Big Data: Inequality by Design?

Prietl, Bianca

Erstveröffentlichung / Primary Publication
Konferenzbeitrag / conference paper

PROCEEDINGS OF THE WEIZENBAUM CONFERENCE 2019
CHALLENGES OF DIGITAL INEQUALITY
DIGITAL EDUCATION | DIGITAL WORK | DIGITAL LIFE

BIG DATA: INEQUALITY BY DESIGN?

Bianca Prietl
Department of Sociology
Technical University Darmstadt
Darmstadt, Germany
prietl@ifs.tu-darmstadt.de

ABSTRACT

This paper proposes to tackle the problem of digital inequality by introducing digital technologies of knowledge generation and decision-making to a feminist critique of rationality that is informed by discourse theory and intersectional perspectives on gender and gendered relations of inequality. Therefore, it takes a closer look at the epistemological foundations of Big Data as one prominent representation of digital technologies. While Big Data and Big Data-based results and decisions are generally believed to be objective and neutral, numeral cases of algorithmic discrimination have lately begged to differ. This paper argues that algorithmic discrimination is neither random nor accidental; on the contrary, it is – amongst others – the result of the epistemological foundation of Big Data – namely: data fundamentalism, post-explanatory anticipatory pragmatics, and anti-political solutionism. As a consequence, a critical engagement with the concepts and premises that become materialized in the design of digital technologies is needed, if they are not to silently (re)produce social inequalities.

KEYWORDS

Big Data; Algorithmic Discrimination; Feminist Critique of Rationality; Epistemology; Intersectionality

DOI: 10.34669/wi.cp/2.11

1 INTRODUCTION

The algorithmically processed and (partially) autonomous generation and analysis of mostly heterogeneous and unstructured large-scale data sets, so-called Big Data, for the production of knowledge is a prominent and much debated example of current developments in digital technology.[1] Advocates of Big Data promise the production of more, better, and most importantly predictive knowledge that is said to improve the lives of all human beings and solve the great problems of mankind (Anderson 2008; Mayer-Schönberger and Cukier 2013; Geiselberger and Moorstedt 2013). Critics, contrariwise, warn against potential privacy breaches and surveillance risks should Big Data turn into Big Brother (van Dijk 2014; Zuboff 2015). Lately, a growing number of research, mostly informed by science and technology studies (STS), refuses to take a deterministic or essentialist stance and calls for differentiated analyses of the historical, sociocultural, political, and economic preconditions and effects of Big Data (early: boyd and Crawford 2012; Gitelman 2013).[2] The latter strand of research tends to be skeptical of the promises made in the name of Big Data, and emphasizes the disparities between the programmatic discourses surrounding Big Data and the material phenomenon called Big Data.[3] All parties alike, however, seem to agree that Big Data constitutes a "computational turn in thought and research" (boyd and Crawford 2012: 665; also Kitchin 2014), whereas Big Data can be understood as an "emerging *Weltanschauung* grounded across multiple domains in the public and private sectors, one that is need of [sic] deeper critical engagement." (Crawford et al. 2014: 1664; original emphasis)

Big Data and Big Data-based results and decisions are generally believed to be objective and neutral; however, numeral cases of algorithmic discrimination have lately begged to differ (O'Neil 2018 [2016]; Eubanks 2018). Google's ad posting algorithm has, for example, been demonstrated to display advertisements for jobs in management positions as well as for executive training programs significantly more often to persons whose browser profile identifies them as male, than to those whose profile identifies them as female (Datta et al. 2015). Data-based risk assessment tools, which are widely employed in the US criminal justice system, to name a second example, have been shown to systematically attest African Americans a higher risk of committing a future crime than so-called white Americans (Angwin et al. 2016). When these and similar cases of algorithmic discrimination gain the attention of the wider public, the resounding outcry testifies to the broken promises of objective knowledge production and neutral decision-making. The vocabulary used to understand what is going on, when algorithms are sexist or racist, often refers to terms such as bias, error or distortion, suggesting that objective results are possible once all errors are eliminated (Zweig 2018) and, thereby, maintaining the modern ideal of "mechanical objectivity" (Daston and Gallison 1992).

