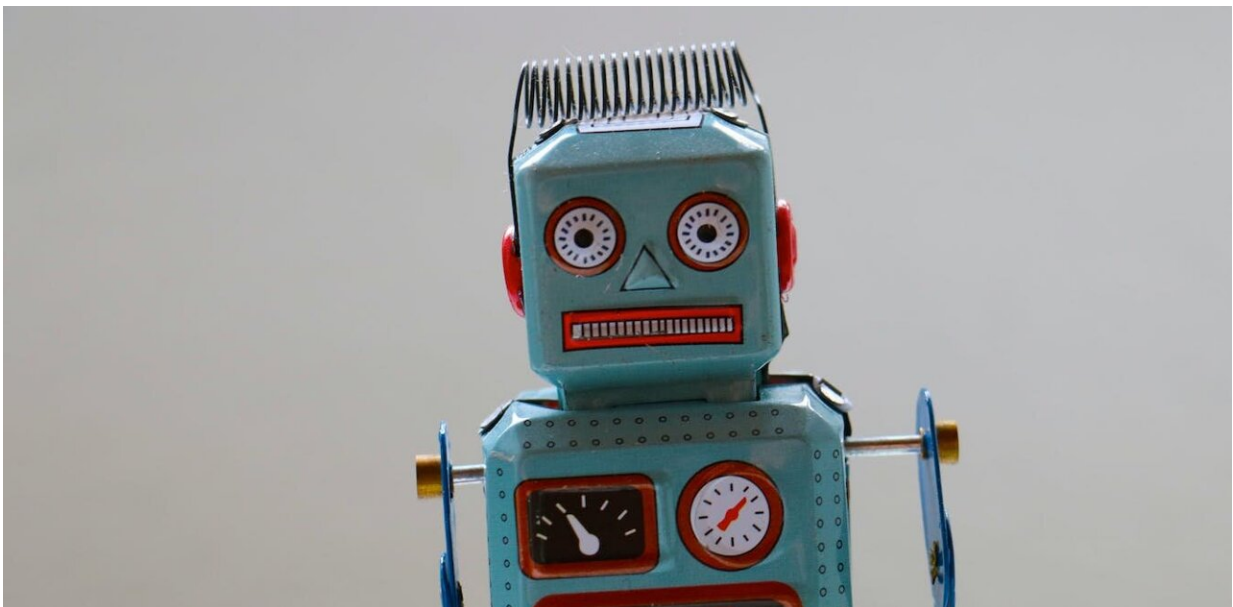# Can a machine be racist? AI has shown troubling signs of bias, but there are reasons for optimism

March 7 2023, by Charles Barbour

Credit: Rock n Roll Monkey/Unsplash

One day in mid-2013, four people, including two police officers and a social worker, arrived unannounced at the home of Chicago resident [Robert McDaniel](#).

McDaniel had only ever had minor run-ins with the law—street gambling, marijuana possession, nothing even remotely violent. But his

visitors informed him that a computer program had determined that the person living at his address was unusually likely to be involved in a future shooting.

Perhaps he would be the perpetrator, perhaps the victim. The computer wasn't sure. But due to something called "predictive policing," the social worker and the police would be visiting him on a regular basis.

McDaniel was not enthusiastic about either prospect, but the computer had made its decision, so they were offers he could not refuse.

The social worker returned frequently with referrals to mental health programs, violence prevention programs, job training programs, and so forth. The police also returned frequently—to remind him that he was being watched.

The official attention did not go unnoticed in McDaniel's neighborhood. Rumors spread that he was a police informant. In 2017, those rumors led to him being shot. In 2020, it happened again.

Thus, in a bizarre sense, the computer's prediction could be said to have caused the tragedy it claimed to predict. Indeed, it could be said to have caused it twice.

Racist machines?

We would not be wrong to interpret McDaniel's story as a Kafkaesque nightmare about a man caught in an inexorable bureaucratic machine, or a Faustian parable about what happens when technology escapes the bounds of human control.

