Why the Data Revolution Needs Qualitative Thinking

Anissa Tanweer[1] Emily Kalah Gade[2] P.M. Krafft[3] Sarah K. Dreier[4]

[1]Data Science for Social Good Program, eScience Institute, University of Washington, Seattle, Washington, United States of America,
[2]Department of Political Science, Emory University, Atlanta, Georgia, United States of America,
[3]Creative Computing Institute, University of Arts London, London, England, United Kingdom,
[4]Department of Political Science, University of New Mexico, Albuquerque, New Mexico, United States of America

ABSTRACT

This essay draws on qualitative social science to propose a critical intellectual infrastructure for data science of social phenomena. Qualitative sensibilities—interpretivism, abductive reasoning, and reflexivity in particular—could address methodological problems that have emerged in data science and help extend the frontiers of social knowledge. First, an interpretivist lens—which is concerned with the construction of meaning in a given context—can enable the deeper insights that are requisite to understanding high-level behavioral patterns from digital trace data. Without such contextual insights, researchers often misinterpret what they find in large-scale analysis. Second, abductive reasoning—which is the process of using observations to generate a new explanation, grounded in prior assumptions about the world—is common in data science, but its application often is not systematized. Incorporating norms and practices from qualitative traditions for executing, describing, and evaluating the application of abduction would allow for greater transparency and accountability. Finally, data scientists would benefit from increased reflexivity—which is the process of evaluating how researchers' own assumptions, experiences, and relationships influence their research. Studies demonstrate such aspects of a researcher's experience that typically are unmentioned in quantitative traditions can influence research findings. Qualitative researchers have long faced these same concerns, and their training in how to deconstruct and document personal and intellectual starting points could prove instructive for data scientists. We believe these and other qualitative sensibilities have tremendous potential to facilitate the production of data science research that is more meaningful, reliable, and ethical.

Keywords: data science, qualitative methods, data ethics, critical data studies, reproducibility, computational social science

1. Introduction

The data revolution has made its mark on academia (National Science Foundation, 2019). Data science methods are becoming ever more broadly adopted and deeply entrenched at universities, with new data science options being added to rosters at a steady clip (De Veaux et al., 2017; Parry, 2018; Song & Zhu, 2017; Tate, 2017). Our universities are at the forefront of figuring out how to access newly available data sources, harness a new generation of powerful computing resources, and develop novel methods that take advantage of both.

Yet, despite the tremendous opportunities of data science, our daily newsfeeds are reminders that the data revolution has also enabled applications that violate expectations of consent (Meyer, 2014), compromise public discourse (Granville, 2018), perpetuate discredited social theories (Colaner, 2020), sow confusion among decision makers (Davey et al., 2020), and adversely impact minority populations (Evans & Mathews, 2019). Addressing this situation requires heeding Sabina Leonelli's (2021) call "to abandon the myth of neutrality that is attached to a purely technocratic understanding of what data science is as a field—a view that

depicts data science as the blind churning of numbers and code, devoid of commitments or values except for the aspiration toward increasingly automated reasoning."

In this essay, we build on recent work from a wide range of academic communities—including science and technology studies (STS); critical data studies; digital sociology; critical geography; statistics; and fairness, accountability, and transparency (FAccT)—proposing that ideas and approaches from qualitative social sciences and the humanities can help address a number of concerns that commonly arise when data science is applied to the production of social knowledge (Bates, 2018; Bates et al., 2020; Cao, 2017; D'Ignazio & Klein, 2020; Dumit & Nafus, 2018; Iliadis & Russo, 2016; Lindgren, 2020; Marres, 2021; Meng, 2021; Moats & Seaver, 2019; Moss et al., 2019; Neff et al., 2017; Pink & Lanzeni, 2018; Richterich, 2018; Selbst et al., 2019; Sloane & Moss, 2019). Broadly, we argue that quantitative and qualitative approaches should be seen as complementary, mutually reinforcing, and co-constitutive of data science when applied to the production of social knowledge. Certain qualitative sensibilities—specifically, interpretivism, abductive reasoning, and reflexivity—can be combined with quantitative computational approaches to produce more reliable, more thorough, and more ethical research than would be produced without integrating these qualitative approaches. Qualitative traditions can provide a critical intellectual infrastructure for data scientists seeking to advance and extend the frontiers of knowledge generation and address new, complex, and systemic social problems.

While computer scientists have begun thinking critically about the social implications of data science, especially with regard to bias and discrimination (e.g., Basta et al., 2019; Bolukbasi et al., 2016; Caliskan et al., 2017; Garg et al., 2018; Gonen & Goldberg, 2019; Sap et al., 2019), this development has been largely divorced from perspectives on the history, philosophy, and sociology of science. Scholars of science have long recognized that distinct epistemologies underlie different disciplinary and paradigmatic uses of data (e.g., Knorr Cetina, 1999; Leonelli, 2014; Rosenberg, 2015), and critics have argued that the problems with data-intensive computational methods have epistemological roots (Burns, 2015; Burns et al., 2018; Taylor & Purtova, 2019). Therefore, rather than starting with particular *techniques* that are typically associated with qualitative social science, we instead focus on a broader set of *concepts* that are intrinsically informed by particular epistemological and ontological positions common in qualitative social sciences—positions that seek to understand the contingently and subjectively constructed nature of the social world. We refer to these concepts as 'sensibilities' because we intend them to intervene on methodology in a *sensitizing* rather than prescriptive way. In other words, while the three sensibilities we discuss—summarized in Table 1—may lend themselves to certain kinds of methodological practices, they are also flexible enough to be coupled with multiple modes of data collection and analysis. In suggesting practical methodological changes for better incorporating interpretivism, abductive reasoning, and reflexivity into data science, we join ongoing calls for data scientists to learn new skills and collaborate with social scientists and humanists in order to mitigate the harms of data-intensive computational methods (Moats & Seaver, 2019; Neff et al., 2017; Pink & Lanzeni, 2018; Resnyansky, 2019; Selbst et al., 2019). Our central contribution is to undertake translational work, laying out a path for moving from critical data *studies* to critical data *science* (Agre, 1997; Mayer & Malik, 2019).

Table 1. Summary of qualitative sensibilities.

Sensibility	Working definition	Example of related methods
Interpretivism	An epistemological approach probing the multiple and contingent ways that meaning is ascribed to objects, actions, and situations.	Trace ethnography (Geiger & Ribes, 2011; Geiger & Halfaker, 2017)
Abductive reasoning	A mode of inference that updates and builds upon preexisting assumptions based on new observations in order to generate a novel explanation for a phenomenon.	Iterations of open coding, theoretical coding, and selective coding (Thornberg & Charmaz, 2013)
Reflexivity	A process by which researchers systematically reflect upon their own positions relative to their object, context, and method of inquiry.	Brain dumps, situational mapping, and toolkit critiques (Markham, 2017)

2. Interpretivism

The staunchest proponents of data science present it as a revolutionary new paradigm (Hey et al., 2009) that, when applied to social questions, will reveal human behavior to be highly predictable and subject to the laws of 'social physics' (Pentland, 2014). "Who knows why people do what they do?" Chris Anderson famously asked in 2008. "The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves" (Anderson, 2008). One purported advantage of digital trace data is that, instead of being able to see only what people do or say when they know they are being observed, traces of digital interactions tell us what people *really do* in their day-to-day lives, such as where they go, what they buy, and who they talk to (Lazer et al., 2009; Mayer-Schönberger & Cukier, 2013; van Atteveldt & Peng, 2018).

Yet much work has demonstrated that findings based on digital traces are not easily generalized. This is not only because of demographic skew and selection bias (Blank & Lutz, 2017; Hargittai, 2020; Lewis & Molyneux, 2018; Mellon & Prosser, 2017) but also because digital traces are so intimately entangled with their contexts of production that it is difficult for researchers to understand what exactly the data represent and to extrapolate their meaning onto the broader social world (boyd & Crawford, 2012; Crawford, 2013; Hargittai, 2015; Hill & Shaw, 2020; Jungherr, 2019; Jungherr et al., 2017; Marres, 2021; Selbst et al., 2019; Zook et al., 2017)—a phenomenon that Offenhuber (2018) refers to as the "stickiness" of digital traces. For example, an analysis of Facebook data purported to show that "weak" social ties did not help people find jobs (Burke & Kraut, 2013), a finding that went against the grain of conventional wisdom and prior research. Eszter Hargittai

(2015) later critiqued the study for not being circumspect enough in the interpretation of results by considering, for example, that perhaps Facebook is simply not the preferred vehicle for mobilizing superficial social ties.

Digital traces enticingly appear to offer unprecedented, uncensored, unadulterated glimpses of social reality and, as such, their meanings are too often taken to be self-evident. Dean Freelon (2014) conducted a review of highly cited literature from the fields of communication and social computing that analyzed digital traces of online behavior such as hyperlinks, retweets, and follows on Twitter. Freelon (2014) found that researchers commonly took these traces to represent complex social constructs such as influence, trust, and credibility while rarely supplying empirical evidence or justification for those imputations. Conducting research based on such assumptions without making careful, empirically supported, rigorous linkages between conceptualization (e.g., the concept of influence) and operationalization (e.g., a retweet as evidence for influence) leads to limited, if not impoverished, understandings of the social world (Jungherr, 2019). Different approaches are sorely needed in a field where digital records are too often seen as an exact and objective representation of social reality (Resnyansky, 2019).

