Ramón Reichert (Hg.)

# BIG DATA

Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie

transcript

Digitale Gesellschaft

# Die Konstruktion von Big Data in der Theorie

Tom Boellstorff

#### EINLEITUNG

»Wegbereitende viktorianische historische Darstellungen der britischen Presse, die zwischen 1850 und 1887 erschienen [...], knieten am Altar der Technik, als sie die Anfangsbuchstaben von Newspaper Press groß schrieben [...] Klügere und skeptischere [spätere Autoren] hörten auf, Newspaper Press mit Großbuchstaben zu schreiben. Vielleicht sollten wir das auch in Bezug auf das Internet tun.« (Curran, 2012, S. 34, 60)

Wir leben in einer Zeit, in der »Big Data« die Gesellschaft verwandeln werden. So behauptet es zumindest der Hype.

Wie jeder Mythos ist auch das derzeitige Trara um Big Data überzogen, enthält aber ein Körnchen Wahrheit. Dagegen ließe sich relativ leicht einwenden: Es gibt kein einheitliches Phänomen »Big Data«. Es gibt keine einzigartige Form von »Gesellschaft«. Und so weiter. Doch Big Data haben ganz reale Auswirkungen und verdienen unsere ganze Aufmerksamkeit.

Vielleicht haben Sie bemerkt, dass ich »Big Data« nicht mit Großbuchstaben schreibe. Dies ist mein erster theoretischer Einstieg in einen theoretischen Aufsatz, der so tief von der Geschichte geprägt ist. Wie James Curran oben festgestellt hat, gibt es einen historischen Background für Praktiken der Großschreibung in Bezug auf die Technik – und es hat einen gewissen Wert, Big Data als Gattungsnamen zu behandeln, der für eine Umgestaltung anfällig ist. Mein Ziel ist es, vier Möglichkeiten für eine Umgestaltung zu erkunden, die zur »großen Theorie« beitragen können, die wir im Hinblick auf Big Data so dringend benötigen. Ich gehe bei diesen Erkundungen von meinem primären fachlichen Background, der Anthropologie, aus, ziehe dazu aber auch andere Gebiete zu Rate: die Etymologie ebenso wie die Ethnografie, die Philosophie ebenso wie die Wissenschaftsforschung.

Dieser Aufsatz hat seinen Ursprung im Interesse an Begriffen, das mich bei meiner Arbeit über digitale Kultur und auch bei meiner früheren Forschung über die Sexualität in Indonesien stets begleitet hat (z.B. Boellstorff, 2005, 2007, 2008,

107

2012). Es gibt jedoch auch ein näher liegendes Motiv: die Affäre um Edward Snowden 2013. Seine Enthüllungen über das Ausspähen von »Metadaten« durch die National Security Agency entfachten eine lebhafte Debatte über Big Data, Überwachung und das Gemeinwohl - eine Debatte, die zu dem Zeitpunkt, da ich dies schreibe, noch nicht verstummt ist (beim ersten Entwurf dieses Aufsatzes saß Snowden noch am Moskauer Flughafen fest). An verschiedenen Stellen werde ich auf Aspekte der Snowden-Affäre und die damit verbundenen umfassenden Diskussionen zurückkommen.

Dieser Aufsatz hat viel mit Begriffen zu tun: Ich bin weniger daran interessiert, abschließende Urteile anzubieten, als Gespräche zu eröffnen. Ich werde Begriffe entwickeln wie ȟberholte Theorie«, »metastasierende Daten«, »Dialektik von Überwachung und Anerkennung« und »faule Daten«. Der Titel »Erstellen von big data - in der Theorie« verweist auf die Themen »Erstellen« und »Theorie«, von denen hier immer wieder die Rede sein wird. Eine Analyse von über 27.000 sozialwissenschaftlichen Artikeln, die zwischen 2000 und 2009 erschienen, fand heraus, dass »nur etwa 30 Prozent von Internetstudien eine oder mehrere theoretische Quellen zitieren, was die Vermutung nahelegt, dass Internetstudien im vergangenen Jahrzehnt nur eine bescheidene theoretische Basis hatten« (Peng et al., 2012, S. 655). Gerade dann besteht ein erheblicher Bedarf an Theoretisierung, wenn sich entwickelnde Datenkonfigurationen Begriffe scheinbar überflüssig machen könnten – nämlich um zu unterstreichen, dass es außerhalb der Welt der Begriffe keinen archimedischen Punkt der reinen Daten gibt. Daten haben stets theoretische Rahmenbedingungen als Grundvoraussetzung ihres Erstellens. Wer konkret mit Big Data arbeitet, weiß: Auch wenn sie »aufschlussreich sein können, sind sie doch nicht unproblematisch. Jeder Datensatz erweist sich als begrenzte Darstellung der Welt.« (Loudon et al., 2013, Unveröffentl. Manuskript. o.S.)

Es steht viel auf dem Spiel. Die algorithmische Lebensweise löst gerade die künstliche Intelligenz als die Modalität ab, nach der die Computertechnik die Gesellschaft gestaltet: Aus einem Paradigma der Semantik, des Verstehens wird ein Paradigma der Pragmatik, der Suche. Beim Übersetzen mit Hilfe von Computerprogrammen etwa versucht man keineswegs einen Computer dazu zu bringen, eine Sprache auf intelligente Weise zu verstehen - Systeme wie Google Translate beispielsweise gleichen Texte aus einem riesigen Korpus ab, ohne dass der Computer jemals »weiß«, was da gesagt wird. Historisch gesehen war dieser Mangel an Wissen ein Problem, über das man diskutierte, etwa beim Gedankenexperiment des »Chinesischen Zimmers«. Es stellte in Frage, dass ein Mensch, der in einem Zimmer eingesperrt ist und Anweisungen auf Englisch bekommt, um chinesische Schriftzeichen zu verwenden, Chinesisch versteht (Searle, 1980; siehe z.B. Collins, 1990). Aber während man natürlich noch immer über die Möglichkeit von künstlicher Intelligenz debattiert, ist es doch überaus erstaunlich, wie sehr solche Fragen ad acta gelegt worden sind. Darüber hinaus machen aufkommende Paradigmen einer algorithmischen Lebensweise es möglich auf, dass Pragmatik und Semantik konvergieren könnten, dass das »Nutzen« der »Sinn« sein werde, auf den es in einem angeblich neuen Zeitalter der Big Data ankomme.

Die Big Data – dieser riesige und sich ständig verändernde Korpus – stehen somit im Zentrum der Vorstellung, dass eine Hinwendung zur algorithmischen Lebensweise nahe sei, auch wenn dieser Begriff relativ neu sein mag. Der Begriff »Big Data« geht wahrscheinlich informell auf die 1990er Jahre zurück, taucht erstmals in einer akademischen Publikation von 2003 (Lohr, 2013) auf, erlangt aber breitere Legitimität erst um 2008 (Lohr, 2012; siehe Bryant et al., 2008). Gleichwohl haben die Big Data in weniger als einem Jahrzehnt eine Vormachtstellung in vielen Bereichen des technischen Sektors, des Wissenschaftsbetriebs und darüber hinaus erlangt. Gewaltige Subventionen, private und öffentliche Arbeitskraft sowie unternehmerisches, staatliches und militärisches Kapital fließen inzwischen in die Generierung, Erfassung und Analyse von Big Data. Die Geistesund Sozialwissenschaften werden mit Gefahren wie Chancen konfrontiert, nicht zuletzt weil die »Ethnografie« oft als das Andere gegenüber den Big Data dargestellt wird, was wiederum faszinierende Fragen hinsichtlich ihrer Rekombination aufwirft (Manovich, 2011). Es ist von großer Bedeutung, dass eine lebhafte theoretische Diskussion diese aufkommenden Paradigmen gestaltet, denn die »Big Data« stehen im Begriff, eine wichtige Rolle bei der wechselseitigen Konstitution von Technik und Gesellschaft im 21. Jahrhundert zu spielen,

# ÜBERHOLTE THEORIE

Auffällig sind räumliche Metaphern von Mobilität und Omnipräsenz in den Diskussionen über Big Data, aber Big Data sind auch ein überaus zeitliches Phänomen, mit dem sich Debatten über Zeit, Technik und Theorie befassen. Denken Sie nur daran, dass ich über die Snowden-Affäre rede, obwohl sich diese Ereignisse zu dem Zeitpunkt, wenn dieser Aufsatz erscheint, verändert haben werden. Was bedeutet es zu sagen, dass dieser Artikel zu dem Zeitpunkt, da Sie ihn lesen, veraltet sein wird? Habe ich damit seine Nützlichkeit eingeschränkt - wird er 2014 oder gar 2024 interessant sein? Könnte es irgendwie relevant sein hinterherzuhinken?

Ich möchte gern ein wenig bei diesem Problem des Zeitgemäßen verweilen, einem Problem, das stets hinter Diskussionen über Big Data und das Digitale generell hinterherhinkt und den Einfluss der Analyse gefährdet. Ich möchte gern über den Wert von Argumenten nachdenken, deren Zeit abgelaufen ist, die unzeitgemäß sind (Grosz, 2004) – die überholt sind. Um die Zusammenhänge zwischen Big Data, Darstellung, Überwachung und Anerkennung anzusprechen, ist es wichtig, dass ich den Begriff ȟberholte Theorie« entwickle.

Meine Vorstellung von ȟberholter Theorie« basiert auf der Begriffsgeschichte von Daten. Als er dem Aufkommen des Begriffs im 17. und 18. Jahrhundert

<sup>1 |</sup> Mehr über die "algorithmische Lebensweise" bei Mainwaring und Dourish, 2012.

nachging, stellte Daniel Rosenberg fest, Daten seien »der Plural des lateinischen Wortes datum, das wiederum das Neutrum des Partizips Perfekt des Verbs dare, geben, ist. Ein ›Datum‹ [...] ist somit etwas Gegebenes, etwa in einem Argument, etwas, das für selbstverständlich gehalten wird« (Rosenberg, 2013, S. 18). Wenn ich aber im Englischen einen Zusammenhang zwischen »data« und »dated« (deutsch: veraltet, überholt) erkenne, ist das durchaus nicht falsch. »Date« geht auch auf das lateinische Wort datum gleich »Datum« zurück – die gemeinsame Etymologie verweist auf eine buchstäbliche Entsprechung:²

»Im klassischen Latein wurde das Datum eines Briefs durch eine Formulierung wie *data xviii K. Maias de Tarentino*, abgeschickt von Tarentum am 18. Aprik, ausgedrückt. Somit wurde *data*, das erste Wort der Formel, als Begriff für Zeit und Ort verwendet, die darin genannt wurden.« (OED, 2013a)

»Data«, Daten, werden also seit langem mit »Zeit und Ort« des Absenders eines Briefes verbunden – mit dem, was man heute »Metadaten« nennt, die ja in der Snowden-Affäre ein zentrales Thema darstellen. Mit dem Begriff »überholte Theorie« (dated theory) möchte ich betonen, dass Daten immer auch ein zeitliches Gebilde sind – »Daten« haben stets ein »Datum«, das Einfluss auf ihren Wahrheitsanspruch hat.

