

BOOK

More than a Glitch

Confronting Race, Gender, and Ability Bias in Tech

Meredith Broussard | MIT Press © 2023 | 248 pages

Buy the book

getAbstract Rating

★ 9 | Eye Opening | Bold | Inspiring

What You Will Learn

- Why machine bias is a structural issue needing complex solutions.
- How technochauvinism ignores machine bias and social fairness.
- Why AI perpetuates bias in law enforcement and medical diagnostics.
- How algorithmic auditing can address bias and promote equity.

Supported Skills

Use AI Ethically | Establish AI Governance | Understand AI Ethical Impact | Understand AI Risks

About the Author

Meredith Broussard is a data journalist and the author of multiple books, including *Artificial Unintelligence: How Computers Misunderstand the World*. She's also an associate professor at New York University's Arthur L. Carter Journalism Institute and a research director at the NYU Alliance for Public Interest Technology.

Recommendation

You may be tempted to dismiss instances of machine bias as “glitches.” However, they’re structural and reflective of real-world racism, sexism and ableism, says data journalism professor Meredith Broussard. Technology should work for everyone – nobody should feel barred from using technology based on their skin color, gender, age or ability. Broussard presents several case studies of machine bias, detailing the harm it’s caused in areas including policing and health care. She urges Big Tech to embrace accountability and work toward the public interest.

Take-Aways

- Machine bias is a structural problem requiring complex solutions.
- Machines “learn” to uphold the status quo and replicate oppressive systems.
- Law enforcement’s faith in biased algorithms has a human cost.
- Big Tech sometimes fails to design for disability and foster inclusive cultures.
- Computer systems often enforce the gender binary, showing bias against LGBTQIA+ people.
- Using AI as a medical diagnostic tool can be unhelpful and even dangerous.
- Embed an algorithmic review process into your business model.
- Build a better world with algorithmic auditing and accountability.

Summary

Machine bias is a structural problem requiring complex solutions.

People assume computers can solve social problems, but this isn’t always true. Machines can only calculate mathematical fairness, which is different from social fairness. The person programming a computer may program it to create a solution mathematically, but that doesn’t mean their algorithms lack bias, leading to neutral decision-making. Programmers are humans who bring their biases, such as those rooted in racism, privilege, self-delusion and greed, to work with them. The belief that technology will solve social problems indicates

"technochauvinism," as it ignores the fact that machine bias exists and that equality often differs from justice or equity.

> *"Digital technology is wonderful and world-changing; it is also racist, sexist and ableist."*

People rarely build biased technology intentionally. Most engineers probably incorrectly assume they're building a "neutral" technology. As an example of machine bias, consider the video of racist soap dispensers that went viral in 2017: A darker-skinned man found that the soap dispenser didn't recognize his hands as human hands, as it only recognized lighter-skinned hands and thus didn't dispense soap for him. When technology shows biases like this, it often occurs because engineers use a homogenous group of test subjects. Machine bias isn't a "glitch" – it's a structural problem demanding complex solutions.

Machines "learn" to uphold the status quo and replicate oppressive systems.

When a machine learns, it doesn't do so as a human would but rather detects and replicates patterns in data. If an algorithmic system is "trained" using data that are reflective of racist policies and actions, it's going to replicate those patterns, maintaining the status quo. For example, if you train an AI model with real data about past recipients of loans in the United States, the model will continue to give Black and Brown people fewer loans, perpetuating the financial services' history of bias.

> *"Let's not trust in it when such trust is unfounded and unexamined. Let's stop ignoring discrimination inside technical systems."*

When giving AI models statistical data, one must consider that the current statistical methodologies people rely on today were developed by the outspoken eugenicists and racists Karl Pearson, Ronald Fisher and Francis Galton. Tukufu Zuberi, a sociologist at the University of Pennsylvania, calls for a "deracialization" of statistical analysis. Zuberi says it's time to replace studies that use racial data without context with those that more fully capture all dimensions of identity and the broader social circumstances of people's lives. In today's racist society, simply saying "I'm not racist" isn't enough. Endeavor to be anti-racist, which entails critically examining your assumptions about race and working to eliminate racist practices when you see them. Embedding anti-racism into machine learning systems means going beyond simply implementing mathematically fair solutions. It requires actively building technologies that challenge the current systems of oppression and end white supremacy.