This paper, by contrast, draws on the insights of STS in assuming that processes and practices of knowledge production as well as the technical artifacts employed within these processes are

---

[1] Within Big Data two theoretically and technically distinct phenomena – algorithmic processing and data – come together, whose distinction and relation demand a thorough examination that cannot be accomplished here.

[2] Research concerned with these kind of questions is predominantly conducted under the labels *Critical Data Studies*, *Critical Algorithm Studies,* or *Critical Code Stud-ies* displaying the yet to be closed terminological discus-sions. In the German-speaking context the discussion is only just beginning, with pioneering publications such as Mämecke et al. 2018, and Houben and Prietl 2018.

[3] Discursively powerful buzzwords such as 'digital transformation' and 'data revolution' suggest that Big Data has no history, whereas the material phenomenon of Big Data is far older than its discursive popularity (Barnes 2013; Barnes and Wilson 2014: 1-2). At the same time, current manifestations of Big Data fall far behind the promises made its name (Beer 2016: 2).

neither neutral, nor objective, but highly political (Winner 1980). From this perspective it is of paramount importance to critically analyze the epistemological foundations and premises of digital technologies such as Big Data. This paper contributes to this endeavor by introducing Big Data to a feminist critique of rationality that is informed by discourse theoretical perspectives on the relation of knowledge and power and intersectional perspectives on gender and gendered relations of inequality (see section 2). It, therefore, focusses on the epistemological foundation and premises of Big Data-based knowledge production and decision-making as they are articulated within the discourses surrounding Big Data. Following David Beer (2016: 5), it is vital to better understand the discourses produced in the name of Big Data as "it is also the very concept of Big Data itself that shapes decisions, judgments and notions of value – as it brings with it a vision for particular types of calculative or numerical knowing about individuals, groups and the social world".[4] Synthesizing the promises made in the name of Big Data as well as the critique brought against Big Data, three epistemological premises are portrayed as central to Big Data-based knowledge production and decision making – namely: *data fundamentalism*, *post-explanatory anticipatory pragmatics*, and *anti-political solutionism*. In order to introduce this epistemological triad of Big Data to a feminist critique of rationality, it is asked how these assumptions underlying Big Data are related to gender and gendered relations of power and inequality.[5] Put differently, this paper is concerned with how Big Data is gendered on the level of its epistemological foundation (see section 3). Finally, the results of this analysis are summarized and discussed (see section 4).

2 THEORETICAL APPROACH

This paper is situated in the tradition of critical engagements with (Western, masculine) rationality and modern objectivity within STS, but also more generally within the social sciences and humanities.[6] It, thus, assumes that (scientific) knowledge production is a social endeavor of utmost political significance, whereat technical artifacts play a crucial role and do themselves have politics (for an overview from a feminist perspective see Singer 2005). This paper's *feminist* critique of rationality is especially inspired by Donna Haraway's work that is known for its early posthumanist and neomaterialist perspectives on the power dynamics embedded in information and communication technologies (ICTs). The analysis proposed further draws on Foucault'ian discourse theoretical perspectives on knowledge production and truth claims as well as on intersectional approaches to gender, according to which gender is always intersecting with other categories of social differentiation and inequality.

Haraway describes ICTs as essential for the development of hybrid assemblages called "technosciences" within which the boundaries between technical and natural sciences, applied and basic research, science, economy, and politics become blurred (Singer 2005: 21). Technosciences establish a new mode of reasoning that is no longer based on the Newtonian logics of deduction and induction, but promotes a reflexive trial-and-error approach. Instead of searching for the universal laws of nature, research and knowledge production shift to finding (technical) solutions and real-world applications of knowledge (Weber 2017: 350-353). Following Haraway, technosciences such as Big Data are to be understood as concrete representations of

[4] As discursive phenomenon the discourses surrounding Big Data are neither congruent with the socio-material phenomenon called Big Data, nor can the one be directly deduced from or reduced to the other.

[5] Thus, the focus of this paper is neither on the gendered inequalities of data-based resource allocation (Fourcade and Healy 2013), nor on the gender aspects of the so-called "digital divide" or "digital inequality" (DiMaggio et al. 2004).