But according to the professor of data journalism and accomplished computer scientist Meredith Broussard, it is also, and perhaps more

importantly, a story about racism.

For when the police arrived at his door in 2013, Robert McDaniel was not just any man. He was a young Black man living in a neighborhood that had been shaped by [a shameful history of racist redlining](). The neighborhood was, as a result, the home of a disproportionate level of both criminal violence and police surveillance. McDaniel was thus all but destined to become the target of the kind of technologically driven predictive policing that led to his being shot.

And, Broussard maintains, what happened to Robert McDaniel is but one example of the many ways that AI is augmenting and exacerbating the inequalities that characterize modern social life.

Do not worry that machines will rise up, take power, and create a completely new world, Broussard argues. Worry that they will silently reproduce and reinforce the world that already exists.

Technochauvinism

At first glance, the notion that a machine might be racist, sexist, ableist or biased in any fashion seems a little strange.

Science, technology, and especially mathematics are presented to us as the gold standards of neutrality. They don't judge. They calculate. And calculation is by definition above the messy world of bigotry and intolerance, hatred and division.

On Broussard's account, this line of thought is a convenient deception. Its purpose is to paper over an increasingly pervasive way of thinking Broussard calls "technochauvinism." Technochauvinism, she explains, "is a kind of bias that considers computational solutions to be superior to all other solutions. Embedded in this bias is an a priori assumption that

computers are better than humans."

More accurately, the assumption is "that the people who make and program computers are better than other humans."

Mathematics on its own might be neutral. But as soon as it is put to any use whatsoever, it becomes a vehicle for human values, human prejudices and human frailties.

Critical AI studies

More than a Glitch contributes to a rapidly expanding field of scholarship and activism that has variously been dubbed Critical Algorithm Studies, Critical Data Studies, and Critical AI Studies.

Here we might include important works like Safiya Umoja Noble's Algorithms of Oppression (2018), which shows how seemingly impartial information sorting tools perpetuate systematic racism, Soshana Zuboff's Age of Surveillance Capitalism (2018), which argues that big data is transforming human experience itself into a surplus that modern capitalism can extract as profit, and Kate Crawford's Atlas of AI (2021), which suggests that we approach AI not as a collection of computer programs, but as an integrated ecology of material relations between humans and the world.

There is even a popular documentary called Coded Bias (2020), directed by Shalini Kantayya and featuring, among others, Broussard herself.

Amid this impressive company, Broussard's book is distinguished by at least two elements: its extraordinarily expansive scope on the one hand, and its no-nonsense approach to both the problem and its solutions on the other.

Baked in bias

The expansiveness of Broussard's approach is discernible in her thesis, which she states directly at the outset: "The biases embedded in technology are more than mere glitches; they are baked in from the beginning."

At least part of the reason for this "baked in" bias can be found in the demographics of those who work in the field. Google's 2019 annual report, for example, showed that only 3% of the tech giant's employees are Black, a deficiency that is common across the industry.

More personally, Broussard notes that, as an undergraduate at Harvard, she was one of only six women majoring in computer science, and the only Black woman.

There has been a great deal of well-meaning discourse around the need to make technology "more ethical" or "fairer," but Broussard contends that real change will require a much more systematic "audit" designed to determine "how it is racist, gender-biased, or ableist":

"We should not cede control of essential civic functions to these tech systems, nor should we claim they are 'better' or 'more innovative' until and unless those technical systems work for every person regardless of skin color, class, age, gender, and ability."

Broussard proceeds to explain the essential principles of AI and machine learning, and the mathematics on which they are based. She is by no means a technophobe; she clearly has enormous knowledge of and respect for the science in question. But she also insists that science cannot be purely mathematical. It relies on "storytelling" as well.

