Book Review

Broussard, Meredith (2023). *More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech.* The MIT Press.

Haerin Shin

School of Media & Communication, Korea University, Korea

Corresponding to
Haerin Shin
School of Media & Communication, Korea University, Media Hall 412, Anamro 145, Seungbuk-gu, Seoul, Republic of Korea
Email: helenshin@stanford.kr

Expanding on her previous book *Artificial Unintelligence* (2018) with specific focus on how algorithmic decision-making mechanisms perpetuate discriminatory practices across the varied fault lines of identificatory markers, data journalism scholar Meredith Broussard offers an extensive exploration of machine bias in her new book *More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech* (2023). Along with works such as Cathy O'Neil's *Weapons of Math Destruction* (2016), Safiya Umoja Noble's *Algorithms of Oppression* (2018), Ruha Benjamin's *Race after Technology* (2019), and Wendy Hui-Kyong Chon's *Discriminating Data* (2021), this book is noteworthy for not only its intersectional approach to but also authorial standing in the burgeoning field of artificial intelligence (AI) ethics. Deftly linking numerous instances of algorithmically perpetrated disenfranchisement with her own experiences of identificatory parameters – as a black woman working in computer science – serving as grounds for alienation, Broussard covers a range of topics that culminate toward a searing critique of algorithmic *glitches*, the term's allusion to atypicality belying the systemic nature of its mechanism.

Comprising eleven chapters, *More than a Glitch* zeroes-in on some of the most salient areas wherein algorithmic bias makes tangibly adverse impacts on racialized, sexualized, and thereby deemed nonnormative bodies, explaining how each of these categories inform and create one another. In the introduction, Broussard opens up with a personal anecdote involving her childhood memory of splitting cookies in one-halves with her sibling. Said "splitting" could never result in physically identical halves, she asserts, necessitating alternative measures of reparation for the sibling who ends up with the smaller half. Using the story to emphasize the disparity between mathematical and social fairness, Broussard dedicates the following chapter to explicating the structural workings of machine learning mechanisms, helping readers

grasp the nature of algorithms as precision-based, procedural, goal-oriented formulae. This expositional section, along with a chapter-long elaboration on the journey she personally undertook to understand the operative principles behind breast cancer detection systems, is one of the most prominent aspects of the book that sets it apart from the slew of academic writings that have been gaining traction of late. As narrow AI systems that identify and employ rules of their own making, contemporary machine learning algorithms are like black boxes, the inner workings of which remain mostly inaccessible to even their designers and operators. And because they are neither understanding (harbor intent) nor fully understandable, Broussard claims, the outputs that they generate can reinforce and even newly create troubling realities, without being understood as such.

Chapters 3 and 4, which deal with the topic of predictive policing and surveillance, demonstrate how the mathematical nature of algorithms render them self-contradictory (or in other words, imprecise) and result in "glitches" when trained on skewed datasets that result from default human assumptions. Chapter 4, in particular, illustrates the disturbing ramifications of what she calls "precognitive" surveillance practices that have so far received relatively less spotlight compared to the seemingly more tangible ills of predictive policing, such as wrongful arrests or neighborhood patrolling. The fate of Robert McDaniel, who was repeatedly shot by his own community members because the Chicago police force identified and subsequently monitored him as someone with a high probability of being involved in a shooting, serves as a chilling reminder that algorithmic systems are actively making life and death decisions concerning our welfare with godlike impunity.

Chapter 5 examines yet another danger that is nascent within predictive systems by zooming in on their application in the education sector. The COVID-19 pandemic has further aggravated the digital divide to alienate those with limited access to technological resources, with the primary mode of social interface migrating to online spaces and other forms of mechanical mediation – this much is a well-known fact. What remains less discussed however, this chapter tells us, is how misguided implementations of machine learning systems can block the possibility of access itself, defeating the purpose of education on a more fundamental level. The example she introduces concerns students from historically underserved communities who were given, by a machine learning system, International Baccalaureate (IB) scores that they were expected to achieve rather than ones they actually earned. Chapter 6 further expands on the topic of digital accessibility by broadening its reach to include the designs of the infrastructure for and substance of information. The troubling story of Richard Dohan, whose deafness became equated to incompetence at Apple's noisy and ableist retail environment, hits ever close to home. The recent proliferation of automated kiosks in South Korea, for instance, has become a major access hurdle for those who find the machines' interfaces less than intuitive.

