

Jan-Felix Schrape

Big Data: Informatisierung der Gesellschaft 4.0

Abstract: *Der Beitrag gibt einen Überblick über vier Phasen der Informatisierung der Gesellschaft seit der Etablierung der ersten digitalen Computer, die von intensiven Diskussionen um die anwachsenden Daten- und Informationsfluten begleitet worden sind. Dabei zeigt sich, dass sich im aktuellen Diskurs um „Big Data“ viele der dystopischen und utopischen Erwartungen widerspiegeln, die bereits ab den 1960er-Jahren mit Computern und digitale Massendaten verknüpft wurden – von der Angst vor einem Totalverlust der Privatsphäre bis hin zu vielfältigen Hoffnungen auf Demokratisierung und Dezentralisierung. Angesichts dieser Kontinuitäten bleibt ein synthetisierender sozialwissenschaftlicher Blick unabdingbar, um den gegenwärtigen Wandel kritisch beobachten und Technikphantasmen entzaubern zu können.*

Der Begriff der Informatisierung wurde durch Simon Nora und Alain Minc (1978) bereits Ende der 1970er-Jahre eingeführt, um die zunehmende Durchdringung des menschlichen Lebens mit Kommunikations- und Informationstechnologien zu beschreiben. Anders als gängige Kompaktermini wie „Web 2.0“ oder „Big Data“ fasst „Informatisierung“ den Übergang zur Daten- und Informationsgesellschaft nicht als radikalen Bruch in kurzer Frist, sondern als vielschrittigen und *graduellen Transformationsprozess* (Dolata/Schrape 2013), der sich über mehrere Jahrzehnte erstreckt (vgl. für eine historische Einordnung: Baukrowitz u. a. 2001). Dieser Prozess lässt sich seit den ersten kommerziell vertriebenen Computern in den 1950er-Jahren in vier Phasen einteilen, die jeweils von intensiven Debatten um die anwachsenden Datenfluten und die damit verknüpften Chancen bzw. Risiken begleitet worden sind. Eine Reflexion dieser Diskursgeschichte kann dazu beitragen, gegenwärtige soziotechnische Trends wie auch die künftige Rolle der Sozialwissenschaften unaufgeregter zu taxieren.

Emergenz der „Informationsgesellschaft“ (1960/70er-Jahre)

Die ersten digitalen Computer wurden in enger Kooperation zwischen staatlichen und privatwirtschaftlichen Organisationen entwickelt, mussten durch ein Team von Spezialisten bedient werden und waren äußerst kostspielig. Der ab 1951 verkaufte raumfüllende Großrechner UNIVAC I etwa schlug mit bis zu 1,5 Mio. US-Dollar zu Buche (Ceruzzi 1998). Ab 1960 verbreiteten sich neben diesen Mainframes schrankgroße „Minicomputer“, die deutlich günstiger waren, von Einzelpersonen bedient werden konnten und durch neue Ein- und Ausgabeschnittstellen (z. B. Bildschirme) die Interaktion mit dem Nutzer vereinfachten. Vor allen Dingen in nordamerikanischen Universitäten boten campusöffentlich zugängliche Minicomputer einen idealen Nährboden für interdisziplinäre Projektgruppen (Levy 1984), wurden so im akademischen Milieu sichtbar und stießen einen ersten sozialwissenschaftlichen Diskurs zu den soziokulturellen Folgen der Computerisierung an.

Nach frühen literarischen Annäherungen an die Informatisierung wie Jorge L. Borges' „Bibliothek zu Babel“ ([1941] 1959) oder Philip K. Dicks' „The Minority Report“ (1956) spannte der Medienphilosoph Marshall McLuhan (1962: 31f.) Anfang der 1960er-Jahre das *Zukunftsszenario einer elektronisch vernetzten Gesellschaft* auf, die von totaler Überwachung geprägt ist: „Die neue elektronische Interdependenz redefiniert die Welt entlang der Vorstellung eines globalen Dorfes (*global village*). Statt zu einer großen alexandrinischen Bibliothek zu werden ist die Welt zu einem Computer geworden, einem elektronischen Gehirn [...]. Und gleichermaßen wie sich unsere Sinne externalisieren, dringt ‚Big Brother‘ in uns ein.“² Neuen Medientechnologien schrieb er vielfältige Effekte auf menschliche Wahrnehmungs- und Organisationsweisen zu, deren Bewertung aber je nach Beobachtungsperspektive variiere. Dementsprechend ging es McLuhan (ebd.: 254) zunächst einmal darum, überhaupt ein Bewusstsein für technikinduzierten Wandel und die damit verbundenen Kausalitäten zu schaffen.

Im selben Jahr prägte der US-amerikanische Ökonom Fritz Machlup (1962: 362) das Konzept der „knowledge economy“ und führte vor Augen, dass Ende der 1950er-Jahre bereits knapp 30 Prozent des Bruttosozialprodukts der USA in Wirtschaftssektoren generiert worden waren, die sich mit der Produktion von Wissen oder der Verarbeitung von Information beschäftigten. Zeitgleich veröffentlichte der japanische Anthropologe Tadao Umesao (1963) seine „Information Industry Theory“, in der er die These aufstellte, dass die Menschheit nach dem Agrar- und dem Industriezeitalter nun an der *Schwelle zur Informationsgesellschaft* stehe, in der Daten und Informationen zu zentralen Wirtschaftsfaktoren werden sollten. Offenkundig lag diese Idee in der Zeit, denn in den Folgejahren entwickelte eine Vielzahl von Sozialwissenschaftlern eigene Ideen zur postindustriellen und informationsbasierten Gesellschaft – so etwa auch Alain Touraine (1969) und Daniel Bell (1973), die wesentlich zur Popularisierung des Begriffs der Informationsgesellschaft jenseits der Wirtschaftsforschung beigetragen haben.

