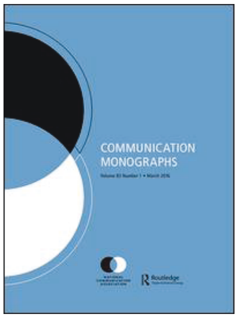


Communication Monographs

ISSN: 0363-7751 (Print) 1479-5787 (Online) Journal homepage: http://www.tandfonline.com/loi/rcmm20


# Situating methods in the magic of Big Data and AI


M. C. Elish & danah boyd



**To cite this article:** M. C. Elish & danah boyd (2017): Situating methods in the magic of Big Data and AI, Communication Monographs, DOI: 10.1080/03637751.2017.1375130

**To link to this article:** http://dx.doi.org/10.1080/03637751.2017.1375130

Published online: 19 Sep 2017.

COMMUNICATION MONOGRAPHS, 2017
https://doi.org/10.1080/03637751.2017.1375130



# Situating methods in the magic of Big Data and AI

M. C. Elish [a] and danah boyd [b]

[a]Data & Society Research Institute, New York, NY, USA; [b]Microsoft Research and Data & Society Research Institute, New York, NY, USA

**ABSTRACT**
"Big Data" and "artificial intelligence" have captured the public imagination and are profoundly shaping social, economic, and political spheres. Through an interrogation of the histories, perceptions, and practices that shape these technologies, we problematize the myths that animate the supposed "magic" of these systems. In the face of an increasingly widespread blind faith in data-driven technologies, we argue for grounding machine learning-based practices and untethering them from hype and fear cycles. One path forward is to develop a rich methodological framework for addressing the strengths and weaknesses of *doing* data analysis. Through provocatively reimagining machine learning as computational ethnography, we invite practitioners to prioritize methodological reflection and recognize that all knowledge work is situated practice.

**ARTICLE HISTORY**
Received 9 August 2017
Accepted 18 August 2017

**KEYWORDS**
Methodology; Big Data; AI; machine learning; epistemology; ethnography



> I found that the people who ascribe the most power to statistics and data are not people who do statistics and data science. They are executives who give the vision talks about the power of data. … I've seen so many cringe-inducing assertions … In my head, I'm listening to all these things and am like, I remember that conversation, and the data on which that is based is so utterly flawed and unlikely to be true. But it supports the mythos of this particular executive. So let's just repeat it until it is true. (Hammerbacher 2016, data scientist)

In 2015, the M.D. Anderson Cancer Center in Texas made a bet that IBM Watson was going to play a vital role in the Center's core mission: to eradicate cancer. Both IBM and M.D. Anderson promised that Watson would revolutionize cancer care, allowing for more personalized treatment and enhanced clinical decision-making (IBM News Room, 2013). In the year following its deployment, a scathing University of Texas audit documented organizational and resource mismanagement (University of Texas, 2016), while critics pointed out that the "cognitive intelligence" technology itself overpromised and under-delivered (Jaklevic, 2017). Costing over $62 million, the project was abandoned in early 2017, never having been fully implemented (Ackerman, 2017).

IBM Watson is just one of an emerging class of technologies being branded as "artificial intelligence" (AI). These technologies have risen to prominence in the last year as the latest game-changer in the tech industry. Only a few years ago, the same might have been said about Big Data, and indeed according to a recent *New York Times* article, AI has been dubbed the "new Big Data" in many circles (Hardy, 2016).

**CONTACT** M. C. Elish mcelish@datasociety.net

Sparkling, spotless, and new, the imaginaries and connotations of these two intertwined technologies promise a future that is scientifically perfectible, assuring a pot of gold at the end of a digitally coded rainbow. The purportedly neutral collection and analysis of large quantities of data promise to present insights that can transcend human limitations. Yet, Big Data and AI must be understood as *socio*-technical concepts. That is, the logics, techniques, and uses of these technologies can never be separated from their specific social perceptions and contexts of development and use. As work in the social sciences has demonstrated, the techniques and practices underpinning Big Data reveal the ways in which social values are encoded into mathematical processes and automated through techniques that scale normative logic (Ambrose, 2015; O'Neil, 2016).

Still, through the manufacturing of hype and promise, the business community has helped produce a rhetoric around these technologies that extends far past the current methodological capabilities. As we argue in this paper, the ability to manufacture legitimacy has far-reaching implications. Not only does it trigger innovation and bolster economies, but it also provides cover for nascent technologies to potentially create fundamentally unsound truth claims about the world, which has troubling implications for established forms of accountability (Barocas & Selbst, 2016; Citron, 2008; Crawford & Schultz, 2013; Zarsky, 2016).

This article begins by unpacking the histories and sets of cultural claims at stake in Big Data and AI, elaborating how these terms are interconnected. To build our argument, we draw on our ethnographic experience interrogating how socio-technical systems are designed, developed, deployed, and integrated in society. We show that dominant conceptions of how Big Data and AI work do not align with the actual techniques of machine learning. Part of what makes the phenomena of Big Data and AI so compelling is the hyped imagination of what is possible, not what is realistic. It is precisely this slippage that produces an epistemological hazard and that requires concerted attention to the methods of data science and machine learning. The article concludes with a provocation, envisioning machine learning metaphorically as computational ethnography. Through this lens, we draw on the methodological knowledge that has emerged from anthropological ethnography in order to argue for the development of reflexive practices in machine learning.

While many argue that this is the dawning of the age of Big Data and AI, those who have lived through previous hype cycles cannot help but echo the mantra that "winter is coming." The key to grounding machine learning-based practices is untethering the work from hype and fear cycles and developing a rich methodological framework for addressing the strengths and weaknesses of both the practices and the claims that can be produced through technical analysis. In other words, it is time for technical practice to develop a reflexive approach.

## The making of a field

### *Origin stories*

In many regards, there is nothing new about either Big Data or AI. As techniques of quantification ordered by the logics of (neo)liberal governance and capitalism, both technologies take shape from within long histories of operationalizing statistics for business profit, population control, and governance (Foucault, 2009, 2010; Rose, 1991). More specifically, the techniques that comprise Big Data and AI reflect long histories of slow, technical development aimed at achieving discrete outcomes (Jones, 2016). Still, far from being precise in public parlance, these terms stir up a set of mythologies about automated and data-driven technologies (boyd & Crawford, 2012; Gillespie, 2014). In order to fully understand the innovations and imaginaries that underpin Big Data and AI, it is necessary to situate the relevant technical orientations and practices historically and socially. Doing so requires drilling into the ways in which those invested in such a phenomenon embrace cultural fantasies and encode particular agendas into technology.

Throughout this article, we often purposefully integrate Big Data and AI into one concept to focus on the phenomenon we are interrogating. Nonetheless, these terms have different roots and many practitioners would take issue with how these terms are used in public discourse. Moreover, the boundaries of what constitutes "AI," as a field of research as well as an aspirational goal, are nebulous and often contested. Rather than relying on rigid definitions, we begin by providing brief histories of these terms in order to draw out the social contexts and research cultures from which they emerged.

## *The rise and fall of Big Data*

Big Data was born of big business. The specific techniques of Big Data date back to at least the 1990s, but the term entered business discourse through a 2001 Gartner report that defined Big Data as the "3Vs": volume, velocity, and variety (Laney, 2001). It is worth pointing out that while the term Big Data was a neologism, collecting data, and using statistics to measure and manage populations dates back centuries (Hacking, 1982; Igo, 2007). Nevertheless, since the early 2000s, numerous scholars and pundits have attempted to offer alternate definitions that scope out both the technology and practices underpinning the phenomenon of Big Data. For example, after listing 10 different operating definitions used in different contexts, Gil Press (2014) offers two of his own: (1) "the belief that the more data you have the more insights and answers will rise automatically from the pool of ones and zeros" and (2) "a new attitude by businesses, non-profits, government agencies, and individuals that combining data from multiple sources could lead to better decisions."[1] More than the familiar business rhetoric of "volume, velocity, and variety," Press's definitions are useful because they highlight aspects of the mythologies that underpin Big Data.

