

**Aus:**

*Ramón Reichert (Hg.)*

## **Big Data**

### Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie

September 2014, 496 Seiten, kart., 29,99 €,  
ISBN 978-3-8376-2592-9

NSA, Google, Facebook & Co.: Sie alle basieren auf »Big Data«, der ungeahnten Möglichkeit, riesige Datenmengen zu sammeln. »Big Data« beschreibt damit nicht nur neuartige wissenschaftliche Datenpraktiken, sondern steht für eine tektonische Verschiebung von Wissen, Medien, Macht und Ökonomie.

Im Unterschied zum Medienhype um »Big Data« schafft der Band einen Reflexionsraum zur differenzierten Auseinandersetzung mit dem datenbasierten Medienumbruch der Gegenwart. International führende Theoretiker der *Digital Humanities*, darunter David M. Berry, Jean Burgess, Alexander R. Galloway, Lev Manovich, Richard Rogers, Daniel Rosenberg, Bernard Stiegler, Theo Röhle und Eugene Thacker, stellen einen fachübergreifenden Theorierahmen zur Verfügung, der es erlaubt, »Big Data« in seiner gesamten sozialen, kulturellen, ökonomischen und politischen Bandbreite zeitdiagnostisch zu thematisieren.

**Ramón Reichert** (Dr. phil.) ist Professor für Neue Medien am Institut für Theater-, Film- und Medienwissenschaft der Universität Wien. Bei transcript u.a. erschienen: »Amateure im Netz« (2008), »Das Wissen der Börse« (2009) sowie »Die Macht der Vielen« (2013).

Weitere Informationen und Bestellung unter:  
[www.transcript-verlag.de/978-3-8376-2592-9](http://www.transcript-verlag.de/978-3-8376-2592-9)

# Inhalt

---

## **Einführung**

Ramón Reichert | 9

## **1. BIG DIGITAL HUMANITIES**

### **Licht und Schatten im digitalen Zeitalter**

Programmatische Vorlesung auf dem Digital Inquiry Symposium  
am Berkeley Center for New Media

Bernard Stiegler | 35

### **Die Computerwende – Gedanken zu den Digital Humanities**

David M. Berry | 47

### **Trending**

Verheißungen und Herausforderungen der Big Social Data

Lev Manovich | 65

### **Eine neue Betrachtung der Digital Humanities im Kontext originärer Technizität**

Federica Frabetti | 85

## **2. GESCHICHTE UND THEORIE DER DATEN**

### **Die Konstruktion von Big Data in der Theorie**

Tom Boellstorff | 105

### **Daten vor Fakten**

Daniel Rosenberg | 133

### **Big Data – Big Humanities?**

Eine historische Perspektive

Theo Röhle | 157

### **Nach dem Cyberspace: Big Data, Small Data**

Richard Rogers | 173

## **3. DIGITALE METHODEN**

### **Twitter-Archive und die Herausforderungen von »Big Social Data« für die Medien- und Kommunikationswissenschaft**

Jean Burgess/Axel Bruns | 191

### **Data-Mining von einem Prozent Twitter – Sammlungen, Basislinien, Stichproben**

Carolin Gerlitz/Bernhard Rieder | 203

### **Der Wert von Big Data für die Erforschung digitaler Medien**

Merja Mahrt/Michael Scharrow | 221

### **Twitter und Wahlen**

Zwischen 140 Zeichen und Milliarden von Tweets

Katrin Weller | 239

### **From Mind to Document and Back Again**

Zur Reflexivität von Social-Media-Daten

Johannes Paßmann | 259

## **4. DATAVEILLANCE: ALGORITHMEN, GRAPHEN UND PROTOKOLLE**

### **Protokoll, Kontrolle und Netzwerke**

Alexander Galloway/Eugene Thacker | 289

### **Der italienische Operaismo und die Informationsmaschine**

Matteo Pasquinelli | 313

### **Infodemiologie – von ›Supply‹ zu ›Demand‹**

*Google Flu Trends* und transaktionale Big Data in der epidemiologischen Surveillance

Annika Richterich | 333

## **You cannot not Transact – Big Data und Transaktionalität**

Christoph Engemann | 365

## **5. DIGITALE TECHNOLOGIEN UND SOZIALE ORDNUNGSVORSTELLUNGEN**

### **»All Watched Over by Machines of Loving Grace«**

Öffentliche Erinnerungen, demokratische Informationen und restriktive Technologien am Beispiel der »Community Memory«

Stefan Höltgen | 385

### **(Very) Nervous Systems. Big Mobile Data**

Regine Buschauer | 405

### **Facebooks Big Data**

Die Medien- und Wissenstechniken kollektiver Verdattung

Ramón Reichert | 437

### **Sozio-technische Imaginationen**

Social Media zwischen »Digital Nation« und pluralistischem Kosmopolitismus

Martin Doll | 453

### **Autorinnen und Autoren | 489**

# Einführung

---

*Ramón Reichert*

In der jüngsten Gegenwart ist »Big Data« zum populären Schlagwort aufgestiegen und wird oft als Sammelbegriff für digitale Technologien verwendet, die in technischer Hinsicht für eine neue Ära digitaler Kommunikation und Verarbeitung und in sozialer Hinsicht für einen gesellschaftlichen Umbruch verantwortlich gemacht werden. Die weit verbreitete Popularität des Buzzwords »Big Data« kann als Hinweis verstanden werden, dass digitale Technologien nicht nur als neutrale Übermittler und passive Objekte firmieren, sondern in einer engen Verbindung mit Diskursen, Akteuren und Netzwerken stehen, die Gesellschaft und Subjekte verändern.

Bisher haben vor allem die Branchen und Anwendungsbereiche von Wirtschaft, Marktforschung, Vertriebs- und Servicesteuerung, Medizin, Verwaltung und Nachrichtendienst die digitalen Methoden und Technologien, die zur Erhebung, Analyse und Modellierung von großen Datenmengen eingesetzt werden, weiterentwickelt und nutzbringend eingesetzt. Dem Sammeln großer Datenmengen ist selbst eine Machtgeschichte der möglichen Herstellung sozialprognostischen Wissens inhärent. An der Schnittstelle von konzernorientierten Geschäftsmodellen und gouvernementalem Handeln experimentieren Biotechnologie, Gesundheitsprognostik, Arbeits- und Finanzwissenschaften, Risiko- und Trendforschung in ihren Social Media-Analysen und Webanalysen mit Vorhersagemodellen von Trends, Meinungsbildern, Stimmungen oder kollektivem Verhalten.

Im Unterschied zum Medienhype um »Big Data«, der seine Entstehung der Möglichkeit zur kollektiven Adressierung durch Massenmedien verdankt, versucht der vorliegende Sammelband, einen Reflexionsraum zur differenzierten Auseinandersetzung mit dem datenbasierten Medienumbruch der Gegenwart zu schaffen. Der Band stellt einen fachübergreifenden Theorierahmen zur Verfügung, der es erlaubt, Big Data in seiner gesamten sozialen, kulturellen, ökonomischen und politischen Bandbreite zu reflektieren. Eine medien- und kulturwissenschaftliche Betrachtung der digitalen Großforschung distanziert sich von den normativen Diskursen des Daten- und Informationsmanagements, welche

die Ansicht vertreten, dass die Entwicklung der spätmodernen Gesellschaften einerseits von der zunehmenden Verbreitung und alltäglichen Nutzung Sozialer Medien und andererseits von der Aussagefähigkeit von Großdaten abhängig ist.

Die Big Data-Research differenziert sich innerhalb der letzten Jahre erheblich aus: Zahlreiche Studien betreiben mithilfe maschinenbasierter Verfahren wie der Textanalyse (quantitative Linguistik), der Sentimentanalyse (Stimmungserkennung), der sozialen Netzwerkanalyse oder der Bildanalyse vielschichtige Social Media-Analysen. Damit sollen *Wahlentscheidungen* (Gayo-Avello, 2012), *politische Einstellungen* (Conover et al., 2012), *Finanztrends und Wirtschaftskrisen* (Gilbert und Karahalios, 2010; Zhang, 2010), *Psychopathologien* (Wald et al., 2012) und *Aufstände und Protestbewegungen* (Yogatama, 2012) frühzeitig vorhergesagt werden können: »Analysts and consultants argue that advanced statistical techniques will allow the detection of on-going communicative events (natural disasters, political uprisings) and the reliable prediction of future ones (electoral choices, consumption).« (Burgess und Puschmann, 2013, S. 4) Von einer systematischen Auswertung der Big Data erwarten sich die Prognostiker eine effizientere Unternehmensführung bei der statistischen Vermessung der Nachfrage- und Absatzmärkte, individualisierte Serviceangebote und eine bessere gesellschaftliche Steuerung. So wertet etwa Google Flu in Zusammenarbeit mit der US-amerikanischen Seuchenbehörde CDC die Häufigkeit bestimmter Suchbegriffe aus, um mit diesem Data Mining die Wahrscheinlichkeit von Grippeerkrankungen und deren Ausbreitung vorherzusagen zu können. Zahlreiche Studien beschäftigen sich mit den Auswertungen von Tweets, die Korrelationen zwischen realen Preiserhöhungen und Stimmungen vermessen, Mobilfunkkonzerne und Computerhersteller erfassen die Bewegungsdaten von Kunden, um neue Ansätze für Dienste und Werbevermarktung zu erschließen, und das Trendmining der Social Media-Analysen versucht, Entwicklungen auf den Absatz-, Finanz- oder Arbeitsmärkten frühzeitig zu erkennen. Big Data resultieren nicht nur aus den Onlinedaten, den Kundendaten und den nutzergenerierten Inhalten als »unstrukturierte Daten« (Inhalte von Emails und SMS-Nachrichten etc.), sondern entstehen gleichermaßen an den Schnittstellen der automatisierten Datenverarbeitung. In diesem Sinne spricht man in der angewandten Netzforschung von transaktionalen Nutzerdaten, die durch Webtracking, Handy-Monitoring oder Sensorerfassung entstehen (Boble, 2011).