Der Begriff der überholten Theorie ist nützlich aufgrund gewisser Ängste, nämlich »dass wir alle in der Zeit, die erforderlich ist, um ein bedeutsames Forschungsprojekt zu formulieren, zu finanzieren, durchzuführen, zu überarbeiten und zu veröffentlichen, befürchten müssen, dass Veränderungen im Medienmilieu unsere Arbeit obsolet machen« (Karpf, 2012, S. 640). Ich teile zwar Karpfs Skepsis gegenüber dieser Ansicht, doch eins trägt eindeutig zu einer derartigen Besorgnis bei, nämlich der Gedanke, dass der Wert einer Analyse von ihrer Antizipation abhängig sei. Dieser Gedanke wird von positivistischen Traditionen, die wissenschaftlichen Wert mit vorausschauenden Gesetzen gleichsetzen, ebenso wie von der vom Hype erfüllten Rhetorik der Selbstdarsteller im Silicon Valley geprägt, die so sehr auf »Trending« stehen. Diese Rhetorik wirkt sich auch auf Wissenschaft und Forschung aus: »Die vorherrschende Zeitform [...] ist die der allernächsten Zukunft. Das heißt, Motive und Rahmenbedingungen [...] bilden eine allernächste Zukunft ab, die gleich hinter der nächsten Ecke lauert.« (Dourish und Bell, 2011, S. 23) Das erinnert an die rhetorische Figur der Prolepsis, das literarische Mittel der »Vorausblende«, wie es in einer Formulierung wie »Du bist ein toter Mann« verwendet wird, die ein Killer gegenüber dem Menschen gebraucht, den er gleich ermorden wird (Boellstorff, 2011).

Gegenüber den proleptischen zeitbezogenen Fantasievorstellungen, die Big Data oft begleiten, kann es hilfreich sein, sich daran zu erinnern, dass die Wissensproduktion sich nie vom Wissensproduzenten trennen lässt. Eine Auseinandersetzung mit überholter Theorie ist zugleich eine Auseinandersetzung mit überholten Theoretikern. Denken Sie nur an die gut dokumentierte zeitgebundene Politik der Anthropologie. Die großenteils der Kolonialzeit entstammende Anthropologie wurde beherrscht von einer »Leugnung der Zeitgenossenschaft [...], einer beharrlichen und systematischen Tendenz, die Belege der Anthropologie in eine andere Zeit als die Gegenwart des Produzenten des anthropologischen Diskurses zu versetzen« (Fabian, 1983, S. 31). Innerhalb der Anthropologie ist diese Tendenz entschieden kritisiert worden: Der Begriff der »Rettungsanthropologie« stammt von 1970 (Gruber, 1970), und es gibt viele Forderungen, die »Wildennische« zu überwinden, der die Anthropologie traditionell ihr Forschungsobjekt zuweist, und »eine Anthropologie der Gegenwart besser zu verankern» (Trouillot, 1991, S. 40). Diese selbstbezogenen Debatten prägen eine Kritik der vorherrschenden zeitbezogenen Fantasievorstellung ihrer Fachleute. Dies ist der Tropus, den man vielleicht am ehesten durch das mythische Bild des Wegbereiters Bronisław Malinowski kennt: Ein einsamer Anthropologe landet auf einer tropischen Insel und »entdeckt« einen »abgeschiedenen« Stamm, dessen Lebensweise sich anscheinend seit Jahrhunderten nicht verändert hat (ein Bild, mit dem ich gespielt und das ich in Frage gestellt habe, siehe Boellstorff, 2008, S. 3-4). Der berühmte Cartoon »Far Side« von Gary Larson, auf dem eine Gruppe von »Eingeborenen« in panischer Hast ihre Fernseher und andere technische Geräte versteckt, als ein Kanu mit zwei Anthropologen landet, ist nur ein weiteres Beispiel dafür, wie hartnäckig sich dieser Tropus hält.

Noch immer werden einige Ethnografen von dieser zeitbezogenen Fantasievorstellung motiviert. Aber die Kritik, so unvollständig sie sein mag, zeitigt bereits Konsequenzen: Um von seinem Beruf leben zu können, muss ein Ethnograf inzwischen auch Phänomene wie Immigration, Wissenschaftler oder schwule Indonesier studieren. Vor allem ist man sich weitgehend darin einig, dass die Einbeziehung der Subjektivität des Forschers seine Forschungen wissenschaftlicher, stabiler und ethischer macht.

Dagegen hat man sich noch kaum mit der zeitbezogenen Fantasievorstellung von Forschern auseinandergesetzt, die sich mit Big Data befassen. Wie also prägt die Zeit ihre Subjektivität und das Erstellen von Big Data? Möglicherweise besteht ihre zeitbezogene Fantasievorstellung nicht darin, dass sie in der Vergangenheit herumgraben, sondern in eine gar nicht so nahe Zukunft vorausschauen – eine ferne Zukunft, die sich vorhersagen und sogar proleptisch antizipieren lässt.

Und wie sieht die paradigmatische Figur dieses Forschertyps aus? Ein Kandidat könnte Hari Seldon sein, der Protagonist in Isaac Asimovs klassischer Science-Fiction-Romantrilogie *Foundation*. Seldon, der bedeutendste »Psychohistoriker«, wird zwanzigtausend Jahre in der Zukunft vom Galaktischen Imperium vor Gericht gestellt. Sein Verbrechen ist die Antizipation: Er droht eine Panik auszulösen, weil er das, was wir Big Data nennen können, dazu benutzt, um den Untergang des Imperiums »auf der Grundlage der Mathematik der Psychogeschichte«

**<sup>2</sup>** | In manchen indogermanischen Sprachen wie Deutsch hat *Datum* noch diese zeitliche Bedeutung. Ich danke Axel Bruns, dass er mich daran erinnert hat.

(Asimov, 1951, S. 26) vorherzusagen. Seldon bringt zu seiner Verteidigung vor, er versuche eine »Encyclopedia Galactica« zu schaffen, um die anschließende Zeit der Anarchie abzukürzen:

"Indem ich das Wissen der Rasse rette. Die Summe menschlichen Wissens reicht weit über einen einzelnen Menschen, ja über tausend Menschen hinaus. Mit der Zerstörung unserer sozialen Struktur wird die Wissenschaft in Millionen Teile zerbrochen [...]. Doch wenn wir jetzt eine riesige Zusammenfassung allen Wissens vorbereiten, wird es nie verloren gehen.« (Asimov, 1951, S. 28)

Asimovs Vorwegnahme von Wikipedia, seine Vision dessen, was wir so anachronistisch wie treffend als »Big Data als Social Engineering« bezeichnen können, steht im Einklang mit einem zeitgenössischen Kontext, in dem »die Entwicklung algorithmischer Berechnungen [...] einen wichtigen Zug signalisiert – vom Bemühen, künftige Trends auf der Basis starrer statistischer Daten vorherzusagen, zu einem Mittel, der Zukunft zuvorzukommen« (Amoore, 2009, S. 53). Nun möchte ich freilich nicht behaupten, dass alle, die mit Big Data arbeiten, Asimov gelesen haben oder Hari Seldon sein wollen, genauso wenig wie alle Ethnografen einen »unberührten« Eingeborenenstamm entdecken möchten. Ich meine vielmehr, dass wir nicht bloß Big Data »datieren« müssen, sondern auch die zeitbedingten Fantasievorstellungen, welche diejenigen prägen, die diese Daten nutzen.

So ist beispielsweise in Forschungsprojekten über das Digitale der Gebrauch von Begriffen wie »Zeitalter«, »Äras« und »Epochen« nahezu allgegenwärtig. Umstritten etwa sind solche »Formulierungen, mit denen man sich das Digitale im Sinn von epochalen Veränderungen vorstellt« (Ruppert et al., 2013, S. 22) weil z.B. »der Begriff Digitales Zeitalter Medienhierarchien für jene klassifiziert, die nicht an der Macht teilhaben« (Ginsburg, 2008, S. 139). Diese Bedenken sind berechtigt (ebenso wie der Einwand, der Wortgebrauch von »Zeitaltern« könne dazu führen, die Geschichte zu ignorieren), doch ebenso auch die Behauptung, dass »die Ära von big data begonnen hat« (Boyd und Crawford, 2012, S. 662). Die Periodisierung ist ja nicht bloß ein Silicon-Valley-Hype – für Sozialtheoretiker kann sie genauso wie für Geologen analytisch nützlich und empirisch genau sein. Wir wollen doch nicht das Kind mit dem zeitbedingten Bad ausschütten, denn viele wissenschaftliche Arbeiten über das Digitale wenden Periodisierungen auf einfühlsame Weise an. Natürlich sind Formulierungen wie »das digitale Zeitalter« inzwischen selbst unter leitenden Mitarbeitern in Unternehmen üblich geworden (siehe z.B. Schmidt und Cohen, 2013). Aber der übertriebene Gebrauch oder Missbrauch von Periodisierungen bedeutet ja nicht, dass sie sinnlos sind. Periodisierungen sind heuristische Verfahren, keine Invariablen. Wir können über das genaue Ende des Mesozoikums diskutieren, ebenso wie über den genauen Beginn der industriellen Revolution. Aber gerade weil Periodisierungen umstritten sind, haben sie einen Wert - sie stellen somit ein wichtiges Mittel für das Entstehen von ȟberholter Theorie« dar.

## METADATEN ERSTELLEN

Die Snowden-Affäre rückte eine vermeintliche Untermenge von Big Data in den Vordergrund, nämlich die »Metadaten« – d.h. Dinge wie die Zeit, in der ein Handyanruf erfolgte, seine Dauer und der Standort des Anrufers, im Gegensatz zum Telefongespräch selbst. Vertreter der US-Regierung beharrten darauf, die Überwachungen durch die NSA seien harmlos, weil es dabei nur um Metadaten gehe (siehe Mayer, 2013). Solche Versuche einer Entpolitisierung von Metadaten hingen somit von der Behauptung eines offensichtlichen Unterschieds zwischen Daten und Metadaten ab, und kritische Reaktionen stellten diese Behauptung daher oft in Frage.³ Dies zeigt, dass es im Hinblick auf eine Theorie der Big Data dringend geboten ist, Metadaten zu historisieren und ihre begrifflichen Implikationen darzulegen.

Der Begriff Metadaten ist älter als der Begriff Big Data – er wurde 1968 vom Informatiker Philip R. Bagley (1927–2011) geprägt: »Mit jedem Datenelement lassen sich gewisse Datenelemente verbinden, die Daten >über« das betreffende Element darstellen. Wir bezeichnen solche Daten als >Metadaten«.« (Bagley, 1968, S. 91)

Als Bagley zu Beginn des Internetzeitalters die Vorsilbe *meta*-verwendete, bezog er sich auf eine vielschichtige Vorgeschichte mit impliziten Voraussetzungen. Insbesondere enthält die Vorsilbe *eine uneingestandene Spannung zwischen Lateralität und Hierarchie.* Diese Spannung ist unauflöslich und hat entsprechende Folgen für Macht, Selbstsein und Gemeinschaft.

Vor Bagleys Begriffsprägung reicht die Verwendung der Vorsilbe *meta*- in den Sprach-, Informations- und Kommunikationswissenschaften bis in die 1950er Jahre zurück, als der Linguist Roman Jakobson den Begriff der »Metasprache« zu entwickeln begann. Jakobson ging es um eine

»Unterscheidung, die eindeutig von der altgriechischen und indischen Tradition vorweggenommen wurde [...], ein Bedürfnis, zwischen zwei Sprachebenen zu unterscheiden, nämlich der ›Objektsprache‹, die von Dingen spricht, die der Sprache als solcher fremd sind, und einer Sprache, in der wir über den verbalen Code selbst sprechen. Der letztere Aspekt von Sprache heißt ›Metasprache‹, eine Lehnübersetzung des in den 1930er-Jahren von Alfred Tarski eingeführten polnischen Begriffs.« (Jakobson, 1980, S. 86)

<sup>3 |</sup> Ich will damit nicht behaupten, dass eine staatliche Überwachung nicht Daten wie Metadaten umfasse; ich will damit nur sagen, dass es Versuche gibt, letztere so zu behandeln, als seien sie sicherer zu überwachen, und dass der Unterunterschied zwischen beidem nicht a priori besteht. Ein Beispiel für eine populäre Reaktion auf die versuchte Entpolitisierung von Metadaten bildet der Cartoon »Nothing to worry about, it's just metadata« von Jeff Parker vom 12. August 2013; siehe www.truthdig.com/cartoon/item/nsa\_its\_just\_metadata\_20130812/.

Metasprache ist somit Sprache über Sprache – wenn ich z.B. sage: »Ich träumte von einem Einhorn«, dann ist das Sprache und die Aussage »Ein Einhorn ist ein imaginäres Pferd mit einem einzelnen Horn auf dem Kopf« ist Metasprache. Mit einer pragmatischen Formel stellt ein Geistlicher fest: »Ich erkläre euch zu Mann und Frau.« Daraus folgt, dass »die Aussage: »Wenn in unserer Gesellschaft ein dazu berechtigter religiöser oder gerichtlicher Funktionär gegenüber einem Mann und einer Frau ernsthaft äußert: ›Ich erkläre euch zu Mann und Frau«, dann gelten beide als verheiratet«, eine metapragmatische Aussage ist, die den effektiven Gebrauch dieser Formel beschreibt« (Silverstein, 2001, S. 383).

Jakobsons Anspielung auf Tarski verweist auf eine zweite Traditionslinie der Vorsilbe *meta*-, nämlich in Verbindung mit dem Begriff »Metamathematik«. Er wurde von David Hilbert in den 1920er Jahren entwickelt, geht aber zurück auf Denker wie Russell, Frege, Gödel und Whitehead (siehe Lee, 1997). Diese Tradition hat anscheinend die Kommunikationswissenschaft und damit die Big Data ganz direkt beeinflusst. Die Autoren des Buchs *Menschliche Kommunikation* beispielsweise leiten ihre Theorie der »Metakommunikation« mit der Feststellung ein:

»Wenn wir Kommunikation nicht mehr ausschließlich zur Kommunikation verwenden, sondern um über die Kommunikation selbst zu kommunizieren, [...] so verwenden wir Begriffe, die nicht mehr Teil der Kommunikation sind, sondern (im Sinne des griechischen Präfix meta) von ihr handeln. In Analogie zum Begriff der Metamathematik wird dies Metakommunikation genannt [...].« (Watzlawick et al., 1969, S. 47, Hervorhebung von mir)

Während des 20. Jahrhunderts wurde die Verwendung der Vorsilbe *meta-* immer mehr ausgeweitet, etwa in Begriffen wie »Metawissen« (Watzlawick et al., 1969), »Metaindexikalität« (Lee, 1997), ja sogar »Metakultur« (Urban, 2001). Dennoch ist die Vorsilbe grundlegend instabil geblieben. Einerseits wird sie im hierarchischen Sinn verwendet, sodass es »zwei Ebenen von Sprache« (Jakobson, 1980, S. 86) geben, Metawissen »ein Wissen zweiter Ordnung« (Watzlawick et al., 1969, S. 287) sein kann oder Metadaten »transzendent und allumfassend« (Beer und Burrows, 2013, S. 51) sein können. Dies ist das System einer Null-Grad-Referenz (z.B. Sprache, Wissen oder Daten) und dann von »Metaphänomenen«, die darüber oder darunter liegen. Andererseits wird die Vorsilbe *meta-* (zuweilen vom selben Autor) im lateralen Sinn verwendet, und dann ist Metasprache »eine Sprache, in der wir über den verbalen Code selbst sprechen« (Jakobson, 1980, S. 86), oder Metakultur ist »Kultur, in der es um Kultur geht« (Urban, 2001, S. 3).

Nur selten hat man sich mit dem Ursprung dieser Doppelbedeutung der Vorsilbe *meta*- befasst: »meta« ist noch keine überholte Theorie. Denken wir nur an die lange Geschichte von μετα (auf die Jakobson mit seinem Hinweis auf die »altgriechische und indische Tradition« anspielt), in der die Vorsilbe ursprünglich nur eine laterale Bedeutung hatte: »Im Altgriechischen und im hellenistischen Griechischen dient die Vorsilbe dazu, Vorstellungen von Teilen, Handeln im All-

gemeinen, Streben, Suche und vor allem von Veränderung (von Ort, Ordnung, Zustand oder Wesen) auszudrücken« (OED, 2013b). Diese ursprüngliche Bedeutung liegt den Begriffen »Massenmedien« oder »digitalen Medien« zugrunde: »Aristoteles [...] spricht von zwei Elementen, nämlich Luft und Wasser, im Sinne von >Zwischenelementen<. Mit anderen Worten: Er hat als Erster eine verbreitete griechische Präposition – metaxú, zwischen – in ein philosophisches Substantiv, in einen Begriff umgewandelt: tò metaxú, das Medium.« (Kittler, 2009, S. 26) Nach dieser ursprünglichen Bedeutung der Vorsilbe meta- ist »Metamedien« somit ein redundanter Begriff. Diese Lateralität ist mittlerweile nahezu vergessen und taucht nur in ein paar Begriffen auf, nämlich »Metapher« (wörtlich »hinübertragen«), »Metathese« und interessanterweise »Metastase«, seit der Spätrenaissance ein medizinischer Begriff, der die Funktionsübertragung zwischen Organen bezeichnet. Sein Gegenteil war Redux, die Rückkehr eines erkrankten Organs in seinen ursprünglichen Zustand (Maurer, 1997). Die Metastase bewirkt eine Zustandsänderung, nicht einen Zustand darüber; somit kann es nützlich sein, sich theoretisch mit Daten zu befassen, die »metastasieren«.

Aber wenn sich μετα ursprünglich auf Lateralität bezog – »vorher« und »nachher« –, wieso kam es dann zu seiner hierarchischen, abstrahierenden Bedeutung? Das lag an einem Klassifizierungsfehler, einer falschen Zuordnung von Büchern. Andronikos von Rhodos, der im 1. Jahrhundert v. Chr. die erste Gesamtausgabe der Werke von Aristoteles zusammenstellte, »fasste eine Reihe von Aristoteles' Schriften in einem einzigen Band zusammen und ordnete diesen nach (meta-)den naturwissenschaftlichen Abhandlungen (physika) ein. Somit bedeutete der Begriff metaphysika, der diesen Band der Werke von Aristoteles bezeichnete, nicht das, was er später bedeutete – ein Untergebiet der Philosophie« (Anagnostopoulos, 2009, S. 18). Aufgrund der Überlieferung »der zwölf Bücher, die leider den Titel >Metaphysik« tragen« (Kittler, 2009, S. 24), wurde der Begriff »als Bezeichnung für den in diesen Büchern behandelten Wissenszweig verwendet und schließlich fehlinterpretiert als >die Wissenschaft von Dingen, die das, was physisch oder natürlich ist, transzendieren« [...], ungeachtet der Tatsache, dass μετα sich keineswegs im Sinn von >jenseits« oder >transzendierend« interpretieren lässt« (OED, 2013c).

Die Geschichte dieser Fehlinterpretation ist keineswegs unbekannt – man kann sie sogar im Eintrag »meta« bei Wikipedia nachlesen. Wenn ich sie hier nacherzähle, will ich damit weder einem etymologischen Determinismus noch dem Eurozentrismus das Wort reden. Wie viele gelehrte Kenner von Aristoteles festgestellt haben, bedeutet die Konstruiertheit des Begriffs »Metaphysik«, der »in keinem uns bekannten Werk von Aristoteles vorkommt« (Merlan, 1968, S. 175), durchaus nicht, dass es überhaupt keinen Begriff von Metaphysik geben kann. Zur Debatte steht vielmehr das beredte Schweigen im Hinblick auf die zufällige Entstehung des Begriffs sowie die Spannungen, die in der Vorsilbe *meta-* an sich enthalten sind. Insbesondere ist die Vorstellung von »Metadaten« von dieser Fehlübersetzung von »Metaphysik« abgeleitet, die sich von der Lateralität entfernt und zur Hierarchie tendiert. Insbesondere seit dem 17. Jahrhundert war »Metaphy-

sik« stark christlich geprägt: Möglicherweise entstammt der einzige neuartige Gebrauch von *meta-* als Vorsilbe vor dem 19. Jahrhundert John Donnes Konzept einer »Metatheologie« über den persönlichen Göttern der Reformation von 1615 (Aronson, 2002; *OED*, 2013c). Dass Metadaten nach diesen Vorstellungen von Metaphysik gestaltet sein könnten, ist doch beachtenswert in einem Gebiet, in dem sich manche Leute als »Technologieevangelisten« bezeichnen und »Avatare« verwenden, zuweilen auf Apple-Computern, deren Logo an den Biss in den Apfel vom Baum der Erkenntnis im Paradies erinnert (Halberstam, 1991).

Ausschlaggebend ist aber, dass der Schwellenwert, der bewirkt, dass sich etwas von einer Nullgrad-Kategorie hin zu seinem »Meta-Analogon« bewegt, nicht a priori existiert. Es handelt sich um einen Akt des Klassifizierens, der als solcher »einen Standpunkt aufwertet und einen anderen unterdrückt« (Bowker und Star. 1999, S. 5). Ich möchte Annahmen einer sauberen Trennung zwischen Daten und Metadaten nicht einfach deshalb in Frage stellen, weil Metadaten eindringlicher sein können als Daten, sondern weil gerade die Einteilung der Welt der Informationen in zwei Bereiche - Nullgrad und meta - Systeme einer impliziten Kontrolle errichtet. Ja, sobald eine solche Unterteilung in null Grad und meta akzeptiert wird, weiß man einfach nicht mehr, wann das aufhört. Wenn wir z.B. annehmen. dass man eine »Metasprache« verwenden muss, um etwas über Sprache zu sagen, dann benötigen wir »eine Metametasprache, wenn wir über die Metasprache reden wollen, und so fort in theoretisch unendlichem Progress« (Watzlawick et al., 1969, S. 217). Daten über Metadaten ließen sich also als »Metametadaten« bezeichnen, aber der Umstand, dass ein »theoretisch unendlicher Progress« in die Vorsilbe meta- eingebaut ist, verweist auf eine Schwäche des Begriffs - nämlich dass er die anfechtbaren sozialen Praktiken, durch die Daten als Gegenstand des Wissens konstituiert werden, überflüssig macht.

