




# Critical computational social science

Sarah Shugars<sup>1\*</sup> 

\*Correspondence:

[sarah.shugars@rutgers.edu](mailto:sarah.shugars@rutgers.edu)

<sup>1</sup>Department of Communication,  
Rutgers University, 4 Huntington St.,  
New Brunswick, NJ, USA

## Abstract

In her 2021 IC2S2 keynote talk, “Critical Data Theory,” Margaret Hu builds off Critical Race Theory, privacy law, and big data surveillance to grapple with questions at the intersection of big data and legal jurisprudence. As a legal scholar, Hu’s work focuses primarily on issues of governance and regulation—examining the legal and constitutional impact of modern data collection and analysis. Yet, her call for Critical Data Theory has important implications for the field of Computational Social Science (CSS) as a whole. In this article, I therefore reflect on Hu’s conception of Critical Data Theory and its broader implications for CSS research. Specifically, I’ll consider the ramifications of her work for the scientific community—exploring how we as researchers should think about the ethics and realities of the data which forms the foundations of our work.

**Keywords:** Computational Social Science; Data governance; Data ethics; Privacy; Surveillance

## 1 Introduction

When the U.S. Census first launched in 1790, it included only three categories for race: “Free white males, free white females” was one category, accompanied by “All other free persons” and “Slaves.” However, these categories weren’t fixed and were updated with each decade’s census. By 1890, for example, four of the eight available racial categories were dedicated to classifying a person’s percentage of “Black blood” down to the level of “one-eighth or any trace” [1]. This official, government-recorded determination was made by a census enumerator who would assume a subject’s race by looking at them. It wasn’t until 1960 that people in the U.S. could self-report their own race and only in 2000 that people could indicate multi-racial identities. Even now, ethnoracial self-identification doesn’t always align with legal labels, as in the case of Middle East and North African populations who are legally classified as white in the United States [2].

The ever-changing racial parameters of the U.S. Census illustrates of one of the deepest challenges to computational work: “data” is not a fixed, unbiased reflection of reality. It is always the product of a given time and place, of a particular way of conceptualizing and operationalizing the world. Indeed, the term “race” itself is rare in censuses globally—only an estimated 63% of countries collect this type of information, with most asking specifically about “ethnicity” or “national origin” [3]. Furthermore, there is wide global variation in whether this data is collected through closed or open-ended responses [3], leading to notable differences in what identities are captured or can even be captured.

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This variation is not just a matter of limitations or incomplete data. Rather, as Kevin Guyan writes, “Decisions made about who to count, what to count, and how to count are not value-neutral but bring to life a particular vision of the social world” [4]. Guyan’s insight stems from his extensive work tracing data collection about queer populations in the U.K., and points to the same underlying challenge: “Numbers do not speak for themselves—they always speak for someone” [4]. Importantly, these numbers, and the decisions that generate them, can have lasting, real-world impact. For example, the racial categories used in a census may contribute to the systematic undercounting of racially marginalized populations [5]. How a census is conducted and what constitutes a “fair” allotment of resources based on census data is a social and political question that cannot be answered with mathematical neutrality [6, 7].

As scientists, we want to think of our data as neutral; as reasonably accurate with predictable and known dimensions of error. But in reality, data is always a reflection of the social context which generated it. What is measured and how it’s measured are social decisions, not universal constructs. To further complicate matters, we as researchers are often embedded in the same social systems which generated the data we study, meaning that we may not always be aware of the ways in which our social context affects our data. A given strategy for collecting racial data may seem obvious in one time and place while seeming nonsensical in another.

In her 2021 IC2S2 keynote talk, “Critical Data Theory,” Margaret Hu, Professor of Law at William & Mary Law School and Research Affiliate with the Institute for Computational and Data Sciences at Penn State University grapples with such challenges in the context of big data and legal jurisprudence [8]. In coining the term “Critical Data Theory,” Hu builds upon a century of work in Critical Theory, a school of thought which has inspired efforts in Critical Legal Studies, Critical Race Theory, Feminist Studies, and more [9].