As Broussard sees it, "we understand the quantitative through the

qualitative." Sheer numbers will always privilege the established order. And they will always subordinate or miss entirely what Broussard calls the "edge cases."

But it is the edge cases, or those cases that statistics and probability cannot help but push to the margins, that represent the potential for both oppression and change.

Everyday algorithms

What follows is a catalog of examples of AI technology failing to account for human diversity and reproducing social inequalities.

Facial recognition and bio-metric identification technologies, for example, have repeatedly been shown to be farcically inept when dealing with anyone other than white, cis-gendered people.

Here digital imaging sits neatly within a long history of discriminatory photographic and film technology designed by and for a small fraction of humans. The result is not just a lack of representation. The effects are concrete and very destructive, particularly when the technologies in question are placed in the hands of law enforcement.

But if we focus exclusively on the more sensational uses (and failures) of AI, we will miss the extent to which it has infiltrated nearly every aspect of our daily lives, including marketing and politics, of course, but also education, medicine, employment, economics, transportation, and more or less everything we do with our mobile phones—which means more or less everything we do full stop.

To explore the everyday use of AI, Broussard weaves together stories, anecdotes, and vignettes drawn from both her research and her personal experience.

Through these stories, she shows how AI is currently being used to, among other things, assign students imaginary grades—or grades based not on their achievements, but on what a statistically trained algorithm predicts they will achieve. It is also being used to determine which job applicants will be granted an interview and run medical diagnostics that presuppose antiquated conceptions of race, gender and ability.

No matter what we do in the modern world, it seems, there is almost always an algorithm churning away in the background, generating results that heavily determine our actions and decisions. And it is almost always doing so to the disadvantage of already disadvantaged groups.

Reforming AI

This brings us to Broussard's no-nonsense approach and her optimism regarding the possibility of changing the systems she describes.

It is undeniably hard not to be overwhelmed by the seemingly unstoppable course of technological development in the contemporary world, especially around AI and everything it appears poised to reinvent. But cynicism can also be a hiding place for privilege, and those who say despairingly that nothing can be done are often those who stand to benefit most from nothing being done.

With this in mind, Broussard is keen to distinguish between "imaginary" or "general AI" and "real" or "narrow AI."

The former is "the AI that will take over the world, the so-called singularity where robots become uncontrollable and irreversible—killer robots—and so on." This, she says, is not real. "Real AI, which we have and use every day," is nothing more than "math."

AI, in other words, is not magical. It is a sophisticated pattern-detection

machine. And while it might be able to detect "patterns that humans can't easily see," and thus function as a kind of "black box," that does not mean that it is "*impossible* to describe."

For the same reason, Broussard is confident that humans can and should treat AI like a tool. Ultimately, she thinks, it is nothing more than a reflection of those who use it. If it is biased, that is only because it is the product of a biased society. And it will change precisely insofar as society changes.

Tech, Broussard concludes, "is racist and sexist and ableist because the world is so. Computers just reflect the existing reality and suggest that things will stay the same—they predict the status quo. By adopting a more critical view of technology, and by being choosier about the tech we allow into our lives and our society, we can employ technology to stop reproducing the world as it is, and get closer to a world that is truly more just."

One question

This is the only place I would be inclined to ask a question. For while Broussard's political agenda is unimpeachable, her approach to technology does seem rather humanist and instrumental.

For a very long time now, science and technology scholars—Donna Haraway, Friedrich Kittler and Bruno Latour, to name just a few—have been suggesting that, as much as we make and use technologies, technologies also make and use us.

The question of how humans will get AI under our control, or how we will guide it towards our ethical and political ends, is quite distinct from the question of how the same technology will transform what it means to be human and how we live in relation to each other and our worlds.

Neither line of inquiry should be considered superior. But it is hard to imagine pursuing one very far without at least encountering the other. A richer conversation between the approach represented by Broussard, and the approach represented by those of us in the tradition of figures like Harraway, Kittler and Latour, would seem to be in order.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation