Automation, Big Data, and Politics:
A Research Review

SAMANTHA SHOREY (1)
Samantha Shorey: sshorey@uw.edu

PHILIP N. HOWARD
Oxford Internet Institute, UK
Philip N. Howard: philip.howard@oii.ox.ac.uk

Date submitted: 2016–08–23

1) This work was supported by the Open Society Foundation. The authors would like to thank Samuel Woolley and Mengjun Guo for their collaboration on a previous version of this manuscript (Howard, Shorey, Woolley, & Guo, 2016). The authors also gratefully acknowledge the feedback from the day-long workshop "Algorithms, Automation and Politics,' organized by the European Research Council–funded Computational Propaganda project of the Oxford Internet Institute and held as a preconference to the International Communication Association Meeting in Fukuoka, Japan, in June 2016. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the European Research Council.
University of Washington, USA

We review the great variety of critical scholarship on algorithms, automation, and big data in areas of contemporary life both to document where there has been robust scholarship and to contribute to existing scholarship by identifying gaps in our research agenda. We identify five domains with opportunities for further scholarship: (a) China, (b) international interference in democratic politics, (c) civic engagement in Latin American, (d) public services, and (e) national security and foreign affairs. We argue that the time is right to match dedication to critical theory of algorithmic communication with a dedication to empirical research through audit studies, network ethnography, and investigation of the political economy of algorithmic production.

Keywords: literature review, critical, big data, algorithms, politics, automation

There is growing concern about the degree to which digital media and device networks can be used as tools of social control. Bots—the focus of this Special Section—are imbedded in larger questions about the place of algorithms, automation, and big data in public life. We often talk about bots in terms of their front-end activities: who they talk to, what they say, and what they do. They are framed more in terms of their actions and outcomes and less in terms of what makes them run.

Bots are computer scripts that act autonomously based on platform data. In this article, we explore how algorithmic control manifests in the creation and subsequent use of big data. We do not offer an exhaustive catalogue of critical big data research. Rather, this research review traces the currents, intersections, and openings for research on computational processes in contemporary political life.

Critical Big Data, Critical Algorithms, and Automation

In this research review, we use the term *big data* to refer to large amounts of information collected about many people using many devices (Howard, Shorey, Woolley, & Guo, 2016). More than size, it characterizes data sets that can be searched, aggregated, and triangulated with other data sets (boyd & Crawford, 2012). While an increasing number of communication scholars embrace big data methods in their research, others working in the discipline have started to think critically about the implications of big data in the academy and beyond. Because communication as a discipline focuses on the exchange of information (Schramm, 1983), big data is a natural object of analysis for communication scholars, as it is generated by interaction with communication information technologies, such as social media, search engines, and the Internet. Big data takes the form of communication artifacts, such as photographs, microtargeting of profiles, social network content, and metadata.

In light of this, communication scholars have embarked on critical big data studies in an effort to demonstrate how flaws—ethical or methodological—in the collection and use of big data may reproduce social inequality (Crawford, Gray, & Miltner, 2014). In particular, Dalton and Thatcher (2014) pose five questions:

☐ "What historical conditions lead to the emergence of big data as a form of knowledge? (Barnes & Wilson, 2014; Dalton, 2013)

☐ Who controls big data, its production, and its analysis? What motives and imperatives drive their work? (Thatcher, 2014)

☐ Who are the subjects of big data, and what knowledges are they producing? (Haklay, 2013)

☐ How is big data actually applied in the production of spaces, places, and landscapes? (Kitchin & Dodge, 2011)

☐ What is to be done with big data, and what other kinds of knowledges could it help produce?" (Shah, 2013)

Concern about the political impact of big data has led social and computer scientists to investigate how algorithmic control can be exercised and abused. In its most straightforward sense, the term *algorithm* can be used to describe any set of steps used to accomplish a task (Gillespie, 2016; Gurevich, 2011). If a computer is performing these steps, then algorithms automate the process. Once built, algorithms run autonomously and perform tasks with little oversight from humans (Zarsky, 2015).

Algorithms can be relatively straightforward. However, the term is often invoked to describe extremely complex computational processes that are difficult for everyday users to understand (Tufekci, 2015). Scholars critically studying algorithms are especially attentive to the subjective decisions made by algorithms: classification, prioritization, association, and filtering (Diakopoulos, 2013a). These decisions are methods of analyzing big data, making it meaningful and useful. They transform information, and they have social consequences (Scannell, 2015).

Domains of Inquiry

The task of studying big data critically can be interpreted in two ways. One way is to examine work that uses big data to engage with—and ideally solve—social problems. But work that mobilizes big data for social good, although important, is not necessarily critical. The excitement of doing something called "big data" has led many universities and think tanks to announce big data initiatives to organize resources, provide a home for big data scientists, and publicly appear to be advancing big data science. Many such initiatives foster the cross-disciplinary collaboration necessary to make big data methodologies available to those working outside of fields of inquiry already driven by statistical methods. However, some of these initiatives fail to engage their own research teams with the questions about the ethics of using personal information, the access and ownership of data sets, and the impact of research outcomes. Few big data projects successfully and fully integrate information ethics in their research efforts.

A second way is to focus on the use of big data itself: its role in either directly harming the research subjects or indirectly harming the public through poor generalizations. Big data studies done in this vein ask about the implications of big data and use a lens of critique to think about the effects of this research on individual autonomy and social equity. Our review is of the existing scholarship in this second domain.

Critiques of big data can also take a few different forms. Gillespie and Seaver's (2015) reading list on critical algorithm studies provides a helpful typology for understanding the different kinds of arguments used to critique methods of big data production. First, big data research can be criticized for removing the complexity and context of social systems. As people are reduced to numbers, we lose sight of the hows and whys of actions in favor of measurable behaviors and outcomes. These critiques are not unique to big data but are also aimed at quantitative work generally. In this way, big data has the same problems as small data. Research of this type can be simply summarized as a critique of big data's accuracy and the validity of inferences made from it.

Second, big data can be criticized because the methods used to create these enormous data sets are still reliant on personal information. Academics, policy workers, lawyers, and journalists regularly point out that businesses and organizations across numerous sectors continue to gather personal data, whether from a credit check or an online search, without individuals' consent. Even data that is anonymized can be linked, with some effort, back to individuals (de Montjoye, Radaelli, Singh, & Pentland, 2015). Critique in this area often takes the form of legal and policy responses to data-gathering practices that infringe on personal autonomy. For example, big data allows for the accumulation of detailed personal profiles, enabling advertisers or political campaigns to microtarget based on information collected through Internet browsing or purchasing habits (Auerbach, 2013). This problem has persisted since astroturfing and political redlining were identified and defined as having a contemporary basis in digital networks (Howard, 2006). Research of this type can be summarized as a critique of surveillance and our right to control our own personally identifiable information.

