

Seven intersectional feminist principles for equitable and actionable COVID-19 data

Big Data & Society
July–December: 1–6
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2053951720942544
journals.sagepub.com/home/bds

Catherine D'Ignazio¹ and Lauren F. Klein²

Abstract

This essay offers seven intersectional feminist principles for equitable and actionable COVID-19 data, drawing from the authors' prior work on data feminism. Our book, *Data Feminism* (D'Ignazio and Klein, 2020), offers seven principles which suggest possible points of entry for challenging and changing power imbalances in data science. In this essay, we offer seven sets of examples, one inspired by each of our principles, for both identifying existing power imbalances with respect to the impact of the novel coronavirus and its response, and for beginning the work of change.

Keywords

Data science, feminism, counterdata, data collection, coronavirus, antiracism

This article is a part of special theme on Viral Data. To see a full list of all articles in this special theme, please click here: <https://journals.sagepub.com/page/bds/collections/viraldata>

Our book, *Data Feminism* (D'Ignazio and Klein, 2020), which outlines an approach to data science and data ethics that is informed by intersectional feminism, was published just as the novel coronavirus began to take hold around the world. In the book, we offer seven principles for challenging power and privilege using data science. The principles are as follows: examine power, challenge power, elevate emotion and embodiment, rethink binaries and hierarchies, embrace pluralism, consider context, and make labor visible. In articulating these principles, we could not have anticipated the spread of the coronavirus. But our rationale for constructing the principles, as well as the lessons they prompt, teach us to anticipate the impact of this deadly virus. Our regional and national support systems are broken, with decisions of national consequence being made by powerful individuals and moneyed corporations, with the most harmful effects experienced by the least powerful among us. In what follows, we offer seven examples drawn from the coronavirus emergency, one inspired by each of our principles. Each principle helps to identify existing power imbalances with respect to the impact of the

coronavirus, and to begin the work of righting the balance of power in the world.

Before beginning, we want to note several authorial decisions: we the authors, Catherine and Lauren, speak as a “we” in this essay, which sometimes expands to include our perspectives as scholars who make use of data in our work. We do not intend this “we” to be a universal address. We also made an intentional decision to focus primarily on COVID-19 in a US context. While we attempt to bring in examples from around the world, the majority of the examples we cite derive from our own locations in (and expertise about) the United States. Finally, we should be clear that our feminism is intersectional, and takes at its starting point that feminism is about more than women or gender: it is about

¹Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA, USA

²Departments of English and Quantitative Theory and Methods, Emory University, Atlanta, GA, USA

Corresponding author:

Lauren F. Klein, Emory University, Atlanta, GA 30322-1007, USA.
Email: lauren.klein@emory.edu



power, about those who have it and those who don't, and how that imbalance of power can be challenged and changed.¹

Principle 1: Examine power

Data feminism begins by analyzing how power operates in the world

The first principle of data feminism relates to how power operates in the world. By power, we mean the current configuration of structural privilege and structural oppression in which some groups experience unearned advantages—because systems have been designed by people like them and work for people them—and other groups experience systematic disadvantages—because those same systems were not designed by them or with people like them in mind. We see this structural inequality quite starkly in the disproportionate impact of COVID-19 on Black communities in the US (Crenshaw, 2020). The underlying cause in this case is deep-seated, longstanding, institutionally entrenched racism. This racism is not new, and many of these deaths could have been prevented with decision-making, including decisions about data collection and data analysis, guided by goals of equity and justice. These are lessons that we, the authors, have learned from the Black feminist scholars and activists we cite in our book.²

Data feminism asks three questions about power in data science: *Data science for whom? Data science by whom?* and *Data science with whose interests and goals in mind?* When asked of the datasets, models, and other data projects surrounding the novel coronavirus, these questions help to expose the governments, corporations, and other powerful institutions that are currently controlling the terms of data collection, as well the decisions being made (or, in the case of the US, not being made) on the basis of the models presented to them. As an obvious example, we might consider how, early in the pandemic, the US President expressed a desire not to allow a cruise ship with passengers afflicted by COVID-19 to dock. As he said, “I like the numbers being where they are. I don't need to have the numbers double because of one ship that wasn't our fault.” (McFall-Johnsen, 2020).