By 2010, technology companies and other enterprises began focusing on Big Data as a new business paradigm (Manyika et al., 2011). Consulting firms emerged to help companies wrangle their data, while technology companies focused on selling their "cloud" server and "software as a service" offerings to help companies store and manage their data. Non-profits and government agencies began to feel as though they too needed to use data to "get smart." To address these needs, education institutions and funding agencies began rebranding statistics and computer science efforts as "data science" (Lohr, 2015). Meanwhile, less well-intended companies emerged to prey on anxious organizations by selling Big Data solutions that were little more than vaporware.[2]

As more organizations and people embraced Big Data, critics started questioning the value and purpose of such analytics. In particular, journalists, civil society advocates, and scholars began raising questions about data-driven approaches in sectors such as criminal justice, education, and employment (Barocas, Rosenblat, boyd, & Gangadharan, 2014; Ramirez, Brill, Ohlhausen, & McSweeny, 2016; Rosenblat, Wikelius, boyd, Gangadharan, & Yu, 2014). They also turned to question long-standing but ever-increasing data-centric practices in credit, insurance, and advertising (Poon, 2016; Turow, 2011). The changing responses of the Obama White House to these technologies offers an example of how attitudes about Big Data slowly shifted. Initially, enthusiastic about the economic potential of Big Data, the White House convened a series of experts in 2014 who ended up highlighting both the opportunities and concerns related to this emerging field of technologies. The White House's first report in 2014 on "seizing opportunities and preserving values" was quite optimistic (Podesta, Pritzker, Moniz, Holdren, & Zientz, 2014). Yet, by 2016, their second Big Data report focused on "algorithmic systems, opportunity, and civil rights" and painted a much more concerning portrait about the potential of data discrimination (Muñoz, Smith, & Patil, 2016). Concerns focused on the amount of personal data being collected and sold, as well as the potential misuse of these techniques to increase inequality and do harm.

Some journalists began framing Big Data as a new form of "big brother" (Cellan-Jones, 2015) – and because the term Big Data seemed to focus on the data rather than the models or analysis involved in the practice, the phenomenon began to lose its sheen within business communities, many of whose leaders feared being associated with surveillance and discriminatory uses of data (boyd, 2016). Furthermore, because of the hype surrounding Big Data, any data analytic practice, regardless of its technical sophistication, was being shoved under the umbrella of hype. Technology-centric companies who were primarily interested in using large quantities of data to power advanced machine learning algorithms, which they felt could be tremendously useful in doing sophisticated analysis and predictive work, began using new rhetoric to differentiate themselves (Levy, 2016). By late 2015, technology companies that were once seen as being at the forefront of Big Data began rebranding their efforts as "AI."

## *Artificial Intelligence: old and new*

While "AI" represented something new to those looking to repackage Big Data, AI itself is decades old. The concept of AI, in its contemporary sense, first came into use during the 1950s, and crystallized during the Dartmouth Summer Research Project on Artificial Intelligence (McCorduck, 2004).[3] Bringing together the latest advances in the "system sciences" (Mindell, 1998), including cybernetics, information theory, systems theory, and cognitive science, researchers during this time predicted rapid advancements in solving "the artificial intelligence problem" (McCarthy, Minsky, Rochester, & Shannon, 1955). Indeed, the unbridled optimism that currently surrounds machine learning and new AI technologies recalls these first decades of AI research, when predictions about future greater-than-human capabilities of AI dominated public discussions of the technologies (Dreyfus, 1972).

The above predictions, however, were far from the reality of the slowly developing software and hardware of AI. During the 1950s and 1960s, millions of dollars, mostly originating from various arms of the Department of Defense, were directed toward AI research "centers of excellence" at universities such as MIT, Stanford, and Carnegie Mellon University. By the mid-1970s, funding for AI research began to dry up, a period known in computer science departments as "the AI winter." A scathing British government report (Lighthill, 1973) essentially declared the project of AI a failure. Moreover, as priorities shifted within American defense agencies that had been funding AI research (such as DARPA, the Defense Advanced Research Projects Agency), less resources were available for unrestricted basic research, and the latitude with which researchers could experiment contracted (Edwards, 1996; Mirowski, 2003).

While the majority of AI research during the first decades was theoretical or limited to experiments in academic labs, work in the area of "expert systems" in the late 1970s marked the first time AI research could be clearly and successfully applied in commercial industry settings (Russell & Norvig, 1995, pp. 21–22). These "expert systems," also called "knowledge systems" or "knowledge-based systems," were conceived as supplements or sometimes replacements for complex decision support in professional settings, such as medical diagnostics. Information was gathered from human experts (usually only one or two) and encoded into rules and procedures that made up the computer system (Forsythe, 1993). In this way, expert systems were intended to emulate human expert decision-making in complex contexts. Proliferating throughout the 1980s, their popularity faded by the mid-1990s. Expert systems came to be perceived as "brittle," working only in limited contexts with less than perfect results (Agre, 1995; Forsythe, 2002; Suchman, 2007).

As hope for widespread adoption of expert systems receded, a new set of techniques to achieve computer intelligence gained prominence.[4] This work involved a different set of assumptions about intelligence compared with the previously reigning approach that centered on the role of logic and reason in abstractly represented models of the world. In addition to developments in "behavior-based" robotics (Brooks, 1991), techniques in the field of machine learning attracted research and development attention, including natural language processing, computer vision, and neural networks. Rooted in cybernetic conceptualizations of command and control, neural networks had in fact been proposed as one of the primary approaches to the Artificial Intelligence problem in the 1950s, and were so named because the concept behind how they work was inspired loosely by how neurons in the brain are thought to function. However, the technique was quickly derided by leading researchers at the time, and branded as an unfeasible approach to AI (Olazaran, 1996). Gaining renewed interest in the 1980s, new research demonstrated the ways in which it could be effectively put to use in certain kinds of problems, such as object and speech recognition (Olazaran, 1996). While more discrete than "expert systems," research in the fields of machine learning and neural nets showed promise for effective transitions into successful products like optical character recognition. Machine learning and deep learning, what many have described as the "rebranding" of neural nets (Elish & Hwang, 2016, p. 13), are driving the renewed attention to AI. Combined with large datasets and concentrated human talent, machine learning is accomplishing what seemed impossible a few years ago. These advances have been closely integrated with commercial companies, and are possible only in the context of the vast data sets, increased computing power, and widespread business commitments to Big Data.

Our intention is not to provide a full accounting of the histories of Artificial Intelligence or Big Data, nor to implicitly suggest a teleological or techno-deterministic narrative of innovations in technology. Rather, by providing a sketch of Big Data and AI development, we want to call attention to the distinct epistemologies at stake in machine intelligence (Kitchin, 2014). The different paradigms of intelligence within AI research have different implications for how knowledge, truth, and fact can be articulated and leveraged in specific social contexts. Specifically, early AI research, which focused on abstracted symbolic representations of human knowledge and procedural logic and is now termed "good, old fashioned AI" (Haugeland, 1985), stands in contrast to the techniques of machine learning, in which knowledge is derived by crunching vast amounts of data, detecting patterns, and producing probabilistic results. In these equations, "meaning" is beside the point; the algorithm "knows" in the sense that it can correlate certain relevant variables accurately. It does not matter if a system thinks like a human – as long as it appears to be as knowledgeable as a human.

### *Blurry edges*

AI, as a category of technology, always waivers between the real and the imaginary. On the one hand, Western perceptions of what AI is – what it can and cannot do, and what it might yet do – are informed by long-standing cultural imaginaries of machines that escape the control of their creators, and the promises and perils of automata and artificial life (Franchi & Guzeldere, 2005; Riskin, 2007).[5]

On the other hand, as we will argue further below, contemporary discourses around AI rely on the *potentials* of such technologies as much, if not more, than current functionalities. Popular media coverage often, albeit inadvertently, reinforces a blurring of the line between fantasy and reality. For instance, news coverage of the deployment of predictive policing in American cities inevitably references the science fiction thriller, *Minority Report* (Koepke, 2016). A particular example illustrates the point: a CBS Sunday Morning (2015) news segment covering the DARPA robotics challenge showcase was reported on and discussed through an interview with the director of the science fiction film, *Ex Machina*, which featured a sentient robot (CBS, 2015). That is, the morning news show used the director of a popular science fiction movie as the main commentator for the culmination of the most prestigious federally funded robotics challenge. Such a framing is only possible if there is an assumption that the science fictions of AI are relevant to the actual functioning of AI. If the topic was medicine, would such a segment have been produced? Would it be appropriate to have an actor who plays a doctor on television be the main commentator on the latest medical advances?

## Manufacturing legitimacy

### *Technologies of magic*

X.ai is a company that has created the popular "AI powered personal assistant" named Amy. The home page of the company promises that using Amy will simplify and streamline the stressful "email ping pong" of scheduling a meeting (X.ai, n.d.). According to the company, after you have signed up, all you have to do is "Cc: Amy" on the email exchange, "handing the job over to her." In a section entitled "How It Works," we are told: "Amy emails with your guest to find the best time and location, knowing your schedule and preferences. Like magic, the meeting invite arrives in your inbox."

This description of how Amy works – like magic – is a common refrain in the marketing materials of new technologies, especially those involving AI.[6] When technologies are said to "work like magic," a recognizable English idiom, we might understand this to connote the ideas of impressive and seamless functionality, in which the end effect or experience is amazing, and the means by which the effect was achieved is irrelevant or even secret. This reinforces Arthur C. Clarke's often-repeated axiom that "any sufficiently advanced technology is indistinguishable from magic" (1973: p. 21). Suggesting that a technology "works like magic" in casual speech or marketing copy serves as a way to express praise, while also reinforcing a sense that how the technology works is unknowable and inscrutable (Selbst, 2017, pp. 89–93).