Diese Probleme im Zusammenhang mit Metadaten sind nicht auf das Internet beschränkt. Während der Snowden-Affäre waren viele überrascht, als sie von einem seit langem existierenden System der Überwachung physischer Metadaten erfuhren, nämlich »dem Mail Isolation Control and Tracking Programm, bei dem Computer der US-Post alle Briefe, die in den USA verschickt werden, fotografieren« (Nixon, 2013). Hier könnte die Unterscheidung zwischen Daten und Metadaten scheinbar außer Frage stehen (genauso wie der Hinweis auf die historische Verbindung zwischen dem Schreiben von »Daten« auf einen Briefumschlag und der »Datierung«). Was ließe sich schließlich leichter unterscheiden als die Adresse, die auf einem Briefumschlag geschrieben steht, und der Brief, der darin steckt?

Doch genau das und alle Unterscheidungen zwischen Nullgrad und metamöchte ich in Frage stellen. Was wäre denn, wenn wir nicht zwischen Briefen in Umschlägen, zwischen Innen und Außen unterscheiden, sondern zwischen den beiden Seiten einer Postkarte? Postkarten waren umstritten, als sie im späten 19. Jahrhundert aufkamen, weil ihr »Inhalt« von jedem gelesen werden konnte (Cure, 2013) – hier wird es nämlich schwierig, zwischen Form und Inhalt zu unterscheiden (Boellstorff, 2013). Könnte nicht die analoge Behandlung der Post-

karte eine Möglichkeit darstellen, diesen Binarismus zu überdenken? Wenn ich nämlich aus einer Postkarte eine Möbiusschleife formen könnte, würde mich das noch mehr freuen: Das würde dann eine Vorstellung davon vermitteln, wie verschränkt Form und Inhalt auf der fundamentalsten Ebene sind, sodass Akte der Zuschreibung von »meta-« eindeutig die kulturellen und politischen Akte sind, die sie eigentlich sind, und nicht irgendwelche vorgegebenen Merkmale.

Bemerkenswerterweise haben Informatiker und Ingenieure das schon lange erkannt. So bestehen beispielsweise Dateisysteme generell aus mindestens zwei »Schichten«, wobei die eine Portionen der »Daten« und die andere »Metadaten« speichert (z.B. wem die Datei gehört oder wann sie angelegt wurde). Aber das ist eine begriffliche Unterscheidung, und eine Operation eines Dateisystems (etwa das Schreiben einer Datei) wird als eine einzige Operation angesehen, die einige Informationen in die »Datenportion« und einige Informationen in die »Metadatenportion« des Dateisystems eingibt. Die »Metadaten« werden nicht wegen irgendwelcher »Metamerkmale« als speziell behandelt, sondern wegen ihrer Nutzung als Daten (z.B. unterschiedliche Zugriffsgeschwindigkeiten oder unterschiedliche Speichereigenschaften). Diese praktischen Erkenntnisse im Hinblick auf die Konstruiertheit der Unterscheidung von Nullgrad und meta werden in wissenschaftlichen Arbeiten über die digitale Technologie aufgegriffen, die die oft verborgene »mühsame Arbeit der Erstellung von Metadaten« betonen (Losh, 2009, S. 266). So hat z.B. ein leitender Bibliothekar an der British Library hingewiesen auf

»die Gefahren des teilweisen Outsourcing der mühsamen Erstellung von Metadaten nach Indien, wo selbst die am besten Englisch sprechenden Operators, die an den digitalen Kopien von Zeitungen arbeiten, vielleicht nicht einmal übliche englische Ortsnamen erkennen. Somit beweist selbst die angeblich neutralste Tätigkeit der Erstellung von Metadaten [...], dass ein kulturell orientiertes Wissen noch immer wichtig sein kann für ein sinnvolles Auszeichnen im digitalen Zeitalter.« (Ebd.)

Diese Beispiele unterstreichen die praktischen und politischen Konsequenzen einer Theorie. Es ist eben nicht einfach so, dass das Bezeichnen von Dingen als »Daten« bloß ein Akt des Klassifizierens ist – Dinge als »Metadaten« zu bezeichnen ist nicht weniger ein Akt des Klassifizierens und nicht weniger in Machtund Kontrollprozesse verstrickt. Dies beruht auf einer langen und verwickelten Geschichte der Spannungen zwischen hierarchischem und lateralem Denken, die alles gestalten, von Dateisystem bis hin zu Gesellschaften. <sup>5</sup> Diese Geschichte

<sup>4 |</sup> Ich danke Mic Bowman für diese Erkenntnisse über Dateisysteme.

**<sup>5</sup>** | Ironischerweise sind es ausgerechnet Markierungen an der Außenseite alter mesopotamischer Tongefäße, die auf deren Inhalt hinweisen, die vielleicht »den genauen Beginn der Technik des Schreibens darstellen« (Schmandt-Besserat, 1980, S. 357). Was man heute als Metadaten versteht, war vielleicht zuerst da.

117

stellt jeden Versuch in Frage, die Unterscheidung zwischen Nullgrad-Daten und Metadaten für selbstverständlich zu halten.

Scherzhaft erkläre ich gern, dass das Definieren von »meta« wie das Definieren eines anderen englischen Four-letter word ist, nämlich von »porn«: Man erkennt es, wenn man es sieht. Aber diese scherzhafte Parallele ist überraschend genau, denn im Hinblick auf das Obszöne gilt weithin – fast nie aber im Hinblick auf die Vorsilbe meta--, dass beides (wie alle sozialen Phänomene) von einer gemeinsamen Praxis definiert wird. Was als Obszönität betrachtet wird, hängt von bestimmten zeitlichen und lokalen Normen ab, und ähnlich kontextuell ist auch »meta«. Linguistisch gesprochen ist es irreführend, eine Parallele zwischen Nullgrad und meta im Sinne einer strukturellen Unterscheidung wie der zwischen gesprochener Sprache und Grammatik zu suchen. Effektiver ist es, dies im Sinne eines »Sprachwechsels« zwischen Englisch und Spanisch zu tun, der Bewegung zwischen formellen und informellen Registern einer Sprache oder gar des »Tagging« - sich entwickelnden Akten des Auszeichnens, die im Lauf der Zeit allgemein akzeptierte »Hashtag-Kategorien« werden.

Es ist somit empirisch korrekt und politisch geboten, »Metasprachen nicht bloß als Analysesysteme, sondern als Kommunikationspraktiken zu behandeln« (Jensen 2013) Man denke nur daran, wie Suchbegriffe, ein prototypisches Beispiel für »Metadaten«, durch die soziale Praxis »Daten« werden können. Dies geschieht etwa, wenn solche Suchverfahren (oft ungenau) dazu benutzt werden, mögliche Grippeausbrüche aufzuspüren, indem man verstärkt nach Begriffen wie »Grippebehandlung« sucht (Crawford, 2013). Ein weiteres Beispiel ist die Reaktion von LGBT-Aktivisten auf die heterosexistische Einstellung von Rick Santorum, eines ehemaligen Kongressabgeordneten aus Pennsylvania: Sie benutzten seinen Namen als Teil einer Kettensuche nach Begriffen für sexuelle Flüssigkeiten, indem sie vorübergehend die eigens geschaffene Internetseite »spreadingsantorum.com« an die Spitze der Google-Ergebnisse für eine Suche nach dem Begriff »santorum« manipulierten (Gillespie, 2012). In derartigen Beispielen fungieren Phänomene, die typischerweise als Metakommunikation eingestuft werden, als Formen von Kommunikation.

Dies ist vielleicht der wichtigste theoretische Aspekt im Hinblick auf das Erstellen von Metadaten, da er sich allgemeiner auf die Sozialtheorie auswirkt. Die Tatsache, dass der Akt der Zuschreibung von »meta« in einem kulturellen Kontext steht, ist relevant für jeden Gebrauch der Vorsilbe, von der Metaphysik bis zur Metapragmatik, von der Metakommunikation bis zu den Metamedien, vom Metawissen bis zur Metakultur. Ja, Gregory Bateson - einer der von Fachwissenschaftlern für Kommunikation und digitale Kultur meistzitierten klassischen Anthropologen - hat die Theorie aufgestellt, Spiel und Fantasie seien eine Art Metakommunikation, die ganz entscheidend für die Entwicklung von Kommunikation an sich sei, nämlich dann, wenn »ein bestimmter Organismus in der Lage ist [...], die von ihm und anderen Angehörigen der Spezies ausgehenden Zeichen als Signale zu erkennen« (Bateson, 1972, S. 243). Ein besseres Verständnis des Entstehens von meta- wird daher von zentraler Bedeutung sein, um sich entwickelnde Formen von Big Data und ihre sozialen Implikationen zu begreifen.

# DIE DIALEKTIK VON ÜBERWACHUNG UND ANERKENNUNG

Die Snowden-Affäre verstärkte die bereits geführten Debatten über Big Data, Überwachung und die Macht des Staates. Um diese Macht in Frage zu stellen, ging Snowden schließlich all diese Risiken ein und betonte: »Die größte Angst bereitet mir das, was sich nach diesen Enthüllungen für Amerika daraus ergibt, nämlich dass sich nichts ändern wird.« (Edward Snowden, in Rodriguez, 2013) Diese Angst war durchaus begründet – öffentliche Umfragen ergaben, dass viele Amerikaner im Hinblick auf diese staatliche Überwachung zwiespältig, gleichgültig oder gar begeistert waren (Ohlheiser, 2013). Debatten über das Erstellen von Big Data beziehen sich eindeutig auf eine umstrittene kulturelle Logik von Überwachen, Privatsphäre und Enthüllung. Ich bezeichne diese Logik vorläufig als »die Dialektik von Überwachung und Anerkennung«.

Nach Snowdens Enthüllungen über eine von Big Data ermöglichte staatliche Überwachung griffen viele Leute auf Präzedenzfälle, Vergleiche, Metaphern zurück. Am häufigsten wurde Orwell herangezogen, wobei man bereits zuvor festgestellt hatte, dass Big Data »eine finsterere Konnotation aufweisen, gewissermaßen als linguistischer Vetter des Großen Bruders« (Lohr, 2012) Doch viele hielten dies für einen begrenzten Tropus (z.B. M. Jensen, 2013), nicht zuletzt weil »Orwell nicht erkannte, dass die Technik [...] Gruppen von Menschen ungeheure Macht verlieh, Dinge aufzubauen, die so kompliziert und wunderbar sind wie [...] Wikipedia« (Cory Doctorow, interviewt in Porzucki, 2013). Mit anderen Worten: Die orwellsche Metapher greift zu kurz, weil das Konzept der »Big Data« sowohl relativ ungewollte Daten (wie von einem bewegten Smartphone generierte GPS-Daten) wie relativ gewollte Daten (wie ein Posten bei Facebook) umfasst.