Wenn man es so betrachtet, ist das Social Web zur wichtigsten Datenquelle zur Herstellung von Regierungs- und Kontrollwissen geworden. Auch die politische Kontrolle sozialer Bewegungen hat sich in das Netz verschoben, wenn Soziologen und Informatiker gemeinsam etwa an der Erstellung eines *Riot Forecasting* mitwirken und dabei auf die gesammelten Textdaten von Twitter-Streams zugreifen: »Due to the availability of the dataset, we focused on riots in Brazil. Our datasets consist of two news streams, five blog streams, two Twitter streams (one for politicians in Brazil and one for general public in Brazil), and one stream of 34 macroeconomic variables related to Brazil and Latin America.« (Yogatama, 2012, S. 3)

Seit dem späten 20. Jahrhundert zählen die digitale Großforschung und ihre großen Rechnerzentren und Serverfarmen zu den zentralen Bausteinen der Herstellung, Verarbeitung und Verwaltung von informatischem Wissen. Damit einhergehend rücken mediale Technologien der Datenerfassung und -verarbeitung und Medien, die ein Wissen in Möglichkeitsräumen entwerfen, in die Mitte der Wissensproduktion und der sozialen Kontrolle. In ihrer Einleitung in das »Routledge Handbook of Surveillance Studies« knüpfen die Herausgeber Kirstie Ball, Kevin Haggerty und David Lyon einen Zusammenhang zwischen technologischer und sozialer Kontrolle auf der Grundlage der Verfügbarkeit großer Datenmengen: »Computers with the Power to handle huge datasets, or ›big data«, detailed satellite imaging and biometrics are just some of the technologies that now allow us to watch others in greater depth, breadth and immediacy.« (2012, S. 2) In diesem Sinne kann man sowohl von *datenbasierten* als auch von *datengesteuerten* Wissenschaften sprechen, da die Wissensproduktion von der Verfügbarkeit computertechnologischer Infrastrukturen und der Ausbildung von digitalen Anwendungen und Methoden abhängig geworden ist.

Damit einhergehend verändern sich auch maßgeblich die Erwartungen an die Wissenschaft des 21. Jahrhunderts und es werden zunehmend Forderungen laut, die darauf bestehen, die historisch, kulturell und sozial einflussreichen Aspekte der digitalen Datenpraktiken systematisch aufzuarbeiten – verknüpft mit dem Ziel, diese in den künftigen Wissenschaftskulturen und Epistemologien der Datenerzeugung und -analyse zu verankern. Diesen fachübergreifenden Ansprüchen sieht sich der vorliegende Band verpflichtet und versammelt interdisziplinäre Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie, die den aktuellen Debatten zum Stellenwert der Big Data eine theoretische, kritische und historische Tiefe verleihen wollen.

Die heterogenen Anwendungsfelder der datengetriebenen Wissenschaftskulturen zeigen auf, dass das Datenwissen aus konkreten Praktiken der Erzeugung entsteht, die eine nuancierte Genealogie der datenintensiven Formen der Wissensproduktion nahelegt. Andererseits muss auch eingeräumt werden, dass die Repräsentation und die popularisierende Vermittlung der datengenerierten Forschung auch Anschlüsse auf frühere materielle Datenkulturen eröffnen, denen sowohl historische Kontinuitäten als auch mediale Umbrüche inhärent sind und die nur dann verständlich werden, wenn sie im historischen, sozialen und kulturellen Kontext reflektiert werden können (vgl. Gitelman und Pingree, 2004). Eine vergleichende Analyse der Datenverarbeitung unter Berücksichtigung der materiellen Kultur von Datenpraktiken vom 19. bis zum 21. Jahrhundert vermag aufzuzeigen, dass bereits im 19. Jahrhundert die mechanischen Datenpraktiken das taxonomische Erkenntnisinteresse der Forscher maßgeblich beeinflussten – lange bevor es computerbasierte Methoden der Datenerhebung gab (vgl. Driscoll, 2012). Weiterführende Untersuchungen erarbeiten die sozialen und politischen Bedingungen und Auswirkungen des Übergangs von der mechanischen Datenauszahlung der ersten Volkszählungen um 1890 über die elektronischen Daten-

verarbeitungen der 1950er Jahre bis zum digitalen Social Monitoring der unmittelbaren Gegenwart (Beniger, 1986; Berry, 2011).

Das *erste Kapitel* »Big Digital Humanities« thematisiert die kulturellen Transformationen und die medialen Umbrüche der digitalen Medienkultur. Die Digital Humanities stellen ein heterogenes Forschungsfeld an der Schnittstelle zwischen der Informatik und den Kultur- und Geisteswissenschaften dar. In den letzten Jahren haben sie vor dem Hintergrund der gestiegenen Verfügbarkeit digitaler Daten stark an Bedeutung gewonnen. Der Begriff der »Digital Humanities« setzte sich erst mit der Verbreitung des Internets durch und ersetzte die seit dem Aufkommen des Computers in den 1960er Jahren gebräuchlichen Begriffe der »Computational Science« und »Humanities Computing«, die sich hauptsächlich mit der methodologischen und praktischen Entwicklung von digitalen Werkzeugen, Infrastrukturen und Archiven beschäftigt haben.

In Anknüpfung an die wissenschaftstheoretischen Sondierungen von Davidson (2008), Svensson (2010), Burdick et al. (2012) und zuletzt Gold (2012) können die Digital Humanities dementsprechend in drei richtungsweisende Theorieansätze gegliedert werden, mit welchen gleichermaßen die historischen Entwicklungen und wissenspolitischen Verschiebungen des Forschungsfeldes gut abgesteckt werden können:

(1) Im Zentrum der Digital Humanities steht die geistes- und kulturwissenschaftliche Computeranwendung und die Digitalisierung von »Ausgangsdaten«. Die Digitalisierungsprojekte umfassen dabei einerseits digitalisierte Portfolios und andererseits die computerphilologischen Nutzungstools zur Anwendung der Sekundär- oder Ergebnisdaten. Diese elementaren Methoden der Digital Humanities speisen sich bis heute aus der Tradition der Philologie, die das evidenzbasierte Sammeln und Ordnen der Daten zur Grundlage von Hermeneutik und Interpretation erklärte. Über die engere Methodendiskussion hinausführend, reklamiert das computergestützte Vermessen der Geistes- und Kulturwissenschaften mediale Objektivitätspostulate der modernen Wissenschaften. Im Unterschied zum textwissenschaftlichen Fächerkanon der 1950er und 1960er Jahre des »Humanities Computing« (McCarty, 2005) hat sich das Forschungsfeld der Fachdisziplinen indes heute erheblich ausdifferenziert und auf die Kunst-, Kultur- und Sozialgeschichte, die Medienwissenschaft, die Ethnologie, die Archäologie und die Geschichts- und Musikwissenschaften ausgeweitet (Gold, 2012).

(2) Entsprechend der zweiten Phase entwickeln sich zusätzlich zur quantitativen Digitalisierung von Texten Forschungspraktiken, die sich mit den Methoden und Verfahren zur Herstellung, Analyse und Modellierung von digitalen Forschungsumgebungen für die geisteswissenschaftliche Arbeit mit digitalen Daten beschäftigen. Dieser Ansatz wird mit dem Begriff »Enhanced Humanities« umschrieben und versucht, neue methodische Ansätze der qualitativen Nutzung von erzeugten, aufbereiteten und gespeicherten Daten für die Rekonzeptionalisierung traditioneller Forschungsgegenstände zu finden.



(3) Die Entwicklung der »Humanities 1.0« zu den »Humanities 2.0« (Davidson, 2008, S. 707-717) markiert den Übergang von der digitalen Methodenentwicklung im Bereich der »Enhanced Humanities« zu den »Social Humanities«, welche die Möglichkeiten des Web 2.0 zum Aufbau vernetzter Forschungsinfrastrukturen nutzen. Die »Social Humanities« nutzen zur Vernetzung und Interdisziplinarität des wissenschaftlichen Wissens Open-Access-, Social-Reading- und Open-Knowledge-Software, indem sie die Möglichkeiten kooperativer und kollaborativer Forschungs- und Entwicklungsarbeit für die Wissenschaft nutzbar machen und online zur Verfügung stellen. Auf der Basis der neuen digitalen Infrastrukturen des Social Web (Hypertextsysteme, Wikitools, Crowdfunding-Software u.Ä.) überführen diese Projekte die rechnergestützten Verfahren der früheren Entwicklungsgeschichte der Digital Humanities in die digitale Vernetzungskultur der »Social Humanities«. Heute sind es die *Blogging Humanities* (digitale Publikations- und Vermittlungsarbeit in Peer-to-Peer-Netzwerken) und die *Multimodal Humanities* (Wissensinszenierung- und repräsentation in multimedialen Softwareumgebungen), die für die technische Modernisierung des akademischen Wissens eintreten (McPherson, 2008) und dafür gesorgt haben, dass die *Digital Humanities* beanspruchen, eine paradigmatisch alternative Form der Wissensgenerierung darzustellen. Vor diesem Hintergrund ist es angebracht, die kultur- und medientechnischen Grundlagen der computergesteuerten Erkenntnisverfahren in der geistes- und kulturwissenschaftlichen Forschung datenkritisch, wissensgenealogisch und medienhistorisch zu reflektieren, um ihr Rollenverständnis unter den Vorzeichen der digitalen Wissensproduktion und -distribution angemessen beurteilen zu können (Thaller, 2012, S. 7-23).

In dem hier abgedruckten Vortrag über digitale Techniken am *Berkeley Center for New Media* vertritt der Medienphilosoph Bernard STIEGLER die These, dass man die digitalen Technologien als das letzte Stadium der Schriftkultur verstehen könne.<sup>1</sup> Diese neue Ära der Schriftkultur nennt Stiegler in Abgrenzung zur ideografischen und alphabetischen Schrift die digitale Schrift, die er als mnemotechnisches Instrument begreift, das in der Lage sei, in uns eine neuartige soziotechnisch vermittelte Gehirnfunktion zu konstituieren. Insofern schreibt er in seiner *Digital Studies-Anthropologie* der digitalen Technik eine organologische Wirkungsmacht zu.