Advanced by scholars of the Frankfurt School in the early 20th century [9, 10], Critical Theory sought to interrogate the foundational assumptions of society and to conduct social scientific inquiry aimed at improving society rather than just describing it [11, 12]. The movement responded in part to the empiricist ideals of the Enlightenment, which assumed that a neutral quantification of reality was both possible and desirable. Perhaps epitomized by Francis Bacon’s famous axiom that “knowledge is power,” Enlightenment thinkers held measurement as the pinnacle of human achievement. However, the flaw in such empiricist thinking became painfully clear amidst the rise of European fascism as the mantle of scientific inquiry was used to justify eugenics and other abhorrent social projects [13]. In response, Critical Theory agrees that knowledge is indeed power, but further adds, as Bent Flyvbjerg put it, that “power is knowledge” [14]. In this way, Critical Theory and its descendants can perhaps best be understood as a philosophical orientation towards knowledge production and application. It is not a testable or falsifiable theory comparable to theories found in the natural sciences. Rather, it is a lens through which to monitor, evaluate, and be critical of, the scientific process itself.

Critical Theory has particularly flourished within the legal domain under the name Critical Legal Studies. Initially articulated in the 1970s, Critical Legal Studies focuses on the social dimensions of law, with particular attention to the role of *power* [9]. Those with power write the law and, intentionally or not, the law upholds the interests and perspectives of those who write it [15, 16]. From this broad conception of the law as a social construct, legal theories around the role of race in law [17–19] and the role of gender in law

[20, 21] began to emerge. The first of these literatures constitutes the legal philosophy better known as Critical Race Theory.

In defining Critical Data Theory, Hu draws on these traditions to argue that “big data must be subjected to critical theoretical treatment” [9]. As a legal scholar, Hu’s work focuses primarily on issues of governance and regulation, examining the legal and constitutional impact of modern data collection and analysis. Yet, her call for Critical Data Theory has important implications for the field of Computational Social Science (CSS) as a whole.

In this article, I therefore first reflect on Hu’s conception of Critical Data Theory and then examine its broader implications for Computational Social Science research. Building off my own background in political communication, civic studies, and computational social science, I’ll consider the ramifications of her work for the scientific community, exploring how we as researchers should think about the ethics and realities of the data that forms the foundations of our work.

## 2 The call for critical data theory

In her keynote talk, Hu argues for a Critical Data Theory that would serve to “deconstruct the relationship between law, power, and emerging technological developments” [8]. She explains that tools of big data, AI, and data science have fundamentally shifted legal, scientific, socio-economic, and political frameworks of power—particularly as they relate to the concept of “self.” Increasingly, a person’s identity can be measured and monitored through the trace data [22] they generate as they live life in a digital society [23]. This passively collected data of clicks, scroll time, location, and more has been central to the ‘big data revolution’ and is tied to the emergence of a concept alternately called digital personhood [24], the cyber self [25], the networked self [26], or the data self [27]. Hu’s work finds that, within the U.S., these digital identities are increasingly interpreted as meaningful measures of the “self” and may be used as mechanisms for governing [23]. For example, through mass cybersurveillance programs, U.S. intelligence agencies construct “digital avatars” which serve as targets for investigation [23]. Such an “amalgamation of data” [23] may be used to authorize a drone strike though it may not represent an actual or known person [23].

Pointing to Critical Race Theory as a foundational inspiration, Hu argues that the process of navigating legal understandings of “self” in the context of digital identities parallels the history of “race negotiation and definition” [9]. As a legal philosophy, Critical Race Theory contends that concepts of “race” are highly dynamic and actively shaped by social “hierarchies of law and power” [9]. Just as we saw through the changing history of racial data collection in the U.S. Census, neither law nor data itself is truly neutral, but rather emerge from social values, understandings, and ideologies [9]. As a result, regimes of privilege are maintained by the rule of law despite legal guarantees of equality [9, 18]. Critical Race Theory therefore examines and critiques the construction of these racial hierarchies and the subordination of racialized populations through policy, governance, and jurisprudence [9, 17, 18].

