großen Teilen: Weder knapper und damit wertvoller Beschreibstoff, noch knapper Platz bei Transport, Lagerung und Präsentation limitieren das Erzeugen und Distribuieren von Schriftmedien. Das seit dem mechanischen Druck bekannte Zentralproblem des ‚mismatch‘ zwischen Produktionsmenge und Nachfragemenge existiert nicht mehr, da die Produktionsmenge exakt der Distributionsmenge entspricht: Ein Exemplar eines Schriftmediums wird durch den Rezipienten per Download der Datei oder Seitenaufruf erzeugt. Zudem ist die Digitaltechnologie eine unspezifische Technologie,²⁰ die nicht auf spezifische – vorab bekannte Verwendungszwecke – limitiert ist und stattdessen für jedwede Problemstellung und für jeden Aufgabenbereich eingesetzt werden kann. Auf einmal geschaffene technologische Basisinfrastrukturen können Applikationen aufgesetzt werden, die relativ einfach von jedermann selbstständig zu bedienen sind, der in der Lage ist, mit Tastatur, Tablet oder Smartphone umzugehen. Clay Shirky spitzt daher zu, dass ‚Publizieren‘ kein komplexer Beruf mehr sei, sondern nur ein Knopf, auf dem das Wort ‚publizieren‘ stünde. Drückt man diesen, so hat das Publizieren – und damit auch die bisher in Bezug auf Produktion und Distribution recht anspruchsvolle schriftbasierte Kommunikation – stattgefunden:

Publishing is not evolving. Publishing is going away. Because the word ‘publishing’ means a cadre of professionals who are taking on the incredible difficulty and complexity and expense of making something

¹⁹ Vgl. Müller, „Verlagswesen“, 171.

²⁰ ‚General Purpose Technology‘ nach Elhanan Helpman, hg., *General Purpose Technologies and Economic Growth* (Cambridge, Mass.: MIT Press, 1998).

public. That's not a job anymore. That's a button. There's a button that says 'publish', and when you press it, it's done.²¹

Die ubiquitäre Verfügbarkeit von Publikationsinfrastrukturen sowie deren einfache Bedienung und Verwendung erzeugen nach Einschätzung von Medienforschern ganz neue Dimensionen von Quantität: Die kulturelle Entwicklung reduziert sich nicht mehr auf einige privilegierte Produzenten, sondern wird durch die Arbeit von zig Millionen nicht-professionellen Medienproduzenten (*user generated content*) geschaffen; das Wachstum dieser Inhalte verläuft exponentiell.²² Die bisher etablierten Grenzen zwischen den Leistungsrollen der aktiven Produzenten einerseits und der passiven Rezipienten andererseits verwischen aufgrund der jüngeren technologischen Entwicklungen,²³ die entstehende neue Sozialfigur wird als ‚*prosumer*‘ –²⁴ eine Verschmelzung der Begriffe Produzent (*producer*) und Konsument (*consumer*) – bezeichnet. Charles Wright Mills hat bereits 1956 in ähnlicher Weise die Konzepte von ‚Publikum‘ und ‚Masse‘ unterschieden:²⁵ Für ihn zeichnet sich *public communication* dadurch aus, dass prinzipiell genauso viele Personen Meinungen äußern wie empfangen können. Jede öffentlich geäußerte Meinung kann unmittelbar in der Öffentlichkeit kommentiert werden und vollzieht sich ohne autoritative Institutionen. ‚*Mass communication*‘ hingegen ist dadurch gekennzeichnet, dass deutlich weniger Personen Kommunikatoren sind als Empfänger von Kommunikation. Eine direkte Antwort auf das

²¹ Clay Shirky, *How We Will Read*. (online). (Für diese und alle folgenden Angaben von Internetquellen in den Fußnoten wird auf die vollständige Nennung im Literaturverzeichnis verwiesen.)

²² Vgl. Lev Manovich, „Auf den Spuren der globalen digitalen Kulturen: Kulturanalytik für Anfänger“, *Dossier der Bundeszentrale für Politische Bildung*, 2011 (online); zur Menge der produzierten und digital vorliegenden Informationen vgl. Peter Lyman and Hal Varian, „How Much Information?“, *The Journal of Electronic Publishing* 6.2 (2000). (online).

²³ Vgl. Ulrich Dolata und Jan-Felix Schrape, „Zwischen Individuum und Organisation: Neue kollektive Akteure und Handlungskonstellationen im Internet“, *Stuttgarter Beiträge zur Organisations- und Innovationsforschung*, Nr. 02 (Stuttgart: Universität Stuttgart: 2013). Die Möglichkeit der Aufhebung der Trennung dieser Rollen wurde schon von Enzensberger in seiner Medientheorie mit Bezug auf Bertold Brechts Radiotheorie den elektronischen Medien zugesprochen; vgl. Hans Magnus Enzensberger, „Baukasten zu einer Theorie der Medien“, *Kursbuch* 20 (1970): 159-186.

²⁴ Differenziert zu diesem Begriff siehe Dolata und Schrape, „Zwischen Individuum und Organisation“.

²⁵ Vgl. Charles Wright Mills, *The Power Elite* (New York: Oxford University Press, 1959), 303f.

Empfangene ist nicht möglich, da Kommunikation in eine Richtung vom Kommunikator zum Empfänger durch Massenmedien stattfindet, die als autoritative Institutionen die Kommunikationskanäle und ihre Befüllung beherrschen.²⁶

Fluten und Kulturkritik als ständiger Begleiter

Das Kommunikations- und Publikationssystem hat – wie oben verdeutlicht wurde – im Laufe der Zeit in Bezug auf zwei Aspekte größere Wandlungsprozesse durchlaufen: Zum einen haben technologische Fortschritte sowie Fortschritte in den Organisationsleistungen der Produktion und Distribution tendenziell zu einer Verbilligung von und zur besseren Zugänglichkeit zu Schrift- und Lesemedien sowie zu Beschleunigungen des Kommunikationssystems geführt. Zum anderen kann über die Zeit eine fortschreitende De-Institutionalisierung des gesamten Gefüges des Kommunikations- und Publikationssystems beobachtet werden. Damit einher geht ein Verlust des Einflusses der bis dahin etablierten Vermittlungsautoritäten, wie Klöster, Universitäten, Adel, gebildete Großbürger, Kulturhüter und Berufsstände.²⁷ Kulturkritik durch Akteure, die als ‚gatekeeper‘ entschieden haben, wer was kommunizieren und rezipieren darf oder soll, ist zwangsläufig ein ständiger Begleiter,²⁸ ohne dass diese Kritiken im Laufe der Zeit substantielle neue Einsichten erzeugt hätten:

Der Gelehrte Petrarca ereifert sich im Mittelalter über ein Zuviel an Büchern, erzeugt durch laienhafte Schreiber, die nicht approbiert und deswegen zum anspruchsvollen Amt des Bücherschreibens nicht befugt sind: „Yederman will yetz buecher machen / yederman will die freyhayt haben / so treibt ein narr jmmer den andern“.²⁹

²⁶ Hierzu auch Lewis A. Coser, „Publishers as Gatekeepers of Ideas“, *The Annals of the American Academy of Political and Social Science* 421 (1975): 14-22.

²⁷ Hierzu z.B. Dieter Wellershoff, „Literatur, Markt, Kulturindustrie“, *Merkur: Deutsche Zeitschrift für europäisches Denken* 21.236 (1967): 1013-1026, hier 1014; Hannes Siegrist, „Geschichte des geistigen Eigentums und der Urheberrechte“, in: *Wissen und Eigentum: Geschichte, Recht und Ökonomie stoffloser Güter*, hg. Jeanette Hofmann (Bonn: Bundeszentrale für Politische Bildung, 2006), 64-80, hier 65f; Rüdiger Schnell, „Handschrift und Druck: Zur funktionalen Differenzierung im 15. und 16. Jahrhundert“, *IASL* 32.1 (2007): 66-111, hier 89.

²⁸ Ähnlich Thomas Christian Bächle, *Digitales Wissen, Daten und Überwachung zur Einführung* (Hamburg: Junius, 2016), 48.

²⁹ Differenziert dargelegt und zitiert bei Giesecke, *Der Buchdruck*, 173.

Der mechanische Druck hat erneut alle Schleusen geöffnet: Der Mainzer Domherr Bernhard von Breydenbach empört sich in der kulturkritischen Vorrede eines seiner Bücher 1486 über die neuen Kommunikationsmöglichkeiten für Jedermann: „[...] ja all me(n)schen vermessen sich zuo schreiben.“³⁰ Aufgrund der stark ansteigenden Buchproduktion im ersten Drittel des 19. Jahrhunderts stellte Heinrich Zschokke misslaunig die „Fluth“ schriftstellerischer Werke fest:

In jenen Zeiten, als man noch keine anderen Bücher hatte als solche, welche von einzelnen Händen geschrieben [...] werden mußten, wagten es nur vorzügliche Männer, ihre Gedanken aufzuzeichnen und durch die Schrift auszubreiten. In unsern Tagen, da durch einfache Druckwerkzeuge das schlechteste, wie das beste Werk mit wunderbarer Schnelligkeit vertausendfacht und in die Welt ausgestreut werden kann, [...] erhält und verbreitet sich das Schlechtere länger und mehr als ehemals, und nimmt an Zahl an gleicher Menge zu, wie es der mittelmäßigen Köpfe, der Halbgelehrten, der Leute mit unedeln Nebenabsichten überhaupt mehr gibt, als der ausgezeichneten [...] Geister, denen es um nichts als das Gute zu thun ist. Daher rührt die zahllose Fluth schriftstellerischer Werke, [...].³¹

Nach dem Zweiten Weltkrieg wurde dann die Erfindung des modernen Taschenbuchs damit kritisiert, dass es keine normativen Vorstellungen von Kanon mehr verkörpern würde, sondern aufgrund entgrenzter und beliebiger Produktion lediglich Bestandteil eines unüberschaubaren Warenhauses des Lesbaren wäre.³² Manfred Dworschak bezeichnete jüngst den größeren Teil des Outputs der heutigen Buchindustrie als Schüttgut, das genauso gut auch der Algorithmus der ‚Datenheinis‘ – gemeint ist *Amazon* – auf den Markt kippen könnte.³³

³⁰ Zitiert nach ebd., 174.

³¹ Heinrich Zschokke, *Stunden der Andacht zur Beförderung wahren Christenthums und häuslicher Gottesverehrung* (Aarau: Heinrich Remigius Sauerländer, 1821), hier 5. Band: *Die Lesesucht*, 130–139. Besprochen auch bei Eduard Kaeser, „Google oder Was Technik aus Menschen macht“, *Merkur: Deutsche Zeitschrift für europäisches Denken* 63.717 (2009): 168–172, hier 168f.

³² Dargelegt bei Wellershoff, *Literatur, Markt, Kulturindustrie*, hier 1017. Habermas unterscheidet die Erleichterung des ökonomischen Zugangs zu Kulturgütern durch Verbilligung der Produkte von der Erleichterung des psychologischen Zugangs durch Adaption des Inhalts an die Bedürfnisse von Verbrauchergruppen mit niedrigem Bildungsstand. Jürgen Habermas, *Strukturwandel der Öffentlichkeit: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft* (Frankfurt a.M.: Suhrkamp, 2015), hier 254.

³³ Vgl. Manfred Dworschak, „Bestseller aus der Retorte“, *Der Spiegel* 12 (2014): 108.

In Bezug auf elektronisch geprägte Welten beginnt der Diskurs um die Effekte von Informations- und Datenfluten mitnichten erst mit dem massiven Erfolg der so bezeichneten Daten- oder Überwachungskapitalisten *Google*,³⁴ *Amazon*, *Facebook* und *Apple* (*GAF*A). Anfang der 1970er Jahre wurde dem elektronischen Medium Fernsehen der „gründlichste und schnellste Wandel, den die Menschheit in ihrem Kommunikationswesen je erlebt hat“ zugesprochen.³⁵ Ben Bagdikian stellt fest, dass das Medium die gesellschaftliche Reaktionszeit massiv beschleunigt und dabei zudem die etablierten Systeme der Informationskontrolle umgeht. Das Fernsehen liefert – wie zuvor der Telegraph – den Autoritäten wie dem Volk die gleichen Nachrichten zur selben Zeit.³⁶ Der Autor verdeutlicht, wie fortschreitende Technologisierung, insbesondere Automatisierung und Elektronifizierung, die Menge des Kommunizierbaren pro Zeiteinheit massiv hat ansteigen lassen: So hat sich die Menge der gesetzten Druckzeilen von einer Zeile pro Minute zu Zeiten Gutenbergs auf 15 000 Zeilen gegen Ende der 1960er Jahre gesteigert, wobei exponentielle Produktivitätsschübe insbesondere durch die Elektronifizierung ab den 1960er Jahren erreicht wurden.³⁷ Neil Postman benutzt die Formulierung „Sturzflut an Informationen“, mit der er auch die Menge an Zeitungen (in den USA: 17 000), Zeitschriften (12 000), Radiogeräten (400 Mio.) und jährlich neu geschaffenen Fotografien (41 Mio.) beklagt.³⁸ Auch für die frühen Ausprägungen des Web (ab Mitte der 1990er Jahre) wurde einerseits festgestellt, dass sie das Ende der Programmdiktate der Massenmedien-*gatekeeper* bedeuten würden,³⁹ andererseits werden Zustände

³⁴ Vgl. Sascha Tamm, „Google: Ein Gespenst geht um in Europa“, *eigentlich frei*, 30. Juni 2014; Der Begriff Datenkapitalisten wurde erstmals von Peter Glotz im Zuge der ersten Internetökonomie-Welle der 1990er Jahre geprägt (*online*); vgl. Peter Glotz, „Digitaler Kapitalismus“, *Der Spiegel* 39 (1999): 82-86; vgl. auch Shoshana Zuboff, „The Secrets of Surveillance Capitalism“, *Frankfurter Allgemeine Zeitung*, 15. September 2014. (*online*).

³⁵ Anonymous, „Die Elektronen haben keine Moral“, *Der Spiegel* 17 (1972): 158-164.

³⁶ Vgl. Bagdikian, *The Information Machines*, 6-15 und 39-45.

³⁷ Vgl. ebd., 95.

³⁸ Vgl. Neil Postman, „Wir informieren uns zu Tode“, *Die Zeit*, 02. Oktober 1992. (*online*).

³⁹ Vgl. Josef Wehner, „Interaktive Medien: Ende der Massenkommunikation?“, *Zeitschrift für Soziologie* 26.2 (1997): 96-114, hier 99. Auch Bagdikian hat in seiner Abhandlung bereits die Vision formuliert, dass zukünftige neuartige Kommunikationssysteme und Endgeräte (er hatte bereits 1971 eine Art Tablet vor Augen) dem Rezipienten einen jederzeitigen und einfachen individuellen Zugriff auf Informationen ermöglichen würden und der Rezipient so die Selektionsentscheidung des *gatekeepers* nicht länger würde akzeptieren müssen; vgl. Bagdikian *The Information Machines*, 113.

fehlender Vernunft konstatiert (etwa bei Suchmaschinen).⁴⁰ Stanislaw Lem beschreibt die zu erwartende ‚Info-Sintflut‘ so:

Etwa drei Fünftel sind Unsinn und ‚vermischter Unsinn‘, den ich ‚Trübkunde‘ nenne; ein Fünftel ist zwar sinnvoll, aber vergängliche Info, und kaum ein Fünftel besteht aus ernstesten Denkfrüchten. Dieses vorhandene Gemisch wird das Internet noch verwässern, Wichtiges weiter mit Unwichtigem vermischen.⁴¹

Die über die Jahrhunderte vorgebrachten Kritiken lassen sich mithilfe des von Niklas Luhmann geprägten Begriffs des Überschusssinns erklären.⁴² Neue Kommunikationsmedien oder Verbreitungstechnologien erzeugen in Bezug auf verschiedene Aspekte neue Möglichkeiten: Sie attrahieren neue Kreise von Kommunikatoren und Adressaten und schließen gegebenenfalls andere Kreise aus, sie erfordern neue Formen der Kommunikation (zum Beispiel ‚Telegrammstil‘) oder neue Zeichensysteme (zum Beispiel *icons*), häufig ist auch zunächst unbekannt, wie das Neue sinnvoll zu nutzen sei. Das bisher Bekannte und Gelernte wird somit zunächst in seiner Ordnung bedroht, in Konsequenz bedarf es gesellschaftlicher Aushandlungs- sowie Aneignungsprozesse, um den Überschusssinn gegebenenfalls in einer neuen Ordnung aufzufangen.⁴³ Hartmut Rosa bezeichnet diesen Prozess als einen Kulturkampf, dem Hindernisse, Widerstände, Gegenbewegungen und nostalgische Sehnsüchte nach dem Verlorenen innewohnen.⁴⁴ Möglich ist die schlussendliche Ablehnung des Neuen, wenn sich eine neue Ordnung nicht verhandeln lässt.⁴⁵

⁴⁰ Vgl. Frank Hornig *et al.*, „Die Datensucht“, *Der Spiegel* 33 (2008): 80-92, hier 88.

⁴¹ Stanislaw Lem, „Zu Tode informiert“, *Der Spiegel* 11 (1996): 108-109, hier 109.

⁴² Vgl. Niklas Luhmann, *Die Gesellschaft der Gesellschaft* (Frankfurt a.M.: Suhrkamp, 1997), hier 405ff.

⁴³ Vgl. Dirk Baecker, „Ausgangspunkte einer Theorie der Digitalisierung“, *Catjects* (2015), (*online*). In Bezug auf die Facebook-Kommunikation differenziert diskutiert von Sascha Dickel, „Im Netz der Selbstreferenz: facebook-Kommunikation als Antwort auf die ‚Katastrophe‘ des Internet“, in: *Internet, Mobile Devices und die Transformation der Medien: Radikaler Wandel als schrittweise Rekonfiguration*, hg. Ulrich Dolata und Jan-Felix Schrape (Berlin: Ed. Sigma, 2013), 331-356.

⁴⁴ Vgl. Hartmut Rosa, *Beschleunigung: Die Veränderung der Zeitstrukturen in der Moderne* (Frankfurt: Suhrkamp, 2005), hier 80-82.

⁴⁵ Vgl. Martina Heßler, *Kulturgeschichte der Technik* (Frankfurt a.M.: Campus Verlag, 2012), hier 111.

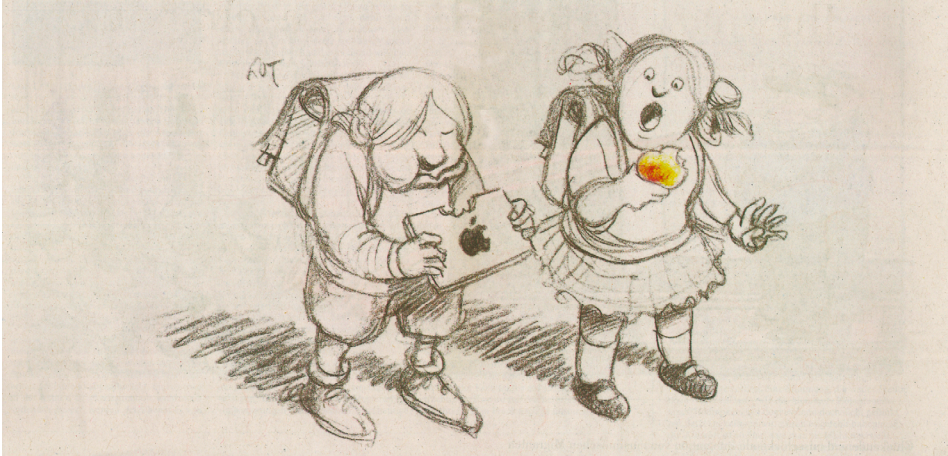


Abbildung 1: Neue Technologien sind schneller entwickelt, als sich die Strukturen der Gesellschaft ändern. Bild aus: *Neue Zürcher Zeitung* (28.05.2016).

Flut in Menge und Gewicht am Beispiel des Deutschen Buchmarkts

Das Buch ist ein Medium der schriftbasierten Kommunikation, das in der uns heute bekannten Form – also gedruckt und als Codex gestaltet –⁴⁶ als das älteste Massenmedium gilt.⁴⁷ Alleine in der aktuell vertrauten Form mit einer über 500jährigen Geschichte beladen, wurde und wird ihm zwar ständig sein baldiger Untergang aufgrund neu aufkommender Kommunikationsmöglichkeiten prophezeit („schon wieder tot“), gleichwohl erfreut sich das Totgesagte einer großen Lebendigkeit und zeigt sich zudem recht resilient: Die deutsche Buchwirtschaft erlöst mit dem Verkauf von Büchern jährlich knapp 10 Milliarden Euro Umsatz, eine Größe, die seit über zehn Jahren im Wesentlichen konstant ist, während zum Beispiel die

⁴⁶ Unter einem Codex versteht man eine Buchform, bei der mehrere Lagen Beschreibstoff, wie Papier oder Pergament, zwischen zwei Deckel gebunden oder verklebt werden; vgl. hierzu Ursula Rautenberg, „Codexform des Buchs“, in: *Reclams Sachlexikon des Buches*, hg. Ursula Rautenberg (Stuttgart: Reclam, 2015), 106-107. Der Codex wird geblättert. Eine alternative Buchform, die im Laufe der Zeit vom Codex abgelöst wurde, ist die Buchrolle (auch: Volumen), bei der der Beschreibstoff nicht geschichtet, sondern ein- und ausgerollt wurde; vgl. hierzu Arno Mentzel-Reuters, „Buchrolle“, in: *Reclams Sachlexikon des Buches*, 97. Die Rolle hat im Digitalen eine Renaissance erfahren: Sowohl Browserseiten wie andere elektronische Dokumente (z.B. PDF) werden *gescrollt*, also gerollt, und stellen somit eine Revolution im wahrsten Wortsinne dar: Zurückwälzen auf den Ausgangszustand.

⁴⁷ Vgl. Gerhard Maletzke, *Psychologie der Massenkommunikation*, hier 33-35.

Erlöse auf den Märkten für Zeitungen sowie für Zeitschriften im gleichen Zeitraum kontinuierlich gesunken sind.

Die in Deutschland lieferbaren Bücher sind im *Verzeichnis der lieferbaren Bücher (VLB)* gelistet. Es umfasst heute knapp 1,5 Millionen Titel, jährlich kommen zwischen 60 000 (im Jahr 2002) und 85 000 (2007) Novitäten, das heißt neue, noch nie dagewesene Buchtitel, hinzu. Der Schnitt über die letzten 14 Jahre in Deutschland liegt bei 77 000 Novitäten, pro Tag eines Jahres waren das durchschnittlich 210 Stück. Zum Vergleich: Postman beklagt für den US-amerikanischen Markt – der gemessen am Umsatz der weltweit größte Buchmarkt ist –⁴⁸ die Anzahl von 40 000 jährlichen Novitäten, dies aber bei circa 300 Millionen Einwohnern. Auf 100 000 Einwohner in Deutschland kommen statistisch 100 Novitäten, in den USA lediglich circa 13. Erschaffen werden die Buchtitel in Deutschland von circa 2 500 statistisch erfassten Verlagen, und verkauft werden sie von circa 4 000 Buchhandlungen. In Deutschlands erstem Buchkaufhaus, Hugendubel am Münchner Marienplatz, standen selbst zum Zeitpunkt der Auflösung der Filiale Anfang 2016 noch circa 100 000 Buchexemplare in den Regalen.⁴⁹ Bei den sogenannten Barsortimentern – das sind Zwischenhändler, die zwischen den Verlagen und dem Einzelhandel auf der Distributionsstufe stehen – sind 400 000 bis 500 000 der 1,5 Millionen lieferbaren Titel auf Abruf eingelagert;⁵⁰ sie sind in der Regel binnen 24 Stunden in einer Buchhandlung vor Ort, nachdem dort eine Bestellung ausgelöst wurde. Das Gewicht der täglichen Buchauslieferungen eines solchen Barsortimenters entspricht dem Gewicht von acht Elefanten.⁵¹ Der jährliche Absatz an Buchexemplaren in Deutschland beträgt zwischen 300 und 400 Millionen, geschätzt stehen in deutschen Haushalten knapp 9 Milliarden Buchexemplare.⁵²

⁴⁸ Deutschland steht je nach Statistik an zweiter oder dritter Stelle.

⁴⁹ Vgl. Katja Riedel und Florian Peljak, „Hugendubel: Die letzte Seite“, *Süddeutsche Zeitung*, 29. Januar 2016. (*online*).

⁵⁰ Vgl. Thomas Bez, hg., *ABC des Zwischenbuchhandels* (Norderstedt: Börsenverein des Deutschen Buchhandels, 2014), hier 11.

⁵¹ Vgl. *100 Jahre Umbreit*. Nicht mehr zugängliche Firmenbroschüre zum 100. Geburtstag des Buchzwischenhändlers Umbreit im Jahr 2012.

⁵² Vgl. Michael Allmaier, „Unsere Bücher sind nichts mehr wert: Warum trennen wir uns nicht einfach von ihnen?“, *Die Zeit*, 28. Januar 2016. (*online*).

Filtern als Kernaufgabe, gleichzeitig Ausblick

Um in der Menge an Informationen diejenigen zu finden, die dem Rezipienten Nutzen stiften und somit kein Rauschen darstellen, sind Selektionsmechanismen in Form von Filtern erforderlich. Die nachstehende Abbildung verdeutlicht die grundsätzliche Idee des Abgleichs des Bedarfs an Informationen mit dem Angebot:

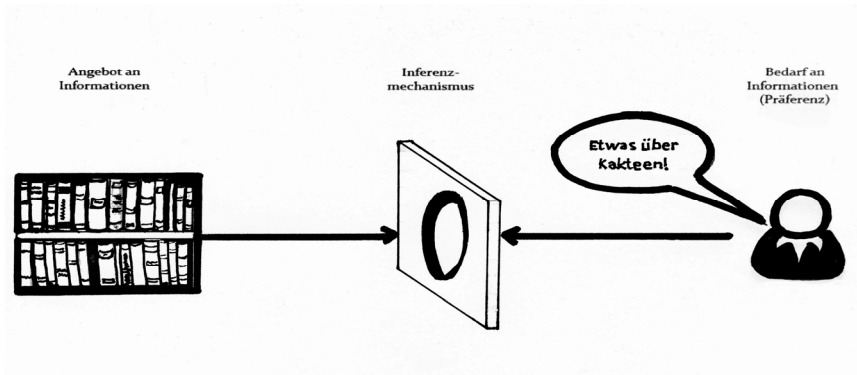


Abbildung 2: Filter zwischen Angebot an und Nachfrage nach Informationen.

Das Thema des Filterns erfährt jüngst im Zuge der Diskussion um die Konsequenzen der Digitalisierung ein hohes Maß an (emotionalisierter) Aufmerksamkeit, ausgelöst durch Eli Parisers Buch *Filter Bubble*.⁵³ Suchmaschinen, Soziale Medien und Empfehlungsmechanismen von elektronischen und ‚digitalen Marktplätzen‘ stehen im Fokus der Aufmerksamkeit der Behandlung datenbasierter Selektions- und Filterprozesse in der digitalen Welt.⁵⁴ Die Ablehnung solcher Mechanismen durch etablierte Vertreter der Medienbranche geht soweit, dass der Buchhändler *Thalia* mit ‚*ThaliaNext*‘ eine Applikation eingeführt hat, in der Buchhändler Buchtitel manuell als ähnlich zu anderen Buchtiteln klassifizieren und

⁵³ Eli Pariser, *Filter Bubble: Wie wir im Internet entmündigt werden* (München: Hanser, 2012).

⁵⁴ Zu ‚digitalen Foren‘ vgl. Svenja Hagenhoff, „Digitaler Marktplatz“ und „Elektronischer Marktplatz“, in: *Reclams Sachlexikon des Buches*, hg. Ursula Rautenberg (Stuttgart: Reclam, 2015), 142. Zum Prozess der Filterung vgl. Pascal Jürgens, Birgit Stark und Melanie Magin, „Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen“, in: *Die Googleisierung der Informationssuche: Suchmaschinen zwischen Nutzung und Regulierung*, hg. Birgit Stark, Dieter Dörr und Stefan Aufenanger (Berlin: de Gruyter, 2014), 98-135.

somit dem potenziellen Leser auf Basis menschlicher statt künstlicher Intelligenz das ‚nächste passende Buch‘ empfohlen wird.⁵⁵

In der Regel wird in den Diskursen sowohl des Feuilletons, der Branchenpresse und häufig auch in den wissenschaftlichen Beiträgen vernachlässigt, dass die Selektion von Inhalten und Informationen den Normalfall der mediengestützten öffentlichen Kommunikation darstellt und somit auch in vor-digitalen Zeiten ständig stattgefunden hat.⁵⁶ Gegebenenfalls werden aber die meisten der selektierenden Filter – wie Buchhändler, Rezensent, Redakteur, Journalist –⁵⁷ im Alltag der Medienutzung von den meisten Rezipienten und Diskutanten kaum noch als solche wahrgenommen. Nur so sind die in der aktuellen Algorithmen-debatte – teils verwundert, teils empört – hervorgebrachten Feststellungen erklärbar, datenbasierte automatisierte Selektionsmechanismen seien nicht neutral, sondern von soziokulturellen Vorannahmen geprägt.⁵⁸ Dieses muss so sein, solange davon auszugehen ist, dass Algorithmen Artefakte, also menschengemacht und somit nicht naturgegeben, sind.⁵⁹ Bagdikian argumentiert dann auch in seiner Abhandlung *The Information Machines* 1971, dass kein *gatekeeper* als selektierende Instanz wertfrei in einem sozialen Vakuum agiert. Vielmehr sind die Selektionsentscheidungen aller *gatekeeper* geprägt von ihren jeweiligen persönlichen sowie

⁵⁵ Vgl. Anonymous „Das Mensch-Maschine-Duell: Thalia baut an Buchempfehlungsplattform: Thalianext“, *Buchreport*, 11. September 2015. (online); Graf hingegen fordert von der Branche mit der filter- und algorithmenbasierten Literaturvermittlung im Digitalen produktiver umzugehen. Vgl. Guido Graf, „Ökosystem aus Filtern: Wie Literaturvermittlung im Digitalen funktioniert (not!)“, *Buchreport*, 15. Juli 2016. (online).

⁵⁶ Hierzu Christoph Kappes, „Menschen, Medien und Maschinen: Warum die Gefahren der ‚Filter Bubble‘ überschätzt werden“, *Merkur: Deutsche Zeitschrift für europäisches Denken* 66.754 (2012): 256-263.

⁵⁷ Siehe Beispiele ebd.

⁵⁸ Vgl. Thomas Christian Bächle, *Digitales Wissen, Daten und Überwachung zur Einführung*, hier 17, konkret zum Phänomen Algorithmus.

⁵⁹ Mowshowitz führt aus, dass die unterstellte Neutralität von Computern schlicht ein Mythos ist; vgl. Abbe Mowshowitz, „Computers and the Myth of Neutrality“, in: *Proceedings of the ACM 12th Annual Computer Science Conference on SIGCSE Symposium*, ed. Frank L. Friedman (New York: ACM, 1984), 85-92. Auch die Kernbotschaft des Computerwissenschaftlers und -kritikers Weizenbaum in seinem viel beachteten Standardwerk lautet, dass Computer wie alle Instrumente nicht wertfrei sind, sondern ihre Werte von der Gesellschaft erben, in die sie eingebettet sind; vgl. Joseph Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft* (Frankfurt: Suhrkamp, 1977).

professionellen Werten wie auch von allem Druck, der sie bei ihrem Handeln umgibt.⁶⁰

In der wissenschaftlichen Literatur existieren insbesondere in der Kommunikationswissenschaft verschiedene theoretische Ansätze, mit denen Selektionsprozesse sowohl auf Erzeugerseite (zum Beispiel Auswahl berichtenswerter Nachrichten durch Journalisten) wie auch auf Rezipientenseite erklärt werden,⁶¹ und auch die Wirkung einzelner, sehr spezifischer Instrumente, wie der Einfluss von Literaturkritik,⁶² sozialem Lernen oder Empfehlungsmechanismen auf Kaufentscheidungen sind verschiedentlich untersucht worden.⁶³ Um jedoch einen fruchtbaren und erkenntnisfördernden Diskurs über Filter, ihre Wirkweisen und damit auch ihre Bedeutung für die mediennutzende Gesellschaft führen zu können, müssten die vorhandenen, etablierten Filter ebenso wie die (vermeintlich) neuartigen Filter identifiziert, klassifiziert und in ihrer Wirkweise analysiert und verstanden werden. Diese Arbeit ist bisher von der Wissenschaft noch nicht geleistet worden. Das Desiderat sollte in transdisziplinärer Zusammenarbeit verschiedener Disziplinen, wie Kommunikationswissenschaft, Medienökonomie, Literaturwissenschaft sowie Informatik bedient werden.

⁶⁰ Vgl. Bagdikian, *The Information Machines: Their Impact on Men and the Media*, hier 106 und 265f.

⁶¹ Übersichten bei z.B. Thilo Hartmann, *Media Choice: A Theoretical and Empirical Overview* (New York: Routledge, 2009); oder Ines Engelmann, *Gatekeeping* (Baden-Baden: Nomos, 2016); Darstellung des Selektionshandelns im Zeitungsgeschäft bei Bagdikian, *The Information Machines*, Kap. 5. Für einen historiografischen Überblick Jürgen Wilke, *Nachrichtenauswahl und Medienrealität in vier Jahrhunderten: Eine Modellstudie zur Verbindung von historischer und empirischer Publizistikwissenschaft* (Berlin, New York: Walter de Gruyter, 1984).

⁶² Michel Clement *et al.*, „Do Critics Make Bestsellers? Opinion Leaders and the Success of Books“, *Journal of Media Economics* 20.2 (2007): 77-105.

⁶³ Zum Einfluss des ‚sozialen Lernens‘ auf Kaufentscheidungen vgl. Marc Keuschnigg, *Das Bestseller-Phänomen: Die Entstehung von Nachfragekonzentration im Buchmarkt* (Wiesbaden: Springer VS, 2012); zur Rolle von ‚Empfehlungen‘ vgl. Yiu-Kai Ng, „Recommending Books for Children Based on the Collaborative and Content-Based Filtering Approaches“, *Computational Science and its Applications - ICCSA 2016. 16th International Conference*, Beijing, China, July 4-7, hg. Osvaldo Gervasi *et al.* (Berlin: Springer, 2016), 302-317.

Literaturverzeichnis

- Anonymous. „Die Elektronen haben keine Moral“. *Der Spiegel* 17 (1972): 158-164.
- Bächle, Thomas Christian. *Digitales Wissen, Daten und Überwachung zur Einführung*. Hamburg: Junius, 2016.
- Bagdikian, Ben H. *The Information Machines: Their Impact on Men and the Media*. New York: Harper & Row, 1971.
- Behringer, Wolfgang. *Im Zeichen des Merkur: Reichspost und Kommunikationsrevolution in der Frühen Neuzeit*. Göttingen: Vandenhoeck & Ruprecht, 2003.
- . „Von der Gutenberg-Galaxis zur Taxis-Galaxis: Die Kommunikationsrevolution: Ein Konzept zum besseren Verständnis der Frühen Neuzeit“. *Historische Zeitschrift. Beihefte* 41 (2005): 39-54.
- Bez, Thomas. Hg. *ABC des Zwischenbuchhandels*. 7. neu bearb. Auflage. Norderstedt: Börsenverein des Deutschen Buchhandels, 2014.
- Clement, Michel, Dennis Proppe und Armin Rott. „Do Critics Make Best-sellers?‘ Opinion Leaders and the Success of Books.“ *Journal of Media Economics* 20.2 (2007): 77-105.
- Coser, Lewis A. „Publishers as Gatekeepers of Ideas“. *The Annals of the American Academy of Political and Social Science* 421 (1975): 14-22.
- Dickel, Sascha. „Im Netz der Selbstreferenz: facebook-Kommunikation als Antwort auf die ‚Katastrophe‘ des Internet“. In: *Internet, Mobile Devices und die Transformation der Medien: Radikaler Wandel als schrittweise Rekonfiguration*. Hg. Ulrich Dolata und Jan-Felix Schrape. Berlin: Ed. Sigma, 2013: 331-356.
- Dolata, Ulrich und Jan-Felix Schrape. „Zwischen Individuum und Organisation: Neue kollektive Akteure und Handlungskonstellationen im Internet“. *Stuttgarter Beiträge zur Organisations- und Innovationsforschung* Nr. 02. Stuttgart: Universität Stuttgart, 2013.
- Dworschak, Manfred. „Bestseller aus der Retorte“. *Der Spiegel* 12 (2014): 108.
- Engelmann, Ines. *Gatekeeping*. Baden-Baden: Nomos, 2016.
- Enzensberger, Hans Magnus. „Baukasten zu einer Theorie der Medien“. *Kursbuch* 20 (1970), 159-186.
- Friebel, Alexander. „Ohne Papier keine Zeitung: Die Erfindung des Holzschliffpapiers als Meilenstein in der Entwicklung des deutschen Preswesens“. *Jahrbuch für Kommunikationsgeschichte* 3 (2001): 132-156.

- Giesecke, Michael. *Von den Mythen der Buchkultur zu den Visionen der Informationsgesellschaft: Trendforschungen zur kulturellen Medienökologie*. Frankfurt a.M.: Suhrkamp, 2002.
- . *Der Buchdruck in der frühen Neuzeit: Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien*. 4. durchgesehene Auflage. Frankfurt a.M.: Suhrkamp, 2006.
- Glötz, Peter. „Digitaler Kapitalismus“. *Der Spiegel* 39 (1999): 82-86.
- Habermas, Jürgen. *Strukturwandel der Öffentlichkeit: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft*. 14. Auflage. Frankfurt a.M.: Suhrkamp, 2015.
- Hagenhoff, Svenja. „Digitaler Marktplatz“. In: *Reclams Sachlexikon des Buches*. Hg. Ursula Rautenberg. 3. verbesserte Auflage. Stuttgart: Reclam, 2015. 142.
- . „Distribution“. In: *Reclams Sachlexikon des Buches*. Hg. Ursula Rautenberg. 3., verbesserte Auflage. Stuttgart: Reclam, 2015. 124.
- . „Elektronischer Marktplatz“. In: *Reclams Sachlexikon des Buches*. Hg. Ursula Rautenberg. 3., verbesserte Auflage. Stuttgart: Reclam, 2015. 142.
- Hartmann, Tilo. *Media Choice: A Theoretical and Empirical Overview*. New York: Routledge, 2009.
- Helpman, Elhanan. Hg. *General Purpose Technologies and Economic Growth*. Cambridge, Mass.: MIT Press, 1998.
- Heßler, Martina. *Kulturgeschichte der Technik*. Frankfurt a.M.: Campus Verlag, 2012.
- Hornig, Frank, Martin U. Müller und Susanne Weingarten. „Die Daten-sucht“. *Der Spiegel* 33 (2008): 80-92.
- Jürgens, Pascal, Birgit Stark und Melanie Magin. „Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen“. In: *Die Googleisierung der Informationssuche: Suchmaschinen zwischen Nutzung und Regulierung*. Hg. Birgit Stark, Dieter Dörr und Stefan Aufenanger. Berlin: de Gruyter, 2014.
- Kaeser, Eduard. „Google oder Was Technik aus Menschen macht“. *Merkur: Deutsche Zeitschrift für europäisches Denken* 63.717 (2009): 168-172.
- Kappes, Christoph. „Menschen, Medien und Maschinen: Warum die Gefahren der ‚Filter Bubble‘ überschätzt werden“. *Merkur – Deutsche Zeitschrift für europäisches Denken* 66.754 (2012): 256-263.
- Keller, Rudi. *Zeichentheorie: Zu einer Theorie semiotischen Wissens*. Tübingen: Francke, 1995.

- Keuschnigg, Marc. *Das Bestseller-Phänomen: Die Entstehung von Nachfragekonzentration im Buchmarkt*. Wiesbaden: Springer VS, 2012.
- Kovarik, Bill. *Revolutions in Communication: Media History from Gutenberg to the Digital Age*. New York: Continuum, 2011.
- Lem, Stanislaw. „Zu Tode informiert“. *Der Spiegel* 11 (1996): 108-109.
- Luhmann, Niklas. *Die Gesellschaft der Gesellschaft*. Frankfurt a.M.: Suhrkamp, 1997.
- Lyons, Martyn. *Das Buch: Eine illustrierte Geschichte*. Hildesheim: Gerstenberg Verlag, 2012.
- Maletzke, Gerhard. *Psychologie der Massenkommunikation: Theorie und Systematik*. Hamburg: Hans Bredow-Institut, 1963.
- Mentzel-Reuters, Arno. „Buchrolle“. In: *Reclams Sachlexikon des Buches*. Hg. Ursula Rautenberg. 3. verbesserte Auflage. Stuttgart: Reclam, 2015. 97.
- Mills, Charles Wright. *The Power Elite*. 3. Auflage. New York: Oxford University Press, 1959.
- Mowshowitz, Abbe. „Computers and the Myth of Neutrality“. In: *Proceedings of the ACM 12th Annual Computer Science Conference on SIGCSE Symposium*. Hg. Frank L. Friedman. New York, NY: ACM, 1984. 85-92.
- Müller, Helen. „Verlagswesen und europäische Massenkommunikationsgesellschaft um 1900“. *IASL* 27.1 (2002): 170-197.
- Ng, Yiu-Kai. „Recommending Books for Children Based on the Collaborative and Content-Based Filtering Approaches“. In: *Computational Science and its Applications – ICCSA 2016*. 16th International Conference, Beijing, China, July 4-7. Hg. Osvaldo Gervasi et al. Berlin: Springer, 2016. 302-317.
- North, Michael. Hg. *Kommunikationsrevolutionen: Die neuen Medien des 16. und 19. Jahrhunderts*. Köln: Böhlau, 1995.
- Pariser, Eli. *Filter Bubble: Wie wir im Internet entmündigt werden*. München: Hanser, 2012.
- Rautenberg, Ursula. „Codexform des Buchs“. In: *Reclams Sachlexikon des Buches*. Hg. Ursula Rautenberg. 3. verbesserte Auflage. Stuttgart: Reclam, 2015. 106-107.
- Rosa, Hartmut. *Beschleunigung: Die Veränderung der Zeitstrukturen in der Moderne*. Frankfurt: Suhrkamp, 2005.
- Schneider, Ute. „Wozu lesen? Persistente Funktionen des Lesens im sozialen Kontext“. *IASL* 39.1 (2014): 268-283.
- Schnell, Rüdiger. „Handschrift und Druck: Zur funktionalen Differenzierung im 15. und 16. Jahrhundert“. In: *IASL* 32.1: 66-111.

- Siegrist, Hannes. „Geschichte des geistigen Eigentums und der Urheberrechte“. In: *Wissen und Eigentum: Geschichte, Recht und Ökonomie stoffloser Güter*. Hg. Jeanette Hofmann. Bonn: Bundeszentrale für Politische Bildung, 2006. 64-80.
- Spiekermann, Uwe. *Basis der Konsumgesellschaft: Entstehung und Entwicklung des modernen Kleinhandels in Deutschland 1850-1914*. München: Beck, 1999.
- Toffler, Alvin. *Future Shock*. New York: Random House, 1970.
- . *Der Zukunftsschock*. Bern: Scherz, 1971.
- Wehner, Josef. „Interaktive Medien: Ende der Massenkommunikation?“ *Zeitschrift für Soziologie* 26.2 (1997): 96-114.
- Weizenbaum, Joseph. *Die Macht der Computer und die Ohnmacht der Vernunft*. Frankfurt: Suhrkamp, 1977.
- Wellershoff, Dieter. „Literatur, Markt, Kulturindustrie“. *Merkur: Deutsche Zeitschrift für europäisches Denken* 21.236 (1967): 1013-1026.
- Wilke, Jürgen. *Nachrichtenauswahl und Medienrealität in vier Jahrhunderten: Eine Modellstudie zur Verbindung von historischer und empirischer Publizistikwissenschaft*. Berlin, New York: Walter de Gruyter, 1984.
- Zschokke, Heinrich. *Stunden der Andacht zur Beförderung wahren Christenthums und häuslicher Gottesverehrung*. 6. verbesserte Auflage. Aarau: Heinrich Remigius Sauerländer, 1821.

Internetquellen

- Allmaier, Michael. „Unsere Bücher sind nichts mehr wert: Warum trennen wir uns nicht einfach von ihnen?“. *Die Zeit* (28. Januar 2016). <http://www.zeit.de/2016/03/buecher-wert-entruempelung-antiquariate-buechermann/komplettansicht> [zuletzt besucht 27.12.2016].
- Anonymous. „Das Mensch-Maschine-Duell: Thalia baut an Buchempfehlungsplattform: Thalianext“. *Buchreport Online* (11. September 2015). <https://www.buchreport.de/2015/09/11/das-mensch-maschine-duell/> [zuletzt besucht 27.12.2016].
- Baecker, Dirk. „Ausgangspunkte einer Theorie der Digitalisierung“. *Catjects*. Ohne Datumsangabe. <https://catjects.wordpress.com/2015/06/10/ausgangspunkte-einer-theorie-der-digitalisierung/> [zuletzt besucht 27.12.2016].

- Graf, Guido. „Ökosystem aus Filtern: Wie Literaturvermittlung im Digitalen funktioniert (not!)“. *Buchreport* (15. Juli 2016).
<https://www.buchreport.de/2016/07/15/guido-graf-oekosystem-aus-filtern-wie-literaturvermittlung-im-digitalen-funktioniert-not/>
[zuletzt besucht 27.12.2016].
- Hagenhoff, Svenja. „Buch und Buchsachgruppen“. In: *Handbuch Medienökonomie*. Hg. Jan Krone und Tassilo Pellegrini.
http://link.springer.com/referenceworkentry/10.1007/978-3-658-09632-8_30-2 [zuletzt besucht 16.03.2017].
- Lyman, Peter and Hal R. Varian. „How Much Information?“ *The Journal of Electronic Publishing* 6 (2000): 2. <http://quod.lib.umich.edu/j/jep/3336451.0006.204?view=text;rgn=main> [zuletzt besucht 27.12.2016].
- Manovich, Lev. „Auf den Spuren der globalen digitalen Kulturen: Kulturanalytik für Anfänger“. *Dossier der Bundeszentrale für Politische Bildung*. 2011. <http://www.bpb.de/gesellschaft/medien/politik-des-suchens/75921/digitale-kulturen> [zuletzt besucht 27.12.2016].
- Postman, Neil. „Wir informieren uns zu Tode“. *Die Zeit* (02. Oktober 1992). <http://www.zeit.de/1992/41/wir-informieren-uns-zu-tode/komplettansicht> [zuletzt besucht 27.12.2016].
- Riedel, Katja und Florian Peljak. „Hugendubel: Die letzte Seite“. In: *Süddeutsche Zeitung* (29. Januar 2016). <http://www.sueddeutsche.de/muenchen/hugendubel-die-letzte-seite-1.2840540>
[zuletzt besucht 08.01.2017].
- Shirky, Clay. „How we will read“. Ohne Datumsangabe.
<http://genius.com/Clay-shirky-how-we-will-read-annotated>
[zuletzt besucht 27.12.2016].
- Tamm, Sascha. „Google: Ein Gespenst geht um in Europa“. *eigentlich frei*. 30. Juni 2014. <http://ef-magazin.de/2014/06/30/5473-google-ein-gespenst-geht-um-in-europa> [zuletzt besucht 27.12.2016].
- Zuboff, Shoshana. „The Secrets of Surveillance Capitalism“. *Frankfurter Allgemeine Zeitung* (15. September 2014). <http://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshan-zuboff-on-big-data-as-surveillance-capitalism-13152525.html> [zuletzt besucht 27.12.2016].