Given the well-established problems with decontextualization in data-intensive computational methods, we argue that an interpretivist lens could address many such shortcomings and greatly enrich analyses of digital traces and other data science–based research of social phenomena. Rather than seeking truths that are universal and determinate (an epistemic goal to which analyses of digital trace data often gravitate), interpretivist social scholars probe the multiple and contingent ways that meaning is ascribed to objects, actions, and situations. For interpretivists, the ultimate question is not, 'Can we predict behavior Y given condition X?' or 'What factor X causes outcome Y?' but 'What does context X and behavior Y mean or represent to the actors involved?'

To illustrate the difference an interpretivist approach can make, consider a comparison between two different studies analyzing the use of bots in Wikipedia. The first study, titled "Even Good Bots Fight," aimed to measure the extent of conflict between Wikipedia's bots, or computer programs that automatically carry out specific tasks (Tsvetkova et al., 2017). The authors measured conflict by tracking 'reverts,' situations in which one bot undoes the action of another bot. According to this operationalization, the authors found a large extent of conflict between Wikipedia bots. The study concludes that "a system of simple bots may produce complex dynamics and unintended consequences," which has "important implications for Artificial Intelligence research" (Tsvetkova et al., 2017).

In response to Tsetkova et al.'s article, Stuart Geiger and Aaron Halfaker (2017) similarly looked at reverted bot actions on Wikipedia, but drew on an approach called "trace ethnography" (Geiger & Ribes, 2011) to develop a more nuanced characterization of that phenomenon. As Geiger and Halfaker (2017) describe it, trace ethnography "is based on a researcher learning how to follow and interpret transaction log data as part of the lived and learned experience of a community." This means Geiger and Halfaker (2017) did not assume that all reverts constituted conflict. Instead, they drew on their firsthand knowledge of the Wikipedia community and closely examined the trajectory of particular revert cases in order to understand what kinds of work the bots

were doing in those instances and what their developers had intended the bots to do. As a result, the study identified instances of reverts where bots were not in conflict with each other at all, but were appropriately and uncontroversially executing tasks that were assigned to them in the context of ongoing changes within Wikipedia.

Take, for example, a situation in which one bot adds an 'orphan' flag to an entry, indicating that the article does not contain any links to other Wikipedia pages; when a link is eventually added, another bot comes along and reverts the original orphan flag because it is no longer relevant. Geiger and Halfaker (2017) used close examinations of such cases to develop various categories of revert activities and determine which situations constituted conflict and which did not. They found that the overwhelming majority of reverts reflected bots not acting in conflict, but rather updating the content of Wikipedia to reflect new formatting conventions, undoing changes that were intended to be temporary in the first place or completing other noncontroversial tasks. Ultimately, they found that only about 1% of all revert actions could be construed as conflict, and they described how human bot developers typically resolve that small fraction of conflicting interactions.

In short, both Wikipedia bot studies (Geiger & Halfaker, 2017; Tsvetkova et al., 2017) used computational and statistical methods to examine the same phenomenon. And both ultimately found at least some evidence of conflict between bots in Wikipedia. However, while one study (Tsvetkova et al., 2017) makes a coarse assumption about what certain transaction logs mean (revert = conflict), the other (Geiger & Halfaker, 2017) qualitatively explores those transaction logs in their broader context to develop a more fine-grained characterization of what they represent (revert = many different things). The contrast between the resulting takeaways in these two studies is stark. The former projects ominous implications for artificial intelligence run amok. The latter frames bots as constructive tools for extending human agency that can be properly managed with appropriate human supervision.

What this comparison shows us is that interpretivist qualitative approaches can be meaningfully incorporated into data science. Promisingly, a number of scholars are increasingly exploring avenues for doing just that: Noortjes Marres (2021) has demonstrated how an interpretivist approach known as "situational analysis" (Clarke, 2003) can be applied to data generated from computational technologies; Laura Nelson (2020) combines interpretive "deep reading" with computational pattern recognition in textual data; Simon Lindgren (2020) explains how the methodological commitments of Actor Network Theory can be coupled with computational approaches to produce interpretive analyses. The research community needs more development of such methodological hybrids that explore how interpretivist approaches can be wed to computational analyses at scale. Only then can we finally dispel the problematic assumption that qualitative interpretation is unnecessary for the quantitative production of knowledge—what boyd and Crawford (2012) have characterized as the "mistaken belief that qualitative researchers are in the business of interpreting stories and quantitative researchers are in the business of producing facts."

3. Abductive Reasoning

Those who prefer deductive approaches to generating knowledge sometimes critique data science for embracing inductive reasoning (Marcus & Davis, 2014) through approaches like data mining and unsupervised machine learning. After all, inductively searching for patterns in data without being driven by a theory-informed question can easily lead to spurious correlations (Mayo, 2020). Tyler Vigen (2015) memorably demonstrated this by showing, for example, that the consumption of mozzarella cheese corresponds to the number of civil engineering doctorates awarded in a given year.

In reality, though, much of data science inquiry actually relies less on induction and more so upon abduction (Goldberg, 2015; Miller, 2010; Thatcher, 2014; Wagner-Pacifici et al., 2015). Whereas deduction tests what must logically occur in order to substantiate a predefined theory, and induction proposes a de novo theory based solely on a preponderance observed evidence, *abduction* is often described as "inference to the best explanation" (Douven, 2011). Many people are passingly familiar with how abduction works from the beloved stories of Sherlock Holmes (Carson, 2009). Abductive reasoning updates and builds upon preexisting assumptions (in other words, theories) based on new observations in order to generate a novel explanation for a phenomenon. As such, it demarks "a creative outcome which engenders a new idea," (Reichertz, 2010). As Charles Sanders Peirce (2013) has put it:

> But suddenly, while we are poring over our digest of the facts and are endeavoring to set them into order, it occurs to us that if we were to assume something to be true that we do not know to be true, these facts would arrange themselves luminously. That is abduction.

Our point is not that data scientists should start relying more heavily on abduction, as this mode of reasoning is already quite prevalent in data-intensive computational analysis (Goldberg, 2015; Miller, 2010; Thatcher, 2014; Wagner-Pacifici et al., 2015). Rather, we wish to point out that the field lacks widely accepted norms and processes for acknowledging, executing, describing, and evaluating the application of abduction. This is why media scholar Warren Sack (2019) recently argued that large-scale algorithmic data analysis necessitates the development of new rhetorical practices for abductively demonstrating the linkages between opaque computational outputs and the meanings assigned to those results.

Qualitative traditions can help in this regard. When using abductive reasoning, qualitative researchers have developed ways of addressing the relationships between prior assumptions, new observations, and newly derived explanations—something that is often sorely needed in data science. This distinction can be illustrated by contrasting the practices of 'labeling' in data science versus 'coding' in qualitative approaches. Take, for example, the common data science approach of supervised machine learning. Human arbiters 'label' a sample of data that will be used to 'train' a machine learning algorithm in applying those labels beyond the sample—a process that often remains opaque (Geiger et al., 2020). The application of a label implies the mere categorization of indisputable facts. Indeed, sometimes the process of labeling data involves tacit knowledge

requiring little explanation, such as tagging photos of fruit in a bowl as 'apple,' 'banana,' 'peach,' or 'kiwi.' But classifications are laden with social, political, and moral consequences, serving to amplify certain perspectives while silencing others (Bowker & Star, 2000; Gitelman, 2013), and in many cases, 'labeling' in supervised machine learning is informed by underlying assumptions that serve to advance a particular theoretical perspective, whether that theoretical framework is acknowledged or not.

Labeling certain social media posts as hate speech is a case in point. In a recent study, Maarten Sap et al. (2019) demonstrate that many widely used hate speech training data sets contain a correlation between the 'toxicity' or 'hatefulness' of the language and whether or not the speaker used linguistic markers of African American vernacular English. They likewise demonstrate that studies using these data sets to train their models then propagate and extend these biases to such an extent that "tweets by self-identified African Americans are up to two times more likely to be labeled as offensive compared to others" (Sap et al., 2019). In short, annotators who 'label' hate speech training corpora are informed by their own assumptions about what speech that is hateful looks like. If the annotator is white, they might find speech by other racial demographics to be 'more hateful' compared to speech from their own demographic group. These unarticulated heuristics guide how the annotators label data and introduce hidden biases into research. This is particularly important because machine learning algorithms trained on data that contain small-scale, latent biases can then amplify those biases when the algorithms are applied to other corpora at scale.

In contrast, qualitative approaches that incorporate abductive reasoning would acknowledge 'labeling' as an intellectual contribution in itself—not a self-evident application of fact but a theoretically consequential process that should be described and justified in the explication of methods. The judgments that go into labeling data would not simply disappear as hidden bias, but get explicitly integrated into the interpretation of patterns that emerge through analysis. As a researcher analyzes their data, they work simultaneously to fit a piece of evidence into existing frameworks and also to *update* those frameworks as necessary to better accommodate the real world as depicted by the data. In other words, the labeling of data in qualitative methods (what qualitative researchers would instead call 'coding') is not a matter of mere assumption, but rather a systematic part of the theory-building process.