Der englische Medientheoretiker David M. BERRY ist Herausgeber des vielbeachteten Sammelbandes »Understanding Digital Humanities«, der 2012 erschienen ist und in fachübergreifenden Studien die Umbruchpotentiale der Verwendung quantitativer Methoden in den Geisteswissenschaften sondiert. Sein Beitrag mit dem Titel »Die Computerwende – Gedanken zu den Digital Humanities« untersucht die Medienspezifität der digitalen Technologien und vertritt die These, dass die medialen Umbrüche des »datenzentrierten Zeitalters« nicht

---

1 | Jede(r) Autor(in) verfügt selbst über Verwendung und Art der geschlechtersensiblen Schreibung.

nur epistemische Veränderungen erzeugen, sondern auch eine neue Art von computergestützter Subjektivität konstituieren, die er in der Schlüsselidee einer ›digitalen Universität‹ verortet. Berrys Konzept computergestützter Subjektivität distanziiert sich aber von der Annahme einer hypostasierten Selbstbezüglichkeit, in der Medien lediglich als Werkzeuge zur Darstellung eines lebensweltlich bereits gegebenen Subjekts betrachtet werden, und lenkt den Blick auf das, was in den medialen Analysen der Subjektivität mit dem Analysebegriff »Dispositiv« beschrieben wird. Die Frage nach dem Code ist für ihn von zentraler Bedeutung, um die Digital Humanities als mediale Anordnung und als eine »Möglichkeitenbedingung für die vielen computerbasierten Formen, die unsere Erfahrung von zeitgenössischer Kultur und Gesellschaft vermitteln,« verstehen zu können.

Im Unterschied zu STIEGLER und BERRY, die sich mit den anthropologischen Transformationen digitaler Medienkultur auseinandersetzen, reflektiert der US-amerikanische Theoretiker der Digitalen Medien, Lev MANOVICH, in seinem vielzitierten Essay »Trending: Verheißungen und Herausforderungen der Big Social Data« die sozialen Auswirkungen der Datenforschung und warnt vor einem ›Digital Divide‹, der das Datenwissen einseitig verteilt und zu Machtasymmetrien zwischen Forschern *innerhalb* und *außerhalb* der Netzwerke führen könnte. Manovich hat den limitierten Zugang zu sozialstatistischen Daten bereits 2008 kritisiert, der von vornherein eine monopolartige Verfügbarkeit von Datenwissen schafft: »Nur Social-Media-Unternehmen haben Zugang zu wirklich großen sozialen Daten – insbesondere zu Transaktionsdaten. Ein Anthropologe, der für Facebook arbeitet, oder ein Soziologe in den Diensten von Google wird Zugang zu solchen Daten haben, die übrige wissenschaftliche Community hingegen nicht.« (Abgedruckt im vorliegenden Band.) Dieses ungleiche Verhältnis festigt die Stellung der sozialen Netzwerke als computerbasierte Kontrollmedien, die sich Datenwissen entlang einer vertikalen und eindimensionalen Netzkommunikation aneignen: (1) Sie ermöglichen einen kontinuierlichen Fluss von Daten (digitale Fußabdrücke); (2) sie sammeln und ordnen diese Daten; und (3) sie etablieren geschlossene Wissens- und Kommunikationsräume für Expert/innen und ihre Expertisen, welche die kollektiven Daten zu Informationen verdichten und interpretieren.

In der Nachfolge haben danach boyd und Kate Crawford die von Manovich eingebrachte Kritik am technisch-infrastrukturellen Aufbau der Netzwerkkommunikation weitergeführt und auf die hierarchisch und pyramidal angeordnete Schichtung der Sozialen Medien bezogen: »The current ecosystem around Big Data creates a new kind of digital divide: the Big Data rich and the Big Data poor. Some company researchers have even gone so far as to suggest that academics shouldn't bother studying social media data sets – Jimmy Lin, a professor on industrial sabbatical at Twitter argued that academics should not engage in research that industry ›can do better‹.« (boyd und Crawford, 2011) Das asymmetrische Verhältnis zwischen dem Kommunikationsraum gewöhnlicher Nutzer/innen innerhalb der Anwendungsschicht (Benutzeroberfläche im ›Frontend‹) und dem

hauptsächlich von Informatikern verwalteten Datenraum im Backend-Bereich (mit dem Begriff ›Backend‹ wird die auf dem Server laufende Administrationsoberfläche verstanden, die ausschließlich der Dateneingabe, -verarbeitung und -weitergabe durch den Seitenbetreiber vorbehalten ist) wird in Anlehnung an die genannten Autoren und Autorinnen auch als »Participatory Gap« (vgl. Taewoo und Stromer-Galley, 2012, S. 133-149) diskutiert.

Eine kritische Sondierung des Stellenwerts der digitalen Methoden in den Geisteswissenschaften nimmt auch die an der *Brookes University* in Oxford lehrende Medientheoretikerin Federica FRABETTI vor. Sie kritisiert im Anschluss an Johanna Drucker (2009) das Objektivitätspostulat der Digital Humanities und stellt in diesem Zusammenhang die oft behauptete Bipolarität von Technik und Geist in Frage. Ihre Hauptthese ist, dass computerbasierte Wissensproduktion und Geisteswissenschaften nicht als zwei getrennte Sphären aufgefasst werden müssen, deren Beziehungen sich ein für allemal definieren lassen. Im Zentrum ihrer erkenntnis- und wissenskritischen Reflexion der Digital Humanities steht die Kritik an der instrumentellen Ausrichtung der digitalen Methoden in den Geisteswissenschaften: »Wenn akademische Arbeit der Instrumentalität widerstehen muss, um politisch zu bleiben, dann werden die Digital Humanities zum idealen Ort für eine beharrliche Kritik an allen instrumentellen Denkweisen.« (Abgedruckt im vorliegenden Band.) In diesem Zusammenhang arbeitet sie in produktiver Auseinandersetzung mit den Schriften von Derrida, Leroi-Gourhan, Kittler und Stiegler heraus, dass es eine der Stärken der Digital Humanities sein könnte, überkommene begriffliche Bezugssysteme im Hinblick auf die digitale Technik zu überdenken und die Frage nach den Beziehungen zwischen Geisteswissenschaften und Digitalität und allgemeiner zwischen Mensch und Technik offen zu halten.

Im *zweiten Kapitel* werden die geschichtlichen Aspekte der Daten und die Theorien des Datenbegriffs erörtert. Der an der *University of California* lehrende Anthropologe Tom BOELLSTORFF beschäftigt sich mit den theoretischen Grundlagen von Big Data und untersucht die begriffsgeschichtlichen, rhetorischen und metaphorischen Kontexte der Data Studies. In diesem Zusammenhang untersucht er die forschungspolitisch-strategischen Verwendungsweisen der Diskursfigur »Big Data« und zeigt die Vormachtstellung von »räumlichen Metaphern der Mobilität und der Omnipräsenz« in den einschlägigen Diskussionen auf. Seine etymologisch und wissenschaftsgeschichtlich motivierten Begriffsanalysen von Daten und Metadaten zeigen auf, dass die jeweiligen konkreten Verbindungen zwischen Technologien, Kommunikationspraktiken und Epistemologien stets ihre eigene Komplexität aufweisen. Während die Technologien sich ändern, bleiben die Praktiken zur Erzeugung von wissenschaftlichem Wissen und die Diskurse über die Organisation und die Repräsentation von Daten länger konstant. Seine wissensgeschichtlichen und -kritischen Einwände weitet Boellstorf auch auf datenkritische Positionen aus und kann aufzeigen, dass auch Datenkritik von rhetorischen Vereinfachungen und traditionellen Dichotomien geprägt ist. In die-

sem Zusammenhang setzt er sich mit dem einflussreichen Sammelband »Raw data< is an oxymoron« auseinander, der 2013 von Lisa Gitelman und Virginia Jackson herausgegeben wurde. Erst in den letzten Jahren sind neben überwiegend affirmativen Aufladungen des Schlagworts Big Data auch einige kritische Aufsätze und Anthologien zur Klärung des Datenbegriffs vorgelegt worden (vgl. exemplarisch Ballsun-Stanton, 2010, S. 119-24; Gitelman und Jackson, 2013; Hagner und Hirschi, 2013; Floridi, 2014). Die von Gitelman und Jackson normativ geäußerte Einschätzung, dass »Rohdaten sowohl ein Oxymoron wie eine schlechte Idee sind« und daher »Daten sorgfältig gekocht werden sollten« (2013, S. 3), nimmt Boellstorf zum Anlass, sich näher mit Lévi-Strauss' Begriffsoption »Rohes/Gekochtes« (1964) zu beschäftigen. Dabei versucht er, die Dichotomie von rohen und gekochten Daten zu überwinden und in einer Relektüre des Aufsatzes »Das kulinarische Dreieck«, der ein Jahr nach »Das Rohe und das Gekochte« erschien, vielmehr ein präzise definiertes Dreiecksverhältnis zwischen »Rohem«, »Gekochtem« und »Verrottetem« in Aussicht zu stellen, welches er für besonders befähigt hält, als begriffliches Werkzeug bei der Erstellung von Big Data dienen zu können. Mit dem Begriff des »Verrotteten«, der die »Materialität von Daten« anzeigt, wäre es nach Boellstorf möglich, dass im »Kontext von roh und gekocht das ›Verrottete‹ Verwandlungen außerhalb der typischen Konstruktionen des handelnden Menschen als Koch versammelt – als das Ungeplante, Unerwartete und Zufällige. Bit Rot beispielsweise entsteht aus dem Zusammenwirken von Speicher- und Verarbeitungstechnologien, während diese sich durch die Zeit bewegen.«

Die medien- und wissenschaftsgeschichtliche Aufarbeitung der Digital Humanities wurde erst vereinzelt in den letzten Jahren eingeleitet (Hockey, 2004). Eine Historisierung der Wechselbeziehungen zwischen den Geistes- und Kulturwissenschaften und dem Einsatz der computergestützten Verarbeitungsverfahren relativiert den Evidenz- und Wahrheitsanspruch der digitalen Methoden und stärkt die Argumentation, dass sich die Digital Humanities aus einem Geflecht historischer Wissenskulturen und Medientechnologien entwickelt haben, deren Anfänge bis in das ausgehende 19. Jahrhundert reichen. Auch der Beitrag von Daniel ROSENBERG mit dem Titel »Daten vor Fakten« konfrontiert die Big Data-Forschung mit historischen Perspektivierungen und stellt die digitale Modernität des Datenbegriffs in Frage. Er verlagert die Diskussion um den Datenbegriff in das 17. Jahrhundert, die er eng verknüpft mit der »Entwicklung moderner Konzepte von Wissen und Argumentation« sieht. Seine Ausgangsfrage lautet: »Was also waren Daten vor dem 20. Jahrhundert?« Im Zentrum seines Aufsatzes steht also der vielversprechende Anspruch, die Geschichte des Datenbegriffs zu rekonstruieren, um seine entscheidende Rolle bei der Herausbildung neuer Wissensräume und -praktiken aufzeigen zu können. In diesem Zusammenhang interessiert er sich für die Frage der historischen Bedingungen, die für die moderne Semantisierung der Daten als ›Wahrheit‹, ›Faktum‹ oder ›Beweis‹ in Betracht gezogen werden können. Schließlich führt er den interessanten Nachweis, dass die zeitgenössischen digitalen Techniken der Gegenwart (Google Ngram Viewer, Google

Books) als ein Echo der historischen Verfahren der Verarbeitung und Visualisierung von Wissens seit dem 17. und 18. Jahrhundert aufgefasst werden können.