Erstmal einfach alles speichern: *Big Data* als Aufgabe für die Informatik

KLAUS MEYER-WEGENER

Abstract: *Big Data* ist derzeit ein Modebegriff in der Informatik, ein ‚Hype‘, wie man auch sagt. Eine präzise Definition gibt es nicht; einige Forscher betrachten den Begriff auch als unwissenschaftlich. Dass immer mehr Daten produziert werden, ist unstrittig, aber die Verwendung des Begriffs suggeriert auch eine sinnvolle Nutzung der Daten, und da herrscht eher ein wildes Durcheinander. Man vermutet, dass viele interessante Informationen in den Daten stecken, aber man weiß nicht so recht, wie man an sie herankommen könnte. Der Beitrag wird zunächst eine Charakterisierung von *Big Data* versuchen und dann diskutieren, was vorhandene technische Konzepte wie Datenbanken, *Data Warehouses*, *Complex-Event-Processing*, *Data Mining* und *Map-Reduce* für die Auswertung der Daten leisten können – und was auch nicht.

Einleitung

Es kommt nicht so oft vor, dass ein Informatik-Thema die Tagespresse erreicht. Und wenn, dann geht es meist um Einbrüche von Hackern in Datenzentren. Mit *Big Data* hat sich das geändert. So titelte beispielsweise die *New York Times* in ihrem „Sunday Review“: „The Age of Big Data“.¹ Wahre Wunderdinge werden nun erwartet; auch Industrie 4.0 kann als Ableger oder Teilgebiet davon angesehen werden.

Wir werden mit der Größe beginnen, doch das allein ist es nicht. Man kann heute sehr viel speichern, aber dadurch haben die Daten erst einmal noch keinen Nutzen. Es geht auch nicht um die sogenannten operationalen Daten, also solche, die man benötigt, um administrative Vorgänge abzuwickeln. Diese werden schon ganz selbstverständlich und zuverlässig genutzt, und sie sind auch gar nicht so groß. Es geht vielmehr um Protokolldaten, Spuren im Netz (sogenannte *Weblogs*), Aufzeichnungen aller Art durch Sensoren und ähnliches. Damit erzeugen wir immer mehr Daten, und wir haben inzwischen auch die Speicherkapazität dafür.

¹ Steve Lohr, „The Age of Big Data“, *The New York Times* (11 February 2012). (online). (Für diese und alle folgenden Angaben von Internetquellen in den Fußnoten wird auf die vollständige Nennung im Literaturverzeichnis verwiesen.)

Der englischsprachige *Wikipedia*-Artikel über *Big Data* sagt dazu:

The world's technological per-capita capacity to store information has roughly doubled every 40 months since the 1980s; as of 2012, every day 2.5 exabytes (2.5×10^{18}) of data is generated.²

Auf Deutsch sind das 2,5 Trillionen Bytes pro Tag.

Was steckt in diesen Daten? Wir werden uns kurz mit der Größe beschäftigen, dann aber auf die Nutzung übergehen. Nach einem Blick auf herkömmliche Systeme zur Speicherung und Auswertung wenden wir uns dem *Data Mining* zu, das als Oberbegriff für eine Sammlung von Verfahren verwendet wird, die aus großen Datenmengen interessante Muster herausfinden können. Dabei muss oft Parallelverarbeitung eingesetzt werden, um in akzeptabler Zeit zu Resultaten zu kommen. Es bleibt das grundlegende Problem der Datenqualität und Datenintegration anzusprechen, die die Voraussetzung für eine sinnvolle Verwendung sind und die leider außerhalb der Wissenschaft noch zu wenig Aufmerksamkeit erhalten. In den meisten Fällen stammen die auszuwertenden Daten aus verschiedenen Quellen. Bevor man sie auswerten kann, muss man sie zusammenführen. Dazu müssen sie vereinheitlicht werden, was sich als verblüffend schwierig erweist und ganz schlicht voraussetzt, dass man die Bedeutung der Daten genau kennt. Wer schon jemals fremde Daten erhalten hat mit der Aufforderung, sie weiterzuverarbeiten, also zum Beispiel auszuwerten, der hat schnell bemerkt, wie schwierig das sein kann. Eng verwandt damit ist das Problem der Datenqualität: Was kann an fehlerhaften Einträgen so alles vorkommen, und wie ist damit umzugehen? All das sind auch bei kleineren Datenbeständen schon beachtliche Herausforderungen; mit der Vergrößerung nimmt das Problem noch deutlich zu.

Wie groß ist „Big“?

Beginnen wir mit den Größenordnungen. Das ist vermutlich auch das erste, an das man bei *Big Data* denkt. Im Folgenden sind einige typische Größenordnungen aufgeführt, die mit den in der Informatik üblichen Bezeichnungen versehen sind. Von einer zur nächsten ist es immer tausendmal mehr:

1 Megabyte = 1.000.000 Bytes = 10^6 Bytes = eine Million.

² „Big Data“, *Wikipedia*. (online).

Das ist heute die normale Größe von einzelnen Dateien. Dieser Text und die zugehörige Präsentation liegen bei weniger als 1 MB, weil nicht so viele Bilder enthalten sind.

1 Gigabyte = 1.000.000.000 Bytes = 10^9 Bytes = eine Milliarde.

Das entspricht sehr großen einzelnen Dateien, also zum Beispiel digitalen Videos.

1 Terabyte = 1.000.000.000.000 Bytes = 10^{12} Bytes = eine Billion.

Das ist die typische Größe einer heutigen Festplatte. Da kann man schon ein paar Tausend Dateien abspeichern.

1 Petabyte = 1.000.000.000.000.000 Bytes = 10^{15} Bytes = eine Billiarde.

Hier sind wir bei großen wissenschaftlichen Datenbanken (*CERN*) und auch manchen *Web-Shops* wie *Amazon* angelangt.

1 Exabyte = 1.000.000.000.000.000.000 Bytes = 10^{18} Bytes = eine Trillion.

Das ist nun so etwas wie alles zusammen, wobei die Tendenz, wie oben schon erwähnt, ja längst zu immer größerem Wachstum führt. Deshalb hat man auch die nächsten beiden Größenordnungen schon benannt:

1 Zettabyte = 10^{21} Bytes = eine Trilliarde.

1 Yottabyte = 10^{24} Bytes = eine Quadrillion.

Diese braucht man im Moment noch nicht, aber wir sind auf dem Weg dahin.

Die Dimensionen großer Datenbestände

Damit ist schon das erste ,*V*' der fünf ,*V*'s' benannt, mit denen man *Big Data* zu definieren versucht, nämlich das Volumen (,volume'). Diese ,*V*'s' sollen die besonderen Herausforderungen zeigen, mit denen man es zu tun hat. Dabei ist die erste Herausforderung vielleicht die am wenigsten spektakuläre, denn speichern kann man die Daten mit den heutigen technischen Mitteln durchaus.

Das zweite ,*V*' aber steht für Geschwindigkeit (,velocity'), und damit meint man, dass ständig neue Daten dazukommen. Oft sind es zu viele,

um sie direkt angemessen verarbeiten und analysieren zu können. Dann greift die Devise, die in den Titel dieses Beitrags eingeflossen ist: „Erstmal speichern...“. Dabei stehen längst Techniken zur Verfügung, mit denen man das intelligenter angehen kann: Datenstromverarbeitung und *Complex-Event-Processing*. Auf sie wird später noch ausführlicher eingegangen. Hier soll nur erwähnt werden, dass man Daten damit direkt und schneller verarbeiten kann, was eine vollständige Speicherung eventuell sogar überflüssig macht. Stattdessen speichert man nur noch das, was man als relevant einordnet. Bei Bedarf kann man auch schneller auf die eingehenden Meldungen reagieren, sozusagen noch vor der Speicherung.

Das dritte ‚V‘ steht für Vielfalt (*variety*), und das stellt nach Ansicht einiger Experten die größte Herausforderung dar. Man kann diese Vielfalt grob in zwei Kategorien aufteilen, einmal die Multimedialität der Daten, die neben Zahlen und kurzen Zeichenketten auch Texte, Fotos, Grafiken, Tonaufnahmen, Videos und viele Arten von Messwerten und Aufzeichnungen aufweisen kann. Gerade Texte entstehen heute in sehr großer Zahl, man denke nur an die sozialen Netze. Leider ist es für Computer immer noch eine Herausforderung, Information aus Texten zu extrahieren – was dem Menschen so leichtfällt.

Zum zweiten sind die Daten aber auch inhaltlich sehr verschieden, selbst wenn sie sich auf Ähnliches beziehen. Was ist eigentlich der Preis eines bestimmten Produkts? In welcher Währung ist er angegeben? Ist die Mehrwertsteuer schon enthalten oder noch nicht? Bevor man die Preise verschiedener Hersteller zum selben Produkt miteinander vergleicht, sollte man das besser geklärt haben. Diese Herausforderung muss einerseits sehr ernst genommen werden, erweist sich aber andererseits als ausgesprochen schwierig und anspruchsvoll, was leider auch bedeutet, dass sie viel Zeit kostet. Wir werden ihr später noch einige Aufmerksamkeit zuwenden (müssen).

Das vierte ‚V‘ wird als Validität (*veracity*‘ oder auch als *validity*‘) bezeichnet und meint den Aspekt der Aufrichtigkeit, Wahrhaftigkeit, Glaubwürdigkeit oder Gültigkeit von Daten. Wo kommen die Daten eigentlich her? Besonders wenn man sie im Web gefunden hat, muss man diese Frage stellen und beantworten. Aber selbst bei Daten aus der eigenen Organisation ist eine gewisse Skepsis angebracht. Wer hat die Daten eingegeben und in welcher Situation? Es gibt haufenweise Geschichten von Pflichteingabefeldern, bei denen Benutzer nicht wissen, was sie dort eingeben sollen. Aber sie müssen mit ihrer Arbeit weitermachen, und sehr bald lernen sie, dass das möglich ist, wenn sie da einfach einmal „9999“

eingeben. So etwas später zu entdecken und dann auch noch zu bereinigen, kann recht teuer werden.

Das fünfte und letzte ‚V‘ steht für Wert (*value*): Die Daten sollten auch einen Nutzen haben. Das ist erst später dazugekommen, als man erkannte, dass Speichern allein eben doch nicht genügt. Oft wird dieser Nutzen erst einmal nur vermutet. Er muss dann ja auch durch geeignete Analysemethoden erst noch realisiert werden.

Möglichkeiten der Auswertung

Die entscheidende Frage ist also nun, was man mit diesen großen Datenbeständen anfangen kann, die man gespeichert hat. Man vermutet, dass darin wichtige Informationen verborgen sind, die man gern herausholen würde. Man muss also die Datenbestände auswerten, analysieren, durchforsten.

Bevor man hier der Fantasie freien Lauf lässt, sollte man sich erst einmal anschauen, was existierende Systeme denn schon an Auswertungsmöglichkeiten bieten. Das mag dann noch nicht ausreichen, und es mag auch auf derart großen Datenbeständen viel zu langwierig sein, aber es gibt doch eine klarere Vorstellung von Machbarkeit, die man nicht ignorieren darf.

Die einfachste Form der Abspeicherung überhaupt, nämlich Dateien, kann man schnell *ad acta* legen, denn sie bietet ohne zusätzliche Software überhaupt keine Auswertungsmöglichkeiten. Die bekommt man erst mit *Datenbanken*, zu denen so ein (ziemlich umfangreiches) Software-Paket gehört, das sogenannte Datenbank-Management-System. Das kauft man von großen Firmen wie *Oracle*, *IBM* oder *Microsoft*, um nur einige zu nennen, und installiert es auf seinen Rechnern. Dann übergibt man diesem System seine Daten (den wichtigeren Teil davon) und erhält deutliche Unterstützung bei der Datenverwaltung. Neben der Sicherung gegen Fehler und Ausfälle sowie der Organisation des gleichzeitigen Zugriffs auf die Daten bekommt man schon recht ordentliche Auswertungsmöglichkeiten. Man kann effizient suchen, Daten filtern und kombinieren und schließlich auch verschiedene Daten miteinander verknüpfen; die Fachleute nennen das einen Verbund oder auf Englisch *join*. Dafür müssen die Daten in Tabellen strukturiert sein, aber das sind viele Daten sowieso schon. Auch einfache statistische Auswertungen sind leicht zu formulieren und effizient auszuführen. Für die operationalen Daten hat sich das seit mehr als drei Jahrzehnten bewährt; die Systeme sind inzwischen sehr robust und effizient geworden. Anwendungsgebiete sind typischerweise

Buchhaltung, Auftragsverwaltung, Rechnungstellung, Lagerverwaltung und Produktionssteuerung. Da braucht man gar nicht so viele Daten, aber es sind die wichtigeren. Bei sehr großen Datenmengen stoßen diese Datenbanksysteme heute an ihre Grenzen, weil sie dafür nicht entwickelt worden sind. Was für die Benutzung als Komfort empfunden wird, kostet in der Realisierung eben auch Ressourcen, und da kommt bei sehr großen Datenbeständen mehr Ressourcenbedarf zusammen, als selbst von den größten Rechnern geboten werden kann. Deshalb gibt es inzwischen eine Alternative, die viel weniger Komfort bietet, dafür aber mit sehr großen Datenbeständen fertig werden kann, die *NoSQL*-Systeme. Auf sie werden wir später noch einmal zurückkommen.

Was nun die Auswertungen angeht, so müssen als nächstes Software-Paket die *Data Warehouses* genannt werden. Sie sind aus den Datenbanksystemen heraus entstanden, weil die mit der gleichzeitigen Abwicklung von laufenden Buchungen und umfangreichen Auswertungen oft überlastet waren. Sie konnten das zwar sauber ausführen, aber es dauerte einfach zu lange. Deshalb hat man die *Data Warehouses* daneben gestellt. Sie bekamen Kopien der Daten von den Datenbanken, auf denen dann die Auswertungen laufen konnten, ohne das Tagesgeschäft der Datenbanken zu stören. Man brauchte mehr Speicherplatz, das ist klar, aber der war zu diesem Zeitpunkt schon recht günstig zu haben. Man konnte noch etwas tun: Man konnte Daten aus mehreren verschiedenen Datenbanksystemen, die inzwischen längst nebeneinander entstanden waren, zusammenfassen und gemeinsam auswerten. Ziel waren umfassende Berichte (*reports*) mit vielen statistischen Kennzahlen und auch Grafiken. Man nennt das *Online Analytical Processing* oder kurz *OLAP*. Damit kann man Umsatzentwicklungen oder auch das Kaufverhalten der Kunden an verschiedenen Orten verfolgen. Zwei zentrale Auswertungsoperationen heißen *Roll-up* und *Drill-down*. *Roll-up* verdichtet Detaildaten zu kompakteren Übersichten. Wenn man zu viele einzelne Zahlen im Bericht hat, kann man sie zusammenfassen, also zum Beispiel aus monatlichen Angaben dann solche pro Quartal oder Jahr generieren lassen. *Drill-down* macht es genau umgekehrt: Wenn eine Zahl merkwürdig erscheint, kann man sich anzeigen lassen, aus welchen Teilen sie zusammengesetzt ist. Aus der Summe der Verkäufe im ganzen Bundesland werden dann eben die Verkäufe in einzelnen Bezirken oder Städten. Beide Operationen lassen sich mit einem *Data Warehouse* sehr einfach ausführen. Die Datenbestände sind hier schon deutlich größer als in den typischen Datenbanksystemen, aber sie sind andererseits auch schon ein wenig aggregiert, also zusammengefasst. Damit geht Information

verloren, und wenn man noch nicht genau weiß, wonach man eigentlich sucht, dann möchte man das vielleicht nicht. Im Prinzip eignet sich ein *Data Warehouse* aber auch für *Big Data*, wenn man seine Auswertungen schon kennt und wenn sie mit den Operationen dieser Software durchgeführt werden können. Dann ist die Aufgabe ‚nur‘ noch, die Daten in das *Data Warehouse* zu übernehmen.

Ein drittes Software-Paket geht noch einmal ganz anders an das Problem heran. *Complex-Event-Processing* und Datenstromsysteme zielen darauf ab, nicht einfach nur alles zu speichern, sondern neu ankommende Daten sofort auszuwerten und dabei zu entscheiden, ob sich das Speichern lohnt. Die sofortige Auswertung hat einen Geschwindigkeitsvorteil; man kann schnell reagieren, wenn eine kritische Situation erkannt wird. Das kann ein beginnender Börsen-Crash sein oder ein Tsunami oder auch eine Havarie in einem Produktionsprozess. Man kann entscheiden, ob sich das Speichern überhaupt lohnt; vor allem dann, wenn keine kritische Situation vorliegt. Man nennt das auch ‚*process and forget*‘ – und das ist ganz sicher auch eine Möglichkeit, mit sehr großen Datenmengen fertig zu werden. Der Gesetzgeber kann sogar vorschreiben, dass so etwas zu tun ist. Man denke nur an die Diskussion über die Vorratsdatenspeicherung.

Data Mining

Am interessantesten erscheint mir aber für die Auswertung von *Big Data*, was man heute als *Data Mining* bezeichnet. Das ist im Grunde auch nur Statistik, aber mit effizienten Algorithmen, sodass es eben auch auf großen Datenbeständen in vernünftiger Zeit ausführbar ist. Der Begriff *Data Mining* hat sich dafür etabliert; er wird aber inzwischen als etwas unglücklich angesehen. Schließlich will man ja nicht Daten herausholen, sondern vielmehr Information oder Wissen aus den Daten. Die Assoziation war damals in den neunziger Jahren wohl das Goldwaschen, bei dem man aus viel Schlamm einige wenige Nuggets gewinnen konnte. Daneben wird auch noch ‚*Knowledge Discovery in Databases*‘ (*KDD*) verwendet, aber auch dieser Begriff ist nicht besser und wirkt eher sperrig.

Es geht dabei um das Auffinden von bestimmten Mustern in großen Datenbeständen. Die Muster sind dabei: Gruppenbildungen, das gemeinsame Auftreten von Ereignissen oder Dingen, Klassifikationen und Ausreißer. Sie werden alle im Folgenden genauer betrachtet.

Cluster-Analyse

Die *Cluster-Analyse* beschäftigt sich mit der Frage, welche Gruppierungen es in den Daten geben könnte. Man möchte also eine Gruppenbildung (*cluster*) in den Daten erkennen. Das kann bei Kunden sehr interessant sein, aber auch bei Tierbeobachtungen und vielen anderen Aufzeichnungen.

Eine entscheidende und oft sogar vernachlässigte Voraussetzung dafür ist die präzise Definition von Ähnlichkeit zwischen den einzelnen Datensätzen. Das *Cluster-Analyse-Verfahren* ermittelt die Gruppen immer nur in Bezug auf diese Ähnlichkeitsdefinition. Ändert man die Bestimmung der Ähnlichkeit, ändern sich auch die *Cluster*, die das Verfahren findet. Es ist also ganz wichtig, sich zuerst intensiv Gedanken über die Ähnlichkeit zwischen den verschiedenen Datenobjekten zu machen.

Das erklärte Ziel aller *Cluster-Analyse-Verfahren* ist dann, die Gruppen (oder eben *cluster*) so zu finden, dass die Ähnlichkeit innerhalb der Gruppen möglichst hoch ist und die Ähnlichkeit zwischen den Gruppen möglichst gering.

Es gibt zahlreiche Verfahren, deren Vorstellung hier viel zu weit führen würde. Ein Klassiker ist das sogenannte *k-Means-Verfahren*. Man kann auf einen großen Fundus zurückgreifen, und es ist auch schon viel verbessert worden, sodass diese Verfahren auf großen Datenbeständen in akzeptabler Zeit ausführbar sind. Wie so oft bei diesem Thema muss man aber auch damit rechnen, dass sich ein neuer Datenbestand doch einmal als zu groß erweist. Das merkt man aber erst, wenn man es einmal versucht. Was dann noch getan werden kann, wird weiter unten dargestellt.

Assoziationsregeln

Die zweite Art der Analyse von Datenbeständen sucht nach etwas ganz anderem, nämlich dem gemeinsamen Auftreten von Ereignissen im allgemeinsten Sinne. Wenn das eine Ereignis eintritt, dann ist das andere oft auch noch dabei. Klassisches Anwendungsgebiet – und auch für jeden leicht vorstellbar – ist die Warenkorb-Analyse beim Einkauf, zum Beispiel im Supermarkt, aber auch im Internet. Wenn man zur Kasse geht und bezahlt, dann dokumentiert man damit seinen Einkauf; das ist ein Datensatz. Die Frage ist nun, welche Artikel oft zusammen eingekauft werden. Ein beliebtes Beispiel ist hier: Wer Windeln kauft, kauft auch noch Bier. So etwas zu wissen, ist für den Betreiber eines Supermarkts durchaus von

Interesse, weil er die Waren, die oft zusammen gekauft werden, dann im Regal nebeneinander platzieren kann – oder auch gerade nicht.

Wie findet man nun so ein häufiges gemeinsames Auftreten? Auch hier muss man erst selbst wieder etwas festlegen, so wie oben bei der *Cluster-Analyse* die Ähnlichkeit. In diesem Fall sind es zwei Schwellenwerte. Der erste ist die Unterstützung oder als Fachbegriff der ‚*support*‘. Das ist für die beiden zu prüfenden Ereignisse – im Beispiel also ‚Windeln‘ und ‚Bier‘ – die Frage, wie oft sie überhaupt zusammen im Datenbestand vorkommen, also wie oft jemand überhaupt Windeln und Bier kauft. Als Schwellenwert kann man dabei zum Beispiel 20% angeben. Wenn die Häufigkeit darunter liegt, wird dieses Paar als uninteressant eingestuft und nicht weiter betrachtet. Der zweite Schwellenwert ist die Konfidenz. Sie wird auch in Prozent angegeben und stellt fest, wie oft jemand, der Windeln in seinem Warenkorb hat, auch noch Bier kauft. Hier wird man vielleicht so etwas wie 70% sagen. Und nur dann, wenn sowohl der Schwellenwert für die Unterstützung als auch der für die Konfidenz überschritten sind, wird man konstatieren: Jawohl, wer Windeln kauft, kauft meistens auch noch Bier. Warum das so ist, hat man damit allerdings noch nicht herausgefunden ...

Das klingt erst einmal ganz einfach, aber man muss beachten, dass die Erreichung der Schwellenwerte für jedes mögliche Paar von Artikeln überprüft werden muss, um eine vollständige Liste dieses gemeinsamen Auftretens zu ermitteln. Das kann auf großen Datenbeständen dann schon wieder ziemlich lange dauern. Man nennt diese Feststellungen gemeinsamen Auftretens übrigens ‚*Assoziationsregeln*‘. Das stand schon in der Überschrift, ist aber ohne Erläuterung wohl nicht zu verstehen.

Auch hier gibt es längst Verfahren, die die Überprüfung aller Paare so schnell wie möglich durchführen; da sind in den letzten Jahren viele Verbesserungen vorgeschlagen und erprobt worden. Trotzdem stellen auch hier die nun nochmals größeren Datenbestände eine neue Herausforderung dar.

Klassifikation

Die dritte Form der Analyse führt auf einen Klassifikationsmechanismus, der es erlaubt, neue Daten nach den Erfahrungen aus der Vergangenheit, die in dem großen Datenbestand stecken, zu klassifizieren. Dabei kann es zum Beispiel um die Kreditwürdigkeit eines neuen Kunden gehen oder auch um die medizinische Diagnose auf der Basis von Symptomen.

Hier teilt man die Daten wieder in Gruppen ein; insofern ist das der *Cluster-Analyse* vergleichbar. Der entscheidende Unterschied liegt aber darin, dass man diese Gruppen (nun ‚Klassen‘ genannt) vorher schon kennt. Sie haben auch schon ihre ganz spezifische Bedeutung.