Law enforcement's faith in biased algorithms has a human cost.

Facial recognition technology (FRT) uses biased algorithms, so when law enforcement relies on the technology, it can result in harm. This is because the technology works better on people with lighter skin tones than on those with darker tones. FRT is also better at recognizing men than women and frequently misgenders nonbinary, trans and gender-nonconforming people. FRT doesn't calculate definite matches and only detects similarities. Yet police use it to make arrests, routinely using it against communities of Black and Brown people the technology fails to work for. Although humans check the accuracy of the matches, this doesn't prevent bias, as humans have biases and can confirm a false match – for example, perhaps they think all people of a certain race look alike. There have been multiple cases of police arresting Black people for crimes they didn't commit when using FRT to identify suspects.

> *"Intelligence-led policing is not data science. It's not scientific. It is just an oppressive strategy that upholds white supremacy."*

Crime-predicting algorithms also show racial bias. For example, the Philadelphia company HunchLab's crime-predicting software maps neighborhoods where it claims crime is most likely to occur, prompting the Philadelphia Police to maintain a greater presence in these areas for public safety. Often, these maps depict communities where Black people live, and increased police patrols can actually make communities less safe, given the prevalence of police brutality in the United States against Black people. HunchLab's maps fail to send patrols to neighborhoods where white-collar financial crimes, such as tax evasion, occur. A mapping project called "White Collar Crime Risk Zones" situates in heavily white areas, such as Manhattan and Wall Street, demonstrating bias regarding the types of crimes police view as worth policing. Police forces across the United States may claim to be doing intelligence-led policing, but they're presenting a false image of objectivity to maintain the status quo.

Big Tech sometimes fails to design for disability and foster inclusive cultures.

People with disabilities need an outlet for their concerns in today's discussions about equity and technology. For one, tech companies could start using the appropriate language when referring to users with disabilities. For example, tech companies may refer to Deaf people as hearing-impaired, but calling someone impaired can make them feel like they're broken when they're just different from those companies perceive as typical users. Richard Dahan, who used to work at a Maryland Apple Store, found that Apple failed to honor its commitments to inclusivity because he had a manager who failed to properly accommodate his needs as a Deaf person (for example, making information accessible via a video interpreter).

Dahan says this manager also allowed customers to refuse to be served by him because he was Deaf. He started the viral movement #AppleToo, highlighting that Apple had some aspects of disability bias within its culture.

> *"Technology is a crucial component of accessibility, but more of what is currently called cutting-edge technology is not necessarily the answer."*

Big Tech often uses images of people with disabilities to promote inclusive technologies when their technologies rarely are accessible. For example, there have been reports of delivery robots making life harder for people with disabilities – for instance, blocking the path of a Blind woman and her guide dog. Yet, "inspiration porn" is rampant, which activist Stella Young refers to as images companies use showcasing people with disabilities to inspire people without disabilities (as if simply existing were an inspirational feat with a disability). Rather than objectify people with disabilities, Elise Roy, a Deaf "human-centered designer," urges people to start designing for disability. Universal design benefits everyone, not just people with disabilities, as you create innovative solutions that make technology more user-friendly, she stresses. Ruha Benjamin, in her book *Race After Technology* takes the conversation a step further, calling for "design justice," meaning that people from marginalized communities should lead design initiatives – design shouldn't be reproducing structural inequalities.

Computer systems often enforce the gender binary, showing bias against LGBTQIA+ people.

Technology often enforces the gender binary, displaying bias toward nonbinary, trans and gender-nonconforming individuals. This is because most computer systems today encode gender as a fixed, rather than changeable, binary value. The way computers encode gender hasn't changed much since 1951 when UNIVAC – a digital computer designed to create databases – gave people only two gender options (M/F). In many ways, this 1950s ideology remains due to "hegemonic cis-normativity and math." It's easier to put people into clean, simple categories than embrace their complex, shifting identities when doing data analysis, so programmers use gender binary and binary representation when coding.

