Data Conscience: Algorithmic Siege on our Humanity

By Dr. Brandeis Hill Marshall

With a foreword by Dr. Timnit Gebru

Wiley September 21, 2022
Paperback: 352 pages

Inclusive Actionable Data Ethics
For Practitioners
A Book Review

By
Alyssa Jones
Stanford University

And
Alexis Mack
Stanford University

Among the increasing number of AI Ethics conversations, leading computer scientist and data equity strategist, Dr. Brandeis Hill Marshall's important new book, *Data Conscience: Algorithmic Siege on our Humanity* offers a unique perspective: actionable, how-to strategies for engineers and AI ethicists for mitigating AI harms. Unlike many ethics conversations in tech, which tend either toward heavily theorical or technical jargon, Marshall addresses her readers with an accessible, welcoming style articulating both the problems of social impact and the pitfalls of building algorithms. In a clear, optimistic tone, she motivates her readers to listen and learn more. Most importantly, she tells them they belong. Non-technical beginners will find it easy to join the discussion, more advanced audiences to rethink their assumptions about the potentials of algorithms, and everyone will become more aware of the necessity of including marginalized groups in the end-to-end algorithmic development process.

Marshall's overview of the social impact of algorithms combines Black history, civil rights history, and women's history, with engineering history, education as well as an overview of the full data lifecycle. She simultaneously instructs engineers in questions of social justice and ethicists in the need for more practical solutions. Indeed, Marshall has little time for superficial gestures of virtue signaling, whether corporate or academic. She asserts provocatively at the opening of Chapter 2: Morality: "ethics of any kind is of limited utility. Understanding the significance and influence of how morality shows up (or doesn't) in tech innovations isn't a new phenomenon" (p. 27). For too long, ethics has failed to "show up" for marginalized communities most disproportionately harmed by tech. She's not against all ethics, however, just mindful of the historic failure of formal ethics to address real problems of exclusion and harm. Instead, *Data Conscience* helps readers uncover substantive approaches to actionable ethics. Reading this book, ethicists will learn about the concrete problems of data and engineers will reflect more critically on their work. All audiences will clearly understand that the Silicon Valley ethos of "move fast and break things" has proven itself privileged, corrosive, and empty. A slower, more thoughtful, self-critical approach to technology demonstrates a greater

capacity to serve humanity. Most of all, moving away from what she calls “monolithic tech innovation” (p. 237) requires centering data in all thinking and strategy, because its misuse leads to far reaching harm.

Aimed at both junior-level to senior-level software developers, *Data Conscience* elaborates a series of ethical challenges for algorithms, teaching the importance of transparency, encouraging computational thinking in practice, offering strategies for encouraging accountability in tech, methods to avoid double-edged data visualization—potential misinformation in an attractive package—and schemes for governing data structures with law and algorithms, and deciding what helps or hinders those most impacted. With clear direction for developers about how to gauge the fit of their data sets for their intended use, Marshall emphasizes the importance of documentation or even the imperative to “halt everything until you’ve reached non-flimsy, non-hand-wavy responses,” (p. 195). Knowing well that such flimsy reasoning and hand waving is the dominant business model of the tech industry, she asks her readers to practice caution as she demonstrates ways to address problems, including social network analysis and linear regression.

Throughout all her first-rate computing ethics computing pedagogy, Marshall offers a uniquely Black perspective. When teaching her audience how algorithms work and the need to work through the most tedious detail of instruction, her examples are Black. Teaching an algorithm how one washes hair? Marshall delineates a step-by-step method for washing her locs, including a concept map for visualization (p. 91).

Centering the Black experience in programming goes a long way toward rethinking tech development. Marshall’s approach tells Black women “…we do fit and helps us develop toolkits for responsible computing,” (p. xix) affirms Dr. Timnit Gebru in her foreword. Helping readers see that humanity need not be reduced to data points for surveillance and profit, Marshall charts a path for preserving humanity and personal identity.

Marshall understands the structural racism and industry culture that makes it difficult for people of color and gender minorities to ask the 12 questions she raises in chapters 5-8. Asking *why* or *how* data is being collected requires standing up to tech industry claims that collection is somehow “to make things ‘better’ for us,” (p. 120) when we know in fact our communities are being disproportionately harmed by data collection. With a critical eye toward the *how* question, Marshall demystifies the whole idea that machine models actually “learn” anything in the data analysis cycle. The actual learning is done by data laborers who perform the data analysis so that models can be trained and fine-tuned (p. 166). To avoid what she calls the “circus of misguided analysis,” data analysts must ask a series of how questions: *How* does one’s team select, verify, and evaluate data, each of which also have important follow-up procedures to avoid misusing the data in a misguided fashion (p. 167-8).

Marshall is strongest in her methods of ensuring responsibility and promoting what she calls “data civics.” For her, there should be legal recourse to prohibit harmful algorithmics with real consequences that would threaten a company’s reputation, profit, and survival. In her vision of all the stages of data civics, from you and your data, to your home, workplace, community and society, one must choose to be proactive and invite others to actions so that we can build a world different from the corporate tech monoculture where power remains in the hands of the few, to a “multiracial, multigender, multiclass cooperative society where its people participate and influence its direction,” (p. 283).

To us students, we think of Dr. Marshall’s *Data Conscience* as the BIBLE for becoming the developers and data scientists we hope to be and building a future we hope to see. AI Ethics courses need this book. Get it, read it, learn.