In a brief essay on the correspondences between magic and technology, anthropologist Gell (1988) proposed that a defining feature of magic, as an orientating framework of actions and consequences in the world, is that it is "'costless' in terms of the kind of drudgery, hazards, and investments that actual technical activity inevitably requires. Production 'by magic' is production minus the disadvantageous side-effects, such as struggle, effort, etc." (Gell, 1988, p. 9). To evoke magic is not only to provide an alternative regime of causal relations, but also to minimize the attention to the methods and resources required to carry out a particular effect.

## AI games

High profile business performances around machine intelligence are often used to sell products by selling a vision of success. In 1997, IBM triggered confidence in machine intelligence when its Deep Blue algorithm beat the reigning world chess champion, Garry Kasparov. The company garnered headlines again in 2011 with "Watson," a "cognitive intelligence" program designed to succeed at *Jeopardy!*. Beginning in 2015, Google's DeepMind took center stage through performances of AlphaGo, a program leveraging new techniques in neural networks that were designed to win at the ancient game of Go, a game that was long thought to be harder to beat than chess. Computer scientists had predicted that it would be decades before a computer would win at Go, and so the relatively unexpected and utter success of AlphaGo amplified the promises that AI would soon surpass the confines of straightforward machines, and even recreate intuition, that most human and ineffable element of human intelligence (Nielsen, 2016).[7] In the business section, a *New York Times* article reporting on another recent and highly publicized match in China opened with the statement, "It isn't looking good for humanity." Later the article declared: "the victory … showed yet another way that computers could be developed to perform better than humans in highly complex tasks, and it offered a glimpse of the promise of new technologies that mimic the way the brain functions" (Mozur, 2017).

While Google has not parlayed the success of AlphaGo into a direct product release, such high profile experiments engender confidence in the state of a company's technology, help attract technical talent to the company, and help solidify the importance of AI in the future. Watson, discussed in further detail below, has since been commercialized and narrated in business conversations as a tool for finding a cure for cancer, supporting customers in retail, and solving long intractable problems in education (Captain, 2017). These highly publicized experiments-as-performances also help shape the public perception that machine intelligence is better or more advanced than human intelligence, and that it works perfectly, every time.

Employing games, such as Chess and Go, is a significant and recurring theme in the history of Artificial Intelligence research. The influence of games has shaped research agendas as well as implicitly prioritized certain kinds of intelligence over others (Ensmenger, 2012). Moreover, and key to our argument here, is that the narratives around such games, when they are performed for a public audience, serve to obfuscate the true state of the field. Underneath the sheen of performativity is a stark reality that the current capabilities of AI systems, like Watson or AlphaGo, are quite narrow. Tasks must be discretely defined and the analytics within these systems are only as good as the data upon which the analysis depends. Although new data sets are increasingly available, the quality of these data vary tremendously and, all too often, limitations in the data mean that cultural biases and unsound logics get reinforced and scaled by systems in which spectacle is prioritized over careful consideration of the implications of long-term deployment (Crawford et al., 2016).

When AI proponents and businesses produce artefacts and performances to trigger cultural imaginaries in their effort to generate a market and research framework for the advancement of AI, they are justifying and enabling a cultural logic that prioritizes what is technically feasible over what is socially desirable. This is not to say that advancements in Big Data and AI are themselves societally damaging or that all who are using hype to drive the development of technology are misguided or ill intended. For example, data scientists like Jeff Hammerbacher, quoted at the beginning of this article, recognize that hype is part of what makes efforts in medicine, biology, and genetics possible. While the advances promised by "precision medicine," an approach that tailors medical care to individual patients through individually specific datasets, are only at beginning to be possible, this set of techniques is driving ground-breaking new research and treatments. Yet, what's typically at stake in the hype surrounding Big Data and AI is not the pursuit of knowledge, but the potential to achieve a performative celebration of snake oil.

## Watson at work

American audiences were introduced to IBM's Watson in 2011 through the game show, Jeopardy! One of the most notable aspects of the appearance was the extent to which Watson was personified. The program produced speech in a male synthetic voice with a standard American accent and was effectively embodied in a small flat screen with an image of a globe constantly orbiting small circles and radiating lines. In turn, IBM's "cognitive intelligence" computer program, effectively became a character, certainly not human, but also somehow more than simply machine.

During the contest, the text of each question was transmitted electronically to Watson. After the program had searched for and calculated a correct response, the program would signal and, when "called upon," vocalize a response. Viewers at home were able to watch Watson "think," meaning that as Watson responded to a question, a horizontal bar graph at the bottom of the screen would indicate the top three most likely responses that Watson had computed (Figure 1). While this visual aid underscored that a computer program was calculating and assessing the probability of correct responses, the overall performance and scene-setting of the *Jeopardy!* challenge produced a set of implicit, if slippery, equivalences between humans and Watson. Moreover, as a male character Watson was shaped by humorous moments during the contest, such as when Watson asked for the category "Chicks dig me" (about female archaeologists) without any self-conscious irony. Moreover, although Watson was named after the founding family father and son who ran IBM, it is a happy coincidence that the name Watson also can be an allusion to Sherlock Holmes' trusted sidekick. Watson ultimately beat the two reigning (human) champions of *Jeopardy!* in a landslide victory.

Four years later, IBM's cognitive computing platform was again center stage for American audiences through a series of prime-time TV commercials. In one of the commercials, Watson, with a similar synthesized male voice and black rectangular screen with a dynamic globe, holds a conversation with Bob Dylan and offers to collaborate with him. In another commercial, Watson interviews filmmaker Sir Ridley Scott. Several months later, another IBM commercial aired during the 2016 Academy Awards featuring Watson. The scene opens, in this highly viewed and widely circulated commercial, on what appears to be a support-group for evil robots, led by a therapist played by Carrie Fisher (an actress who became famous for her role in the science fiction epic *Star Wars*). After complaining about how they, the evil robots, are no longer needed or revered in the world, we learn that Fisher has invited Watson to confront the group, and Watson explains, in the synthesized voice, "I can understand, reason, and learn *with* humans" (Mark Summers, 2016). Watson goes on, in the first person, to explain the many useful things it can do. The other robots interrupt with disgust. The commercial is humorous and, like other commercials from the IBM ad campaign, communicates that Watson is not like any other AI or robot that has come before.

IBM's earlier chess-playing program DeepBlue was a public relations success but never resulted in business products. With Watson, the teams behind the system development always had in mind that *Jeopardy!* was to be just the first stage of a longer and more expansive product development initiative (Angelica, 2011). While Watson was designed specifically to play *Jeopardy!*, developers of the Watson system intentionally focused on techniques that could also be applied to a range of contexts and problems (Thompson,


**Figure 1.** IBM Watson competes on US TV show Jeopardy, 11 January 2013 (Photo credit: IBM).

2010). Today, dozens of Watson products work in fields ranging from customer service chat agents to healthcare diagnostics.

A common misperception is that Watson is one entity, like its on-screen portrayals, and can be applied potentially anywhere and everywhere. The marketing and business narratives that animate Watson as an artificially intelligent system – although IBM uses the term "cognitive intelligence" – produces an imagined mobility in which Watson has enough intelligence to move around and between domains. However, moving seamlessly between contexts is precisely one of the unsolved challenges facing AI research (Stanford, 2016). Watson is best understood as a platform upon which specific datasets can be analyzed in order to produce "intelligent" responses (Talbot, 2009). Watson products are developed and fine-tuned to specific domains and business problems, resulting in a different system every time; the Watson that provides answers to customer queries on the Geico website is not the same as the Watson built from medical information used at the Cleveland Clinic.

Yet, the marketing of Watson, from its debut on *Jeopardy!* to recent advertisements, personifies the product and encourages a specific interpretation of what Watson is, calling upon and perpetuating a series of mythologies about automation and AI. In this instance, not only is technology awe-inspiring, but it is also endowed with a unique form of agency, or set of capacities, that is generally considered the domain of human beings. The language that has emerged around Big Data and machine learning further encourages an equation between human and machine intelligence by invoking human or biological activities: data are "fed" to a computer and "digests" information and machines "learn" and "think." This, in turn, sets up new kinds of equivalencies that are worth interrogating. The aspect worth consideration is not so much that we are willing to attribute agency to non-human entities, but rather, what kinds of agency and with what expectations do such attributions emerge (Suchman, 2007)?