Other government agencies are implicated in this deliberate undercounting. For 11 weeks from March through May, the US CDC published no data about how many people were being tested for the virus (Meyer and Madrigal, 2020). “All of a sudden those stats vanished,” explained Rep. Mark Pocan, a Democrat from Wisconsin. At a certain point, a new CDC website was quietly released that did track national tests, but based on independent verification by *The Atlantic*, the numbers did

not match those that the states themselves were reporting (Meyer and Madrigal, 2020). Deaths from and cases of COVID-19 are going underreported as well. We know from anecdotal reports that men are dying at higher rates than women, and that Black, Native, and Latinx communities are hit harder (Curtis and Choo, 2020; Lawton, 2020, Watson-Daniels et al., 2020). But states are not reliably tracking sex, gender, or race/ethnicity in COVID-19 cases. Even in cases where states are tracking race/ethnicity, they are still collapsing Native populations into the “Other” category, making it impossible to disaggregate any racialized effects of the disease on Native people (Nagle, 2020). Recognizing how each of these decisions relating to data collection are questions of power can help us begin to address them.

Principle 2: Challenge power

Data feminism commits to challenging unequal power structures and working toward justice

The second principle of data feminism follows from the first, and it is to challenge the unequal distributions of power that we encounter in the world. In other words: it is not enough to unmask and expose unequal power relations in data collection and data science; we must commit to righting and rebalancing the distribution of power. In this chapter of the book, we propose several methods of challenging power in datasets and data projects, as well as for using data science to directly confront corporations and governments. One of these methods involves collecting *counterdata*, especially when our counting institutions are failing to collect the data that may help to quantify the problem at hand.

With respect to the coronavirus, we see this exemplified in the work of Data for Black Lives, and the group's efforts to compile a comprehensive list of datasets relating to the impact of the coronavirus on Black people in each state in the US. “Data is not often collected about Black communities when it's needed the most,” the group writes on its website (Data for Black Lives). Likewise, the only national database reporting Native American tribal affiliation is being published by the newspaper *Indian Country Today* (Nagle, 2020). And journalist Alisha Haridasani Gupta has worked to list out all of the knowledge that is missing because the US is not collecting sex-disaggregated data about COVID-19: Why are men dying at higher rates? Are pregnant people more at risk (as they were for SARS)? Gupta (2020) lists what we don't know and won't know until we have data and analysis to fill in the blanks. While collecting counterdata on its own does not dissipate all inequality, it can be an important part of a comprehensive strategy to hold powerful institutions accountable.

Principle 3: Elevate emotion and embodiment

Data feminism teaches us to value multiple forms of knowledge, including the knowledge that comes from people as living, feeling bodies in the world

The third principle of data feminism, to elevate emotion and embodiment with respect to data, begins by challenging the premise that data communication should always be neutral, rational, and “just the facts, ma’am.” Emotion is often exiled from data science, seen as a suspicious element that may introduce subjectivity into an otherwise “objective” process. Challenging this false binary comes from the work of feminist philosopher Donna Haraway (1988). Her concept of *situated knowledges*—the idea that all knowledge originates at a particular time, in a particular place, and from within a particular set of social and political contexts—helps us to recognize how all knowledge is shaped by the particular perspective of the person or group who produces it.

Valuing emotion as part of a data science project grounds us back in the world and among the bodies from which data are derived. This is particularly important if we aspire to undertake responsible data visualization in a health crisis that is having disproportionate impacts on specific populations. A tick on a death count may be a simple number for some people, but for others it represents the loss of a beloved family, friend, or community member. Because there is so much trauma embedded in these data, information studies scholar Faith Day from the COVID Black project has begun working on a guide for researchers working with COVID-19 data about Black people that is grounded in a feminist ethics of care (COVID Black, 2020). Data alone also fail to recognize the people who are on the frontlines giving care and sustaining society. Here we are inspired by artist Aya Brown’s series of COVID-19 portraits (2020). They honor the caregivers and essential workers—all women of color—looking straight at the viewer, caught in a brief exchange with the viewer before they proceed with their work.

Principle 4: Rethink binaries and hierarchies

Data feminism requires us to challenge the gender binary, along with other systems of counting and classification that perpetuate oppression

The fourth principle of data feminism derives from the false binary that our culture has constructed between the category of “man” and “woman.” There are more

than two genders, of course, and a fundamental commitment of feminist thought is gender equality for all genders. As we write in our book, binaries are often hiding hierarchies. The gender binary is no different—it hides a hierarchy in which men are on top, dominating social institutions from corporate boards to government leadership positions.