Die Diskussion um die gesellschaftlichen Folgen der digitalen Datenaggregation im engeren Sinne begann Ende der 1960er-Jahre: Der Jurist Arthur Miller (1967: 54) vermutete, dass die positiven Effekte der Computertechniken den Blick auf ihre negativen Folgen verstellten: Früher sei die *ubiquitäre Anhäufung personenbezogener Daten* schlicht nicht möglich gewesen, nun aber wachse das individuelle Datendossier kontinuierlich an. „Unser Erfolg oder Misserfolg im Leben wird mehr und mehr davon bestimmt sein, was ausgerichtet an den Entscheidungen anderer Eingang in unsere Dateien findet und inwieweit der damit befasste Programmierer in der Lage ist, diese Informationen zu evaluieren, zu verarbeiten und zu verknüpfen.“ Und ähnlich wie Juristen schon Ende des 19. Jahrhunderts befürchteten, dass das subjektive „Recht alleine gelassen zu werden“ durch die Mechanisierung der Gesellschaft zunehmend bedroht würde (Warren/Brandeis 1890: 193), vermutete auch der Techniksoziologie Lewis Mumford ([1967] 1977: 650), dass die künftige Zentralstellung des Computers in einer „Invasion der Privatsphäre“ und „Zerstörung der menschlichen Autonomie“ münden könnte.

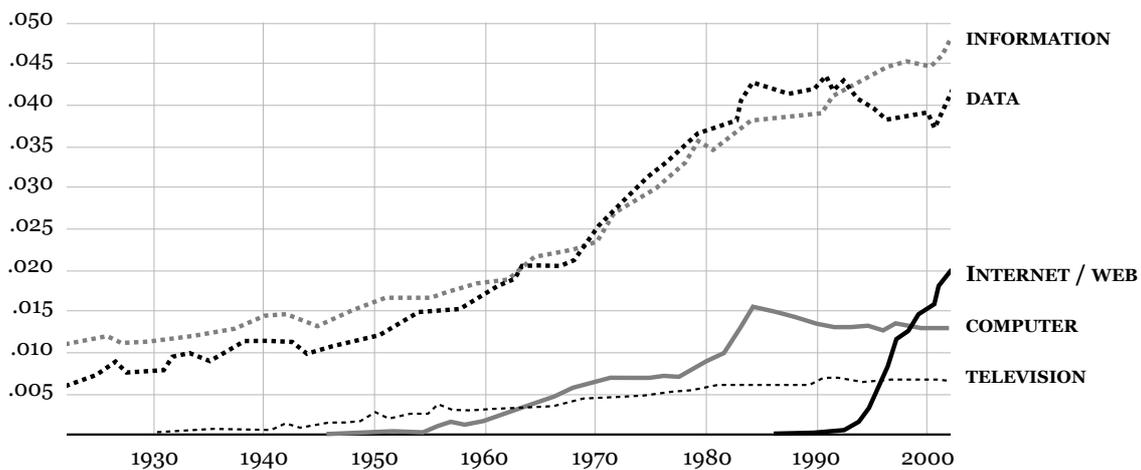
Der Journalist Ben Bagdikian (1971) sah das Unheil der elektronischen Medien hingegen weniger in diesem Überwachungspotential, sondern vielmehr in einer „Überschwemmung des Individuums mit Informationsfluten frei Haus“: Sie ermutigten „eher zur Reaktion auf aktuell wichtige Ereignisse als auf große Trends“ und erweckten nur „die Illusion umfassenden Wissens“ (zit. n. Spiegel o.V. 1972: 164). Auf vergleichbare Weise wies der Futurologe Alvin Toffler (1970: 350) auf die Gefahr der *kognitiven Überstimulation* durch neue Medien hin. Während Toffler dieses Risiko des „information overload“ indes auf zunehmend dezentralere Medienstrukturen zurückführte und befürchtete, dass immer weniger Rezipienten in der Lage sein könnten, sich eigenständig ein Gesamtbild der Lage zu erarbeiten, warnten einige

² Längere englischsprachige Zitate wurden zur Steigerung der Lesbarkeit durch den Autor übersetzt.

Informatikpioniere bereits in dieser Zeit vor beständig zentralisierteren Informationssystemen, die der Kontrolle weniger Anbieter überlassen würden (z.B. Steinbuch 1971).

Nach den Markterfolgen der ersten persönlichen Mikrocomputer und der Herausbildung einer frühen Amateur-Computing-Szene verbreiteten sich ab Mitte der 1970er-Jahre neben diesen Schreckensvisionen freilich auch *positive Propheetien zur informatisierten Gesellschaft*, die sich in kondensierter Form in dem TIME-Artikel „Pushbutton Power“ (TIME o.V. 1978: 46f.) wiederfinden: „Das Papierwirrwarr wird verschwinden, sobald Heiminformationssysteme die Aufgaben von Notizblöcken, Merkheften, Akten, Belegen und Küchenpinnwänden übernehmen. [...] Der Computer mag als entmenschlichender Einflussfaktor erscheinen, tatsächlich ist mithin genau das Gegenteil der Fall. [...] In vollem Gegensatz zu George Orwells düsterer Vorstellung von ‚1984‘ stimuliert der Computer die Verstandeskräfte [...] und führt die Menschheit auf eine höhere Stufe ihrer Existenz.“

Abbildung 1: Wort-Gebrauchsfrequenzen im Korpus von Google Books (englisch, in %)³



Datenquelle: Google Books Ngram Viewer (<https://books.google.com/ngrams>, Stand: 8/2016). Bücheranzahl/Jahr normalisiert, Singular- und Pluralformen

Die Informatisierung des Alltags (1980/90er-Jahre)

In den frühen 1980er-Jahren wurden erste Formen der elektronischen Vernetzung in der Alltagswelt sichtbar und die Idee der „Informationsgesellschaft“ kam im allgemeinen Diskurs an. Dies zeigt sich nicht zuletzt in den Gebrauchsfrequenzen der Begriffe „data“, „information“ und „computer“ (Abb. 1). Zwar wurden die entsprechenden Medieninnovationen von der Bevölkerung anfänglich eher zurückhaltend aufgenommen (Kellermann 1999); zugleich aber entwickelte sich auch hierzulande eine rege *Debatte über die Folgen neuer Informationstechnologien*, die medial eingehend reflektiert wurde: Der Bildschirmtext etwa sollte den Abschied von Druck bzw. Papier einläuten und den Bürgern neuartige Optionen bieten, „an