Critical Data Theory similarly aims to interrogate hierarchies of power, but this legal scholarship focuses on the constructs that emerge from digital data. Just as Critical Race Theory aims to highlight that “race is not a static phenomenon or fixed concept” [9], Critical Data Theory aims to complicate and interrogate the construction of the digital self [23].

Hu argues that Critical Data Theory has become necessary as we move from a “small data world” to a “big data world” [23, 28]. Until recently, our social, legal, and technical understanding was governed entirely by “small” data: “knowledge that humans can see, touch, analyze, and perceive without the assistance of supercomputing capabilities” [23]. While both “small” and “big” data are subject to human bias and risk misrepresenting social constructs as natural facts, Hu argues that the size and scope of this challenge has grown significantly with the advent of “big” data. In the legal domain, for example, Hu finds that small data surveillance tends to be technologically and logistically limited, relying on human capacity, judgement, and interpretable evidence that can be seen, shared, discussed, and debated [23]. Big data cybersurveillance, on the other hand, allows for vast troves of biometric and bibliographic data to be automatically collected, stored, and analyzed [28, 29]. In other words, as Hu argues, referencing the work of boyd & Crenshaw, big data creates new forms of knowledge as well as new processes for producing that knowledge [23, 30].

This suggests that the challenges of big data go beyond the need to articulate standards for AI ethics or to address algorithmic discrimination. Such work plays an important role in minimizing harms, but Critical Data Theory more fundamentally aims to “deconstruct the legal and constitutional impact of big data” and systems of power which enable a big data world [9]. Specifically, Critical Data Theory interrogates the ways in which big data and computational analysis “normalize surveillance technologies” within our “day-to-day governance” [9, 31]. It actively encourages “counterintuitive counternarratives” [9] that challenge presumed truths about data, data governance, and our digital selves [9]. Hu argues that this critical approach to data theory works to protect individuals by providing needed friction between the interests of big tech companies and the State [32]. In other words, by insisting on the thoughtful development of laws and norms around growing technologies, Critical Data Theory works to support proper oversight and regulation in digital spaces.

As a legal scholar, Hu’s conception of Critical Data Theory focuses largely on the legal and governance implications of advances in data collection and computation. Her work pays particular attention to issues of cybersurveillance and privacy and she closed her IC2S2 keynote talk by illustrating how Critical Data Theory can be applied in this space. Building on Balkin & Levinson’s conception of the National Surveillance State [31], Hu argues that the era of big data has given rise to a new Cybersurveillance State in which our digital selves are the primary target for governance [28]. For example, after the 9/11 terrorist attacks the U.S. government began collecting and centralizing massive amounts of human data in an effort to identify and track potential threats [28]. This continued state of emergency later gave rise to an immigration policy of “extreme vetting” [33] in which the U.S. government aimed to build a fully automated tool which would sweep all social media and internet data in order to anticipate the future intentions of individuals crossing the border [9]. More recent crises, such as the COVID-19 pandemic, have given further cause for governments around the world to collect, centralize, and analyze vast quantities of human trace data [32].

In confronting these massive systems of cybersurveillance, Critical Data Theory reminds us that the “digital avatars” [23] constructed from a person’s online presence are merely a proxy for—and not a true measure of—the person who generated that data. There may be meaningful signal in that collection of trace data, but we must remain critical of

what that online persona represents and should resist the temptation to interpret a collection of passive outputs as a reflection of the true self. Perhaps most critically, we should be skeptical of efforts which aim to treat this “data self” as an object of governance. A citizen is not synonymous with the data they generate. One reason this distinction is important is because it is not possible for a person to know all of the secondary and tertiary ways in which their online identity can be captured, used, and potentially abused [9]. This means that even as people become aware of large-scale efforts to leverage their trace data, they can’t fully curate their online identity and often aren’t aware of micro-targeting efforts conducted on the basis of that perceived identity.

