Routledge
Taylor & Francis Group

Information, Communication & Society

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rics20

When data became big: revisiting the rise of an obsolete keyword

Christian Pentzold & Charlotte Knorr

To cite this article: Christian Pentzold & Charlotte Knorr (22 Jun 2023): When data became big: revisiting the rise of an obsolete keyword, Information, Communication & Society, DOI: 10.1080/1369118X.2023.2227673

To link to this article: https://doi.org/10.1080/1369118X.2023.2227673

INFORMATION, COMMUNICATION & SOCIETY
https://doi.org/10.1080/1369118X.2023.2227673

Routledge
Taylor & Francis Group

When data became big: revisiting the rise of an obsolete keyword

Christian Pentzold and Charlotte Knorr

Department for Communication and Media Studies, Leipzig University, Leipzig, Germany

ABSTRACT

This article unpacks the short-lived but momentous buzz around big data. Although talk about big data was once widespread, little is known about the efforts animating its semantics. Tracing this sociotechnical imaginary, we revisit how business insiders and IT commentators fueled the ephemeral yet potent excitement around the term. Our genealogical examination rests on a selection of publications from 2013 to 2017. We employ methods from critical discourse analysis to interrogate how big data was written into being and hyped into a topic of concern. In this aspirational discourse, tech evangelists and writers extrapolated from contexts in which large troves of data were already being harnessed to suggest that inescapable transformations were imminent. They sought to concretize abstract and unfathomable quantities while simultaneously overwhelming their readers with a sense of vastness that exceeds all contexts and outruns the most exuberant expectations. The term may have lost this luster, but big data technologies and practices are an integral part of today's technological infrastructures.

ARTICLE HISTORY
Received 5 December 2022
Accepted 17 May 2023

KEYWORDS
Big data; sociotechnical imaginaries; keyword; genealogy; critical discourse analysis; critical data studies

Remember big data? Around 2012, journalists, IT professionals, and academics were enthralled by it. The term seemed to pop up everywhere, and for a couple of years it dominated public expectations, financial investments, and policymaking around the use of digitally stored, computationally analyzable information before it gave way to the next buzzwords. This public hype somewhat paradoxically coincided with Edward Snowden's 2013 revelations of the enormous surveillance of internet traffic by the Western intelligence apparatus and the ensuing data scandal. In the concomitant yet disjointed aspirational discourse of business insiders and tech commentators, the term 'Big Data' (often but not always capitalized) was used to draw attention to the development of data-intensive tools and instruments as well as to hint at their possible effects, benign or not. Thus, the term denoted technological innovation and carried certain cultural connotations.

Overall, big data encapsulates the idea of exploiting data. More than simply labeling large amounts of information as 'big,' the term refers to the increasing capabilities for

CONTACT Christian Pentzold christian.pentzold@uni-leipzig.de Department for Communication and Media Studies, Leipzig University, Nikolaistrasse 27-29, 04109, Leipzig, Germany

collecting and harvesting large volumes of information in real time (Kitchin, 2014). Besides being a technology, big data was associated with 'the widespread belief that large data sets offer a higher form of intelligence and knowledge … with the aura of truth, objectivity, and accuracy,' as boyd and Crawford (2012, p. 663) put it.

Returning to the era of enthusiasm for big data, our study examines how it was employed in a discourse that accompanied the swift diffusion and uptake of diverse technologies and practices collected under the rubric of big data. We thereby address the lack of knowledge around big data and its genealogy, and we seek to advance endeavors to problematize the assumptions inherent in discourses of datafication (Michael & Lupton, 2016). Moreover, the study contributes to scholarship that troubles the myths, intentions, and politics undergirding taken-for-granted vocabularies such as 'platforms,' 'cyberspace,' 'cloud,' and 'internet.' Like other notions before it and after, big data has conjured a momentous hype, sense of newness, and profound impact. Yet such hyperbole, or 'cyberbole,' as Woolgar (2002) dubbed it, implies neither total novelty nor any straightforward effect because there is a long and multifaceted prehistory of technologies and ideologies underpinning current datafication schemes.

Concepts: imagining new technologies

Our work is inspired by Williams's (1976) analysis of cultural keywords as well as subsequent scholarship scrutinizing the terms and conditions of changing media environments (Katzenbach & Bächle, 2019). Keywords, Williams (1976) explicates, are 'significant, binding words in certain activities and their interpretation; they are significant, indicative words in certain forms of thought' (p. 15). In short: 'A keyword is a socially significant word that does socially significant work' (Peters, 2016, p. xx).

Keywords are an expression of the vital conjunction between discourse and materiality where the construction of technology is inextricably linked with systems of thought and social institutions. Yet the significance of a given keyword like 'Big Data' does not presuppose a definite meaning. Rather, its openness to investment with many different interpretations facilitates its popularity. Intermediaries such as journalists, commentators, and advocates play a powerful role in this context since they channel and propagate the beliefs and ideas associated with big data (Droog et al., 2020).

Sociotechnical imaginaries

Far from being an insignificant choice of words, the nomenclature we use to describe digitally networked services and infrastructures engenders tangible consequences for the formation of contemporary societies. There is no technological precedence here in which technology comes first and discourses second. To the contrary, conceptions of upcoming or nascent technologies and projections of their desirable or undesirable potential are intimately linked to design affordances, scientific discovery, commercial prospects, perceptions of customers, and policy decisions (Mager & Katzenbach, 2021). In principle, that pertains to all kinds of technology, established or emergent and it is this entanglement of visions and socio-technical situations that has been at the heart of classic science and technology studies (Latour, 2005). Arguably, the 'social construction of facts and artifacts,' to use Pinch and Bijker's phrase (1984, p. 47), is

particularly effective in the early stages of a pioneering technology and while its paths of development are still undetermined. Hence, Natale and Balbi (2014) differentiate between imaginaries preceding an invention and those that accompany a new technology when there is 'interpretative flexibility' of its uses, character, and implications.

The concept of sociotechnical imaginaries offers one way of engaging with the interplay between semantics and materiality. At the core, imaginaries refer to 'the ways people imagine their social existence, how things go on between them and their fellows, the expectations that are normally met, and the deeper normative notions and images that underlie these expectations' (Taylor, 2004, p. 23). An imaginary is thus culturally embedded and exists in a collective of people; it is a form of common understanding that also enables a shared practice.

Due to this broad definition, the notion of imaginary has been used to describe many new technologies and the mindsets of those constructing and using them. Both Flichy (2007) and Mansell (2012) gesture to an 'internet imaginary.' Lesage and Rinfret (2015) examine imaginaries of the web, while Bareis and Katzenbach (2022) study AI imaginaries. All of these contributions emphasize that the ideas, visions, and beliefs encapsulated in an imaginary have tangible effects on the course of technological development, policymaking, and public expenditure. To reflect this conjunction terminologically, Jasanoff and Kim (2009) speak of sociotechnical imaginaries that animate technoscientific enterprises, prefigure social constellations, and foreshadow regulatory decisions. Sociotechnical imaginaries are 'collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology' (Jasanoff, 2015, p. 4).