Die einschlägige Forschungsliteratur zum historischen Kontext und zur Genese der Digital Humanities sieht in Roberto Busas lochkartenbasierter Thomas von Aquin-Konkordanz einer der ersten Projekte genuin geisteswissenschaftlicher Computeranwendung (Vanhoutte, 2013, S. 126). Der italienische Jesuitenpater Roberto Busa (1913-2011) gilt vielen als der Pionier der Digital Humanities, was bereits früh zu einer weitgehenden Vereinheitlichung der Geschichtsschreibung der »Computational Science« geführt hat (Schischkoff, 1952). Busa, der seit 1949 gemeinsam mit Thomas J. Watson, dem Gründer von IBM, den korpuslinguistischen »Index Thomisticus« entwickelte (Busa, 1951; 1980, S. 81-90), wird in der einschlägigen Historiografie als »Gründungsvater« der Schnittstelle von Geisteswissenschaften und Informatik angesehen. Die erste digitale Edition auf Lochkarten initiierte eine Reihe von textphilologischen Folgeprojekten: »In den sechziger Jahren folgte eine erste elektronische Fassung der ›Modern Language Association International Bibliography‹ (MLAIB), einer periodischen Fachbibliografie aller modernen Philologien, die mit Telefonkopplern durchsucht werden konnte. Dann setzte die retrospektive Digitalisierung des kulturellen Erbes ein, die immer mehr Werke und Lexika wie etwa das Deutsche Wörterbuch der Brüder Grimm, historische Wörterbücher wie den Krünitz oder Regionalwörterbücher umfasste.« (Lauer, 2013, S. 104)

Im Vorfeld haben sich aber zahlreiche andere Disziplinen und nichtphilologische Bereiche herausgebildet, wie die Literatur-, Bibliotheks- und Archivwissenschaften, die eine längere Wissensgeschichte im Feld der philologischen Case Studies und der praktischen Informationswissenschaft aufweisen und sich seit dem Aufkommen der Lochkartenmethode mit quantitativen und informatikwissenschaftlichen Verfahren für wissensverwaltende Einrichtungen befassen. Kevin Driscoll hat die Genealogie der Datenverarbeitung großer Datenmengen untersucht und unterscheidet dabei drei historische Perioden:<sup>2</sup>

»The first period begins in the late-19th century with the development of mass-scale information processing projects and the electro-mechanical punched card systems that made them possible. Although these early machines were gradually replaced by programmable computers in the 1950s and 1960s, the organizational logic embedded in such systems persisted more or less unchanged until the 1970s. The second period is marked by the rise of database populism and the increasing availability of microcomputers in the late-1970s. Implementations of the relational data model enabled the production of more accessible interfaces for non-specialists and large institutional databases were increasingly accompanied by small personal databases built by individuals and stored on microcomputers. In the third period, however, small personal databases receded from the desktop with the increasing sophistication of spreadsheet software and the diffusion of internet access.

---

2 | Diesen Hinweis verdanke ich Johannes Paßmann.

In the early 21<sup>st</sup> century, the demanding task of tracking millions of users through highly-centralized communication systems such as Facebook brought about new approaches to database design that departed significantly from the previous four decades.« (Driscoll, 2012, S. 6f.)

Die von Driscoll etablierte Geschichtsperiodisierung der Big Data bildet einen vielversprechenden Ansatz, um die sozialen und historischen Auswirkungen der Informationsverarbeitung mit sehr großen Datenmengen aufzeigen zu können. So gesehen sind weder die Forschungsfrage noch die methodischen Verfahren von Busa voraussetzungslos und können in einer breiter angelegten Wissensgeschichte und Medienarchäologie verortet werden. So finden sich bereits in den Schriften des Schweizer Archivars Karl Wilhelm Bührer (1861-1917) Modellentwürfe eines mechanisch kombinierungsfähigen Wissensapparates (Bührer, 1890, S. 190-92). Diese Denkfigur flexibilisierter und modularisierter Informationseinheiten wurde später zum konzeptionellen Inbegriff der mechanischen Datenverarbeitung. Die Archiv- und Bibliothekswissenschaften waren auch direkt am historischen Paradigmenwechsel der Informationsverarbeitung beteiligt. Denn es ist der Arzt und spätere Direktor der *National Medical Library*, John Shaw Billings, der die von Hermann Hollerith im Jahr 1886 entwickelte Apparatur zur maschinellen Verarbeitung von statistischen Daten mitentwickelte (Krajewski, 2007, S. 43). Folglich verdankt die Lochkartentechnik ihre Entstehung also auch teilweise der technischen Pragmatik der bibliothekarischen Wissensorganisation; auch wenn erst später – unter den Vorzeichen der Rationalisierungsbewegung der 1920er Jahre – die bibliothekarischen Arbeitsverfahren gezielt mechanisiert wurden. Die weiterführenden Projekte der Datenverarbeitung zeigen, dass die maschinelle Erstellung eines *Index* oder einer *Konkordanz* für die Lexikografie und für den Katalogapparat der Bibliotheken den Beginn der computergestützten Geistes- und Kulturwissenschaften markiert. Im Zentrum der ersten Anwendungen der maschinellen Methode zur Erschließung großer Textmengen stand bis in die späten 1950er Jahre das Hollerith-Verfahren der Lochkarteisysteme. Das medientechnische Verfahren der Lochkarte veränderte die Lektürepraktiken der Texterschließung, indem es das Medium Buch in eine Datenbank transformierte und die linear-syntagmatische Ordnung des Textes in eine sachliche und begriffliche Systemordnung übersetzte. Bereits im Jahr 1951 etablierte sich eine zeitgenössische Debatte in akademischen Zeitschriften. Sie beurteilte die Anwendungsmöglichkeiten der Lochkarteisysteme überwiegend positiv und verknüpfte sie mit ökonomisch motivierten Rationalitätsvorstellungen.

In seinem Aufsatz »Big Data – Big Humanities? Eine historische Perspektive« nutzt Theo RÖHLE die oben angesprochene historische Perspektivierung der Debatten im Umfeld von Big Data und Humanities Computing für eine kritische Auseinandersetzung mit den Ansprüchen und Versprechen der Digital Humanities. Vor diesem Hintergrund beschäftigt er sich mit historischen Debatten, die bereits in der naturwissenschaftlich dominierten Ära der »Big Science« (de Price,

1963) der 1960er Jahre über die Verwendung quantifizierender Methoden in den Geschichtswissenschaft geführt wurden. Für ihn bietet eine quellenkritische Auseinandersetzung mit dieser Streitfrage die Möglichkeit, einerseits Parallelen zum aktuellen Diskurs in den Digital Humanities zu ziehen und andererseits problematische Behauptungen, die von einem Paradigmenwechsel (Kuhn, 1962) in der Geschichtswissenschaft sprechen, in die Schranken zu weisen.

Die Historisierung von Big Data vermag aufzuzeigen, dass der historische Stellenwert der Arbeit mit Großdaten und der Diskurs der Großforschung nicht alleine aus der Tradition der Humanities abgeleitet werden kann, da die Entstehung und die Genese der Großforschung mit militärstrategischen und nationalstaatlichen Interessen und Institutionen verknüpft war (Leslie, 1993). Während des Kalten Kriegs erhielt die »Big Science« einen entscheidenden Impuls und erfreute sich in den USA der frühen 1960er Jahre großer Popularität. In dieser Aufbruchsituation wurden im Rahmen des Internationalen Geophysischen Jahres (1957-1958) Weltdatenzentren (World Data Centers) geschaffen, die bis heute einen richtungsweisenden Stellenwert für die Erzeugung weltweiter Datennetze und die Organisation großer Datenbanken einnehmen (vgl. Aronova, Baker und Oreskes, 2010, S. 183-224).

Die automatischen Auswertungsverfahren von Texten für die literaturwissenschaftliche Editorik wurde in der frühen Phase des »Humanities Computing« (vorangetrieben durch dessen Teilbereiche »Computerphilologie« und »Computerlinguistik«) auf der Grundlage von zwei zentralen Diskursfiguren beschrieben, die bis heute gültig sind. Die eine Diskursfigur beschreibt die Errungenschaften des neuen Werkzeuggebrauchs mit der instrumentellen Verfügbarkeit der Daten (»Hilfsmittel«), die andere fokussiert die ökonomische Erschließung der Daten und betont das Effiziente und Effektive der maschinellen Methode bei der Dokumentation. Die mediale Figur der Automatisierung wurde schließlich in beiden Fällen mit der Erwartung zusammengeführt, interpretative und subjektive Einflüsse aus der Informationsverarbeitung und Informationserschließung systematisch auszuschließen. In den 1970er und 1980er Jahren etablierte sich die Computerlinguistik als ein institutionell verankerter Forschungsbereich inklusive universitärer Einrichtungen, Vereinszeitschriften (*Journal of Literary and Linguistic Computing*, *Computing in the Humanities*), Diskussionsforen (HUMANIST) und Konferenzaktivitäten. Das rechnergestützte Arbeiten in der historisch-sozialwissenschaftlichen Forschung erlebte damit einen ersten großen Aufschwung, wurde aber in den Arbeitsberichten weniger als eigenständige Methode, sondern hauptsächlich als ein Hilfsmittel bei der Textkritik und als eine Erleichterung bei der Quantifizierung der jeweiligen Gegenstände angesehen (Jarasch, 1976, S. 13).