Consider, for example, how the capabilities of a Watson product are described in the context of healthcare:

> A human doctor will go through 16 years of education before entering medical school, and even then he or she still has to pass through internship and residency before practicing medicine. Once past that threshold, "Dr. Human" will still need several decades of experience to become a qualified cancer expert. Watson will probably need two to five years of "learning" to become as knowledgeable. It is the complexity of the domain that defines how long Watson will take to learn it. (Terdoslavich, 2015)

This narration of Watson's capabilities and the ways in which its "intelligence" is achieved creates substantial elisions around what constitutes "learning" and becoming "knowledgeable." Implicitly, "a qualified cancer expert" is reduced to the amount of digital data that can be processed. Elsewhere in the article, and in most marketing materials, we are assured that this is the future of healthcare. Such elisions characterize media coverage and general discussions around AI, compounding the notion of technological inscrutability with a glossing over of technological limitations. Furthermore, "promissory rhetorics" of AI (Weber & Suchman, 2016) suggest that any shortfalls in the system will be solved in the near future. However, these shortfalls are constituent of how current AI systems work (Ekbia & Nardi, 2014; Irani, 2015). By calling upon a future that is imminent but always just beyond reach, what technologies can currently do is not as important as what they might yet do in the future. It is enough that they appear to work, just like magic.

### *Faith in prediction*

Shortly following the 2016 U.S. Presidential election, journalists began reporting on the role of a little known company called Cambridge Analytica in shaping public perception through the use of Big Data. In the opinion section of *The New York Times*, McKenzie Funk described how the company built psychological profiles based on responses to Facebook quizzes, enabling direct targeting of advertisements, including news stories, to Facebook users (Funk, 2016). Alongside her article, the *Times* published an image credited to Yoshi Sodeoka of a white hand puppet controlling websites and people (Figure 2).

Profiling the psychologist whose work on psychometrics undergirds Cambridge Analytica, *Motherboard* ran a story with the headline: "The Data That Turned the World

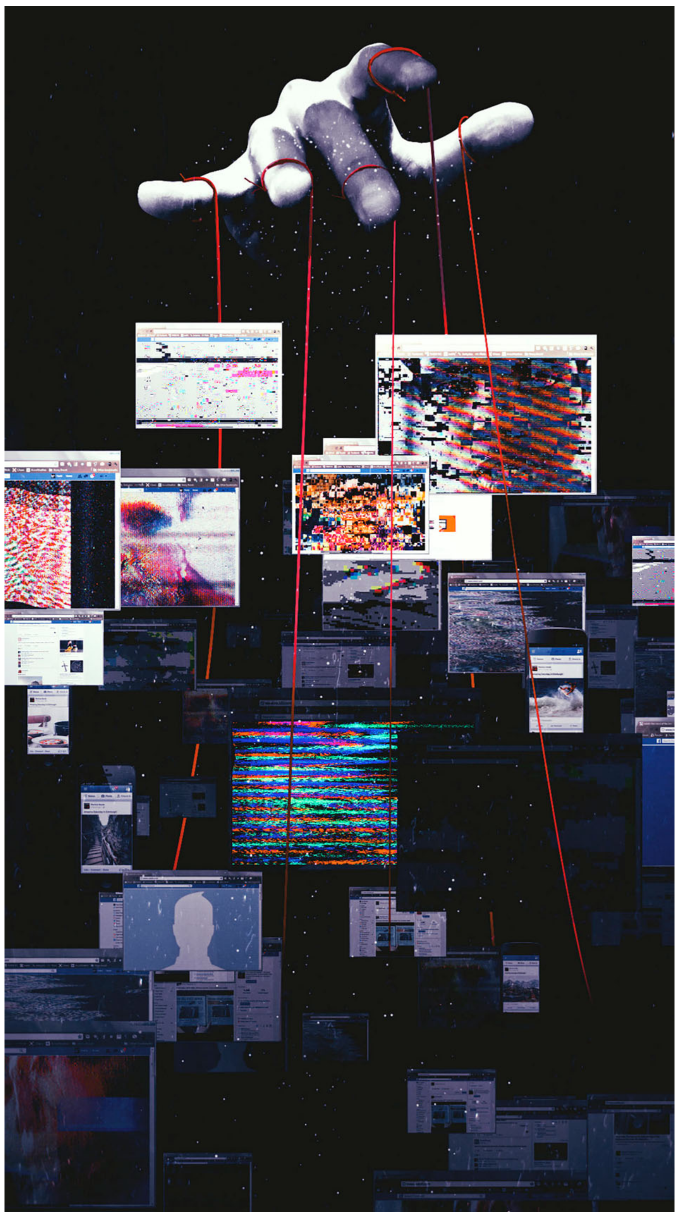

**Figure 2.** *New York Times* Illustration, Yoshi Sodeoka (Funk, 2016).

Upside Down" (Grassegger & Krogerus, 2017). *The Guardian* produced a tell-all piece about "how our democracy was hijacked" that reads almost like a conspiratorial accounting, linking together different funders, political operatives, and companies into a web of actors with Cambridge Analytica at the center; both the company and others named in the story have sued the newspaper (Cadwalladr, 2017). Many more articles have focused on the fact that Robert Mercer – a hedge fund billionaire and computer scientist known for conservative philanthropy – had invested in the service, along with providing funding directly to Donald Trump's presidential campaign and investing in sites like *Brietbart*, known best for its deeply conservative pro-Trump news coverage (Mayer, 2017). Articles like these helped trigger panic throughout Democratic circles that the Republicans had developed a "dark campaign arts" (Confessore & Hakim, 2017) approach to using data and algorithms to effectively profile and manipulate people.

Wary of the salacious articles and doubting the company's marketing hype, tech reporters and scholars began interrogating the claims put forward (Sauter, 2017). Lacking any public evidence of efficacy or ability to scrutinize the technical work of the company, *MIT Technology Review* turned to the top researchers in the field, all of whom raised serious concerns about the company's statements, effectively arguing that the state of the field was simply not that advanced (Talbot, 2016). Even Republican political consultants started noting that, "You get a lot of snake oil like this in data work" (Taggart, 2017). Although advertising networks like Google Adsense have long used machine learning techniques to target ads, data scientists within the advertising industry note that their techniques are better than demographic categorization, but they are not nearly as good as people believe them to be (C. Perlich, personal communication, September 4, 2014). For their part, Cambridge Analytica proclaims that it "uses data to change audience behavior."

Critiques did not dent the widespread view that Cambridge Analytica helped develop a "Weaponized AI Propaganda Machine" (Anderson, 2017). Months later, in June 2017, *The Atlantic* ran a piece titled "Hillary Clinton Was the First Causality in the New Information Wars" which drew on Clinton's statements at a public technology conference where she argued that she lost because of "weaponized" technology that was used against her. Alongside claims about Russian bots and "fake news," she explicitly called out Cambridge Analytica, highlighting the investment that Mercer and other conservative operatives made in the company as proof that it could not be snake oil. In doing so, she helped propagate the view that Cambridge Analytica had the Big Data tools necessary to manipulate the election.

The widespread belief that Cambridge Analytica's approach can (and did) work builds on a more general notion that advanced technology companies have both the data and knowledge to accurately model people and provide targeted interventions. Whether the topic at hand is "personalized" learning, "precision" medicine, or "predictive" policing, an increasingly large number of non-technical experts are devoted to implementing technical solutions to address long-standing and seemingly intractable problems in fields as varied as education, medicine, and criminal justice. Yet, for the technical experts working on those efforts, there is widespread awareness that the computational reality is far from the idealized narrative. Among computer scientists, Ps like personalization, precision, and prediction are goals that motivate and drive their work, not accurate depictions of the state of the art.

The endemic hype and fear surrounding Cambridge Analytica highlights how convincing rhetoric around AI systems can be. Lost are the technical realities and analytic possibilities. As with any potent frame, the story that an AI system has been built to manipulate the public's hive mind is hard rhetoric to combat.

## Epistemological duct tape

Public discourses surrounding AI and Big Data may help generate interest in the field, spur financial investment, and trigger research and development. But they also obfuscate the actual practices involved in *doing* machine learning and data science. As with any nascent and emerging field, what is behind the curtain is full of contested boundaries and uncertainties, methodological challenges and epistemological pitfalls. From machine learning research to industry practice, there are tremendous gaps between analytic ideals and current state of the art. Those who work to build AI systems must simultaneously push against the science fiction narratives that allow outsiders to value their work, while also trying to celebrate and promote the contributions that such analysis can offer.

AI has the potential to collapse under the weight of the hype that surrounds it, even though the analytic contributions it has to offer have significant potential. Caught in the middle of a spectacle, there is limited room for the field to develop strategic methodological sensibilities that can ground its analysis and claims. In this section, we turn to examine the actual work of doing machine learning and data science in order to offer a different way of understanding the practice itself; if we reimagine AI as a new form of qualitative inquiry, the power and potential of the field look quite different. More importantly, such a reframing suggests new opportunities for methodological development.