In short, Critical Data Theory provides important friction around big data systems. It doesn’t claim this data can’t or shouldn’t be used at all, but rather persistently asks critical questions of *what* is being measured and *why*. Critical Data Theory interrogates *who* benefits from these systems and calls for us to constantly remember that these methods and measures are social outputs, “not any transcendent or a priori truth” [9].

### 3 Critical computational methods

While conceived of as a legal philosophy, Critical Data Theory raises important questions for computational work more broadly. How can we measure and analyze the world when the data themselves are subjective? How can we interpret social constructs such as race and gender while acknowledging that these identities are malleable and shift over time? How can we aim to understand the world without perpetuating the systems of power which have defined that world? These questions are not new and are core to critical approaches employed by social scientists for over a century [9–11].

In thinking about broader implications of Critical Data Theory, I therefore begin by returning to the core insight of Critical Theory that has stretched through the social sciences for generations: data is never objective, it is merely imbued with the air of impartiality by systems of power. The role of power in controlling, manipulating, and defining what is then seen as “neutral” information has been documented in cases around the world [34]. For example, while studying urban planning in Denmark, Bent Flyvbjerg repeatedly found instances where people with power subtly shaped the terms of debate: not only determining what information was shared, but more fundamentally controlling what even counted as information [14]. In the U.S., John Gaventa found similar dynamics playing out in the Appalachian Valley, as those with power continuously defined the terms of reality for those without power [35]. Importantly, in all these cases, power doesn’t merely dominate as the loudest voice in the room: it more perniciously shapes reality [35], determining who gets to be in the room and what even counts a voice.

The critical lens then presents a fundamental challenge to all efforts of measurement and quantification. One might be tempted accuse critical approaches of advocating for an entirely relativistic interpretation of the world, rendering all attempts at standardization and generalization useless. Yet such a claim oversimplifies and belittles the contributions of Critical Theory. Mirroring statistician George E. P. Box’s famous phrase “all models are wrong, but some are useful,” Critical Theory doesn’t end with the claim that “all data is wrong.” Rather, while wrong, imperfect, and reflections of power, Critical Theory similarly concedes that some data is useful.

The challenge is that the usefulness of data is often apparent while the “wrongness” can be easy to forget—particularly for those who benefit from the established systems

of power. Furthermore, ignoring the role power plays in constructing data and making meaning can lead to real and lasting harm. James C. Scott argues that many of the worst human tragedies of the 20th century came about through the uncritical application of administrative tools designed to document and standardize society [13]. Scott concedes that administrative ordering itself—the standardization of names, practices of landownership, documentation of income, and more—are not inherently bad. These are the necessary tools of statecraft required to “make a society legible” to the State that governs it [13]. Yet the forcible application of those measures—attempts by the State to impose a presumed proper order onto citizens—has led to terrible human atrocities time and time again.

Perhaps this is why Critical Theory has persisted so firmly within the context of legal studies. Historically, the State has been the primary instrument for quantifying society; for arbitrating what categories exist and articulating how those categories can be interpreted and measured. For example, in 1976 the U.S. Supreme Court heard a case in which five Black women sued General Motors for employment discrimination [36]. The case not only asked whether the plaintiffs themselves had been subject to discrimination, but implicitly asked the State to determine whether their doubly-marginalized identities as both Black and women could simultaneously be taken into account. The court ultimately refused to see these identities as overlapping: although all the Black people General Motors hired were men and all the women they hired were white, the company did not seem to discriminate on the basis of race alone nor on the basis of gender alone. While individuals in the dual category of “Black women” did seem to face discrimination, this was not an identity the State was prepared to recognize and thus the women lost their suit. This example is one of the cases Kimberlé Crenshaw pointed to in coining the term “intersectionality” [36] to express the social and legal erasure of Black women and their experiences. The law upheld the dominant social conception of “race and gender as mutually exclusive categories of experience and analysis” [36], and the failure to examine this conception with a critical eye caused real material harm for Black women.