Note that Jasanoff refers to sociotechnical imaginaries in the plural, thus implying that varied imaginaries may be sponsored by different collectives for either concurring or conflicting ventures. To capture the admixture of knowledge, heterogenous collectives of actors in different positions, including experts and citizens, and the prefiguration of socio-technical situations, Irwin and Michael (2003) coin the notion of 'ethno-epistemic assemblages' (p. 111). They not only enable the generation and dissemination of knowledge but are invested in the formation of concrete socio-technical situations. Because these constellations have usually long-term implications for the allocation of resources, enforcing and sustaining a particular sociotechnical imaginary against rival imaginaries is an excise of power, be it political, economic, or cultural (Brown et al., 2000).

Big data discourse

Although the origins of the term 'Big Data' itself remain murky, it has most likely been in use since the mid-1990s and gained traction in 2012 (Lohr, 2013). It has signaled an expansion in the datafication of activities and processes (Kitchin, 2014). Its ideological underpinnings have been described as a kind of 'dataism' that, in van Dijck's (2014) words, 'betrays a belief in the objectivity of quantification and in the potential of tracking all kinds of human behavior and sociality' (p. 201). McAfee and Brynjolfsson (2012) confirm such grand expectations when they stated that 'the big data of this revolution is far more powerful than the analytics that were used in the past.'

At its core, dataism treats data as a sort of 'raw material' that is to be exploited for analytical purposes. Data becomes a natural resource and is thereby turned into a thing that can be renewed and appropriated without limitation (Taffel, 2021). This idea informs the suggestive metaphor of big data as the 'new oil' primed for mining or harvesting (Awati & Shum, 2015; Nolin, 2019). The rationale of data as an asset and valuable resource surfaced not only in business reports and marketing but also informed EU policymaking and the documents and speeches of Brussel officials (Rieder, 2018). Another metaphor views big data as a force of nature to be controlled (Puschmann & Burgess, 2014). Here, often aquatic notions like 'data flood' or 'data deluge' are deployed to capture the sense of abundance and the associated necessity of containment. Existing studies examining the sensemaking around big data either looked at official documents (Couldry & Yu, 2018; Nolin, 2019; Paganoni, 2019; Rieder, 2018), corporate statements (Beer, 2018; Couldry & Yu, 2018; Paganoni, 2019), or journalistic articles (Paganoni, 2019; Pentzold & Fischer, 2017; Pentzold, Brantner, & Fölsche, 2019; Puschmann & Burgess, 2014). Yet missing from scrutiny are the pathbreaking books marketed to a broad audience. They promoted the recently coined notion 'Big Data' and made the concept familiar beyond expert circles. Arguably, it was through nonfiction bestsellers like *The New Digital Age* by Google's then-Executive Chairman Eric E. Schmidt and political advisor Jared Cohen (2014), or US entrepreneur Christian Rudder's *Dataclysm* (2016), that the term and the associated imaginary entered the public agenda. These publications and others like them helped prepare the ground for today's pervasive datafication.

With respect to this body of literature, the task then is to trace how the notion of 'Big Data' came to feature in public discourse and which perspectives it represented. Therefore, we ask: How was big data defined and which frameworks were offered to make sense of it (*RQ1*)? Which expectations of more or less desirable consequences were formulated (*RQ2*)? Which contexts of application were invoked (*RQ3*)? Which implications were deduced and which demands were raised (*RQ4*)?

Materials and method

Our genealogy of the moniker 'Big Data' rests on a selection of influential publications that came out between 2013 and 2017. This timeframe captures big data's heyday when it made headlines in public discourse and appeared on general interest book titles geared toward a mainstream non-expert audience. There are still plenty of books published on 'big data' every year in addition to many other publications. They spell out business strategies or data analytics but they are not making a case for or against big data anymore. Take, for example, Gilder's 2017 publication *The Fall of Big Data and the Rise of the Blockchain Economy* (2017). Today's bestselling books shelved under the tech rubric like *Surveillance Capitalism* by Zuboff (2019) or O'Neil's *The Shame Machine* (2022) have stopped using the keyword despite scrutinizing the implications of datafication, too. In Zuboff's contribution, the notion is not even appearing in the index. Of course, the books still contribute to the public understanding of big data but they have relinquished attracting their readers with the buzzword. When we turn from this supply-side, so to say, to the demand-side, we find that for instance Google searches for 'big data' also peaked between 2015 and 2019. Concentrating on the apex of the broad public reckoning with big data, our sample includes books categorized as public interest

that popularized both the expression and the ideas it represents. While many of these titles were published for English-speaking audiences only, some were translated into other languages.

As a genealogy, the analysis follows Foucault's (1991) idea of writing a history of the present. It traces how current practices and institutions emerge from power struggles that play out in discourses with far-reaching societal and material consequences. A genealogy thus emphasizes the contingency and contentiousness of conditions that are either taken for granted or deemed imperative and constant (Garland, 2014). Whereas an alternative to the present is often difficult to imagine, genealogy points us to the formation of present conditions and their openness to change. For big data, this means scrutinizing the formation of the dominant ideas and expectations that have become associated with the term.

Our selection of books included contributions covering a wide range of topics, that were widely recognized, and that catered to a broad public. Searching for the keyword 'big data,' we relied on book review sections in newspapers and on bestseller lists, and we sourced book recommendations from journalistic offerings and the social cataloging sites Goodreads, LibraryThing, and StoryGraph. The aim was to survey contributions that were shaping and giving meaning to public sensemaking around big data. Collectively, they distilled the expectations and frameworks of understanding, the contexts of application, and implications of big data. These were taken up by media outlets and journalistic writing, with reverberations in the sphere of administration and business. The books, most of which featured the keyword in either the title or subtitle, thus not only reflected core aspects of big data discourse; they also actively provoked further debate and provided points of reference and argumentative cues.

We chose 17 books encapsulating big data discourse based on the public response, critical reception, and appreciation they elicited. They were published between 2013 and 2017. This is a deliberate selection – while other books on the topic existed, these were narrower in scope or more scholarly. Almost all were written by male authors with a professional background in business analytics, management, IT strategy, innovation, cybersecurity, or technology. Of the 22 business insiders and tech commentators included in our selection, three were women. The authors' job titles include consultant, analyst, founder, leader, data scientist, and entrepreneur; others work as journalists, public speakers, writers, legal advisors, or in academia (Table 1). In their biographies, some of the authors style themselves as 'visionary,' 'inventor,' 'cyber enthusiast,' 'thought leader,' or 'Big Data Guru.'