### *Building and interpreting models*

When the glitz of AI hype is brushed aside, a great deal of mundane work underlies the practices of doing machine learning. This work includes collecting, cleaning, and curating data, managing training datasets, choosing or designing algorithms, and altering code based on outputs. In addition, as with any development process, engineers must grapple with the practical tasks of debugging and optimization, not to mention making sense of poorly documented code written by others. Such work may appear "purely technical." However, it is through this minutia where cultural values are embedded into systems. Every step requires countless decisions and trade-offs. In an imaginary if ideal world, code is bug-free, data are straightforward, and algorithms are perfect fits for the desired task. Reality is much messier.

In computer science, terms like "information" and "communication" lose their colloquial sensibilities alongside concepts like "learning" and "intelligence." In everyday speech, information is presumed to have meaning and communication is assumed to involve interpretation. Yet, in technical practice, these notions are decoupled (Gleick, 2012). For instance, when working with data and building models, the substantive value of the data is often irrelevant. As computer scientist and roboticist Bill Smart points out, while a facial recognition algorithm is called a face detector, it is more accurately a "set-of-pixel-values-that-often-correlate-well-with-the-presence-of-faces-in-the-training-data-that-you-collected-detector, not a 'face detector'" (Smart, 2016). A facial recognition system does not *know*, as a human would, what is or is not a face. Rather, it is a system that is designed to categorize incoming data based on a model that was produced using previously tagged data. The mechanism of validation is not rooted in teaching a computer the intrinsic meaning of what is a face. Rather, the processes of validation include: (1) providing training data that humans have associated with being a face, (2) developing an algorithm that learns to detect which features of that data reliably are associated with faces, and (3) evaluating the features of new data to infer whether or not the data fits the model.

This is not to say that semantic meaning is not part of the process. Because the data can be correlated with socially significant categories (e.g., "a face"), the system is being designed to associate data features to a human concept. Moreover, as Seaver (2015b) highlights, algorithmic systems, such as recommendation systems, are crafted through precise attention to the context of data. However, what context is relevant, and even what *constitutes* context is where opinions diverge. Still, the actual implementation of computing the meaning and the significance of associations within datasets requires converting a set of contextual and socially embedded relationships into data that is machine-readable. Algorithms then operate over data that is disconnected from meaning, seeking signals that convey information about the mathematical properties of the encoded content. This process is not about a search for meaning, but about the construction and depiction of statistical models.

Any quantitative scholar working on tabulations can lose track of the holistic view of the data when navigating data and statistical questions. Regressions must be turned into mathematical questions while data about people are boiled into values sitting in rows and columns. Rigorous statisticians know how to move between the mathematical constructs and the conceptual analysis. Yet, machines do not manipulate social constructs – they manipulate numbers. When computers serve as tools to support a human's analysis, the ideal scenario is one in which an analyst responsibly narrows the appropriate questions based on the data and has the tools to understand the limitations of the statistical results (Leonelli, 2014). The larger and more complex the data, the less practical that is. Computers may be able to computationally analyze multi-dimensional data sets with thousands of intertwined features, but in doing so, the task escapes the cognitive capacity of any human, rendering unique forms of "algorithmic opacity" (Burrell, 2016, p. 3) and limiting forms of accountability.

Data scientists know that data are never perfect. Missing data, input errors, and conflict data schemas all plague practitioners. They must clean the training data to address weaknesses, while also assessing how constructed categories and data outliers might contort the model. Yet, classic work by Bowker and Star (2000) offers a clear and profound elaboration of the ways in which the very practices of articulating categories and deploying classification standards are thoroughly socially and historically contingent, irrespective of intention. Purportedly neutral and self-evident categories, from race to types of disease, emerge from complex social and political contexts and have consequences far beyond the specific, isolated categorization at stake.[8] Yet, for a machine learning system to work, data scientists must make choices about how to provide discrete labels and generate bounded categories for sensitive topics or in cases where such boundaries are far from solidified. In designing the system, they must also account for how unexpected or even malicious user behavior could produce problematic or unexpected data that then is used to re-train a system through feedback loops. Increasingly, researchers in machine learning are looking to techniques in "adversarial machine learning" to better understand vulnerabilities in technical models where intentionally designed training data could corrupt the results (Papernot et al., 2017).

Not only does the construction of models hold significant epistemological implications, but also the ways in which models may be interpreted generates epistemological fault lines in the kinds of truth and knowledge that AI systems produce. Because machine learning results can be difficult to interpret, there is a danger that data scientists might inappropriately use the results when converting them back into conceptual information for decision-making. One problem, common to all statistical practices, is the production of spurious correlations, or results that suggest connections between data that have no real connection (Calude & Longo, 2016). There are many websites dedicated to showing entertaining but spurious correlations in statistical datasets, such as a connection between "U.S. spending on science, space, and technology" and "suicides by hanging, strangulation, and suffocation" (Vigen, 2015). Another common risk is overfitting, such that the model is designed to explain noise in the data rather than the underlying phenomenon. Because it is impossible for a system to detect whether or not the correlation it has identified is meaningful or if the model it is using has been overfit, these common mistakes require a qualitative understanding of what the quantitative data means.

The decisions that data scientists must make when analyzing data do not just require understanding the context of the data, but also require a deep understanding of how the data may be used or transformed by the algorithmic system as it is deployed in the social world. In designing and testing a system, a programmer may examine different outputs to see if the results seem reasonable. Yet, once deployed for public use, recommendations, predictions, and classifications produced by technical systems are often accepted as uncontroversial until a result challenges socially constructed assumptions. For example, it is the "uncanny valley"[9] of a recommendation that is too good – or one that is absolutely absurd – that renders visible the automated nature of a recommendation. In addition, when recommendations appear offensive or have unexpected cultural significance, the public is quick to challenge the decision-making of the system. For example, the Android store recommended that Mike Ananny download a "Sex Offender Search" app after he installed "Grindr," a gay male dating app. As he wrote in *The Atlantic*, there are many conceivable explanations for this statistical connection, but the actual outcome is offensive to the human eye because of the homophobic belief that gay men are sex offenders (Ananny, 2011). Algorithmic systems do not inherently compute the range of culturally specific interpretations that are acceptable or those to avoid; they must be given machine-readable information telling them when associations may appear problematic.

The choices that inform these systems and the challenges in interpreting their results reveal the limitations when we construct technological tools to solve inherently social problems. Because computational systems require precise definitions and mathematically sound logics, sociocultural phenomena that are typically nuanced and fuzzy are rendered in coarse ways when implemented into code, formalizing boundaries and erecting divisions where none previously existed. Fundamentally, the practices of building AI systems – of doing machine learning or data science – cannot be divorced from the social contexts in which these technologies are situated. Seemingly straightforward categories and mundane assumptions stack up with unanticipated ripple effects. Models that made sense in one instance are incorrect in another, or undermined by malicious or unwitting actors. As a result, a sophisticated developer of an AI system cannot simply build the perfect system and let it loose in the wild with full confidence that it will work as expected. Rather, the practices of machine learning and data science require many of the methodological tools that are required to understand cultural practices more generally.

### *Machine learning as computational ethnography*

Machine learning is not a science, at least not in the traditional sense. Unlike disciplines that leverage the scientific method as a tool for interrogating phenomena, machine learning techniques do not require formulating hypotheses rooted in earlier theories to test for validity. Most machine learning models are constructed based on an initial exploration of the data and evolve through supervised or unsupervised processes to fit the data. While the decisions involved in fitting the data require systematically drawing on an understanding of theory and earlier discoveries to make strategic choices for analysis, those deploying machine learning systems are often asked to explain commonsense correlations, justify spurious connections produced by the system, or contend with how strategic business decisions may have led to overfitting. However, effective prediction – not interpretability – has thus far been the expectation and rewarded goal for machine learning models. Because of decisions made during the labeling, cleaning, and modeling of data, compounded by how the models change in response to new data, the results from most systems cannot be easily reproduced. Researchers have begun examining what they perceive to be a growing crisis in validity in AI and machine learning research.[10]

The challenges that machine learning faces as a field are not unique. Ethnography has been there. Although the epistemological frameworks appear incongruent, the ways of building knowledge have striking parallels (Seaver, 2015a). Of course, there are significant differences, including the scale and amount of data examined. Nevertheless, similar to those doing machine learning, ethnographers surround themselves with data ("a field site"), choose what to see and what to ignore, and develop a coherent mental model that can encapsulate the observed insights. They identify and piece together meaningful data from particular instances, and then seek to generalize insights gleaned from the particulars and unspoken details of everyday life, or what Bronislaw Malinowski, a founding figure of ethnographic methods, once termed, "the imponderabilia of actual life" (Malinowski, 1984, p. 18). The articulation of (cultural) logics are formulated iteratively as researchers interrogate whether or not their models resonate with other analyses or understandings of the research topics. At every stage, decisions must be made about how to interpret what is observed with an eye toward constructing a model of others' internally coherent worldview.