The modern computational context—in which massive amounts of data can be collected, automatically analyzed, and acted upon—has not only raised the stakes of failing to apply a critical lens but has shifted where and how the quantification of society is navigated. Increasingly corporations, as well as the individual researchers and data scientists who conduct computational analysis, hold the power to determine what information goes into the model and how the resulting algorithmic output is used. We are living in an era not just of the Cybersurveillance State [28], but of the corporate surveillance state—as companies gather increasingly detailed data about our lives and use that information to influence our experiences, outcomes, and decisions [37, 38]. In developing curation algorithms and recommender systems, private companies now have tremendous power over what information people come into contact with [39–41]. Private data science models hold power over who qualifies for loans or gets access to other resources [42]. Public-private partnerships have been used to develop automated tools to help judges and other officials with sentencing decisions [43, 44]. Nearly every aspect of our lives is monitored through data collection and that data is then used to shape our experiences and perceptions of reality. And much of this is done beyond the purview of the State.

Much like pre-computational efforts to standardize categories and measures, these big data mechanism systematically erase some identities—choosing, for example, not to collect or report data on racial and gender minorities who comprise a relatively small share of

the population [4, 45]. Arguably, this is out of necessity: it is simply not practical nor possible to develop models which fully account for the complexity of human existence. Yet, if all models consistently erase these populations—and do so without a critical assessment of who is being erased and why—this approach truly can cause harm to the people who are systematically not counted or are otherwise misrepresented by measurements. Furthermore, the lasting reification of measurement from Enlightenment-era thinking can compound this problem, as the algorithmic output generated by simplified data and models is erroneously interpreted as reflecting some natural truth [46–48].

Consider, for example, the case of Large Language Models (LLMs) which have consistently shown to reflect gender bias through the terms, occupations, and assumptions assigned to the binary genders of “male” and “female” while simultaneously failing to account for any other genders. In itself, such “algorithmic bias” is not inherently bad. Indeed, the output of these models is merely a reflection of the data which went into them. This means that bias in algorithmic output can actually be a useful tool for interrogating bias in the society which generated that model’s data. This is the critical approach to computational analysis—an ongoing process of continually questioning why and how data was generated and examining the structures of power which influenced that genesis. The real problem of algorithmic bias occurs when these outputs are interpreted *uncritically*—when a society’s assumptions and bias are codified into technological systems which both repeat and reinforce that bias. This non-critical approach occurs when socially-generated data is implicitly assumed to reflect the natural order of society, suggesting that it can be unquestionably used in search engines, predictions tasks, classification efforts, or other systems which serve to support the established ordering of society.

All of this suggests that Critical Data Theory is needed not only as a legal philosophy, but as a computational philosophy: as a pro-active standard for constantly questioning what is measured, how its measured, and why its measured. As Hu suggests, it is absolutely essential for legal scholars to examine how systems of big data seep into our governance and to push back on efforts to treat our “digital avatars” [23] as the object of governance. But the promise of Critical Data Theory goes beyond such legal efforts. As private interests increasingly control the mechanisms through which society measures, organizes, and standardizes, Critical Data Theory ought to expand to broadly encompass all sites of computational analysis. Building off the roots of Critical Theory, such a Critical Data Theory would work to ensure that neither data nor output is unquestionably accepted at face value. It would add, as Hu puts it, a needed element of “friction” to this work [32]—demanding researchers, public and private, consistently think critically about their work and its potential impact on individuals and society.