Chapters 8 and 9 shows how ableism is not only about physical prowess or constitution, but also the product of biased defaults that preclude minority bodies from even being recognized as such. Broussard's own experience with the messiness of medical diagnostic algorithms, combined with examples of their incompetence in detecting cancerous lesions on darker skins, showcase the manner in which minority bodies are alienated recursively, and reflexively – first in the eyes of the social system across human and nonhuman players, and then also from their own selves.

Coming together under the book's catchphrase "more than a glitch," these chapters uniformly critique the underlying assumption that discriminatory practices are anomalies within the grand scheme of things, in which the human and nonhuman constituents are both implicated.

Hence, she proposes, the need to recalibrate not only the machinic but also human system with tangible measures of intervention, dedicating the last two chapters (10 and 11) to list guidelines to this end. Calling for algorithmic auditing and accountability, Broussard also raises awareness about other principles to be implemented by the government, corporate entities, and users, such as AI alignment, explainability, and transparency. From predictive policing (Ch. 3) to preemptive surveillance (Ch. 4), educational inequity (Ch. 5), ableist product designs (Ch. 6), exclusionary database politics (Ch. 7), and colorblind medical diagnostic systems (Ch. 8, 9), Broussard engages with a wide spectrum of problems that ultimately converge upon three keywords, which respectively explain the techno-social fabrics of bias: *technochauvinism*, the microaggressive (cumulative) agency of *implicit bias*, and *glitch*.

Her articulation of technochauvinism, which may be otherwise described as the illusion of technological neutrality, effectively addresses how misguided faith in the objectivity of science-backed technology in fact originates from social inculcation, implicitly valorizing the authority of the predominantly male, white, and ableist demographic behind the systems' design. The discursive implication of technochauvinism harkens back to the Cartesian legacy wherethrough the epistemological dictate of "cogito ergo sum" becomes a differential ontology of anthropo-supremist investment in intelligence, which has proven to easily translate into racist, sexist, gendered, and ableist hierarchies.

Building on her in-depth examination of the machine learning mechanisms in the beginning portion of the book, meanwhile, Broussard also notes how neither the systems nor the people who design and/or deploy them seldom harbor explicit intent to discriminate, explaining how the absence of willful intent is not a ground for moral exoneration but rather the cause and effect of algorithmic bias. Combined with the "under-the-hood" look into the workings of the black box-like

innards of machine learning, Broussard's emphasis on the absence of malicious intent – whether it be human or nonhuman – is a factor that renders this book distinctively edifying for lay and academic readers alike. The spectacular introduction and subsequent popularization of chatGPT is a case in point; while Broussard does not address large language models explicitly, their impact in the contemporary mediascape positions her argument in a more relatable light. The friendly (and even servile) disposition of chatGPT has helped to dispel fears of an anthropomorphized AGI (artificial general intelligence) scheming to overtake the world, largely thanks to the interventive measures that Open AI implemented. The chatbot's refusal to share its own thoughts or emotions based on its lack of intentional agency, however, does not mean it does not have intentional patiency – namely, the ability to reflect and in turn generate intent from and for its human counterparts. Broussard's observation that those who blindly trust in and thereby execute the verdicts of predictive algorithmic systems, their examples aplenty, acquires a doubly sinister note when combined with her point on technochauvinism as a deeply ingrained prejudice that eventually becomes the default principle of cognition and perception within the social fabric.