False binaries and hidden hierarchies permeate all of our data collection and categorization decisions. For instance, the initial decision not to collect sex-disaggregated data on COVID-19 deaths meant that, for several weeks, the worse outcomes that men experience was not initially known. Interestingly, but not surprisingly, since men tend to fare worse than women, many states and countries have since moved to track sex-disaggregated data on this issue (Global Health 5050, 2020). But they have not paid as much attention to other areas in which gender matters: for instance, the fact that healthcare workers, domestic workers, and other care workers are disproportionately women, and therefore more vulnerable to exposure to the virus (Robertson and Gebeloff, 2020). Grocery workers and shopkeepers, who are also predominantly women, are similarly more exposed. Looking at the situation through the lens of gender also leads to consideration of the gendered effects of policy decisions—how women and children are more at risk for domestic violence as a result of state-issued lockdowns, for example.³

Viewing the coronavirus from outside the gender binary also necessitates a stronger focus on its disproportionate effects on trans and gender nonconforming people. These risks include access to health care, food, and housing security for the most vulnerable populations (like sex workers), and the impact of binary gender-segregated quarantines in places like Panama and Peru. Of course, tracking these data carries with it its own risks and potential for harm—what we term in the book a *paradox of exposure*—because of the increased visibility that collecting data on these populations might bring them. Counting is always complicated, and attending to context, which we discuss more below, is essential when making decisions about what and who to count, and how to do so.

Principle 5: Embrace pluralism

Data feminism insists that the most complete knowledge comes from synthesizing multiple perspectives, with priority given to local, Indigenous, and experiential ways of knowing

The fifth principle of data feminism has to do with the importance of bringing together multiple perspectives in any knowledge-making process. Following thinkers

such as Donna Haraway (1988) and Kim Tallbear (2013), the underlying premise of this principle is that we can gain better, more detailed, more accurate, and ultimately more truthful knowledge if we pool perspectives from a wide range of individuals and groups, especially those who are most directly impacted by the issues at hand. For example, there has been a rush to implement digital contact tracing apps to try to track the spread of the virus. In his detailed breakdown of the risks of Google and Apple's proposed implementations, Ali Alkhatib (2020) states plainly that "digital contact tracing will exclude the poor, children, and myriad other uncounted groups". This excludes precisely those groups who are at the most risk for contracting the disease and whose needs should be prioritized. But even if digital contact tracing were to count those in most need of being counted, there is no guarantee that these apps would not become another form of state surveillance—another line in the long history of the over-surveillance of, for example, Black people in the US (Browne, 2015).

Contact tracing is being implemented in some places in the tried and tested way: hiring people as contact tracers to track down connections after an individual tests positive. But collecting these data in a human way requires a unique focus on already having (or building) relationships of trust with those most impacted. Contact tracing can only succeed if the people being interviewed trust public health representatives, tell the truth, and follow the guidelines for quarantining (Chen, 2020). In April, the state of New York announced its contact tracing program, funded by Bloomberg Philanthropies and administered by Johns Hopkins University, based in Maryland, and Vital Strategies, a global health organization with ties to Bloomberg. New York-based community health groups decried it for failing to build on years of deep ties in Black and Latinx communities in the state. "It's astounding that the state seems unaware that in the past, local community groups, have been phenomenally successful at contact tracing in exceedingly difficult situations, from HIV/AIDS to Hepatitis," said Chris Norwood, Executive Director of Health People based in the Bronx (Health People and Commission on the Public's Health System, 2020). Norwood joined with representatives from other groups to push back on how the New York program is structured and call for a Community Review Board.

This highlights another COVID-19 missed opportunity around data. Instead of embracing pluralism—looking to organizations led by and based in communities, which have the relational infrastructure to be successful—the government went with another familiar route: the elitism and cronyism that preserve the status quo.

Principle 6: Consider context

Data feminism asserts that data are not neutral or objective. They are the products of unequal social relations, and this context is essential for conducting accurate, ethical analysis

The principle of considering context is, in some ways, the most plainly applicable to the coronavirus pandemic. Almost every country affected by the virus has been releasing data on the number of cases, fatality rates, the percentage of the population affected, and so on. Yet each country's data are subject to the particular conditions of their collection. How many tests were conducted in each country, which populations were being sampled, and how truthful were the countries being in reporting their numbers, are only some of the questions being asked in order to understand just how much uncertainty there is in the COVID-19 data. The visualizations of COVID-19 cases at the country level created by the *Financial Times*, which became an early reference point in the pandemic, included annotations that called out unusual testing conditions in certain countries so that viewers understood not to assume too much by comparing one country's curve to another.

But the problems are not only about missing data and uncertain data. Just having precise data does not mean that a problem will be addressed. We must also consider the role of context in interpreting COVID-19 data. For example, the groups Data for Black Lives, COVID Black, and others have emerged to push back against "any use of COVID-19 data to reinforce the narratives about Black people that have made race a risk factor, while blatantly ignoring the central role of racism" (Data for Black Lives, 2020). Understanding the geographic, environmental, and economic conditions that contribute to negative health outcomes experienced in Black communities helps to identify those communities as more at risk, rather than as themselves riskier. As Joia Crear-Perry (2018) has explained with respect to issues of maternal health, "Race Isn't a Risk Factor . . . Racism Is".

