Book Review

Data Feminism by Catherine D'Ignazio & Lauren F . Klein (The MIT Press)

Reviewed by

Rafaan Daliri-Ngametua
The University of Queensland, Brisbane, Australia

Madeline Good
University of Missouri, Columbia

The contribution of *Data Feminism* by Catherine D'Ignazio and Lauren Klein (2020) to the contemporary data science landscape is a timely invitation to imagine the possibilities of data science and artificial intelligence informed by intersectional feminist thought. In an impassioned call to action, D'Ignazio and Klein challenge commonplace approaches to data design and utilization. They expose what they consider is the predominantly white, male and "technoheroic" undercurrent of data (D'Ignazio & Klein, 2020, p. 9). In the current political and social climate, where divisive rhetoric and harmful policies continue to support the privileged and powerful to further marginalize people of color, women and non-binary people, the dismantling of such systemic oppression in and with the data science field becomes a key tenet of the activist work the authors are calling on their readers to ignite. The arguments in this book are guided by three main questions: Data science by whom? Data science for whom? Data science with whose interests in mind? Applying the theoretical concept of intersectionality along with Patricia Hill Collins' (1990) matrix of domination, power is challenged and reimagined as these questions are deconstructed and addressed. D'Ignazio and Klein's incorporation of local and grassroots cases of data collection and modeling provide concrete yet emotive examples for how data science could be transformed and how marginalized peoples have been and continue to utilize data to challenge power in their own communities.

In what follows, we review the seven core principles that D'Ignazio and Klein have conceptualized through their engagement with the foundations of intersectional feminist thinking. These principles serve as both suppositions and interrogations of what they view as the oppressive, sexist and racist systems entangled between the interplay of data and justice.

D'Ignazio and Klein aptly begin with issues of power and how it influences data science in commonly ignored or unidentified ways. Here, they introduce their main theoretical framework, Patricia Hill Collins' (1990) matrix of domination, to provide a lens through which to understand how power works across structural, hegemonic, disciplinary, and interpersonal domains. Importantly, this early chapter utilizes the matrix of domination to discuss who typically gets to 'do' data science, who data science is usually for, and whose interests are prioritized when data science is implemented. Data science is described as an "extractive system" dominated by elite white men and without the participatory processes expected in neutral and democratic work (D'Ignazio & Klein, 2020, p. 45). It is this interrogation of the asymmetrical distribution of power between those who extract the data and those whose data is extracted from that positions the first principle *Examine Power* as fundamental in understanding and challenging systems of oppression within data science.

D'Ignazio and Klein note that traditionally, data science as a field, prioritizes power-securing concepts such as ethics, bias, fairness, and accountability over power-challenging concepts like justice, oppression, equity, and co-liberation. To address this concern, "counterdata" is introduced as a concept that can liberate oppressed populations through the quantification and visualisation of structural oppression. Examples of power-challenging projects in this book such as the Detroit Geographic Expedition and Institute and those implemented by individuals such as Laurie Rubel, a mathematics teacher from New York City, demonstrate tangible liberatory approaches to data science created within and for marginalized communities. The second principle *Challenge Power* is thus perhaps less a principle and more of a call to recognize and commit to data justice that centers co-liberation and equity. Pushing back against institutional inequity and oppression through counterdata, can expose the nature of embedded race in the algorithms and data of systems and technologies that further perpetuate disadvantage under the "allure of objectivity" (Benjamin, 2019, p. 42).

D'Ignazio and Klein continue by problematizing the common view that plain and unemotional data representations, like bar graphs, are "neutral" and "objective." They argue that these traditional

approaches of data visualization are largely persuasive and far from impartial. The authors challenge what they see as a false binary between reason and emotion, urging data scientists to see the power behind incorporating embodiment and uncertainty within their data visualizations. The authors draw on an example from the data visualization firm, Periscopic, which conducted a project that visualizes gun deaths in the United States in 2013. This haunting representation captures not only the number of gun deaths (11,419) but also how that figure equates to "stolen years" (502,025). D'Ignazio and Klein recognise that this work is framed around the emotion of "loss," ensuing the reality of death be felt and seen as stolen time from those lives lost and from those left behind who continue to suffer from that loss. They argue that "[…]emotion and affect, embodiment and expression, embellishment and decoration" are often what is excluded in traditional visualisations of data (p. 96). This exclusion is noted as a repudiation of aspects of the human experience, experiences that are often associated with women. The third principle *Elevate Emotion and Embodiment* encourages us to re-examine the false assumption that data can be from "no body" (p. 95).

At this thought-provoking juncture, the authors turn to *binaries* and *hierarchies* to demonstrate that, "what gets counted counts" (D'Ignazio & Klein, 2020, p. 97). They discuss how using data for classification of categories inherently includes, excludes, labels, and mislabels those who are being classified. They also critique gender classification and its intersection with race and class. The examples selected by the authors illustrate clearly how data categorization can be made more inclusive and representative of the people behind the categories. The fourth principle *Rethink Binaries and Hierarchies* demonstrates how counting and measuring can be used to hold to account power and the powerful, rather than remain as tools of oppression.

Building on inclusivity and representation in data categorisations, D'Ignazio and Klein argue for the emancipation of suppressed voices in all stages of data work. They claim that through valuing and foregrounding multiple perspectives, we can paint a more complex picture and resist the temptation to rely on one "loud" and "technical" voice. Inspired by the work of Donna Haraway, they position knowledge as only ever partial stance, to bust the myth of a "capital T Truth" (D'Ignazio & Klein, 2020, p. 136). D'Ignazio and Klein theorize that the space where transparency and reflexivity meet opposes ideas of objectivity, allowing for deeper and more robust understandings of the world. The authors assure that problematizing objectivity is not an anti-science stance; it simply allows for multiple voices to make knowledge, to produce situated and embodied positions and standpoints. Their fifth principle *Embrace Pluralism* encourages data scientists to intentionally and actively seek multiple perspectives, including the voices of those most marginalized, so that "enduring and asymmetrical" power relations can be dismantled, and the transformative power of data science can work to co-liberate all (p. 141).

D'Ignazio and Klein continue by problematizing any notion of "raw" data. They assert that the numbers do not and should not speak for themselves, engaging feminist standpoints to urge the necessity to place data within the contexts in which they were produced. In this way, the potential for power and privilege to obscure the truth is mitigated. A key tenet here is that historical, political and social conditions and circumstances cannot be stripped so that decontextualised "raw" data remains. This means, the data is already "cooked" when it enters a research project (D'Ignazio & Klein, 2020, p. 159). D'Ignazio and Klein argue that their 6th principle of *Consider Context* will allow those who engage with data to take a much deeper approach, one that considers culture, context and the many nuances that are inescapable or rather inextricable from the data itself. In this way, the colonial influences of racism and sexism that enter data can be recognised as digitally reinforced oppression (Noble, 2018). "Letting the numbers speak for themselves" is viewed as unethical and undemocratic, with potential to do more harm than good by reinforcing the unjust status quo (p. 159).

The seventh and final principle *Make Labor Visible* contextualizes the many points of uncomfortable reflection throughout the book as the "[…]true cost and planetary consequences of data work" (D'Ignazio & Klein, 2020, p. 201). This is an appropriate final principle as it advocates for the recognition of the "many hands" that work to deliver data to end users (p. 176). The discussion here involves a recognition of the injustice and arduous labor involved with data science in a capitalist society, work that is often "invisible and uncredited" (p.178). D'Ignazio and Klein highlight historical examples of invisible labor and slavery and liken them to current practices of outsourcing and crowd sourcing data work to countries outside of the Unites States and to women and people of color who work for less than minimum wage in countries with less opportunities. Through recognizing the oppressive nature of such practices, the authors advocate for acknowledgement, value and credit to be attributed to the vast network of people who contribute to data science work in all stages of production. Drawing again on the feminist practices of giving credit to a broader range of work, D'Ignazio and Klein advocate for resistance against the certain *types* of people who take up space, "screening out others" (p. 185). Through seeking to make visible the work involved in the full life cycle of any data project, the invisible labor of data science that encourages "[…]the exploitation of Black and Brown bodies so that white bodies can thrive" can be brought to account (p. 184).

D'Ignazio and Klein recognize that there are simultaneous and multiple starting points in the

contestation of oppression in data science and make note of the varied groups from community, research and business fields that are transforming their work as a means of resistance to oppression in all its forms. This recognition is proceeded by a call for the mobilization of efforts, a chorus of voices and an assembly of links between the emerging forces that promote the use of data to expose systemic bias and oppression. *Data Feminism* brings important and much needed discussions about concepts such as power, race, class, and gender into the field of data science. However, there were moments where further theoretical and methodological depth may have bolstered the arguments presented. Scholars such as Jennifer C. Nash, Anna Julia Cooper, and bell hooks are cited when intersectionality is drawn upon, and Patricia Hill Collins' matrix of domination is foregrounded, but no further contextualization is provided. This potentially wades into the murky water of "flattening" and commoditizing Black women's work, stripping it of situational history and undermining the purpose behind the creation of such concepts in the first place (Alexander-Floyd, 2012, p. 5; duCille, 1994). Additionally, D'Ignazio and Klein do not explain how this scholarship on which they drew was integrated to create the seven principles of *Data Feminism*, methodologically speaking. They note that the principles emerged from "the foundation of intersectional feminist thought" yet little elaboration was provided henceforth (D'Ignazio & Klein, 2020, p. 17). Thus, incorporating more background on the theories themselves and offering further insight into the process of creating the framework for *Data Feminism* would further legitimize the priority of transparency throughout the book as well as provide a framework for those who want to feminize other fields in a similar fashion.

In the end, D'Ignazio and Klein encourage us to take action in ways that they may not have included or even imagined. Whether you are an activist, academic, artist, or community organizer, challenging oppression may quite legitimately look completely different. As D'Ignazio and Klein urge, effective and sustaining resistance within data science requires a collective effort or rather, a *multiplication.* It is important to note how while doing this work, D'Ignazio and Klein attempted to use data to hold themselves accountable, acknowledging the harm that well-meaning white folks can cause when trying to help with the effort to make change. Reflexivity is key in making progress, no matter your positionality, but especially when you have the potential to engage in ways that perpetuate harm. Thus, the authors urge us to move forward, creatively, inclusively, and reflexively. In this way, we can build upon *Data Feminism* by thinking about the ways in which we can transform data science through centring narratives, voices and ideas from marginalized communities, and especially from women of color (Alexander-Floyd, 2012). Ultimately, Data Feminism stands out in a field

positioned as unbiased and neutral by carving out space for the incorporation of justice and liberation, challenging readers to see how data is much more than numbers.

References

Alexander-Floyd, N. G. (2012). Disappearing acts: Reclaiming intersectionality in the social sciences in a post-Black feminist era. *Feminist Formations*, 24(1), 1-25.

Benjamin, R. (2019). Race after Technology: Abolitionist Tools for the New Jim Code. Medford, MA: Polity.

D'Ignazio, C., & Klein, L. F. (2020). Data feminism. Cambridge, MA: MIT Press.

duCille, A. (1994). The occult of true Black womanhood: Critical demeanour and Black feminist studies. *Signs*, 19(3), 591-629.

Noble, S.U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press.