³ Die Abbildung bezieht sich auf den englischsprachigen Korpus von Google Books (eng_2012). Der Ngram Viewer zeichnet sich durch einen Überhang wissenschaftlicher Literatur aus, reflektiert keine Bedeutungsverschiebungen und erfährt nicht zuletzt aus diesen Gründen Kritik (Pettit 2016); dennoch bietet er eine erste Übersicht darüber, ab wann ein Ausdruck in einem Sprachraum Verwendung findet.

wesentlichen Entscheidungen unmittelbar teilzunehmen“ (Haefner 1984: 290); das Kabelfernsehen sollte je nach Beobachtungsperspektive „alle freie Information und Kommunikation [...] überwuchern“ (Ratzke 1975: 104) oder aber zur Schaffung basisdemokratischer Strukturen beitragen (Modick/Fischer 1984).

Parallel dazu konkretisierte sich in der Informatik die Diskussion um die erhöhte Produktion von Daten und Informationen im Zeitalter elektronischer Medien. Ithiel de Sola Pool (1983: 610) stellte in seiner Auswertung medienvermittelter Informationsflüsse in den USA zwischen 1960 und 1977 fest, dass Medien zum individuellen Abruf gegenüber Rundfunkmedien mehr und mehr an Einfluss gewannen. Hal Becker (1986) schätzte, dass auf einer Steintafel ein Zeichen pro Zoll, mit dem Buchdruck 500 Zeichen pro Zoll und auf Halbleiterspeichern im Jahr 2000 ca. 1.25x10¹¹ Bytes pro Zoll gespeichert werden könnten und überschrieb seinen Artikel vor diesem Hintergrund mit dem Titel: „Can users really absorb data at today's rates? Tomorrow's?“ Und Peter Denning (1990: 402) stellte schließlich die Frage nach adäquaten *Möglichkeiten der Automatisierung in der Datenstrukturierung* und sah einen Ausweg aus der genannten Überlastungssituation in Maschinen, die Datenmuster bzw. -korrelationen erkennen und vorhersagen können, ohne notwendigerweise deren Sinn verstehen zu müssen.

Auch Tim Berners-Lee (1989: 2) ging es in seinem Projektantrag zum World Wide Web zunächst nicht um ein weltumspannendes Informationssystem, sondern um eine typische Problemstellung großer Organisationen, deren hierarchisch angelegte Dokumentationssysteme den stetig komplexeren Verweisen zwischen Projekten, Menschen, Kennziffern, Text- und Bildmaterialien nicht mehr gerecht werden konnten: „Die aktuell beobachtbare Arbeitsstruktur einer Organisation ist ein mannigfaltig verknüpftes ‚Netz‘, dessen Interrelationen sich beständig fortentwickeln. [...] In vielen Fällen ist die entsprechende Information gespeichert worden; sie kann bloß nicht mehr aufgespürt werden.“ Das originäre Problem, das Berners-Lee angehen wollte, war also das der Verknüpfung von Daten, und er schuf mit dem *World Wide Web als eingängiges Organisationsprinzip* und ‚Interface‘ des Internets zugleich die Basis für ihre erneut beschleunigte Produktion. Diesen Wechselprozess zwischen wachsenden Datenmassen und erweiterten Austauschoptionen beschrieb John Mashey bereits 1998 als „Big Data and the Next Wave of InfraStress“. Ab 1996 war es günstiger, Daten digital statt auf Papier zu speichern, der Datenverkehr im Internet wuchs um 100 Prozent per anno und erstmals in der Geschichte wurde das Gros der Daten durch die Nutzer selbst generiert (Lyman/Varian 2000).

Einerseits wurde das World Wide Web daher rasch als freies Medium gepriesen, das eine „Verschiebung der Intelligenz vom Sender zum Empfänger“ (Negroponte 1995: 29) befördere, weil „es im Netz keine totalitären Instrumente mehr gibt, die Kontrolle über das Denken ausüben können“ (Bollmann/Heibach 1996: 473) und sich „die Rollentrennung von Kommunikator und Rezipient auflöst“ (Höflich 1996: 13). Andererseits monierte Stanislaw Lem (1996: 108), dass das Netz seine Tore jedem öffne, der „Daten stehlen und Geheimnisse aushorchen will“. Neil Postman (1999: 124) merkte an, dass nicht mehr die Verbreitung von Information das Problem der Gegenwart sei, sondern „wie man Information in Wissen verwandelt und wie Wissen in Erkenntnis“. Und Hans Magnus Enzensberger (2000: 96) verwarf seine „Prophezeiung von der emanzipatorischen Kraft“ neuer Medien: „Nicht jedem fällt etwas ein, nicht jeder hat etwas zu sagen [...]. Die viel beschriebene Interaktivität findet hier ihre Grenze. [...] Da kein Zentrum vorhanden ist, kann sich jeder einbilden, er befinde sich, wie die Spinne in ihrem Netz, im Mittelpunkt der Welt.“

Der Aufstieg der Datenkonzerne (2000er-Jahre)

Nach einer kurzen Phase der Desillusion gegenüber den neuen Informationstechnologien in Folge der geplatzten „Dotcom-Blase“ im Frühjahr 2000 führte der IT-Unternehmer Tim O’Reilly die Debatte in seinem ursprünglich primär internetökonomisch angelegten Artikel „What is Web 2.0“ (2005) wieder auf das Grundproblem der Digitalisierung – den Umgang mit Daten – zurück: „Datenbankmanagement gehört zu den Kernkompetenzen von Web 2.0-Unternehmen. Das führt zu einer Schlüsselfrage: Wem gehören die Daten? [...] Im Zeitalter des Internets lässt sich bereits eine Reihe an Fällen erkennen, in denen Kontrolle über den Datenbestand zur Kontrolle über den Markt und sehr hohen Erlösraten geführt hat.“ Ein Hauptkennzeichen des Web 2.0 bestand O’Reilly zufolge also (neben neuen Spielarten des Cloud Computings und Crowdsourcings) in einer so noch nie dagewesenen *Zentralstellung von Daten in der Geschäftswelt* und den damit verbundenen Fragen ihrer Kontrolle und Auswertung.