While we did not predetermine the topics in our sample, the available Cataloging in Publication data revealed that most of the books were similar in scope and classified as business intelligence, information technology, consumer behavior, or management. Our sample includes neither academic publications nor the genre of books that denounce data-driven surveillance and control yet avoid the keyword. Moreover, some publications may topically foreground big data to sell copies but actually focus on narrow cases of data-driven practices like Tanner's *What Stays in Vegas* (2016), Sawchik's *Big Data Baseball* (2016), or Mitnick's *The Art of Invisibility* (2019).

The selection favored publications that took a positive view of big data, the only exceptions to this were Payton and Claypool's *Privacy in the Age of Big Data* (2014) and *Weapons of Math Destruction* by O'Neil (2016). There was no decision to exclude other

Table 1. Book selection.

Year of publication	Title	Author(s)	Publisher
2013	Numbersense: How to Use Big Data to Your Advantage	K. Fung	McGraw Hill
2013	Big Data: A Revolution That Will Transform How We Live, Work and Think	V. Mayer-Schönberger & K. Cukier	John Murray
2013	Too Big to Ignore. The Business Case for Big Data	P. Simon	Wiley
2014	Privacy in the Age of Big Data: Recognizing Threats, Defending Your Rights, and Protecting Your Family	T. Payton & T. Claypoole	Rowman & Littlefield
2014	big data@work: Dispelling the Myths, Uncovering the Opportunities	T.H. Davenport	Harvard Business Press
2014	The New Digital Age: Transforming Nations, Business, and Our Lives	E. Schmidt & J. Cohen	Vintage
2015	Big Data: Using Smart Big Data Analytics and Metrics to Make Better Decisions and Improve Performance	B. Marr	Wiley
2015	Humanizing Big Data	C. Strong	KoganPage
2016	Big Data: Does Size Matter?	T. Harkness	Bloomsbury
2016	Data-ism: Inside the Big Data Revolution	S. Lohr	Harper Business
2016	The Internet of Us: Knowing More and Understanding Less in the Age of Big Data	M. P. Lynch	W. W. Norton
2016	The Industries of the Future	A. Ross	Simon & Schuster
2016	Dataclysm: What Our Online Lives Tell Us About Our Offline Selves	C. Rudder	Crown
2016	Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy	C. O'Neil	Penguin
2017	Big Data: How the Information Revolution is Transforming Our Lives	B. Clegg	Icon
2017	What To Do When Machines Do Everything: How to Get Ahead in a World of AI, Algorithms, Bots, and Big Data	B. Pring, P. Roehrig, & M. Frank	Wiley
2017	Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are	S. Stephens-Davidowitz	Harper

dystopian views by design. Yet although the books chosen in our sample cover only some aspects of a wider spectrum of definitions, expectations, contexts, and demands around big data, it was their positive attitude that informed public sensemaking and motivated political and commercial strategy. For sure, there was and is an equally flourishing market for more pessimistic fiction and non-fiction books about the fears and risks of data analytics and pervasive data collection. They too had a share in influencing public sentiments and the critical attitude toward datafication but laid outside of our sampling as they eschewed the keyword, potentially because it was already claimed by tech evangelists

In a sense, the analytical task was to work 'from texts to zeitgeists,' as Streeter (2011, p. 7) called it. To this end, we employed methods from critical discourse analysis (CDA) to examine how big data was written into being and made into a topic of political, technological, economic, and academic concern. Our analysis treated pronouncements on big data as vital elements in public sensemaking without aiming for a post hoc separation of correct claims from incorrect ones. Rather, in line with CDA, we saw the books as strategic elements of persuasion and legitimation that are powerful in shaping policy and technology.

Following a CDAs methodology, we considered not only the content of the books but the underlying rationalities and imaginaries entailed by big data. As recommended by Reisigl and Wodak (2016), we hermeneutically reconstructed rationalities that rest on predication and argumentation. In our case, we focused on the discursive qualifications

of technology and other phenomena, as well as on the justification and questioning of claims. We especially looked for passages with evaluative attributions, predicates, collocations, comparisons, allusions, implicatures, and tropes.

Practically, this meant reading the publications in teams of two before extracting and discussing all suitable passages (which could range from a few words up to entire sentences or even longer paragraphs). The resulting overview of text excerpts, a total of $N = 630$ passages, was first itemized in a table and second semantically clustered (available from the authors upon request). This involved joint meetings that served to validate our analysis and align diverging readings. It was the basis for identifying core issues and themes. Because that was an inherently interpretative procedure, no reliability measures could be calculated. Clustering the issues and themes according to similarities and differences yielded four core aspects of the unfolding sociotechnical imaginary based on the most pertinent rhetorical devices (Table 2). These will be presented following the four research questions about definitions, expectations, contexts, and implications associated with the keyword.

Results: big data hyperbole

In the analysis, it emerges a general – though neither consistent nor necessarily cogent – idea of big data as a revolutionary force. This revolutionary force, in turn, is imagined as unfathomable in size and as having widespread yet inestimable effects in a kaleidoscope of contexts. Consequently, thinking about expectations and implications does not stop with technology but inevitably entails social and political considerations, too.

Due to its focus on overarching lines of reasoning, our analysis cannot do justice to the texts' nuances, and there are differences between the publications that should not be reduced to simplistic caricature. The emphasis instead is on tone and tendency; it looks beyond singular arguments and insights proposed by each of the pieces. To showcase the particular lingo used in the texts, we use exemplary quotes from some of the books.

Definitions: what is big data?

Given the selection criteria, big data featured prominently in the titles of the books in our sample. Yet even though the term was making headlines, its evangelists were conspicuously apologetic about its utility and meaning. In this somewhat ironical situation, they embraced the term 'Big Data' to draw attention to those phenomena and processes that strengthened their argument while simultaneously distancing themselves from it. 'Big Data is undeniably big, but it's also a bit misnamed,' Davenport (2014, p. 1) declared at the start of his book. Similarly, Strong (2015) admitted that 'there is no single agreed academic or industry definition of big data' (p. 7).

In their critique, they and others problematized two things. First, 'Big Data' is a relational term because the size and manageability of datasets are always determined by the available storage and processing facilities (Driscoll, 2012). Hence, data can be either big or small at any point, depending on computational capacity. 'Big' is therefore a vacuous attribute. Second, the notion is transient and will become 'a victim of Silicon Valley's notorious hype cycle,' as Mayer-Schönberger and Cukier (2013, p. 7) put it. Despite

Table 2. Rhetorical devices.