#### **4 Committing to critical computational social science**

The call for critical computational social science is not new [45, 49–52]. Indeed, one might argue that the “social science” part of CSS implies or even demands taking a critical perspective. After all, one can’t truly study the social world without attention to the human systems which shape and define that world [53]. Personally, I would like to see critical theory assumed as part of the definition of CSS. Work that happens to take a computational approach to social systems is not automatically “Computational Social Science.” Rather, true CSS requires a respect for social science theory and an understanding of the social mechanisms which underlie the methods and data we use.

Yet, in the current time and place—where the technological capacity to use big data and computational methods doesn't require social science training—it is worth explicitly noting the “critical” piece of Computational Social Science. Data scientists who work with data generated by or impacting humans should be trained in critical methods. Questioning the provenance and implications of such social data should be expected as a matter of course. Again, these data are unarguably *useful*, but we must remember that they are also *wrong*. And ignoring the ways in which they are wrong can lead to real harm [13, 36, 46].

A key challenge here is that it may not be readily apparent to researchers how to go about applying a critical approach, or what such an approach might look like. Particularly for scholars trained in empiricist traditions, the admonition to “think critically” about your data risks sounding trite and underspecified. An empiricist approach would expect rules, tests, and procedures for identifying and rectifying data issues. For example, there is a growing literature around how to “correct for bias” by minimizing its impact on downstream tasks [54, 55]. Such literature applies an important band-aid and helps interrupt a cycle in which biased human-generated data is used to train biased algorithmic systems. Yet, critical approaches intentionally push back on the very notion of being able to “correct” for bias.

At its heart, critical methods are about examining, confronting, and interrogating systems of power—a term which Julie E. Cohen eloquently refuses to define in her book, *Between Truth and Power*. “The essence of power lies precisely in its ability to shape-shift—to elude the perfect, crystalline characterizations,” she writes [38]. “Power in operation is pragmatic, seeking and finding paths of least resistance and mobilizing the practical and conceptual resources that appear ready to hand” [38]. Critical approaches acknowledge this “shape-shifting” nature of power and similarly refuse to define precise antidotes to its influence. Rather, researchers must continually interrogate—based on their own positionality and research context—how systems of power may influence their data, their methods, and their results.

For example, there are many studies for which it is entirely reasonable and appropriate to include a demographic analysis—examining behaviors by race, gender, or other personal characteristics. Yet, as we have seen, both race [17–19] and gender [4, 20, 21, 45] are social constructs, shaped by systems of power. Any effort to classify people by race and gender will necessarily simplify what categories of race and gender are considered, and will likely misclassify some people into categories they wouldn't, or would prefer not to, self-select. Furthermore, as we saw in the discussion of intersectionality [36], treating “race” and “gender” as separate constructs further obscures some identities by implying, for example, that all women share a collective experience.

A critical approach to computational social science wouldn't argue that such racial or gender classification is never appropriate but rather would encourage researchers to be continuously critical of the role power plays in such analysis. What categories of race or gender do you analyze? Why? Do you assume these are the only categories available? That everyone must fit into a single category? How do you treat observations that don't fit into these categories? How do you talk about these categories—as natural truths or as a pragmatic simplification? These are the questions of critical computational social science, and they must be consistently and methodically interrogated [50].



## 5 Conclusions

Hu's argument for Critical Data Theory provides a valuable framework for articulating the need for critical approaches in computational social science; for encouraging "counterintuitive counternarratives" [9] and consistently interrogating the systems of power which shape our assumptions and perspectives. In short, Critical Data Theory outlines a philosophy that can help us all better commit to conducting critical computational social science.

### Acknowledgements

The author would like to thank Termeh Shafie and Christoph Stadtfeld for their leadership of this special issue. I would also like to thank Margaret Hu for sharing her excellent scholarship, and the anonymous reviewers for their thoughtful and helpful feedback.

### Funding

Not applicable.

### Abbreviations

CSS, Computational Social Science; IC2S2, International Conference on Computational Social Science; LLM, Large Language Model.

### Data availability

Not applicable.

## Declarations

### Competing interests

The author declares no competing interests.