One very common approach to 'coding' a corpus of qualitative data falls under the guise of what is known as grounded theory development. Grounded theory spans both objectivist and constructivist approaches (Charmaz, 2000), but all take the categorization and organization of data to be not merely the matter of labeling a fact, but of developing a particular ontological perspective. While some grounded theorists (especially its early champions) claim this method of theory development to be purely inductive (Glaser & Strauss, 1967), we draw here on methodologists who acknowledge and embrace abduction in grounded theory (Coffey & Atkinson, 1996; Richardson & Kramer, 2006; Thornberg & Charmaz, 2013). This position recognizes that researchers are never 'blank slates' when analyzing their data. Instead, they have assumptions, expectations, and preexisting theories about how the world works that are iteratively interrogated and incorporated into

emergent explanations of their data. Because iteration is a key feature of grounded theory, and most abductive qualitative analyses more generally, here we briefly describe a common, idealized procedure for qualitatively coding textual data.

The process of iteration in many approaches to grounded theory begins with a round of 'open coding' in which researchers tag segments of their data (e.g., sentences, paragraphs, quotes, etc.) with summative keywords or phrases that typically stay very close to the language used in the original text. Next, there is often a second round of coding in which the researchers draw relationships between their open codes; for example, clustering them together and creating several overarching thematic categories. This step is sometimes referred to as 'theoretical coding' (Thornberg & Charmaz, 2013) when researchers begin to draw on concepts and theories from preexisting literature to craft salient categories for their data. This is usually followed by another round of coding in which the researchers return to their data corpus and selectively apply the newly crafted coding schema to test and further refine it. In this way, qualitative coding entails tacking back and forth between preexisting assumptions and emergent theories, and documenting each step in that process.

Throughout these rounds of coding, qualitative researchers take great care to interrogate and revisit the appropriateness of their codes, frequently through some sort of collaborative process. Instead of generating a quantitative measurement of intercoder reliability to purportedly demonstrate the absence of subjective bias, interpretive grounded theorists often discuss with others why and how they made particular coding decisions in "peer debriefing" sessions (Barbour, 2013) geared toward arriving at "dialogic intersubjectivity" (Saldaña, 2009), which can be thought of as "agreement through a rational discourse and reciprocal criticism between those interpreting a phenomenon" (Brinkmann & Steiner, 2018). Through this process, a team of researchers discusses why each researcher arrived at the decisions they made and the team deliberates together on differences in their interpretations. This dialogue prompts qualitative coders to acknowledge and articulate the assumptions and logics they employed in developing and applying codes. Importantly, dialogic intersubjectivity is not limited to a practice among research peers, but can also be pursued between researchers and the subjects of their inquiry. One way to do this is through a 'member check,' which entails sharing preliminary coding schemas, ideas, or analyses with some of the people who are represented in the data in order to solicit their feedback.

While it is impossible to guarantee that participants in such dialogues do not share the same 'blind spots' (Barbour, 2014), these processes nonetheless can provide important occasions for surfacing and recording the biases, assumptions, and logics involved in qualitative coding. In data science, such steps could go a long way in evaluating if and when it is appropriate to computationally scale an annotation procedure. For instance, in the previous example about racial bias in hate speech detection algorithms, the process of pursuing dialogic intersubjectivity among researchers or between researchers and subjects represented in the text corpus could highlight how the application of labels might rely on bankrupt misconceptions about race.

If acknowledged and systematized, abductive reasoning is a powerful approach that allows for theories to be updated based on real-world evidence and helps scholars reduce the extent to which their own biases and assumptions shape how they measure, interpret, and extrapolate from a piece of evidence. As we have argued, part of the problem is that many computationally mediated quantitative traditions lack established norms for articulating, systemizing, documenting, and evaluating the process of abduction in their work. What if we thought, then, about certain data science techniques, like supervised machine learning, as qualitative approaches at scale? This means that data science researchers need not reinvent the wheel when grappling with how to soundly integrate and develop theory. They can and should draw on the expertise qualitative researchers have developed in exercising abductive reasoning and describing its process.

4. Reflexivity

A variety of data science techniques, when applied to social questions, are commonly critiqued in academic scholarship (e.g., Mittelstadt et al., 2016), policy analysis (e.g., United States Executive Office of the President, 2014), trade books (e.g., O'Neil, 2016), and journalistic investigations (e.g., Marconi et al., 2019) for their potential to exacerbate inequality, undermine democratic processes, violate norms of privacy, and circumvent due process. For example, several years ago, ProPublica famously exposed that algorithms used in the criminal justice system for predicting recidivism are less accurate for people of color, leading to people of color being denied bail with disproportionate frequency (Angwin et al., 2016).

This outcome is plainly and profoundly unjust, and significant ongoing work in computer science is dedicated to making algorithms 'more fair' (ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT), n.d.). But for many critics, tweaking mathematical models and algorithmic decision systems to make them 'less bad' is insignificant (Gangadharan & Niklas, 2019) if the larger social systems they operate within are inherently unjust or oppressive (Eubanks, 2018). As Cathy O'Neil (2016) has asked, why do we make predictions of recidivism in order to decide who should be denied release from incarceration? This question assumes that punitive measures against individuals are the most appropriate ways to address crime (to say nothing of the fact that policies based on these predictions would constrain a person based on others' actions). Why do we not instead try to predict what kinds of programs and experiences in the criminal justice system lead to less recidivism, a question that stems from the view that criminal justice should be rehabilitative? Or why not use data to interrogate the basis of the concepts of crime and criminality? We concur with David Leslie's (2021) statement that "Where data scientists, who view themselves simply as socially disembodied, quantitative analysts, engineers, or code-churners go wrong is that they are insufficiently attentive to the commitments and values that undergird the integrity of their knowledge practices and the ethical permissibility of the projects, enterprises, and use-contexts in which they involve themselves."

Data scientists should constantly ask themselves questions about why they study what they study, what the social ramifications for their work will be, and what assumptions are going unremarked in their work—exercises that are core to a reflexive practice (D'Ignazio & Klein, 2020; Leurs, 2017). Although there is no

consensus on how to define or enact reflexivity in qualitative research (Day, 2012; Mauthner & Doucet, 2003), here we understand reflexivity to be a process by which researchers systematically reflect upon their own position relative to their object, context, and method of inquiry. For many qualitative social scientists, reflexivity means spending time thinking about and disclosing how their own biases, identities, experiences, and premises influence their work. This is important because, as Donna Haraway (1989) has demonstrated, our personal starting points (for example, our experiences of class, gender, race or ethnicity, training, entry point into a given project, etc.) can all influence what we study and what we find. Failure to acknowledge and discuss these starting points is problematic in all research. But such lapses may have particularly problematic outcomes in data science, which researchers, policymakers, and industries routinely use to make sweeping generalizations about large swaths of society and to develop public interest applications with particularly high stakes (Stone, 2017), as in the incarceration example above.

You might be a data scientist participating in a competition to predict where crime will occur (National Institute of Justice, 2019) so that police patrols can be more efficiently assigned—something that dozens of U.S. cities have recently tried to do (Haskins, 2019). Without being trained in practices of reflexivity, you may not stop to consider that you have embraced this task based on an assumption you hold that the criminal justice system can and should be made more efficient rather than fundamentally reformed or abolished. You may hold this view because you were not raised in a community with a historically fraught relationship with law enforcement, or because you are not deeply aware of the racialized nature of the current criminal justice framework (Alexander, 2012). If you were trained in reflexivity, however, you would be equipped to recognize and critically interrogate these assumptions. Perhaps after engaging in a reflexive process, you would arrive at the decision that you cannot, in good conscience, participate in building such a system (Barocas et al., 2020). Or perhaps you would still decide to participate in the competition, but disclose in writing your concerns and the thought process that led to that decision.

Recently there have been a number of promising efforts to establish reflexive norms within data-intensive computational practices. For example, the 2020 Association of Computing Machinery Conference on Fairness Accountability and Transparency (ACM FAccT, n.d.) featured several interactive sessions that introduced reflexivity and related concepts as a way of addressing transgressions in machine learning, artificial intelligence, and algorithmic technologies (Goss et al., 2020; Kaeser-Chen et al., 2020; Wan et al., 2020). In *Data Feminism*, Catherine D'Ignazio and Lauren Klein (2020) not only describe the importance of reflexivity and provide examples of data science projects that center reflexive practices, but also model what reflexive disclosure in research can entail.