Aspects	Sample passages
Definitions: Big Data is	
… a wave/flood/deluge	'The allusion has a dual resonance: there is, of course, that data as unprecedented deluge. … But there's also the hope of a world transformed – of both yesterday's stunted understanding and today's limited vision gone with the flood.' (Rudder, 2016, p. 31) 'The big numbers behind big data, and the power inherent in those numbers, are impressive. Not long ago, it was said we were living in a time of information 'glut'; we were 'flooded'' (Lynch, 2016, p. 9)
… the new oil	'Big Data is like Big Oil' (Lynch, 2016, p. 9) 'Data is often compared to oil, as the raw material that will power the next industrial revolution.' (Harkness, 2016, p. 25)
… a buzzword	'Big Data is undeniably big, but it's also a bit misnamed' (Davenport, 2014, p. 1) 'unavoidable buzzword' (Lohr, 2016, p. 11) 'there is no single agreed academic or industry definition of big data' (Strong, 2015, p. 7) 'big data won't be a buzz phrase any longer. It will have permeated parts of our lives that we do not think of today as being rooted in analytics' (Ross, 2016, p. 157)
Expectations: Big data is going to be	
… the grandest of technological revolutions	'a transition on par with the invention of writing or the Internet' (Lohr, 2016, p. 15) 'the paradigm-destroying new paradigm' (Rudder, 2016, p. 15) 'The marriage of data and technology is radically changing our world and making it smarter' (Marr, 2015, p. 7) 'The revolution is just getting started' (Simon, 2013, p. 5) 'Big Data allows us to finally see what people really want and really do' (Stephens-Davidowitz, 2017, p. 66) 'More effective communication across borders and languages will build trust and create opportunities for hardworking and talented individuals around the world' (Cohen/Schmidt, 2014, p. 19)
… a business bonanza	'Big Data can revolutionize your business' (Marr, 2015, p. 20) 'allow organizations to interpret previously unimaginable amounts and types of data, and the most progressive organizations are harnessing significant value in the process' (Simon, 2013, p. 24) 'This industry faces a bonanza of big data opportunities' (Davenport, 2014, p. 48) 'reaping the big data bonanza' (Clegg, 2017, p. 89) 'Big data is transitioning from a tool primarily for targeted advertising to an instrument with profound applications for diverse corporate sectors and for addressing chronic social problems' (Ross, 2016, p. 13)
Contexts: Big data is going to happen	
… everywhere	'a veritable Chinese menu of possibilities' (Davenport, 2014, p. 8) 'moving well beyond Internet incubators in Silicon Valley' (Lohr, 2016, p. 9) 'sciences to healthcare, government, education, economics, the humanities, and every other aspects of society' (Mayer-Schönberger & Cukier, 2013, p. 11)
… with global pioneers	'buzz over to Southern California and see big data at work' (Harkness, 2016, p. 26) 'big companies … travel industry' (Davenport, 2014, p. 3) 'Wall Street … investment banks' (Ross, 2016, p. 166) 'The age of big data is coming of age, moving well beyond Internet incubators in Silicon Valley … It began in the digital only-world of bits, and is rapidly marching into the physical world of atoms, into the mainstream' (Lohr, 2016, p. 9)
Implications: Big data requires	
… the right mindset	'in fact, with the right mindset, data can be cleverly reused to become a fountain of innovation and new services' (Mayer-Schönberger & Cukier, 2013, p. 5) 'There is a huge opportunity for brands to make use of the big data but it requires a change of mindset' (Strong, 2015, p. 13) 'numbersense' (Fung, 2013, p. 13) 'to start thinking big' (Simon, 2013, p. xxi) 'open-minded inquisitiveness' (Lohr, 2016, p. 25) 'What is needed most is the vision and determination of organizations to build and deploy these innovations' (Davenport, 2914, p. 41)

championing the notion, the books agreed that it will ultimately cease to be relevant. However, they did not take this prognosis as a reason to stop using it. To the contrary: 'Big Data' as a term may be an obsolete misnomer but big data as a new way of harnessing large troves of data is here to stay, these authors suggested While they assume that the moniker itself will vanish, what it denotes will not. In that respect, Ross (2016) stated that 'big data won't be a buzz phrase any longer. It will have permeated parts of our lives that we do not think of today as being rooted in analytics' (p. 157). This implies

that with big data becoming ever-present, the word itself will fade from use. Transforming from something new to something ubiquitous and taken for granted, the notion of big data may lose its appeal while data themselves become implicated in an increasing number of contexts. In effect, the publications asserted that topic of their inquiry is a profound matter of long-term concern, even if the term 'Big Data' remains a dubious keyword.

The hesitation extends to the definition of big data and thus *RQ1,* which cannot be answered straightforwardly. Again, the books in our sample were elusive and stressed that there is no rigorous meaning. This is not the fault of big data advocates, they suggested, but the consequence of a highly volatile and dynamically emerging field of data-based practices that are difficult to capture. As a result, the term remains an ill-defined catch-all phrase that escapes further attempts at clarification or straightforward specification. It is an umbrella term, and a 'bundle of technologies fly under the banner of big data' (Lohr, 2016, p. 3).

Fathoming big data is challenging, evidenced by the fantastic quantities referenced by the authors. They spoke of terabytes, zettabytes, exabytes, petabytes, or 2.5 quintillion bytes, which is '2.5 followed by eighteen zeros bytes,' as Davenport (2014, p. 11) broke it down. As the writers admitted, these volumes are difficult if not impossible to apprehend in their sheer enormity. Yet these staggeringly large numbers also serve to inspire a sense of awe that may overwhelm readers more than it helps them to grasp their actual volume. On this note, recurrent metaphors depict big data as a deluge, flood, avalanche, or ocean in which things drown in data. 'The big numbers behind big data, and the power, are impressive. Not long ago, it was said we were living in a time of information 'glut'; we were 'flooded,' as Lynch (2016, p. 9) reasoned. In their pioneering analysis of big data discourse, Puschmann and Burgess (2014) likewise maintained that data is often likened to a force of nature, something that needs to be controlled but is difficult to be tamed due to its sheer magnitude. Besides these cataclysmic water-based associations, proponents of big data came up with a variety of other idiosyncratic metaphors to tease out certain aspects of big data practices and technologies. Thus, big data was compared to weed, the yeti, a golf club in a bag, or 'my grandmother' (Stephens-Davidowitz, 2017, p. 32).

Beyond these creative and somewhat absurd comparisons, a number of other images were invoked to describe big data. For instance, many authors discussed the notion of 'raw' data as a valuable resource in need of processing and refinement.

Expectations: what is going to happen?

