



# 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](mailto: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

# **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.<sup>1</sup> 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?<sup>2</sup> 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.<sup>3</sup> 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,<sup>4</sup> Verfahren und Möglichkeiten der Datenana-

---

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

<sup>2</sup> 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).

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

<sup>4</sup> 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 herausfordert.<sup>5</sup> 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.<sup>6</sup> 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).

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

<sup>6</sup> 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,<sup>7</sup> 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.<sup>8</sup> 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

---

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

<sup>8</sup> 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.<sup>9</sup>

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.<sup>10</sup> Die Formel „Wissen ist Macht“,<sup>11</sup> 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.<sup>12</sup>

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

---

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

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

<sup>11</sup> 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.

<sup>12</sup> 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.<sup>13</sup> 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.<sup>14</sup>

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‘.<sup>15</sup>

---

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

<sup>14</sup> Vgl. ebd., 31-36.

<sup>15</sup> 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.<sup>16</sup> 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.<sup>17</sup>

### **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

---

<sup>16</sup> 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).

<sup>17</sup> 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.<sup>18</sup>

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.<sup>19</sup> 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.<sup>20</sup> 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

---

<sup>18</sup> 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).

<sup>19</sup> 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.

<sup>20</sup> 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.<sup>21</sup>

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.<sup>22</sup> 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.

---

<sup>21</sup> 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).

<sup>22</sup> 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.<sup>23</sup> In der Folge erleben wir einen großen Zuwachs des ‚Volumens‘ und der ‚Varietät‘ von Daten sowie der ‚Geschwindigkeit‘ der Datengenerierung und der Datenanalyse.<sup>24</sup> 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

---

<sup>23</sup> 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).

<sup>24</sup> 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.<sup>25</sup>

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.<sup>26</sup> 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

---

<sup>25</sup> Vgl. ebd., 80-99.

<sup>26</sup> 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.<sup>27</sup>

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.<sup>28</sup>

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

---

<sup>27</sup> 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.

<sup>28</sup> 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.<sup>29</sup> 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.<sup>30</sup>

Auch hier ließe sich vielleicht noch einwenden, dass dies im Prinzip nichts Neues ist. Auskunftseiten 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

---

<sup>29</sup> 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.

<sup>30</sup> 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.<sup>31</sup> 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

---

<sup>31</sup> 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.<sup>32</sup>

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-

---

<sup>32</sup> 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.<sup>33</sup>

---

<sup>33</sup> 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.
- Schirmmacher, 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](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.sozio-polis.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.sozio-polis.de/beobachten/wissenschaft/artikel/big-data-als-erwartungsraum/> [zuletzt besucht 15.01.2017].