[6] Important contributions to this discussion stem also from the eponym of this conference, Joseph Weizenbaum (e.g. 1990 [1976]).

globally dominant technologies, but also as a specific approach to the world that transports certain possibilities of generating knowledge and of political engagement.

Haraway further offers a feminist and anti-racist perspective on technosciences that sensitizes to structures and processes of patriarchal, colonial, and capitalist power embedded within scientific practices and technical artifacts, without demonizing technology in general or claiming innocence for feminist approaches.[7] Emphasizing that every knowledge or truth claim is "situated", Haraway calls for taking responsibility for one's truth claims and making one's standpoint visible (2017 [1995]). This includes technical artifacts such as technologies of measurement or visualization that Haraway understands as agents in the discursive-material processes of knowledge production. According to Haraway, the technical reconfiguration of the world cannot be understood as a neutral project, but as a highly contested political endeavor. Therefore, it is important for a feminist critique of rationality as well as for an engaged intervention in the technological development to scrutinize the mode of reasoning, the rationalities, and the powerful norms of producing knowledge and making decisions promoted by the concept of Big Data.

Drawing on the discourse theoretical work of Michel Foucault (2012 [1976]) the interdependencies between knowledge and power can be conceptualized in some more detail. According to Foucault, there is no such thing as objective truth, but only knowledge claims that are acknowledged to be true. Thereby, power unfolds by means of knowledge, by "developing, organizing, and circulating a certain knowledge or rather knowledge apparatuses" (1978: 87). These knowledge apparatuses work as historically contingent "regimes of truth" that (pre)structure the acknowledged modes of reasoning as well as the norms according to which someone can come to know something at all. As indicated by Beer (2016), Big Data can be understood as such a regime of truth and, thus, needs to be confronted with the question of *who* can become a producer of true knowledge within the concept of Big Data, *how* can truth claims be made, and *what* can consequently be known and what not.

Last but not least, this paper takes an intersectional perspective on gender and gendered relations of power and inequality. According to the concept of intersectionality gender is always intersecting with other categories of social differentiation such as class, age, sexuality or race/ethnicity. Instead of taking one form of domination/marginalization as prior to others, it is an empirical question, how different modes of domination/subordination intersect, reinforcing each other or suspending one another (for an overview see Davis 2008; Bührmann 2009). Accordingly, this analysis is not limited to binary forms of gendered power relations or relations of inequality, such as men vs. women or masculine vs. feminine, but takes into account more complex forms of intersecting axes of power and inequality.

3 THE EPISTEMOLOGICAL TRIAD OF BIG DATA

Data Fundamentalism

"Before big data, our analysis was usually limited to testing a small number of hypotheses that we defined well before we even collected the data. When we let the data speak, we can make connections that we had never thought existed." (Mayer-Schönberger and Cukier 2013)

The concept of Big data promotes the idea of a "data-driven rather than knowledge-driven science" (Kitchin 2014: 1). The key to this supposedly strictly inductive mode of reasoning is the idea that (self-learning) algorithms search 'freely' – that is without recourse to theoretical

[7] In her famous *Cyborg Manifesto* she highlights the possibilities of overcoming the modern hierarchies between male and female or culture and nature by technical means (Haraway 2004 [1985]).

models or hypotheses – for patterns in large data sets, uncovering connections between different variables that would not have been foreseeable, thus producing new knowledge in its purest form. Instead of testing theoretical models and hypotheses, and thereby proceeding deductively, the concept of Big Data idealizes "the primacy of inductive reasoning in the form of a technology-based empiricism" (Mazzocchi 2014: 1250).

Big Data's "data fundamentalism" (Crawford 2013) apparently resides on two equally controversial epistemological premises: first "the belief that life can be captured and modeled by data or even fully transformed into it" (Thatcher 2014: 1768), and second the assumption that objectivity is the result of subject-free and therefore neutral production of knowledge. Both ideas have been heavily criticized within STS and shown to form specifically modern ideals of science.