While drawing parallels between these two practices may in itself be intellectually intriguing, the reason to pull on this metaphorical thread is because those who do ethnography have spent the better part of a century grappling with methodological challenges. After a long history of blindspots, paternalistic agendas, and colonialist orientations – and a contentious attempt to turn ethnography into a science (Lende, 2010) – ethnographers began focusing on their own role in the production of knowledge. Through this process, those who practice ethnography, as a method as well as a mode of knowledge production, have developed rich frameworks for reflexivity, fully aware that any model of social behavior is inseparable from the social context and research methods from which it was produced. That is, ethnographers must always account for how their research practices might influence or distort the knowledge that results from their work.

This fundamental methodological orientation is more than an acknowledgment of the "observer effect" referring to the fact that the observation of a system necessarily changes that system; it is an acknowledgement that what can be observed and claimed to be known is "always a view from somewhere." Donna Haraway (1988) used this phrase in the context of critiquing claims to the idealized rational objectivity proposed by and continued on since the European Enlightenment, which seems to offer a "view from nowhere," to argue that what is known is always particular, partial, and incomplete. Knowledge claims are always already embodied and socio-historically situated. In ethnographic practices, this conceptualization of the limits of knowledge production manifests in different ways, depending on the disciplinary or institutional context, but consistently involves a dimension of methodological reflexivity, ranging from research agendas and areas of focus (Asad, 1973; Faubion & Marcus, 2009; Hymes, 1974) to the cultural and geographical delineations of those areas (Gupta & Ferguson, 1997), to the very modes of representation and engagement at stake in ethnographic research (Cefkin, 2010; Clifford & Marcus, 1986; Taussig, 2011). What results, ideally, is not a form of glorified navel gazing, but rather a richer understanding of the reality being observed because the ethnographer has attempted to understand her place in that reality and the nature of the tools at her disposal.

The discipline of anthropology, and the many disciplines that count ethnography as a core method, have not found consensus about the best, or even right, way to do ethnography or be reflexive. Consequently, our provocation here is not about the adoption of any methods in particular, but rather about the kinds of methodological orientations that might guide the future of machine learning, AI, and data science. It is about developing and embracing a practice of the unresolvedness at stake in producing models about the world. In their own terms and particular contexts, technical practitioners need to be exploring and developing what it means to be reflexive in the methods of data science and statistics. In other words, they need to ask themselves what would it mean, and what are all the ways it could mean, to develop an algorithmic or AI system reflexively, and to communicate the truth claims at stake as limited and partial. Ethnography does not offer the answers so much as it offers a historical example of a field navigating these questions to grapple with an iterative and interpretive way of knowing.

One approach, as Matthew Jones writes in the context of data science methods, would be to contend with "a valorization of the muddling through, the wrangling, the scraping, the munging of poorly organized, incomplete, likely inconclusive data" (Jones, 2014, p. 358). Acknowledging the limits of Big Data and AI should not result in their dismissal, but rather enable a more grounded and ultimately more useful set of conversations about the appropriate and effective design of such technologies.

## Conclusion

Vast and consequential resources are being mobilized around Big Data, and now AI. The resultant technologies are frequently invoked as the solution to otherwise intractable social, political, and economic problems, and seem to promise efficiency, neutrality, and fairness – ideals that are often viewed as impossible to achieve through individual human or organizational decision-making processes. However, as we have argued, the fantasies and promises of Big Data and AI often obscure the limitations of the field and trade-offs involved in doing technical work under the rubric of AI. These new phenomena must be taken seriously and interrogated not only as modes of adjudicating in the world, but also and in their very essence, modes of *knowing* about the world.

When proponents of Big Data, machine learning, and AI rely on mobilizing imaginaries of AI as working like magic and glossing over the limitations of technological systems, they run the risk of undermining the power and potential of the systems they are building. In the long run, the biggest challenge for a hype-driven ecosystem where countless public and private sector actors feel the need to implement AI systems is the plethora of poorly constructed models produced through methodologically unsound practices. As long as those models are overconfidently produced and viewed as infallible, there is limited space for interrogating how cultural logics get baked into the very practice of machine learning. For the technological advancements to endure, it is imperative to ground both the practice and rhetoric of AI and Big Data. Doing so requires developing the methodological frameworks to reflexively account for the strengths and weaknesses of both the technical practices and the claims that can be produced through machine learning-based systems.



## Notes

1. For an expanded list of definitions see Press (2014).
2. In the technology sector, the term "vaporware" refers to publicized hardware or software that does not exist or does not do what is promised. Because developers in the tech industry create software based on platform specifications, false advertising is seen as harmful to the ecosystem. Yet, in a more practical sense, vaporware is also duplicitous in the same sense as snake oil.
3. Computer scientists Russell and Norvig (1995) have argued that the history of AI, far from centering on any particular definition of "intelligence," can be seen as orienting around four interrelated but distinct goals: "systems that think like humans, systems that act like humans, systems that think rationally, systems that act rationally."
4. This is not to say that all artificial intelligence research assumed (or assumes) the same form during the periods under discussion, and the field of AI research has always been characterized by multi-disciplinary and divergent approaches. See (Olazaran, 1996) for an analysis of one of the most well-known controversies over methods and techniques in AI.
5. The idea that otherwise inanimate objects might become self-animating can be found in ancient Greek and Chinese texts (Mazlish, 1995), and throughout history, especially during periods of social upheaval or rapid technological change, new mythologies emerged and circulated about the paradise or peril that new technologies would bring. Contemporary American imaginaries about AI are a palimpsest of previous mythologies as well as a particular formulation rooted in the rapidly developing technological cultures of the mid-twentieth century. Contemporary imaginaries of AI and robotics necessarily vary between cultures, and our focus on this paper is particular to Anglo-European cultures and histories. For recent work on the animating imaginaries of robots and other embodied form of "artificial intelligence" in Japan and Korea, see Jeon (2016) and Robertson (2017).
6. It is also significant, and typical, that Amy is gendered female. While X.ai offers the option to have "Andrew" as an assistant, feminized artificial and robotic agents abound. It is beyond the scope of this paper to analyze these dynamics in detail. For a critical analysis of the

gendered aspects of artificial intelligence, see Adam (1998) and Suchman (2011); see also Robertson (2010) for an analysis of gendered-robots in Japan.

7. While news media pronounced that a new frontier had been crossed, it should be noted that the formal article published in *Nature* (Silver et al., 2016) was reasonably contained in its extrapolations.
8. In the United States, the history of the Census (Anderson, 2015) offers a parallel insight into the challenges and perils of articulating categories.
9. The "uncanny valley" refers to the hypothesis, first originating in robotics design, that there is a negative emotional response to a machine that is very close to being human-like, but not an exact representation. This not-quite-perfect likeness creates a sense of the uncanny, and creates a valley, a low point, in emotional responses which are otherwise positive for machines either are perceptibly poor representations of humans or perfectly accurate. See Mori (2012).
10. The intertwined topics of validity, reproducibility, and replicability are frequently emerging as topics of workshops and discussions at top machine learning conferences. See, for example, https://sites.google.com/view/icml-reproducibility-workshop/home.



## Acknowledgements

We are grateful to Robyn Caplan, Andrew Selbst, and Caroline Jack for their insightful comments on an early draft, and to Kinjal Dave for research assistance. In addition, the attendees of the Data & Society Workshop: Eclectic (2016) provided valuable feedback on thinking through the histories and epistemologies at stake in AI.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

*M. C. Elish* http://orcid.org/0000-0002-9647-1178

## References

Ackerman, T. (2017, January 17). Touted IBM supercomputer project at MD Anderson on Hold after audit finds spending issues. *Houston Chronicle*. Retrieved from http://www.houstonchronicle.com/news/houston-texas/houston/article/Touted-IBM-supercomputer-project-at-MD-Anderson-10941783.php

Adam, A. (1998). *Artificial Knowing: Gender and the Thinking Machine*. New York, NY: Routledge.

Agre, P. (1995). Constructions of the mind: Artificial intelligence and the humanities. In G. Guzeldere & S. Franchi (Eds.), *Mechanical bodies, computational minds: Artificial intelligence from automata to cyborgs* (pp. 153–174). Cambridge, MA: MIT Press.