Tabelle 1: Jahresumsätze ausgewählter IT-Unternehmen in Mrd. US-Dollar

	2003	2006	2009	2012	2015
<i>Apple</i>	6,2	19,3	41,5	156,5	233,7
<i>Google (ab 2015: Alphabet)</i>	1,5	10,6	23,7	46,0	74,9
<i>Amazon</i>	5,3	10,7	24,5	61,1	107,0
<i>Facebook</i>	–	0,05	0,7	5,1	17,9
<i>Microsoft</i>	32,2	44,3	58,4	73,7	93,6

Datenquelle: Jahresberichte der genannten Unternehmen

In der Tat wurden bis Mitte der 2000er-Jahre die Grundsteine für den *Aufstieg global agierender Technologieunternehmen* gelegt, deren Kerngeschäft zu wesentlichen Teilen auf der Aggregation und Urbarmachung von Daten bzw. dem Verkauf entsprechender Geräte und Services basiert: Apple, Google, Amazon und Facebook stellen inzwischen die zentralen infrastrukturellen Grundlagen der Online-Welt bereit und prägen den Erfahrungsraum der meisten Mediennutzer auf elementare Weise mit (Dolata 2015). Die führenden Internetkonzerne (Tab. 1) machen inzwischen Jahresumsätze, die mit klassischen produzierenden Unternehmen wie Bosch (Umsatz 2015: 71 Mrd. Euro), Airbus (Umsatz 2015: 65 Mrd. Euro) oder Daimler (Umsatz 2015: 149 Mrd. Euro) vergleichbar sind.

Die öffentliche wie sozialwissenschaftliche Diskussion entfernte sich in dieser Phase indes zügig von dem eigentlichen Topos des Datenmanagements und der Begriff „Web 2.0“ avancierte schnell zum Synonym für eine erneute allumfassende Aufbruchsstimmung, die sich vor allen Dingen auf die kommunikationsermöglichenden und koordinationserleichternden Eigenschaften des Netzes bezog. Dabei lassen sich drei interagierende Veränderungserwartungen unterscheiden, welche letztlich alle auf einen *technikinduzierten Rückbau eingespielter sozialer Differenzierungen und Einflussasymmetrien* hinauslaufen (Dickel/Schräpe 2015):

- Eine Dezentralisierung in der Herstellung und Distribution von Inhalten bzw. „the implosion of production and consumption“ (Ritzer/Jurgenson 2010: 19) sowie das Aufbrechen starrer Rollenverteilungen zwischen Produzenten und Konsumenten;
- ein Relevanzverlust massenmedialer Anbieter (one-to-many) in der Nachrichtenverteilung zugunsten nutzerzentrierter Austauschprozesse im Social Web (many-to-many) – „transforming the news from a lecture to a conversation“ (Gillmor 2006: I);

- und „eine weltweite Demokratisierung [...] der Willens- und Bewusstseinsbildung“ (Sury 2008: 270) durch die Etablierung erweiterter Partizipationsmöglichkeiten sowie einen Relevanzverlust klassischer intermediärer Organisationen in der Politik.

Kritische Anmerkungen fanden in der Diskussion um das „Web 2.0“ hingegen lange kaum Gehör und zurückhaltende Stimmen, die wie Jürgen Habermas (2008: 161) auf die ambivalenten Folgen fragmentierter Publika für die politische Öffentlichkeit hinwiesen, wurden zunächst als rückwärtsgewandt eingestuft (z. B. in Bruns 2007). Seit 2010 freilich werden neben den utopischen auch wieder dystopische Erwartungen erörtert: Evgeny Morozov (2011: 118) sprach von der „Falle der Self-Empowerment-Diskurse – die kaum mehr darstellten als eine ideologische List, die die Unternehmensinteressen verschleiern“. Zygmunt Bauman (2013: 54f.) vermutete, die meisten Nutzer seien sich der Überwachung im Internet durchaus bewusst, allerdings werde heute „die alte Angst vor Entdeckung von der Freude darüber abgelöst, dass immer jemand da ist, der einen wahrnimmt“. Und Jaron Lanier (2012: 246) warnte vor dem unreflektierten Glauben an die schiere „Schwarmintelligenz“: „Ein Kollektiv auf Autopilot kann ein grausamer Idiot sein“.

Big Data und die „digitale Gesellschaft“ (2010er-Jahre)

Trotz dieser intensiven Debatte um das Internet und seine gesellschaftlichen Wirkungen rückten *Daten* an sich erst in den letzten Jahren in den Fokus der öffentlichen wie sozialwissenschaftlichen Aufmerksamkeit: 2008 lancierte der routinierte Internetvisionär Chris Anderson, der jüngst auch die Vorstellung einer neuerlichen industriellen Revolution durch 3D-Drucker popularisiert hat, seinen provokanten Text „Das Ende der Theorie. Die Datenschwemme macht wissenschaftliche Methoden obsolet“ (Anderson [2008] 2013); 2010 widmete die britische Wochenzeitung „The Economist“ digitalen Datenfluten eine erste große Sonderausgabe; 2011 beschrieb das Marktforschungsunternehmen Gartner „Big Data“ als die größte ökonomische Herausforderung unserer Tage (Genovese/Prentice 2011) und 2014 prognostizierte die International Data Corporation, dass das digitale Datenuniversum von 0,8 Zettabyte im Jahr 2009 auf 44 Zettabyte im Jahr 2020 anwachsen werde (EMC/IDC 2014).