Principle 7: Make labor visible

The work of data science, like all work in the world, is the work of many hands. Data feminism makes this labor visible so that it can be recognized and valued

Data science, like all work in the world, depends upon the labor of numerous people—not nearly enough of them valued or even named. In the early days of the coronavirus, when workplaces had been closed but work-from-home conditions had not yet been

established, much of the world was exposed to the human labor that underlies many of our most commonly used platforms (Matsakis and Martineau, 2020). Google search results became less accurate and Facebook posts featured more spam. This was, of course, because there is usually an army of people moderating the Facebook feed, and hand-tweaking Google's search algorithm. With them at home, they could not perform this crucial work. But the lessons of invisible labor extend out from the domain of data to our post-Covid world. Those deemed "essential workers" are, with the exception of well-compensated medical professionals, those who work for minimum wage (or less) at grocery stores and at gas stations, or in private homes as health aids or caregivers. As they were required to report for work, those with more elite and more highly compensated jobs were able to stay safely at home, protecting themselves and their families (Scheiber et al., 2020). If the coronavirus has done one positive thing, it has exposed the labor on which our global economy depends. It is now our job, as scholars and students of data science as it is our job as people in the world, to work to see that labor is properly compensated, valued, and named.

Coda: Organize against oppression

In the conclusion to *Data Feminism*, we take inspiration from the Google Walkout, and other movements from within the tech industry, such as #NoTechforICE, that have used their visibility and value in our data-driven society to advocate for social change. We are not the first to observe that we have an opportunity to emerge from this crisis with a greater sense of the inequality that surrounds us, and a greater commitment to working towards justice. On 1 May, workers celebrated International Workers Day with strikes at Instacart, Amazon, Target, and Whole Foods, calling attention to how the coronavirus had put them on the front lines without hazard pay, sick leave, or any of the workplace protections that should accompany their essential work. How do we build on this momentum?

We drafted this essay in early May 2020. As we write now, in mid-June 2020, to prepare the final copy for submission, the US has been catalyzed into action following the murder of George Floyd by members of the Minneapolis Police Department. Among the conversations that Floyd's death and the subsequent protests have prompted is the recognition that police violence against Black people constitutes another longstanding pandemic, one which demands an equally urgent response. Our thinking about COVID-19 through the lens of intersectional feminism—a body of work created by Black feminists—should prepare us for this moment. As scholars operating in the privileged and

predominantly white world of academia, and as white scholars ourselves, we must listen and learn from and cite the Black scholars and activists who have come before us. They teach us that these pandemics are interconnected and that these twin crises lay bare deep-rooted structural inequalities at the intersection of race, gender, and class. For white academics reading this essay, it is our responsibility to recommit to anti-racist work in our institutions, to decolonize our own teaching and training in order to engage authentically with impacted communities, and to follow where Black womxn lead.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Lauren F. Klein  <https://orcid.org/0000-0002-1511-0910>

Notes

1. For a more detailed articulation of our positionalities as well as our feminism(s), see *Data Feminism* (D'Ignazio and Klein, 2020).
2. Black feminists have long advocated for the idea that sexism is compounded by other structural oppressions, including racism, classism, colonialism, and more. The term "intersectionality," coined by legal scholar Kimberlé Crenshaw, is often used to explain how social inequality cannot be explained by only one form of oppression, such as sexism or racism. But the idea was described by others before her—for example, the Combahee River Collective described systems of oppression as "interlocking" in 1977 (Taylor, 2017). Prior to that, in the nineteenth century, Black women scholars and activists like Anna Julia Cooper, Frances Ellen Watkins Harper, and Sojourner Truth, also described intersectionality in practice if not by name (Nash, 2018). In *Data Feminism*, we trace these links forward, drawing on an intersectional model of power—the matrix of domination—proposed by Patricia Hill Collins (2002) as well as the swell of critical data work from Black women scholars that challenges that matrix of domination as it is enacted in data and artificial intelligence (e.g. Benjamin, 2019; Broussard, 2018; Buolamwini and Gebru, 2018; Noble, 2018).
3. The Dutch Society for Gender and Health has been publishing an exhaustive spreadsheet monthly of all the research on gender and COVID-19 that they have been

able to track down. Additionally, there is an open library on Mendeley collecting interdisciplinary research on COVID-19 and gender.