Ein nachhaltiger Medienumbruch sowohl im Bereich der Produktion als auch im Bereich der Rezeptionsästhetik ergab sich mit der Verwendung von standardisierten Textauszeichnungen wie der 1986 etablierten *Standard Generalized Markup Language* und softwarebasierten Textverarbeitungsprogrammen. Sie stellten

eine zusätzliche Reihe digitaler Module, analytischer Werkzeuge und Textfunktionen zur Verfügung und transformierten den Text in ein Datenbankmodell. Texte konnten unter diesen Vorzeichen als strukturierte Informationen abgerufen werden und wurden optional als (relationale) Datenbanken verfügbar. In den 1980er und 1990er Jahren wurde die technische Entwicklung und die Textwahrnehmung also weitgehend vom Datenbank-Paradigma beherrscht.

Mit der schrittweisen Durchsetzung des World Wide Web veränderten sich die Forschungs- und Lehrpraktiken der Kultur- und Geisteswissenschaften auf radikale Weise: Die Fachkommunikation erhielt durch die digitale Vernetzungskultur öffentlich zugänglicher Online-Ressourcen, Email-Verteiler, Chats und Foren eine breite Dynamik und ist durch die mediale Feedbackmentalität der Rankings und der Votings hochgradig responsiv geworden. Mit ihrem Anspruch, die hierarchischen Strukturen des akademischen Systems durch ein Re-Engineering des wissenschaftlichen Wissens zu überwinden, sorgen die Digital Humanities 2.0 in diesem Sinne für eine Wiederbelebung der Ideale von Egalität, Freiheit und Allwissenheit.

Im Unterschied zu ihren Anfängen in den 1950er Jahren vertreten die Digital Humanities heute auch den Anspruch, das gesellschaftliche Wissen neu zu organisieren, und verstehen sich daher »sowohl als wissenschaftliches wie auch als sozialutopisches Projekt« (Hagner und Hirschi, 2013, S. 7). Mit dem Einzug der Social Media in die Geistes- und Kulturwissenschaften haben sich nicht nur die technologischen Möglichkeiten und die wissenschaftlichen Praktiken der Digital Humanities weiterentwickelt, sondern sie haben auch neue Phantasmagorien wissenschaftlicher Distribution, Qualitätsprüfung und Transparenz im World Wide Web angeregt (Haber, 2013, S. 175-190). Vor diesem Hintergrund haben Bernhard Rieder und Theo Röhle in ihrem 2012 veröffentlichten Aufsatz »Five Challenges« fünf zentrale Herausforderungen für die aktuellen Digital Humanities identifiziert: die Verlockung der Objektivität, die Macht der visuellen Evidenz, Black-Boxing (Unschärfen, Problematik der Stichprobenziehung etc.), institutionelle Turbulenzen (konkurrierende Serviceeinrichtungen und Lehrfächer) und das Streben nach Universalität. Computerunterstützte Forschung wird in der Regel von der Auswertung der Daten dominiert, und manche Forscher sehen fortgeschrittene Analyseverfahren im Forschungsprozess gar als Ersatz für eine gehaltvolle Theoriebildung. Forschungsinteressen könnten so fast ausschließlich durch die Daten selbst getrieben werden. Diese evidenzbasierte Konzentration auf das mit den Daten Mögliche kann Forscher aber dazu verleiten, heuristische Aspekte ihres Gegenstandes zu vernachlässigen. Diese Problematik hat der englische Medientheoretiker Richard ROGERS zum Anlass genommen, sich intensiv mit der Theorie und der praktischen Anwendung digitaler Methoden auseinanderzusetzen. Seine Aufsätze und Bücher haben maßgeblich die einschlägigen Debatten geprägt und die digitale Medienanalyse entscheidend beeinflusst. In dem hier abgedruckten Text mit dem Titel »Nach dem Cyberspace: Big Data, Small Data« setzt er sich mit der digitalen methodologischen Perspektive auseinander, das



Internet als Datensatz für die Sozial- und Kulturforschung zu betrachten, und beschäftigt sich vor diesem Hintergrund eingehend mit Methodenproblemen, die bei der Analyse von Suchmaschinen-Anfragedaten und bei der Formatierung der Daten in Metadaten auftreten können. Forschungsstrategisch gesehen zielt seine Argumentation darauf ab, die Datenkonstruktionen der sozialen ›Wirklichkeit‹ methodenkritisch in Frage zu stellen. Seine digitale Methodenkritik beschränkt sich hier weniger auf die methodischen Mängel im Prozess des Forschens (z.B. auf Fehlerquellen, die bei der Erhebung und Auswertung von Daten vorkommen könnten), sondern befragt darüberhinausgehend die technologischen Möglichkeitsbedingungen und theoretischen Vorannahmen, die dem Untersuchungsobjekt vorausgehen und die Modellierung sozialer Datenaggregate mitbestimmen.

In ihrem Textbeitrag über die ›Twitter-Archive und die Herausforderungen von ›Big Social Data‹ für die Medien- und Kommunikationswissenschaft‹ beschäftigen sich Jean BURGESS und Axel BRUNS mit den technischen, epistemologischen und politischen Verfahren und Praktiken, die bei der Nutzung von Twitter-Archiven in der medien- und kommunikationswissenschaftlichen Forschung zur Anwendung kommen. In ihrem Beitrag zeigen sie auf, dass die Forschungsinfrastrukturen der Big Data-Forschung wesentlich von der – kostspieligen und zeitaufwendigen – Weiterentwicklung mit technischen Experten und einer bisher stark eingeschränkten Verfügbarkeit experimenteller Werkzeuge und Methoden zur Social Media-Datenanalyse abhängig sind. Angesichts dieser strukturellen Beschränkungen plädieren sie dafür, die Agenda der ›Code Literacy‹ für die Konzeptentwicklung der Großdatenforschung zu stärken, um diese in das Feld der Medien-, Kommunikations- und Kulturwissenschaften transferieren zu können. In dieser Hinsicht könnten auch die narrativen, fiktionalen und bildkulturellen Anteile der Big Data in den Blick genommen werden, nämlich in dem Sinn, dass sie als etwas Gemachtes, Hergestelltes und Konventionelles aufgefasst werden können und nicht einfach als faktisch-gegebenes Fenster zur sozialen Welt der Tatsachen in einer passiven Rezeptionshaltung akzeptiert werden müssen (vgl. Hayles, 2005).

Bernhard RIEDER und Carolin GERLITZ befassen sich in ihrer Untersuchung ›Data-Mining von einem Prozent Twitter – Sammlungen, Basislinien, Stichproben‹ mit den materialtechnischen Voraussetzungen von sozialen Netzwerken, indem sie die performativen Fähigkeiten von Twitter, spezifische Aktivitäten zu ermöglichen und zu strukturieren, untersuchen. Auf der Grundlage eines breit gestreuten Überblicks über die Methodenliteratur zur Social Media-Research entwickeln sie für das Mining von sozialen Medien ein Problemverständnis für die Techniken der Stichprobenziehung, Messung und Analyse von großen Datenmengen und interpretieren Twitter in erster Linie als eine mediendispositive Anordnung. Hashtags, Tweets, Retweets, Favoriten und Follows sind Medien des Sozialen und stellen Sozialität her. In dieser Sichtweise können die Funktionen der digitalen Kommunikation in Anlehnung an Deleuze als Sozialisierungslinien, die Sozialität hervorbringen, gekennzeichnet werden (vgl. Deleuze, 1991, S. 159).

Mit dieser Sichtweise ausgerüstet können sie weiterführende Fragen nach dem Stellenwert der sozialen Vernetzungstechnologien von Online-Plattformen aufwerfen. Aufgrund der hohen Anforderungen an die schnelle Verfügbarkeit der großen Datenmengen können die Rohdaten nicht rechtzeitig verarbeitet werden. Folglich haftet den Datensamples eine technisch bedingte Unsicherheit oder Ungenauigkeit an, die in ihrer Überlagerung aufzeigen, dass den jeweiligen konkreten Verbindungen zwischen Technologien, Praktiken, Materialitäten und Epistemologien eine unauflösliche Mehrdeutigkeit anhaftet.

Auch die Studie »Der Wert von Big Data für die Erforschung digitaler Medien« von Merja MAHRT und Michael SCHARKOW setzt sich mit den methodologischen Aspekten der Big Data-Forschung auseinander und befragt ihre Anwendbarkeit und ihren Nutzen bei der Erforschung der digitalen Medien. In ihrer theoretisch fundierten und empirisch ausgerichteten Argumentation plädieren sie dafür, dass die rezente Onlineforschung im Bereich der Sozialen Medien im Web 2.0 den Aussagewert von sehr großen Datenanalysen konsequenter theoretisch reflektieren sollte, da deren Validität durch infrastrukturelle Rahmenbedingungen stark eingeschränkt sein kann. Katrin WELLER hat sich in ihrer Analyse »Twitter und Wahlen: Zwischen 140 Zeichen und Milliarden von Tweets« intensiv mit dem wissenschaftlichen Feld der Twitter-Forschung auseinandergesetzt und 49 Publikationen aus dem Jahr 2013 untersucht, die mit Twitter-Daten im Kontext von Wahlanalysen operieren. Das erklärte Ziel ihres innovativen Beitrags ist es, den aktuellen Forschungsstand hinsichtlich seiner Ergebnisse, verschiedenen Herangehensweisen, angewandten Methoden und unterschiedlichen Datentypen methodenkritisch zu vergleichen.