Many books used superlatives to give expression to the enormous expectations surrounding big data. Its transformative force was likened to the grandest of technological revolutions, and it was dubbed 'a transition on par with the invention of writing or the Internet' (Lohr, 2016, p. 15). Others drew comparisons to the printing press. These technologies and the major transformations heralded by them serve as historical analoges to big data, which, the authors claimed, will unleash similarly profound changes and present massive opportunities to those availing themselves of it. In this rhetorical dramatization, big data is treated as both potent and inescapable; it is 'the paradigm-destroying new paradigm,' as Rudder (2016, p. 15) purported. Disruption is imminent and

unprecedented, and it upends all aspects of life. Big data, the authors concurred, is a game changer. This mirrors the common rhetoric of rupture and upheaval that goes along new technologies, like for instance AI (Bareis & Katzenbach, 2022).

While commentators and visionaries were also keen to concede potentially problematic aspects of this cultural and technological shift, the overall stance was affirmative. It was carried by a fascination with novelty and the promise of 'reaping the big data bonanza,' as Clegg (2017, p. 89) asserted. Most of the improvements made by the use and implementation of big data followed clichéd values of efficiency, effectiveness, and speed (Beer, 2018; boyd & Crawford, 2012). There were widespread expectations that these improvements will help to save time, cut costs, and arrive at more precise decisions. Consequently, many books saw big data's prime areas of application in commerce and business, i.e., in fields where it will boost revenues, curb costs, and guide strategy. Although O'Neil's (2016) criticism is in direct opposition to the thrust of argument, it too started from acknowledging (and denouncing) the exuberance of the big data economy.

To answer *RQ2* – big data is imagined as both omnipresent and consequential. As a result, assumptions of its beneficence are applied to other sectors as well. In these contexts, data-driven or data-fueled decisions are believed to trump intuition. In this imagined competition between datafied analysis and human discernment, the latter is denigrated to a gut feeling and guesswork, while the former is praised as the hallmark of precision and facticity. In some of the pieces, these expectations were tied to the notion of smartness. 'The marriage of data and technology is radically changing our world and making it smarter' (Marr, 2015, p. 7). Not only does big data lead to smarter decisions. 'Smartness' is also attributed to people who know how to harness the potential of data and to technologies that are bolstered by data-driven operations. Overall, big data features as a benign and unthreatening force. While most of the books admitted that there may be payoffs, backlashes, and 'ominous consequences' (Rudder, 2016, p. 15), these were deemed manageable and simply a matter of regulatory adjustment and technical design if one proceeds with a 'clever use of data' (Clegg, 2017, p. 3).

Contexts: where is it going to happen?

Because big data is all-encompassing, virtually no area can be left untouched by it. Responding to *RQ3*, we may argue that big data is treated to be everywhere, becoming part and parcel of a myriad of concrete applications. The publications listed a litany of use cases to illustrate this sense of vastness and the innumerable contexts in which big data plays a decisive role. For instance, Clegg (2017) declared that 'applications of big data are multiplying rapidly and possess huge potential to impact us for better or worse' (p. 5). By stressing the scope of big data, the authors highlighted the importance of examining several cases. As with the use of overwhelmingly big numbers, the message transported by these supercharged statements is that big data's hold on society is increasing and will ultimately become total. Hence, although the lists provided by the authors are limited, they convey a sense that the real number of application contexts shaken up by big data is far bigger and must exceed any overview. 'For every idea I have talked about in this book, there are a hundred ideas just as important,' Stephens-Davidowitz (2017, p. 345) wrote. Indicating the kaleidoscope of potential applications for big data

also demonstrates its versatility. As Davenport (2014) noted, big data is 'a veritable Chinese menu of possibilities' (p. 8). As a result, the multitude of use cases attests to both big data's eclectic potential and the multi-purpose adaptability of data-driven operations.

Some of the pronouncements made in reference to the use cases of big data placed them in a peculiar spatiotemporal setup that is neither fully in the future nor in the present. Here, these authors locate big data in specific venues, most notably Silicon Valley and other sites of technological innovation such as metropolitan centers or leading global businesses. Big data, Simon (2013), for instance, stressed, 'allow[s] organizations to interpret previously unimaginable amounts and types of data, and the most progressive organizations are harnessing significant value in the process' (p. 24). The underlying idea is that these environments are the first sites in which big data can be extensively exploited. These pioneers are spearheading practices and technologies that will eventually become omnipresent. While the potential of big data is currently restricted to certain pilot contexts, so the argument goes, it will become widespread soon. This interjacent spatiotemporal orientation is reminiscent of a quote attributed to William Gibson stating that the future is already here, it is just not very evenly distributed yet. Lohr (2016) expressed a similar sentiment: 'The age of big data is coming of age, moving well beyond Internet incubators in Silicon Valley … It began in the digital only-world of bits, and is rapidly marching into the physical world of atoms, into the mainstream' (p. 9). Relatedly, Harkness (2016) invited readers to 'buzz over to Southern California and see big data at work' (p. 26). The contexts evoked in the books are both test beds and beacons of a big data future that will at one point touch lives across the globe.

However, in their texts mixing reportage with anticipation, big data enthusiasts overlooked questions of agency and authority. They mainly focussed on enterprises and organizations in the Global North that collect and utilize data. These entities were presented as the promoters and profiteers of the big data hype, while large portions of the general public have only few opportunities to either avoid or confront them. While some authors remarked on this discrepancy – 'many people will gain hugely. But many will also be displaced' (Ross, 2016, p. 6) – this was usually a side note to the rejoicing over opportunities. As such, publications tended to marginalize problems and instead concentrated on those who benefit from exploiting data, mostly in terms of control, monitoring, and prediction.

Implications: what has to happen?

The books' bottom line to *RQ4* is that the changes brought by big data are substantial. Yet because they are so enormous and dramatic, the implications are difficult to envision both in their entirety and in detail. In other words, the alleged unprecedented and disruptive force of big data obstructs our ability to account for the many interrelated changes left in its wake. Consequently, the works conceived of themselves as harbingers of momentous transformation. They marked out a frontier from which they looked out onto some of big data's potential future pathways. Yet because this future is still uncertain, the authors highlighted that a shift in mindset is paramount – or as Mayer-Schönberger and Cukier wrote, 'in fact, with the right mindset, data can be cleverly reused to become a fountain of innovation and new services' (2013, p. 5). Lacking a clear roadmap or concrete guidance, readers should adopt a forward-looking stance that embraces

agility, discovery, and the pressure to innovate. 'There is a huge opportunity for brands to make use of the big data but it requires a change of mindset' (Strong, 2015, p. 13). In this regard, the books spoke of an 'open-minded inquisitiveness' (Lohr, 2016, p. 25) and a 'numbersense' (Fung, 2013, p. 13), and they asked their readers 'to start thinking big' (Simon, 2013, p. xxi).