Historians of science have described how the idea that 'nature should speak for itself' became dominant throughout the 19th century in modern Western societies. Whereas personal judgment was considered an important prerequisite for any scientist in the 18th century, the new notion of "mechanical" or "non-interventionist" objectivity (Daston and Galison 1992) disavowed the scientist as the subject of knowledge production. In contrast to the then spreading machines and technical apparatuses of observation and measurement the scientist was portrayed as a source of prejudice and misinterpretation and, thus, as a threat to the supposedly pure image of nature. With the replacement of the human body with technical artifacts, numerical data became increasingly important for the production and communication of scientific knowledge. Since numbers can be communicated independently, or so it seems, from the persons, places, times, and contexts of their production, they became swiftly regarded as the ideal manifestation of neutral objectivity (also Heintz 2007; Singer 2005: 62-67).

Numerous contributions within STS have pointed out that there is no such thing as a subject-free, neutral discovery of the laws of nature. Neither can reality be simply depicted by or transferred into data. Recently, Critical Data Studies have pointed out with regards to Big Data that the notion of "raw data is an oxymoron" (Gitelman 2013), as data are always already 'cooked'. Consequently, also Big Data have to be understood as the product of numerous practices of categorization and classification, of the production of comparability, and of the demarcation between what gets included and what is not, between what is considered as relevant and what is not (see also Heintz 2010; Mau 2017: 30; Busch 2017).

Feminist work in STS has further demonstrated that the modern Western ideal of science resides on the notion of a rational, non-situated, and bodyless subject of knowledge that has been constituted in contrast to the notion of the emotionally bound and physically situated 'others', namely: women and people of color (Singer 2005: 83). Thus, the notion of objectivity as a 'view from nowhere' has to be considered to be androcentric as well as Eurocentric. It has long served to legitimize the exclusion of women and people of color from academia, and continues to marginalize forms of knowledge and modes of reasoning that are based on lived bodily experience or oral traditions (Haraway 2017 [1995]; also Bath 2009).

In the context of Big Data a revival of this modern ideal of western, masculine rationality and subject-free objectivity can be witnessed that potentially reopens the doors for diverse gendered inequalities. Renyi Hong (2016), for example, observes a double marginalization of women in the course of profiling Big Data-methods within human resources (HR): First the association of computing, programming, and analytical skills with masculinity paths the way for discriminating against women professionals in HR. Second, the demand for 'hard numbers' in HR tends to neglect emotional work and other work mostly done by women that is difficult to quantify and model by data. Others draw comparable lessons from historic experiences of quantification efforts in human geography or social physics: The call for numerical representation is considered to favor mechanistic conceptions of the world that tend to be unsuitable to grasp power relations, inequalities, and cultural or symbolic phenomena (Barnes and Wilson 2014: 10; Kitchin 2014: 8; Mazzocchi 2014).

Post-Explanatory Anticipatory Pragmatics

"Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity." (Anderson 2008)

The concept of Big Data shifts the prime object of knowledge production from understanding or explaining a phenomenon – that is asking the *why*- or *how*-questions – to generating probabilistic predictions about a phenomenon that allow for describing or predicting its future appearance. Reasoning then moves increasingly from "data gathered about the past to simulations or probabilistic anticipations of the future that in turn demand action in the present" (Adams et al. 2009: 255), as can currently be observed in fields as diverse as the criminal justice system or credit scoring, where "post-explanatory pragmatics" (Andrejevic 2014: 1675) meet with a "regime of anticipation" (Adams et al. 2009).

The method of choice to implement this new purpose of knowledge production is processing large amounts of data with the help of regression analysis (boyd and Crawford 2012). Regression analysis searches for patterns in the relationship between different variables by calculating how they correlate in a given data sample; proposing a model of this relationship, it then allows for predicting how these variables co-develop in the future. Put differently, big data analysis operates on the assumption that patterns found in data of the past allow for approximating the future.