Here, we build on these urgent calls for greater reflexivity by exploring how this concept might be practically incorporated into the day-to-day work of data science as an integral part of the research method. After all, "reflexivity can be thought of as a method of meta-analysis," according to the qualitative methodologist Annette Markham (2017):

> The basic position of reflexivity is analyzing the self recursively and critically in relation to the object, context, and process of inquiry. It's more than just reflection, which is what we get when we look in a mirror. Rather, it's like trying to look at yourself looking in the mirror. (Markham, 2017)

To make that rather abstract idea more concrete, we present several suggested exercises that could be incorporated into data science practice, all of which have been adopted and modified from Markham's (2017) web essay, "Reflexivity: Some Techniques for Interpretive Researchers":

1. *Brain Dump*. This is a timed writing exercise in which researchers reflect on certain prompting questions about their work. Examples of prompts that invite reflexive thinking include: 'What do I already know about this subject?' and 'Why am I studying this?' and 'What do I expect or hope to find, and why?' In answering these questions during a 'brain dump,' one should very intentionally avoid consulting or referencing external sources. Although the prompts may be similar to those questions that are typically answered during a literature review, the point of this exercise is not to identify 'the state of knowledge in the field' or 'lacunas in the extant literature,' but to articulate and examine the ideas and assumptions that the researcher has internalized in their own head. Timing the exercise helps to ensure that the insights generated are honest, raw, and unfiltered. For example, a real answer to the question, 'What do I already know about this subject?' might involve some personal, first-hand experience with the phenomenon of study. And a complete answer to the question, 'Why am I studying this?' might not hinge purely on intellectual curiosity—it may also involve some variation of motivations such as, 'because my advisor needs me to do it,' 'because there is funding to study it,' 'because this is an issue that impacts someone I care about,' or 'because this will confirm my worldview.'

2. *Situational Mapping*. This exercise explores the researcher's position with respect to other relevant entities, including persons, organizations, and objects. The goal is not only to surface links between the self and others, but also to expose variations and interrogate asymmetries that exist in these relationships. For example, the first author of this essay leads an internship program in which students learn data science through projects intended to have societal impact. Team members conduct a power-mapping exercise at the outset of each project, allowing program participants to position each stakeholder (including themselves) relative to how much influence the stakeholder has over the work, and how the stakeholder will be affected by the conclusions the project produces. Importantly, this process frequently forces practitioners to acknowledge that they do not know very much about some of the persons or organizations that are affected by their work, which in turn prompts them to learn more about the positions and perspectives of those entities and to think about the broader ramifications of their work.

3. *Toolbox Critique*. In this exercise, researchers interrogate the suite of resources, ideas, approaches, and technologies at their disposal—this may include everything from theories to data, software packages, and methods of analysis. The researcher asks, 'Am I using this data set because it is the best possible data set for understanding the phenomenon I am interested in, or because it is the data that is readily available to me? Am I using this method because it is the most reliable option or because it is an approach I am interested in

learning? Am I using this programming language because it is the most efficient for the job, or because it is the most prevalent in my field?' Answering these questions honestly can help surface the personal values and cultural norms that typically go unstated in research, but nonetheless shape it in powerful ways.

These three exercises may be put to best use if incorporated at the outset of a new project or at key inflection points in the research process. However, another reflexive practice involves carefully and continuously documenting the entire research process (Watt, 2007). This includes not only recording each decision point or judgment call (the 'what'), but also the basis for that decision (the 'why'). This daily, reflexive aspect of research documentation can take the form of a journal, akin to a traditional lab notebook:

> Rather than erasing one's previous thoughts, one simply notes new additions or modifications. Keeping dates on each entry can help illustrate how the researcher is changing through the course of the study. During this process, it is useful to ask questions of oneself such as the following: What led me to that perception? How do I know that? So what? Why did I conclude that? (Markham, 2017, n.p.)

We argue that the field of data science is fertile ground for incorporating reflexive practices such as those described here. Data scientists are often quite self-aware and self-critical of their methods and techniques (Hahn et al., 2018; Moats & Seaver, 2019; Neff et al., 2017; Pink, Lanzeni, & Horst, 2018; Pink, Ruckenstein et al., 2018; Ribes, 2019; Tanweer et al., 2016). Indeed, this very journal, *Harvard Data Science Review*, is a testament to the introspection that has characterized the emergence of data science. Moreover, in part because data science relies on the circulation and reuse of data and code (Meng, 2016), many academic data scientists have been 'first responders' (of sorts) to the so-called reproducibility crisis by building a movement to introduce greater levels of transparency in scientific research (Nosek et al., 2015). This includes establishing norms for preregistration of hypotheses, publication of data, and open access code. These important practices are intended to overcome the incentive in quantitative research to 'cover up' mistakes, dead ends, and research limitations (Brookshire, 2016). But these open science practices do little to address how a researcher's own subjective experiences shape every step of the inquiry process. A reflexive stance acknowledges that subjectivity and bias are not aberrations that can ever be fully eradicated from research but inherent aspects of human inquiry that should be acknowledged and accounted for. As such, we see reflexivity as a complement to the push for transparency that is already underway in data science—a complement that is necessary to fulfill the potential of data science methods for understanding the social world. Feminist epistemologists have long argued that it is impossible for a researcher to erase their own subjectivities, and that it is only through acknowledging and articulating these subjectivities that a researcher can approach a more complete understanding of the world (Harding, 1992). Qualitative traditions have long encouraged reflexive accounting of the complete process of studying a phenomenon, from the inception of a research question to the interpretation of findings, from theory building to theory testing, from the influence of the researcher on the phenomenon of study to the limitations of the study design. Data scientists are well-equipped to adopt this process of reflexive accountability, and, if they do so, their resulting conclusions would better represent,

understand, and support the social world. Therefore, we propose shifting the conversation to be about *explicitness* in the data science research process, which would encompass both the emerging norms around transparency in data science and the reflexivity practices that have emerged in qualitative research.

5. Example Application

For the sake of clarity, we have presented interpretivism, abduction, and reflexivity as distinct concepts. In reality, they are often entwined and mutually reinforcing in qualitative research. Interpretive insights are gained through an abductive process that integrates reflexive exercises. How does this look in practice? How could data science practitioners update their existing approaches with an eye for centering qualitative thinking—specifically by incorporating interpretivism, abductive reasoning, and reflexivity?

5.1. Case Background

We model practical steps to help integrate qualitative sensibilities into a data science project through a description of ongoing research conducted by two authors of this article and their collaborators (Dreier et al., 2021). The project asked, "How do government officials internally rationalize policies that violate the rights of their citizens?" During times of real, perceived, or constructed security crises, liberal democracies routinely deny rights protections to certain subsets of their citizenry, claiming those restrictions are necessary to maintain or reestablish national security. Britain's "Troubles in Northern Ireland'' provide a historical case in point. In an effort to quell escalating sectarian violence, and in impudent disregard for due-process rights, Britain in 1971 authorized internment without trial for those suspected of violence. More than 1,800 nationalists were interned. Publicly, Britain rationalized internment as a necessary response to a dire security situation. But were these the true motivations for internment? How did officials internally rationalize these violations to themselves and their colleagues? To answer these questions, Dreier et al. (2021) consulted digitized archives of British prime ministers' security-related correspondence during the early years of the Troubles (1969–73).

5.2. Context on Methods

Traditional means of analyzing historical collections involve painstaking qualitative coding procedures, which limit researchers to a relatively small universe of relevant documents but allow them to develop deep, textured, and complex understandings of the processes they study. Today, if the universe of relevant data exceeds a human coder's reasonable capacity to qualitatively analyze it, the researcher can turn to computational text-analysis tools to automate the identification of concepts of interests within text. In doing so, researchers dramatically expand the amount of text they can analyze. However, this can come at the expense of nuance, interpretability, or recognition that policy processes take place within—and are shaped by—extended historical time periods (Rast, 2012, p. 6; see also, Pierson, 2004).

The rapidly evolving field of natural language processing (NLP) offers promising advances for computationally recognizing the complex, multifaceted ideas that pervade the social world. NLP's family of

transformer-based approaches (e.g., BERT or RoBERTa) contextualize a word's vector representation and are trained on hand-coded data to accomplish a task. If and when NLP models achieve an acceptable level of agreement with human coders on a set of training data, the model could then be used to annotate the remaining unlabeled ('held-out') text in a corpus. Under the right circumstances, these models may be used to bring qualitative methods to scale while capturing some degree of complexity.

In order for NLP model outputs to be useful, however, the researcher must develop contextually meaningful classification schemes (for annotating the training data), and even then, models are shaped by social contexts and biases built into language. Therefore, to most effectively implement NLP's state-of-the-art technologies, Dreier et al. (2021) integrated interpretivist, abductive, and reflexive qualitative sensibilities into the NLP pipeline, particularly as the research team developed and coded for concepts that would later be used to train an NLP model. In this sense, this research approached NLP tools as augmenting and amplifying (rather than replacing) qualitative methods and thinking.

5.3. Qualitative Thinking to Develop Meaningful Classification Schemes

Before turning to NLP, researchers must first develop the categories they are interested in examining, establish the boundaries between those categories, and code segments of the data according to those coding guidelines. However, identifying classification categories in real text is an inevitably complicated and subjective process. Concepts can be difficult to pinpoint and distinguish from one another. They are often only identifiable by a researcher who has detailed case-study knowledge. And the process of establishing and coding classification categories is shaped by researchers' interpretive understanding of the case, their preconceived theoretical understandings of the concepts of interest, their methodological understandings about the relationship between temporal sequences and causation (Grzymala-Busse, 2011), and how the researchers abductively update those understandings as directed by the evidence they observe.