Vielleicht berief Snowden sich deshalb nicht auf George Orwell, sondern auf Michel Foucault, als er erklärte: »Wenn ein Überwachungsprogramm wertvolle Informationen liefert, ist es legitim [...] In einem Schritt ist es uns gelungen, die Operation des Panopticon zu rechtfertigen.« Snowden spielte hier auf Foucaults Darstellung des Panopticon in Überwachen und Strafen an, wie es der utilitaristische Philosoph Jeremy Bentham als Teil einer Gefängnisreform vorschlug.<sup>6</sup> Ein Gefängnis sollte danach aus ringförmig angeordneten Zellen bestehen, die auf einen zentralen Turm ausgerichtet sind, das Panopticon, der es einem einzigen Aufseher ermögliche, das Gefängnis zu überwachen. Außerdem wäre das

<sup>6 |</sup> Foucault wird zwar nicht namentlich genannt, doch die Anspielung ist wahrscheinlich bewusst. Dazu ein Beobachter: »Mit Sicherheit hat Snowden seinen Michel Foucault genau gelesen (so hat er auch seinen Abscheu angesichts der Fähigkeiten dieser Architektur der Unterdrückung betont.) « (Escobar, 2013)

Panopticon so konstruiert, dass die Gefangenen niemals wissen, ob sich jemand im Turm befindet. Sie würden den Blick des Panopticons verinnerlichen und ihr eigenes Verhalten überwachen: »Daraus ergibt sich die Hauptwirkung des Panopticon: die Schaffung eines bewussten und permanenten Sichtbarkeitszustandes beim Gefangenen [...] Die Wirkung der Überwachung ist permanent, auch wenn ihre Durchführung sporadisch ist.« (Foucault, 2008b, S. 906) Das Panopticon stellt eine optische Metapher dar, die geradezu vorausschauend zu sein scheint, wenn ein Überwachungsprogramm der NSA den Codenamen »Prisma« bekommen kann: »Zu ihrer Durchsetzung muss sich diese Macht mit einer ununterbrochenen, erschöpfenden, allgegenwärtigen Überwachung ausstatten [...]. Tausende von Augen, die überall postiert sind; bewegliche und ständig wachsame Aufmerksamkeiten; ein weites, hierarchisiertes Netz [...].« (Foucault, 2008b, S. 920)

Doch aus der Sicht Foucaults müsste die Hauptmetapher für das Erstellen von Big Data nicht das Panopticon sein, sondern das Geständnis. Der erste Band von Sexualität und Wahrheit erschien zwar nur ein Jahr nach Überwachen und Strafen und griff viele Themen des früheren Werks auf, doch Foucault widmete hier noch mehr Aufmerksamkeit der Frage, wie Macht, Wissen und Selbstsein in spezifischen historischen Kontexten zusammenkommen. In dem Kapitel Die Anreizung zu Diskursen heißt es: »Um das 18. Jahrhundert herum entsteht ein politischer, ökonomischer und technischer Anreiz, vom Sex zu sprechen [...], in Form von Analyse, Buchführung, Klassifizierung und Spezifizierung, in Form quantitativer oder kausaler Untersuchungen.« (Foucault, 2008a, S. 1045) Sex wurde in Daten umgesetzt, und das hatte zwei entscheidende Konsequenzen. Zum einen waren diese Daten Teil eines staatlichen Projekts: »Der Sex, das ist nicht nur eine Sache der Verurteilung, das ist eine Sache der Verwaltung. Er ist Sache der öffentlichen Gewalt, er erfordert Verwaltungsprozeduren, er muss analytischen Diskursen anvertraut werden.« (Foucault, 2008a, S. 1046) Zum andern wurden diese Daten produziert durch das Diskursritual des »Geständnisses«, das auf das Christentum und die psychoanalytische Begegnung zwischen Therapeut und Patient zurückgeht:

»Für uns verbinden sich Wahrheit und Sex im Geständnis mittels des obligatorischen und erschöpfenden Ausdrucks eines individuellen Geheimnisses [...]. [Z]ugleich ist es ein Ritual, das sich innerhalb eines Machtverhältnisses entfaltet, denn niemand leistet sein Geständnis ohne die wenigstens virtuelle Gegenwart eines Partners, der nicht einfach Gesprächspartner, sondern Instanz ist, die das Geständnis fordert, erzwingt [...].« (Foucault, 2008a, S. 1074f.)

Das Geständnis ist eine moderne Möglichkeit, Daten zu produzieren, ein Anreiz zum Diskurs, den wir nun einen Anreiz zur Enthüllung nennen könnten. Er ist zutiefst dialogisch: Einer gesteht einem Mächtigen/Anderen. Dies kann technisch vermittelt sein – wie Foucault feststellte, denn es kann in der »virtuellen Gegenwart« einer Instanz stattfinden. Das ist die einzige Stelle im ersten Band von Se-

xualität und Wahrheit, an der das Wort »virtuell« vorkommt, und seine Verwendung ist bedeutsam. Um der Frage nachzugehen, welche Rolle dieser Anreiz zur Enthüllung in zeitgenössischen Kontexten spielt, ist es hilfreich, sich mit Charles Taylors Darstellung der »Politik der Anerkennung« zu befassen, die für ihn von zentraler Bedeutung für die Moderne ist:

»Die These lautet, unsere Identität werde teilweise von der Anerkennung oder Nicht-Anerkennung, oft auch von der Verkennung durch die anderen geprägt, so dass ein Mensch oder eine Gruppe von Menschen wirklichen Schaden nehmen, eine wirkliche Deformation erleiden kann, wenn die Umgebung oder die Gesellschaft ein einschränkendes, herabwürdigendes oder verächtliches Bild ihrer selbst zurückspiegelt.« (Taylor, 2009, S. 13)

Wenn ich von der »Dialektik von Überwachung und Anerkennung« spreche, versuche ich, den Gedanken des Geständnisdiskurses mit der Politik der Anerkennung zu verknüpfen.<sup>7</sup> Zur weiteren Erforschung stelle ich die These auf, dass das Aufkommen von Big Data von einem Diskurs begleitet wird, der Überwachung mit Anerkennung verbindet, der Überwachung als eine Form von Zugehörigkeit strukturiert. Kein Diskurs ist einzigartig, und mit Sicherheit gibt es umgekehrte Diskurse, Gegendiskurse und alternative Diskurse. Die Frage ist nicht, ob die von der Snowden-Affäre ans Licht gezerrte staatliche Überwachung unumstritten ist (weil sie offensichtlich umstritten ist), sondern warum so viele Menschen Überwachung akzeptabel und sogar angenehm finden: »Das Spiel ist ganz entscheidend für das Verständnis der neuen sozialen Daten.« (Beer und Burrows, 2013, S. 51). Eine der wichtigsten politischen Lehren von Foucaults Werk war die Erkenntnis, dass Widerstand oft im Rahmen eines Diskurses auf eine komplexe Weise erwächst, der puristische Vorstellungen von Gegnerschaft kaum gerecht werden. Im Hinblick auf den Umgang der Politik mit Homosexualität hat Foucault festgestellt, die Diskurse im 19. Jahrhundert über die Arten und Unterarten der Homosexualität hätten

»auch die Konstitution eines Gegen-Diskurses ermöglicht: die Homosexualität hat begonnen von sich selber zu sprechen, auf ihre Rechtmäßigkeit oder auf ihre ›Natürlichkeit‹ zu pochen – und dies häufig in dem Vokabular und in den Kategorien, mit denen sie medizinisch disqualifiziert wurde. Es gibt nicht auf der einen Seite den Diskurs der Macht und auf der andern Seite den Diskurs, der sich ihr entgegensetzt.« (Foucault, 2008a, S. 1105)

Noch ist nicht klar, was für Gegen-Diskurse im Hinblick auf Big Data und ihre Dialektik von Überwachung und Anerkennung aufkommen werden. Doch eine gewisse Vorstellung davon vermittelt die Tatsache, dass viele Reaktionen auf das Erstellen von Big Data implizit nicht etwa Aufrufe für ihre Abschaffung, sondern für

**<sup>7</sup>** | Weitere wichtige Ausführungen zu Anerkennung und Zugehörigkeit finden sich bei Fraser, 2000; Povinelli, 2002.

ihre Ausweitung sind. In ihrer Kritik an Big Data bemerkte Kate Crawford: »Daten spiegeln angeblich die soziale Welt wider, doch da gibt es erhebliche Lücken, weil von bestimmten Gemeinschaften nur geringe oder keine Signale kommen«, sodass »wir bei jedem Satz von big data fragen müssen, welche Menschen ausgeschlossen worden sind. Welche Orte sind weniger sichtbar? Was geschieht, wenn man im Schatten von Big-Data-Sätzen lebt?« (Crawford, 2013) Auch viele andere Wissenschaftler äußern diese Bedenken und fordern, dass wir uns »der Zweifel am repräsentativen Charakter von Daten bewusst sein sollen, wenn von Suchmaschinennutzern pauschale Schlüsse auf eine ganze Bevölkerung gezogen werden» (Trevisan, 2013, S. 2). Ich teile diese Bedenken und wünsche mir, dass mehr Menschen einbezogen werden. Es geht darum, dass auf eine beinahe homöopathische Weise das Heilmittel innerhalb des begrifflichen Horizonts der Krankheit liegt, die es lindern soll – innerhalb der Dialektik von Überwachung und Anerkennung.

# VERROTTETE DATEN, DICHTE DATEN

Snowden rechtfertigte seine Enthüllungen über die NSA-Überwachung mit dem Argument, dass eine solche Datensammlung stets in einem Interpretationsrahmen stattfinde, der sogar nachträglich angelegt werde, sodass eine Regierung »zeitlich zurückgehen und jede Entscheidung, die Sie je getroffen haben, überprüfen« könne (Edward Snowden, in Rodriguez, 2013). Damit verknüpfte er Behauptungen über Daten und Zeitlichkeit mit »Überprüfung« – mit der in einem kulturellen Kontext stehenden Arbeit der Interpretation. Dies spiegelt sich in einer gerade entstehenden Literatur wider, die die Vorstellung von »Rohdaten« in Frage stellt. In ihrer Einleitung zu »Raw data« is an oxymoron bemerkten Gitelman und Jackson, der Titel des von ihnen herausgegebenen Bandes spiele auf eine Bemerkung von Geoffrey Bowker an (Bowker, 2013, S. 1). Diese Bemerkung lautet: »Rohdaten sind sowohl ein Oxymoron wie eine schlechte Idee – im Gegenteil, Daten sollten sorgfältig gekocht werden.« (Bowker, 2005, S. 184). Dies wiederum spielt natürlich auf Claude Lévi-Strauss' Buch Das Rohe und das Gekochte (1964) an:

»Der Begriff ›Rohdaten entspricht Levi-Strauss' Gebrauch des Begriffs ›roh [...] [zur Beschreibung] eines riesigen mythologischen Feldes [...]. Er argumentierte, eine Reihe von Zweiteilungen würde dieses Feld charakterisieren, wobei viele eine Variante dessen wären, was wir die Kluft zwischen Natur und Gesellschaft nennen würden. Das Natürliche wäre dann das Rohe (Honig), das Soziale das Gekochte (Asche). « (Bowker, 2013, S. 168)

Überraschenderweise ist diese Zweiteilung von roh und gekocht sowohl etisch (aus der Sicht eines Außenseiters) wie emisch (im Alltagsgebrauch). So fanden Räsänen und Nyce beispielsweise in ihrer Studie über die schwedischen Geheimdienste heraus, dass »Geheimdienstangehörige den Begriff ›Rohdaten‹ als Ver-

nunftkategorie verwenden«, und versuchten, »dieses praktische Verständnis zentraler Kategorien wie des Rohen und des Gekochten in Frage zu stellen« (Räsänen und Nyce, 2013, S. 656, 660). Diese Kategorien sind nämlich im Hinblick auf Big Data unglaublich wichtig. Ein Grund dafür ist die Folgerung, die »Größe« von Daten bedeute, dass sie vor jeder Interpretation gesammelt werden müssten – also »roh«. Darauf verweisen Metaphern wie das »Datenscraping«, die an das Abschaben von Fleisch vom Knochen erinnern, an das Entfernen von etwas, das für ein augenscheinliches Oberflächenphänomen gehalten wird. Eine weitere Folgerung lautet, dass in einer schönen neuen Welt der Big Data die Interpretation dieser Daten, ihr »Kochen«, zunehmend von Computern selbst geleistet werde.

Doch wie die oben genannten Autoren (und andere) festgestellt haben, ist dies ein weiteres Beispiel dafür, wie die klassische Anthropologie wertvolle Erkenntnisse auf anderen Gebieten vermittelt. Lévi-Strauss betonte zu Beginn von Das Rohe und das Gekochte die Bewegung zwischen emisch und etisch: »Das Ziel dieses Buches ist es, aufzuzeigen, auf welche Weise empirische Kategorien [...], die nur durch die ethnographische, jeweils den Standort einer besonderen Kultur einnehmende Beobachtung, präzise definierbar sind, [...] dennoch als begriffliches Werkzeug dienen können.« (Lévi-Strauss, 1969, S. 1) In diesem Buch behandelte Lévi-Strauss das Rohe und Gekochte oft im Sinne einer Dichotomie. Doch in dem Aufsatz »Das kulinarische Dreieck«, der ein Jahr nach Das Rohe und das Gekochte erschien, stellte er diese Kategorien in ein Dreiecksverhältnis zum »Verrotteten«.8 Nach dieser Theorie stellen »roh« und »gekocht« keine Zweiteilung dar, in der roh gleich Natur und gekocht gleich Kultur ist. Vielmehr werden sie als Elemente eines »kulinarischen Dreiecks« dargestellt, das von der Schnittstelle der Zweiteilungen von »Natur« und »Kultur« sowie von »normal« und »verwandelt« gebildet wird (Abb. 1). Hier »stellt das Rohe den unbetonten Pol dar, während die anderen beiden Pole stark betont sind, aber in unterschiedlichen Richtungen: Tatsächlich ist das Gekochte eine kulturelle Verwandlung des Rohen, während das Verrottete eine natürliche Verwandlung ist« (Lévi-Strauss, 1997, S. 29).

**<sup>8</sup>** | Lévi-Strauss befasste sich mit der Kategorie des Verrotteten nur gelegentlich in *Das Rohe und das Gekochte* (z.B. S. 176, 254). *Das Rohe und das Gekochte* erschien zuerst 1964, »Das kulinarische Dreieck« 1965 (siehe Lévi-Strauss, 1997).

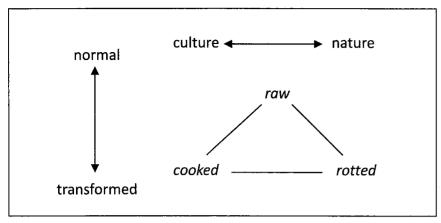


Abbildung 1: Lévi-Strauss' kulinarisches Dreieck (Zeichnung des Autors) (Kultur – Natur, normal – verwandelt, roh, gekocht, verrottet)

Diese Dreiteilung ist überaus anregend für eine Theorie des Erstellens von Big Data. Wie »roh« und »gekocht« ist auch die Kategorie des »Verrotteten« sowohl emisch wie etisch, wie es im Begriff des »Bit Rot« zum Ausdruck kommt. Damit kann die Materialität von Daten gemeint sein – die Art und Weise, wie verrottendes Magnetband und beschädigte Festplatten zu Datenverlust führen. Doch seit langem hat man festgestellt, dass beim »Bit Rot« nicht so sehr der Datenverlust auf dem Spiel steht, sondern die Fähigkeit zur Interpretation: »Lange bevor die Festplatte sich abnützt oder einem Bit Rot erliegt, ist die Maschine, die die Festplatte liest, ein Museumsstück geworden. Die unmittelbare Aufgabe besteht also nicht darin, die Informationen zu erhalten, sondern die Mittel, um an sie zu gelangen.« (Hayes, 1998, S. 410)

Im Kontext von roh und gekocht ermöglicht das »Verrottete« Verwandlungen außerhalb der typischen Konstruktionen des handelnden Menschen als Koch – das Ungeplante, Unerwartete und Zufällige. Bit Rot beispielsweise entsteht aus dem Zusammenwirken von Speicher- und Verarbeitungstechnologien, während diese sich durch die Zeit bewegen. Aber »Verrotten« bewegt sich zwischen Natur und Gesellschaft ebenso wie zwischen absichtlich und unabsichtlich. Verrotten kann »entweder spontan oder kontrolliert« (Lévi-Strauss, 1997, S. 29) sein, wobei es im letzteren Fall oft als »Gärung« oder »Destillation« bezeichnet wird und alles Mögliche erzeugt, von Brot und Käse bis zu Bier und Wein.

Da es die Vorstellung des »Destillierens« von Bedeutung aus Big Data schon lange gibt (z.B. Frankel und Reid, 2008), könnte die Vorstellung von verrotteten Daten somit eine begriffliche Lupe darstellen, um Überschneidungen des Absichtlichen und des Zufälligen in der Erzeugung, Interpretation und Anwendung von Big Data zu betrachten. Ein »kulinarisches Datendreieck« stellt zwischen dem Rohen und dem Gekochten eine logische statt eine zeitliche Beziehung her. Bei drei Begriffen in einem Dreieck statt zwei Begriffen in einer Reihe lässt sich das

Einfügen eines Zeitstrahls leichter vermeiden – ebenso wie die Annahme, dass das Rohe vor dem Gekochten komme, und damit lässt sich der Machtanspruch leichter in Frage stellen, der im zeitlichen Argument enthalten ist, dass Daten vor der Interpretation kommen. Der Gedanke von »verrotteten Daten« spiegelt somit wider, wie sich Daten in parahumane, komplex materielle und zeitlich emergente Möglichkeiten umwandeln lassen, die nicht immer einem vorgegebenen algorithmischen »Rezept« folgen.

Auch auf anderen anthropologischen Erkenntnissen außer denen von Lévi-Strauss lässt sich aufbauen. Indem sie die Vorstellung von »Rohdaten« in Frage stellen, knüpfen Gitelman und Jackson auch an Snowden, Bowker und viele andere an, wenn sie betonen, dass »man sich Daten als Daten vorstellen muss, die als solche existieren und funktionieren, und die Vorstellung von Daten ist mit einer Interpretationsgrundlage verbunden« (Gitelman und Jackson, 2013, S. 3). Dieser Hinweis auf eine »Interpretationsgrundlage« erinnert an Clifford Geertz Einmischungen in eine andere Debatte über das Erstellen von Daten, die vor fast einem halben Jahrhundert geführt wurde. In Dichte Beschreibung, Beiträge zum Verstehen kultureller Systeme reagierte Geertz (ein Anthropologe, der oft zitiert wird von denen, die sich mit Big Data auseinandersetzen) auf kognitive Ethnologen wie Ward Goodenough und Charles Frake, deren wissenschaftliche Arbeiten wiederum zu den algorithmischen Strukturen beitrugen, die für heutige Big Data von so zentraler Bedeutung sind.9 Geertz fasste zunächst das Gedankenexperiment des Philosophen Gilbert Ryle im Hinblick auf das zusammen, was Ryle als »dichte Beschreibung« bezeichnete:

»Stellen wir uns, sagt er, zwei Knaben vor, die blitzschneil das Lid des rechten Auges bewegen. Beim einen ist es ein ungewolltes Zucken, beim anderen ein heimliches Zeichen an seinen Freund. Als Bewegungen sind die beiden Bewegungen identisch [...]. [Es] besteht jedoch ein gewichtiger Unterschied [...]. Der Zwinkerer teilt etwas mit, und zwar auf ganz präzise und besondere Weise: (1) er richtet sich absichtlich (2) an jemand Bestimmten, (3) um eine bestimmte Nachricht zu übermitteln, (4), und zwar nach einem gesellschaftlich festgelegten Code und (5) ohne dass die übrigen Anwesenden eingeweiht sind [...]. Sobald

**<sup>9</sup>** | Diese historischen Zusammenhänge sind komplex und noch unzureichend erforscht. Doch klar ist, dass viele heutige algorithmische Methoden zur Analyse von Big Data auf die Beschäftigung mit der Kognition um die Mitte des 20. Jahrhunderts zurückgehen, die ausgesprochen interdisziplinär war und auf Forschungen von Psychologen wie Amos Tversky zurückgriff (so Nick Seaver in einem Gespräch mit mir). Dies prägte eine ganze Generation von kognitiven Anthropologen für die es möglich war, »die ganze Kultur als Information zu betrachten und jede einzelne Kultur als Informationsökonomie« zu verstehen« (Roberts, 1964, S. 438). Ein solches Paradigma führte z.B. zu Formen der mathematischen Konsensanalyse, die auf der Prämisse basierte, dass man bei »jedem systemischen Kulturmuster davon ausgehen kann, dass es einen damit verbundenen semantischen Bereich hat« (Romney et al., 1986, S. 315).

es einen öffentlichen Code gibt, demzufolge das absichtliche Bewegen des Augenlids als geheimes Zeichen gilt, so *ist* das eben Zwinkern.« (Geertz, 1983, S. 10f.; Hervorhebung im Original)

Während Geertz' Argument komplexer ist, als ich es hier darstellen kann (da gibt es noch einen dritten Jungen, der das Zwinkern des zweiten parodiert), ist es in diesem Zusammenhang relevant, wenn er betont, dass die materielle Geste und das semiotische System auf der gleichen ontologischen Ebene liegen - einer Ebene der öffentlichen, von anderen geteilten Bedeutung. Und genau das macht Daten »dicht«: Was »wir als unsere Daten bezeichnen, [sind] in Wirklichkeit unsere Auslegungen davon [...], wie andere Menschen ihr eigenes Tun und das ihrer Mitmenschen auslegen [...]« (Geertz, 1983, S. 14). Geertz setzte sich hier zwar mit ethnografischen Daten auseinander, doch »Dichte« ist für jede Form von Daten relevant. Ja, Geertz stellte fest, einer Auffassung von Kultur wie die von Goodenough folge »eine ebenso eindeutige Auffassung davon, was es heißt, eine Kultur zu beschreiben: nämlich ein System von Regeln aufzustellen, [einen] ethnografischen Algorithmus« (Geertz, 1983, S. 17, Hervorhebung von mir). Gegenüber dieser früheren Wiederholung von Datenmustern und algorithmischen Lebensweisen betonte Geertz den Wert einer Interpretationsgrundlage: »Wir haben die Triftigkeit unserer Erklärung nicht nach der Anzahl uninterpretierter Daten und radikal verdünnter Beschreibungen zu beurteilen, sondern danach, inwieweit ihre wissenschaftliche Imagination uns mit dem Leben von Fremden in Berührung zu bringen vermag.« (Geertz, 1983, S. 24)