References

- Alkhatib A (2020) We need to talk about digital contact tracing. Available at: <https://ali-alkhatib.com/blog/digital-contact-tracing> (accessed 15 June 2020).
- Benjamin R (2019) *Race after Technology: Abolitionist Tools for the New Jim Code*. New York, NY: John Wiley & Sons.
- Broussard M (2018) *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.
- Brown A (2020) Essential Workers. Instagram. Available at: <https://www.instagram.com/ayabrown.tiff/> (accessed 15 June 2020).
- Browne S (2015) *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.
- Buolamwini J and Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR 81: 77–91.
- Chen C (2020) You don't need invasive tech for successful contact tracing. Here's how it works. *ProPublica*, 19 May. Available at: <https://www.propublica.org/article/you-dont-need-invasive-tech-for-successful-contact-tracing-heres-how-it-works> (accessed 15 June 2020).
- Collins PH (2002) *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. New York, NY: Routledge.
- “COVID Black” (2020) COVID Black: A taskforce on black health and data. Available at: <https://www.cla.purdue.edu/academic/sis/p/african-american/covid-black/> (accessed 15 June 2020).
- Crear-Perry J (2018) Race isn't a risk factor in maternal health. Racism is. *Rewire News*, 11 April.
- Crenshaw KW (2020) When blackness is a preexisting condition. *The New Republic*, 4 May. Available at: <https://newrepublic.com/article/157537/blackness-preexisting-condition-coronavirus-katrina-disaster-relief> (accessed 15 June 2020).
- Curtis K and Choo E (2020) Indigenous populations: Left behind in the COVID-19 response. *The Lancet* 395 (10239): 1753.
- Data for Black Lives (2020) Action: COVID-19 open data by state. Available at: <http://d4bl.org/action.html> (accessed 15 June 2020).
- D'Ignazio C and Klein LF (2020) *Data Feminism*. Cambridge, MA: MIT Press.
- Global Health 5050 (2020) Sex, gender, and COVID-19: overview and resources. Available at: <https://globalhealth5050.org/covid19/> (accessed 15 June 2020).
- Gupta AH (2020) Does Covid-19 hit women and men differently? U.S. isn't keeping track. *The New York Times*, 3 April.
- Haraway D (1988) Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies* 14(3): 575–599.
- Health People and Commission on the Public's Health System (2020) *Exclusion of community partnerships in Bloomberg and Johns Hopkins-controlled New York COVID-19 contact tracing plan decried* [Press release], 24 April. https://www.newswise.com/coronavirus/ny-state-contact-tracing-plan-dead-end-for-communities/?article_id=730579 (accessed 15 June 2020).
- Lawton G (2020) Why are men more likely to get worse symptoms and die from COVID-19?. *The New Scientist*, 16 April. Available at: <https://www.newscientist.com/article/2240898-why-are-men-more-likely-to-get-worse-symptoms-and-die-from-covid-19/> (accessed 15 June 2020).
- McFall-Johnsen M (2020) Trump said he wants to keep Grand Princess cruise passengers on the ship so that coronavirus numbers don't 'double.' That strategy failed in Japan. *Business Insider*, 6 March.
- Matsakis L and Martineau P (2020) Coronavirus disrupts social media's first line of defense. *WIRED Magazine*, 18 March. Available at: <https://www.wired.com/story/coronavirus-social-media-automated-content-moderation/> (accessed 15 June 2020).
- Meyer R and Madrigal AC (2020) State and Federal Data on COVID-19 Testing Don't Match Up. *The Atlantic*, 17 May.
- Nagle R (2020) Native Americans being left out of US coronavirus data and labelled as 'Other.' *The Guardian*, 24 April.
- Nash JC (2018) *Black Feminism Reimagined: After Intersectionality*. Durham, NC: Duke University Press.
- Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Robertson C and Gebeloff R (2020) How millions of women became the most essential worker in America. *The New York Times*, 18 April.
- Scheiber N and Schwartz N and Hsu T (2020) 'White-Collar Quarantine' over virus spotlights class divide. *The New York Times*, 27 March.
- Tallbear K (2013) *Native American DNA: Tribal Belonging and the False Promise of Genetic Science*. Minneapolis: University of Minnesota Press.
- Taylor KY (ed.) (2017) *How we Get Free: Black Feminism and the Combahee River Collective*. Chicago, IL: Haymarket Books.
- Watson-Daniels J, Milner, Y, Triplett, N., et al. (2020) *Data for Black Lives COVID-19 movement pulse check and roundtable report*. Available at: <http://d4bl.org/reports.html> (accessed 15 June 2020).