Third, big data can be criticized because the methods used to analyze these data sets are embedded with values and reflect existing biases (Barocas & Selbst, 2015). The same predictive analytics that harvest data for product recommendations can be used to select job candidates or make predictions about the likelihood that one may commit a crime based on one's social network (boyd, Levy, & Marwick, 2014; Stroud, 2014). These practices are, at their core, exclusionary. Metrics may rely on existing categorizations such as *cultural fit*, which would ultimately make an organization more homophilous. Research of this type offers a critique of algorithmically structured discrimination, focusing on big data's power to systematically favor groups of people. Critical research into the politics of algorithms, automation, and big data often mobilize multiple forms of critique. In conversation with a large network of researchers, we identify 13 domains of inquiry.

Banking and Credit

Modern banking is driven by complex data-driven algorithmic trading, often without close oversight from humans (MacKenzie, 2014). Intermediaries, such as floor brokers, have been replaced with automated matching engines for rapid exchanges—despite the fact that the technology contributes to "flash crashes" in the market (Beunza & Millo, 2014). Outdate regulatory systems do little to mitigate the immense effects of these fluctuations (Snider, 2014).

Sociological research in this area has shown that these numeric systems are, nevertheless, still reflective of human judgment. Banking algorithms are designed to replicate human trading patterns and are informed by economic theories (Lenglet, 2011; MacKenzie, 2006; Muniesa, 2014). They also respond to social data, sometimes impulsively—as was the case in 2013, when a hacked @AP tweet sent the S&P into a $136.5 billion downward spiral (Karppi & Crawford, 2015).

Data collection by credit and insurance brokers presents an early example of data-driven discrimination. Throughout the 1990s, Janet Ford (1988) flagged the practice and potential future harms of basing credit availability on dehumanized data. Neighborhood zip codes serve as shorthand for discriminatory redlining based on race. Today, "digital redlining" could incorporate demographic data gathered from social media sites (Noyes, 2015; Wihbey, 2015). Private companies also use data generated from everyday transactions—bills, subscriptions, prepaid cards—to build extensive user profiles of far more depth than even the credit bureaus (Mui, 2011). Banking and credit systems once relied on communities and segments to determine credit; big data allows them to judge "quantified individuals" (Turow, McGuigan, & Maris, 2015).

Business

Since the mid-2000s, the same logics used in Wall Street trading algorithms have been applied to a range of online businesses (Steiner, 2013). Advertising space is bought in real-time ad auctions, microtargeting individual customers based on browser histories (Auerbach, 2013). Strategies such as cookie tracking have become accepted elements of business online (Leyshon, French, Thrift, Crewe, & Webb, 2005).

Consumers, regularly trade their data for digital services, often without fully understanding the terms-of-service agreements that govern the trade (Singer, 2015; Turow, Hennessy, & Draper, 2015). Marketing firms aggregate this data and categorize users into desirable or less desirable consumer segments (Keller & Neufeld, 2014; Turow, 2011). The same features used to "personalize" product recommendations can also be used to manipulate prices and steer users toward more expensive products (Hannak, Mislove, Soeller, Wilson, & Lazer, 2014). Even banks use social media data to build customer profiles, using significant life events to market their offerings (Crosman, 2015).

Discrimination and Civil Rights

Social science research has documented the rise of data-driven discrimination—wherein social decisions derived from big data analysis lead to unfair treatment of minorities (Upturn, 2014). One of the ways this happens is through automated classification. For example, hiring decisions based on similarity algorithms may reproduce existing disparities in the workforce (Barocas & Selbst, 2014; boyd, Levy, & Marwick, 2014). Classification systems are not necessarily discriminatory, and with intentional design, researchers can build systems that classify both effectively and fairly (Dwork, Hardt, Pitassi, Reingold, & Zemel, 2011).

Discrimination can be introduced or reinforced through algorithms both in their design and in their use (Bozdag, 2013). Filtering algorithms that learn from user input may be replicating larger societal biases. For example, Google search results tend to reflect occupational gender stereotypes—returning images of men in male-associated professions, even if women are an equal or majority share of that workforce (Kay, Matuszek, & Munson, 2015). When question stems regarding race are typed into the Google search bar, they also elicit autocomplete answers that are associated with negative stereotypes (Baker & Potts, 2013). These types of search results reinforce racial and gender stereotypes and perpetuate destructive representations, especially for women of color (Noble, 2012).

Big data classification can also lead to discriminatory targeting. Leading up to the 2008 recession, triangulated data were used to target minorities for subprime loans (Gangadharan, 2014). Users who searched for non-White-associated names were more likely to be targeted for advertisements about arrest records than those who searched for White-associated names—despite the fact that this ad copy was generated regardless of an actual arrest record being present (Sweeney, 2013).

Self-provided data from users can also lead to user discrimination based on race—especially in sharing economy marketplaces like AirBnB (Edelman & Luca, 2014; Edelman, Luca, & Svirsky, 2016). Reviews on Yelp also show bias based on the racial identities of businesses' neighborhoods (Zukin, Lindeman, & Hurson, 2015). Though analyses of big data actually revealed these biases, they also reveal the ways that user biases are incorporated into automated systems. Hart and Case (2014) provide an evocative, interactive example of this in *Parable of the Polygons: A Playable Post on the Shape of Society*—showing how even a small amount of bias can lead to complete segregation of populations. It is easy to imagine how these user preferences are learned by automated classification systems. The cause of systematic biases can be very difficult to determine and remedy, as the algorithms that make associations are completely inaccessible to researchers and to the broader public (Pasquale, 2015).

Democracy, Elections, and National Security

Data-driven polls, social media bots, and campaign protocols are of concern to those focused on the critical study of big data in politics. Political campaigns in Western democracies now operate via data-focused systems for voter outreach and categorization. Beginning in 2008, the Obama campaign received widespread attention for innovative and extensive use of voter data to reach communities and individual voters (Issenberg, 2012; Kreiss, 2016). Effectively mobilizing so much data requires immense financial resources, only available to the most established political candidates. Storage alone cost billions of dollars (Pearce, 2013).

During the same election, Facebook launched the "I'm voting" button—a nudge to promote voting that generated voting behavior data for millions of people (Sifry, 2014). Later, big data research based on 61 million Facebook users indicated that the positive social pressure of the voting button encouraged friends to do the same (Bond et al., 2012). In light of this, scholars have raised concerns about the possibility of other automated technologies manipulating elections. Social bots attack activists and spread propaganda worldwide (Woolley, 2016; Woolley & Howard, 2016). Research on search engines also demonstrates their influence on candidate selection (Epstein, 2015). Certainly the experience of the United States with algorithms, automation and politics is not universal, and research on political conversation in the United Kingdom has demonstrated that electronic petitions and other forms of online engagement through social media platforms have long-term, somewhat positive, consequences for civic engagement (Margetts, John, Hale, & Yasseri, 2015; Vaccari, Chadwick, & O'Loughlin, 2015).

Internationally, Edward Snowden's revelations made it clear that intelligence services in many countries, particularly in the United States and United Kingdom, build and use large data sets in spying missions and among many sectors of domestic and foreign affairs (Lyon, 2014). Those that criticize this practice often do so under the banner of privacy, but it is crucial that researchers better contextualize the role of data in these practices. Big data, and the algorithms that make it meaningful, has played a key role in modern warfare: creating associations, tracking bodies, and producing targets (Amoore, 2009; Howard, 2015).