When used as the basis for an NLP automation, subjective coding decisions can substantially shape model outcomes and researcher conclusions. When coding text data for NLP implementation, therefore, we urge researchers to ground their work in qualitative sensibilities and to move away from the idea that they are 'labeling' true instances, instead embracing the idea that they are 'coding' carefully defined but inevitably subjective concepts. The exemplary research of Dreier et al. (2021) took the following steps to integrate interpretivism, abduction, and reflexivity into the process of developing and coding the concepts of interest in their study:

Investing in case study knowledge to develop an interpretive intuition. Understanding that identifying and making inferences about rationalizations for internment in Northern Ireland was highly context-specific, Dreier et al. (2021) leveraged insights from interpretive methods, spending about three months mapping the conflict and developing an understanding about the context in which their archive data were created and curated. Archival data (like many text-based data sources) are unavoidably incomplete (Decker, 2014),

underrepresent or omit certain actors (Decker, 2013), and often prioritize the worldviews of those in power (Stoler, 2008). By acknowledging these empirical realities, the study's research team was able to consider the relative importance of different types of evidence, identify subtle clues in the data, discuss the potential directions of bias in the study's analysis, and update the study's coding ontology accordingly.

Furthermore, the study team's rich understanding of the case study allowed the researchers to interpret meaning and connect ideas that would have otherwise appeared unrelated, and to adapt the study's coding scheme to include these connections. For example, the government in Northern Ireland initiated internment alongside banning public marches. A coder with detailed case-study knowledge of the sectarian conflict may recognize that those bans were directly related to internment: Northern Ireland banned marches (which would disproportionately affect Protestants) in an effort to prevent Catholics from feeling singularly targeted (by internment). Based on this understanding, Dreier et al. (2021) treated bans on marches as part of the government's efforts to publicly rationalize internment.

More broadly, understanding the historical processes of change through case study inquiry requires scholars to temper impulses to treat data points as ahistorical, generalizable demonstrations of causality. Instead, researchers should adopt an interpretative appreciation for how social processes occur over extended periods of time and are shaped by context-specific junctures and processes (Grzymala-Busse, 2011; Howlett & Jayner 2006; Pierson, 2004; Rast, 2012;). Indeed, Britain's efforts to rationalize internment make little sense when interpreted outside the context of the post–World War II 'liberal consensus' that held states accountable to honoring individuals' rights.

Using an abductive approach to developing coding ontology. A purely deductive framework might encourage researchers to develop predefined categories and maintain those categories even if the data reveals flaws in that approach. Instead, the research team in Dreier et al. (2021) intentionally developed coding stages that allowed the team to abductively and systematically update the study's coding categories—as indicated by the case study and data—before executing the bulk of coding. First, the researchers surveyed a sample of documents to ground their intuitions and then developed the study's coding scheme. Next, the researchers coded a small subset of documents, carefully recording their coding decisions, judgments, uncertainties about boundaries between categories, unanticipated categories, and evident flaws in their assumptions. The study's authors then modestly updated their classification categories based on these initial coding observations.

Researchers will inevitably encounter their own blind spots as they begin coding new text data, and abductive approaches allow researchers to update their parameters accordingly. In the case of Dreier et al. (2021), the authors initially conflated political motivations for internment and military benefits of internment into one rationalization category representing the government's strategic motivations for internment. By building abduction into their pipeline, Dreier et al. (2021) were able to update their categories to distinguish between what became two of the most conflicting attitudes about internment's necessity: political pressures to impose internment versus skepticism about its military advantages.

Integrating reflexive tactics into coding and analysis. Dreier et al. (2021) maintained an extensive research journal, which they referred to as "field notes," following each coding session. These field notes helped to identify changes in how the research team thought about categories, to identify potential disagreements between coders, to identify coding rules that were unclear, to note any individual biases or starting points that could shape how each coder uniquely reacted to a given piece of evidence, and to record textured observations about the case and relationships between variables of interest. These field notes became an invaluable source of data for the study's substantive analysis; captured critical events, meta-shifts over time, and contextual meaning; allowed the research team to consider and confront their own biased reactions to the data; and yielded a methodological appendix that comprehensively detailed the study's qualitative construction of quantitative data (a form of research transparency that is too often omitted in the publication process, to the detriment of downstream users and the scientific process).

By adopting and carefully documenting the qualitative sensibilities of interpretivism, abductive reasoning, and reflexivity into this study's coding pipeline, Dreier et al. (2021) accomplished at least four things that would have otherwise been impossible: updated the study's coding scheme to meaningfully reflect the data and context; discerned the relative importance of different types of evidence; considered the potential biases in the research team's analysis and reactions; and, ultimately, provided the study's downstream NLP models with more accurate, systematically developed annotated data. (See Dreier and Gade, 2021, for further details on a step-by-step process for incorporating interpretation, abduction, and reflection into a data science pipeline.)

5.4. Acknowledging Biases

We close discussion of the Troubles in Northern Ireland example by encouraging scholars implementing NLP (or data science tools in general) to broadly apply a critical qualitative lens to acknowledge the biases within the computational models they use. Language is complicated and constantly changing. Words and their meanings are idiosyncratic to an industry, geography, and time ('IRA' might refer to a retirement account in one text collection and a cadre of political and paramilitary groups in another). And how people use and interpret words is shaped by dominant or privileged voices in a given context. State-of-the-art NLP models adapt a word's vector representation based on its context, but word vectors will inevitably retain social biases and other undesirable associations that are present in the text on which vectors are pretrained. It is not computationally feasible to fully address—or even fully discover—these issues, and when used as starting points for NLP analysis, such associations run the risk of reinforcing social inequalities (Bender et al., 2021; Blodgett et al., 2020; Sap et al., 2019).

These concerns are particularly salient to the corners of data science that analyze text data to map ideology, flag hate speech, contain the spread of misinformation, anticipate protests against injustice, or track plans for violence or insurrection. However, understanding that context shapes meaning, and that biases that reinforce privilege hide within our 'objective' data sources, are concerns with which *all* data scientists must contend. Adopting qualitative sensibilities of interpretivism, abduction, and reflexivity will help position data scientists

to take these concerns seriously by carefully reconsidering their assumptions, qualifying their results, and attending to the possible biases embedded within their projects.

6. Conclusion

We have argued that many of the current problems with data science as it is applied to the production of social knowledge could be mitigated through the integration of qualitative approaches. But qualitative ways of understanding the world have tremendous value beyond what they can do to systemize data science research practices. So let us be clear: We do not believe qualitative methods should be co-opted as the handmaiden (Hesse-Biber, 2010) of data science. Nor are we arguing for a qualitative methods toolbox that data scientists can casually dip into, deploying an interview or two here, some field observations there. While such dabbling may prove valuable in certain cases, we envision a more fundamental shift in the way we practice data science as it is applied to social research, so that certain qualitative sensibilities are substantively integrated into data-intensive social science. We have argued that, to integrate qualitative approaches in a manner not decoupled from the epistemological positions undergirding them, data science of social phenomena should draw on qualitative practices related to interpretivism, abductive reasoning, and reflexivity.

Although we have demonstrated certain ways in which these sensibilities are compatible with trends and norms in data science, we also realize that their integration into data science practice is likely to encounter some friction. For example, we described an intensive process of iterative coding and dialogue among qualitative researchers engaged in grounded theory development. Such a method can result in less biased and more nuanced findings but does not necessarily lend itself to being reproduced—a gold standard for data science research. One could rightly argue that the naive 'labeling' of training data with which we contrasted the grounded theory coding approach is also not reproducible. But it is worth acknowledging more broadly that qualitative research typically aspires toward justifiability of findings rather than reproducibility or replicability of methods precisely because it is geared toward understanding the nuances of particular social contexts rather than producing universal claims. We do not recommend that data science abandon its invaluable commitment to reproducibility any more than we would advise surrendering the nuance that qualitative sensibilities can uniquely generate. Rather, data scientists must evaluate the tradeoffs between contextual integrity, reproducibility, and scalability to determine the appropriate approach for any given project. This does not mean defanging data science. To the contrary, Leonelli (2021) has compellingly argued that data science could be rendered *more* incisive if coupled with qualitative approaches. Assessing the role of data science in understanding COVID-19, Leonelli argues that data scientists can inform a more tailored, effective, and sustainable response to the pandemic by eschewing a narrow focus on predictive models and embracing investigations into the relationships between disease and socio-environmental conditions within localized contexts—inquiries that necessitate the inclusion of qualitative questions, qualitative data, and qualitative expertise.