### Author contributions

Solo author made all contributions to paper. The author read and approved the final manuscript.

### Authors' information

Sarah Shugars: Assistant Professor, Department of Communication, School of Communication & Information, Rutgers University, 4 Huntington St., New Brunswick, NJ 08901, [sarah.shugars@rutgers.edu](mailto:sarah.shugars@rutgers.edu).

Received: 2 January 2023 Accepted: 30 November 2023 Published online: 26 February 2024

## References

1. Brown A (2020) The changing categories the U.S. census has used to measure race. *Pew Res Cent*
2. d'Urso AS (2022) A boundary of white inclusion: the role of religion in ethnoracial assignment. *Perspective Polit*, 1–18. <https://doi.org/10.1017/S1537592722003309>
3. Morning A (2015) Ethnic classification in global perspective: a cross-national survey of the 2000 census round. In: *Social statistics and ethnic diversity*, pp 17–37
4. Guyan K (2022) *Queer data: using gender, sex and sexuality data for action*. Bloomsbury Publ., New York
5. Sifre E (2021) The role of census data in policy and racial equity. Technical report, Institute on Taxation and Economic Policy. <https://itep.org/the-role-of-census-data-in-policy-and-racial-equity/>
6. Bouk D, Boyd D (2021) Democracy's data infrastructure: the technopolitics of the U.S. census. Technical report, Knight First Amendment Institute. <https://knightcolumbia.org/content/democracys-data-infrastructure>
7. Duchin M (2022) Introduction. In: Duchin M, Walch O (eds) *Political geometry: rethinking redistricting in the US with math, law, and everything in between*. Birkhäuser, Basel, pp 1–28
8. Hu M (2021) Critical data theory. 7th international conference on computational social science (IC2S2)
9. Hu M Critical data theory. *William & Mary Law Review* (Forthcoming)
10. Arato A, Gebhardt E, Piccone P (1978) *The essential frankfurt school reader*. Urizen Books, New York
11. Horkheimer M (1937) Traditional and critical theory. In: *Critical theory: selected essays*, pp 1–11
12. Horkheimer M, Adorno TW, Noeri G (1947) *Dialectic of enlightenment*. Stanford University Press, Redwood City
13. Scott JC (2008) *Seeing like a state*. Yale University Press, New Haven
14. Flyvbjerg B et al (1998) *Rationality and power: democracy in practice*. University of Chicago press, Chicago
15. Unger RM (1983) The critical legal studies movement. *Harvard Law Rev* 96(3):561–675
16. Kelman M, Kelman MG (1987) *A guide to critical legal studies*. Harvard University Press, Cambridge
17. Bell D (1991) Racial realism. *Conn Law Rev* 24:363
18. Carbado DW, Roithmayr D (2014) Critical race theory meets social science. *Ann Rev Law Soc Sci* 10:149–167
19. Crenshaw KW (2017) Race, reform, and retrenchment: transformation and legitimation in antidiscrimination law. In: *Law and social movements*. Routledge, New York, pp 475–531
20. MacKinnon CA (1989) *Toward a feminist theory of the state*. Harvard University Press, Cambridge
21. Bartlett K (2018) *Feminist legal theory: readings in law and gender*. Routledge, New York
22. Salganik MJ (2019) *Bit by bit: social research in the digital age*. Princeton University Press, Princeton
23. Hu M (2014) Small data surveillance v. big data cybersurveillance. *Pepperdine Law Rev* 42:773