Computational Journalism and News Production

According to recent research from the Pew Research Center for Journalism and Media, the majority of Americans get news from social media (Gottfried & Shearer, 2016). Social media is a prevalent source of and space for political discussion, representing the possibility of a modern public sphere (Caplan & Reed, 2016). This makes the design of sites and apps that deliver content for users especially important (Ananny & Crawford, 2014; Benthall, 2015). Search engine producers find themselves caught between market factors and the values of fairness and representativeness that motivate journalists (Van Couvering, 2007). News-filtering algorithms serve a gatekeeping function, editing what social media users see (Tufekci, 2015). Search engines serve a similar function (Introna & Nissenbaum, 2000). Personalization through algorithms has the potential to create "filter bubbles" in which algorithms favor information that users find agreeable and eliminate other types of information (Pariser, 2013). Big data scholars often acknowledge that algorithms have immense power when they make unknown and unexpected patterns of social inequality or public opinion apparent. However, equally important is the threat of invisibility as algorithms make content or users disappear from view (Bucher, 2012). Invisibility caused by deliberate exclusion is seen as censorship, but invisibility may also be the product less insidious forms of algorithmic curation (Gillespie, 2012). Both forms have political consequences (Granka, 2010).

Journalists are increasingly informed by audience metrics and granular data on viewers. Although viewers, and now page views, are established performance metrics for media, the success of news stories is also measured in terms of interaction and integration on social media sites (Lichterman, 2016). Clicks and comments provide almost instant audience feedback, leading to new levels of responsiveness (Anderson, 2011). Big data and algorithms have shaped journalistic production, ushering in an era of "computational journalism" (Anderson, 2013; Lewis, 2015).

Education

The protection of student data was identified as a top priority in a report from the Obama administration's Big Data and Privacy working group (White House, 2014). Big data is used in educational settings for algorithmic student placement, testing, aptitude evaluation—for states, regions, districts, and students—and other areas. Critical researchers study outcomes and effects of these data-reliant education systems. Policy makers acknowledge the potential positive futures of big data in school systems. However, in education, big data has often been among the most inaccurate and ineffective data (O'Neil, 2013; Strauss, 2014). Educational technologies, which aim to provide individualized learning for students, also produce individualized data. Though the selling of data is controversial in many contexts, educational data are of specific concern. If integrated into algorithmic systems—similar to the ones built for credit scoring or professional hiring—it could declare and reinforce a child's aptitude for the rest of his or her life (Chideya, 2015).

Health

Research concerned with big data used in health care has grown at pace with the industry's switch from paper to digital records. The massive amount of health-care data in the world leaves pundits concerned with leaks or discriminatory outcomes (American Association for the Advancement of Science, Federal Bureau of Investigation, & United Nations Interregional Crime and Justice Research Institute, 2014). Moreover, scientists and companies now use big data generated from online platforms in attempts to predict disease outbreaks and health-care crises—with mixed results (Butler, 2013; Lazer, Kennedy, King, & Vespignani, 2014). Individuals also produce significant amounts of data through health devices such as Fitbits. These data are primarily generated for personal monitoring or to be shared with a community—part of a larger trend toward self-tracking (Neff & Nafus, 2016; Reigeluth, 2014). Yet, these little data have the potential to be combined into aggregate big data because of unclear terms-of-service agreements and the need for updated privacy policies.

Work and Labor

Algorithms and data increasingly serve the functions that middle management once did. They identify job candidates through personality tests and algorithms based on estimates of work efficiency or labor potential (Peck, 2013; Weber & Dwoskin, 2014). They assign and review tasks for workers. In the case of Uber, for example, an algorithm assigns drivers to passengers partially based on location—and passengers then rate drivers to ensure system quality through data collection from users (Lee, Kusbit, Metsky, & Dabbish, 2015). Ratings systems favor consumers, often having no system of appeal should a worker be given a rating unfairly.

Along with these developments come a host of other ethical quandaries. Geolocation puts workers under constant surveillance, allowing employers to know their whereabouts at all times to maximize productivity (Levy, 2015). This may extend even beyond work hours, as employers use wearable devices to reward healthy lifestyle choices like exercise and sleep (O'Connor, 2015). The activities of workers and consumers generate valuable, uncompensated, and often personally identifiable data to improve algorithmic systems.

Public conversation about automation typically frames it as a threat to employment, but these discussions obscure the tangled relationship between people and automated labor. Human employees often perform both the initial and final steps of a task that is otherwise fully automated, grooming and censoring the enormous amounts of data that flow through social media and e-commerce platforms (Chen, 2014; Ekbia & Nardi, 2014). This labor is distributed internationally, relying on computational systems for organization and governance (Aneesh, 2009).

Urban Life, Smart Cities, and the Internet of Things

Data and computation are imbedded in our everyday environment (Greenfield, 2015). Smart cities are wired with environmental sensors, which amplify already existing techniques for monitoring the activities of citizens (Howard, 2015; O'Reilly, 2013). Critical projects focused on this arena trace how technology is used in cityscapes and address potential power imbalances, discriminatory practices, and other sociocultural outcomes of data-supported cities (Powell, 2014).

Mobile technologies and social media produce an immense amount of location-based data, most obviously though geotags. Combined with the fact that much of these data are produced and analyzed in real time, spatial and temporal data contribute to surveillance (Graham & Wood, 2003). For example, in a since-deleted blog post, Uber wrote about rides they suspected to be one-night stands based on overnight stays at destinations other than home (Tufekci & King, 2014). Though seemingly innocuous, this post highlighted the kinds of information that can be gleaned simply from location-based data.

Spatial and temporal data are also being produced by more devices than ever before. The Internet of things (IoT) refers to the multitude of physical devices, automobiles, climate control systems, and appliances connected to the Internet and thus producing and requiring large swaths of data (Bessis & Dobre, 2014; Greengard, 2015). Critical research studying the IoT looks at the ways information gathered from these device systems are used in ways unexpected by owners or operators (Howard, 2015). Scholars, pundits, and professionals concerned with a globally connected physical world make security and privacy key arenas of focus. Is the IoT a form of media? What will it mean if our physical world is governed by digital-rights-management software and algorithms that evaluate our rights of use for material goods?

Policing and Incarceration

Data analysis techniques that use big information streams are now essential in many states' considerations of sentencing, parole, and other aspects of incarceration. Big data and the prison system is a growing field, with risk-assessment software making computational decisions about lives of incarcerated citizens (Calabresi, 2014). Risk assessment tools incorporate factors such as educational attainment and employment history, which are strong indicators of socioeconomic status (Palacios, 2014). This raises red flags for scholars concerned with punishments based on poverty (Starr, 2013). Additionally, similar survey tools are shown to underestimate the recidivism rates for White inmates (Larson, 2016). Much of the sensitive information about prisoners' backgrounds is stored online—leading to questions about the security and privacy of the data.