Deshalb geht man auch anders vor. Man will aus den vorhandenen Daten einen ‚Klassifikator‘ ableiten, mit dem man dann anschließend neue Daten klassifizieren kann. Das geht nur mit Daten, die schon klassifiziert sind, den sogenannten ‚Trainingsdaten‘. Sie haben ihre Klasse durch die Erfahrungen der Vergangenheit erhalten oder durch die Zuordnung eines Experten. Und dann versucht man, den Zusammenhang zwischen den verschiedenen Merkmalen des Datenobjekts – den ‚Symptomen‘ – und der Klasse – der ‚Diagnose‘ – zu erfassen und in einem Klassifikator nutzbar zu machen.

Eine mögliche Form eines solchen Klassifikators ist der ‚*Entscheidungsbaum*‘, von dem in Abbildung 1 ein Beispiel zu sehen ist. Es geht um die Frage, ob ein neu eintreffender Kunde einen Computer kaufen wird (Klasse ‚yes‘) oder nicht (Klasse ‚no‘). Man kennt von seinen Kunden das Alter und die Kreditwürdigkeit und weiß auch, ob sie Studenten sind oder nicht. Meist kennt man noch sehr viel mehr Merkmale, aber für das Beispiel genügt das. Die Erfahrung der Vergangenheit, die in den vielen aufgezeichneten Käufen von Kunden steckt, liefert nun einen Zusammenhang zwischen den Merkmalen eines Kunden und seiner Kaufentscheidung, die der Entscheidungsbaum greifbar machen kann. Wie man ihn aus dem großen Datenbestand extrahiert, kann an dieser Stelle nicht gezeigt werden. Aber dafür gibt es wieder eine ganze Reihe von effizienten Verfahren:

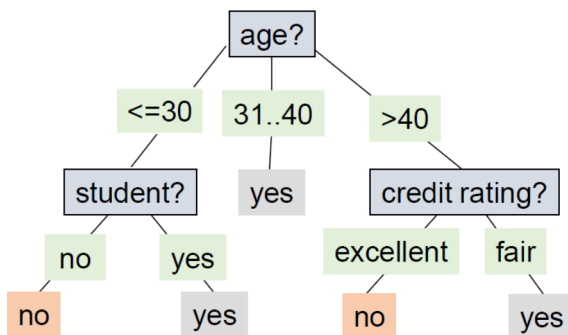


Abbildung 1: Ein Entscheidungsbaum.

Mit den Merkmalen eines neuen Kunden kann man nun von oben in den Entscheidungsbaum hineingehen. Zuerst wird das Alter überprüft. Liegt es zwischen 31 und 40, kann man sofort die Klasse feststellen: Ja, dieser Kunde wird wohl einen Computer kaufen.³ Andernfalls kann die Entscheidung noch nicht getroffen werden; es sind noch weitere Merkmale zu prüfen. Ist der Kunde nicht mehr als 30 Jahre alt, muss geschaut werden, ob er Student ist. Wenn ja, wird er wohl den Computer kaufen, wenn nein, eher nicht.

Diese Entscheidungsbäume sind also ganz einfach zu benutzen. Es ist aber überhaupt nicht einfach, sie aus großen Datenbeständen abzuleiten, auch wenn es dafür längst Verfahren gibt. Wieder kann und muss man vieles einstellen, und wenn man diese Einstellungen ungünstig wählt, wird der erzeugte Entscheidungsbaum auch immer mal falsche Entscheidungen treffen – oder besser: vorschlagen.

Es gibt neben den Entscheidungsbäumen noch viele andere Klassifikatoren, die hier nicht gezeigt werden können; es wird auf diesem Gebiet auch noch intensiv geforscht, sodass mit weiteren Verbesserungen zu rechnen ist.

Ausreißer

Das letzte Verfahren, das ich hier vorstellen möchte und das auch mit zum Kern des *Data Minings* gerechnet wird, ist die Suche nach Ausreißern oder auf Englisch ‚*Outliers*‘. Das sind Datenobjekte, die aus dem Rahmen fallen. Es können einfach Fehler in den Daten sein, aber auch die besonders interessantesten Fälle. Man denke nur an Betrugsversuche und andere Formen der Computer-Kriminalität, oder auch an Neuigkeiten und wissenschaftliche Entdeckungen.

Tatsächlich kann man hier sogar die oben schon eingeführten Verfahren des *Data Minings* wiederverwenden. So lässt sich die Entdeckung von Ausreißern als eine spezielle Form der *Cluster*-Analyse auffassen, bei der man das Augenmerk auf die ungewöhnlich kleinen *Cluster* richtet. Diese bieten kaum eine Ähnlichkeit mit anderen Datenobjekten. Oder man richtet eine Klassifikation mit den beiden Klassen ‚Ausreißer‘ – ‚kein Ausreißer‘ ein. Beides ist in Wahrheit nicht so einfach, wie es hier klingt, und es gibt daneben auch spezielle Verfahren, die nur auf die Entdeckung

³ So etwas nutzen Internet-Händler wie *Amazon*, indem sie dann Computer mit auf der Web-Seite anbieten.

von Ausreißern hin entwickelt worden sind. Es würde aber zu weit führen, diese hier vorzustellen.

Grenzen

Die eben beschriebenen Verfahren des *Data Minings* sind heute erfreulicherweise in großen Software-Paketen verfügbar. Man kann sie also relativ komfortabel benutzen. Wenn man das dann aber für sehr große Datenbestände versucht, stellt man fest, dass die Auswertungen sehr lange dauern. Das können durchaus auch schon einmal Stunden oder Tage für eine Statistik sein. Wenn die Software-Pakete sich auf ein herkömmliches Datenbanksystem stützen, um die Daten einfach abrufen zu können, stellt man womöglich auch noch fest, dass diese Datenbanksysteme es gar nicht mehr schaffen, alle diese Daten überhaupt zu speichern.

Deshalb nutzt der Ansatz ‚Erstmal einfach nur speichern‘ auch neue, schlankere Systeme, die man als *NoSQL*-Systeme bezeichnet. ‚*SQL*‘, steht dabei stellvertretend für die Leistung herkömmlicher Datenbanksysteme, die hier gar nicht mehr angestrebt wird. Stattdessen legt man den Fokus auf die Nutzung sehr vieler Rechner, sodass man die Daten gleichzeitig und damit insgesamt deutlich schneller verarbeiten kann. Allerdings sind die Möglichkeiten der Verarbeitung sehr eingeschränkt; im Wesentlichen kann man nur auf einzelne Datensätze zugreifen. Will man mehr, muss man es selbst außerhalb des Speicherungssystems realisieren, also programmieren.

Ein Ansatz dafür, der einen hohen Bekanntheitsgrad erreicht hat, ist *Map-Reduce*. In der Abbildung 2 sieht man den Ablauf nur für zwei Rechner, aber in Wirklichkeit sind es Tausende, die jeweils für einen kleinen Teil der Daten zuständig sind und die Schritte von links nach rechts gleichzeitig durchlaufen. Zuerst wird die *Map*-Funktion auf die Daten des jeweiligen Rechners angewendet. Diese muss man selbst programmieren, und zwar so, dass sie in diese Umgebung passt. Das macht etwas Arbeit, aber man hat auch alle Freiheiten, seine eigenen Auswertungen zu realisieren. Die Ergebnisse der *Map*-Funktion müssen kategorisiert werden, und das System ordnet sie dann so um, dass Ergebnisse derselben Kategorie zusammengefasst werden. Jede Kategorie wird dann noch mit einer *Reduce*-Funktion ausgewertet, die man wieder selbst programmieren muss. Auch das erfolgt auf allen Rechnern gleichzeitig. Anschließend muss man die Ergebnisse nur noch von den vielen Rechnern einsammeln. Dieses Verfahren hat *Google* erfunden, weil man anders der großen Datenmengen nicht mehr Herr werden konnte:

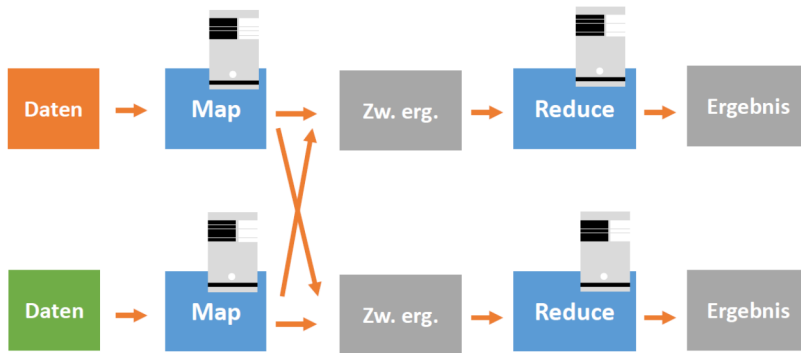


Abbildung 2: Map-Reduce.

Wie schon erwähnt, muss man zentrale Arbeiten dabei immer noch selbst erledigen, und manche der genannten und gewünschten Auswertungen passen auch nicht so recht in diesen Rahmen. Es besteht hier also durchaus noch Forschungsbedarf.

Eine wichtige Voraussetzung für alles: Qualität der Daten

Ein Problem muss zum Abschluss noch angesprochen werden, weil es eigentlich die Voraussetzung für alle Auswertungen ist: die Qualität der Daten. Bisher sind wir einfach davon ausgegangen, dass diese schon in Ordnung sein wird – diese Haltung ist auch in der Praxis der Datenanalyse durchaus noch anzutreffen.

Tatsächlich aber sind reale Daten fast immer fehlerhaft, unvollständig, irreführend und widersprüchlich, um nur ein paar der Probleme zu nennen. Ich vermute sehr stark, dass jeder schon einmal Erfahrung mit solchen Daten gesammelt hat.

Es gibt natürlich viele Vorschläge, wie man so etwas vermeiden könnte. Am besten ist es immer, die Daten gleich richtig zu erfassen und die Entstehung falscher Daten von vorneherein zu vermeiden. Das meint: Man braucht gute Benutzungsoberflächen. Jeder weiß, dass das oft nicht gelingt. Die Benutzer wissen einfach nicht, was sie eingeben sollen, und dann geben sie eben irgendetwas ein – nur damit es weitergeht. Später wundern sich andere Benutzer, warum in diesem Feld „9999“ eingetragen ist, was doch gar nicht sein kann. Außerdem können Daten auch dadurch, dass sie gerade *nicht* geändert – nämlich aktualisiert – werden, ihre Qualität verlieren.

Man muss sich also darauf einstellen, dass man nachträglich prüfen und verbessern muss. Das geschieht oft bedarfsgetrieben, also erst dann, wenn man es braucht, weil es sonst zu aufwendig wäre. Es geschieht mangels Alternative auch immer wieder von Hand, trotz der damit verbundenen Kosten. Bei sehr großen Datenmengen mit rein statistischen Auswertungen hat man zumindest noch die Option, als fehlerhaft erkannte Daten einfach wegzulassen.

Es ist aber festzuhalten, dass die Herstellung einer ausreichenden Datenqualität kein Luxus ist, sondern eine notwendige Voraussetzung, wenn man Auswertungsergebnisse erhalten möchte, mit denen man auch etwas anfangen kann. Durch *Big Data* verschärft sich dieses Problem eher noch.

Ein weiterer Aspekt muss hier auch noch angeführt werden: Es ist eher die Regel als die Ausnahme, dass die Daten, die man auswerten möchte, aus unterschiedlichen Datenquellen stammen, die alle unabhängig voneinander entstanden sind und betrieben werden, mit eigenen Zielen und Konzepten. Nun möchte man aber über alle hinweg auswerten, also zum Beispiel Verkaufszahlen eines Biergartens mit dem Wetter vergleichen oder Verkaufszahlen im Internet mit den Zeiten eines Werbespots im Fernsehen. Dergleichen Querbezüge gibt es zuhauf. Zusätzlich zu den eigenen Daten gibt es da auch noch riesige Datenmengen im Internet, die frei verfügbar sind; sie heißen dort etwa *Linked Open Data* oder *Google Web Tables*. Es würde zu weit führen, diese auch noch zu beschreiben.

Es geht also darum, verschiedene Datenbestände zusammenzubringen, und das heißt zuallererst einmal, sie zu vereinheitlichen. Das ist alles andere als trivial. Die Merkmale heißen unterschiedlich, auch wenn sie dasselbe meinen, und auch der Inhalt kann unterschiedlich dargestellt sein. Man kann das schon am noch relativ einfachen Beispiel von Datumsangaben demonstrieren:

„15.03.2016“ – „03/15/16“ – „2016-03-15“ – „Luise“.

Das sind die deutsche, die amerikanische und die international genormte Schreibweise für den 15. März. Und ja, auch „Luise“ ist ein Datum, wie Historiker sofort vehement bestätigen werden, nämlich ebenfalls der 15. März, wie er in historischen Dokumenten durchaus vorkommt. Das alles muss vereinheitlicht werden, bevor die erste Auswertung der Datenangaben durchgeführt werden kann.

Letztlich führt das alles zurück auf die Frage, was die Daten eigentlich bedeuten. Wenn man das nicht weiß oder mit einigem Aufwand herausfindet, wird man mit allen Auswertungen ziemlichen Schiffbruch erleiden.

Ein zweites Problem ist damit verwandt und ebenso dringlich: die Identifikation von Dubletten. Wenn man Daten aus unterschiedlichen Quellen zusammenführt und sich dabei auch die Bedeutung der Daten klargemacht hat, kann es immer noch vorkommen, dass zwei Einträge aus verschiedenen Quellen ein und denselben Gegenstand beschreiben, nur mit teilweise anderen Merkmalen. Erkennt man das, kann man die Beschreibung dieses Gegenstands vervollständigen und hat wahrscheinlich einen Nutzen davon. Erkennt man es aber nicht, zählt man in allen Auswertungen den Gegenstand doppelt. Dieses Problem wird in der wissenschaftlichen Literatur unter ganz verschiedenen Namen geführt, als Beispiele seien nur ‚Record Linkage‘, ‚Entity Resolution‘ und ‚Data Matching‘ angeführt. Es ist unbestritten, dass das Problem dringlich ist, und es gibt auch Vorschläge, wie man einer Lösung näherkommen kann. Leider sind die Möglichkeiten der Verschiedenheit aber so groß, dass eine allgemeine Lösung bisher nicht gefunden werden konnte. Man braucht immer noch Speziallösungen für seine eigenen Daten, und diese zu entwickeln, kostet Zeit und Geld. Hier besteht weiterer Forschungsbedarf.

Sowohl für die Vereinheitlichung als auch für die Dubletten-Erkennung muss man sehr viele *Regeln* definieren und im System nutzbar ablegen, damit sie auch für neu ankommende Daten angewendet werden können. Diese Regeln definieren zum Beispiel Umbenennungen, Umrechnungen und Umcodierungen, um nur die einfachsten Maßnahmen zu nennen. Sie definieren auch die Entscheidung, wann zwei Einträge als Dubletten zu betrachten sind. Es macht viel Arbeit, diese Regeln zu finden, und leider wird man dabei auch nie fertig, denn sie müssen laufend fortgeschrieben werden.

Fazit

Wie ich in diesem Beitrag vielleicht zeigen konnte, ist das Speichern der Daten noch der einfachste Teil; erst danach beginnt die eigentliche Arbeit.

Da gibt es bereits ein umfangreiches Spektrum an Auswertungsmöglichkeiten, und besonders das *Data Mining* erscheint für die Datenbestände, die man gern mit *Big Data* assoziiert, interessant. Leider sind es trotz aller Bemühungen zur Verbesserung der Verfahren oft einfach noch

zu viele Daten, sodass die Auswertungen zu lange dauern. Hier gibt es den Ansatz, eine gleichzeitige, parallele Verarbeitung auf vielen Rechnern vorzunehmen. Man ist gerade dabei, die Verfahren des *Data Minings* auf diese Umgebung zu übertragen. Es wird sicher nicht mehr so lange dauern, bis sie dadurch auch auf sehr großen Datenmengen in akzeptabler Zeit ausgeführt werden können.

Ein oft unterschätztes Problem ist damit aber auch noch nicht gelöst: das der Datenqualität. Die ist eigentlich Voraussetzung für eine sinnvolle Auswertung. Sie ist aber leider oft nicht in ausreichendem Maße vorhanden. Man sieht es einem Auswertungsergebnis nicht an, ob es auf der Basis hinreichend guter Daten erzeugt wurde. Die Datenqualität zu verbessern, ist sehr aufwendig, und die Aufgabe wird noch verschärft durch die Notwendigkeit einer Datenintegration, also der Zusammenführung von verschiedenartigen Daten aus mehreren Quellen. Hinweise und Anregungen für die Vorgehensweise sind inzwischen verfügbar, aber es gibt noch keine allgemeine Lösung für alle Daten, oder anders ausgedrückt: keinen Automatismus. Man braucht also Zeit und Geduld. Man muss seine Daten und die Daten anderer, die man benutzen möchte, sehr genau kennenlernen. Das mag mühsam sein, langfristig zahlt es sich aber aus.

Wir an der Universität sitzen ja schon lange nicht mehr im Elfenbeinturm. Wir sprechen oft und gern mit der Wirtschaft. Dabei hören wir in der letzten Zeit immer wieder so einen Satz wie: ‚Wir machen jetzt auch *Big Data*!‘ Ich hoffe sehr, dass dieser Beitrag darlegen konnte, dass *Big Data* an sich keine Lösung für irgendein Problem darstellt, sondern eine Aufgabe, die man methodisch angehen kann und sollte.

Literaturverzeichnis

- Christen, Peter. *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. New York: Springer Publishing Company, Inc., 2012.
- Doan, AnHai, Alon Y. Halevy and Zachary G. Ives. *Principles of Data Integration*. Amsterdam: Elsevier, 2012.
- Han, Jiawei, Micheline Kamber and Jian Pei. *Data Mining: Concepts and Techniques*. 3rd ed. Amsterdam: Elsevier, 2012.
- Heath, Tom and Christian Bizer. *Linked Data: Evolving the Web into a Global Data Space*. Synthesis Lectures on the Semantic Web, Morgan & Claypool Publishers, 2011.
- Luckham, David C. *The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems*. Boston: Addison-Wesley Longman, 2001.
- Rahm, Erhard, Gunter Saake und Kai-Uwe Sattler. *Verteiltes und Paralleles Datenmanagement: Von verteilten Datenbanken zu Big Data und Cloud*. Berlin: Springer Vieweg, 2015.
- Redman, Thomas C. *Data Quality: The Field Guide*. Boston et al.: Digital Press, 2001.
- Stonebraker, Michael, Daniel Bruckner, Ihab F. Ilyas, George Beskales, Mitch Cherniack, Stanley B. Zdonik, Alexander Pagan and Shan Xu. „Data Curation at Scale: The Data Tamer System“. *Proceeding 6th Biennial Conference on Innovative Data Systems Research*, Asilomar, CA, USA, January 6-9, 2013, Online Proc., www.cidrdb.org.

Internetquellen

- „Big Data“. *Wikipedia*. https://en.wikipedia.org/wiki/Big_data [zuletzt besucht 04.10.2016].
- Lohr, Steve. „The Age of Big Data“. *The New York Times* (11 February 2012). <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html> [zuletzt besucht 04.10.2016].

Der vermessene Mensch: Wissen wir, was wir alles über uns preisgeben?

HANS-ULRICH PROKOSCH

Abstract: ‚*Quantified Self*‘ – die Selbstvermessung des Menschen – eine Bewegung, die immer mehr um sich greift. Die Miniaturisierung macht es möglich, dass Sensoren und Aktoren in Alltagsgeräte und Kleidung als sogenannte ‚*Wearables*‘ verknüpft mit Anwendungsprogrammen auf Smartphones (Neudeutsch als ‚*Apps*‘ bezeichnet) von immer mehr Menschen in ihr tägliches Leben integriert werden. Wir vermessen uns selbst, leben nach Zahlen, stehen im Wettstreit mit uns selbst oder mit Freunden und Verwandten. Meist beschränkt sich diese Datengenerierung heute noch auf gelaufene Schritte oder Kilometer, die Herzfrequenz oder Schlafzeiten sowie Schlafphasen. Die technologische Entwicklung auf diesem Gebiet ist aber rasant. In einzelnen Anwendungen findet man heute auch schon Sensoren in Schuhen, T-Shirts, Kontaktlinsen oder Brillen. Sie zeichnen genaue Gangprofile, EKGs, Blutzuckerwerte, Atemparameter oder unser Ernährungsverhalten auf. Implantierbare Sensoren stellen technologisch kein Problem mehr dar und werden sicherlich auch bald in verschiedenen Fitness- und auch medizinischen Szenarien zum Einsatz kommen. Doch denken die Selbstvermesser unter uns auch darüber nach, wo ihre gemessenen, ‚patientengenerierten‘ Daten abgespeichert werden? Sind wir uns darüber im Klaren, welche Wege diese Daten über das Internet gehen, wer sie alles mitlesen kann, und was damit wirklich gemacht wird? Fast nie dienen diese Daten nur dem einen Zweck, für den wir sie generieren – sei es zur Steigerung unserer Fitness oder auch in einem medizinischen Vorsorge- oder Monitoringszenario. Wie schützt der Hersteller unserer *Smartwatch* die damit generierten Daten? Auf welcher *Cloud*, in welchem Land, unter welchen Sicherheitsmaßnahmen und unter der Gültigkeit welcher Datenschutzgesetze werden diese Daten verwaltet? Welche ‚*Big-Data-Analysen*‘ macht wer mit welcher Absicht mit unseren persönlichen Daten?

In diesem Kapitel werden die Optionen und Möglichkeiten der neuen sensorbasierten *Wearables*, die Vielfalt ihrer Anwendungsmöglichkeiten, ihr Potenzial für die Nutzung in der medizinischen Versorgung, aber auch die damit assoziierten Risiken und Gefahren illustriert und beleuchtet. Das Kapitel soll weder ein negatives noch ein zu positives Bild der Selbstvermessung beschreiben. Es soll aber dazu anregen, einen kritischeren Umgang mit all diesen neuen Technologien zu pflegen und deren Einsatzzwecke und Gefahren etwas bewusster zu hinterfragen.

Prolog

Es ist Dienstag, 1. März 2016, 15 Uhr – Zeit, die Fahrt zu meinem Vortrag nach Amberg im Rahmen der Erlanger Universitätstage zu planen. Eigentlich wollte ich mit dem Zug fahren. Meine digitale Kalender-App erinnert mich, weist mich darauf hin, dass eine Bahnfahrt problematisch werden könnte. Denn über ihre Verbindung zu ‚bahn.de‘ konnte sie feststellen, dass es Bauarbeiten auf der Strecke gibt und dass die Züge deshalb häufig größere Verspätungen haben. Als Alternative schlägt meine Kalender-App die Fahrt mit dem Pkw vor und hat auch schon automatisch mit *Google Maps* Kontakt aufgenommen und den Verkehr überprüft. Keine Probleme – die App hat keine Staus auf der Strecke gefunden und fragt mich, ob ich lieber mit dem Auto fahren möchte. Ich bestätige dies. Schon hat die App aber auch den Online-Wetterdienst befragt und informiert mich über die aktuellen Temperaturen. Es ist März – und kalt: nur mal gerade 5° Celsius. Mein Auto parkt vor der Tür – in der Kälte. Meine App schlägt vor, um 17 Uhr los zu fahren – ich bin einverstanden.

16.45 Uhr: Meine Kalender-App schlägt Alarm – es heißt zusammenpacken und die Fahrt vorzubereiten. Draußen ist es noch kälter geworden, nur noch 3° Celsius. Längst hat sich aber meine Kalender-App mit meinem Auto verbunden und dort die Sitz- und die Innenraumheizung eingeschaltet. Wenn ich gleich einsteige, wird es zum Glück schön warm sein. Mein Navigationsgerät im Auto muss ich gar nicht mehr programmieren, denn die digitale Kalender-App hat diesem natürlich schon mitgeteilt, wohin die Fahrt gehen soll.