Ich bin hier bewusst anachronistisch. Natürlich ist Geertz' Formulierung »ethnografischer Algorithmus« nicht identisch mit dem heutigen Gebrauch von Algorithmen in der Analyse von Big Data. Aber sie sind eben auch nicht völlig voneinander zu trennen, denn ein historisches Erbe verbindet sie. Dazu gehören auch Kritiken am Strukturalismus von Lévi-Strauss und anderen Anthropologen wie Lacan. Dagegen erkennt das, was Daten »dicht« macht, ihre unreduzierbare Kontextualität an: Was »wir niederschreiben (oder niederzuschreiben versuchen) [ist] kein roher sozialer Diskurs« (Geertz, 1983, S. 29, Hervorhebung von mir). Für Geertz waren »Rohdaten« bereits in den frühen 1970er Jahren ein Oxymoron. Ob gekocht oder verrottet – Daten entstehen aus Interpretationssystemen: »Ebenso wenig haben mich [...] Behauptungen überzeugt, wonach es uns die strukturale Linguistik, die Computertechnik oder irgendeine andere fortgeschrittene Form des Denkens möglich macht, Menschen zu verstehen, ohne sie zu kennen.« (Geertz, 1983, S. 43)

Wenn ich Lévi-Strauss und Geertz in zeitgenössische Debatten über »Rohdaten« einbeziehe, möchte ich damit weder sie verkürzt darstellen noch behaupten, sie würden uns eine Lösung liefern. Vielmehr erfordert der Begriff »Rohdaten« wie der Begriff »Metadaten« eine weitergehende theoretische Beschäftigung damit. Wichtig in diesem Zusammenhang kann es sein, die Rhetorik des Noch-nie-Dagewesenen und des Beschleunigten zu überdenken, die unter-

stellt, wir hätten nichts aus der Geschichte der Sozialtheorie zu lernen – dass Big Data »das Ende der Theorie« bedeuten könnten (Anderson, 2008). Entgegen dem »dünnen« Begriff von Rohdaten können wir uns Daten nicht bloß als gekocht oder verrottet vorstellen, sondern auch als »dicht«. Damit wird betont, dass Big Data niemals einer Interpretation ontologisch vorausgehen – »Interpretation steht im Mittelpunkt der Datenanalyse« (boyd und Crawford, 2012, S. 668) – und dass die Interpretation innerhalb kultureller Horizonte stattfindet, die in Kontexte von Macht eingebettet sind.

## SCHLUSSBEMERKUNG: DAS ERFINDEN VON BIG DATA

In diesem Aufsatz habe ich versucht, mehrere Untersuchungsmöglichkeiten im Hinblick auf das Erstellen von Big Data darzulegen. Aufbauend auf einer Reihe von wissenschaftlichen Gesprächen, habe ich die Zeitlichkeit und die Möglichkeiten der ȟberholten Theorie« erforscht, die impliziten Geschichten, die Metadaten gestalten, »die Dialektik von Überwachung und Anerkennung« sowie Fragen der Interpretation, die sich im Zusammenhang mit Begriffen wie »verrotteten Daten« und »dichten Daten« stellen. Mein Ziel war es, Systeme zu erweitern, mit denen sich Probleme von Zeit, Kontext und Macht ansprechen lassen. Als Ethnograf schätze ich den Wert der fokussierten und lokalisierten Erklärung, doch wir dürfen ein verallgemeinertes Theoretisieren nicht nur einigen Disziplinen und methodologischen Vorgehensweisen überlassen. Notwendig ist eine »plattformagnostische Theorie«, wie ich das nennen möchte – eine Theorie, die Behauptungen über Muster und Dynamik jenseits der Fallstudie und des individuellen Forschungsgebiets aufstellt, selbst wenn diese Besonderheiten das Gebäude der Theorie wie dessen kontextuelle Modifikation bilden. Wenn wir die Big Data einen Gang herunterschalten (indem wir sie nicht mehr mit Großbuchstaben schreiben), können wir sie als konzeptuelle Rubrik, aber auch als ein Forschungsgebiet verstehen, das der Kulturkritik und der ethnografischen Interpretation zugänglich ist.

Es ist ebenso bezeichnend wie beunruhigend, dass Edward Snowden, während ich dies schreibe, noch immer am Moskauer Flughafen ausharren muss – an einem Durchgangsort, der inzwischen zu einem Ort der Nichtzugehörigkeit geworden ist. Außerhalb des Zugriffs der Staatsmacht, die er herausgefordert hat, ist seine Notlage ausgerechnet ein Teil des Systems des Erstellens von Big Data, das er enthüllt hat. Er und viele andere bemühen sich darum, zu zeigen, wie die Gefahren und Verheißungen von Big Data von der Erkenntnis abhängen, dass Big Data nicht bloß »gemacht«, sondern »ausgedacht« werden, im Sinne des »Ausdenkens von Menschen«, von dem Ian Hacking gesprochen hat – demnach beeinflussen Klassifizierungen »die klassifizierten Menschen, und [...] die Beeinflussungen der Menschen verändern wiederum die Klassifizierungen« (Hacking, 2006, S. 23).

Wie Ethnografien sind auch ausgedachte Big Data »Fiktionen, und zwar in dem Sinn, dass sie ›etwas Gemachtes‹ sind, ›etwas Hergestelltes‹ — die ursprüngliche Bedeutung von fictio —, nicht in dem Sinn, dass sie falsch wären, nicht den Tatsachen entsprächen oder bloße Als-ob-Gedankenexperimente wären« (Geertz, 1983, S. 23). Sie sind mehr als bloß »Schnipsel« der Wirklichkeit, sondern wesentliche Bestandteile dieser Wirklichkeit, dem Menschsein immanent. Big Data sind stets auch »große Theorie«, ob sie nun anerkannt wird oder nicht. Wie diese Informationssysteme die Gesellschaften in die sich entfaltende Zukunft hinein gestalten, hängt in nicht geringem Maß von unserer Fähigkeit ab, das Ausdenken von Big Data an sich zu verstehen und darauf zu reagieren.

#### DANK

Mit vielen Menschen habe ich über die Themen diskutiert, die ich in diesem Aufsatz anspreche. Insbesondere danke ich Ken Anderson, Geoffrey Bowker, Mic Bowman, Axel Bruns, Tarleton Gillespie, Klaus Bruhn Jensen, Elizabeth Losh, Annette Markham, Bill Maurer und Nick Seaver. Geistige Anregungen für diese Arbeit verdanke ich dem Intel Science and Technology Center for Social Computing.

## LITERATUR

- Anagnostopoulos, Georgios: »Aristotle's works and the development of his thought«, in: Georgios Anagnostopoulos (Hg.): *A companion to Aristotle*, Malden, Mass., S. 14-27.
- Amoore, Louise: »Algorithmic war: Everyday geographies of the War on Terror«, in: *Antipode*, Bd. 41, Nr. 1, 2009, S. 49-69. http://dx.doi.org/10.1111/j.1467-8330. 2008.00655.x.
- Anderson, Chris: »The end of theory: The data deluge makes the scientific method obsolete«, in: Wired, Bd. 16, Nr. 7, 2008. www.wired.com/science/discoveries/magazine/16-07/pb\_theory.
- Aronson, Jeff: »When I use a word: Meta-«, in: *British Medical Journal*, Bd. 324, Nr. 7344 (27. April 2002), S. 1022.
- Asimov, Isaac: Foundation, New York 1951.
- Bagley, Philip R.: Extension of programming language concepts. Philadelphia 1968.
- Bateson, Gregory: »A theory of play and fantasy«, in: Gregory Bateson: Ökologie des Geistes, Frankfurt a.M. 1973, S. 150-166.
- Beer, David, und Roger Burrows: »Popular culture, digital archives, and the new social life of data«, in: *Theory, Culture & Society*, Bd. 30, Nr. 4, 2013, S. 47-41. http://dx.doi.org/10.1177/0263276413476542.

- Boellstorff, Tom: »An afterword in three postcards«, in: Dominic Power und Robin Tiegland (Hg.): The immersive Internet: Reflections on the entangling of the virtual with society, politics and the economy. Houndmills, Basingstoke, Hampshire 2013, S. 247-252.
- Ders.: »Rethinking digital anthropology«, in: Heather A. Horst und Daniel Miller (Hg.): Digital anthropology. London 2012, S. 39-60.
- Ders.: »But do not identify as gay: A proleptic genealogy of the MSM category«, in: *Cultural Anthropology*, Bd. 26, Nr. 2, 2011, S. 287-312. http://dx.doi.org/10.1111/j.1548-1360.2011.01100.x.
- Ders.: Coming of age in Second Life: An anthropologist explores the virtually human. Princeton, N.J. 2008.
- Ders.: A coincidence of desires: Anthropology, queer studies, Indonesia. Durham, N.C. 2007.
- Ders.: The gay archipelago: Sexuality and nation in Indonesia. Princeton, N.J. 2007. Collins, Harry M.: Artificial experts: Social knowledge and intelligent machines. Cambridge, Mass., 1990.
- Cure, Monica: »Tweeting by mail: The postcard's stormy birth«, in: Los Angeles Times (22. Juni 2013). http://articles.latimes.com/2013/jun/22/opinion/la-oecure-postcards-twitter-20130623.
- Bowker, Geoffrey C.: »Data flakes: An afterword to ›Raw Data‹ is an oxymoron«, in: Lisa Gitelman (Hg.): »Raw data« is an oxymoron, Cambridge, Mass., 2013, S. 167-171.
- Ders.: Memory practices in the sciences. Cambridge: Mass. 2005.
- Ders. und Susan Leigh Star: Sorting things out: Classification and its consequences, Cambridge, Mass., 1999.
- boyd, danah und Kate Crawford: »Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon«, in: *Information, Communication & Society*, Bd. 15, Nr. 5, 2012, S. 662-679. http://dx.doi.org/10.1080/1369118X.2012.678878.
- Bryant, Randall E., Randy H. Katz und Edward D. Lazowska: »Big-Data computing: Creating revolutionary breakthroughs in commerce, science, and society«, in: Computing Research Consortium, 2008. www.cra.org/ccc/resources/ccc-led-white-papers/.
- Crawford, Kate: »The hidden biases in big data«, in: HBR Blog Network (1. April 2013). http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/.
- Curran, James: »Rethinking Internet history«, in: James Curran, Natalie Fenton und Des Freedman: *Misunderstanding the Internet*, London 2012, S. 34-65.
- Dourish, Paul und Genevieve Bell: Divining a digital future: Mess and mythology in ubiquitous computing. Cambridge, Mass., 2011.
- Escobar, Pepe: »Digital Blackwater rules«, in: *Asia Times* (11. Juni 2013). www. atimes.com/atimes/World/WOR-03-110613.html.
- Fabian, Johannes: Time and the other: How anthropology makes its object, New York 1983.