In seinem Textbeitrag »From Mind to Document and Back Again. Zur Reflexivität von Social-Media-Daten« entwickelt auch Johannes PASSMANN eine vielversprechende theoretische Perspektive, die es erlaubt, das wissenschaftliche Wissen von Big Data-Analysen einer kritischen Revision zu unterziehen. Er zeigt auf, inwiefern nutzer- und plattformspezifische Reflexionsformen spezifische Social Media-Plattformdaten erzeugen, die nicht immer eindeutig lesbar sind, nicht auf die formale Genauigkeit repräsentativer Aussagen hochgerechnet werden können und eher auf die Erzeugung spieltaktischer Effekte und experimenteller Situationen verweisen. Bei 4chan und YouTube »besteht oft gerade der Reiz darin, Dokumente zu erzeugen, die absolut unerwartbare Konsequenzen zeitigen, wie bspw. dass der eigene Tweet in den Tagesthemen zitiert wird. Man könnte es so als ein Kernmoment des Social Web bezeichnen, aus der Reflexivität möglicher Konsequenzen einen spielähnlichen Vorgang zu machen, bei dem die Spielhaftigkeit dieser Konsequenzen stets in Frage steht.« (Abgedruckt im vorliegenden Band.) Der theoriekritische Ansatz von Paßmann kann als wertvoller Hinweis dafür genommen werden, die theoretisch in Aussicht gestellte Vermessung des Sozialen kritisch zu hinterfragen, wenn in Betracht gezogen wird, dass Soziales nicht einfach wie ein lebloses und bewusstloses Ding vermessen werden kann, da soziale Akteure über Reflexivität verfügen. Um Reflexivität auszubilden, müssen

soziale Akteure auf sozialen Netzwerkseiten nicht unbedingt alle Prozeduren und Hintergründe ihrer Überwachung und Kontrolle durchschauen; den Akteuren genügt bereits ein vager Verdacht und ein nicht spezifizierbares Unbehagen, um eine Reflexivität möglicher Ordnungen oder möglicher Konsequenzen auszubilden.

Das *vierte Kapitel* mit dem Titel »Dataveillance: Algorithmen, Graphen und Protokolle« widmet sich den sozialen Steuerungsprozessen und machtpolitischen Aspekten, die der Erzeugung von datengenerierter Forschung und materiellen Datenkulturen inhärent sind (vgl. Gillespie, 2010, S. 347-364). Auf welche Weise beschleunigen und verstärken Algorithmen, Protokolle und Netzwerke die Organisation, Analyse und Repräsentation einer digitalen Kontrollgesellschaft? Welche Analysemodelle bieten sich an, um die datenintensiven Formen der Wissensproduktion kritisch zu befragen? Die in diesem Kapitel zusammengeführten Texte arbeiten die politische Relevanz der Software und der digitalen Medientechnologien heraus. Diese den Critical Code Studies (Hayles, 2004, S. 67-90; Marino, 2006; Kirschenbaum, 2004, 2008; Berry, 2011; Chun, 2011; Wardrip-Fruin, 2011) nahestehenden Textbeiträge gestehen der Vernetzungstechnologie selbst eine strukturbildende Macht zu und interpretieren die digitalen Informations- und Kommunikationstechnologien als eine rechner- und softwarebasierte Ermöglichungsmacht sozialer und kultureller Praktiken, die weite Bereiche der Alltagskommunikation dominieren.

Geht man von der Annahme aus, dass die kollektive Datenkommunikation in Rechnernetzen durch die Netzwerkinfrastruktur der Netzwerkprotokolle organisiert wird und in funktionale Internetschichten zergliedert ist, dann können die Netzwerkprotokolle als Kulturtechniken der sozialen Regulierung angesehen werden, mit denen kollektive Prozesse als technisch bedingte Effekte von Netzwerktechnologien angeschrieben werden können (Wagner, 2006, S. 26f.). Mit der von Alexander Galloway projektierten Medien- und Machttheorie der »protokollgischen« Kontrolltechniken kann der liberale Machttypus der verteilten Kontrollgesellschaft treffend als nicht statische Kräfterelation beschrieben werden. Die Protokolle stehen folglich für eine liberale Regierungstechnologie, die hochgradige Spielräume an unregulierter Kommunikation und flexibilisierter Distribution von Inhalten für einen taktischen Mediengebrauch ermöglicht. Protokolle operieren unterhalb der sichtbaren Anwendungsschichten im Verborgenen und werden daher nicht als mediale Beschränkung von Informationsflüssen wahrgenommen, sondern als herrschafts- und machtneutrales Tool. Diesem blinden Fleck im Mediengebrauch schenkt der in diesem Sammelband erstübersetzte Essay »Protokoll, Kontrolle und Netzwerke« von Alexander GALLOWAY und Eugene THACKER seine volle Aufmerksamkeit. In der Tradition medienmaterialistischer Ansätze und vor dem Theoriehintergrund der Science Studies fragen sie nach dem Stellenwert von informatischen Konzepten und Benutzerschnittstellen bei der Herausbildung sozialer Formationen und politischer Figuren des Wissens: »Das Set von Verfahren zur Überwachung, Regulierung und Modulierung von

Netzwerken als lebendigen Netzwerken ist auf der grundlegendsten Ebene auf die Produktion von Leben in seinen biologischen, sozialen und politischen Kapazitäten ausgerichtet.« (Abgedruckt im vorliegenden Band.) Sie interpretieren die algorithmischen Standards, Normen und Protokolle also in erster Linie als vermittelnde Instanz zwischen den kulturellen Praktiken und den technischen Infrastrukturen. Galloway und Thacker verstehen Netzwerke nicht bloß als technische Systeme, sondern als sozial dynamische und lebendige Netzwerke, die sich in Echtzeit organisieren. Vor diesem Hintergrund untersuchen sie nicht nur die technischen Möglichkeiten der politischen Kontrolle durch Algorithmen und Protokolle, sondern befragen die politischen Handlungsmöglichkeiten von netzwerk-basierten Bewegungen.

Auch Matteo PASQUINELLI versteht Software als einen historischen Wissensbestand, der seine eigene Geschichte aufweist und somit nicht nur technologischen Normen und Standards unterliegt, sondern ebenso durch soziale, institutionelle und kulturelle Rahmenbedingungen bedingt ist. In seinem Aufsatz »Der italienische Operaismo und die Informationsmaschine« verknüpft Pasquinelli den Maschinenbegriff von Gilles Deleuze und Felix Guattari mit der Algorithmik des digitalen Codes, um beide Konzepte auf die Marx'sche Theorie der Maschinen zu beziehen. In diesem Sinn schlägt Pasquinelli vor, die Extraktion von Metadaten als einen wertschöpfenden Faktor zu deuten und Netzwerke als Maschinen zur Kontrolle, Akkumulation und Steigerung des Mehrwerts auszulegen: »Abschließend – als ein Set von vorläufigen Hypothesen innerhalb der entstehenden ›Big Data‹-Gesellschaft – kann von den Metadaten gesagt werden, dass sie dazu genutzt werden: 1) um den Wert von sozialen Beziehungen zu messen; 2) das Design von Maschinen und maschineller Intelligenz zu verbessern; und 3) Massenverhalten zu überwachen und vorherzusagen.« (Abgedruckt im vorliegenden Band.)

Dieser strategische Zusammenhang von politischer Ökonomie, Wissensproduktion und Datenbanktechnik steht auch im Zentrum der Analyse von Christoph ENGEMANN, der sich in seinem Textbeitrag »You cannot not Transact – Big Data und Transaktionalität« mit den gemeinsamen Bezugspunkten zwischen der ökonomischen und der formalisierten Transaktion in der Datenbanktechnik beschäftigt. In diesem Kontext vertritt er die These, dass sich am Transaktionsbegriff die genealogische Verschränkung von relationalen Datenbanken und Ökonomie ablesen lässt. In einer ersten Annäherung untersucht er die Begriffsgeschichte der Transaktionalität, um zu klären, was unter der bislang medienwissenschaftlich ungenügend untersuchten Kulturtechnik der Transaktion verstanden werden kann. In einem zweiten Schritt arbeitet er heraus, inwiefern Transaktionalität ein Kernelement von Big Data bildet, um schließlich die historisch fundierte und politisch motivierte Frage nach der ökonomischen Dimension von Big Data stellen zu können.

Die Aggregation, Archivierung und Auswertung von transaktionalen Daten zählen heute zum Kernbereich erfolgreicher Geschäftsmodelle. Auch das epi-

demiologische Data Mining nutzt große Verhaltensdatenbanken, um auf der Grundlage von transaktionalen Daten zukünftige Krankheitsentwicklungen vorhersagen zu können. Annika RICHTERICH untersucht in ihrem Aufsatz »Infodemiologie – von ›Supply‹ zu ›Demand‹. Google Flu Trends und transaktionale Big Data in der epidemiologischen Surveillance« den Stellenwert von transaktionalen Daten im epidemiologischen Monitoring und Mapping von Nutzerverhalten und beschäftigt sich mit den sozial- und netzpolitischen Problemen dieser neuen Datenpraktiken. Ihre Forschungsperspektive auf die Software als eine sozial konstituierende Macht bleibt aber von der Prämisse begleitet, dass Geschichte nicht lückenlos geplant, sondern in unterschiedlichen sozialen Räumen und Kontexten nur als ein offener Prozess verstanden werden kann. So kreist etwa die Übertragung von Befragungs- und Nutzerdaten in Kombination mit dem Surfverhalten der User auf die Gesamtheit aller Nutzer statistischen Berechnungen durch das Social Media Targeting und das Predictive Behavioral Targeting um mathematische Optimierungsprobleme, da die sozialen Beziehungsgefüge stets wandelbar, schwach strukturiert und unzyklisch in ihrer Mediendynamik erscheinen.