Ambrose, M. L. (2015). Lessons from the avalanche of numbers: Big data in historical perspective. *I/S: A Journal of Law and Policy for the Information Society*, *11*(2), 201–277. Retrieved from http://hdl.handle.net/1811/75477

Ananny, M. (2011, April 14). The curious connection between apps for gay men and sex offenders. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340/

Anderson, B. (2017, February 12). The rise of the weaponized AI propaganda machine. [Blog post]. Retrieved from https://medium.com/join-scout/the-rise-of-the-weaponized-ai-propaganda-machine-86dac61668b

Anderson, M. J. (2015). *The American census: A social history* (2nd ed.). New Haven, CT: Yale University Press.
Angelica, A. D. (2011, January 31). How Watson works: A conversation with Eric Brown, IBM Research Manager [Blog post]. Retrieved from http://www.kurzweilai.net/how-watson-works-a-conversation-with-eric-brown-ibm-research-manager
Asad, T. (Ed.). (1973). *Anthropology & the colonial encounter*. Ithaca, NY: Ithaca Press.
Barocas, S., Rosenblat, A., boyd, d., & Gangadharan, S. P. (2014, October). Data & civil rights: Technology primer. Data & Civil Rights Conference. doi:10.2139/ssrn.2536579
Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*, 671–732. doi:10.15779/Z38BG31
Bowker, G., & Star, L. (2000). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.
boyd, d. (2016, February 5). It's not cyberspace anymore [Blog Post]. *Data & Society Points*. Retrieved from https://points.datasociety.net/it-s-not-cyberspace-anymore-55c659025e97
boyd, d., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679. doi:10.1080/1369118X.2012.678878
Brooks, R. (1991). Intelligence without Representation. In J. Haugeland (Ed.), *Mind Design II, 1997* (pp. 395–421). Cambridge, MA: MIT Press.
Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data& Society*, *3*(1), 1–12. doi:10.1177/2053951715622512
Cadwalladr, C. (2017, May 7). The Great British Brexit Robbery: How our democracy was hijacked. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy
Calude, C. S., & Longo, G. (2016). The deluge of spurious correlations in big data. *Foundations of Science*, 1–18. doi:10.1007/s10699-016-9489-4
Cambridge Analytica [Home page]. Retrieved from http://https://cambridgeanalytica.org/
Captain, S. (2017, January 1). Can IBM's Watson do it all? *Fast Company*. Retrieved from https://www.fastcompany.com/3065339/can-ibms-watson-do-it-all
CBS Sunday Morning. (2015, June 14). The future of robots and artificial intelligence [video interview]. Retrieved from http://www.cbsnews.com/videos/the-future-of-robots-and-artificial-intelligence/
Cefkin, M.2010). *Ethnography and the Corporate Encounter*. New York, NY: Berghahn.
Cellan-Jones, R. (2015, June 5). When big data becomes big brother. *BBC*. Retrieved from http://www.bbc.com/news/technology-33019405
Citron, D. (2008). Technological Due Process. *Washington University Law Review*, 85, 1249–1313. Retrieved from http://openscholarship.wustl.edu/cgi/viewcontent.cgi?article = 1166&context=law_lawreview
Clarke, A. C. (1973). *Profiles of the future: An enquiry into the limits of the possible*. New York, NY: Harper & Row.
Clifford, J., & Marcus, G. E. (Eds.). (1986). *Writing culture: The poetics and politics of ethnography*. Berkeley: University of California Press.
Confessore, N., & Hakim, D. (2017, March 6). Data firm says 'Secret Sauce' aided trump; Many Scoff. *The New York Times*. Retrieved from https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html?_r=0
Crawford, K., & Schultz, J. (2013). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review*, *55*, 93–128. Retrieved from http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4/
Crawford, K., Whittaker, M., Elish, M. C., Barocas, S., Plasek, A., & Ferryman, K. (2016). The AI now report: The social and economic implications of artificial intelligence technologies in the near-term. Report prepared for the AI Now public symposium, hosted by the White House and New York University's Information Law Institute. Retrieved from https://artificialintelligencenow.com/media/documents/AINowSummaryReport_3.pdf
Dreyfus, H. (1972). *What computers still can't do: A critique of artificial reason*. Cambridge, MA: MIT Press.

Edwards, P. (1996). *The closed world: Computers and the politics of discourse in Cold War America*. Cambridge, MA: MIT Press.

Ekbia, H., & Nardi, B. (2014). Heteromation and its (dis)contents: The invisible division of labor between humans and machines. *First Monday*, *19*(6), doi:10.5210/fm.v19i6.5331

Elish, M. C., & Hwang, T. (2016). *An AI pattern language*. New York, NY: Data & Society Research Institute.

Ensmenger, N. (2012). Is chess the drosophila of artificial intelligence? A social history of an algorithm. *Social Studies of Science*, *42*(1), 5–30. doi:10.1177/0306312711424596

Faubion, J., & Marcus, G. E. (Eds.). (2009). *Fieldwork is not what it used to be: Learning anthropology's method in a time of transition*. Ithaca, NY: Cornell University Press.

Forsythe, D. E. (1993). Engineering knowledge: The construction of knowledge in artificial intelligence. *Social Studies of Science*, *23*(3), 445–477. doi:10.1177/0306312793023003002

Forsythe, D. E. (2002). *Studying those who study us: An anthropologist in the world of artificial intelligence*. Stanford, CA: Stanford University Press.

Foucault, M. (2009). *Security, territory, population: Lectures at the Collège de France 1977–1978 (Lectures at the College de France)*. (G. Burchell, Trans.). New York, NY: Picador.

Foucault, M. (2010). *The Birth of Biopolitics: Lectures at the Collège de France, 1978–1979 (Lectures at the College de France)*. (G. Burchell, trans.). New York, NY: Picador.

Franchi, S., & Guzeldere, G. (Eds.). (2005). *Mechanical bodies, computational minds: Artificial intelligence from automata to cyborgs*. Cambridge, MA: MIT Press.

Funk, M. (2016, November 19). The secret agenda of a Facebook quiz [opinion editorial]. *New York Times*. Retrieved from https://www.nytimes.com/2016/11/20/opinion/the-secret-agenda-of-a-facebook-quiz.html?_r=0

Gell, A. (1988). Anthropology and magic. *Anthropology Today*, *4*(2), 6–9. doi:10.2307/3033230

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). Cambridge, MA: MIT Press.

Gleick, J. (2012). *The information: A history, a theory, a flood*. New York, NY: Pantheon.

Grassegger, H., & Krogerus, M. (2017). The data that turned the world upside down. *Vice Motherboard*. Retrieved from https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

Gupta, A., & Ferguson, J. (1997). *Anthropological locations: Boundaries and grounds of a field science*. Berkeley: University of California Press.

Hacking, I. (1982). Biopower and the avalanche of printed numbers. *Humanities in Society*, *5*, 279–295.

Hammerbacher, J. (2016). Data science from Wall Street to startups to academic biomedicine [lecture]. Retrieved from https://datasociety.net/events/data-no-93-jeff-hammerbacher/

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575–599. doi:10.2307/3178066

Hardy, Q. (2016, March 27). Silicon Valley looks to artificial intelligence for the next big thing. *The New York Times*. Retrieved from https://www.nytimes.com/2016/03/28/technology/silicon-valley-looks-to-artificial-intelligence-for-the-next-big-thing.html

Haugeland, J. (1985). *Artificial intelligence: The very idea*. Cambridge, MA: MIT Press.

Hymes, D. (Ed.). 1974). *Reinventing anthropology*. New York, NY: Vintage.

IBM News Room. (2013, October 18). *MD Anderson Taps IBM Watson to Power* "Moon Shots" *Mission Aimed at Ending Cancer, Starting with Leukemia* [Press Release]. Retrieved from https://www-03.ibm.com/press/us/en/pressrelease/42214.wss

Igo, S. (2007). *The averaged American: Surveys, citizens, and the making of a mass public*. Cambridge, MA: Harvard University Press.

Irani, L. (2015, January 15). Justice for 'Data Janitors'. *Public Books*. Retrieved from http://www.publicbooks.org/justice-for-data-janitors/

Jaklevic, C. (2017, February 23). MD Anderson Cancer Center's IBM Watson project fails, and so did the journalism related to it. *Health News Review*. Retrieved from https://www.