- Foucault, Michel: Sexualität und Wahrheit, in: Die Hauptwerke, Frankfurt a.M. 2008(a), S. 1021-1582.
- Ders.: Überwachen und Strafen, in: Die Hauptwerke, Frankfurt a.M. 2008(b), S. 701-1019.
- Frankel, Felice und Rosalind Reid: »Big data: Distilling meaning from data«, in: *Nature*, Bd. 455, Nr. 7209 (4. September 2008), S. 30. http://dx.doi.org/10.1038/455030a.
- Fraser, Nancy: »Rethinking recognition«, in: *New Left Review*, Bd. 3, 2000, S. 107-120. http://newleftreview.org/II/3/nancy-fraser-rethinking-recognition.
- Geertz, Clifford: »Thick description: Toward an interpretive theory of culture«, in: Clifford Geertz: The Interpretation of Culture: Selected Essays, New York 1973 (deutsch »Dichte Beschreibung. Bemerkungen zu einer deutenden Theorie von Kultur«, in: Clifford Geertz: Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme, Frankfurt a.M. 1983, S. 7-43).
- Gillespie, Tarleton L.: »The relevance of algorithms: The case of »spreading santorum««, in: *Internet Research* 13.0, Association of Internet Researchers (AoIR), Salford, GB (Oktober 2012).
- Ginsburg, Faye: »Rethinking the digital age«, in: David Hesmondhalgh und Jason Toynbee (Hg.): *The media and social theory*, New York 2008, S. 127-144.
- Gitelman, Lisa und Virginia Jackson: »Introduction«, in: Lisa Gitelman (Hg.): »Raw data« is an oxymoron, Cambridge, Mass., 2013, S. 1-14.
- Grosz, Elizabeth: The nick of time: Politics, evolution, and the untimely, Durham, N.C., 2004.
- Gruber, Jacob: »Ethnographic salvage and the shaping of anthropology«, in: *American Anthropologist*, Bd. 72, Nr. 6, 1970, S. 1,289-1,299. http://dx.doi.org/10.1525/aa.1970.72.6.02a00040.
- Hacking, Ian: »Making up people«, in: London Review of Books, Bd. 28, Nr. 16 (17. August 2006), S. 23-26. www.lrb.co.uk/v28/n16/ian-hacking/making-up-people.
- Halberstam, Judith: »Automating gender: Postmodern feminism in the age of the intelligent machine«, in: *Feminist Studies*, Bd. 17, Nr. 3, 1991, S. 439-460.
- Hayes, Brian: »Bit rot«, in: American Scientist, Bd. 86, Nr. 5, 1998, S. 410-415. http://dx.doi.org/10.1511/1998.5.410.
- Jakobson, Roman: »Metalanguage as a linguistic problem«, in: Roman Jakobson: *The framework of language,* Ann Arbor 1980, S. 81-92.
- Jensen, Morten Høi: »What everybody gets wrong about Orwell,« Salon (19 June), at http://www.salon.com/2013/06/19/big\_brother\_is\_the\_wrong\_meta phor\_for\_our\_time/, accessed 12 July 2013. (12. Juli 2014)
- Jensen, Klaus Bruhn: How to do things with data: Meta-data, meta-media, and meta-communication. First Monday, [S. 1.], sep. 2013. Available at: http://first monday.org/ojs/index.php/fm/article/view/4870, Date accessed: 28 Jul. 2014. doi:10.5210/fm.v18i10.4870.

- Karpf, David: »Social science research methods in Internet time«, in: Information, Communication & Society, Bd. 15, Nr. 5, 2012, S. 639-661. http://dx.doi.org/10.1080/1369118X.2012.665468.
- Kittler, Friedrich: »Towards an ontology of media«, in: Theory, Culture & Society, Bd. 26, 2009, Nr. 2-3, S. 23-31. http://dx.doi.org/10.1177/0263276409103106.
- Lee, Benjamin: Talking heads: Language, metalanguage, and the semiotics of subjectivity, Durham, N.C., 1997.
- Lévi-Strauss, Claude: »The culinary triangle«, in: Carole Counihan und Penny Van Esterik (Hg.): Food and culture: A reader, London 1997, S. 28-35.
- Ders.: Mythologica. 1. Das Rohe und das Gekochte, Frankfurt a.M. 1976.
- Lohr, Steve: »The origins of >Big Data«: An etymological detective story«, in: *New York Times* (1. Februar 2013). http://bits.blogs.nytimes.com/2013/02/01/the-origins -of-big-data-an-etymological-detective-story/.
- Ders.: »How big data became so big«, in: *New York Times* (11. August 2012). www. nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed. html/.
- Losh, Elizabeth: Virtualpolitik: An electronic history of government media-making in a time of war, scandal, disaster, miscommunication, and mistakes, Cambridge, Mass., 2009.
- Loudon, Melissa, Bill Maurer, Jessamy Norton-Ford, Martyn Fricker und Joshua Blumenstock: »Big data in ICT4D: What can we learn from prepaid mobile airtime transactions?« Proceedings of ICTD 2013: Sixth International Conference on Information and Communication Technologies and Development (7.-10. Dezember 2013, Kapstadt).
- Mainwaring, Scott und Paul Dourish: »Intel Science and Technology Center for Social Computing: White paper«, 2012. http://socialcomputing.uci.edu/ISTC-Social-Whitepaper.pdf.
- Manovich, Lev: »Trending: the promises and the challenges of big social data« (28. April 2011). www.manovich.net/DOCS/Manovich\_trending\_paper.pdf.
- Maurer, Bill: »Creolization redux: The plural society thesis and offshore financial services in the British Caribbean«, in: *New West Indian Guide*, Bd. 71, Nr. 3-4, 1997, S. 249-264.
- Mayer, Jane: »What's the matter with metadata?«, in: *New Yorker* (6. Juni 2013). www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html.
- Merlan, Philip: »On the terms >metaphysics< and >being-qua-being<«, in: Monist, Bd. 52, Nr. 2, 1968, S. 174-194. doi: http://dx.doi.org/10.5840/monist196852216.
- Nixon, Ron: »U.S. Postal Service logging all mail for law enforcement«, in: New York Times (3. Juli 2013). www.nytimes.com/2013/07/04/us/monitoring-of-snail-mail.html.
- Ohlheiser, Abby: »Twitter's jaded reaction to the NSA's phone records collection program«, in: Atlantic Wire (5. Juni 2013). www.theatlanticwire.com/technology

- /2013/06/twitters-jaded-reaction-nsas-phone-records-collection-program /65951/.
- Oxford English Dictionary (OED): »date, n.2«, 2013a. www.oed.com/.
- Oxford English Dictionary (OED): »theory, n.1.«, 2013b. www.oed.com/.
- Oxford English Dictionary (OED): »metaphysics, n. pl. [1989 edition]», 2013c. www. oed.com/.
- Peng, Tai-Quan, Lun Zhang, Zhi-Jin Zhong und Jonathan J.H. Zhu: »Mapping the landscape of Internet studies: Text mining of social science journal articles 2000-2009«, in: *New Media & Society*, Bd. 15, Nr. 5, 2012, S. 644-664. http://dx.doi.org/10.1177/0263276409103106.
- Porzucki, Nina: »NSA leak: Did George Orwell get it right in 1984?«, in: PRI's The World (12. Juni 2013). www.theworld.org/2013/06/nsa-leak-orwell-1984/.
- Povinelli, Elizabeth A.: The cunning of recognition: Indigenous alterities and the making of Australian multiculturalism. Durham, N.C., 2002.
- Räsänen, Minna und James M. Nyce: »The raw is cooked: Data in intelligence practice«, in: *Science, Technology, & Human Values*, Bd. 38, Nr. 5, 2013, S. 655-677. http://dx.doi.org/10.1177/0162243913480049.
- Roberts, John M.: »The self-management of cultures«, in: Ward Goodenough (Hg.): Explorations in cultural anthropology: Essays in honor of George Peter Murdock, New York 1964, S. 433-454.
- Rodriguez, Gabriel: »Edward Snowden interview transcript full text: Read the Guardian's entire interview with the man who leaked PRISM« in: Policymic 2013. www.policymic.com/articles/47355/edward-snowden-interview-transcript-full-text-read-the-guardian-s-entire-interview-with-the-man-who-leaked-prism.
- Romney, A. Kimball, Susan C. Weller und William H. Batchelder: »Culture as consensus: A theory of culture and informant accuracy«, in: *American Anthropologist*, Bd. 88, Nr. 2, 1986, S. 313-338. http://dx.doi.org/10.1525/aa.1986.88.2.02a00020.
- Rosenberg, Daniel: »Data before the fact«, in: Lisa Gitelman (Hg.): »Raw data« is an oxymoron, Cambridge, Mass., 2013, S. 15-40.
- Ruppert, Evelyn, John Law und Mike Savage: »Reassembling social science methods: The challenge of digital devices«, in: Theory, Culture & Society, Bd. 30, Nr. 4, 2013, S. 22-46. http://dx.doi.org/10.1177/0263276413484941.
- Schmandt-Besserat, Denise: »The envelopes that bear the first writing«, in: *Technology and Culture*, Bd. 21, Nr. 3, 1980, S. 357-385.
- Schmidt, Eric und Jared Cohen: The new digital age: Reshaping the future of people, nations and business. New York 2013.
- Searle, John R.: »Minds, brains, and programs«, in: Behavioral and Brain Sciences, Bd. 3, Nr. 3, 1980, S. 417-457.
- Silverstein, Michael: »The limits of awareness«, in: Alessandro Duranti (Hg.): Linguistic anthropology: A reader. Oxford 2001, S. 382-401.
- Taylor, Charles: "The politics of recognition", in: Charles Taylor: Multiculturalism: Examining the politics of recognition, Princeton, N.J., 1994, S. 25-73

- (deutsch: »Die Politik der Anerkennung«, in: Charles Taylor: Multikulturalismus und die Politik der Anerkennung, Frankfurt a.M. 2009, S. 11-66).
- Filippo Trevisan: »Social engines and social science: A revolution in the making« (15. Mai 2013). http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2265348.
- Trouillot, Michel-Rolph: »Anthropology and the savage slot: The poetics and politics of otherness«, in: Richard G. Fox (Hg.): Recapturing anthropology: Working in the present, Santa Fe, N.M., 1991, S. 17-44.
- Urban, Greg: Metaculture: How culture moves through the world, Minneapolis 2001. Watzlawick, Paul, Janet Helmick Beavin und Don D. Jackson: Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes, New York 1967 (deutsch: Menschliche Kommunikation. Formen, Störungen, Paradoxien, Bern, Stuttgart 1969).