> *"The rigid requirements of 'elegant code' enforce cis-normativity, the assumption that everyone's gender identity is the same as the sex they were assigned at birth."*

Big Tech companies that claim to support people with LGBTQIA+ identities often fail to make algorithms inclusive. For example, trans people using Google Photos often find that the software identifies them as different people before and after their transition. Facebook may have been one of the first social media companies that allowed users to change their gender identity. Still, the system didn't encode nonbinary identities, saving only the identities of male, female and null. Rather than "nullify" those who don't fit into the gender binary, it's time to extensively update computer systems to correct past biases.

Using AI as a medical diagnostic tool can be unhelpful and even dangerous.

Some envision a future in which AI diagnostics replace those of human doctors, but research suggests AI has a long way to go (if such a future is even possible). In a 2021 study of three different deep learning models, researchers found that each AI only performed well as a diagnostic tool when using the datasets from the specific hospital the model trained on (which included the National Institutes of Health Clinical Center, Stanford Health Care and Beth Israel Deaconess Medical Center). When researchers introduced data from different hospitals, the AI results went "haywire." In hospitals that use diagnostic AI, doctors often ignore it. For example, a study of AI usage amongst diagnostic radiologists found that the doctors making bone age determinations and diagnosing breast cancer viewed the AI results as "opaque and unhelpful."

> *"The fantasy of computerized medicine sounds a lot like the fantasy of the self-driving car: fascinating, but impractical."*

AI diagnostic tools aren't just unreliable – they can also display harmful bias. For example, Google launched an AI-powered dermatology tool in 2021 designed to help people detect various skin issues. Google had its own goals when releasing the tool, as it wanted to motivate users to search more. The tool contained bias, as Google trained its AI on images from patients in only two US states, and the vast majority of these images featured patients with light and medium-toned skin, while only 3.5% featured patients with darker-toned skin. Given that skin cancer appears different in people with different skin tones, rolling out a product without training the AI model on a wide range of skin tones demonstrates a lack of concern for the public interest.

Embed an algorithmic review process into your business model.

Moving forward, organizations should embrace the following organizational process to embed responsible AI into their business processes:

1. **Take inventory** – Examine your company's algorithms and system vendors.
2. **Audit a single algorithm** – Just as engineers inspect roads and bridges, it's essential that you audit algorithms, start by auditing one.
3. **Remediate "harms"** – If you discover biases, take corrective action and reshape your business processes to avoid future harm.
4. **Learn from your mistakes** – Learn from this process, proactively looking for similar problems.
5. **Make algorithmic review an ongoing process** – Auditing may have less marketing hype than innovation, but it's a vital business process. Update your business processes, ensuring you have funding in place for infrastructure – not just innovation.
6. **Repeat these steps** – Audit more algorithms, remediating harms where necessary. Work with people from diverse backgrounds, as people with varying perspectives will be better equipped to flag potential issues.

Build a better world with algorithmic auditing and accountability.

Tech companies often use the "lean start-up" frame, searching for pain points, solutions and target markets on their journey toward scaling up. When attempting to find solutions to complicated problems, organizations would be best served by the frame of "public interest technology." Those creating public interest technology aspire to work toward to public good and advance collective well-being.

> *"Algorithmic auditing and algorithmic accountability reporting are two strains of public interest technology that I think show the most promise for remedying algorithmic harms along the axes of race, gender and ability."*

In particular, creating public interest technology that supports algorithmic accountability reporting and auditing could remedy some harms caused by machine bias. In the United States, new bills limit the activities of Big Tech, such as the Algorithmic Accountability Act of 2022, which would legally require companies to check algorithms for both effectiveness and bias before introducing them to the public. It's time to forgo technochauvinist optimism for a more discerning, critical view of technology while working toward a more just, equitable world.

Did you like this summary?

Buy the book

https://getab.li/47369

This document is restricted to personal use only.

getAbstract maintains complete editorial responsibility for all parts of this abstract. All rights reserved. No part of this abstract may be reproduced or transmitted or used for training any machine learning system in any form or by any means – electronic, photocopying or otherwise – without prior written permission of getAbstract AG (Switzerland).