Emphasizing the need for a 'big-data mindset' (Mayer-Schönberger & Cukier, 2013, p. 139) foregrounds the progressiveness and agility of those intending to make use of big data's potential. Technological, regulatory, or ethical problems are ultimately 'people topics' (Davenport, 2014, p. 15) that require courage and commitment to leaving well-trodden paths and expecting constant vicissitudes. Yet the books' emphasis on imagination and anticipation is diminished by the call to seize the critical and opportune moment. Whereas imagination was seen as important, taking action was essential. Being decisive was deemed particularly critical at this junction because such decisions would determine future fortunes and failures. Ironically, these contributions advise their readers to take their time to be imaginative, yet they must do so quickly. This call for leadership and action is similarly at the core of discourse around AI where latent opportunities demand firm initiative (Bareis & Katzenbach, 2022).

Discussion: apodictic aspirations of big data

Despite all the uncertainties involved in handling and envisioning big data's potential, its evangelists convey a sense of futurity: They might not know the particulars of what is to come but they clearly know that change is here and that big data will bring many beneficial opportunities, if harnessed properly. In other words, the publications claim to know what is going to happen, at least the broad strokes of general directions and obvious implications. They support their apodictic aspirations by gesturing toward different spatiotemporal zones. The proponents thus extrapolate from places where large troves of data have already been exploited to potentially global transformations. The either auspicious or daunting data-driven future is already here but it is not evenly distributed. In effect, change is not about to happen but is already happening somewhere, and soon it will be everywhere. Arguably, it is this hyperbolic exuberance and solutionist conviction that prevent the evangelists to grapple with big data's shortcomings. In that somewhat perplexing pattern, they acknowledge the downsides yet refuse to seriously engage with them. The only stark contrast is provided by O'Neil (2016) who welcomed her readers to 'the dark side of Big Data' (p. 13). Her account has urged many more critical responses which however avoided the keyword, arguably because of its affirmative connotations established by big data evangelists.

Essentially, the sociotechnical imaginary evoked by the books sees big data as a powerful and instrumental entanglement of large troves of numerical information, technological capacities, analytical ambitions, and human audacity. While many technicalities and specifications remain unclear or unfold dynamically, the books transport the idea that big data is mainly about a mindset of utilizing computational data to the maximum. In this ambitious plan, big data is not only associated with fields of application like election campaigning, marketing, retail, state intelligence, or logistics (Kitchin, 2021). Big data also gestures toward solutions for problems affecting society as a whole, such as climate

change, pandemics, or war. Big problems require big data, and big data can solve big problems, so the solutionist quintessence.

Yet all the crises big data is meant to resolve are bedeviled by an unprecedented level of opacity and complexity. Hence, the future big data is meant to steer is mired in intractable mysteries. Unsurprisingly, many of the publications also indicate an ambiguous understanding of big data as either threat or promise depending on the course of action, thus reiterating common templates of emerging technologies as either utopian or dystopian. Even though our sample mainly included publications that promoted big data, indicated by the prominent use of the keyword in the books' titles, they still acknowledged its detrimental effects and consequences. For instance, as Marr put it, big data has 'a seedy underbelly that is so disturbing I don't even want to write about it' (2015, p. 143). Nevertheless, the overall tone of the books is aspirational, and the arguments foreground opportunities and beneficial data usage.

It is indeed surprising that the publications retain this position after the Snowden revelations in 2013 and the public reckoning with massive data-driven surveillance. Although the disclosures spotlighted the misuse of big data and its potential adverse side effects, the books in our sample maintained a positive attitude and highlighted big data's many advantages. While the authors are not ignorant of the problematic aspects of datafication and data-based control, these are mostly reframed as resolvable regulatory issues and technological tasks. The problems become bugs, not innate features; a stance which echoes the solutionist beliefs endorsed by the IT industry, particularly Silicon Valley conglomerates (Levina & Hasinoff, 2017). As such, these beliefs also animate the ideology of dataism, that prides itself on not faltering in the face of emerging troubles (van Dijck, 2014). Rather, problems are taken as motivation for repair and improvement. As Clegg (2017) summed it up, '[w]e can talk of the pros and cons of big data. We can worry about the bad and ugly while appreciating the good. But we have to accept that big data is not going away' (p. 149). Harkness (2016) seconded this, stating that '[w]e should not let our preoccupations … distract us from its real potential' (p. 285).

This ties in with Beer's (2018) notion of a 'data imaginary,' which he specifies as 'powerful visions of what data can achieve, what they can solve, how they might help us to thrive, what they are able to reveal and how they are able to make us more informed, efficient or better at things' (p. 14). He reconstructs the data imaginary from public marketing material and frames data analytics as speedy and accessible. These data analytics are meant to reveal hitherto untapped value. Furthermore, the data imaginary depicts data analytics as panoramic, all-embracing, and prophetic, i.e., as forward-looking and prognostic.

In the books we read, little thought is, however, spend on the unequal prerequisites that may hinder a potentially diffusion. In the metaphorical analogy to forces of nature, big data seem to blaze abroad naturally thus glossing over the many divides in technology access and availability as well as in use patterns and capital. Attending to harbingers, the books instigate to follow their example and make sure that big data can spread. And by adopting an affirmative stance of persuading and convincing their readers of big data's precedence, the authors diminish the downsides and forms of data activism (D'Ignazio & Klein, 2020; Milan & van der Velden, 2016). What is more, the idea of pilot sites ignores to acknowledge that datafication is from the start a global affair that predicates on extractive capitalism (Couldry & Mejias, 2019; Milan & Treré, 2019).

Furthermore, the solutionist reasoning in the books resonates with a more or less firmly pronounced determinism. Although big data's success is said to hinge on smart people embracing opportunities, the books emphasize its tangible and straightforward outputs. 'Data is the input, and the output is smarter choices and wiser judgments,' as Lohr (2016, p. 35) concluded. Like a giant machine, the entanglement of big data sets with practices, attitudes, and technologies is seen to process information and produce outputs. This resonates with the notion of 'raw' data that is fed into machine operations and yields results. In the publications headlining big data, most effects are presented as beneficial, that is, they are rendered in terms of effectiveness, efficiency, and precision.

More pernicious consequences are again side-lined. In their effort to stress how big data empowers organizations and actors, its enthusiasts lose sight of questions of data power, of inequalities in terms of data access and computational equipment, and of the limited ability to opt out of datafication schemes (Kennedy & Bates, 2017). Consequently, the books spend little time on calls for data justice and claim that anyone embarking on data analytics for whatever purpose is accountable for their own actions. The elitist discourse of makers and pioneers addresses readers as adepts who can use big data to full capacity; it passes over those who are nevertheless implicated either by laboring in global supply-chains, from mines to retail, or in data handling and monitoring. Their enormous contribution remains unacknowledged. Fascinated with the exploitation of data, the books tend to overlook the exploitation of humans upon which their audacious visions rest.