18.05 Uhr: Ankunft in Amberg. Wir stehen vor dem Parkhaus. Ich hasse es, in einem fast vollen Parkhaus nach freien Parkplätzen zu suchen. Doch zum Glück fahre ich ja das neueste Forschungsfahrzeug von *BMW*. Dieses hat vier Laserscanner, jede Menge Computer-Intelligenz und natürlich ein hochpräzises *GPS*, das die Position des Fahrzeugs auch im Parkhaus exakt bestimmen kann. Ich steige also aus und aktiviere die Parkplatzsuchfunktion meines *BMW* mit meiner *Smartwatch*. Und wie von magischer Hand gesteuert, fährt er los und sucht sich seinen Parkplatz.¹

¹ Vgl. hierzu Axel Kannenberg, „BMW: Autonom Einparken per *Smartwatch*-Befehl“, *heise* (16.12.2014). (*online*). (Für diese und alle folgenden Angaben von Internetquellen in den Fußnoten wird auf die vollständige Nennung im Literaturverzeichnis verwiesen.) Im Februar 2016 hat sich die diesbezügliche Vision von *BMW* allerdings schon wieder etwas relativiert und das nun beschriebene Szenario verspricht nur noch das automatische Einparken per App für die letzten Meter in die Parklücke; vgl. Fabian

20.30 Uhr: Mein Vortrag ist vorüber, ich gehe zurück zum Parkhaus. Auf dem Weg aktiviere ich die Abholfunktion meines *BMW* über meine *Smartwatch*. Da der *BMW* seine eigenen und meine *GPS*-Koordinaten, aber auch meine Gehgewohnheiten und Gehgeschwindigkeit kennt, weiß er, wann er losfahren muss, um genau zur gleichen Zeit wie ich am Parkhausausgang zu sein.

20.40 Uhr: Ich bin eingestiegen und kann losfahren. Jetzt hätte ich Lust auf ein schönes Glas gekühlten Weißwein, wenn ich nach Hause komme. Und meine Wohnung soll natürlich angenehm warm sein. Also noch schnell die Kühlschranks-App auf dem Smartphone aufgerufen: Die digitale Kamera in meinem Kühlschrank zeigt mir, dass tatsächlich ein guter Pfälzer Riesling auf mich wartet; mit der Heizungs-App erhöhe ich die Heiztemperatur im Wohnzimmer und fahre dann los.

Das Internet der Dinge

Wir können uns heute fast alle ein Leben ohne das Internet nicht mehr vorstellen. Unser Zugang zum Internet aber hat sich rasant verändert. Vor 30 Jahren noch hatten nur einige Privilegierte einen solchen Zugang, konnten in ein Rechenzentrum gehen und dort an einem klobigen Bildschirmgerät und über eine zeilenbasierte Benutzeroberfläche E-Mails schreiben. Das *World-Wide Web* (*WWW*) entstand erst zwei Jahre später im Jahr 1989.² Tim Berners-Lee hatte damals die geniale Idee, verschiedene Informationsquellen, egal auf welchen Rechnersystemen dieser Welt diese auch liegen würden, mit sogenannten ‚*Links*‘ miteinander zu verknüpfen. Mit einem Klick konnte man von da an, aus einem Text oder auch von einem Bild, zu immer neuen Texten springen – quer durch die ganze Welt. Von einem Computer zum anderen. Computer wurden ‚vernetzt‘ und bekamen Adressen (die sogenannten IP-Adressen = Internet Protokoll Adressen). Man ‚surfte im Internet‘.

„The WorldWideWeb (*W3*) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents“,³ so hatte es Tim Berners-Lee damals beschrieben.

Mit dem Einzug der ‚*Personal Computer*‘ kamen dann irgendwann auch Privatpersonen in diesen Genuss. Die Informationsmenge wurde immer größer, zu Texten und Bildern kamen Musik und Videos hinzu.

Hohberg, „Meine Smartwatch parkt noch schnell den Wagen ein“, *PS-Welt* (04.02.2016). (online).

² Vgl. hierzu „Tim Berners-Lee’s Proposal“. (online).

³ Vgl. CERN, „World Wide Web“. (online).

Der Begriff des *World-Wide Web* geriet in Vergessenheit – man sprach einfach nur noch vom Internet. Die Computer wurden kleiner und immer mobiler. Irgendwann benötigte man dann auch gar kein Telefonkabel mehr, um ins Internet zu kommen – *WLAN* (*Wireless Local Area Network*) machte es möglich. Mit einem Notebook oder – noch schicker – mit einem Tablet kommt man heute über *WLAN* problemlos ins Internet. Und natürlich wurden auch unsere Mobiltelefone immer schlauer, wandelten sich zum Smartphone und boten damit auch Zugang zum Internet. Schon für Schulkinder ist es damit heute überhaupt kein Problem mehr, immer und überall ‚ins Internet zu kommen‘.

Mit dem Fortschreiten der Endgeräte, über die wir ins Internet gelangen, entwickelte sich aber auch unsere Art, das Internet zu nutzen, immer weiter. Während es Tim Berners-Lee anfangs nur darum ging, der Menschheit einen lesenden Zugriff auf die Vielzahl der in der Welt verfügbaren Dokumente zu ermöglichen, entwickelten andere in den nächsten Jahren immer interaktivere Internetdienste. Heute buchen wir im Internet Reisen, schauen online Kinofilme oder kaufen im Internet alle Weihnachtsgeschenke, ohne auch nur einen Fuß vor die Haustüre zu setzen.

Und dennoch – für die meisten Menschen gehört zur Vorstellung des Internets heute immer noch ein Bildschirm und auch eine (oft schon virtuelle) Tastatur, auch wenn diese immer kleiner werden und – in Form eines Smartphones – in die Handtasche passen. Im Hintergrund gibt es natürlich große Rechenzentren mit riesigen Servern, Datenspeichern und Softwareprogrammen. Noch sind es nur wenige, die ihren Kühlschrank, ihre Heizung, ihr Auto oder gar ihre Turnschuhe und ihre Haarbürste mit dem Internet in Verbindung bringen würden.

Die Miniaturisierung der Computerchips macht es aber möglich, ‚Rechner‘ mit einer *CPU* (*Central Processor Unit*) und eigenem Speicher so klein zu bauen, dass sie mittlerweile in jedes andere Gerät hinein passen und längst nicht mehr wie ein Computer aussehen. Sensoren, die aus ihrem Umfeld Signale aufnehmen und diese in digitale Daten umwandeln können, passen heute in einen Stecknadelkopf oder auch in eine Kontaktlinse. Und die heutige ‚*WLAN*-Karte‘ ist eine Miniatur-*SIM*-Karte, die kleiner ist als ein Fingernagel. Die Ära des ‚Internet der Dinge‘ (*IdD*, *IoT*, *Internet of Things*) hat längst begonnen. Die Vision der Techniker, aber auch der Politiker, ist es, dass jedes Ding in Zukunft einen Netzanschluss und eine *IP*-Adresse bekommen soll. Der Begriff ‚Internet der Dinge‘ steht dafür, dass der (*Personal*) Computer, so wie wir ihn heute noch kennen, in der zukünftigen digitalen Welt immer mehr mit

‚intelligenten Gegenständen‘ ergänzt und irgendwann von diesen abgelöst wird. Anstatt selbst Gegenstand der menschlichen Aufmerksamkeit zu sein, soll das ‚Internet der Dinge‘ die Menschen zukünftig bei ihren Tätigkeiten unmerklich unterstützen. Mag diese Vorstellung für viele von uns noch etwas befremdlich anmuten, so ist diese Entwicklung doch unabwendbar. ‚*BI Intelligence*‘, der Forschungsservice des Online Magazins *Business Insider* geht davon aus, dass es bereits im Jahr 2020 auf der Erde mehr als 24 Milliarden IoT-Geräte geben wird. Das sind ungefähr vier Geräte für jedes menschliche Wesen auf unserem Planeten.⁴

„Was eigentlich ist das Internet der Dinge?“ Diese Frage stellte auch Lena Schipper in der *Frankfurter Allgemeine Zeitung* anlässlich der Computermesse *Cebit* im März 2015. Und ihre Antwort darauf lautete „Im ‚Internet der Dinge‘ kommuniziert alles mit allem. Bloß ohne Menschen.“⁵

Auch wenn meine kleine Geschichte im Prolog dieses Kapitels noch nicht der Wahrheit entsprach, so wäre dies dennoch alles heute schon technisch realisierbar. Die Geschichte illustriert an einem einfachen Beispiel, wie stark das Internet der Dinge unser Leben in Zukunft verändern wird. Die gleichen Technologien sind es nun, die auch das ‚Vermessen des Menschen‘ immer mehr beschleunigen werden.

Das Vermessen nimmt früh seinen Anfang

Das Vermessen eines Menschen beginnt früh. Bereits bei der Geburt verkünden Eltern glücklich die Körpergröße, das Gewicht und vielleicht auch noch den Kopfumfang des Neugeborenen. Letzterer ist später nicht mehr so wichtig, doch Größe und Gewicht sind Parameter, die uns unser Leben lang begleiten. Sie lassen sich ja auch problemlos mit einfachen analogen Messinstrumenten bestimmen. Wenn wir älter werden, verändern sich diese beiden Parameter meist nicht mehr so stark, und wir messen diese dann auch nicht mehr so häufig. Bei Erwachsenen verlagert sich das Messen dieser (und weiterer) Parameter dann normalerweise mehr auf Aktivitäten von Ärzten und Pflegekräften und reduziert sich auf Krankheitszeiten. Chronisch Kranke müssen allerdings einige Parameter, wie zum Beispiel den Blutdruck (Hypertoniker) oder den Blutzucker (Diabetiker), kontinuierlich kontrollieren und messen diese auch zu Hause. In der Regel führen diese dann händisch ein Blutdruck- oder

⁴ Andrew Meola, „What is the Internet of Things (IoT)?“, *Business Insider Deutschland* (19.12.2016). (online).

⁵ Lena Schipper, „Was eigentlich ist das Internet der Dinge?“, *Frankfurter Allgemeine Zeitung* (17.03.2015). (online).

Blutzucker-Tagebuch und nehmen dies zu jedem Arzttermin zur Kontrolle mit. Das Internet der Dinge und die Digitalisierung haben in den letzten Jahren aber begonnen, auch in diese Bereiche ihren Einzug zu halten. Längst bleibt es bei den täglichen digitalen Messungen nicht mehr alleine beim Schrittzähler. Tragbare Sensoren – integriert in Armbänder oder Uhren – werden als sogenannte *Fitnessstracker* eingesetzt und belegen, dass heute auch gesunde Personen wieder beginnen, ihre Vitalwerte, wie Blutdruck und Herzfrequenz, regelmäßig zu messen und diese dann auch digital zu speichern und zu verwalten. Derartige Sensorik in Verbindung mit *Bluetooth*, Smartphones, einer Internetanbindung und ‚medizinischen Apps‘,⁶ die Daten auf dem Smartphone oder direkt in der ‚Cloud‘ ablegen, ermöglichen das digitale Verwalten und Visualisieren der früher papierbasierten Tagebücher. Ein einfaches Beispiel einer unmittelbaren medizinischen Anwendung ist das Smartphone-kompatible Blutzuckermessgerät. Dabei funktioniert die Koppelung mit dem Messgerät über den sogenannten Multiport des iPhones. Die Übertragung der gemessenen Blutzuckerwerte auf einen *Internet-/Cloud*-basierten Datenspeicher erfolgt automatisch. Dies entlastet die Patienten von der bisherigen aufwendigen händischen Dokumentation. Die *Companion App* der Firma *mySugr* ermöglicht Diabetikern, zusätzlich zu den gemessenen Blutzuckerwerten auch Daten zu Ernährung, Sport und Gesundheitszustand in einem digitalen Tagebuch zu speichern, zu verwalten und zu analysieren.⁷ Dieser technologische Fortschritt kommt nun auch einer Bewegung zu Gute, die 2007 in der San Francisco Bay Area gegründet wurde und seit 2012 auch in Deutschland vertreten ist. Es handelt sich um die *Quantified-Self-Community*:

eine Gemeinschaft von Anwendern und Anbietern von Self-Tracking Lösungen. Ziel dieser Gemeinschaft ist der Austausch von Wissen über die Nutzung persönlicher Daten. Dies umfasst die Mittel und Methoden zur Erfassung von Daten aus allen Lebensbereichen. Im Vordergrund stehen jedoch die persönlichen Erkenntnisse, welche aus den Daten abgeleitet werden können, sowie die Veränderungen welche sich mit ihnen nachvollziehen lassen.⁸

⁶ Zur Vertiefung und kritischen Aufarbeitung dieses Themas vgl. Urs-Vito Albrecht, „Chancen und Risiken von Gesundheits-Apps (*CHARISMHA*)“ (2016). (*online*).

⁷ Vgl. hierzu Caspar Grote, „Diabetes-App von mySugr: Smartphone überwacht Blutzucker“, *medizin-und-elektronik.de* (11.06.2014). (*online*).

⁸ Vgl. „QS: Quantified Self: Self Knowledge Through Numbers (Deutsche Community)“ (*online*) sowie auch „igrowdigital: Deutschland ist quantifiziert“. (*online*).

Das gemeinsame Motto dieser Gemeinschaft lautet ‚Self knowledge through numbers‘.⁹ Anhänger dieser Bewegung belassen es meist nicht nur beim Vermessen einiger weniger Parameter. Insbesondere Frauen sind zum Beispiel begeistert von Schlafphasenweckern, welche es versprechen, die Schlafphase des Nutzers zu erkennen und ihn somit zum idealen Zeitpunkt aufzuwecken. Hierzu muss man lediglich das Smartphone auf die Matratze legen, um über Mikrofon und Gyroskop (Kreiselstabilisator) die von den Bewegungen des Schläfers erzeugten Geräusche und die Lageänderungen des Smartphones zu erfassen. Andere Schlafsensoren bestehen aus einem dünnen Sensorstreifen, der mit Hilfe eines Klebebands auf der Matratze fixiert wird. Der Sensor übermittelt die Messwerte mittels einer winzigen Elektronikeinheit per *Bluetooth* an das Smartphone. Die zugehörige Smartphone-App wertet die Daten des Sensors aus, um aus diesen auf das Schlafverhalten zu schließen. Zusätzlich können mit Mikrofon und Helligkeitssensor des Smartphones die Helligkeit im Schlafzimmer und das nächtliche Schnarchen erkannt und aufgezeichnet werden. Nach dem Aufwachen kann man sich dann sofort grafisch die Schlafdauer sowie Kurven der Atem- und Herzfrequenz und einen ‚*Sleepscore*‘ ansehen.¹⁰ Nach dem Gang auf die digitale Waage, die natürlich nicht nur das Gewicht, sondern auch noch Körperfettanteil, Körperwasser, Muskelmasse, Knochenmasse, Viszeral-Fett und den *Body-Mass-Index* bestimmt, legt man sich einen Pulsmesser-Brustgurt um oder seinen *Fitnesstracker* ums Armgelenk und misst damit den ganzen Tag seine Herzfrequenz, überträgt diese drahtlos an seine Smartphone-Applikation, und diese legt die Daten dann auf einem Server im Internet ab. Andere Entwicklungen gehen da mittlerweile schon viel weiter. So hat das Fraunhofer Institut zum Beispiel ein T-Shirt (das ‚*FitnessSHIRT*‘) entwickelt, welches ausgestattet ist mit Messtechnik für ein 1-Kanal-EKG und Atmungserfassung. Die nutzerfreundliche Integration der Sensorik direkt in das Kleidungsstück vereinfacht – laut *Fraunhofer IIS* – die Handhabung und bietet hohen Tragekomfort.¹¹

⁹ Was man wohl übersetzen kann als: „Ich lerne mich besser kennen, indem ich mich selbst vermesse“.

¹⁰ Vgl. zum Beispiel: „igrowdigital: Schlaftracker – der große Überblick“. (*online*).

¹¹ Vgl. Fraunhofer-Institut für Integrierte Schaltungen (IIS), „FitnessSHIRT – Mit Sicherheit ein guter Einsatz“. (*online*).

Die Selbstvermessung kennt kaum noch Grenzen: Wenige Jahre nachdem die ersten *Facebook*-Freunde mit Smartphone-Anwendungen wie *Runtastic* auf dem sozialen Netzwerk mitteilten, wie weit, wie schnell und wo entlang sie gejoggt sind, wird der Trend nun zum Massenphänomen.¹²

Einen entscheidenden Beitrag dazu leisten Technologiefirmen, die mit *Fitnessstrackern* und Co. Geschäfte machen. Allen voran, *Google* und *Apple*: so ist mit dem neuen mobilen *Apple*-Betriebssystem, *iOS 8*, die Anwendung *HealthKit* auf jedem *iPhone* fest vorinstalliert. Diese zeichnet auf Wunsch Daten zu Bewegung, Ernährung und Schlaf auf. Gleichzeitig bietet das Programm eine Entwickler-Schnittstelle, die auf die Daten anderer Gesundheits-Apps zugreift. „Die damit bereitgestellten Technologieplattformen sollen für die Bürger zum Regieraum des Lebens werden.“ So sagte *Apple*-Manager Craig Federighi auf der *World Wide Developer Conference* bereits 2014: „Unsere Plattform ist der Ort, an dem Anwendungen dazu beitragen können, ein zusammenhängendes Profil Ihrer Aktivitäten und Ihrer Gesundheit zu erstellen.“¹³

Doch längst nicht alle sehen diese Entwicklungen nur positiv. So hat sich die Schriftstellerin Juli Zeh bereits mehrmals öffentlich gegen die private Datensammelwut gestellt und ihre Kritik daran mit dem Statement in einem Gastbeitrag für den *Schweizer Tagesanzeiger* deutlich gemacht:¹⁴ „Quantified Self verabschiedet sich von einer Vernunft, die zum Bestimmen des richtigen Lebens keinen Taschenrechner braucht – Selbstvermessung ist das Gegenteil von Selbstvertrauen“. Und auch Martin U. Müller und Kollegen hinterfragen bereits 2013 in *Spiegel Online*: „Die Macht der Algorithmen erreicht die nächste Stufe: Unternehmen, Behörden und Wissenschaftler versuchen, mit Hilfe gigantischer Mengen persönlicher Daten, unser Leben zu lenken. Wollen wir das?“¹⁵

¹² Vgl. Jan Wilmroth, „Regieraum des Lebens: Selbstvermessung am Handgelenk“, *Süddeutsche Zeitung* (online); sowie zur Vertiefung auch Hristio Boytchev, „Quantified-Self-Bewegung: Miss dich selbst!“, *Spiegel* (04.03.2013). (online). Vgl. auch die ‚Quantified-Self‘-Rubrik „Quantified Self: Vermessen, verbessern, verkaufen“, in *DIE ZEIT*, die in Analysen, Interviews und Videos beschreibt, wo die ‚Quantified-Self‘-Bewegung heute steht, und einen Blick in die Zukunft der Selbstvermessung wirft. (online).

¹³ Vgl. Wilmroth, „Regieraum des Lebens: Selbstvermessung am Handgelenk“.

¹⁴ Ebd.

¹⁵ Vgl. Martin U. Müller, Marcel Rosenbach und Thomas Schulz, „Die gesteuerte Zukunft“, *Spiegel* (13.05.2013). (online).

Warum machen Menschen dies? Warum sind sie so versessen darauf, all diese Daten über sich zu messen, zu sammeln und auszuwerten? Kaum einer der ‚Selbstvermesser‘ hat ausreichend medizinische Kenntnisse, um diese Daten auch sinnvoll zu interpretieren. Oft treibt diese Menschen zunächst nur die Neugier oder der technologische *Hype*. „Für viele Selbstvermesser bleibe das Ziel auch unklar“, erläutert der Mitbegründer der *Quantified-Self*-Bewegung Gary Wolf:

Zu Beginn haben sie eine klare Frage im Kopf, setzen dann ihre Messungen aber fort, weil sie ein nicht zu ignorierendes Geheimnis hinter den Zahlen vermuten. Anderen geht es dagegen sehr konkret darum, ihren Alltag, ihr Gewicht, ihre Fitness oder ihre Gesundheit zu optimieren. Zahlen helfen Ziele zu setzen und zu prüfen, ob der Kurs stimmt: Habe ich mich so viel bewegt wie geplant? Habe ich so viel Wasser getrunken, wie ich wollte, und genügend Obst gegessen?¹⁶

Tatsächlich liefert alleine das regelmäßige Messen der Schritte, der zu sich genommenen Kalorien und der gelaufenen Kilometer, das kontinuierliche Vergleichen der eigenen Werte im Zeitverlauf, aber auch im Vergleich mit anderen, einen Anreiz, der gesundheitsförderlich sein kann. ‚*Gamification*‘ nennt man dieses Prinzip:¹⁷ durch spielerische Elemente Wettkampfsituationen erzeugen und damit die Motivation für bestimmte Aktivitäten erhöhen. Für viele ist das Potenzial der *Quantified-Self*-Bewegung damit aber noch längst nicht ausgereizt. Der Sportmediziner Lutz Graumann zum Beispiel traut ihr eine wichtige Rolle im Bereich der individualisierten Medizin zu. Erhalten die unterschiedlichsten Patienten bei der Therapie verschiedener Erkrankungen heute oft nur eine Standarddosierung (obwohl man weiß, dass die Wirkung von Medikamenten stark von individuellen körperlichen und seelischen Gegebenheiten abhängig ist), so ist er der Meinung, dass zukünftig, durch die Berücksichtigung vieler individueller Daten, Therapien optimiert werden können.¹⁸

Andere aber sehen dies deutlich kritischer und beschreiben das *Quantified-Self*-Phänomen vielmehr als Motor eines bislang dynamischen Marktes für Gerätehersteller und als Tummelplatz für Enthusiasten.¹⁹

¹⁶ Vgl. hierzu Janina Horst, „Die Selbstvermesser“, *Spektrum.de* (02.03.2015). (online).

¹⁷ Vgl. zum Beispiel *Gabler Wirtschaftslexikon*, „Gamification“. (online).

¹⁸ Vgl. Horst, „Die Selbstvermesser“.

¹⁹ Ebd.

Unterstützt Selbstvermessung die Gesundheitsvorsorge?

Nicht nur in den USA erhoffen sich Krankenversicherungen von der massenhaften Verbreitung der mobilen sensorbasierten Geräte und den zugehörigen medizinischen Smartphone-Apps einen Fortschritt in der Gesundheitsvorsorge. Dass Krankenversicherungen durch Prämien eine gesunde Lebensweise fördern, ist nicht neu – für die Anmeldung im Fitnessstudio, regelmäßige Routineuntersuchungen beim Arzt oder den Nachweis anderer gesundheitsfördernder Maßnahmen gibt es bei vielen Kassen und Tarifen Boni. Neu ist allerdings, dass einige Kassen auch die Anschaffung von *Fitnesstrackern* bezuschussen – und weil die *Apple Watch* eine solche Funktion hat, fällt auch das neueste Technik-Spielzeug von *Apple* darunter.²⁰

Die mittels dieser *Apps* gesammelten Daten sind natürlich auch für Versicherungen und Krankenkassen interessant. So bietet die *AOK Nordost* ihren Versicherten die Fitness-App *Dacadoo* für eine gewisse Zeit kostenlos an und errechnet aus den aufgezeichneten Werten deren allgemeinen Gesundheitszustand. Ähnliche Angebote gibt es bereits bei der *DAK*, der *Barmer* und der Betriebskrankenkasse des Autoherstellers *Daimler*.²¹

Noch ist die Förderung der Krankenkassen meist nicht an die Übermittlung von Daten geknüpft. „Wir wollen keine Daten sehen, sondern nur, dass unser Mitglied das gekauft hat“, sagt eine Sprecherin der privaten *DKV*.²² Im Februar 2016 äußerte sich Jens Baas, der Chef der *Techniker-Krankenkasse*, diesbezüglich aber bereits etwas anders: „Auch über Fitness-Tracker sollen die Daten der Versicherten gesammelt – und von den Kassen verwaltet werden“.²³

Die *AOK Nordost* betont in diesem Zusammenhang, man erhebe nur anonyme Daten und keine Detailinformationen über einzelne Versicherte – *Dacadoo* aber erhält diese Daten.²⁴ Was *Dacadoo* oder die vielen anderen Anbieter von *Fitnesstrackern* mit den auf ihren Servern gespeicherten Patientendaten machen, wird von den wenigsten Selbstvermessern hinterfragt. Jens Baas ist davon überzeugt, dass ‚*Big-Data-Firmen*‘ – die

²⁰ Stephan Dörner, „Diese Krankenkassen bezuschussen die Apple Watch“, *Die Welt* (04.08.2015). (online).