There has also been a recent surge of interest, especially among academics and media practitioners, about the ways law enforcement agencies use data-driven analytics to inform decisions related to policing (Brayne, Rosenblat, & boyd, 2015; O'Neil, 2016). It has come to light that the Los Angeles Police Department, the Chicago Police Department, and other agencies in dozens of U.S. cities use conclusions drawn from big data for predictive policing (Stroud, 2014; van Rijmenam, 2015). These departments use computational power to predict crimes and identify potential offenders. However, the exact methods used for calculation remain opaque (Eubanks, 2015). The New York Police Department also uses social media to monitor the activity of citizens, specifically young people of color (Hackman, 2015). These tactics raise many questions about how communication systems, from software-based social media algorithms to hardware such as drones, are being used for discriminatory profiling, surveillance, and police abuse (Choi-Fitzpatrick, 2014).

Robotics and Automation

Robotics complicate concepts of big data because robots can be designed to download and execute actions based on cloud-based data. Access to large swaths of data could prove useful for robots run by self-learning software, but automated use of such data could also lead to unexpected or dangerous behavior of technologies such as drones, driverless cars, or medical robotics (Calo, 2014).

Online, autonomous bots collect data to perform routine functions on platforms such as Wikipedia. They also produce data through their interactions on social media platforms, often designed to look and act like human users (Abokhodair, Yoo, & McDonald, 2015). Bots make up almost half of all online traffic (Incapsula, 2015), and their activities are motivated by and imbedded in data logs across the Web. They infiltrate social networks with relative ease (Boshmaf, Muslukhov, Beznosov, & Ripeanu, 2011). For example, an estimated 24 million Instagram users are actually bots—a number that should raise concerns for any researcher using big data to draw conclusions about the communicative practices of human users (Franceschi-Bicchierai, 2015).

Communication Policy

It is difficult to think of an aspect of public life that has not been affected by the use of algorithms, automation, and big data. Yet, the methods that fuel these computational processes often remain in the hands of private companies, inaccessible to researchers or the broader public. In light of these processes' widespread impact and opacity, there is a need for transparency and regulation of algorithms (Medina, 2015). Current laws, such as the Computer Fraud and Abuse Act, actually hinder Internet researchers' ability to investigate their operations (*Sandvig v. Lynch*, 2016).

Social science researchers have called for algorithmic due process in two primary forms. First, algorithms should be evaluated for their accuracy and fairness, if not for their impact on political discourse (Citron & Pasquale, 2014; Mittelstadt, 2016). Algorithms can be variously audited: through evaluating code, observing real users, or creating fictitious users as part of an experiment (Guilbeault, 2016; Sandvig, Hamilton, Karahalios, & Langbort, 2016). Researchers and journalists attempt to reverse engineer these systems to better understand how they work (Diakopoulos, 2013b). Second, people should be notified, and given an opportunity to contest, the conclusions drawn about them from their data (Crawford & Schultz, 2014). This issue is intimately tied to questions of privacy and data ownership. However, should algorithms be open to investigation, questions still remain about responsibility and accountability (Neyland, 2016; Rosenblat, Kneese, & boyd, 2014).

Privacy, Security, and Surveillance

Personal privacy is among the most pressing concerns for those studying issues of big data. Both business-based data collection and government-based surveillance threaten to erode civil liberties and privacy (McQuillan, 2015). Massive databases of private information are vulnerable to attack and theft, and the amalgamation of other data online can pose widespread risks to security. Much of these data are not only personally identifiable information but also visual, using images as biometric data (Gates, 2011). Modern surveillance systems use algorithmic technologies to identify and classify the people depicted (Introna & Wood, 2002). Scholars are concerned with how these data might be abused—what if they were to fall into the wrong hands? (R. T. Ford, 2000). Researchers exploring security and privacy implications of big data seek to understand and illuminate the ways such data not only challenges these ideals but also change them.

Conclusion

Critical data research is flourishing but needs help turning insights into creative applications. Finding fault in the political economy of data, identifying the research and policy projects with questionable ethics, and demonstrating the inadequacies of social research that is not self-reflexive has proved to be relatively straightforward, though not easy. What domains are notably absent from contemporary inquiry on algorithms, automation, and politics?[2]

2 This gap analysis is based on a series of conversations with leading experts in the field. For sharing their perspectives on these questions we are grateful to Nancy Baym, Laura Brandimarte, Finn Brunton, Alberto Cerda, John Cheney-Lippold, Kate Crawford, Greg Elmer, Jeremy Epstein, Brittany Fiore-Gartland, Tarleton Gillespie, Janet Haven, Lance Hoffman, Paul Jacobs, Nathan Jurgenson, Karen Levy, Lori McGlinchey, Eric Meyer, Anita Nikolich, Cathy O'Neil, Eric Sears, David Robinson, Samuel Sinyangwe, Stefaan Verhulst, Janet Vertesi, Taha Yasseri, and Harlan Yu.

- Raising the level of scholarly interest in understanding algorithms, automation, and politics.
- Improving the sophistication of journalists working with big data or writing about it.
- Raising the literacy of public policy makers on the findings of critical and empirical research.
- Drawing popular attention to the impact of algorithms on public life.

China

Our current understanding of algorithms and social control in China is extremely limited. We know that the vast majority of Chinese citizens use a relatively narrow suite of tools that duplicate the technology services and applications offered in other countries (King, Pan, & Roberts, 2016). Yet we also know that these tools are built by state agencies with censorship and surveillance as a core design value. Nonetheless, we know little of how algorithmic manipulation occurs over systems such as Weibo, Renren, and WeChat. China is important for multiple reasons. First, on the question of algorithms and social control, China's information infrastructure will shape the lives of a billion people. Second, China is the source of algorithmic manipulations—such as social media bots that have an impact on public life in democracies. Third, many of the hardware and software innovations by the Chinese state are being sold to other countries hoping to develop their information infrastructures. This means that the tools for algorithmic control are being exported to other authoritarian regimes that also seek an Internet for social control, while Chinese security services retain ultimate control.

What are the specific structures and functions of algorithmic control and big data manipulation in China? How do citizens—and democracy advocates—respond to or circumvent, and how widespread is, critical knowledge of algorithmic control? What are the mechanisms by which the Chinese government uses big data to influence social media and public opinion beyond its borders?

A growing number of authoritarian regimes are using algorithms to manipulate conversations not only in their own countries but also the public spheres of democracies (Howard, 2015). Strategies include attacking civil society groups in democracies, muddying international debate on sensitive security issues, and interfering with public opinion during elections. Which countries try to exercise soft power through algorithms and big data? How often, and in what ways, do governments meddle in the public sphere of other countries using big data and algorithms? How is political discourse and good governance in democracies and open societies threatened by algorithmic manipulation originating outside their borders?

Civic Engagement in Latin America

In several countries in Latin America, big data and the Internet of things actually represent opportunities for civic engagement. Global attention may be focused on political crises and recalcitrant regimes across Asia, Eastern Europe, and the Middle East, but it is in Latin America that we find relatively stable democracies with political interest in investing in public information infrastructure. Latin American civil society groups also have some fairly specific opportunities to engage with other nations and citizens on the horizon. Chile will be rewriting its constitution in the next two years and has signaled interest in crowdsourcing the constitutional process, in addressing privacy issues at the constitutional level, and in investing in e-voting. Cuba, with relatively high levels of engineering education, is opening and transitioning. Argentina is home to an active community of hacktivists. Brazil has a unique history of technology-enabled participatory budgeting, an exceptionally vibrant social-media-user population, a commitment to open source software, a sophisticated level of public interest in Marco Civil da Internet, and broad values of technology use that differ from those in the United States. If there is a region where making the analysis and findings of critical big data work will be welcomed and translated into policy action, that region is Latin America.

Public Services and Security

A growing number of public services, including police, are being caught up in an uncritical drive for big data analysis. There are many kinds of models for making various levels of government more sophisticated in their use of data, but some models must be better than others. One business model, used by the city of Los Angeles, is to crowdsource data gathering using publicly accessible records. Private companies then sell real-time data back to municipal governments in Los Angeles after processing the data through proprietary algorithms. The city of Chicago collects vast amounts of information, ostensibly through policing operations, but releases some of the data through an open data initiative that helps local entrepreneurs develop hyperlocal apps (O'Neil, 2016). It is not known how much policy oversight or ethical review has been extended to such efforts to bring data into city government. A study of best practices or a recommended process for emerging smart cities, perhaps in conjunction with the national conference of mayors, would help to set a high standard for transparent and ethical big data involved with public housing, policing, and other public services.

How should public agencies engage with private data vendors when exploring new big data projects? What kind of big data training should contemporary policy makers have? When should big data projects and data be developed within public agencies, and when should they be contracted out, and under what terms?

National Security, Domestic and Foreign Affairs

The work of Edward Snowden and Julian Assange has brought to light a profusion of new ways in which data, computation, and advanced technology are used in domestic and foreign intelligence operations. These revelations were centered on the idea that new varieties of surveillance were invading the privacy of citizens. Essentially, security practitioners were accused of building massive databases of information containing all sorts of communication—with little attention to nuance or relevance. Because concerns stemming from these various leaks center on more acute questions of surveillance and privacy, the role of big data, and its continued application in national and international security settings, is often obscured or supplanted by generalized conversation. More robust conversation about the way big data research affects both domestic and foreign policy is certainly needed. Although the use of big data by corporations has received increasing critical attention, more research is needed on how this information is collected and used by governments interested in shaping foreign policy outcomes, achieving national security goals, and interfering in the governance of other countries. How much data collection is too much? What kind of public policy oversight would allow national security agencies to meet reasonable collection goals?

Research on algorithms, automation, and politics has the potential to shape public policy and social norms. The amount of attention to this type of work is growing, but much opportunity exists for new lines of creativity and critique. We argue that these lines of inquiry should become mainstream concerns for both critical theorists and social scientists. For critical theorists, algorithms operate on social life—they encode social structure. For social scientists, algorithms govern a growing number of social processes, and thus most modern research questions can include some big data analysis. Moreover, most contemporary social problems have an algorithmic dimension in that computational processes can either exacerbate or diminish social inequalities, depending on how they are designed and applied.

Mainstreaming Algorithmic Research

The next big step for improving our understanding of the political power of algorithms is to mainstream critical big data research. By *mainstreaming*, we mean:

Although many big data research projects are multidisciplinary, research that collaborates across domains is lacking. Research that incorporates individuals situated in businesses, governments, and the academy will foster a more nuanced understanding of how algorithms are used and the mechanisms that may (or may not) be in place to make sure it is used acceptably. It will allow for researchers to arrive at critiques and solutions that take into account the actual practices and constraints of institutions that use big data—rather than merely to critique from ethical ideals. Some of the most rigorous thinking about big data is being done by critical theorists whose powerful ideas are not being integrated into work done social scientists, much less those who are building systems or writing policy. Critical big data research needs teams of researchers to build conceptual bridges and to identify shared terms so that the work being done across domains can be effective.

References

Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). Dissecting a social botnet: Growth, content and influence in Twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 839–851). New York, NY: ACM. http://doi.org/10.1145/2675133.2675208

American Association for the Advancement of Science, Federal Bureau of Investigation, & United Nations Interregional Crime and Justice Research Institute. (2014). *National and transnational security implications of big data in the life sciences*. AAAS.org. Retrieved from http://www.aaas.org/sites/default/files/AAAS-FBI-UNICRI_Big_Data_Report_111014.pdf

Amoore, L. (2009). Algorithmic war: Everyday geographies of the war on terror. *Antipode*, *41*(1), 49–69. http://doi.org/10.1111/j.1467-8330.2008.00655.x

Ananny, M., & Crawford, K. (2014). A liminal press: Situating news app designers within a field of networked news production. SSRN. Retrieved from http://papers.ssrn.com/abstract=2448736

Anderson, C. W. (2011). Deliberative, agonistic, and algorithmic audiences: Journalism's vision of its public in an age of audience transparency. *International Journal of Communication*, *5*, 529–547.

Anderson, C. W. (2013). Towards a sociology of computational and algorithmic journalism. *New Media & Society*, *15*(7), 1005–1021. http://doi.org/10.1177/1461444812465137

Aneesh, A. (2009). Global labor: Algocratic modes of organization. *Sociological Theory*, *27*(4), 347–370. http://doi.org/10.1111/j.1467-9558.2009.01352.x

Auerbach, D. (2013, February 13). You are what you click: On microtargeting. *The Nation*. Retrieved from http://www.thenation.com/article/you-are-what-you-click-microtargeting/

Baker, P., & Potts, A. (2013). "Why do White people have thin lips?" Google and the perpetuation of stereotypes via auto-complete search forms. *Critical Discourse Studies*, *10*(2), 187–204. http://doi.org/10.1080/17405904.2012.744320

Barnes, T. J., & Wilson, M. W. (2014). Big data, social physics, and spatial analysis: The early years. *Big Data & Society*, *1*(1). http://doi.org/10.1177/2053951714535365

Barocas, S., & Selbst, A. (2014, August 6). Losing out on employment because of data mining. *The New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2014/08/06/is-big-data-spreading-inequality

Barocas, S., & Selbst, A. D. (2015). Big data's disparate impact. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2477899

Benthall, S. (2015). Designing networked publics for communicative action. *Interface*, *1*(1). http://doi.org/10.7710/2373-4914.1003

Bessis, N., & Dobre, C. (2014). *Big data and Internet of things: A roadmap for smart environments*. New York, NY: Springer.