While we have discussed data science as applied to the production of *social knowledge*, the main problems that we highlight—including decontextualized data, hidden biases, and an uncritical approach to research topics—are also present in data science that is not applied to the production of knowledge about the social world. Our focus has been on applications in the social domain due to our own interests and expertise, but our argument could no doubt be further extended. For instance, data in medicine and biology is systematically biased, and researchers in those and other fields may also benefit from a more humanistic approach (Stevens et al., 2018). Artificial intelligence and machine learning programs like recommender systems, which are not oriented toward intelligibly producing social knowledge of the sort found in the social sciences, have been one of the central objects of critique recently in critical data studies (Benjamin, 2019; Noble, 2018). Even problems in areas as far afield as physics, pure mathematics, geology, or astronomy are heavily influenced by the positionality of the individuals researching them, and hence would benefit from greater reflexivity. Exploring the potential for the integration of qualitative sensibilities across these various veins of work is an exciting direction for future work.

To be sure, qualitative research encompasses a heterogeneous set of approaches, and there is a difference between adopting sensibilities and radically changing the methodological approaches of data science. The gap between practicing data analytics and undertaking an ethnography, for example, is significant, and we are not suggesting that all data scientists should (or could) become ethnographers. Instead, we have argued that the 'sensibilities' of ethnography and other qualitative methods can influence how questions are formulated, how findings are interpreted, and how implications are framed, and that some aspects of qualitative methods themselves can be integrated into the data science research pipeline. We allow that there is a spectrum in terms of how deeply interpretivism, abduction, and reflexivity may be taken up in data science and how robust the relationship between data science and qualitative traditions might become. At one end of the spectrum, formal exposure to qualitative sensibilities can, at the very least, help illuminate how qualitative thinking is always already implicit at various stages of data science research—from determining that a research question is salient, to defining variables, to drawing conclusions from patterns in the data—even if that fact typically goes unrecognized in quantitative analyses (Meng, 2016). Qualitative sensibilities can be deployed to systematize qualitative thinking inherent in data science, making it more 'methodical' so to speak and better equipped to accurately quantify the social world. At the other end of the spectrum, earnestly engaging with qualitative sensibilities could fundamentally alter the approach to data science research and result in a true blending of quantitative computational and qualitative methods. Here, inspiration can be drawn from work Ograjenšek and Gal (2016) have done to reimagine statistical education in a way that is unshackled from a narrow range of analytical techniques and reoriented toward a 'need to know.' In this scenario, researchers would deploy the modes of thinking and analysis best suited to answering a question of interest, whether those methods be derived from qualitative or quantitative traditions:

> [Q]ualitative and quantitative data and research methods should not be seen as mutually exclusive enterprises. They should be perceived as building blocks that co-exist under the larger umbrella of

> research. (Ograjenšek & Gal, 2016)

We similarly suggest that qualitative and quantitative approaches should not merely be 'mixed' but should be considered as complementary and co-constitutive elements of producing social knowledge through data science. Such an approach can improve data science by tempering its findings, surfacing its modes of failure, and adding nuance to its intellectual contributions. It can also allow data science to ask new and different questions in the first place. We have focused on relatively circumscribed methodological innovations in data science, rather than the kinds of radical shifts called for by authors such as Catherine D'Ignazio and Lauren Klein (2020) and Sasha Costanza-Chock (2020), or projects such as Erase the Database (Erase the Database, n.d.) and Our Data Bodies (Our Data Bodies, n.d.)—all of which explicitly center emancipatory perspectives like antiracism and intersectional feminism. We offer relatively more revisionist suggestions, not because we do not support the same goals (we do), but because we believe that today's data scientists could readily adopt the incremental shifts we outlined here, and that the training and experience required to make these changes will help lay the groundwork for more transformative work.

As a starting point, qualitative scholars must be welcomed into conversations about how the academic community trains future generations of data scientists. And, at the very least, data scientists must be conversant enough in qualitative sensibilities and the subjective realities of knowledge production to understand the strengths and limitations of both quantitative and qualitative methods, to know when qualitative approaches are appropriate, and to collaborate with experts in qualitative research to improve and expand their ability to understand the social world. This means that our current data revolution necessitates not only cultivating increased capacity in quantitative and computational programs but also building up qualitative research in social science and humanities departments rather than continuing to disinvest from them—a troubling side effect of the data-driven turn in the academy. Such an investment will bear the fruit of data science research that is more reliable, ethical, and meaningful.

Acknowledgments

Disclosure Statement

References

ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT). (n.d.). *ACM conference on fairness, accountability, and transparency*. https://facctconference.org/

Agre, P. (1997). Toward a critical technical practice: Lessons learned in trying to reform AI. In G. C. Bowker, L. Gasser, S. L. Star, & W. Turner (Eds.), *Social science, technical systems, and cooperative work: Bridging the great divide* (pp. 131-158). Erlbaum Associates.

Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.

Anderson, C. (2008, June 23). The end of theory: The data deluge makes the scientific method obsolete. *Wired*. http://www.wired.com/2008/06/pb-theory/

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Barbour, R. S. (2013). Quality of data analysis. In U. Flick (Ed.), *The Sage handbook of qualitative data analysis* (pp. 496–509). Sage. https://doi.org/10.4135/9781446282243.n34

Barocas, S., Biega, A. J., Fish, B., Stark, L., & Niklas, J. (2020). When not to design, build, or deploy [Plenary Session]. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (p. 695). https://doi.org/10.1145/3351095.3375691

Basta, C., Costa-Jussà, M. R., & Casas, N. (2019). Evaluating the underlying gender bias in contextualized word embeddings. In *Proceedings of the First Workshop on Gender Bias in Natural Language Processing* (pp. 33–39). http://doi.org/10.18653/v1/W19-3805

Bates, J. (2018). Data cultures, power and the city. In R. Kitchin, T. P. Lauriault, & G. McArdle (Eds.), *Data and the city* (pp. 189–200). Routledge. http://doi.org/10.4324/9781315407388-14

Bates, J., Cameron, D., Checco, A., Clough, P., Hopfgartner, F., Mazumdar, S. Sbaffi, L., Stordy, P., & de la Vega de León, A. (2020). Integrating fate/critical data studies into data science curricula: Where are we going and how do we get there? In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 425–435). https://doi.org/10.1145/3351095.3372832

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 Conference on Fairness, Accountability, and Transparency* (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code*. Polity Press.

Blank, G., & Lutz, C. (2017). Representativeness of social media in Great Britain: Investigating Facebook, LinkedIn, Twitter, Pinterest, Google+, and Instagram. *American Behavioral Scientist*, *61*(7), 741–756. https://doi.org/10.1177/0002764217717559

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *arXiv*. https://doi.org/10.48550/arXiv.2005.14050

Bolukbasi, T., Chang, K. W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In *NIPS'16: Proceedings of the 30th International Conferences on Neural Information Processing Systems* (pp. 4356–4364).

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.

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

Brinkmann, S., & Steiner, K. (2018). *Doing interviews* (2nd ed.). Sage. https://doi.org/10.4135/9781529716665

Brookshire, B. (2016, October 21). Blame bad incentives for bad science. *Science News*. https://www.sciencenews.org/blog/scicurious/blame-bad-incentives-bad-science

Burke, M., & Kraut, R. (2013). Using Facebook after losing a job: Differential benefits of strong and weak ties. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (pp. 1419–1430). https://doi.org/10.1145/2441776.2441936

Burns, R. (2015). Rethinking big data in digital humanitarianism: Practices, epistemologies, and social relations. *GeoJournal*, *80*(4), 477–490. https://doi.org/10.1007/s10708-014-9599-x

Burns, R., Dalton, C. M., & Thatcher, J. E. (2018). Critical data, critical technology in theory and practice. *The Professional Geographer*, *70*(1), 126–128. https://doi.org/10.1080/00330124.2017.1325749

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Cao, L. (2017). Data science: Challenges and directions. *Communications of the ACM, 60*(8), 59–68. https://doi.org/10.1145/3015456

Carson, D. (2009). The abduction of Sherlock Holmes. *International Journal of Police Science and Management, 11*(2), 193–202. https://doi.org/10.1350/ijps.2009.11.2.123

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.) (pp. 509-535). Sage.

Clarke, A. E. (2003). Situational analyses: Grounded theory mapping after the postmodern turn. *Society for the Study of Symbolic Interaction, 26*(4), 553–576. https://doi.org/10.1525/si.2003.26.4.553

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Sage.

Colaner, S. (2020, June 12). AI weekly: AI phrenology is racist nonsense, so of course it doesn't work. *Venture Beat*. https://venturebeat.com/2020/06/12/ai-weekly-ai-phrenology-is-racist-nonsense-so-of-course-it-doesnt-work/

Costanza-Chock, S. (2020). *Design justice: Community-led practice to build the worlds we need*. MIT Press.