Conclusion

Reviewing the trajectory of big data, it seems that the keyword has ceased to enthrall recently. It has been superseded by more mundane notions of datafication that dispense with grandiose gestures toward size, volume, or velocity of data. While the term may have become obsolete, its associated technologies and techniques are an integral part of today's infrastructures: 'Don't doubt it – big data is here to stay,' Clegg (2017, p. 4) therefore urged. Simon (2013) concurred: 'Big Data is no temporary blip or fad. In fact, it is going to intensify in the coming years, and its ramifications for the future of business are impossible to overstate' (p. xxi). This argument may explain the advocates' ambiguous appreciation of the term. Although they continue to use it, they emphasize their personal dislike for the 'umbrella classification' (Simon, 2013, p. 77) and 'vague concept' (Stephens-Davidowitz, 2017, p. 18). However, the analytical attitude and operations associated with big data remain pervasive despite the ill-defined and vacuous nature of the term.

Like other innovations big data technologies gradually transform from the new to the habitual. As such, they seem to follow the usual trajectory of digital media in which ubiquitous implementation coincides with terminological evanescence. While they may disappear both from view and from colloquial language, these practices and tools have become part of the taken-for-granted infrastructure upon which current society rests. Speaking about the comparable demise of the notion of 'cyberspace' that coincided with the omnipresence of networked communications, Mosco (2004) points out how ironic it is that 'as these once-new technologies lost their luster, gave up the promises of contributing to world peace, and withdrew into the woodwork, they gained a power that

continues to resonate in the world' (p. 2). Beyond the utopian or dystopian visions associated with keywords, the protean technologies they denote become crucial elements in social and economic change. The auspicious sociotechnical imaginary surrounding big data thus comes to bear upon concrete commercial investments, political decisions, and cultural sensemaking. It treats big data as the transformative harnessing of large quantities of digital data whose radical potential touches virtually all contexts, frustrates neat definitions while exceeding expectations, and challenges us to acquire a new mindset.

The books we examined were instrumental in creating a discursive atmosphere in which the use of big data became a duty, not a matter of choice. The ambivalent rhetoric used in the pieces is one element of the excitement they prompted. On the one hand, the books promise to unpack the endless possibilities of big data. On the other, they constantly evoke its potential to exceed any imagination and explanation. The proponents of big data were thus confident in expounding its tremendous value and could not exaggerate the superior and unparalleled significance of 'amazing, remarkable, breathtaking, groundbreaking data' (Stephens-Davidowitz, 2017, 21f). Big data, in a sense, is too big to contain in a book.

While discontent and extensive criticism of big data certainly existed, it was largely absent from our limited sample (which was also skewed toward more expectant visions). These critical points form the core of a number of works with a more dystopian message. In a divided book market in which few attempts are made to balance the views, they follow a different publication strategy than the books analyzed in our sample.

Rather than merely reflecting general practice, the authors we studied were actively engaged in pushing the issue and keeping it relevant. They participated in creating the hype that surrounded big data as protagonists – not observers – by repeatedly invoking its phenomenal opportunities. They contributed by 'spreading the big data gospel' (Simon, 2013, p. 219). It is also due to their efforts that some tropes such as big data as the 'new oil' or the 'data deluge' now sound familiar and arouse little scrutiny (Couldry & Yu, 2018). They belong to the cultural repertoire of speaking about big data. It is imbued with a strong sense of revolutionary vigor that no sector of contemporary society can resist. Imagining big data as crucial and indispensable has thus fostered the expansion of today's data-driven operations and prefigured their implementation. In effect, a post-hoc check of their visions against today's reality proves difficult since they were pre-empting the future. Instead of waiting for their prognostications to unfold, the sociomaterial imaginary of big data was implicating more immediate actions which in hindsight render it an almost self-fulfilling prophecy. This is not to say that challenging the status quo is impossible or futile. Quite the contrary: reconstructing the auspicious discourse is in itself a form of critique that shows how the present situation was evoked and is not depending on technological imperatives.

The big data hype may be over, yet it is the well-rehearsed playbook of grand visions and concrete postulations that remains available and forms a principal component of technology diffusion. As of now, it can, for instance, be observed in the momentous fascination for chatbots and it also propelled the excitement around the metaverse a while ago. Such apodictic aspirations nurture an expectant public ready to embrace a new technology's potential which sets the course for further investments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the German Research Foundation under [grant number 447465824/PE 2436/3-1].

Notes on contributors

Christian Pentzold is Chair and Professor of Media and Communications in the Department for Media and Communication Studies at Leipzig University. He is interested in the construction and appropriation of digital media and the roles information and communication technologies play in modern society. In current projects, he looks at the public understanding of big data, humans interacting with robots, the smart home, the organization and governance of peer production, as well as the interplay of time, data, and media.

Charlotte Knorr is a research associate in the Department for Media and Communication Studies at Leipzig University. Since April 2021, she has been working in the Framing Big Data research project. There, she studies the discursive framing of datafied processes both in professional communicative forms and user-generated content. Additionally, Charlotte is a trained mediator with a focus on Transformative Mediation for group disputes and conflicts.

ORCID

Christian Pentzold http://orcid.org/0000-0002-6355-3150
Charlotte Knorr http://orcid.org/0000-0001-6591-3781

References

Awati, K., & Shum, S. B. (2015). *Big data metaphors we live by*. Medium. https://towardsdatascience.com/big-data-metaphors-we-live-by-98d3fa44ebf8

Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values*, *47*(5), 855–881. https://doi.org/10.1177/01622439211030007

Beer, D. (2018). *The data gaze: Capitalism, power and perception*. Sage.

boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Brown, N., Rappert, B., & Webster, A. (2000). *Contested futures*. Ashgate.

Clegg, B. (2017). *Big data*. Icon.

Couldry, N., & Mejias, U. (2019). *The cost of connectivity*. Stanford University Press.

Couldry, N., & Yu, J. (2018). Deconstructing datafication's brave new world. *New Media & Society*, *20*(12), 4473–4491. https://doi.org/10.1177/1461444818775968

Davenport, T.H. (2014). *big data@work*. Harvard Business Press.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Driscoll, K. (2012). From punched cards to "big data": A social history of database populism. *Communication+1*, *1*, 1–33. https://doi.org/10.7275/R5B8562P

Droog, E., Burgers, C., & Kee, K. F. (2020). How journalists and experts metaphorically frame emerging information technologies: The case of cyberinfrastructure for big data. *Public Understanding of Science*, *29*(8), 819–834. https://doi.org/10.1177/0963662520952542

Flichy, P. (2007). *The internet imaginaire*. MIT Press.

Foucault, M. (1991). Nietzsche, genealogy, history. In P. Rabinow (Ed.), *The Foucault reader* (pp. 75–100). Pantheon.
Fung, K. (2013). *Numbersense*. McGraw Hill.
Garland, D. (2014). What is a "history of the present"? On Foucault's genealogies and their critical preconditions. *Punishment & Society*, *16*(4), 365–384. https://doi.org/10.1177/1462474514541711
Gilder, G. (2017). *Book Title*. Regenery.
Harkness, T. (2016). *Big data*. Bloomsbury.
Irwin, A., & Michael, M. (2003). *Science, social theory & public knowledge*. Open University Press.
Jasanoff, S. (2015). Future imperfect. In S. Jasanoff & S. Kim (Eds.), *Dreamscapes of modernity* (pp. 1–33). University of Chicago Press.
Jasanoff, S., & Kim, S.-H. (2009). Sociotechnical imaginaries and nuclear power in the United States and South Korea. *Minerva*, *47*(2), 119–146. https://doi.org/10.1007/s11024-009-9124-4
Katzenbach, C., & Bächle, T. C. (2019). Defining concepts of the digital society. *Internet Policy Review*, *8*, 1–6. https://doi.org/10.14763/2019.4.1430
Kennedy, H., & Bates, J. (2017). Data power in material contexts: Introduction. *Television & New Media*, *18*(8), 701–705. https://doi.org/10.1177/1527476417720034
Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481
Kitchin, R. (2021). *The data revolution*. Sage.
Latour, B. (2005). *Reassembling the social*. Oxford University Press.
Lesage, F., & Rinfret, L. (2015). Shifting media imaginaries of the web. *First Monday*. https://doi.org/10.5210/fm.v20i10.5519
Levina, M., & Hasinoff, A. A. (2017). The silicon valley ethos: Tech industry products, discourses, and practices. *Television & New Media*, *18*(6), 489–495. https://doi.org/10.1177/1527476416680454
Lohr, S. (2013, February 1). The origins of 'big data'. *York Times*. https://archive.nytimes.com/bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/
Lohr, S. (2016). *Data-ism*. Harper Business.
Lynch, M.P. (2016). *The internet of us*. W.W. Norton.
Mager, A., & Katzenbach, C. (2021). Future imaginaries in the making and governing of digital technology: Multiple, contested, commodified. *New Media & Society*, *23*(2), 223–236. https://doi.org/10.1177/1461444820929321
Mansell, R. (2012). *Imagining the internet*. Oxford University Press.
Marr, B. (2015). *Big data*. Wiley.
Mayer-Schönberger, V., & Cukier, K. (2013). *Big data*. John Murray.
McAfee, A., & Brynjolfsson, E. (2012, October). Big data: The management revolution. *Harvard Business Review*. https://hbr.org/2012/10/big-data-the-management-revolution
Michael, M., & Lupton, D. (2016). Toward a manifesto for the 'public understanding of big data'. *Public Understanding of Science*, *25*(1), 104–116. https://doi.org/10.1177/0963662515609005
Milan, S., & Treré, E. (2019). Big data from the south(s): Beyond data universalism. *Television & New Media*, *20*(4), 319–335. https://doi.org/10.1177/1527476419837739
Milan, S., & van der Velden, L. (2016). The alternative epistemologies of data activism. *Digital Culture & Society*, *2*(2), 57–74. https://doi.org/10.14361/dcs-2016-0205
Mitnick, K. (2019). *The art of invisibility*. Little, Brown & Company.
Mosco, V. (2004). *The digital sublime*. MIT Press.
Natale, S., & Balbi, G. (2014). Media and the imaginary in history. *Media History*, *20*(2), 203–218. https://doi.org/10.1080/13688804.2014.898904
Nolin, J. M. (2019). Data as oil, infrastructure or asset? Three metaphors of data as economic value. *Journal of Information, Communication and Ethics in Society*, *18*(1), 28–43. https://doi.org/10.1108/JICES-04-2019-0044
O'Neil, C. (2016). *Weapons of math destruction*. Penguin.
O'Neil, C. (2022). *The shame machine*. Crown.
Paganoni, M. C. (2019). *Framing big data*. Palgrave Macmillan.

Payton, T., & Claypoole, T. (2014). *Privacy in the age of big data*. Rowman & Littlefield.
Pentzold, C., Brantner, C., & Fölsche, L. (2019). Imagining big data: Illustrations of 'Big Data' in US news articles, 2010–2016. *New Media & Society*, *21*(1), 139–167. https://doi.org/10.1177/1461444818791326
Pentzold, C., & Fischer, C. (2017). Framing big data: The discursive construction of a radio cell query in Germany. *Big Data & Society*, *4*(4). https://doi.org/10.1177/2053951717745897
Peters, B. (2016). Introduction. In B. Peters (Ed.), *Digital keywords* (pp. xiii–xlv). Princeton University Press.
Pinch, T., & Bijker, W. (1984). The social construction of facts and artifacts. In W. Bijker, T. Hughes, & T. Pinch (Eds.), *The social construction of technological systems* (pp. 17–50). MIT Press.
Puschmann, C., & Burgess, J. (2014). Metaphors of big data. *International Journal of Communication*, *8*, 1690–1709.
Reisigl, M., & Wodak, R. (2016). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse studies* (pp. 23–61). Routledge.
Rieder, G. (2018). Tracing big data imaginaries through public policy. In A. Sætnan, I. Schneider, & N. Green (Eds.), *The politics of Big data* (pp. 89–109). Routledge.
Ross, A. (2016). *The industries of the future*. Simon & Schuster.
Rudder, C. (2016). *Dataclysm*. Crown.
Sawchik, T. (2016). *Big data baseball*. Flatiron Books.
Schmidt, E., & Cohen, J. (2014). *The new digital age*. Vintage.
Simon, P. (2013). *Too big to ignore*. Wiley.
Stephens-Davidowitz, S. (2017). *Everybody lies*. New York: Harper.
Streeter, T. (2011). *The net effect: Romanticism, capitalism, and the internet*. NYU Press.
Strong, C. (2015). *Humanizing big data*. KoganPage.
Taffel, S. (2021). Data and oil: Metaphor, materiality and metabolic rifts. *New Media & Society*, *25*(5), 980–998. https://doi.org/10.1177/14614448211017887
Tanner, A. (2016). *What stays in Vegas*. New York: PublicAffairs.
Taylor, C. (2004). *Modern social imaginaries*. Duke University Press.
van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776
Williams, R. (1976). *Keywords*. Oxford University Press.
Woolgar, S. (Ed.). (2002). *Virtual society? Technology, cyberbole, reality*. Oxford University Press.
Zuboff, S. (2019). *The age of surveillance capitalism*. PublicAffairs.