Das den Band abschließende *fünfte Kapitel* befasst sich mit den technik- und sozialgeschichtlichen Aspekten der Verarbeitung von Großdaten und erkundet das Wechselverhältnis von technischen Infrastrukturen und sozialen Ordnungsvorstellungen. Es ist aber weniger eine chronologische Geschichtsschreibung der technischen Innovationen, welche die Medienumbrüche und sozialen Transformationen erklären kann. Denn die gegenwärtigen Medien- und Subjektkulturen sind immer auch schon in den Technologien, Prozeduren und medialen Anordnungen enthalten, die ihnen vorausgehen und die sie schließlich voraussetzen – ohne sie damit auch schon kausal zu determinieren. Stefan HÖLTGEN spürt einer dieser technologischen Entwicklungen zur sozialen Rechnernetzung nach, die sich in den frühen 1970er Jahren herausbildete und veränderliche Wissensformen, kulturelle Hybridisierungen und multimediale Anordnungen im Operationsbereich der Großdaten hervorbrachte. Er vertritt in seiner Abhandlung »All Watched Over by Machines of Loving Grace. Öffentliche Erinnerungen, demokratische Informationen und restriktive Technologien am Beispiel der ›Community Memory‹« die These, dass für das Aufkommen großer Datenmengen mehr Maschinen als Menschen verantwortlich sind. Um die medienhistorischen Weichenstellungen der Rechnerverarbeitung von großen Datenmengen freizulegen, untersucht er das »Community Memory«-Projekt, das in den frühen 1970er Jahren in der California Bay Area bei San Francisco entstand. Das »Community Memory« interpretiert Höltgen als eine Basistechnologie sozialer Netzwerke und kann so eine mögliche Genealogie der Entstehung aller modernen Massendaten-Services und Massendatenverarbeitungen nachweisen.

In einem ersten Schritt beleuchtet er die Protagonisten und die institutionelle Einbettung des »Community Memory«-Projektes. Anschließend untersucht er die technischen und historischen Bedingungen seiner Software- und Speichertechnologie und rekonstruiert die medientechnischen Bedingungen der Mög-

lichkeit zeitgenössischer Netzwerkseiten. In diesem Zusammenhang zeigt er nicht nur, dass Massendatenauswertung eine dauernde Forderung an die Weiterentwicklung von Hardware und Software stellt, sondern auch, dass sich seit den ersten rechnerbasierten Netzwerkprojekten eine sich formierende Politik des Informationszeitalters ablesen lässt. In dieser Hinsicht vollzieht der Computer als Medium einen Registerwechsel, insofern er von einer Medienkultur, die das Speichern privilegiert, zu einer Medienkultur der permanenten Übertragung und Vernetzung mutiert und damit die gegenwärtige Vernetzungskultur vorwegnimmt, deren Unterschiede und Gemeinsamkeiten Regine BUSCHAUER in ihrer Untersuchung thematisiert. Sie befasst sich in »(Very) Nervous Systems. Big Mobile Data« mit der medien- und technikgeschichtlichen Reflexion der pervasiven Vernetzung des Alltags durch den Einsatz »intelligenter« Medien. Zunächst wird die Geschichte des technischen »Nervensystems« als die eines historisch grundlegenden Topos der Kommunikationstechnik nachgezeichnet. Schließlich befragt sie die aktuellen Mediendiskurse zur omnipräsenten Kommunikationstransparenz auf ihre techno-imaginären Geschichtsbezüge und sozialen Ordnungsvorstellungen. Ausgehend vom zellulären System des Mobilfunks und einem Location Sensing, das in der Konvergenz mobiler digitaler Medien seit den 1990er Jahren in den Vordergrund tritt, befragt sie in weiterer Folge die Vitalitätsmetaphern »pervasiver« und »ubiquitärer« Medien, die über die Ordnungsvorstellung einer imaginären »God's eye view« hinausweisen. In dieser Sichtweise kann sie die diskursive Ermöglichung von Gegenwartsaussagen freilegen und in ihrem normativen Anspruch problematisieren. Indem sie ihre Untersuchung auf die Medien- und Kommunikationsgeschichte der technischen Systeme konzentriert, vermag sie schließlich aufzuzeigen, dass das Technische immer schon an einem gemeinsamen Vorrat von Gedächtniskulturen und Bilderrepertoires partizipiert. In dieser Engführung erschließt sie sich die imaginären Bilderwelten der Big Data an den Schnittstellen der heterogenen Genealogie eines pervasiven Medienwandels.

Digitale Infrastrukturen wie die von Buschauer erörterte rechnerbasierte Mobilität im Sinne selbstorganisierender ad-hoc-Netze und kontextorientierter Softwaresysteme zeigen auf, dass Mediengebrauch nicht länger als »Unterwerfung« und »Anpassungsleistung« einer »ursprünglichen« Subjektivität angesehen werden muss, weil Subjektivität selbst schon hergestellt ist und das Netz bereits zu den gängigen Metaphern zählt, mit denen Subjekte und ihr Verhalten als zeitgemäß beschrieben werden. In diesem Sinne teilen netzwerktheoretische Ansätze eine konstruktivistische Grundeinstellung und gehen davon aus, dass die neuen »Verstehensformen von Subjektivität« (Paulitz, 2005, S. 40) maßgeblich von den Technologien der Vernetzung und dem Internet als technisches Artefakt geprägt sind.

Ramón REICHERT beschäftigt sich in seinem Textbeitrag »Facebooks Big Data: Die Medien- und Wissenstechniken der kollektiver Verdatung« mit dem Data Mining des *Facebook Data Teams*. Unter Leitung des Soziologen Cameron

Marlow erstellt die Forschergruppe, bestehend aus der Sozialpsychologin Moira Burke, dem Informatiker Danny Ferrante und dem Soziologen Adam Kramer, Trend-, Text- und Sentimentanalysen zur Stimmungserkennung der Facebook-User. Auf der Datengrundlage der Status-Updates erstellt die Facebook Data Science Research seit 2007 ihren »Gross National Happiness Index« (GNH), der das »Bruttonationalglück« großer sozialer Gruppen repräsentieren soll. Das *Facebook Data Team* hat anhand von Status-Updates die Stimmungslage in 22 Ländern analysiert und interpretiert damit die eigene Plattform als relevanten Indikator zur Vermessung der nationalen Stimmungslage. Daher steht zunächst die folgende Kernfrage im Zentrum der Analyse: Welche Medien- und Wissenstechniken werden von Facebook eingesetzt, um »Stimmungsdaten« in Metadaten zu überführen? In einem zweiten Schritt wird aufgezeigt, inwiefern die von Facebook in Aussicht gestellten Verhaltensprognosen bereits in ihren Grundannahmen von Präsuppositionen durchdrungen sind und mit einer Rhetorik und »Poetologie des Wissens« (Vogl, 1998) operieren, welche in die jeweiligen Datenkonstruktionen einfließen.

Neben den neuen Technologien zur Herstellung digitaler Kommunikationsgemeinschaften haben sich in den Peer-to-Peer-Netzwerken und Online-Portalen aber auch traditionelle Leitbilder des gemeinschaftlichen Zusammenlebens und -arbeitens erhalten, die zur Konstruktion sozialer, kultureller und politischer Identität eingesetzt werden. Die Gemeinschaftsidee ist heute zum Angelpunkt der Legitimation der Sozialen Medien im Web 2.0 aufgestiegen (van Dijck, 2012). Vor diesem Hintergrund untersucht Martin DOLL den Bezugsrahmen von technischen Medien und den Ordnungsvorstellungen politischer Imaginationen. In seinem Beitrag »Theorie und Genealogie des Techno-Imaginären: Social Media zwischen »Digital Nation« und pluralistischem Kosmopolitismus« richtet er ausgehend von einer luziden Lektüre von Benedict Andersons »Imagined Communities« seine Aufmerksamkeit auf die Entstehung der gegenwärtig mit den Social Media verbundenen Vorstellungen von Kollektivität. In produktiver Bezugnahme auf die Sozial- und Kommunikationsheorien von James Carey, Clifford Geertz und Bruno Latour entwickelt Doll ein vielschichtiges Interesse für die aktuellen gesellschaftlichen Selbstbeschreibungen der Plattformen Facebook und Twitter. Dabei kann er nicht nur eine Pluralität von bild-, ideen- und mediengeschichtlichen Tradierungslinien sozialer Ordnungsvorstellungen nachweisen, sondern auch gleichermaßen aufzeigen, inwiefern die mit den Sozialen Netzwerken verknüpften Imaginationen von Kollektivität auf künftige politische Handlungen und Entscheidungen Einfluss nehmen.

Mit seiner Sichtweise verfolgt Martin Doll das Ziel, die Aufmerksamkeit für den veränderten Stellenwert der gegenwärtigen Vernetzungskultur zu schärfen. Zusammenfassend lässt sich feststellen, dass sich in der Ära der Big Data nicht nur die Datenpraktiken und Wissenskultur, sondern auch der Stellenwert von sozialen Netzwerken radikal geändert hat, denn sie figurieren zunehmend als gigantische Datensammler für die Beobachtungsanordnungen sozialstatisti-

schen Wissens und als Leitbild normalisierender Praktiken. Es haben sich mit der Durchsetzung der digitalen Kommunikation, der Sozialen Medien und dem Internet der Dinge aber auch die sozialen Akteure selbst verändert. Insofern erfordert das Zeitphänomen Big Data nicht nur eine Neuorganisation der Digital Humanities und die institutionelle Einbindung digitaler Methodik, sondern auch eine weitgefassete politische, kritische und historische Perspektivierung von Daten, die in der Lage ist, die tektonischen Verschiebungen der Gegenwartsgesellschaft in allen Bereichen des Alltags angemessen reflektieren zu können.