24. Solove DJ (2004) *The digital person: technology and privacy in the information age*, vol 1. NYU Press, New York
25. Whitman CN, Gottdiener WH (2016) The cyber self: Facebook as a predictor of well-being. *Int J Appl Psychoanal Stud* 13(2):142–162
26. Boyd D (2010) Social network sites as networked publics: affordances, dynamics, and implications. In: *A networked self*. Routledge, New York, pp 47–66
27. Gordon R (2001) The electronic personality and digital self. *Dispute Resolut J* 56(1):8
28. Hu M (2017) From the national surveillance state to the cybersurveillance state. *Ann Rev Law Soc Sci*
29. Hu M (2022) Biometrics and an ai bill of rights. *Duquesne Law Rev* 60:283
30. Boyd D, Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inf Commun Soc* 15(5):662–679
31. Balkin JM, Levinson S (2006) The processes of constitutional change: from partisan entrenchment to the national surveillance state. *Fordham Law Rev* 75:489
32. Hu M (2022) Frictionless pandemic surveillance and social credit systems. In: *Pandemic surveillance*. Edward Elgar, Cheltenham Glos, pp 64–85
33. Hu M (2017) Algorithmic jim crow. *Fordham Law Rev* 86:633
34. Lukes S (1974) *Power: a radical view*. Palgrave Macmillan, London
35. Gaventa J (1982) *Power and powerlessness: quiescence and rebellion in an Appalachian valley*. University of Illinois Press, Champaign
36. Crenshaw K (1989) Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *Univ Chic Leg Forum* 139
37. Zuboff S (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. PublicAffairs
38. Cohen JE (2019) *Between truth and power*. Oxford University Press, London
39. Noble SU (2018) *Algorithms of oppression*. New York University Press, New York
40. O’neil C (2016) *Weapons of math destruction: how big data increases inequality and threatens democracy*. Crown, New York
41. Robertson RE, Jiang S, Joseph K, Friedland L, Lazer D, Wilson C (2018) Auditing partisan audience bias within Google search. In: *Proceedings of the ACM on human-computer interaction 2(CSCW)*, pp 1–22
42. Eubanks V (2018) *Automating inequality: how high-tech tools profile, police, and punish the poor*. St. Martin’s, New York
43. Angwin J, Larson J, Mattu S, Kirchner L (2016) Machine bias. ProPublica. Available online at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing/>
44. Benjamin R (2019) Race after technology: abolitionist tools for the new jim code. *Soc Forces*
45. Hanna A, Denton E, Smart A, Smith-Loud J (2020) Towards a critical race methodology in algorithmic fairness. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp 501–512
46. Scheuerman MK, Hanna A, Denton E (2021) Do datasets have politics? Disciplinary values in computer vision dataset development. *Proc ACM Hum-Comput Interact* 5(CSCW2):317. <https://doi.org/10.1145/3476058>
47. Buolamwini J, Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: *Conference on fairness, accountability and transparency*, pp 77–91. PMLR
48. Joseph K, Shugars S, Gallagher R, Green J, Mathé AQ, An Z, Lazer D (2021) (Mis)alignment between stance expressed in social media data and public opinion surveys. In: *Proceedings of the 2021 conference on empirical methods in natural language processing (EMNLP)*
49. Shugars S (2022) Core skills of computational social science: data justice, intellectual humility, and algorithmic thinking. <https://doi.org/10.4135/9781529611274>
50. Shugars S (2024) A matter of perspective: computational social science and researcher choice. In: Box-Steffensmeier J, Christenson D, Sinclair-Chapman V (eds) *Oxford Handbook of Methodological Pluralism*
51. Shugars S, Quintana Mathé A, Lange R, David L (2024) Gender visibility. *J Comput-Mediat Commun*
52. Tromble R (2019) In search of meaning: why we still don’t know what digital data represent. *J Digit Soc Res* 1(1):17–24
53. Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M et al (2009) Computational social science. *Science* 323(5915):721–723
54. TeBlunthuis N, Hase V, Chan C-H (2023) Misclassification in automated content analysis causes bias in regression. Can we fix it? Yes we can! arXiv preprint. [arXiv:2307.06483](https://arxiv.org/abs/2307.06483)
55. Jiang H, Nachum O (2020) Identifying and correcting label bias in machine learning. In: *International conference on artificial intelligence and statistics*, pp 702–712. PMLR

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.