healthnewsreview.org/2017/02/md-anderson-cancer-centers-ibm-watson-project-fails-journalism-related/
Jeon, C. (2016). A good life in Asia with Robots? In Digital Asia Hub. *The Good Life in Asia's Digital 21st Century*. Hong Kong: Digital Asia Hub.
Jones, M. (2014). Data and hubris. In R. Schutt & C. O'Neill (Eds.), *Doing data science: Straight talk from the frontline* (pp. 357–360). Sebastopol, CA: O'Reilly Media.
Jones, M. (2016). Calculating devices and computers. In B. Lightman (Ed.), *Companion to the history of science* (pp. 472–487). Malden, MA: Wiley.
Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), doi:10.1177/2053951714528481
Koepke, L. (2016, November 21). Predictive policing isn't about the future. *Slate.com*. Retrieved from http://www.slate.com/articles/technology/future_tense/2016/11/predictive_policing_is_too_dependent_on_historical_data.html
Laney, D. (2001, February 6). 3D data management: Controlling data volume, velocity, and variety [Blog Post]. *Gartner*. Retrieved from https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf
Lende, D. (2010, December 11). Anthropology after the "Science" controversy: We're moving ahead. *PLOS Blogs: Neuroanthropology*. Retrieved from http://blogs.plos.org/neuroanthropology/2010/12/11/anthropology-after-the-science-controversy-were-moving-ahead/
Leonelli, S. (2014). What difference does quantity make? On the epistemology of big data in biology. *Big Data & Society*, *1*(1), doi:10.1177/2053951714534395
Levy, S. (2016, June 22). How Google is remaking itself as a "machine learning first" company. *Wired*. Retrieved from https://www.wired.com/2016/06/how-google-is-remaking-itself-as-a-machine-learning-first-company/
Lighthill, J. (1973). *Artificial intelligence: A general survey*. Artificial Intelligence: a paper symposium. Science Research Council.
Lohr, S. (2015, April 28). Less Noise but more money in data. *New York Times*. Retrieved from https://bits.blogs.nytimes.com/2015/04/28/less-noise-but-more-money-in-data-science/?_r=0
Malinowski, B. (1984). *Argonauts of the Western Pacific*. Prospect Heights, IL: Waveland Press.
Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. Retrieved from http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation
Mark Summers Casting & Management. (2016). 'Coping with Humans' A Support Group for Bots IBM. *Youtube.com*. Digital video commercial. Retrieved from https://www.youtube.com/watch?v=vuHqhR9VpNc&t=20s
Mayer, J. (2017, March). The reclusive hedge-fund tycoon behind the trump presidency. *The New Yorker*. Retrieved from http://www.newyorker.com/magazine/2017/03/27/the-reclusive-hedge-fund-tycoon-behind-the-trump-presidency
Mazlish, B. (2005). The Man-machine and artificial intelligence. In S. Franchi & G. Guzeldere (Eds.), *Mechanical bodies, computational minds: Artificial intelligence from automata to cyborgs* (pp. 175–202). Cambridge, MA: MIT Press.
McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Retrieved from https://www.aaai.org/ojs/index.php/aimagazine/article/view/1904
McCorduck, P. (2004). *Machines who think*. Natick, MA: A. K. Peters.
Mindell. (1998). *Between human and machine: Feedback, control, and computing before cybernetics*. Baltimore, MD: Johns Hopkins University Press.
Mirowski, P. (2003). McCorduck's *machines who think* after twenty-five years: Revisiting the origins of AI. *AI Magazine*, *24*(4), 135–138. doi:10.1609/aimag.v24i4.1739
Mori, M. (2012). The uncanny valley: The original essay by Masahiro Mori. *IEEE Spectrum*. Retrieved from http://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley

Mozur, P. (2017, May 25). Google's A.I. program rattles Chinese go master as it wins match. *New York Times*. Retrieved from https://www.nytimes.com/2017/05/25/business/google-alphago-defeats-go-ke-jie-again.html

Muñoz, C., Smith, M., & Patil, D. (2016). *Big data: A report on algorithmic systems, opportunity, and civil right*. Report prepared for Executive Office of the President, United States of America. Retrieved from https://www.whitehouse.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Nielsen, M. (2016, April 4). How Google's AlphaGo Imitates Human Intuition. *Atlantic Monthly*. Retrieved from https://www.theatlantic.com/technology/archive/2016/04/how-alphago-imitates-human-intuition/476508/

Olazaran, M. (1996). A sociological study of the official history of the perceptrons controversy. *Social Studies of Science*, 611–659. doi:10.1177/030631296026003005

O'Neil, C. (2016). *Weapons of math destruction: How Big Data increases inequality and threatens democracy*. New York, NY: Crown.

Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, B., & Swami, A. (2017, March 19). Practical black-box attacks against machine learning. Proceedings of the 2017 ACM Asia Conference on Computer and Communications. Retrieved from https://arxiv.org/abs/1602.02697

Podesta, J., Pritzker, P., Moniz, J. E., Holdren, J., & Zientz, J. (2014). *Big data: Seizing opportunities and preserving values*. Report prepared for Executive Office of the President, United States of America. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf

Poon, M. (2016). Corporate capitalism and the growing power of big data. *Science, Technology, & Human Values*, *41*(6), 1088–1108. doi:10.1177/0162243916650491

Press, G. (2014, September 3). 12 big data definitions: What's yours? [Blog Post]. Retrieved from http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/

Ramirez, E., Brill, J., Ohlhausen, M. K., & McSweeny, T. (2016). *Big data: A tool for inclusion or exclusion? Understanding the issues*. Report prepared for Federal Trade Commission, United States of America. Retrieved from https://www.ftc.gov/system/files/documents/reports/bigdatatoolinclusionorexclusionunderstandingissues/160106 bigdatarpt.pdf

Riskin, J. (Ed.). (2007). *Essays in the history and philosophy of artificial life*. Chicago, IL: University of Chicago Press.

Robertson, J. (2010). Gendering humanoid robots: Robo-sexism in Japan. *Body & Society*, *16*(2), 1–36. doi:10.1177/1357034X10364767

Robertson, J. (2017). *Robo sapiens Japanicus*. Berkeley: University of California Press.

Rose, N. (1991). Governing by numbers: Figuring out democracy. *Accounting, Organizations and Society*, *16*(7), 673–692. doi:10.1016/0361-3682(91)90019-B

Rosenblat, A., Wikelius, K., boyd, d., Gangadharan, S. P., & Yu, C. (2014, October). Data & civil rights: Criminal justice primer. Data & Civil Rights Conference. doi:10.2139/ssrn.2542262

Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. New York, NY: Simon & Schuster.

Sauter, M. (2017, May 24). Persuasion and the other thing: A critique of big data methodologies in politics [Blog Post]. *Ethnography Matters*. Retrieved from http://ethnographymatters.net/blog/2017/05/24/persuasion-and-the-other-thing-a-critique-of-big-data-methodologies-in-politics/

Seaver, N. (2015a). Bastard algebra. In T. Boellstorff, & B. Maurer (Eds.), *Data: Now bigger and better!* (pp. 27–46). Chicago: Prickly Paradigm Press.

Seaver, N. (2015b). The nice thing about context is that everyone has it. *Media, Culture, and Society*, *37*(7), 1101–1109. doi:10.1177/0163443715594102

Selbst, A. (2017, March 28). A mild defense of our new machine overlords. *Vanderbilt Law Review*, *70*, 87–104. Retrieved from https://ssrn.com/abstract=2941078

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., … Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 484–492. doi:10.1038/nature16961

Smart, W. D. (2016). "The robot revolution has been rescheduled (until we can debug the sensors): Technical aspects of robotics." WeRobot pre-conference Workshop. Presentation. Miami, FL.

Stanford University. (2016, September). Artificial intelligence and life in 2030. Report prepared for One Hundred Year Study on Artificial Intelligence (2015 Study Panel). Retrieved from https://ai100.stanford.edu

Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions* (2nd ed). New York, NY: Cambridge University Press.

Suchman, L. (2011). Subject objects. *Feminist Theory*, *12*(2), 119–145.

Taggart, K. (2017, February 16). The truth about the trump data team that people are freaking out about. *BuzzFeedNews*. Retrieved from https://www.buzzfeed.com/kendalltaggart/the-truth-about-the-trump-data-team-that-people-are-freaking?utm_term = .cxLAwpN1vm#.ucyYzXv25E

Talbot, D. (2009, May 27). How IBM plans to win jeopardy! *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/413605/how-ibm-plans-to-win-jeopardy/

Talbot, D. (2016, April 15). How political candidates know if you're neurotic. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/601214/how-political-candidates-know-if-youre-neurotic/

Taussig, M. (2011). *I swear I saw this*. Chicago, IL: University of Chicago Press.

Terdoslavich, W. (2015, December 25). IBM Watson: AI for the real world. *Information Week*. Retrieved from http://www.informationweek.com/big-data/software-platforms/ibm-watson-ai-for-the-real-world/d/d-id/1323645

Thompson, C. (2010, June 16). What is I.B.M's Watson? *New York Times Magazine*. Retrieved from http://www.nytimes.com/2010/06/20/magazine/20Computer-t.html

Turow, J. (2011). *The daily you. How the new advertising industry is defining your identity and your worth*. New Haven, CT: Yale University Press.

University of Texas. (2016). Special review of procurement procedures related to the M.D. Anderson Cancer Center Oncology Expert Advisor Project. Austin, TX: University of Texas System Audit Office.

Vigen, T. (2015). *Spurious correlations*. New York, NY: Hachette Books.

Weber, J., & Suchman, L. (2016). Human-machine autonomies. In N. Bhuta, S. Beck, R. Geiss, H. Liu, & C. Kress (Eds.), *Autonomous weapons systems: Law, ethics, policy* (pp. 75–102). Cambridge: Cambridge University Press. doi:10.1017/CBO9781316597873.004

X.ai [Home page]. (n.d.). Retrieved from https://x.ai/.

Zarsky, T. (2016). The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology & Human Values*, *41*(1), 118–132. doi:10.1177/0162243915605575