CATALYST
feminism, theory, technoscience

Book Review | *Data Feminism* by Catherine D'Ignazio and Lauren F. Klein (MIT Press, 2020)

Becca Rose Glowacki
Goldsmiths, University of London
b.glowacki@gold.ac.uk

Dominant groups have long used information to maintain power. For example, taxonomies, categorizations, and statistics have historically been used to uphold existing power structures. Today, however, data is faster, its scale is massive, and it is pervasive in our everyday lives—which makes Catherine D'Ignazio and Lauren F. Klein's *Data Feminism* so timely. *Data Feminism* starts from the foundation that power is not distributed equally: "Those who wield power are disproportionately elite, straight, white, ablebodied, cisgender men from the Global North" (page 8). Building on this premise, D'Ignazio and Klein demonstrate how imbalances of power are reflected in examples of data biases and injustices discussed throughout the book. Underpinning the text is a series of intersectional feminist concepts that can be used to address and challenge these relations of power.

Data Feminism does not argue against data science; rather, it offers methodologies, examples, and principles to reimagine and illustrate how data science could advocate for justice. The chapters of the book are organized around seven guiding principles: examine power, challenge power, elevate emotion and embodiment, rethink binaries and hierarchies, embrace pluralism, consider context, and make labor visible. The authors discuss and develop each principle by reflecting on related theories and inspiring practices of data justice. As a whole, *Data Feminism* provides a methodological guide for developing more socially just

Głowacki, Becca Rose. 2021. Review of *Data Feminism*, by Catherine D'Ignazio and Lauren F. Klein (MIT Press, 2020). *Catalyst: Feminism, Theory, Technoscience* 7 (1): 1–5.
http://www.catalystjournal.org | ISSN: 2380-3312

data through its critiques of data practices, theoretical discussions, and practical guidelines on ways to implement change.

In Chapter 1, the authors explain how data feminism, as a methodological approach to interrogating data, is guided by insights from intersectional feminisms. Core to this is Patricia Hill Collins's "matrix of domination," which she first developed in *Black Feminist Thought* in 1990. The matrix of domination is a useful analytic for discussing how power unfolds in and around data through four social domains: structural (law and policies), disciplinary (implementation of law and policy), hegemonic (culture and media), and interpersonal (individual experiences). Attending to power, the authors illustrate who is represented in computing and data analyst jobs (spoiler alert: not many women at all, and very few Black or Latinx women). Thus, people in the data science sector (those who design data systems and artificial intelligence models) have limited lived experience of sexism, racism, and other forms of discrimination, which partially explains how these modes of inequality become hard-coded into digital infrastructures. Illustrating and challenging these relations of power, the authors discuss scholar and activist Joy Buolamwini's research on underrepresentation of Black women in data science, and the problem of gender and skin-type bias within facial recognition software and its associated datasets. Mariá Salguero's efforts to collect missing data on femicides in Mexico through crowdsourcing also simultaneously reveals and works to fill a vacuum in official government records.

Chapter 2 details how the underlying power relations related to data can be further challenged through a commitment to working towards social justice. Data feminism works towards dismantling structural power in data through a co-liberation approach of mutual benefit that understands that the liberation of minoritized groups is also bound up with those in power. The authors contrast this with a "data ethics" approach that is commendable for addressing bias and the need for transparency and accountability but remains inattentive to unsettling relations of power. For example, ProPublica's research into recidivism risk assessment algorithms shows how racism is embedded in the survey questions used to collect data, thus demonstrating the pitfalls of data as evidence for criminal justice decision making and how it limits reform efforts that would benefit all. The chapter ends with examples of data justice and how power is challenged through teaching young people data feminism (rather than data science). The Local Lotto project, a mapping project by young people from New York, exemplifies this through including learners' lived experiences into statistical analysis.

Chapter 3 discusses how data feminism values and includes multiple forms of knowledge. The authors argue that knowledge comes from situated experience and must include feeling, emotion, and embodiment. Feminist objectivity is used as a framework for understanding how data is situated, which includes how it is shaped by culture, history, and geographic location. Critical of the "god trick"—a Zeus-like perspective of data from above— the authors argue for situated data, based on Donna Haraway's concept of situated knowledges, to account for the possibility of multiple perspectives in data. To illustrate this point, the authors introduce work by Margert W. Pearce, who created a map using Indigenous place names. This map produces data that recognizes Canadian cartography and re-envisages geography in terms of Indigenous Peoples' languages and knowledge, including what they're not willing to make public.