Beunza, D., & Millo, Y. (2014). Blended automation: Integrating algorithms on the floor of the New York Stock Exchange. *Systemic Risk Centre*. Retrieved from http://www.systemicrisk.ac.uk/ publications/discussion-papers/blended-automation-integrating-algorithms-floor-new-york-stock

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61 million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. http://doi.org/10.1038/nature11421

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). The socialbot network: When bots socialize for fame and money. In *Proceedings of the 27th Annual Computer Security Applications Conference* (pp. 93–102). Orlando, FL: ACM. http://doi.org/10.1145/2076732.2076746

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

boyd, d., Levy, K., & Marwick, A. (2014). The networked nature of algorithmic discrimination. In S. P. Gangadharan, V. Eubanks, & S. Barocas (Eds.), *Data and discrimination: Collected essays* (pp. 53–57). Washington, DC: Open Technology Institute at The New America Foundation.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227. http://doi.org/10.1007/s10676-013-9321-6

Brayne, S., Rosenblat, A., & boyd, d. (2015). Predictive policing. Data & Society Research Institute. Retrieved from http://www.datacivilrights.org/pubs/2015-1027/Predictive_Policing.pdf

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, *14*(2), 1164–1180. http://doi.org/10.1177/1461444812440159

Butler, D. (2013). When Google got flu wrong. *Nature*, *494*(7436), 155–156. http://doi.org/10.1038/494155a

Calabresi, M. (2014, July 31). Statistics can predict criminal risk. Can they deliver equal justice? *TIME*. Retrieved from http://time.com/3061893/holder-to-oppose-data-driven-sentencing/

Calo, R. (2014). Robotics and the lessons of cyberlaw. *California Law Review*, *103*(3). Retrieved from http://papers.ssrn.com/abstract=2402972

Caplan, R., & Reed, L. (2016). *Who controls the public sphere in an era of algorithms?* Data & Society Research Institute. Retrieved from http://www.datasociety.net/pubs/ap/CaseStudies_PublicSphere_2016.pdf

Chen, A. (2014, October 23). The laborers who keep dick pics and beheadings out of your Facebook feed. *Wired*. Retrieved from http://www.wired.com/2014/10/content-moderation/

Chideya, F. (2015, June 27). No child left un-mined? Student privacy at risk in the age of big data. *The Intercept*. Retrieved from http://theintercept.com/2015/06/27/child-left-un-mined/

Choi-Fitzpatrick, A. (2014). Drones for good: Technological innovations, social movements, and the state. *Journal of International Affairs*, *68*(1), 19.

Citron, D. K., & Pasquale, F. A. (2014). The scored society: Due process for automated predictions. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2376209

Crawford, K., Gray, M. L., & Miltner, K. (2014). Critiquing big data: Politics, ethics, epistemology. *International Journal of Communication*, *8*, 1663–1672. Retrieved from http://ijoc.org/index.php/ijoc/article/view/2167/1164

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review*, *55*(1), 93–128. Retrieved from http://bclawreview.org/review/55_1/03_crawford_schultz/

Crosman, P. (2015, October 19). Tracking bank customers online: Necessary, stalkerish, or both? *American Banker*. Retrieved from http://www.americanbanker.com/news/bank-technology/tracking-bank-customers-online-necessary-stalkerish-or-both-1077310-1.html

Dalton, C., & Thatcher, J. (2014). What does a critical data studies look like, and why do we care? Seven points for a critical approach to "big data." *Society and Space*. Retrieved from https://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/

Dalton, C. M. (2013). Sovereigns, spooks, and hackers: An early history of Google geo services and map mashups. *Cartographica*, *48*(4), 261–274. http://doi.org/10.31.38/carto.48.4.1621

de Montjoye, Y. A., Radaelli, L., Singh, V. K., & Pentland, A. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, *347*(6221), 536–539. http://doi.org/10.1126/science.1256297

Diakopoulos, N. (2013a). Algorithmic accountability reporting: On the investigation of black boxes. Tow Center for Digital Journalism. Retrieved from http://towcenter.org/algorithmic-accountability-2/

Diakopoulos, N. (2013b, October 3). Rage against the algorithms. *The Atlantic*. Retrieved from http://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2011). Fairness through awareness. arXiv. Retrieved from http://arxiv.org/abs/1104.3913

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of Airbnb.com. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2377353

Edelman, B. G., Luca, M., & Svirsky, D. (2016). Racial discrimination in the sharing economy: Evidence from a field experiment. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2701902

Ekbia, H., & Nardi, B. (2014). Heteromation and its (dis)contents: The invisible division of labor between humans and machines. *First Monday*, *19*(6). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/5331

Epstein, R. (2015, August 19). How Google could rig the 2016 election. *Politico.* Retrieved from http://www.politico.com/magazine/story/2015/08/how-google-could-rig-the-2016-election-121548.html

Eubanks, V. (2015, April 30). The policy machine. *Slate*. Retrieved from http://www.slate.com/articles/technology/future_tense/2015/04/the_dangers_of_letting_algorithms_enforce_policy.html

Ford, J. (1988). *The indebted society: Credit and default in the 1980s*. London, UK: Routledge.

Ford, R. T. (2000). Save the robots: Cyber profiling and your so-called life. *Stanford Law Review*, *52*(5), 1573–1584. http://doi.org/10.2307/1229522

Franceschi-Bicchierai, L. (2015, June 30). Up to 24 million Instagram accounts are spambots, study says. *Motherboard*. Retrieved from http://motherboard.vice.com/read/24-million-instagram-accounts-spambots-study

Gangadharan, S. P. (2014, August 6). The dangers of high-tech profiling, using big data. *The New York Times*. Retrieved from http://www.nytimes.com/roomfordebate/2014/08/06/is-big-data-spreading-inequality

Gates, K. (2011). *Our biometric future: Facial recognition technology and the culture of surveillance*. New York, NY: New York University Press.

Gillespie, T. (2012). Can an algorithm be wrong? *Limn*, *1*(2). Retrieved from http://escholarship.org/uc/item/0jk9k4hj

Gillespie, T. (2016). Algorithm. In B. Peters (Ed.), *Digital keywords: A vocabulary of information society and culture* (pp. 18–30). Princeton, NJ: Princeton University Press.

Gillespie, T., & Seaver, N. (2015, November 9). Critical algorithm studies: A reading list. Microsoft Research Social Media Collective. Retrieved from http://socialmediacollective.org/reading-lists/critical-algorithm-studies/

Gottfried, J., & Shearer, E. (2016, May 26). News use across social media platforms 2016. *Journalism.org*. Retrieved from http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/

Graham, S., & Wood, D. (2003). Digitizing surveillance: Categorization, space, inequality. *Critical Social Policy*, *23*(2), 227–248. http://doi.org/10.1177/0261018303023002006

Granka, L. A. (2010). The politics of search: A decade retrospective. *The Information Society*, *26*(5), 364–374. http://doi.org/10.1080/01972243.2010.511560

Greenfield, A. (2015, June 29). Uber, or: The technics and politics of socially corrosive mobility. *Speedbird.wordpress.com*. Retrieved from https://speedbird.wordpress.com/2015/06/29/uber-or-the-technics-and-politics-of-socially-corrosive-mobility/

Greengard, S. (2015). *The Internet of things*. Cambridge, MA: MIT Press.

Guilbeault, D. (2016). New approaches to bot security: An ecological view of bot agency [this Special Section]. *International Journal of Communication*, *10*.