## LITERATUR

- Aronova, Elena, Karen Baker und Naomi Oreskes: »Big Science and Big Data in Biology: From the International Geophysical Year through the International Biological Program to the Long-Term Ecological Research Program, 1957-present«, in: *Historical Studies in the Natural Sciences* 40/2, 2010, S. 183-224.
- Ball, Kirstie, Kevin D. Haggerty und David Lyon: *Routledge Handbook of Surveillance Studies*, London: Routledge 2012.
- Ballsun-Stanton, Brian: »Asking about Data: Experimental Philosophy of Information Technology«, in: *5th International Conference on Computer Sciences and Convergence Information Technology* 1/5, 2010, S. 119-124.
- Beniger, James R.: *The Control Revolution*, Cambridge et al.: Harvard University Press 1986.
- Berry, David M.: *The Philosophy of Software: Code and Mediation in the Digital Age*, Basingstoke: Palgrave Macmillan 2011.
- Bobley, Brett: »Announcing the 2011 Digging into Data Challenge«, 2011. [www.neh.gov/divisions/odh/grant-news/announcing-the-2011-digging-data-challenge](http://www.neh.gov/divisions/odh/grant-news/announcing-the-2011-digging-data-challenge) (20. Mai 2014).
- boyd, danah, und Kate Crawford: »Six Provocations for Big Data«, in: *Conference Paper, A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, September 2011, Oxford: University of Oxford 2011. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1926431](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431) (20. Mai 2014).
- Bührer, Karl Wilhelm: »Ueber Zettelnotizbücher und Zettelkatalog«, in: *Fernschau* 4, 1890, S. 190-192.
- Burgess, Jean, und Cornelius Puschmann: »The Politics of Twitter Data«, 2013. [www.papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2206225](http://www.papers.ssrn.com/sol3/papers.cfm?abstract_id=2206225) (20. Mai 2014).
- Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner und Jeffrey Schnapp: »Digital\_Humanities«, Cambridge, MA: MIT Press 2012. [http://mitpress.mit.edu/sites/default/files/titles/content/9780262018470\\_Open\\_Access\\_Edition.pdf](http://mitpress.mit.edu/sites/default/files/titles/content/9780262018470_Open_Access_Edition.pdf) (20. Mai 2014).
- Busa, Roberto: *S. Thomae Aquinatis Hymnorum Ritualium Varia Specimina Concordantiarum. Primo saggio di indici di parole automaticamentecomposti e stampati da macchine IBM a schede perforate*, Milano: Bocca 1951.



- Ders.: »The Annals of Humanities Computing: The Index Thomisticus«, in: *Computers and the Humanities* 14 (2), 1980, S. 83-90.
- Chun, Wendy H.K.: *Programmed Visions: Software and Memory*, Cambridge, MA: MIT Press 2011.
- Conover, Michael D., Bruno Goncalves, Alessandro Flammini und Filippo Menczer: »Partisan Asymmetries in Online Political Activity«, in: *EPJ Data Science* 1/6, 2012. [www.epjdatascience.com/content/1/1/6](http://www.epjdatascience.com/content/1/1/6) (20. Mai 2014).
- Davidson, Cathy N.: »Humanities 2.0: Promise, Perils, Predictions«, in: *Publications of the Modern Language Association (PMLA)*, Vol. 123 (3), 2008, S. 707-717.
- Deleuze, Gilles: »Was ist ein Dispositiv?«, in: Francois Ewald und Bernhard Waldenfels (Hg.), *Spiele der Wahrheit. Michel Foucaults Denken*, Frankfurt a.M.: Suhrkamp 1991, S. 153-162.
- Dijck, José van (2013), *The Culture of Connectivity: A Critical History of Social Media*, Oxford: Oxford University Press.
- Driscoll, Kevin: »From punched cards to ›Big Data‹: A social history of database populism«, in: *Communication +1* (1), 2012. <http://kevindriscoll.info/> (20. Mai 2014).
- Drucker, Johanna: *SpecLab: Digital Aesthetics and Projects in Speculative Computing*, Chicago und London: University of Chicago Press 2009.
- Floridi, Luciano: *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford: Oxford University Press 2014.
- Gayo-Avello, Daniel: »No, You Cannot Predict Elections with Twitter«, in: *IEEE Internet Computing* 16/6, 2012, S. 91-94.
- Gilbert, Eric, und Karrie Karahalios: »Widespread Worry and the Stock Market«, in: *4th International AAAI Conference on Weblogs and Social Media (ICWSM)*, Washington, DC: George Washington University 2010, S. 58-65.
- Gillespie, Tarleton: »The Politics of ›Platforms‹«, in: *New Media & Society* 12/3, 2010, S. 347-364.
- Ders.: »The Relevance of Algorithms«, in: Tarleton Gillespie, Pablo Boczkowski und Kirsten Foot (Hg.), *Media Technologies*, Cambridge, MA: MIT Press 2012. [www.tarletongillespie.org/essays/Gillespie%20-%20The%20Relevance%20of%20Algorithms.pdf](http://www.tarletongillespie.org/essays/Gillespie%20-%20The%20Relevance%20of%20Algorithms.pdf) (20. Mai 2014).
- Gitelman, Lisa, und Virginia Jackson (Hg.) (2013): *›Raw data‹ is an oxymoron*, Cambridge MA: MIT Press.
- Gitelman, Lisa, und Geoffrey B. Pingree: *New Media: 1740-1915*, Cambridge, MA: MIT Press 2004.
- Gold, Matthew K. (Hg.), *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press 2012.
- Haber, Peter: »›Google Syndrom.« Phantasmagorien des historischen Allwissens im World Wide Web«, in: *Zürcher Jahrbuch für Wissensgeschichte* 9, 2013, S. 175-190.
- Hagner, Michael, und Caspar Hirschi (Hg.), *Nach Feierabend, Digital Humanities – Zürcher Jahrbuch für Wissensgeschichte* 9, Berlin: diaphanes 2013.

- Hayles, Katherine N.: »Print Is Flat, Code Is Deep: The Importance of Media-Specific Analysis«, in: *Poetics Today* 25 (1), 2004, S. 67-90.
- Dies.: *My Mother Was a Computer: Digital Subjects and Literary Texts*, Chicago: University of Chicago Press 2005.
- Hockey, Susan: »History of Humanities Computing«, in: Susan Schreibman, Ray Siemens und John Unsworth (Hg.), *A Companion to Digital Humanities*, Oxford: Blackwell 2004.
- Jaraus, Konrad H.: »Möglichkeiten und Probleme der Quantifizierung in der Geschichtswissenschaft, in: Ders. (Hg.), *Quantifizierung in der Geschichtswissenschaft. Probleme und Möglichkeiten*«, Düsseldorf: Droste 1976, S. 11-30.
- Kirschenbaum, Matthew G.: »Extreme Inscription: Towards a Grammatology of the Hard Drive«, in: *TEXT Technology* (2), 2004, S. 91-125.
- Ders.: *Mechanisms: New Media and the Forensic Imagination*, Cambridge: MIT Press 2008.
- Krajewski, Markus: »In Formation. Aufstieg und Fall der Tabelle als Paradigma der Datenverarbeitung«, in: David Gugerli, Michael Hagner, Michael Hampe, Barbara Orland, Philipp Sarasin und Jakob Tanner (Hg.), *Nach Feierabend. Zürcher Jahrbuch für Wissenschaftsgeschichte*, Bd. 3, Zürich und Berlin: diaphanes 2007, S. 37-55.
- Kuhn, Thomas S.: *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press 1962.
- Lauer, Gerhard: »Die digitale Vermessung der Kultur. Geisteswissenschaften als Digital Humanities«, in: Heinrich Geiselberger und Tobias Moorstedt (Hg.), *Big Data. Das neue Versprechen der Allwissenheit*, Frankfurt a.M.: Suhrkamp 2013, S. 99-116.
- Leslie, Stuart W.: *The Cold War and American science: The military-industrial-academic complex at MIT and Stanford*, Columbia: Columbia University Press 1993.
- Lévi-Strauss, Claude (1964): *Das Rohe und das Gekochte*, Frankfurt a. Main: Suhrkamp.
- McCarty, Willard: *Humanities Computing*, London: Palgrave 2005.
- McPherson, Tara: »Dynamic Vernaculars: Emergent Digital Forms in Contemporary Scholarship«, Lecture presented to HUMlab Seminar, Umeå University, 3. April 2008. <http://stream.humlab.umu.se/index.php?streamName=dynamicVernaculars> (20. Mai 2014).
- Paulitz, Tanja (2005): *Netzsubjektivität/en. Konstruktionen von Vernetzung als Technologien des sozialen Selbst. Eine empirische Untersuchung in Modellprojekten der Informatik*. Münster: Westfälisches Dampfboot.
- Price, Derek J. Solla: *Little science, Big science*, New York: Columbia University Press 1963.
- Schischkoff, Georgi: »Über die Möglichkeit der Dokumentation auf dem Gebiete der Philosophie«, in: *Zeitschrift für Philosophische Forschung* 6 (2), 1952, S. 282-292.

- Svensson, Patrik: »The Landscape of Digital Humanities«, in: *Digital Humanities Quarterly (DHQ)*, Vol. 4 (1), 2010. [www.digitalhumanities.org/dhq/vol/4/1/000080/000080.html](http://www.digitalhumanities.org/dhq/vol/4/1/000080/000080.html) (20. Mai 2014).
- Taewoo, Nam, und Jennifer Stromer-Galley: »The Democratic Divide in the 2008 US Presidential Election«, in: *Journal of Information Technology & Politics* 9/2, 2012, S. 133-149.
- Thaller, Manfred: »Controversies around the Digital Humanities: An Agenda«, in: Ders. (Hg.), *Computing Historical Social Research*, Vol. 37 (3), 2012, S. 7-23.
- Vanhoutte, Edward: »The Gates of Hell: History and Definition of Digital | Humanities«, in: Melissa Terras, Julianne Tyham und Edward Vanhoutte (Hg.), *Defining Digital Humanities*, Farnham: Ashgate 2013, S. 120-156.
- Vogl, Joseph (Hg.), *Poetologien des Wissens um 1800*, München: Wilhelm Fink 1998.
- Wald, Randall, Taghi M. Khoshgoftaar und Chris Sumner: »Machine Prediction of Personality from Facebook Profiles«, in: *13th IEEE International Conference on Information Reuse and Integration*, 2012, S. 109-115.
- Wagner, Kirsten: *Datenräume, Informationslandschaften, Wissensstädte. Zur Veräumlichung des Wissens und Denkens in der Computermoderne*, Freiburg: Rombach 2006.
- Wardrip-Fruin, Noah: *Expressive Processing*, Cambridge: MIT Press 2011.
- Yogatama, Dani: »Predicting the Future: Text as Societal Measurement«, Language Technologies Institute, School of Computer Science, Carnegie Mellon University 2012. [www.cs.cmu.edu/~dyogatam/Home\\_files/statement.pdf](http://www.cs.cmu.edu/~dyogatam/Home_files/statement.pdf) (20. Mai 2014).
- Zhang, Qi: »Cloud computing: state-of-the-art and research challenges«, in: *Journal of Internet Services and Applications* 1/1, 2010, S. 7-18.