Geoffrey Bowker (2014) argues that Big Data, by virtue of this methodological approach, offers a possibility of finding patterns for instance in human behavior that do not originate in stereotypical classifications such as women being more social. The tendency, however, to refrain from understanding the correlations identified, quickly turns this advantage into a disadvantage, as can be exemplified through Google's sexist ad posting algorithm: When the fact that women are underrepresented in leading positions is discovered as a pattern in data analysis – which is highly likely due to the gendered segregation of the labor market – and this pattern then forms the basis for choices in ad posting – what can only be assumed due to the lack of transparency of Google's algorithms –, women will by the very means of this data-based technology be less encouraged to make a career, eventually confirming the initially found pattern – or sociologically speaking, reproducing the existing social inequalities. In other words, not asking *why* there are few women in leadership positions, easily leads to misjudging the patterns discovered by data analysis as positivistic expression of the truth, and consequently confirming the gendered structures of social inequality.

The motto "correlation trumps causation" (Bowker 2014) within the concept of Big Data, therefore, rapidly unfolds a conservative tendency, with social inequalities being reproduced by the very means of the epistemological and methodological design of Big Data-technologies.[8] Whereas this may seem harmless when it comes to ad postings, the same logic applies in

[8] Furthermore, it is difficult to contradict the results of Big Data analyses for at least three reasons: First, equipped with the symbolic authority of data, a successful objection requires either alternative data or a well-founded critique of the available data (see for a similar argument Heintz 2010: 172). Whereas the latter requires insight into how the Big Data-analysis at hand operates, the former re-quires considerable resources to come up with data on one's own. Second, decisions based on Big Data analyses are difficult to criticize, because no reason or explanation is given that qualifies for a certain decision (Andrejevic 2014: 1679; O'Neil 2018 [2016]). Ultimately, Lessing's (1999) dictum "code is law" also applies to Big Data. Where there is no human, but a machine behind a decision, there is also no one to direct criticism or objection to.

more serious cases such as racist risk assessment.

Anti-Political Solutionism

"[T]he most important thing we at Facebook can do is develop the social infrastructure to give people the power to build a global community that works for all of us […] – for supporting us, for keeping us safe, for informing us, for civic engagement, and for inclusion of all." (Zuckerberg 2017)

Research concerned with the digital avant-garde of Silicon Valley as one important birthplace of big data describes a "solutionist ethos" as prevalent amongst Big Data evangelists (Morozov 2014; Nachtwey and Seidl 2017). The utopias, being portrayed around digital technologies, depict the world as being full of 'bugs' that need to be 'fixed'. The preferred means to do so, are technological ones, especially ICTs, digital technologies, and last but not least Big Data. The core idea of the promoted anti-political solutionism is that every problem, including social problems, can ultimately be reduced to a series of small and, therefore, manageable problems, for which technological solutions are then to be found. The optimistic belief in technological progress in combination with libertarian ideals and a deep distrust in established politics draws on the so-called "Californian ideology" that has become prominent throughout the second half of the 20th century.[9] Instead of political debate and public opinion formation, ICTs are supposed to create a virtual agora, a public space of discussion, where everyone can speak freely and equally, thereby, pathing the way for democratization, decentralization, and emancipation (Dickel and Schrape 2015).

To make this vision come true, two things are needed according to high-tech solutionist: Humans need to live up to their full potential, which is supposed to be enabled by networking, the distribution and sharing of information, and, therefore, equal access to knowledge and technology. Additionally, all institutions that hinder or restrict the free unfolding of human potential, such as bureaucracy, are to be removed and a strict meritocracy is to be established (Barbook und Cameron 1996; Dickel and Schrape 2015; Nachtwey and Seidl 2017).

At the same time, the protagonists of a Big Data-based solutionism seem to fail to recognize not only the existing inequalities in access to digital technologies, but also the reproduction of power asymmetries and social inequalities within the virtual space (e.g. Zilien and Hargittai 2009). Likewise, the well documented effect that the meritocratic ideal stabilizes existing social inequalities due to its disregard of the deeply embedded structural inequalities in society (Becker and Hadjar 2017) is not problematized any further. As Barbook and Cameron (1996: 49-50) argue, this may be due to the fact that the protagonists of the New Economy form themselves a well-educated, socio-economically privileged, mostly 'white' "virtual class" that is hardly ever affected by racism, social inequality or poverty. From a gender perspective it is to be added that with the rise of Big Data, activities and professions, such as computing, statistics or programming, gain importance that are structurally dominated by men and symbolically associated with masculinity.[10]