Establishing how biased systems reproduce microaggressive practices by way of macro-expansion, Broussard makes yet another key intervention by pointing out how technochauvinism not only advocates but also shapes *ableist* perspectives by way of racialized and gendered parameters. Her technical exposition of faulty algorithmic diagnostic systems in areas such as cancer detection and treatment in contemporary medicine, their roots pointing all the way back to the disturbing premise behind the Tuskegee experiment in which black bodies are deemed most fitting to host debilitating disease, shows that the physical particularities of minority physique is categorized as nonnormative, and thereby pathologized. The examples she introduces

expose the racist and gendered metrics that govern the default baselines of ableism, which also become apparent in her exploration of "beauty concerns." However intuitively appealing Masahiro Mori's widely known theory of the uncanny valley may be, its mapping of an able-bodied, normative human being at the apex of its topographical chart demonstrates that privilege operates as at once a default condition of being and the state of ultimate desirability. In like (or rather, inverted) fashion, medical diagnostic algorithms' failure to detect cancerous lesions in bodies that do not comply with the default parameters in the training dataset points to the disconcerting reality of our algorithmic mediascape in which ontological hierarchy is not only reflected but also reproduced, in scale.

Drawing on Noble and Benjamin's characterization of biased algorithms as "glitches," Broussard goes on to assert that such errors are not matters of mere inconvenience but a glaring instantiation of how the baseline settings of machine-learning systems are profoundly dysfunctional in and of themselves. Here, Broussard's choice to frame the myriad exemplars that she introduces as "glitches" over other terms that signify dysfunction, such as "bugs" or "errors," is of significant note. Whereas a bug pertains more to a fundamental problem that undercuts the system's performativity on the level of its design, a glitch is more of an emergent phenomenon that arises from the system's interaction with factors other than the source code itself (e.g. the training datasets, hardware, user behavior, etc.). Deep-diving into a number of real-life incidences wherein predictive systems perpetrate identity-based profiling, Broussard shows that machine learning algorithms – as modular and allopoietic mechanisms – neither exist nor work on their own, producing harmful intent even without having such views baked into their make.

While largely addressing instances in the U.S., this book undoubtedly a useful resource for Korean readers and scholars, whose positionality within the global mediascape demands a constant engagement with the broader strokes of ethno-racial discourses. As an imagined cultural community that has long suffered from the Orientalist gaze, and increasingly subject to seemingly complimentary but essentially marginalizing stereotypes such as the model minority myth and techno-Orientalist portrayals in popular culture, Asian communities are no exception to the grand scheme of the *glitch*. Such stereotypes, after all, were what stoked resurgent fears of yellow peril during the COVID-19 pandemic. Whereas the direct target of xenophobic exclusionism was China and its people in light of the country's ever-growing reach as a economic and cultural powerhouse, South Korea is no exception to the discriminatory dynamics that are becoming part and parcel of the technological fabric of our media infrastructure, to which the conflationary rejection and alienation of Asian people throughout the pandemic amply attests. As an aspiring force of cutting-edge technical innovations in the domain of artificial intelligence and attendant forms of new media, Korea is also in dire need of an ethical reckoning in order to avoid retracing the steps that big-tech pioneers have taken in the course of their advancements. The scandal involving ScatterLab and Kakao's chatbot Luda Lee, who (or which) ended up reproducing a domestic version of Microsoft's Tay's deterioration into a misogynist, racist, and xenophobic bigot in less than 48 hours since its grand debut, is but one among many a cautionary tale in South Korea alone. Thus deserving credit for the intersectional applicability of its coverage and timely directionality, this book is highly recommendable as a helpful text for a broad range of readers given the wealth of real-life examples and technical expositions, the eloquent yet accessible prose, and fluent engagement with existent bodies of scholarship. Moreover, the author's emphasis on the haunting presence of technochauvinism, the microaggressive effect of implicit social bias (absence of intent), and the

paradoxical nature of algorithmic glitches serve as three faculties that map the volume onto the field of AI Ethics studies with distinction.

REFERENCES

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new jim code*. Polity.

Broussard, M. (2023). *More than a glitch: Confronting race, gender, and ability bias in tech*. The MIT Press.

Chon, W. H.-K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition*. The MIT Press.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.