²¹ Vgl. Wilmroth, „Regieraum des Lebens: Selbstvermessung am Handgelenk“.

²² Dörner, „Diese Krankenkassen bezuschussen die Apple Watch“.

²³ Guido Bohsem und Ulrich Schäfer, „Krankenkasse wirbt: Fitness-Armband für alle“, *Süddeutsche Zeitung* (08.02.2016). (online).

²⁴ Vgl. Wilmroth, „Regieraum des Lebens: Selbstvermessung am Handgelenk“.

derartige Datenpools auswerten – in den kommenden Jahren mit Macht in das Gesundheitswesen drängen werden. Er deklariert es als sein Ziel, dieses Feld nicht den US-Internetkonzernen zu überlassen. Vielmehr sollen die deutschen Bürger ihre Daten selbst in der Hand behalten. Nach Baas sollen diese Daten in einer elektronischen Patientenakte zusammengeführt werden, in einem persönlichen Datenpool, in dem jeder Patient alle Daten sammelt, die seine Gesundheit betreffen. Gemäß seiner Aussage soll die elektronische Patientenakte klassische medizinische Daten enthalten, und in sie sollen auch Daten einfließen, die zum Beispiel über einen *Fitnessstracker* erhoben werden. Verwalten soll die Daten die jeweilige Krankenkasse, gehören sollen die Daten aber dem Patienten. Wenn er von einer Kasse zu einer anderen wechselt, soll er die Daten mitnehmen können.²⁵

Baas beschreibt es als großen Vorteil der elektronischen Patientenakte und der darin aus *Fitnesstrackern* und anderen tragbaren Geräten gesammelten Daten,

dass Kassen und Ärzte die Patienten viel besser betreuen können: „Wir können über das Risiko einer Erkrankung informieren, wenn wir die Krankheiten, den Puls, das Ausmaß der Bewegung und so weiter zusammen analysieren“.²⁶

Betrachtet man diese, derzeit insbesondere von Krankenkassen angetriebenen Entwicklungen, so ist natürlich auch nicht mehr undenkbar, dass früher oder später die Höhe von Versicherungsbeiträgen an das per *App* aufgezeichnete Verhalten geknüpft wird. Auch wenn entsprechende kleinere Prämienreduktionen von jungen, gesunden Versicherten gerne angenommen werden, so sollte jeder dennoch auch darüber nachdenken, wie die Krankenkasse wohl reagieren wird, wenn man auf einmal wegen eines Beinbruchs – oder eventuell auch einer schwerwiegenderen chronischen Erkrankung – für längere Zeit nicht mehr sportlich aktiv sein kann.

Patientengeführte elektronische Patientenakten

Diese Bemühungen nehmen derzeit schon sehr konkrete Formen an. So ist zum Beispiel die *AOK Nordost*, gemeinsam mit dem Technologiepartner *CISCO*, seit Anfang 2016 dabei, eine solche elektronische

²⁵ Guido Bohsem und Ulrich Schäfer, „Krankenkasse wirbt: Fitness-Armband für alle“.

²⁶ Ebd.

Patientenakte zu entwickeln.²⁷ Die *Vivantes Kliniken* und die *SANA Kliniken AG* haben sich diesem Projekt bereits angeschlossen.²⁸ Jens Baas beginnt seine Ankündigungen ebenfalls umzusetzen. Die ‚*Techniker Krankenkasse*‘ hat 2016 die Entwicklung einer elektronischen Patientenakte für ihre Versicherten öffentlich ausgeschrieben und sucht nach einem Technologiepartner. Auch in Bayern ist man auf diesen Zug aufgesprungen: Die Bayerische Staatsregierung überlegt derzeit die Initiierung eines regional begrenzten Modellprojekts und lässt dessen Machbarkeit analysieren.

Da Bundesgesundheitsminister Hermann Gröhe die Einführung einer patientengeführten elektronischen Patientenakte sowie eines sogenannten Patientenfachs in seinem *E-Health-Gesetz* vom November 2015 fordert,²⁹ ist es absehbar, dass sich 2017 weitere *Player* dieser Herausforderung stellen und ähnliche Entwicklungen anstoßen werden.³⁰ Dazu passt auch das aktuelle Strategiepapier der CDU, in dem die Selbstbestimmung des Patienten durch *E-Health*-Anwendungen betont und darauf hingewiesen wird, dass

die Digitalisierung für Patienten und Ärzte gleichermaßen Vorteile schafft: Mit der ePatientenakte, auf die der Patient über das Internet jederzeit zugreifen kann, hat er selbst einen Überblick über Diagnosen und Therapien und ist wesentlich umfassender informiert als bisher und kann somit auch weit besser in gemeinsame Entscheidungsprozesse eingebunden werden.³¹

²⁷ Vgl. CISCO Pressemitteilung, „Krankenversicherung: Digitale Plattform für Versicherte“ (20.04.2016). (online).

²⁸ Vgl. AOK, „Digitales Gesundheitsnetzwerk will Versorgung optimieren“, AOK Nordost (13.09.2016). (online).

²⁹ „Patienten erhalten ab 2018 einen Anspruch darauf, dass ihre auf der Gesundheitskarte gespeicherten Daten in ein Patientenfach aufgenommen werden. In dieses Patientenfach können auch eigene Daten wie – z.B. ein Tagebuch über Blutzuckermessungen – abgelegt werden. In diesem Patientenfach können Patienten ihre Daten künftig auch außerhalb der Arztpraxis eigenständig einsehen.“ Vgl. Bundesministerium für Gesundheit, „Die elektronische Gesundheitskarte“ (19.02.2016). (online).

³⁰ Für eine etwas detailliertere Betrachtung der AOK-/TK-Projekte und deren Einbettung in die Vorgaben des *E-Health-Gesetzes* sei auf Sebastian Vorberg, „Digitalisierung im Gesundheitswesen“, *Vorberg & Partner* (22.12.2016), (online) verwiesen. Vgl. auch Philipp Grätzel von Grätz, „Gesundheitsregionen: Neue Vorreiter bei der Vernetzung?“, *Ärztezeitung*. (online).

³¹ Vgl. „Digitalisierung im Gesundheitswesen: CDU legt Strategiepapier vor“, *Ärzteblatt* (25.01.2017). (online). Vgl. auch Bundesfachausschuss Gesundheit und Pflege sowie Netzwerk Digitalisierung, „E-Health-Strategie für Deutschland: 12 Punkte für ein digitalisiertes Gesundheitswesen“. *CDU.de*. (online).

Im Zuge der nun initiierten Pilot- und Modellprojekte wird es sicherlich wichtig sein, welche Anbieter sich letztendlich als Betreiber und Datenhalter derartig persönlicher elektronischer Patientenakten herauskristallisieren und welche Datennutzungsszenarien sich daraus entwickeln werden. Langfristig wird das Gewinnen des Vertrauens der Bürger in die sorgsame, den Datenschutzregularien gerecht werdende, Speicherung und Verwendung der Daten, aber auch in den sinnvollen medizinischen Nutzen, der aus der Analyse solcher Daten gezogen wird, wesentlich für den Erfolg dieser Entwicklungen sein.

Kritiker weisen zurecht darauf hin, dass ähnliche Entwicklungen in den USA bereits am Ende des letzten Jahrhunderts und in Deutschland nur wenige Jahre später angestoßen wurden,³² dass aber auch über zehn Jahre später der Mehrwert persönlicher elektronischer Patientenakten für die Verbesserung der medizinischen Versorgung oder die Optimierung von Vorsorge-, sowie Nachsorge-/Monitoringprogrammen noch nicht valide belegt wurde.³³ Vielmehr hatte zum Beispiel die *Barmer Krankenkasse* ihren Versicherten bereits von 2007 bis 2010 eine solche elektronische Patientenakte angeboten und in einem begleitenden Forschungsprojekt „Nutzen und Akzeptanz von elektronischen Gesundheitsakten“ evaluiert.³⁴ Als ein wesentliches Ergebnis dieses Forschungsprojekts berichteten Hanna Kirchner und Kollegen im Jahr 2010:

Der allseits postulierte und erwartete Nutzen einer eGA korreliert also in Deutschland (noch) nicht mit den Erwartungen der Versicherten, den technischen Infrastrukturvoraussetzungen bei den Ärzten sowie der notwendigen flächendeckenden Aufklärung über dieses neue Medium. Ärzte müssen deshalb für eine aktive Mitarbeit gewonnen werden.

³² Vgl. zum Beispiel Roy Schoenberg und Charles Safran, „Internet Based Repository of Medical Records that Retains Patient Confidentiality“, *BMJ* (2000): 321: 1199-1203; vgl. auch Frank Ückert *et al.* „Functions of an Electronic Health Record“, *International Journal of Computerized Dentistry* 5.2-3 (2002): 125-132.

³³ Vgl. zum Beispiel Paul C. Tang, Joan S. Ash, David W. Bates, J. Marc Overhage and Daniel Z. Sands, „Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption“, *Journal of the American Medical Informatics Association* 13.2 (2006): 121-126. Vgl. auch Hanna Kirchner, Hans-Ulrich Prokosch, Joachim Dudeck *et al.*, „Die BARMER Gesundheitsakte: Erste Erfahrungen mit einer patientengeführten Gesundheitsakte“, in: *e-Health 2010*, hg. Frank Duesberg (Solingen: Medical Future Verlag, 2010), 155-160.

³⁴ In Anlehnung an den im englischen verwendeten Begriff eines ‚personal electronic health records‘, oft *pEHR* oder *PHR* abgekürzt, wurde eine solche patientengeführte persönliche Patientenakte über viele Jahre hinweg als ‚elektronische Gesundheitsakte‘ (abgekürzt EGA oder eGA) bezeichnet

Sie schlussfolgerten damals:

Die persönliche Gesundheitsakte wird sich langfristig nur dann durchsetzen, wenn es gelingt, eine Umgebung aufzubauen, in der wichtige gesundheitsbezogene Daten mit einer gewissen Automatik ohne zusätzliche Belastung der Ärzte und Kliniken in die eGA übertragen werden können.³⁵

Die *Barmer eGA* wurde 2011 dann auch wieder eingestellt. Ähnlich erfolglos war die Initiative von *Google Health™*, die 2011 eingestellt wurde.³⁶

An die *Quantified-Self*-Bewegung und sensorbasierte *Fitnessstracker* sowie *Smartwatches* hat zum damaligen Zeitpunkt wohl noch niemand gedacht. Doch erfüllen heute gerade diese tragbaren Geräte die geforderte Voraussetzung, dass sie „mit einer gewissen Automatik ohne zusätzliche Belastung der Ärzte und Kliniken Daten in einen Datenspeicher übertragen“. Verknüpft mit der nun auch gesetzlichen Verankerung der Übernahme solcher patientengenerierter Daten in patientengeführte elektronische Patientenakten, könnte dies der Türöffner sein, der diesen frühen Visionen nun in den kommenden Jahren den Weg ebnet.

Von der Selbstvermessung zum medizinischen Einsatz?

Gewiss ist aber, dass die Selbstvermesserbewegung ihr medizinisches Potenzial bis dato weder bewiesen noch ausgeschöpft hat. Kommen wir noch einmal zurück auf das vom *Fraunhofer IIS* propagierte *FitnessSHIRT*, welches mit der Erfassung eines 1-Kanal-EKGs und der Atembewegung im Brustbereich Messwerte zweier grundlegender medizinischer Vitalparameter liefert. Zur Erfassung des EKGs kommen leitfähige textile Elektroden zum Einsatz, die direkt in das T-Shirt integriert wurden. Ein flexibles Band im unteren Bereich des Thorax erfasst die Atembewegung des Brustkorbs resistiv. Die gemessenen Daten werden in Echtzeit per Funk an eine Einsatzzentrale übertragen und dort ausgewertet. Zusätzlich können die Daten auf einer Speicherkarte zwischengespeichert und zur weiteren Auswertung auf einen PC übertragen werden. Die Elektronik zur Messwerterfassung und Speicherung sowie Funkübertragung befindet sich in einem separaten, abnehmbaren Gehäuse. Das Gehäuse hat etwa die Größe eines Smartphones, wird mit Druckknöpfen am *FitnessSHIRT*

³⁵ Vgl. Kirchner, Prokosch, Dudeck *et al.*, „Die BARMER Gesundheitsakte: Erste Erfahrungen mit einer patientengeführten Gesundheitsakte“, 155-160.

³⁶ Vgl. Richard Brandt and Rich Rice, „Building a Better PHR Paradigm: Lessons from the Discontinuation of Google Health“, *Health Policy and Technology* 3.3 (2014): 200-207.

befestigt und kann zum Waschen des T-Shirts abgenommen werden.³⁷ Technologisch könnte dieses T-Shirt wohl in der Fernüberwachung von Patienten (Telemonitoring) eingesetzt werden. Entsprechend beschreibt das *Fraunhofer IIS* das *FitnessSHIRT* als

ein Kleidungsstück zur mobilen, kontinuierlichen Erfassung von Körpersignalen, dessen mögliche Anwendungsbereiche die Betreuung von Risikopatienten (z.B. Herz-Kreislauf-Erkrankungen), die Erhöhung der Sicherheit für Einsatzkräfte in Gefahrensituationen, die Unterstützung im Rahmen der Leistungsdiagnostik, die Biofeedback-Behandlung und Stressmanagement und die Anwendung bei Sport und Freizeit sind.³⁸

Allerdings ist das *FitnessSHIRT* für den Einsatz in der medizinischen Versorgung noch nicht zugelassen, da es noch nicht CE-zertifiziert ist und die Einhaltung der Anforderungen des Medizinproduktegesetzes (MPG) und der europäischen Richtlinie 93/42/EWG noch nicht nachgewiesen wurde. Diese Hürde zu überwinden, ist nicht einfach, daran scheitern derzeit noch viele im Fitness-Bereich eingesetzte sensorbasierte Geräte. Und so wird das von *Fraunhofer IIS* entwickelte *Fitness-Shirt* wohl vor allem seinem Namen Ehre machen und noch länger nur im Sport und Freizeitbereich eingesetzt werden.

Ähnlich ging es auch *Google* mit einer Entwicklung, die 2014 erstmals Schlagzeilen machte. Auch bei der Blutzuckermessung und der Übertragung der gemessenen Werte auf ein Smartphone ging die Entwicklung nämlich längst weiter. Der Blutzuckerspiegel kann nicht nur im Blut, sondern auch in sämtlichen anderen Körperflüssigkeiten gemessen werden. *Google* macht sich nun die Miniaturisierung der Analysetechnik zu Nutze und nimmt die Tränenflüssigkeit für die Blutzuckerbestimmung ins Visier. Zwischen zwei Schichten aus weichem Kontaktlinsenmaterial sind ein miniaturisierter Glukose-Sensor, der durch eine Pore in der Linse im Kontakt mit der Tränenflüssigkeit steht, und ein Microchip eingebettet, der Informationen an ein externes Gerät übermittelt. Dieses schlägt Alarm, wenn der Zuckerspiegel zu stark nach oben oder unten vom Normwert abweicht. Eine ringförmige Mikroantenne, feiner als ein menschliches Haar, sendet die Daten und empfängt drahtlos Energie von einer nahegelegenen Quelle. Damit möchte *Google* Diabetikern in Zukunft mehrere unangenehme Fingerpiekser täglich ersparen. Auch

³⁷ Vgl. Fraunhofer-Institut für Integrierte Schaltungen IIS, „FitnessSHIRT: Mit Sicherheit ein guter Einsatz“.

³⁸ Ebd.

wenn diese Art der Blutzuckermessung seine medizinische Validität noch nicht bewiesen hat, so hat *Google* sich seine smarte Kontaktlinse doch schon patentieren lassen, und im Rahmen einer Partnerschaft soll das Produkt an das Schweizer Pharmaunternehmen *Novartis* lizenziert werden.³⁹ Und auch hier gilt derzeit, dass diese Kontaktlinse als Medizinprodukt natürlich bisher noch nicht zugelassen ist. Die dafür notwendigen klinischen Studien soll dann *Novartis* durchführen.⁴⁰

Ein weiteres Beispiel für die medizinische Nutzung von Sensoren – in diesem Falle integriert in die Sohle von Sportschuhen – gibt das *eGait* Projekt der Abteilung Molekulare Neurologie am Erlanger Universitätsklinikum gemeinsam mit dem FAU Lehrstuhl für Mustererkennung und dem Industriepartner *Astrum IT*. In die Schuhsohle wurde in diesem Projekt ein Beschleunigungssensor (Accelerometer) und ein Gyroskop integriert, um somit kontinuierlich den Gang von Parkinson-Patienten im Frühstadium überwachen und somit frühzeitig eine Krankheitsverschlechterung identifizieren zu können.⁴¹ Die Forschergruppe erhofft sich aus der Weiterentwicklung und zukünftigen Vermarktung dieser Technologie für die ‚Diagnostik von morgen‘ ein mobiles Sensorsystem zur automatisierten Bewegungsanalyse, das eine objektive und kontinuierliche Datenbasis liefert, sodass aus den aufgezeichneten Gangsignalen eines Parkinson-Patienten dessen Bewegungsstörung frühzeitig erkannt und objektiv beurteilt werden kann.

Allgemein scheint der Schuh aber auch im Sport- bzw. Fitnessbereich ein beliebtes Einsatzgebiet für Sensoren zur Gangvermessung zu sein. So propagierte die Firma *Adidas* bereits 2011 einen Sportschuh mit Sensor, der die Leistungsauswertung über eine *iPhone App* ermöglichte.⁴² Ähnliche Ansätze verfolgte die Firma *beMeister*, wollte sich aber vom Schuhhersteller unabhängig machen und integrierte die entsprechende Sensorik in eine Einlegesohle.⁴³ Wirklich durchgesetzt hat sich aber augenscheinlich keine dieser Technologien, denn die Firma *beMeister* ist

³⁹ Vgl. „Gibt’s bald die Zuckercheck-Kontaktlinse für Diabetiker?“, *EYEGLOSS24 Blog* (16.03.2016). (online).

⁴⁰ Ebd.

⁴¹ Vgl. „Forschungsprojekt eGait“ (online); vgl. auch Jochen Klucken *et al.*, „Unbiased and Mobile Gait Analysis Detects Motor Impairment in Parkinson’s Disease“, *PLoS One* 8.2 (2013): e56956. (online).

⁴² Vgl. „Neuer Adidas-Fußballschuh mit Sensor erlaubt Leistungsauswertung auf dem iPhone“ (16.11.2011). (online). Vgl. auch „Neuer Adidas Fußballschuh merkt sich alles“ (25.11.2011). (online).

⁴³ Vgl. „beMeister: Chip im Schuh misst Trainingsleistung“, *Chip* (05.02.2012). (online).

mittlerweile vom Markt verschwunden, und zum in die Sohle eines *Adidas* Schuhs integrierbaren ‚*adidas* Laufcomputer *Speed Cell*‘ findet man sehr unterschiedliche Bewertungen.⁴⁴

Datenschutz und Datensicherheit

Zum Abschluss dieses Kapitels muss ich noch ein wenig Wasser in den Wein gießen und einige kritische Fragen aufwerfen, die bei aller Euphorie und allem *Hype* nicht unter den Tisch fallen dürfen.

Denken all die Selbstvermesser unter uns überhaupt darüber nach, wo ihre gemessenen, ‚patientengenerierten‘ Daten abgespeichert werden? Sind wir uns darüber im Klaren, welche Wege diese Daten über das Internet gehen, wer sie alles mitlesen kann, und was damit wirklich gemacht wird? Fast nie dienen diese Daten nur dem einen Zweck, für den wir sie generieren – sei es zur Steigerung unserer Fitness oder auch in einem medizinischen Vorsorge- oder Monitoringszenario. Wie schützt der Hersteller unserer *Smartwatch* die damit generierten Daten? Auf welcher *Cloud*, in welchem Land, unter welchen Sicherheitsmaßnahmen und unter der Gültigkeit welcher Datenschutzgesetze werden diese Daten verwaltet? Welche *Big-Data*-Analysen macht wer mit welcher Absicht mit unseren persönlichen Daten?

Diesen und ähnlichen Fragen ist die Firma *Symantec* in ihrer Studie „How safe is your Quantified Self?“ nachgegangen und kam in Bezug auf die Datensicherheit und den Datenschutz vieler *Quantified-Self*-Anwendungen unter anderem zu den Ergebnissen, dass 52% der *Apps* keinerlei Aussagen zu Datenschutz veröffentlichten, die untersuchten *Apps* ihre erhobenen Daten oft über eine Vielzahl von Servern im Internet transportieren, bevor diese auf dem eigentlichen Zielservers gespeichert werden („the maximum number of unique domains contacted by a single app was 14“) und dass 20% der *Apps* die Passwörter der Benutzer unverschlüsselt über das Internet übertragen.⁴⁵ Etwas erstaunlich mutet es in diesem Zusammenhang an, dass potentielle Datenschutzprobleme bei der elektronischen Gesundheitskarte in den Medien sehr schnell und lautstark vermeldet werden,⁴⁶ dass die Datenschutzmaßnahmen bei der

⁴⁴ Vgl. Kundenrezensionen zum ‚*adidas* Laufcomputer *Speed Cell PC/Mac,Black*‘, NS, V42039. (online).

⁴⁵ Vgl. Mario Ballano Barcena, Candid Wueest und Hon Lau, „How Safe is Your Quantified Self?“, *Symantec* (11.08.2014). (online).

⁴⁶ Vgl. zum Beispiel „Massives Sicherheitsleck bei Gesundheitsdaten“, *Deutsche Apotheker Zeitung* (25.06.2015). (online).

privaten Nutzung von *Fitnesstrackern* und anderen *Hype-Technologien* von den wenigsten Bürgern hinterfragt wurden. So findet man beispielsweise bei den Kundenfragen und Kommentaren zum *adidas* Laufcomputer *Speed Cell* zwar eine Menge von Rückfragen zur Batterielaufzeit und der Übertragung der Daten vom Chip per USB-Stick an einen PC, doch das Wort Datenschutz ist dort an keiner Stelle zu finden. Tatsächlich findet man bei *Adidas* die Informationen zum Datenschutz auch auf der Website der zugehörigen *App*,⁴⁷ doch bleibt es fraglich, wie viele der Käufer sich die entsprechenden 19 DIN A4 Seiten jemals durchgelesen haben. Es ist somit nachvollziehbar, dass Justizminister Heiko Maas vor dem Hintergrund der weltweiten *Big-Data*-Aktivitäten den Missbrauch der Daten verhindern will, die von *Fitnesstrackern* erhoben werden, und Regeln für Gesundheitsdaten fordert.⁴⁸ Erstaunlich wiederum ist dann die Tatsache, dass die deutschen Bürger, diese Bedenken offensichtlich weniger teilen. So berichtet *DAZ.Online* im Januar 2016 über eine Studie des *Vodafone*-Instituts und die dort erhobenen Ergebnisse, nach denen zwar eine weit verbreitete Skepsis zu *Big Data* aufgezeigt wurde, die Befragten aber gerade in Bezug auf ihre Gesundheitsdaten doch weniger kritisch waren („Mehr als die Hälfte der Befragten hat keine Bedenken, anonymisierte Daten zur besseren Behandlung von Krankheiten zur Verfügung zu stellen“).⁴⁹ Offensichtlich müssen Krankenversicherungen aber noch ein wenig an der Vertrauensbeziehung zu ihren Versicherten arbeiten, denn nur 25% der Befragten würden nach dieser Studie der Erhebung und Speicherung ihrer Gesundheitsdaten durch die Krankenkassen zustimmen.⁵⁰

Mit welchen ethischen Fragestellungen Institutionen, die im Rahmen der Selbstvermessung immer mehr Daten über uns erheben und sammeln, unabhängig von allen Datenschutzregularien, konfrontiert werden können, zeigte sich im Frühjahr 2016 am Beispiel der Firma *Apple*. Es ging dabei um das *iPhone* des Terroristen, der am 2. Dezember 2015 in San Bernardino (Kalifornien) 14 Menschen erschossen hatte, bevor er selbst von Polizisten erschossen wurde. Natürlich wollte das *FBI* gerne an die Inhalte des Handys des Terroristen kommen, um darüber an Informationen über eventuelle Mittäter zu gelangen. Um den Inhalt des

⁴⁷ Vgl. *adidas* Datenschutzerklärung (17.12.2016). (online).