Crawford, K. (2013, April 1). The hidden biases in big data. *Harvard Business Review*. https://hbr.org/2013/04/the-hidden-biases-in-big-data

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. https://data-feminism.mitpress.mit.edu/

Davey, M., Kirchgaessner, S., & Boseley, S. (2020, June 3). Surgisphere: Governments and WHO changed COVID-19 policy based on suspect data from tiny US company. *The Guardian*. https://www.theguardian.com/world/2020/jun/03/covid-19-surgisphere-who-world-health-organization-hydroxychloroquine

Day, S. (2012). A reflexive lens: Exploring dilemmas of qualitative methodology through the concept of reflexivity. *Qualitative Sociology Review*, *8*(1), 61–84. http://doi.org/10.18778/1733-8077.8.1.04

De Veaux, R. D., Agarwal, M., Averett, M., Baumer, B. S., Bray, A., Bressoud, T. C., Bryant, L., Cheng, L. Z., Francis, A., Gould, R., Kim, A. Y., Kretchmer, M., Lu, Q., Moskol, A., Nolan, D., Pelayo, R., Raleigh, S., Sethi, R. J., Sondjaja, M., . . . Ye, P. (2017). Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application*, *4*(1), 15–30. https://doi.org/10.1146/annurev-statistics-060116-053930

Decker, S. (2013). The silence of the archives: Business history, post-colonialism and archival ethnography. *Management & Organizational History*, *8*(2), 155–173. https://doi.org/10.1080/17449359.2012.761491

Decker, S. (2014). Solid intentions: An archival ethnography of corporate architecture and organizational remembering. *Organization*, *21*(4), 514–542. https://doi.org/10.1177/1350508414527252

Douven, I. (2011). *Abduction*. Stanford Encyclopedia of Philosophy. Retrieved from https://plato.stanford.edu/entries/abduction/

Dreier, S., Serrano, S., Gade, E., & Smith, N. (2021). *Troubles in text: Finetuning NLP to recognize government rationalizations for rights abuses* (Working paper).

Dreier, S., & Gade, E. (2021). *Qualitative sensibilities for data science research pipeline*. https://www.skdreier.com/uploads/7/5/0/0/75005149/qual_data_science_pipeline.pdf

Dumit, J., & Nafus, D. (2018). The other ninety per cent: Thinking with data science, creating data studies—an interview with Joseph Dumit. In H. Knox, & D. Nafus (Eds.), *Ethnography for a data-saturated world* (pp. 252–274). Manchester University Press. https://doi.org/10.7765/9781526127600.00020

Erase the Database. (n.d.). *Home*. Retrieved from http://erasethedatabase.com/

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Evans, M., & Mathews, A. W. (2019, October 25). Researchers find racial bias in hospital algorithm. *The Wall Street Journal*. https://www.wsj.com/articles/researchers-find-racial-bias-in-hospital-algorithm-11571941096

Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. *Journal of Broadcasting & Electronic Media*, *58*(1), 59–75. https://doi.org/10.1080/08838151.2013.875018

Gangadharan, S. P., & Niklas, J. (2019). Decentering technology in discourse on discrimination. *Information Communication and Society*, *22*(7), 882–899. https://doi.org/10.1080/1369118X.2019.1593484

Garg, N., Schiebinger, L., Jurafsky, D., & James, Z. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), E3635–E3644. https://doi.org/10.1073/pnas.1720347115

Geiger, R. S., & Halfaker, A. (2017). Operationalizing conflict and cooperation between automated software agents in Wikipedia: A replication and expansion of "Even Good Bots Fight." *Proceedings of ACM Human-Computer Interaction*, *1*(2), Article 49. https://doi.org/10.1145/3134684

Geiger, R. S., & Ribes, D. (2011). Trace ethnography: Following coordination through documentary practices. In *44th Hawaii International Conference on System Science* (pp. 1–10). IEEE. https://doi.org/10.1109/HICSS.2011.455

Geiger, R. S., Yu, K., Yang, Y., Dai, M., Qiu, J., Tang, R., & Huang, J. (2020). Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from? In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 325–336). https://doi.org/10.1145/3351095.3372862

Gitelman, L. (Ed.). (2013). *"Raw data" is an oxymoron*. MIT Press.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Routledge.

Goldberg, A. (2015). In defense of forensic social science. *Big Data & Society*, *2*(2), 1–3. https://doi.org/10.1177/2053951715601145

Gonen, H., & Goldberg, Y. (2019). Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. *arXiv*. https://doi.org/10.48550/arXiv.1903.03862

Goss, E., Hu, L., Sabin, M., & Teeple, S. (2020). Manifesting the sociotechnical: Experimenting with methods for social context and social justice. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (p. 693). ACM. https://doi.org/10.1145/3351095.3375682

Granville, K. (2018, March 19). Facebook and Cambridge Analytica: What you need to know as fallout widens. *The New York Times*. https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

Grzymala-Busse, A. (2011). Time will tell? Temporality and the analysis of causal mechanisms and processes. *Comparative Political Studies*, *44*(9), 1267–1297. https://doi.org/10.1177/0010414010390653

Hahn, C., Hoffman, A. S., Slota, S. C., Inman, S., & Ribes, D. (2018). Entangled inversions: Actor/analyst symmetry in the ethnography of infrastructure. *Interaction Design and Architecture(S)*, *38*, 124–139. http://www.mifav.uniroma2.it/inevent/events/idea2010/doc/38_7.pdf

Haraway, D. J. (1989). *Primate visions: Gender, race, and nature in the world of modern science*. Routledge, Chapman & Hall.

Harding, S. (1992). Rethinking standpoint epistemology: What is "strong objectivity"? *The Centennial Review*, *36*(3), 437–470.

Hargittai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *The Annals of the American Academy of Political and Social Science*, *659*(1), 63–76. https://doi.org/10.1177/0002716215570866

Hargittai, E. (2020). Potential biases in big data: Omitted voices on social media. *Social Science Computer Review*, *38*(1), 10–24. https://doi.org/10.1177/0894439318788322

Haskins, C. (2019, February 9). Dozens of cities have secretly experimented with predictive policing software. *Vice Motherboard*. https://www.vice.com/en_us/article/d3m7jq/dozens-of-cities-have-secretly-experimented-with-predictive-policing-software

Hesse-Biber, S. (2010). Qualitative approaches to mixed methods practice. *Qualitative Inquiry*, *16*(6), 455–468. https://doi.org/10.1177/1077800410364611

Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009). *The fourth paradigm: Data-intensive scientific discovery*. Microsoft Research. http://research.microsoft.com/en-us/collaboration/fourthparadigm/

Hill, M. B., & Shaw, A. (2020). Studying populations of online communities. In B. F. Welles, & S. Gonzalez-Bailon (Eds.), *The handbook of networked communication* (pp. 173–193). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190460518.013.8

Howlett, M., & Rayner, J. (2006). Understanding the historical turn in the policy sciences: A critique of stochastic, narrative, path dependency and process-sequencing models of policy-making over time. *Policy Sciences*, *39*(1), 1–18. https://doi.org/10.1007/sl

Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, *3*(2), 1–7. https://doi.org/10.1177/2053951716674238

Jungherr, A. (2019). Normalizing digital trace data. In N. J. Stroud, & S. C. McGregor (Eds.), *Digital discussions: How big data informs political communication* (pp. 9–35). Routledge. https://doi.org/10.4324/9781351209434-2

Jungherr, A., Schoen, H., Posegga, O., & Jürgens, P. (2017). Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support. *Social Science Computer Review*, *35*(3), 336–356. https://doi.org/10.1177/0894439316631043

Kaeser-Chen, C., Dubois, E., Schüür, F., & Moss, E. (2020). Translation tutorial: Positionality-aware machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (p. 704). https://doi.org/10.1145/3351095.3375666

Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, *323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Leonelli, S. (2014). What difference does quantity make? On the epistemology of Big Data in biology. *Big Data & Society*, *1*(1), 1–11. https://doi.org/10.1177/2053951714534395

Leonelli, S. (2021). Data Science in Times of Pan(dem)ic. *Harvard Data Science Review*, 3(1). https://doi.org/10.1162/99608f92.fbb1bdd6

Leurs, K. (2017). Feminist data studies: Using digital methods for ethical, reflexive and situated socio-cultural research. *Feminist Review*, *115*(1), 130–154. https://doi.org/10.1057/s41305-017-0043-1

Lewis, S. C., & Molyneux, L. (2018). A decade of research on social media and journalism: Assumptions, blind spots, and a way forward. *Media and Communication*, *6*(4), 11–23. https://doi.org/10.17645/mac.v6i4.1562

Lindgren, S. (2020). *Data theory*. Polity Press.

Marconi, F., Daldrup, T., & Pant, R. (2019, February 14). Acing the algorithmic beat, journalism's next frontier. *Nieman Lab*. https://www.niemanlab.org/2019/02/acing-the-algorithmic-beat-journalisms-next-frontier/

Marcus, G., & Davis, E. (2014, April 6). Eight (no, nine!) problems with big data. *The New York Times*. https://www.nytimes.com/2014/04/07/opinion/eight-no-nine-problems-with-big-data.html

Markham, A. N. (2017, Feb 28). Reflexivity: Some techniques for interpretive researchers. *Annettemarkham*. https://annettemarkham.com/2017/02/reflexivity-for-interpretive-researchers/

Marres, N. (2021). For a situational analytics: An interpretative methodology for the study of situations in computational settings. *Big Data & Society*, *7*(2), 1–16. https://doi.org/10.1177/2053951720949571

Mauthner, N. S., & Doucet, A. (2003). Reflexive accounts and accounts of reflexivity in qualitative data analysis. *Sociology*, *37*(3), 413–431. https://doi.org/10.1177/00380385030373002

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. John Murray.