With its anti-political solutionism the concept of Big Data privileges a focus on allegedly anti-political, purely factual aspects of reality and social

[9] Barbook and Cameron (1996) describe the "Californian ideology", prevalent in Silicon Valley and related high-tech institutions of the US-westcoast, as a bizarre amal-gamation of "cultural bohemianism", „hippie anarchism" (56), and „anti-corporatism" (52) at the one hand and „economic liberalism" (56), „entrepreneurial zeal of the yuppies" (45), and „laissez faire ideology" (52) at the other.

[10] The recently published anti-feminist manifesto by a Google employee and the following global echo on social media (Bovensiepen 2017) suggest further that sexist work cultures are still prevailing within the work spheres surrounding and implementing Big Data.

life, thus ignoring its highly political and, therefore, inequality-relevant notions. Combined with the insensibility towards power asymmetries and social inequalities, Big Data further runs the risk of misjudging the perspective of a privileged view as universal perspective, rendering those in marginalized positions (again) invisible.

4 DISCUSSION

The discourses surrounding Big Data claim for it to establish a new regime of truth (and governance). The aim of this paper was to introduce the epistemological foundations of Big Data as they are articulated within these discourses to a feminist critique of rationality. Systematically sorting the promises made in the name of Big Data and the critique brought against them, three epistemological premises were discussed as central to understanding Big Data – namely: data fundamentalism, post-explanative anticipatory pragmatism, and anti-political solutionism. This epistemological triad has proven to be anything but (gender) neutral: The revival of the modern ideal of rationality and objectivity within the concept of Big Data links the subject of knowledge production once more to Western masculinity and threatens to marginalize modes of reasoning and aspects of reality beyond the androcentric and Eurocentric norm. The primacy of correlation over causation facilitates the misjudgment of social inequalities as expressions of positivistic truth. These tendencies are reinforced by an anti-political solutionist ethos embedded within the concept of Big Data that renders the privileged virtual class of Big Data protagonists insensitive towards gendered relations of power and inequality. In the light of these findings, the alleged biases of Big Data-based analyses prove to be less the result of random distortions or errors than the systematic consequence of the epistemological foundations of Big Data.

Consequently, this paper argues that Big Data constitutes a specific approach to the world that transports certain possibilities of knowing, and is itself not neutral, but favors the reproduction of existing social inequalities. It does so by (1) privileging phenomena that are easily transformed into (numerical) data and (distinct) categories and that are, therefore, more readily algorithmically processable; by (2) promoting the generation of (probabilistic) knowledge about what there is (or will be), instead of the critical engagement with questions of why specific phenomena have (not) come about; and by (3) favoring the presumably non-political analyses of facts over normative discussions. Future empirical research will have to examine whether this analytical argument holds true across diverse areas of Big Data applications and for different forms of data (such as non-numerical data), as well as whether similar arguments can be made with regards to other digital technologies.

In any case, a critical engagement with the concepts and premises that become materialized in the design of digital technologies is needed, if they are not to (re)produce social inequalities. When it comes to Big Data, this might include that their protagonists acknowledge their own situatedness within social relations of power and inequality and the effects this position has on the design of Big Data technologies and the truth claims that they make. This might also include to acknowledge the limitations of Big Data, for example its tendency to underrepresent already marginalized groups such as the elderly or socio-economically disadvantaged persons (Lazer and Radford 2017). Last but not least, this means to confront Big Data with questions such as: Which interests does Big Data (not) serve? Which questions can Big Data-based analyses (not) ask and answer? Which solutions do Big Data-based analyses focus on?

5 REFERENCES

1. Adams, V., Murphy, M., Clarke, A. E. (2009). Anticipation: Technoscience, life, affect, temporality. Subjectivity, 28 (1), 246-265.

2. Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. WIRED magazine, 16 (7), 16-07.
3. Andrejevic, M. (2014). The Big Data Divide. International Journal of Communication, 8 (2014), 1673-1689.
4. Angwin, J., Larson, J., Surya, M., Kirchner, L., Parris, T. Jr. (2016). Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 23.5.2016. https://www.propublica.org/article/machine-bias-riskassess-ments -in-criminal-sentencing.
5. Barbrook, R., Cameron, A. (1996). The Californian Ideology. Science as Culture, 26 (6/1), 44-72.
6. Barnes, T. J. (2013). Big data, little history. Dialogues in Human Geography, 3 (3), 297-302.
7. Barnes, T. J., Wilson, M. W. (2014). Big Data, social physics, and spatial analysis: The early years. Big Data & Society. DOI: 10.1177/2053951714535365.
8. Bath, C. (2009). De-Gendering informatischer Artefakte: Grundlagen einer kritisch feministischen Technikgestaltung. Universität Bremen: Dissertationsschrift.
9. Becker, R., Hadjar, A. (2017). Meritokratie – Zur gesellschaftlichen Legitimation ungleicher Bildungs-, Erwerbs- und Einkommenschancen in modernen Gesellschaften. In Becker, R. (Ed.). Lehrbuch der Bildungssoziologie. Wiesbaden: VS Verlag, 37-62.
10. Beer, D. (2016). How should we do the history of Big Data? Big Data & Society. DOI: 10.1177/2053951716646135.
11. Bovensiepen, N. (2017): Die weißen Männer von Google. Süddeutsche Zeitung, 09.08.2017. https://www.sueddeutsche.de/karriere/gleichberechtigung-die-weissen-maenner-von-google-1.3620497.
12. Bowker, G. (2014). The Theory/Data Thing. International Journal of Communication, 8 (2043), 1795-1799.
13. boyd, d., Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA. Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, 15 (5), 662-679.
14. Busch, L. (2017). A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large-Scale Data Sets. International Journal of Communication, 8 (2014), 1727-1744.
15. Bührmann, A. (2009). Intersectionality – ein Forschungsfeld auf dem Weg zum Paradigma? GENDER, 2009 (2), 28-44.
16. Crawford, K. (2013). The Hidden Bias in Big Data. Harvard Business Review, 01.04.2013. https://hbr.org/2013/04/the-hidden-biases-in-big-data.
17. Crawford, K., Miltner, K., Gray, M. L. (2014). Critiquing Big Data: Politics, Ethics, Epistemology. International Journal of Communication, 8 (2014), 1663-1672.
18. Daston, L., Galison, P. (1992). The Image of Objectivity. Representations, 40(autumn), 81-128.
19. Datta, A., Tschantz, M. C., Datta, A. (2015). Automated Experiments on Ad Privacy Settings. A Tale of Opacity, Choice, and Discrimination. Proceedings on Privacy Enhancing Technologies, 92-112.
20. Dickel, S., Schrape, J.-F. (2015). Dezentralisierung, Demokratisierung, Emanzipation. Zur Architektur des digitalen Technikutopismus. Leviathan, 43 (3), 442-463.
21. Davis, K. (2008). Intersectionality as buzzword. Feminist Theory, 9 (1), 67-85.
22. DiMaggio, P., Hargittai, E., Celeste, C., Shafer, S. (2004). Digital Inequality: From Unequal Access to Differentiated Use. In Neckerman, K. (Ed.). Social Inequality. New York: Sage, 355-400.
23. Eubanks, V. (2018). Automated Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press.
24. Fourcade, M., Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. Accounting, Organizations and Society, 38, 559-572.
25. Foucault, M. (2012 [1976]). Der Wille zum Wissen. Sexualität und Wahrheit I. Frankfurt/Main: Suhrkamp.
26. Foucault, M. (1978). Dispositive der Macht. Über Sexualität, Wissen und Wahrheit. Berlin: Merve.
27. Geiselberger, H., Moorstedt, T. (Eds.). (2013). Big Data. Das neue Versprechen der Allwissenheit. Berlin: Suhrkamp.
28. Gitelman, L. (Ed.). (2013). Raw Data is an Oxymoron. Cambridge: The MIT Press.
29. Haraway, D. (2004 [1985]). A Manifesto For Cyborgs: Science, Technology, and Socialist Feminism in the 1980s. In Haraway, D. (Ed.). The Haraway Reader. Milton Park: Routledge, 7-45.
30. Haraway, D. (2017 [1995]). Situiertes Wissen. Die Wissenschaftsfrage im Feminismus und das Privileg

einer partialen Perspektive. In Bauer, S., Heinemann, T., Lemke, T. (Eds.). Science and Technology Studies. Frankfurt/Main: Suhrkamp, 369-403.