Ähnlich wie im Falle des Labels „Web 2.0“, das die kommunikationserleichternden Eigenschaften der Onlinetechnologien betont hat, legt der Begriff „Big Data“ als semantischer Bezugspunkt heute in nahezu allen gesellschaftlichen Bereichen den Eindruck eines disruptiven Umbruchs nahe, auf den unmittelbar reagiert werden muss. Die Idee eines separaten Kommunikations- und Handlungsraums im Web weicht dabei der *Vorstellung einer „digitalen Gesellschaft“*, die mehr und mehr von einem omnipräsenten Internet der Dinge und Services durchdrungen wird. Big Data wird nun freilich nicht mehr nur als informationstechnisches, sondern als „kulturelles, technologisches und wissenschaftliches Phänomen“ zugleich gefasst (Boyd/Crawford 2012: 663). Dementsprechend reichen die Erwartungen um diesen Kompaktbegriff von dem „Versprechen der Allwissenheit“ (Geiselberger/Moorstedt 2013) über die ubiquitäre Vernetzung der Welt und Vorhersagbarkeit menschlichen Verhaltens bis hin zum Orwellschen Albtraum einer Totalüberwachung durch Staat oder Konzerne.

Dabei wird „Big Data“ als *Chiffre für verschiedenartige soziotechnische Trends* genutzt: Versicherungen erproben individuell anpassbare Preismodelle auf der Basis digitaler Selbstvermessung, die eine gerechtere Beitragsberechnung versprechen, aber auch die Gefahr einer Entsolidarisierung mit sich bringen. In der Verbrechensbekämpfung sollen – ähnlich, wie sich das Philip K. Dick (1956) vorgestellt hat – Delikte und Ausschreitungen auf der Basis aggregierter Be-

wegungsdaten vorhergesagt werden. Im Verteidigungsbereich gilt die automatisierte Auswertung von Massendaten als gängiges „Handwerkszeug für moderne militärische Aufklärung und Lageanalyse“ (Hofstetter 2014: 13). In Wahlkämpfen rekrutieren Parteien eigene Datenanalysten, um personenscharfe Wählerprofile zu erstellen. „Big Data Marketing“ verspricht eine umfassende Erfassung von Konsumentenerfahrungen, um Kunden ein vollpersonalisiertes Einkaufserlebnis zu bieten. In der Medizin findet die Auswertung von Massendaten nicht mehr nur im Gesundheitsmanagement, sondern ebenso in der Früherkennung von Epidemien Anwendung. Und auch in anderen Forschungsbereichen führt die Verfügbarkeit wachsender Datenmengen (z. B. unter dem Stichwort „Digital Humanities“) zur Erprobung neuer Erhebungs- und Auswertungsmethoden (vgl. als Überblick Reichert 2014; Richter 2015).

Am augenfälligsten tritt die Zwiespältigkeit von Big Data im öffentlichen Diskurs bislang auf dem Feld der Mobile Devices (Smartphones, Tablets, Wearables) hervor: Einerseits werden die persönlichen medialen Erfahrungsräume durch die Alltagsintegration dieser Geräte und die situative Echtzeitauswertung passender Datenbestände zwar beständig erweitert. Andererseits werden diese Erfahrungsräume allerdings weitaus eindeutiger als im klassischen Web durch die Architekturen der entsprechenden Betriebssysteme und Anwendungen vorgeprägt, die sich auch abgesehen von den „Allgemeinen Nutzungsbedingungen“ durch markante soziotechnische Strukturierungsmuster auszeichnen. Zudem führt die Anbieterkonzentration auf dem Feld der Mobile Devices zu einer bis dato weitgehend unregulierten und mediengeschichtlich singulären *Konzentration privatwirtschaftlicher Verfügungsmacht* über Interaktionsdaten und Infrastrukturen, mit der letztlich eine Privatisierung des Schutzes persönlicher Daten einhergeht (Schrape 2016b; Dolata/Schrape 2014). Die damit verknüpfte Ambivalenz kommt nirgends besser zum Ausdruck als in einem Zitat von Eric Schmidt (2010: 15. Min.), der Executive Chairman der Alphabet Holding (vormals Google) ist: „Je mehr Informationen wir über dich haben – mit deiner Erlaubnis – umso mehr können wir unsere Suchergebnisse verbessern. [...] Dazu brauchst Du überhaupt nichts mehr tippen, denn wir wissen wo du bist – mit deiner Erlaubnis. Wir wissen, wo du warst – mit deiner Erlaubnis. Wir können mehr oder minder errahnen, worüber du nachdenkst.“

Informatisierung und Soziologie

In der Langzeitbetrachtung wird deutlich, dass sich im aktuellen Diskurs um die gesellschaftlichen Folgen von „Big Data“ viele der dystopischen und utopischen Erwartungen widerspiegeln, die bereits seit den 1960er-Jahren an digitale Massendaten und Datenfluten geknüpft worden sind – von der Gefahr eines „information overload“ über die Angst vor einer Invasion der Privatsphäre bis hin zu vielfältigen positiven Visionen, die entlang von Stichworten wie Dezentralisierung, Demokratisierung oder Emanzipation immer wieder die allgemeine und sozialwissenschaftliche Öffentlichkeit durchkreuz(t)en.

Diese Kontinuitäten rechtfertigen es, die gegenwärtigen soziotechnischen Trends und die damit einhergehende Diskussion um das Chiffre „Big Data“ als *Informatisierung der Gesellschaft 4.0* zu beschreiben (Tab. 2): (1) Nach der Emergenz der Idee einer Informationsgesellschaft im Kontext der akademischen Urbarmachung von Computern in den 1960er- und 1970er-Jahren (*1. Phase*), einer ersten Welle der Informatisierung der alltäglichen Lebenswelt in den 1980er- und 1990er-Jahren (*2. Phase*) sowie dem Aufstieg der Datenkonzerne und der zeitgleichen diskursiven Betonung der ermöglichenden Eigenschaften der Plattformen im Web 2.0 in den 2000er-Jahren (*3. Phase*) rücken seit 2010 explizit Massendaten sowie die Fragen nach ihrer Kontrolle und Auswertung in den Mittelpunkt der gesellschaftsweiten Diskussion um neue Informations- und Kommunikationstechnologien (*4. Phase*).