Mayer, K., & Malik, M. M. (2019). *Critical data scientists at work: Summary report of the ICWSM-2019 Workshop on Critical Data Science*. GitHub. https://critical-data-science.github.io/

Mayo, D. (2020). P-values on trial: Selective reporting of (best practice guides against) selective reporting. *Harvard Data Science Review*, *2*(1). https://doi.org/10.1162/99608f92.e2473f6a

Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research and Politics*, *4*(3), 1–9. https://doi.org/10.1177/2053168017720008

Meng, X.-L. (2016). Discussion: The Q-q dynamic for deeper learning and research. *International Statistical Review*, *84*(2), 181–189. https://doi.org/10.1111/insr.12151

Meng, X.-L. (2021). What are the values of data, data science, or data scientists? *Harvard Data Science Review*, *3*(1). https://doi.org/10.1162/99608f92.ee717cf7

Meyer, R. (2014, June 28). Everything we know about Facebook's secret mood manipulation experiment. *The Atlantic*. https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, *50*(1), 181–201. https://doi.org/10.1111/j.1467-9787.2009.00641.x

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Moats, D., & Seaver, N. (2019). '"You social scientists love mind games"': Experimenting in the '"divide"' between data science and critical algorithm studies. *Big Data & Society, 6*(1), 1–11. https://doi.org/10.1177/2053951719833404

Moss, E., Chowdhury, R., Rakova, B., Schmer-Galunder, S., Binns, R., & Smart, A. (2019). Machine behaviour is old wine in new bottles. *Nature*, *574*(7777), 176. https://doi.org/10.1038/d41586-019-03002-8

National Institute of Justice. (2019). *Real-time crime forecasting challenge*. https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge

National Science Foundation. (2019). *Harnessing the data revolution*. https://www.nsf.gov/news/special_reports/big_ideas/harnessing.jsp

Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and contribute: A practice-based framework for improving critical data studies and data science. *Big Data*, *5*(2), 85–97. https://doi.org/10.1089/big.2016.0050

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3–42. https://doi.org/10.1177/0049124117729703

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. *Algorithms of Oppression*. New York University Press.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M. Dafoe, A., Eich, E., Feese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., . . . Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

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

Offenhuber, D. (2018). Sticky data: Context and friction in the use of urban data proxies. In R. Kitchin, T. P. Lauriault, & G. McArdle (Eds.), *Data and the city* (pp. 99–108). Routledge. http://doi.org/10.4324/9781315407388-8

Ograjenšek, I., & Gal, I. (2016). Enhancing statistics education by including qualitative research. *International Statistical Review*, *84*(2), 165–178. https://doi.org/10.1111/insr.12158

Our Data Bodies. (n.d.). *Our Data Bodies*. https://www.odbproject.org/

Parry, M. (2018, March 4). Data scientists in demand: New programs train students to make honest sense of numbers. *The Chronicle of Higher Education*. https://www.chronicle.com/article/Colleges-Rush-to-Ride/242674

Peirce, C. S. (2013). Quote from "Harvard lectures on pragmatism: Lecture VII, a deleted passage" [1903]. In *Commens: Digital Companion to C.S. Peirce*. http://www.commens.org/dictionary/entry/quote-harvard-lectures-pragmatism-lecture-vii-deleted-passage

Pentland, A. (2014). *Social physics: How good ideas spread-the lessons from a new science*. Penguin Press.

Pierson, P. (2004). *Politics in time: History, institutions, and social analysis*. Princeton University Press.

Pink, S., & Lanzeni, D. (2018). Future anthropology ethics and datafication: Temporality and responsibility in research. *Social Media + Society*, *4*(2), 1–9. https://doi.org/10.1177/2056305118768298

Pink, S., Lanzeni, D., & Horst, H. (2018). Data anxieties: Finding trust in everyday digital mess. *Big Data & Society*, *5*(1), 1–14. https://doi.org/10.1177/2053951718756685

Pink, S., Ruckenstein, M., Willim, R., & Duque, M. (2018). Broken data: Conceptualising data in an emerging world. *Big Data & Society*, *5*(1), 1–13. https://doi.org/10.1177/2053951717753228

Rast, J. (2012). Why history (still) matters: Time and temporality in urban political analysis. *Urban Affairs Review*, *48*(1), 3–36. https://doi.org/10.1177/1078087411418178

Reichertz, J. (2010). Abduction: the logic of discovery of grounded theory. *Forum: Qualitative Social Research*, *11*(1), Article 13. https://www.qualitative-research.net/index.php/fqs/article/view/1412/2902

Resnyansky, L. (2019). Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge. *Big Data & Society*, *6*(1), 1–12. https://doi.org/10.1177/2053951718823815

Ribes, D. (2019). STS, meet data science, once again. *Science, Technology, & Human Values*, *44*(3), 514–539. https://doi.org/10.1177/0162243918798899

Richardson, R., & Kramer, H. E. (2006). Abduction as the type of inference that characterizes the development of a grounded theory. *Qualitative Research*, *6*(4), 497–513. https://doi.org/10.1177/1468794106068019

Richterich, A. (2018). *The big data agenda: Data ethics and critical data studies*. University of Westminster Press. https://doi.org/10.2307/j.ctv5vddsw

Rosenberg, A. (2015). *Philosophy of social science* (5th ed.). Avalon Publishing.

Sack, W. (2019). Rhetoric. In *The software arts* (pp. 145–202). MIT Press. https://doi.org/10.7551/mitpress/9495.003.0009

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage.

Sap, M., Card, D., Gabriel, S., Choi, Y., & Smith, N. A. (2019). The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 1668–1678). http://doi.org/10.18653/v1/P19-1163

Selbst, A. D., Friedler, S. A., Venkatasubramanian, S., Vertesi, J., boyd, d., & Venkatasubrama, S. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency* (pp. 59–68). https://doi.org/10.1145/3287560.3287598

Sloane, M., & Moss, E. (2019). AI's social sciences deficit. *Nature Machine Intelligence*, *1*(8), 330–331. https://doi.org/10.1038/s42256-019-0084-6

Song, I.-Y., & Zhu, Y. (2017). Big data and data science: Opportunities and challenges of iSchools. *Journal of Data and Information Science*, *2*(3), 1–18. https://doi.org/10.1515/jdis-2017-0011

Stevens, M., Wehrens, R., & De Bont, A. (2018). Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society*, *5*(2), 1–21. https://doi.org/10.1177/2053951718816727

Stoler, A. (2008). *Along the archival grain: Epistemic anxieties and colonial common sense*. Princeton University Press.

Stone, A. (2017, March). When big data gets it wrong. *Government Technology (Govtech)*. https://www.govtech.com/data/When-Big-Data-Gets-It-Wrong.html

Tanweer, A., Fiore-Gartland, B., & Aragon, C. (2016). Impediment to insight to innovation: Understanding data assemblages through the breakdown–repair process. *Information, Communication & Society*, *19*(6), 736–752. https://doi.org/10.1080/1369118X.2016.1153125

Tate, E. (2017, March 15). Data analytics programs take off. *Inside Higher Ed*. https://www.insidehighered.com/digital-learning/article/2017/03/15/data-analytics-programs-taking-colleges

Taylor, L., & Purtova, N. (2019). What is responsible and sustainable data science? *Big Data & Society*, *6*(2), 1–6. https://doi.org/10.1177/2053951719858114

Thatcher, J. (2014). Living on fumes: Digital footprints, data fumes, and the limitations of spatial big data. *International Journal of Communication*, *8*(1), 1765–1783. https://ijoc.org/index.php/ijoc/article/view/2174

Thornberg, R., & Charmaz, K. (2013). Grounded theory and theoretical coding. In U. Flick (Ed.), *The Sage handbook of qualitative data analysis* (pp. 153–170). Sage. http://doi.org/10.4135/9781446282243

Tsvetkova, M., García-Gavilanes, R., Floridi, L., & Yasseri, T. (2017). Even good bots fight: The case of Wikipedia. *PLOS One*, *12*(2), Article e0171774. https://doi.org/10.1371/journal.pone.0171774

United States Executive Office of the President. (2014). *Big data: Seizing opportunities, preserving values*. https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf

van Atteveldt, W., & Peng, T. Q. (2018). When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science. *Communication Methods and Measures*, *12*(2–3), 81–92. https://doi.org/10.1080/19312458.2018.1458084

Vigen, T. (2015). *Spurious correlations*. Hachette Books.

Wagner-Pacifici, R., Mohr, J. W., & Breiger, R. L. (2015). Ontologies, methodologies, and new uses of Big Data in the social and cultural sciences. *Big Data & Society*, *2*(2), 1–11. https://doi.org/10.1177/2053951715613810

Wan, E., de Groot, A., Jameson, S., Păun, M., Lücking, P., Klumbyte, G., & Lämmerhirt, D. (2020). Lost in Translation: An interactive workshop mapping interdisciplinary translations for epistemic justice. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* https://doi.org/10.1145/3351095.3375685

Zook, M., Barocas, S., boyd, d., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., Pasquale, F. (2017). Ten simple rules for responsible big

data research. *PLoS Computational Geography, 13*(3), Article e1005399. https://doi.org/10.1371/journal.pcbi.1005399