Gurevich, Y. (2011, June). What is an algorithm? Microsoft Research. Retrieved from http://research.microsoft.com/pubs/155608/209-3.pdf#page=1&zoom=auto,-44,792

Hackman, R. (2015, April 23). Is the online surveillance of Black teenagers the new stop-and-frisk? *The Guardian*. Retrieved from http://www.theguardian.com/us-news/2015/apr/23/online-surveillance-black-teenagers-new-stop-and-frisk

Haklay, M. (2013). Neogeography and the delusion of democratization. *Environment and Planning*, *45*(1), 55–69.

Hannak, A., Mislove, A., Soeller, G., Wilson, C., & Lazer, D. (2014). Measuring price discrimination and steering on e-commerce Web sites. In *The Internet Measurement Conference*. Retrieved from http://www.ccs.neu.edu/home/cbw/pdf/imc151-hannak.pdf

Hart, V., & Case, N. (2014). Parable of the polygons. Retrieved from http://ncase.me/polygons

Howard, P. N. (2006). *New media campaigns and the managed citizen*. New York, NY: Cambridge University Press.

Howard, P. N. (2015). *Pax technica: How the Internet of things may set us free or lock us up*. New Haven, CT: Yale University Press.

Howard, P. N., Shorey, S., Woolley, S. C., & Guo, M. (2016). *Creativity and critique: Gap analysis of support for critical research on big data* (Working Paper No. 2016.2, p. 44). Oxford, UK: Project on Computational Propaganda. Retrieved from http://politicalbots.org/?p=538

Incapsula. (2015). *2015 bot traffic report*. Retrieved from https://www.incapsula.com/blog/bot-traffic-report-2015.html

Introna, L., & Nissenbaum, H. (2000). Shaping the Web: Why the politics of search engines matters. *The Information Society*, *16*(3), 169–185. http://doi.org/10.1080/01972240050133634

Introna, L., & Wood, D. (2002). Picturing algorithmic surveillance: The politics of facial recognition systems. *Surveillance & Society*, *2*(2/3). Retrieved from http://library.queensu.ca/ojs/index.php/surveillance-and-society/article/view/3373

Issenberg, S. (2012, December 19). A more perfect union: How Obama's team used big data to rally voters. Retrieved from https://www.technologyreview.com/s/508836/how-obama-used-big-data-to-rally-voters-part-1/

Karppi, T., & Crawford, K. (2015). Social media, financial algorithms and the hack crash. *Theory, Culture & Society*, *33*(1), 73–92 http://doi.org/10.1177/0263276415583139

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual Conference on Human Factors in Computing Systems* (pp. 3819–3828). Soeul, South Korea. http://doi.org/10.1145/2702123.2702520

Keller, M., & Neufeld, J. (2014, October 30). Terms of service: Understanding our role in the world of Big Data. *Al Jazeera*. Retrieved from http://projects.aljazeera.com/2014/terms-of-service/

King, G., Pan, J., & Roberts, M. E. (2016, July 27). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *Gking.Harvard.edu.* Retrieved from http://gking.harvard.edu/50c

Kitchin, R., & Dodge, M. (2011). *Code/space: Software and everyday life*. Cambridge, MA: MIT Press.

Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. New York, NY: Oxford University Press.

Larson, S. (2016, May 23). The futures of many prison inmates depend on racially biased algorithms. *Dailydot.com*. Retrieved from http://www.dailydot.com/debug/racist-algorithms-predictive-policing-software-bias-propublica-study/

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science*, *343*(6176), 1203–1205. http://doi.org/10.1126/science.1248506

Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1603–1612). New York, NY: ACM. http://doi.org/10.1145/2702123.2702548

Lenglet, M. (2011). Conflicting codes and codings: How algorithmic trading is reshaping financial regulation. *Theory, Culture & Society*, *28*(6), 44–66. http://doi.org/10.1177/0263276411417444

Levy, K. E. C. (2015). The contexts of control: Information, power, and truck-driving work. *The Information Society*, *31*(2), 160–174. http://doi.org/10.1080/01972243.2015.998105

Lewis, S. C. (2015). Journalism in an era of big data. *Digital Journalism*, *3*(3), 321–330. http://doi.org/10.1080/21670811.2014.976399

Leyshon, A., French, S., Thrift, N., Crewe, L., & Webb, P. (2005). Accounting for e-commerce: Abstractions, virtualism and the cultural circuit of capital. *Economy and Society*, *34*(3), 428–450. http://doi.org/10.1080/03085140500112160

Lichterman, J. (2016, March 17). *The Financial Times* is launching a new analytics tool to make metrics more understandable for its newsroom. *Niemanlab.org*. Retrieved from http://www.niemanlab.org/2016/03/the-ft-is-launching-a-new-analytics-tool-to-make-metrics-more-understandable-for-its-newsroom/

Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society*, *1*(2), 1–13. http://doi.org/10.1177/2053951714541861

MacKenzie, D. (2006). *An engine, not a camera: How financial models shape markets*. Cambridge, MA: MIT Press.

MacKenzie, D. (2014, February). A sociology of algorithms: High frequency trading and the shaping of markets. School of Social and Political Science, University of Edinburgh, Edinburgh, Scotland. Retrieved from http://www.sps.ed.ac.uk/__data/assets/pdf_file/0004/156298/Algorithms25.pdf

Margetts, H., John, P., Hale, S., & Yasseri, T. (2015). *Political turbulence: How social media shape collective action*. Princeton, NJ: Princeton University Press.

McQuillan, D. (2015). Algorithmic states of exception. *European Journal of Cultural Studies*, *18*(4/5), 564–576.

Medina, E. (2015). Rethinking algorithmic regulation. *Kybernetes*, *44*(6/7), 1005–1019. http://doi.org/10.1108/K-02-2015-0052

Mittelstadt, B. (2016). Auditing for transparency in content personalization systems [this Special section]. *International Journal of Communication*, *10*.

Mui, Y. Q. (2011, July 16). Little-known firms tracking data used in credit scores. *The Washington Post*. Retrieved from http://www.washingtonpost.com/business/economy/little-known-firms-tracking-data-used-in-credit-scores/2011/05/24/gIQAXHcWII_story.html

Muniesa, F. (2014). *Discovering stock prices*. London, UK: Routledge.

Neff, G., & Nafus, D. (2016). *Self-tracking*. Cambridge, MA: MIT Press.

Neyland, D. (2016). Bearing account-able witness to the ethical algorithmic system. *Science, Technology & Human Values*, *41*(1), 50–76. http://doi.org/10.1177/0162243915598056

Noble, S. U. (2012, March 8). Missed connections: What search engines say about women. *Bitch Magazine*, (54). Retrieved from http://safiyaunoble.com/2012/03/08/bitch-magazine-article/

Noyes, K. (2015, January 15). Will big data help end discrimination—or make it worse? *Fortune*. Retrieved from http://fortune.com/2015/01/15/will-big-data-help-end-discrimination-or-make-it-worse/

O'Connor, S. (2015, June 8). Wearables at work: The new frontier of employee surveillance. *Financial Times*. Retrieved from http://www.ft.com/intl/cms/s/2/d7eee768-0b65-11e5-994d-00144feabdc0.html#axzz3oa0EbvZv

O'Neil, C. (2013, March 31). Value-added model doesn't find bad teachers, causes administrators to cheat. *Mathbabe.org*. Retrieved from http://mathbabe.org/2013/03/31/value-added-model-doesnt-find-bad-teachers-causes-administrators-to-cheat/

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

O'Reilly, T. (2013). Open data and algorithmic regulation. In B. Goldstein & L. Dyson (Eds.), *Beyond transparency: Open data and the future of civic innovation* (pp. 289–300). San Francisco, CA: Code for America Press.