Tabelle 2: Phasen der Informatisierung der Gesellschaft

1. Phase 1960/70er-Jahre	2. Phase 1980/90er-Jahre	3. Phase 2000er-Jahre	4. Phase ca. ab 2010
Emergenz der „Informationsgesellschaft“ als Begriff und Idee	Beginnende Informatisierung der alltäglichen Lebenswelt	Aufstieg der Datenunternehmen und „Web 2.0“-Diskurs	„Big Data“: Soziale Vergegenwärtigung der Informatisierung

(Quelle: Eigene Überlegungen)

Gerade im gegenwärtigen Stadium, in der sich die Gesellschaft ihrer allumfassenden Informatisierung bewusst wird, kommt der *Soziologie als Gesellschaftswissenschaft im klassischen Sinne* eine zentrale Rolle zu, denn anders als in der Industrieforschung und in anwendungsbezogenen Disziplinen geht es ihr nicht um kurzfristige Vorhersagen oder eine möglichst zeitnahe Verwertbarkeit ihrer Forschungsergebnisse, sondern um die Einordnung gegenwärtiger Dynamiken in mittel- bzw. langfristige gesellschaftliche Entwicklungsprozesse, die Identifikation von Kontinuitäten und Brüchen sowie die Dekonstruktion eingeschliffener Beschreibungsweisen (Schrape 2016a). Pointiert formuliert: Ein anwendungsorientierter Forscher will mittels Algorithmen ungeordnete Daten in Form bringen, also aus rohen Daten *strukturierte Informationen* gewinnen, die praktisch verwertbar sind und einen unmittelbaren Neuigkeitswert bieten; ein sozialtheoretisch informierter Gesellschaftswissenschaftler hingegen will *übertragbares Wissen* generieren und generalisierbare Erklärungen erarbeiten (Tab. 3). „Datenwissenschaftler“ und Soziologen arbeiten dementsprechend auf zwei unterschiedlichen Ebenen der Komplexitätsreduktion in der gesellschaftlichen Wirklichkeitsbeschreibung.

Tabelle 3: Verschiedene Forschungskulturen im Umgang mit Daten (idealtypische Darstellung)

	Gesellschaftswissenschaften	Ingenieurwissenschaften / Industrieforschung
Ziele	Erklärung mittel- und langfristiger Entwicklungsprozesse, Abstraktion und generalisierbares Wissen	Kurzfristige Vorhersagen praxisrelevanter Dynamiken; Entwicklung rasch anwendbarer Ordnungsalgorithmen/-kategorien
Fokus	Theoriebildung	Praktische Verwertbarkeit
Grundüberzeugung	Daten und Erhebungsergebnisse bergen prinzipiell Verzerrungspotential	Daten und Erhebungsergebnisse orientieren sich an Grundrealitäten
Praktiken	<ul style="list-style-type: none"> • Lange Publikationsintervalle • Lange Journalartikel / Bücher • Dezidierte theoretische Einordnung • Weit zurückreichende Literaturbasis • Geringer Forschungskonsens 	<ul style="list-style-type: none"> • Kurze Publikationsintervalle • Kurze Konferenzpapiere • Keine theoretische Einordnung • Nur aktuelle Literatur (1–5 Jahre alt) • Hoher Forschungskonsens

Quelle: Orientiert an McFarland et al. 2016: 25

Aus ihrer kritisch-distanzierten und synthetisierenden Beobachtungsperspektive kann die Gesellschaftswissenschaft in der aktuellen Diskussion ein notwendiges Gegengewicht zu der verbreiteten Überzeugung bieten, dass blanke Zahlen bei einer hinreichend breiten Datenbasis für sich sprechen (vgl. Anderson 2013): Zum einen resultiert aus der erhöhten Verfügbarkeit quantifizierter Kennziffern die Gefahr, sich einzig und alleine auf ebensolche Daten zu verlassen und diese ohne weitere Kontextualisierung als unmittelbare Entscheidungsgrundlage zu verwenden. Zum anderen bleiben selbst die detailreichsten Datenmuster das Objekt subjektiver und kontextgebundener Interpretation, zumal sich mittels Algorithmen lediglich ermitteln lässt, welche *Korrelationen* die jeweiligen Daten aufweisen, nicht aber die dahinterliegenden *Kausalitäten*. Je größer indes die verfügbare Datenbasis ist, desto größer wird auch die Wahrscheinlichkeit, sich zufällig ähnelnde Datenreihen zu finden und daraus Zusammenhänge abzuleiten, die realiter gar nicht vorhanden sind, so z. B. eine Korrelation zwischen nichtkommerziellen Raketenstarts weltweit und erfolgreichen Soziologie-Doktoranden in den USA.⁴

Neben der Diskussion der methodischen Fallstricke und Potenziale von „Big Data“ sollte eine weitere Aufgabe der Soziologie überdies darin bestehen, die *gesellschaftlichen Effekte* der entsprechenden utopischen bzw. dystopischen Zukunftsbilder zu eruieren sowie übergreifende Muster und Kontinuitäten herauszuarbeiten, um so „empirisch zu zeigen, wie [...] der durch digitale Medien eröffnete Möglichkeitsraum in Experimente mit konkreten Praktiken und schließlich wieder in stabile – wenn auch möglicherweise andere – Erwartungsstrukturen überführt wird“ (Süssenguth 2015: 116). Denn obgleich die vorgestellten Prophetien in ihrer Radikalität bis dato empirisch kaum eingelöst worden sind, erfüllen sie in ihren jeweiligen Anwendungskontexten elementare kommunikative Funktionen: Sie erzeugen Aufmerksamkeit für neue technologische Möglichkeiten und die damit verbundenen Chancen bzw. Risiken, kanalisieren den Diskurs in eine bestimmte Richtung, eröffnen neue Geschäftsmöglichkeiten und dienen nicht zuletzt als Legitimationsgrundlage in wirtschaftlichen, politischen und wissenschaftlichen Entscheidungsprozessen. In dieser Beziehung bieten die Erwartungen, die an das Chiffre „Big Data“ geknüpft werden, sozialwissenschaftlichen Beobachtern durchaus Orientierung – zwar nicht darüber, was die Zukunft bringen wird, aber umso mehr darüber, durch welche „Veränderungsbedürfnisse, Unsicherheiten und Konfliktlinien die aktuelle gesellschaftliche Kommunikation“ geprägt ist (Dickel/Schrape 2015: 459).