⁴⁸ Vgl. Guido Bohsem, „Regeln für Gesundheitsdaten“, *Süddeutsche Zeitung* (09.02.2016). (online).

⁴⁹ Vgl. „TNS Infratest-Umfrage zu Big Data: Meine Gesundheitsdaten gebe ich gern...“, *Deutsche Apotheker Zeitung* (19.01.2016). (online).

⁵⁰ Ebd.

Mobiltelefons aber nicht durch falsches Eingeben der vier Passwort-Ziffern auf immer zu verlieren, bat das *FBI* die Firma *Apple* um Hilfe. Berechtigterweise verschloss sich *Apple* dieser Forderung und berief sich auf die Privatsphäre des Handy-Besitzers. Dennoch wurde dieses Beispiel in den Medien sehr unterschiedlich und kontrovers diskutiert.⁵¹ Mathias Müller von Blumencron brachte die Diskussion auf den Punkt und wies darauf hin, dass „uns die Argumente von Apple-Chef Tim Cook alle angehen. Es geht um nichts Geringeres als die Freiheit unserer Gedanken.“⁵²

⁵¹ Vgl. unter anderem Helmut Martin-Jung, „Code der Freiheit“, *Süddeutsche Zeitung* (17.02.2016). (online). Vgl. auch Georgij Bowt, „Hält Apple die Verteidigungsposition gegen FBI-Ansprüche wegen der Freiheit oder für Geld?“, *fit4Russland* (25.02.2016). (online).

⁵² Vgl. Mathias Müller von Blumencron, „FBI gegen Apple: Vorstoß ins Ich“, *Frankfurter Allgemeine Zeitung* (26.02.2016). (online).

Literaturverzeichnis

- Brandt, Richard and Rich Rice. „Building a Better PHR Paradigm: Lessons from the Discontinuation of Google Health“. *Health Policy and Technology* 3.3 (2014): 200-207.
- Kirchner, Hanna, Hans Ulrich Prokosch, Joachim Dudeck *et al.* „Die BARMER Gesundheitsakte: Erste Erfahrungen mit einer patientengeführten Gesundheitsakte“. In: *e-Health 2010*. Hg. Frank Duesberg. Solingen: Medical Future Verlag, 2010, 155-160.
- Klucken, Jochen, Jens Barth, Patrick Kugler *et al.* „Unbiased and Mobile Gait Analysis Detects Motor Impairment in Parkinson’s Disease“. *PLoS One* 8.2 (2013): e56956. doi: 10.1371/journal.pone.0056956.
- Schoenberg, Roy and Charles Safran. „Internet Based Repository of Medical Records that Retains Patient Confidentiality“. *BMJ* 321: 1199-1203.
- Tang, Paul C., Joan S. Ash, David W. Bates *et al.* „Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption“. *Journal of the American Medical Informatics Association* 13.2 (2006): 121-126.
- Ückert, Frank, Michael Ataian, Michael Görz and Hans-Ulrich Prokosch. „Functions of an Electronic Health Record“. *International Journal of Computerized Dentistry* 5.2-3 (2002): 125-132.

Internetquellen

Anonymous:

- „beMeister: Chip im Schuh misst Trainingsleistung“. Chip (05. Februar 2012). http://www.chip.de/news/beMeister-Chip-im-Schuh-misst-Trainingsleistung_54210205.html [zuletzt besucht 26.01.2017].
- „Digitalisierung im Gesundheitswesen: CDU legt Strategiepapier vor“. *Ärzteblatt* (25.01.2017). <http://www.aerzteblatt.de/nachrichten/72690> [zuletzt besucht 26.01.2017].
- „Forschungsprojekt eGait“. <http://www.egait.de/> [zuletzt besucht 26.01.2017].
- „igrowdigital: Deutschland ist quantifiziert“. <http://igrowdigital.com/2012/08/deutschland-ist-quantifiziert/> [zuletzt besucht 21.01.2017].
- „igrowdigital: Schlaftracker: der große Überblick“. <http://igrowdigital.com/2015/11/schlaftracker-der-groese-ueberblick/> [zuletzt besucht 21.01.2017].

- „Neuer Adidas Fußballschuh merkt sich alles“. (25. November 2011). <http://www.materna-newmedia.de/blog/allgemein/neuer-adidas-fussballschuh-merkt-sich-alles/> [zuletzt besucht 26.01.2017].
- „Neuer Adidas-Fußballschuh mit Sensor erlaubt Leistungsauswertung auf dem iPhone“. (16. November 2011). <https://www.iphone-ticker.de/adizero-f50-micoach-iphone-27688/> [zuletzt besucht 26.01.2017].
- „QS: Quantified Self: Self knowledge through numbers (Deutsche Community)“. <http://qsdeutschland.de/> [zuletzt besucht 21.01.2017].
- „TNS Infratest-Umfrage zu Big Data: Meine Gesundheitsdaten gebe ich gern...“. *DAZ.Online* (19. Januar 2016). <https://www.deutsche-apotheker-zeitung.de/news/artikel/2016/01/19/weniger-datenschutz-skepsis-bei-gesundheitseinrichtungen> [zuletzt besucht 26.01.2017].
- ADIDAS Datenschutzerklärung (17. Dezember 2016). http://www.adidas.de/help-topics-privacy_policy.html [zuletzt besucht 26.01.2017].
- AOK. „Digitales Gesundheitsnetzwerk will Versorgung optimieren“. *AOK.DE* (13. September 2016). <https://nordost.aok.de/inhalt/digitales-gesundheitsnetzwerk-will-versorgung-optimieren/> [zuletzt besucht 26.01.2017].
- Bundesfachausschuss Gesundheit und Pflege der CDU sowie Netzwerk Digitalisierung. „E-Health-Strategie für Deutschland: 12 Punkte für ein digitalisiertes Gesundheitswesen“. *CDU.de* (25. Januar 2016). https://www.cdu.de/system/tdf/media/dokumente/170125_ehealth_bf_a_gesundheit_digital.pdf?file=1 [zuletzt besucht 26.01.2017].
- Bundesministerium für Gesundheit. „Die elektronische Gesundheitskarte“. (19.02.2016). <https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/e-health-gesetz/allgemeine-informationen-egk.html#c1063> [zuletzt besucht 26.01.2017].
- CERN. „Tim Berners-Lee’s proposal“. <http://info.cern.ch/Proposal.html> [zuletzt besucht 07.01.2017].
- CERN. „World Wide Web“. <http://info.cern.ch/hypertext/WWW/TheProject.html> [zuletzt besucht 20.01.2017].
- CISCO. „Krankenversicherung: Digitale Plattform für Versicherte“. CISCO Pressemeldung (20. April 2016). <http://globalnewsroom.cisco.com/de/de/press-releases/krankenversicherung-digitale-plattform-fur-versic-nasdaq-cisco-1254838> [zuletzt besucht 26.01.2017].

- Fraunhofer-Institut für Integrierte Schaltungen IIS. „FitnessSHIRT – Mit Sicherheit ein guter Einsatz“. <https://www.iis.fraunhofer.de/de/ff/med/prod/fitnessshirt.html> [zuletzt besucht 17.01.2017].
- Albrecht, Urs-Vito. Hg. *Chancen und Risiken von Gesundheits-Apps (CHARISMHA)*. Medizinische Hochschule Hannover, 2016.
urn:nbn:de:gbv:084-16040811167.
<http://www.digibib.tu-bs.de/?docid=60002>.
[zuletzt besucht 16.02.2017].
- Barcena, Mario Ballano, Candid Wueest and Hon Lau. „How safe is your quantified self?“. *Symantec* (11. August 2014).
<https://www.symantec.com/content/dam/symantec/docs/white-papers/how-safe-is-your-quantified-self-en.pdf>
[zuletzt besucht 26.01.2017].
- Bendel, Oliver. „Gamification“. *Gabler Wirtschaftslexikon*.
<http://wirtschaftslexikon.gabler.de/Definition/gamification.html>
[zuletzt besucht 21.01.2017].
- Bohsem, Guido. „Regeln für Gesundheitsdaten“. *Süddeutsche Zeitung* (09. Februar 2016). <http://www.sueddeutsche.de/wirtschaft/big-data-regeln-fuer-gesundheitsdaten-1.2855818?reduced=true>
[zuletzt besucht 26.01.2017].
- Bohsem, Guido und Ulrich Schäfer. „Krankenkasse wirbt: Fitness-Armband für alle“. *Süddeutsche Zeitung* (08. Februar 2016).
<http://www.sueddeutsche.de/wirtschaft/montagsinterview-krankenkassen-chef-wir-muessen-ein-cooles-produkt-anbieten-1.2854002> [zuletzt besucht 17.01.2017].
- Bowt, Georgij. „Hält Apple die Verteidigungsposition gegen FBI: Ansprüche wegen der Freiheit oder für Geld?“ *fit4Russland* (25. Februar 2016). <https://www.fit4russland.com/wirtschaft/1614-haelt-apple-die-verteidigungsposition-gegen-fbi-ansprueche-wegen-der-freiheit-oder-fuer-geld> [zuletzt besucht 26.01.2017].
- Grätzel von Grätz, Philipp. „Gesundheitsregionen: Neue Vorreiter bei der Vernetzung?“ *Ärzte Zeitung* (15. November 2016).
http://www.aerztezeitung.de/praxis_wirtschaft/e-health/telemedizin/article/923727/gesundheitsregionen-neue-vorreiter-vernetzung.html
[zuletzt besucht 26.01.2017].

- Grote, Caspar. „Diabetes-App von mySugr: Smartphone überwacht Blutzucker“. *medizin-und-elektronik.de* (11. Juni 2014).
<http://www.medin-und-elektronik.de/sonstige/artikel/109834/>
[zuletzt besucht 21.01.2017].
- Hohberg, Fabian. „Meine Smartwatch parkt noch schnell den Wagen ein“. *PS-Welt* (04. Februar 2016).
<https://www.welt.de/motor/article151835280/Meine-Smartwatch-parkt-noch-schnell-den-Wagen-ein.html> [zuletzt besucht 07.01.2017].
- Horst, Janina. „Die Selbstvermesser“. *Spektrum.de* (02. März 2015).
<http://www.spektrum.de/news/die-selbstvermesser/1334103>
[zuletzt besucht 17.01.2017].
- Kannenberg, Axel. „BMW: Autonom Einparken per Smartwatch-Befehl“. *Heise Online* (16. Dezember 2014).
<https://www.heise.de/newsticker/meldung/BMW-Autonom-Einparken-per-Smartwatch-Befehl-2498100.html>
[zuletzt besucht 07.01.2017].
- Martin-Jung, Helmut. „Code der Freiheit“. *Süddeutsche Zeitung* (17. Februar 2016). <http://www.sueddeutsche.de/politik/datenschutz-code-fuer-die-freiheit-1.2867468> [zuletzt besucht 26.01.2017].
- Meola, Andrew. „What is the Internet of Things (IoT)?“ *Business Insider Deutschland* (19. Dezember 2016).
<http://www.businessinsider.de/what-is-the-internet-of-things-definition-2016-8?r=US&IR=T> [zuletzt besucht 07.01.2017].
- Müller von Blumencron, Mathias. „FBI gegen Apple: Vorstoß ins Ich“, *Frankfurter Allgemeine Zeitung* (26. Februar 2016).
<http://www.faz.net/aktuell/politik/ausland/amerika/fbi-gegen-apple-vorstoess-ins-ich-14092084.html> [zuletzt besucht 21.01.2017].
- Schipper, Lena. „Was eigentlich ist das Internet der Dinge?“ *Frankfurter Allgemeine Zeitung* (17. März 2015).
<http://www.faz.net/aktuell/wirtschaft/cebit/cebit-was-eigentlich-ist-das-internet-der-dinge-13483592.html> [zuletzt besucht 07.01.2017].
- Vorberg, Sebastian. „Digitalisierung im Gesundheitswesen“. *Vorberg & Partner* (22. Dezember 2016).
<http://www.medizinanwalt.de/2016/12/22/allgemein/die-digitalisierung-im-gesundheitswesen-einfuehrung-einer-elektronischen-patientenakte/> [zuletzt besucht 26.01.2017].

Autoren- und Herausgeberverzeichnis

RUDOLF FREIBURG, Universitätsprofessor Dr., Inhaber des Lehrstuhls für Anglistik, insbesondere Literaturwissenschaft

Anschrift: Bismarckstr. 1, 91054 Erlangen

E-Mail: rudolf.freiburg@fau.de

Rudolf Freiburg studierte Germanistik, Anglistik und Pädagogik an der Westfälischen Wilhelms-Universität Münster. Nach dem Ersten Staatsexamen 1982 promovierte er 1984 mit einer Studie zur ‚Intentionalität literarischer Texte‘. Von 1985 bis 1986 war Freiburg Lehramtsanwärter im Bezirksseminar Rheine, bevor er 1986 als Hochschulassistent an die Georg-August-Universität Göttingen kam. 1992 habilitierte er sich mit einer Arbeit über die Rolle der Essayistik bei Samuel Johnson. Von 1992 bis 1995 lehrte und forschte er als Hochschuldozent am Englischen Seminar der Universität Göttingen, bevor er 1995 einen Ruf auf den Lehrstuhl für Anglistik an die FAU Erlangen-Nürnberg erhielt. In seinen Forschungen interessiert sich Freiburg für die Literatur und Kultur des 18. Jahrhunderts, insbesondere für Texte von Daniel Defoe, Joseph Addison, Jonathan Swift, Samuel Richardson und Samuel Johnson. Ein weiterer Forschungsschwerpunkt liegt im Bereich der englischen Gegenwartsliteratur. Von 2002 bis 2004 war Freiburg Dekan der Philosophischen Fakultät, von 2004 bis 2007 Senator der FAU. Zehn Jahre lang (2006-2016) leitete er das Interdisziplinäre Zentrum für Literatur und Kultur der Gegenwart, an dessen Gründung er auch maßgeblich beteiligt war. Seit 2014 koordiniert er die Erlanger Universitätstage in Amberg und Ansbach.

SEBASTIAN BÜTTNER, Vertretungsprofessor Dr., Vertreter des Lehrstuhls für Allgemeine Soziologie und Kultursoziologie

Anschrift: Institut für Soziologie, Kochstr. 4, 91054 Erlangen

E-Mail: sebastian.buettner@fau.de

Sebastian Büttner studierte Soziologie im Studienschwerpunkt „Globale und europäische Vergesellschaftung“ an der Otto-Friedrich-Universität Bamberg und an der Universität Warschau. Nach dem Abschluss als Diplom-Soziologe 2004 promovierte er im Jahr 2010 als Promotionsstipendiat der *Bremen International Graduate School of Social Sciences* (BIGSSS) mit einer Arbeit zur regionalen Mobilisierung in Europa. Die Promotionsschrift wurde im Jahr 2010 unter dem Titel *Mobilizing Regions*,

Mobilizing Europe: Expert Knowledge and Scientific Planning in European Regional Development bei Routledge veröffentlicht. Seit April 2010 ist Sebastian Büttner als wissenschaftlicher Assistent und seit April 2017 auch als wissenschaftlicher Oberassistent am Institut für Soziologie der Friedrich-Alexander-Universität Erlangen-Nürnberg tätig. Im Februar 2017 erfolgte seine Habilitation im Fach Soziologie. Büttner vertritt derzeit den Lehrstuhl für Allgemeine Soziologie und Kulturosoziologie. Er ist darüber hinaus Mitglied der DFG-geförderten Forschergruppe (FOR 1539) „Europäische Vergesellschaftung“ und im Rahmen dieses Forschungsverbunds Leiter eines Teilprojekts zur „Professionalisierung von EU-Expertise“. Besondere Arbeitsschwerpunkte bilden die Kulturosoziologie und die soziologische Theoriebildung, die Europaforschung, die Soziologie der Transnationalisierung und der Globalisierung sowie die Soziologie der Politik.

GEORG GLASZE, Universitätsprofessor Dr., Inhaber des Lehrstuhl für Geographie (Kulturgeographie und Orientforschung)

Anschrift: Wetterkreuz 15, 91058 Erlangen

E-Mail: georg.glasze@fau.de

Prof. Dr. Georg Glasze studierte Geographie mit den Nebenfächern Soziologie, Öffentliches Recht, Biologie und Physik an den Universitäten Mainz und Dijon (Frankreich) und wurde 2002 mit einer Arbeit zu ‚bewachten Wohnkomplexen‘ (*gated communities*) promoviert. Von 2002 bis 2008 war er wissenschaftlicher Assistent in Mainz und gleichzeitig Lehrbeauftragter am *Institut d'Études Politiques* (Paris/Nancy). 2008 wurde er mit einer politisch-geographischen Arbeit zur diskursiven Konstitution der Frankophonie habilitiert. 2009 nahm er den Ruf auf den Lehrstuhl für Geographie (Kulturgeographie) der FAU Erlangen-Nürnberg an. Gastdozenturen führten ihn unter anderem an die *Université Paris-Est* sowie an das *Oxford Internet Institute*.

SVENJA HAGENHOFF, Universitätsprofessorin Dr., Professur für
Buchwissenschaft, insbesondere E-Publishing und Digitale Märkte
Anschrift: Katholischer Kirchenplatz 9, 91054 Erlangen
E-Mail: svenja.hagenhoff@fau.de

Prof. Dr. Svenja Hagenhoff ist Professorin für Buchwissenschaft, insbesondere *E-Publishing* und Digitale Märkte. Sie folgte dem Ruf nach Erlangen 2011. Zuvor war sie Leiterin des Forschungsinstituts für Medienwirtschaft an der Fachhochschule St. Pölten und Forschungsgruppenleiterin an der Georg-August-Universität Göttingen. Forschungs- und Lehraufenthalte führten sie an Universitäten in Hildesheim, Los Angeles und Lübeck. Die Promotion (2001) und die Habilitation (2007) stammen aus Gebieten der Wirtschaftsinformatik und Betriebswirtschaftslehre. Ihre Kernarbeitsgebiete in Erlangen sind die Medienwirtschaft mit Fokus auf Schriftmedien sowie anwendungsbezogene Technologieaspekte in dieser Domäne.

KLAUS MEYER-WEGENER, Universitätsprofessor Dr., Inhaber des Lehrstuhls
Informatik 6 (Datenmanagement)
Anschrift: Martensstr. 3, 91058 Erlangen
E-Mail: klaus.meyer-wegener@fau.de

Prof. Dr. Klaus Meyer-Wegener ist seit Oktober 2001 Lehrstuhlinhaber am Department Informatik in der Technischen Fakultät der Universität Erlangen-Nürnberg. Er vertritt dort das Fachgebiet Datenmanagement. Von 1975 bis 1980 studierte er Informatik an der Technischen Hochschule Darmstadt und erwarb einen Abschluss als Diplom-Informatiker. Danach wurde er wissenschaftlicher Mitarbeiter am Fachbereich Informatik der Universität Kaiserslautern. 1986 wurde er dort zum Doktor-Ingenieur promoviert. Er blieb als Hochschulassistent zunächst an der Universität Kaiserslautern. Von Oktober 1987 bis Dezember 1988 war er beurlaubt, um als *Adjunct Research Professor* an der *Naval Postgraduate School* in Monterey, Kalifornien, zu arbeiten. 1991 konnte er seine Habilitation an der Universität Kaiserslautern abschließen und wechselte als Professor für Informatik (Datenbanksysteme) an die Universität Erlangen-Nürnberg. 1993 erhielt er einen Ruf auf eine Professur an der Fakultät Informatik der TU Dresden, wo er bis 2001 blieb.

HANS-ULRICH PROKOSCH, Universitätsprofessor Dr., Inhaber des
Lehrstuhls für Medizinische Informatik

Anschrift: Wetterkreuz 13, 91058 Erlangen

E-Mail: ulli.prokosch@imi.med.uni-erlangen.de

Prof. Dr. Hans-Ulrich Prokosch hat als Diplom-Mathematiker an der Universität Gießen in der Medizin zum Medizin-Informatiker promoviert und auch seine Habilitation erlangt. Von 1995 bis 2003 hatte er die Professur für Medizinische Informatik in Münster inne, bevor er 2003 den Ruf auf den Lehrstuhl für Medizinische Informatik an der Friedrich-Alexander-Universität Erlangen-Nürnberg erhielt. Er ist seitdem gleichzeitig als *Chief Information Officer* für die strategische Weiterentwicklung der IT-Landschaft des Universitätsklinikums Erlangen verantwortlich. Seine Forschungsschwerpunkte sind die Medizinische Prozessunterstützung, Telematikanwendungen im Gesundheitswesen und die Wiederverwendung von Daten aus der Elektronischen Krankenakte für die translationale medizinische Forschung. In jüngerer Zeit hat er sich schwerpunktmäßig auch mit den Möglichkeiten, Herausforderungen, aber auch Gefahren der mobilen medizinischen *App*-Anwendungen, der sensorbasierten Gewinnung von Vitaldaten sowie der *Quantified-Self*-Bewegung auseinandergesetzt.