The fourth chapter explores ways to reimagine binaries and hierarchies. *Data Feminism* questions systems of classification. The basis of many binary modes of categorization stem from computing system data types such as a Boolean (which are either true or false). For example, when Facebook requests information from users based upon a binary poll where the choice is only "male" or "female," it is imbued with assumptions on gender and sex. Many genders are not counted and rendered invisible. Thus, the datasets that are produced from such binaries exclude multiple lived experiences. *Data Feminism* continues to outline its key principles in Chapter 5 through a discussion of the need to embrace pluralism. It insists that complete knowledge comes from the synthesis of many voices, with priority given to local, Indigenous, and experiential ways of knowing, and pluralism can be worked towards through actively and deliberately invite multiple perspectives into data analysis and storytelling processes. The authors distinguish between the practices of collecting data "for good," where projects are undertaken in the public interest, with data "for co-liberation" (page 140). It is not enough to make a socially engaged data project. At their core, co-liberation practices are based on partnership and collaboration, and the goal is to develop pluralistic outcomes with data for and by communities themselves. Co-liberation data projects tell local stories through data murals, such as in the work of Groundwork Sommerville, a nonprofit organization in Massachusetts, or support Indigenous rights through developing data-driven cases for land rights, as with the work of Digital Democracy.

Chapter 6 builds on ideas of pluralism by asserting that data is not neutral or objective, and that attention must be given to the context in which data are produced. For example, the Global Database of Events, Language, and Tone

(GDELT), which collects data from news and events, can easily be wrongly interpreted by journalists and researchers. Large sets of raw data are misleading because they lack context or related stories such as political events that shape how and why the data is produced. D'Ignazio and Klein coin the term "Big Dick Data," a tongue-in-cheek academic term to describe the fetishization of data and a corresponding inflated sense of the technical importance of having more data. The numbers do not speak for themselves, and data feminism strives to communicate inequalities that are represented by data though restoring context.

Chapter 7 explores how data feminism values and recognizes the work that goes into data science through making labor visible. The chapter starts with a discussion of invisible labor and the "wages for housework" movement of the 1970s, while arguing for the need to consider all the labor that goes unseen in data systems. The concept of making labor visible extends into the fabric of the book, where the authors have attempted to show the different kinds of labor that has gone into its making. They do this through embedded detailed attribution throughout and a section acknowledging the work of community organizations by distributing royalties to several of these groups.

In many ways, *Data Feminism* is praxis. It is not just making labor visible that is embodied in the text, but the other principles of data feminism are acted out as well. The authors examine and challenge power through reflecting on potential bias in the book itself. For example, they include an appendix chapter that describes their values and metrics for how they hold themselves accountable to the principles of data feminism, which includes their citational practice. D'Ignazio and Klein demonstrate accountability by punctuating the text with their voices, motivations, and self-reflexivity. The authors also released an early draft of the book for open editing in a public online forum (a reflection on the changes during this process is included towards the end of the book). The intersectional feminist methodology is both exemplified *and* practiced in the book. Theories that underpin the book's methodologies, such as Haraway's situated knowledges, or Collins's matrix of domination, are put into action through collaboration with community organizations and the open dialogical approach in the production of the text.

The book gives an overview of many feminist concepts and data practices, and it provides an introduction to ideas such as bias in machine learning, community-led data practices, and research into surveillance capitalism. It is a useful text for data scholars new to intersectional feminist methodology, and could work well as a

course book for both undergraduate or master's degree students. But also, it is an excellent resource for feminist STS scholars, full of inspiring projects and references to explore further.

References

Collins, Patricia Hill. (1990) 2009. *Black Feminist Thought: Knowledge, Consciousness, and Political Empowerment*. New York: Routledge.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspectives." *Feminist Studies* 14, no. 3 (Autumn): 575–99.

Author Bio

Becca Rose Glowacki is a doctoral researcher at Goldsmiths, University of London. Her research and practice sit at the intersection of critical design, computing, and pedagogy. Becca is an associate artist at Knowle West Media Centre, a resident of the Pervasive Media Studio, and co-curator of hybrid art and technology festival, Control Shift.