Der von Chris Anderson ([2008] 2013) für seine bekannte These vom nahen Ende der Theorie als Kronzeuge aufgeführte Statistiker George E. P. Box (1976: 798) beschließt seinen Artikel zu „Science and Statistics“ übrigens wie folgt: „In der Soziologie, Psychologie, Erziehungswissenschaft und sogar [...] im Ingenieurwesen misstrauen Forscher, die nicht selbst Statistiker sind, fälschlicherweise ihrem eigenen Menschenverstand und adaptieren unangemessene Verfahrensweisen, die von Mathematikern ohne wissenschaftliche Erfahrungen entworfen worden sind. [...] Der Bedarf an kompetenten Statistikern, die [belastbare] Fakten aus Datenanalysen herausarbeiten [...] und die notwendigen neuen Theorien und Techniken dafür entwickeln können, wird daher weiter steigen“. Insofern ist Andersons Text das beste Beispiel dafür, dass Aussagen wie Daten stets der Kontextualisierung benötigen, um nicht missverstanden zu werden. Das genuine Kernanliegen soziologischer Forschung – die Vermeidung von Phantasiewissen und die *Entzauberung von Beobachtungsmethoden* (Elias 2006: 32, 100) – hat in der Ära der „Informatisierung 4.0“ mithin nichts an Dringlichkeit verloren.

⁴ Dieses wie viele weitere Beispiele für unterhaltsame Scheinkorrelationen findet sich auf folgender Website: <http://tylervigen.com/spurious-correlations> (Abruf: 31.10.2016).

Literatur

- Anderson, Chris (2013): Das Ende der Theorie. Die Datenschwemme macht wissenschaftliche Methoden obsolet [2008]. In: Geiselberger, Heinrich/Moorstedt, Tobias (Hg.): Big Data: Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp, S. 120-130.
- Bagdikian, Ben H. (1971): The Information Machines. New York: Harper & Row.
- Baukrowitz, Andrea; Boes, Andreas; Schmiede, Rudi (2001): Die Entwicklung der Arbeit aus der Perspektive ihrer Informatisierung. In: Matuschek, Ingo; Henninger, Annette; Kleemann, Frank (Hg.): Neue Medien im Arbeitsalltag. Wiesbaden: Westdeutscher Verlag, S. 219-235.
- Bauman, Zygmunt (2013): Das Ende der Anonymität. Was Drohnen und Facebook verbindet. In: Blätter für deutsche und internationale Politik 10/2013, S. 51-62.
- Becker, Hal B. (1986): Can users really absorb data at today's rates? Tomorrow's? In: Data Communications 15 (8), S. 177-193.
- Bell, Daniel (1973): The Coming of Post-Industrial Society: A Venture in Social Forecasting. New York: Basic Books.
- Berners-Lee, Tim (1989): Information Management: A Proposal. Arbeitspapier. Bern: CERN.
- Bollmann, Stefan; Heibach, Christiane (Hg.) (1996): Kursbuch Internet. Mannheim: Rowohlt.
- Borges, Jorge L. (1959): Die Bibliothek von Babel [1941]. In: Ders.: *Labyrinthe. Erzählungen*, München: Hanser, S. 187-197.
- Box, George E. P. (1976): Science and Statistics. In: Journal of the American Statistical Association 71 (356), S. 791-799.
- Boyd, Dana; Crawford, Kate (2012): Critical Questions for Big Data. In: Information, Communication & Society 15 (5), S. 662-679.
- Bruns, Axel (2007): Habermas and/against the Internet. In: Snurblog vom 18. Dezember 2007. URL: <http://snurb.info/node/621> (Abruf: 31.10.2016).
- Ceruzzi, Paul E. (1998): A History of Modern Computing. Cambridge: MIT Press.
- Denning, Peter J. (1990): Saving all the Bits. In: American Scientist 78, S. 402-405.
- Dick, Philip K. (1956): The Minority Report. In: Fantastic Universe 1/1956, S. 4-36.
- Dickel, Sascha; Schrape, Jan-Felix (2015) Dezentralisierung, Demokratisierung, Emanzipation. Zur Architektur des digitalen Technikutopismus. In: Leviathan 43 (3), S. 442-463.
- Dolata, Ulrich (2015): Volatile Monopole. Konzentration, Konkurrenz und Innovationsstrategien der Internetkonzerne. In: Berliner Journal für Soziologie 24 (4), S. 505-529.
- Dolata, Ulrich; Schrape, Jan-Felix (2013): Medien in Transformation. Radikaler Wandel als schrittweise Rekonfiguration. In: Dies. (Hg.): Internet, Mobile Devices und die Transformation der Medien. Berlin: Edition Sigma, S. 9-37.
- Dolata, Ulrich; Schrape, Jan-Felix (2014): App Economy: Demokratisierung des Software-Marktes? In: Technikfolgenabschätzung – Theorie und Praxis 23 (2), S. 76-80.
- Elias, Norbert (2006): Was ist Soziologie? Frankfurt am Main: Suhrkamp.
- EMC Corporation/IDC (2014): The Digital Universe of Opportunities. Framingham: IDC.
- Enzensberger, Hans Magnus (2000): Das digitale Evangelium. In: Der Spiegel 2/2000, S. 92-101.
- Genovese, Yvonne; Prentice, Stephen (2011): Pattern-Based Strategy: Getting Value From Big Data. Stamford: Gartner Inc.
- Habermas, Jürgen (2008): Ach, Europa. Kleine Politische Schriften XI. Frankfurt am Main: Suhrkamp.
- Haefner, Klaus (1984): Mensch und Computer im Jahre 2000. Basel: Birkhäuser.
- Höflich, Joachim (1996): Technisch vermittelte interpersonale Kommunikation. Opladen: Westdeutscher Verlag.
- Hofstetter, Yvonne (2014): Sie wissen alles. Wie intelligente Maschinen in unser Leben eindringen und warum wir für unsere Freiheit kämpfen müssen. München: C. Bertelsmann.
- Lanier, Jaron (2012): Digitaler Maoismus. In: Kemper, Peter; Mentzer, Alf; Tillmanns, Julika (Hg.): Wirklichkeit 2.0. Stuttgart: Reclam, S. 243-247.
- Kellerman, Aharon (1999): Leading Nations in the Adoption of Communications Media 1975–1995. In: Urban Geography 20 (4), 377-389.
- Lem, Stanislaw (1996): Zu Tode informiert. In: Der Spiegel 11/1996, S. 108-109.
- Levy, Steven (1984): Hackers: Heroes of the Computer Revolution. Garden City: Anchor Press.
- Lyman, Peter; Varian, Hal R. (2000): Reprint: How Much Information? In: Journal of Electronic Publishing 6 (2). URL: <http://dx.doi.org/10.3998/3336451.0006.204> (Abruf: 31.10.2016).
- Machlup, Fritz (1962): The Production and Distribution of Knowledge in the United States. Princeton: Princeton University Press.