Sachregister

- Adidas* 132, 134, 137
adidas Laufcomputer *Speed Cell* 132, 133
ahuman 26
akquisitorische Distribution 81
Algorithmen 8, 12, 14, 19, 23, 24, 26, 29, 30, 31, 32, 35, 39, 49, 70, 92, 105, 124
Algorithmendebatte 92
algorithmisch 49
algorithmische Evolution 29
Allgegenwart 30
Allmacht 30
Allmachtsfantasien 55, 58
Allwissenheit 13, 30, 34, 48, 51, 53, 59
Alphabet Holding 55
Amazon 19, 28, 34, 52, 53, 86, 87, 101, 109
An Inquiry into the Nature and Causes of the Wealth of Nations. 44
Analyse 8, 22, 39, 43, 44, 45, 49, 50, 52, 53, 56, 61, 66, 72, 74, 106, 107, 108, 109, 128
Android 64
AOK 126, 127, 128, 137
App 52, 55, 73, 117, 118, 119, 122, 123, 124, 125, 126, 127, 132, 133, 134, 138, 144
Apple 30, 87, 124, 125, 126, 134, 135, 138, 139
Apple Watch 126
ars artificialiter scribendi 80
Assoziationsregeln 106, 107
Aufklärung 7, 8, 9, 29, 35, 36, 41, 129
Aufrichtigkeit 102
Ausreißer 105, 109
Authentizität 25, 28
automatisierte Datenanalyse 51
Axiom 11
background check 23
Banken 26, 39
Barmer 126, 129, 130
BEIDOU 64
beMeister 132, 136
Benutzerprofil 15
Bevölkerung 42, 43, 46
Bevölkerungsdaten 43, 44, 45
Bevölkerungszahlen 43
Bibel 42, 60
Big Data 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 30, 34, 35, 36, 37, 39, 41, 42, 43, 48, 49, 50, 51, 53, 55, 58, 59, 60, 61, 62, 63, 74, 99, 100, 101, 105, 112, 113, 114, 115, 134, 137
big geospatial data 61
Big-Data-Prognosen 14
biometrische und soziometrische Analysen 22
Blasiertheit 16
Blog 21, 79
Blogger 19
Bluetooth 122, 123
Blutdruck 121
Blutzucker 121, 122, 138
BMW 118, 139
Body-Mass-Index 123
Börsenhandel 49
Buch 6, 8, 25, 41, 45, 80, 82, 89, 91, 96, 98
Buchdruck 78, 80, 81, 82, 85, 95

- Bücher 85, 89
 Buchexemplare 90
 Byte 100, 101
 CISCO 127, 137
 Cloud 16
 Cluster 106, 107, 108, 109
 Cluster-Analyse 106
 Companion App 122
 Complex-Event-Processing 99, 102, 105
 Computerchips 120
 consumer 84
 CPU (Central Processor Unit) 120
 crime preview 23
 critical data studies 61
 crowdsourced geodata 63, 65
 culturomics 21
 Daimler 126
 DAK 126
 Dash-System 28
 Data Matching 113, 115
 Data Mining 99, 100, 105, 113, 115
 Data Warehouses 99, 104
 Datafizierung 7, 14, 20, 27, 29, 35
 Daten- oder
 Überwachungskapitalisten 87
 Datenbanken 8, 23, 53, 99, 101, 103, 104, 115
 Datenbestände 100, 103, 104, 105, 106, 107, 109
 Datenflut 6, 15, 18, 20, 27, 29, 33, 39, 61
 Datenhändler 39
 Datenintegration 100, 114
 Datenkraken 19, 39
 Datenmacht 32
 Datenmenge 12
 Datenqualität 100, 112, 114
 Datensammlung 5, 23, 36, 42, 48, 49, 51, 52, 53, 54, 55, 56, 57
 Datenschutz 133
 Datenschutzmaßnahmen 133
 Datenschutzprobleme 133
 Datenschutzregulieren 128, 134
 Datenschwemme 13, 34, 51, 59
 Datenstromsysteme 105
 Datentransfer 48
 Datenvolumen 12
 Deep Learning 6, 11
 deontologisch 31
 Diffusion von Informationen 78
 Digital Humanities 14, 27, 34, 35
 digitale Demenz 27, 37
 digitale Geodaten 65, 66
 digitale Geographien 61, 72
 digitale Marktplätze 91
 digitale Paranoia 18
 digitale Revolution 5, 6
 digitale Schizophrenie 18
 digitales Grundgesetz 5
 digitales Nervensystem 15
 Digitalisierung 19, 39, 40, 41, 47, 48, 49, 64, 77, 79, 80, 83, 88, 91, 97, 121, 128, 136, 137, 139
 Diktatur 8, 12, 19, 20, 22, 36, 38
 Distribution 78, 81, 83, 85, 95
 Distributionsprozesse 83
 DKV 126
 drill-down 104
 Druckerpresse 81
 Dual-Use-Technologie 64
 Dubletten 113
 E-Book 19
 eGaIT Projekt 132
 E-Health 128, 137
 E-Health-Gesetz 128

- Elefanten 90
 elektronische Patientenakte 127,
 128, 129, 130
 empirische Sozialforschung 40, 41,
 45, 46, 47, 49, 60
 Ende der Theorie 13, 34, 51, 59
 Entgrenzung 79
 Entmündigung 29
 Entscheidungsbaum 108, 109
 Epigone 18
 Epigonentum 18
 Epistemologie 9, 10, 11
 Erkenntnisgegenstand 12
 Erkenntnisinteresse 15
 Erkenntnisobjekt 15
 Erkenntnisprozesse 10, 12
 Erkenntnissubjekt 15
 Erkenntnistheorie 9
Essay on the Principle of Population
 44
 Ethik 9, 29, 30, 31, 36
 Exabyte 100, 101
 Exhibitionismus 18, 19
 Exklusion 14, 71
Facebook 14, 16, 17, 20, 39, 52, 55, 87,
 88, 123
 Fakten 21, 25, 26, 37
FBI 134, 135, 138, 139
 Fernsehen 87, 112
 Filter 15, 36, 49, 77, 91, 92, 93, 95, 96
Filter Bubble 91
 filtern 91, 98, 103
 Finanzbereich 26
Fitness-App Dacadoo 126
FitnessSHIRT 123, 130, 131, 138
Fitnessstracker 122, 123, 126, 130
Flickr 67
 Fluten 22, 79, 85
 Fortschritt 57, 82, 122, 125
Fotocommunity 20
Fraunhofer Institut 123, 130, 131
GAFI 87
GALILEO 64
Gamification 125, 138
gatekeeper 77, 85, 87, 92
Gazeteers 65
 Geheimdienste 22, 39
 Geo-Daten 61, 72
 Geographische Informationssysteme
 (GIS) 66
 georeferenzierte Daten 63, 68, 72
Geoweb 67, 68, 69, 70, 71, 73, 74
 Gesellschaftsanalyse 19, 39, 40, 41,
 42, 43, 47, 49, 56, 57, 58
 Gesellschaftswissenschaft 39
Gesundheits-App 124
 Gesundheitsvorsorge 125
 Gewicht 79, 89, 90, 121, 123, 125
 Gigabyte 48, 101
GIS 66, 67, 69, 74, 75
 gläserner Bürger 5, 39
GLONASS 64
Google 14, 15, 19, 22, 37, 39, 49, 52,
 55, 64, 67, 70, 86, 87, 95, 98, 110,
 118, 124, 130, 131, 136
Google Earth 68
Google Flu 22
Google Health 130
Google Maps 67, 70
Google Web Tables 112
GPS 63, 64, 66, 68, 75, 118, 119
GPS-Tracks 68
grands récits 26
 Größe 89, 99, 100, 101, 121, 130
 Großkonzerne 26
 Gültigkeit 28, 102, 117, 133

- Gutenberg-Galaxis 81, 94
 Handpresse 82
HealthKit 124
Here 67
 Hermeneutik 9, 10, 11, 14
 hermeneutische Spirale 10
 hermeneutischer Zirkel 10, 15
 Herrschaftstechnologie 43
 Herzfrequenz 117, 122, 123
 Holzschliffpapier 82
homo connexus 16
homo rationalis 29
Hype 58, 62, 99, 124, 133
 Hypermnesie 8
 Hypothese 10, 12
 Hypothesenbildung 12
IBM 55, 103
idola foris 9
idola theatri 10
imaginatio 8
 implantierbare Sensoren 117
 Indikator 47
 Individualisierung 51
 Industrie 4.0 7, 99
 Informatik 3, 12, 41, 93, 99, 100, 143, 144
information overload 77
 Informations- und Medienselektion 80
 Informationsflut 77
 Informationskapitalismus 20, 22, 32
 Infrastrukturen 39, 40, 78, 81
 Inklusion 14
 intelligenten Maschinen 15, 28
 Internet der Dinge 7, 28, 40, 60, 119, 120, 121, 139
Internet of Things 120
 Internet-Blase 15
IoT-Geräte 121
 IP-Adressen 70, 119
iPhones 122
 Isolation 15, 27
 Karten 21, 61, 62, 63, 66, 67, 69, 70
 Kartographie 61, 62, 63, 64, 66, 67, 68, 69, 70, 71, 72, 74
 Kartographieforschung 69
 kategorischer Imperativ 30
 Kaufempfehlungen 49
 Kausalität 10, 12
Keyhole 67
 Klassifikation 107, 109
 Klassifikator 108
k-Means-Verfahren 106
Knowledge Discovery in Databases (*KDD*) 105
 Kommerzialisierung 70
 kommerzielle Akteure 41, 51, 52
 Kommunikation 22, 27, 28, 63, 77, 78, 79, 83, 84, 88, 89, 92, 94
 Kommunikatoren 77, 79, 84, 88
 Konstanzer Schule 11
 Kontrollverlust 79
 Konzerne 39, 41
 Korrelation 12
 Krankenversicherung 125, 134
Kritik der reinen Vernunft 10
 Kritische Kartographie 61, 63
 künstliche Intelligenz 6, 12, 35
Link 119
Linked Open Data 112
 Literaturkritik 93
 Literaturwissenschaft 93, 141
location based services 64
 Lukas-Evangelium 42, 43, 60
 Makulieren 81
Mapping Parties 68

- Map-Reduce* 99, 110, 111
 Marginalisierung 6, 70
 Markt- und Meinungsforschung 46,
 47, 50, 51
 Maschine 28, 32, 91, 97
Mass communication 84
 Massenkommunikation 79, 87, 89,
 96, 97
 Massenmarkt 81
Master of the Universe 26, 38
 Mathematik 8, 11, 41, 51
 Medien 16, 20, 21, 27, 34, 36, 63, 72,
 77, 78, 84, 87, 88, 91, 92, 94, 95,
 96, 97, 133, 135
 Medienökonomie 80, 93, 98
 medizinischen Daten 22
 Megabyte 100
memoria 7
Messenger-Dienste 52
 Metadaten 19, 22, 72, 74
 Metaphysik 30
 Microchip 131
Microsoft 55, 103
Minority Report 23
 Monitoringprogramme 129
 Multimedialität 102
 Muster 13, 20, 100, 105
MySpace 17
mySugr 122, 138
 Mythos 63, 92
 Nachsorge 129
 Nationalökonomie 44
*Natural and Political Observations
 Made upon the Bills of Mortality*
 44
 Navigationssysteme 20, 64
*Navstar Global Positional System
 (GPS)* 63, 64
Navteq/Here 67
 Nervensystem 15
Neues Testament 42
 Neurasthenie 15
 neuronal 11
nGram 21
 Normen 57
NoSQL 104, 110
Novartis 131
Novum Organum Scientiarum 9
 Nutzen 90, 99, 103, 113, 128, 129
 Nutzsinal 77
 Objektivitätsversprechen 50
OLAP 104
 Omnipotenz 30
 Omnipräsenz 30
 Omniszienz 30
Online Analytical Processing 104
 Ontologie 9, 51
 ontologischen Blase 27
Open Sky Doctrine 65
OpenData-Bewegung 63, 65, 68
OpenStreetMap 67, 68, 71, 72, 74, 75,
 76
Oracle 103
Ordnance Survey 68
Organon 9
OSM-Geodaten 68
 Ostrazismus 23
Outliers 109
 Papierproduktion 77, 82
 Parallelverarbeitung 100
 Parameter 121, 123
Personal Computer 119
 Personalisierte Information 51
 personalisierte Werbung 70
 Personalisierung 51, 70

- Personalisierung der Datenanalyse
 51
 persönliche Informationen 54
 Persönlichkeitsrechte 57
 Petabyte 101
 physische Distribution 81
Political Arithmetick 44
 Politikwissenschaft 45
 politische Arithmetik 44, 45
 politische Philosophie 44
 posthuman 7, 26, 28
 Postmoderne 25
 poststrukturalistisch 62
 prädiktive Algorithmen 23
Principles of Economics 44
 Privatsphäre 54, 57, 134
process and forget 105
producer 19, 84
 Produktions- und
 Distributionskosten 79
 Prognose 22, 31, 47
prosumer 84
 Protokolldaten 99
 Protokolle 19
 prozessuale Kartographie 69
 Psychologie 9, 45, 51, 79, 89, 96
public communication 84
 Publika (Massenpublika) 79
 Publikationsinfrastruktur 84
 Publikationswesen 19, 77, 79, 80
 Qualität der Daten 111
 Quantencomputer 12
Quantified Self 117, 122, 123, 124, 133,
 137
Quantified-Self-Community 122
 quantitative Sozialforschung 45, 47,
 58
 Radiogeräte 87

ratio 7
 Rationalität 7
 Rauschen 77, 90
Record Linkage 113, 115
 Reizüberflutung 15
report 104
 Rezeptionsästhetik 11
riot forecasting 14
roll-up 104
 Rotationsmaschine 78, 82
 Routenplanung 67
Runtastic 123
SANA Kliniken AG 127
 satellitengestützte
 Navigationssysteme 64
 Satellitenindustrie 65
 Scheinrealität 27
 Schlafphasenwecker 123
 Schnellpresse 82
Schufa Holding AG 54
science and technology studies 69
Second Life 18
 selbstfahrende Autos 31
 Selbstvermesser 117, 124, 125, 133, 139
 Selbstvermessung 117, 123, 124, 125,
 126, 130, 134
 selbstverschuldete Unmündigkeit 8
 Selektion 92
 Selektionsentscheidungen 92
 Selektionsmechanismen 91, 92
 Sensor 63, 65, 99, 117, 120, 122, 132
 Server 55, 120, 126, 133
simulacrum 26
 Sinnverlust 79
 Skript 49
Sleepscore 123
smart car 28
smart house 28

- smart phone* 28
smartwatch 117, 118, 119, 133, 139
social media-Dienste 67
social physics 62
Social-Media-Portale 52
software sorted geographies 72
 Solipsismus 15, 27
 soziale Netzwerke 16, 17, 28
 Sozialprognosen 23
 Soziologie 3, 9, 39, 40, 41, 44, 45, 51, 58, 87, 97, 141, 142
Sputnik 1 65
 SQL 110
Stanhope-Presse 82
 Statistik 45, 46, 90, 105, 110
 Stichprobe 47, 50
 Suchmaschine 52, 87, 91, 95
 Superrechner 7
 Syllogismus 10
Syantec 133, 138
Taxis-Galaxis 81, 94
Techniker Krankenkasse 127
 Teo Tronico 8, 38
 Terabyte 101
Thalia 91, 97
ThaliaNext 91
The Filter Bubble 15, 36
The Information Machine 78, 87, 92, 93, 94
TomTom 67
 Torwarter 14
 totale uberwachung 21
 Totalvermessung von Individuen 52
 Trainingsdaten 108
 Transparenzgesellschaft 56, 58
 Trenderaussagen 40
 Turingmaschine 28
 Twitter 14, 16, 21, 24, 34, 37
Twitterer 19
 Twitternachrichten 21
 uberwachung 6, 21, 22, 85, 92, 94
 Universalisierung 70
 Unmundigkeit 8
 Unternehmen 20, 23, 39, 41, 52, 53, 54, 56, 67, 124
user generated content 84
 utilitaristische Ethik 31
validity 102
value 103
 Varietat 48
variety 48, 102
velocity 48, 101
veracity 102
 Verifikation 11
 Verlags- und Pressewesen 83
 Vermassung 79
 vermessener Mensch 22, 117
 Vermessung 29, 35, 41, 42, 45, 46, 47, 57, 59
 Vermessung der Gesellschaft 41, 42, 57
 Vermessung der Welt 41
 Vermittlungsautoritaten 85
 Vernunft 7, 9, 32, 87, 92, 97, 124
 Versicherungen 39, 126
Verzeichnis der lieferbaren Bucher (VLB) 90
 virtuelle Welten 24, 26
Vivantes Kliniken 127
Vodafone 134
volume 48, 101
 Volumen 89, 101
volunteered geographic information 67, 70, 71, 73
 Vorhersage 22
 Vorsorge 117, 129, 133

Sachregister

- Wahrhaftigkeit 102
Walter-Rotationsmaschine 82
Wandel 4, 14, 18, 19, 24, 34, 35, 36,
37, 40, 41, 42, 48, 50, 51, 57, 58, 59,
87, 88, 94, 156
wearables 20, 22, 117
Web 2.0 64, 67, 68, 72
Weblogs 99
Weltbilder 63, 69, 70, 74, 156
Wertvorstellungen 57
WhatsApp 17
Where2Technologies 67
Whistleblower 22
Wikipedia 67, 68, 99, 100, 115
Wired 50, 61, 75
Wissenschaftstheorie 11
WLAN (*Wireless Local Area
Network*) 120
World-Wide Web (WWW) 119
Yottabyte 101
Youtube 16, 19
Zeitalter der Digitalisierung 80
Zeitalter der handwerklichen
Produktion 80
Zeitalter der industriellen
Produktion 80, 82
Zeitalter der mechanischen
Produktion 80
Zettabyte 12, 101

Namensregister

- Adorno, Theodor W. 7, 35
Anderson, Chris 13, 50, 51, 61
Aristoteles 9
Assmann, Jan 42
Baas, Jens 126, 128
Bacon, Francis 7, 10, 34, 43
Bagdikian, Ben H. 77, 78
Barnes, Trevor 62
Bauders, Marc 26
Baudrillard, Jean 25, 34
Bedford-Strohm, Heinrich 5
Berners-Lee, Tim 119, 120, 137
Biles, Simone 24
Blumenberg, Hans 9
Büttner, Sebastian M. 3, 19, 39, 142
Cook, Tim 135
Di Lorenzo, Giovanni 5
Dilthey, Wilhelm 10
Drösser, Christoph 8
Dworschak, Manfred 86
Federighi, Craig 124
Feynman, Richard 8
Foucault, Michel 43
Freiburg, Rudolf 10, 11, 33, 38, 141, 156
Freud, Sigmund 6
Gadamer, Hans-Georg 10
Gallup, George 46
Gehlen, Dirk von 54
Glasze, Georg 3, 22, 61, 63, 67, 70, 142
Gödel, Kurt 8
Graham, Steven 72
Grass, Günter 5
Graumann, Lutz 125
Graunt, John 44
Gröhe, Hermann 128
Gutenberg, Johannes 78, 81, 82, 94, 96
Habermas, Jürgen 5, 86, 95
Hagenhoff, Svenja 3, 19, 77, 78, 80, 91, 143
Hilbert, David 8
Hirsch, Eric D. 10, 11
Hofstetter, Yvonne 6, 8, 20, 48
Horkheimer, Max 7
Husserl, Edmund 11
Iser, Wolfgang 10, 21
Jones, Leslie 25
Junge, Barbara 53
Kant, Immanuel 8
Kehlmann, Daniel 41
Kirchner, Hanna 129
Kitchin, Rob 42, 61, 62, 74
Laplace, Pierre-Simon 8
Lauriault, Tracey P. 61, 62, 75
Leibniz, Gottfried Wilhelm 8
Luhmann, Niklas 88
Maas, Heiko 134
Mainzer, Klaus 8, 13
Malthus, Thomas Robert 44
Mayer-Schönberger, Viktor 7, 12
Menasse, Eva 5
Meyer-Wegener, Klaus 3, 12, 99, 143
Mills, Charles Wright 84
Morgenroth, Markus 14
Müller, Martin U. 95, 124
Orwell, George 6, 22
Pariser, Eli 15, 91
Parlin, Charles Coolidge 46

Namensregister

- Petrarca 85
Petty, William 44
Plato 8, 9, 27
Popper, Karl 11
Postman, Neil 87, 90, 98
Prokosch, Hans-Ulrich 3, 22, 117, 129,
136, 144
Rautenberg, Ursula 78, 89, 91, 95,
96
Reichert, Ramón 14, 20, 34, 35, 36,
37, 38, 48
Roosevelt, Franklin D. 46
Rosa, Hartmut 88
Schirmmacher, Frank 9, 40
Schleiermacher, Friedrich 10
Schmidt, Eric 55
Schulz, Martin 5
Shirky, Clay 83, 84
Simmel, Georg 16
Smith, Adam 44
Snowden, Edward 22
Spielberg, Steven 23
Spitzer, Manfred 27
Toffler, Alvin 77
Trump, Donald 24
Turtle, Sherry 16, 17
Voltaire 33, 37
Weber, Max 41
Weizenbaum, Joseph 28, 92
Welzer, Harald 8
Williams, Serena 24
Williams, Venus 24
Wolf, Gary 125
Woolf, Virginia 17
Zeh, Juli 5, 124
Zook, Matthew 67, 74
Zuckerberg, Mark 25
Zuse, Konrad 8

Im Rahmen der Erlanger Universitätstage sind bisher erschienen:

Band	Titel	ISBN	Preis
1986	Über den Schmerz. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe B, Bd. 18)	978-3-922135-52-8	10,50 €
1988	Identität. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe B, Bd. 20)	978-3-922135-59-3	10,50 €
1989	Sprache. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 54)	978-3-922135-68-5	10,50 €
1990	Zeit. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 59)	978-3-922135-76-0	11,00 €
1991	Über das Glück. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 64)	978-3-922135-82-1	11,00 €
1992	Natur. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe B, Bd. 23)	978-3-922135-95-1	11,00 €
1994	Altern und Alter. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 71)	978-3-930357-03-1	11,00 €
1995	Über den Zufall. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 75)	978-3-930357-09-3	11,00 €
1996	Die Würde des Menschen. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 80)	978-3-930357-20-8	11,00 €
1997	Gerechtigkeit. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 86)	978-3-930357-27-7	11,00 €
1998	Wertwandel und neue Subjektivität. Hrsg. v. H. Kößler (Erlanger Forschungen, Reihe A, Bd. 91)	978-3-930357-35-2	11,00 €
1999	Über das Experiment. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe B, Bd. 25)	978-3-930357-36-9	11,00 €
2000	Über die Macht. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe A, Bd. 96)	978-3-930357-45-1	12,00 €
2001	Über das Verhältnis der Geschlechter. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe A, Bd. 100)	978-3-930357-49-9	12,00 €
2002	Über die Folgen der Einheit. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe A, Bd. 104)	978-3-930357-57-4	12,00 €
2003	Über das Verhältnis von Kultur und Religion. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe A, Bd. 106)	978-3-930357-63-5	12,00 €

2004	Über die Grenzen von Wissenschaft und Forschung. Hrsg. v. G. Wanke (Erlanger Forschungen, Reihe B, Bd. 28)	978-3-930357-68-0	12,00 €
2005	Was Du ererbt von Deinen Vätern hast ... – Erbe, Erben, Vererben. Hrsg. v. H. Neuhaus (Erlanger Forschungen, Reihe A, Bd. 112)	978-3-930357-79-6	12,00 €
2006	Lauter Anfänge. Hrsg. v. H. Neuhaus (Erlanger Forschungen, Reihe A, Bd. 117)	978-3-930357-89-5	12,00 €
2007	Wasser. Hrsg. v. H. Neuhaus (Erlanger Forschungen, Reihe A, Bd. 119)	978-3-930357-98-7	12,00 €
2008	Licht. Hrsg. v. H. Neuhaus (Erlanger Forschungen, Reihe A, Bd. 120)	978-3-930357-95-6	12,00 €
2009	Das Ende. Hrsg. v. H. Neuhaus (Erlanger Forschungen, Reihe A, Band 122)	978-3-941871-03-8	12,50 €
2010 2011	Klimawandel. Mediengesellschaft. Hrsg. von Karl Möseneder (Erlanger Forschungen, Reihe A, Bd. 124/125)	978-3-941871-07-6	14,00 €
2012	Menschenrechte. Hrsg. von Karl Möseneder (Erlanger Forschungen, Reihe A, Bd. 128)	978-3-944057-04-0	12,50 €
2013	Weltbilder im Wandel. Hrsg. von Karl Möseneder (FAU Forschungen, Reihe A, Geisteswissenschaften, Band 2)	978-3-944057-19-4	12,50 €
2015	Träume. Hrsg. von Rudolf Freiburg (FAU Forschungen, Reihe A, Geisteswissenschaften, Band 4)	978-3-944057-32-3	12,50 €
2016	Unendlichkeit. Hrsg. von Rudolf Freiburg (FAU Forschungen, Reihe A, Geisteswissenschaften, Band 7)	978-3-944057-57-6	15,00 €

Bestellungen erbeten an:
FAU University Press
Universitätsstraße 4, 91054 Erlangen
Tel.: 09131/85-22155, Fax: 09131/85-29309
E-Mail: university-press@fau.de

ISBN 978-3-96147-046-4



9 783961 470464