31. Heintz, B. (2007). Zahlen, Wissen, Objektivität. In Mennicken, A., Vollmer, H. (Eds.). Zahlenwerke. Kalkulation, Organisation und Gesellschaft. Wiesbaden: VS Verlag, 65-85.
32. Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. Zeitschrift für Soziologie, 39 (3), 162-181.
33. Hong, R. (2016). Soft skills and hard numbers: Gender discourse in human resources. Big Data & Society. DOI: 10.1177/2053951716674237.
34. Houben, D., Prietl, B. (Eds.). (2018). Datengesellschaft. Einsichten in die Datafizierung des Sozialen. Bielefeld: transcript.
35. Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. Big Data & Society. DOI: 10.1177/2053951714528481.
36. Lazer, D., Radford, J. (2017). Data ex Machina: Introduction to Big Data. Annual Review of Sociology, 47 (7), 19-39.
37. Lessig, L. (1999). Code and other laws of cyberspace. New York: Basic Books.
38. Mau, S. (2017). Das metrische Wir. Über die Quantifizierung des Sozialen. Berlin: Suhrkamp.
39. Mayer-Schönberger, V., Cukier, K. (2013). Big Data. Die Revolution, die unser Leben verändern wird. München: Redline.
40. Mazzocchi, F. (2014). Could Big Data be the end of theory in science? EMBO reports, 16(10), 1250-1255.
41. Mämecke, T., Passoth, J.-H., Wehner, J. (Eds.). (2018). Bedeutende Daten. Modelle, Verfahren und Praxis der Vermessung und Verdatung im Netz. Wiesbaden: VS Springer.
42. Morozov, E. (2013). To Save Everything, Click Here. Technology, Solutionism and the Urge to Fix Problems that Don’t Exist. New York: Public Affairs.
43. Nachtwey, O., Seidl, T. (2017). Die Ethik der Solution und der Geist des digitalen Kapitalismus. IFS Working Paper, 11. Frankfurt/Main: IfS.
44. O’Neil, C. (2018 [2016]). Angriffe der Algorithmen. Bonn: Bundeszentrale für politische Bildung.
45. Singer, M. (2005). Geteilte Wahrheit. Feministische Epistemologie, Wissenssoziologie und Cultural Studies. Wien: Löcker.
46. Thatcher, J. (2014). Living on Fumes: Digital Footprints, Data Fumes, and the Limitations of Spatial Big Data. International Journal of Communication, 8 (2014), 1765-1783.
47. Van Dijk, J. (2014). Datafication, dataism and dataveillance. Surveillance & Society, 12 (2), 197-208.
48. Weber, J. (2017). Einführung. In Bauer, S., Heinemann, T., Lemke, T. (Eds.), Science and Technology Studies. Frankfurt/Main: Suhrkamp, 339-368.
49. Weizenbaum, J. (1990 [1976]): Die Macht der Computer und die Ohnmacht der Vernunft. Frankfurt/Main: Suhrkamp.
50. Winner, L. (1980). Do Artifacts Have Politics? Daedalus, 109 (1), 121-136.
51. Zillien, N., Hargittai, E. (2009). Digital Distinction: Status-Specific Types of Internet Usage. Social Science Quarterly, 90 (2), 275-291.
52. Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. Journal of Information Technology, 30, 75-89.
53. Zuckerberg, M. (2017). Building Global Community. https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634/.
54. Zweig, K. (2018). Wo Maschinen irren können. Verantwortlichkeiten und Fehlerquellen in Prozessen algorithmischer Entscheidungsfindung. Gütersloh: Bertelsmann Stiftung.