- Mashey, John R. (1998): Big Data and the Next Wave of InfraStress. Präsentation, 25.4.1998. URL: http://static.usenix.org/event/usenix99/invited_talks/mashey.pdf (Abruf: 31.10.2016).
- McFarland, Daniel A.; Lewis, Kevin; Goldberg, Amir (2016): Sociology in the Era of Big Data: The Ascent of Forensic Social Science. In: *The American Sociologist* 47 (1), S. 12-35.
- McLuhan, Marshall (1962): *The Gutenberg Galaxy*. Toronto: University of Toronto Press.
- Miller, Arthur R. (1967): The National Data Center and Personal Privacy. In: *The Atlantic* 11/1967, S. 53-57.
- Modick, Klaus; Fischer, Matthias J. (1984): *Kabelhafte Perspektiven*. Hamburg: Nautilus.
- Morozov, Evgeny (2011): Back to the Roots: Cyberspace als öffentlicher Raum. In: *Blätter für deutsche und internationale Politik* 9/2011, S. 114-120.
- Mumford, Lewis (1967/1977): *Mythos der Maschine. Kultur, Technik und Macht*. Frankfurt am Main: Fischer.
- Negroponte, Nicholas (1995): *Being Digital*. New York: Knopf.
- Nora, Simon; Minc, Alain (1978): *L'Informatisation de la Société*. Paris: La Documentation Française.
- O'Reilly, Tim (2005): What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. In: *O'Reilly Network* vom 30.9.2005. URL: <http://oreilly.com/pub/a/web2/archive/what-is-web-20.html> (Abruf: 31.10.2016).
- Pettit, Michael (2016): Historical Time in the Age of Big Data: Cultural Psychology, Historical Change, and the Google Books Ngram Viewer. In: *History of Psychology* 19 (2), S. 141-153.
- Pool, Ithiel de Sola (1983): Tracking the Flow of Information. In: *Science* 221 (4611), S. 609-613.
- Postman, Neil (1999): *Die zweite Aufklärung. Vom 18. ins 21. Jahrhundert*. Berlin: BvT.
- Ratzke, Dietrich (1975): *Netzwerk der Macht*. Frankfurt am Main: Societats.
- Reichert, Ramón (Hg.) (2014): *Big Data. Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie*. Bielefeld: Transcript.
- Richter, Philipp (Hg.) (2015): *Privatheit, Öffentlichkeit und demokratische Willensbildung in Zeiten von Big Data*. Baden-Baden: Nomos.
- Ritzer, George; Jurgenson, Nathan (2010): Production, Consumption, Prosumption. The Nature of Capitalism in the Age of the Digital ‚Prosumer‘. In: *Journal of Consumer Culture* 10 (1), S. 13-36.
- Schmidt, Eric; Google Inc. (2010): Eric Schmidt at Washington Ideas Forum 2010. Videodokument. URL: <https://www.youtube.com/watch?v=CeQsPSaitL0> (Abruf: 31.10.2016).
- Schrape, Jan-Felix (2016a): Soziologie als „Marke“. Kernkompetenz, gesellschaftlicher Nutzen, Vermittlungswege. In: *Soziologie* 45 (3), S. 279-293.
- Schrape, Jan-Felix (2016b): Social Media, Mass Media and the ‚Public Sphere‘. Differentiation, Complementarity and Co-existence. *Research Contributions to Organizational Sociology and Innovation Studies 2016-01*. Stuttgart: University of Stuttgart.
- Spiegel (ohne Verfasser) (1972): „Die Elektronen haben keine Moral“. In: *Der Spiegel* 26 (17), S. 158-164.
- Steinbuch, Karl (1971): *Mensch, Technik, Zukunft: Basiswissen für die Probleme von morgen*. München: Deutsche Verlags-Anstalt.
- Süssenguth, Florian (2015): Die Organisation des digitalen Wandels. In: Ders. (Hg.): *Die Gesellschaft der Daten*. Bielefeld: Transcript, S. 93-121.
- Sury, Ursula (2008): Internet(o)kratie. In: *Informatik Spektrum* 31 (3), S. 270-271.
- Time Magazine (ohne Verfasser) (1978): Living: Pushbutton Power. In: *Time Magazine* 111 (8), S. 46-49.
- Toffler, Alvin (1970): *Future Shock*. New York: Bantam Books.
- Touraine, Amain (1969): *La société post-industrielle. Naissance d'une société*. Paris: Denoël.
- Umesao, Tadao (1963): Information Industry Theory: Dawn of the Coming Era of the Ectodermal Industry. In: *Hoso Asahi*, S. 4-17.
- Warren, Samuel D.; Brandeis, Louis D. (1890): The Right to Privacy. In: *Harvard Law Review* 4 (5), S. 193-220.