Palacios, L. D. (2014, October 31). The dangers of evidence-based sentencing. *Thegovlab.org*. Retrieved from http://thegovlab.org/the-dangers-of-evidence-based-sentencing/

Pariser, E. (2013). *The filter bubble: How the new personalized Web is changing what we read and how we think*. New York, NY: Penguin Books.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Pearce, R. (2013, May 28). Big data is BS: Obama campaign CTO. *CIO*. Retrieved from http://www.cio.com.au/article/462961/big_data_bs_obama_campaign_cto/

Peck, D. (2013, December). They're watching you at work. *The Atlantic*. Retrieved from http://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/

Powell, A. (2014). Datafication, transparency, and good governance of the data city. In K. O'Hara, C. Nguyen, & P. Haynes (Eds.), *Digital enlightenment yearbook 2014 : Social networks and social machines, surveillance and empowerment* (pp. 215–224). Burke, VA: IOS Press.

Reigeluth, T. B. (2014). Why data is not enough: Digital traces as control of self and self-control. *Surveillance & Society*, *12*(2), 243–254.

Rosenblat, A., Kneese, T., & boyd, d. (2014). *Workshop primer: Algorithmic accountability*. Data & Society Research Institute. Retrieved from http://www.datasociety.net/pubs/2014-0317/AlgorithmicAccountabilityPrimer.pdf

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2016). When the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication*, *10*(Special Section). Retrieved from ijoc.org.

Sandvig v. Lynch. (2016). Retrieved from https://www.aclu.org/legal-document/sandvig-v-lynch-complaint

Scannell, J. (2015). What can an algorithm do? *DIS Magazine*. Retrieved from http://dismagazine.com/discussion/72975/josh-scannell-what-can-an-algorithm-do/

Schramm, W. (1983). The unique perspective of communication: A retrospective view. *Journal of Communication*, *33*(3), 6–17.

Shah, V. (2013, August 28). Map of stop and frisks in New York City show concentration by race and neighborhood. *Untapped Cities.* Retrieved from http://untappedcities.com/2013/08/28/new-map-shows-police-stop-and-frisks-according-to-race-and-neighbourhood-in-new-york-city/

Sifry, M. (2014, October 31). Facebook wants you to vote on Tuesday: Here's how it messed with your feed in 2012. *Mother Jones*. Retrieved from http://www.motherjones.com/politics/2014/10/can-voting-facebook-button-improve-voter-turnout

Singer, N. (2015, June 4). Sharing data, but not happily. *The New York Times*. Retrieved from http://www.nytimes.com/2015/06/05/technology/consumers-conflicted-over-data-mining-policies-report-finds.html

Snider, L. (2014). Interrogating the algorithm: Debt, derivatives and the social reconstruction of stock market trading. *Critical Sociology*, *40*(5), 747–761. http://doi.org/10.1177/0896920513504603

Starr, S. B. (2013). Evidence-based sentencing and the scientific rationalization of discrimination. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2318940

Steiner, C. (2013). *Automate this: How algorithms took over our markets, our jobs, and the world*. New York, NY: Portfolio.

Strauss, V. (2014, November 22). Principal uncovers flawed data in her state's official education reports. *The Washington Post*. Retrieved from https://www.washingtonpost.com/blogs/answer-sheet/wp/2014/11/22/principal-uncovers-flawed-data-in-her-states-official-education-reports/

Stroud, M. (2014, February 19). The minority report: Chicago's new police computer predicts crimes, but is it racist? *The Verge.* Retrieved from http://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist

Sweeney, L. (2013). Discrimination in online ad delivery. Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2208240

Thatcher, J. (2014). Living on fumes: Digital footprints, data fumes, and the limitations of spatial big data. *International Journal of Communication*, *8*, 1766–1783.

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal*, *13*, 203–218.

Tufekci, Z., & King, B. (2014, December 7). We can't trust Uber. *The New York Times*. Retrieved from http://www.nytimes.com/2014/12/08/opinion/we-cant-trust-uber.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.

Turow, J., Hennessy, M., & Draper, N. (2015). *The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation*. University of Pennsylvania. Retrieved from https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf

Turow, J., McGuigan, L., & Maris, E. R. (2015). Making data mining a natural part of life: Physical retailing, customer surveillance and the 21st century social imaginary. *European Journal of Cultural Studies*, *18*(4–5), 464–478. http://doi.org/10.1177/1367549415577390

Upturn. (2014). Civil rights, big data, and our algorithmic future. *Bigdata.fairness.io.* Retrieved from https://bigdata.fairness.io/wp-content/uploads/2015/04/2015-04-20-Civil-Rights-Big-Data-and-Our-Algorithmic-Future-v1.2.pdf

Vaccari, C., Chadwick, A., & O'Loughlin, B. (2015). Dual screening the political: Media events, social media, and citizen engagement. *Journal of Communication*, *65*(6), 1041–1061. http://doi.org/10.1111/jcom.12187

Van Couvering, E. (2007). Is relevance relevant? Market, science, and war: Discourses of search engine quality. *Journal of Computer-Mediated Communication*, *12*(3), 866–887. http://doi.org/10.1111/j.1083-6101.2007.00354.x

van Rijmenam, M. (2015, April 15). The Los Angeles Police Department is predicting and fighting crime with big data. *DataFloq.com*. Retrieved from https://datafloq.com/read/los-angeles-police-department-predicts-fights-crim/279

Weber, L., & Dwoskin, E. (2014, September 30). Are workplace personality tests fair? *The Wall Street Journal*. Retrieved from http://www.wsj.com/articles/are-workplace-personality-tests-fair-1412044257

White House. (2014). Big data: Seizing opportunities, preserving values. *CFR.org.* Retrieved from http://www.cfr.org/technology-and-science/white-house-big-data---seizing-opportunities-preserving-values/p32916

Wihbey, J. (2015, January 14). The possibilities of digital discrimination: Research on e-commerce, algorithms and big data. Retrieved from http://journalistsresource.org/studies/society/internet/possibilities-online-racial-discrimination-research-airbnb

Woolley, S. (2016). Automating power: Social bot interference in global politics. *First Monday*, *21*(4). http://doi.org/10.5210/fm.v21i4.6161

Woolley, S., & Howard, P. (2016, May 15). Bots unite to automate the presidential election. *Wired*. Retrieved from http://www.wired.com/2016/05/twitterbots-2/

Zarsky, T. (2015). 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. http://doi.org/10.1177/0162243915605575

Zukin, S., Lindeman, S., & Hurson, L. (2015). The omnivore's neighborhood? Online restaurant reviews, race, and gentrification. *Journal of Consumer Culture*, 1–21. http://doi.org/10.1177/1469540515611203