

Can We Make Our Robots Less Biased Than We Are?

A.I. developers are committing to end the injustices in how their technology is often made and used

By David Berreby
The New York Times

Chris S. Crawford, a computer scientist at the University of Alabama. “I personally was in Silicon Valley when some of these technologies were being developed,” he said, and more than once, “I would sit down and they would test it on me, and it wouldn’t work. And I was like, You know why it’s not working, right?”

On a summer night in Dallas in 2016, a bomb-handling robot made technological history. Police officers had attached roughly a pound of C-4 explosive to it, steered the device up to a wall near an active shooter and detonated the charge. In the explosion, the assailant, Micah Xavier Johnson, became the first person in the United States to be killed by a police robot.

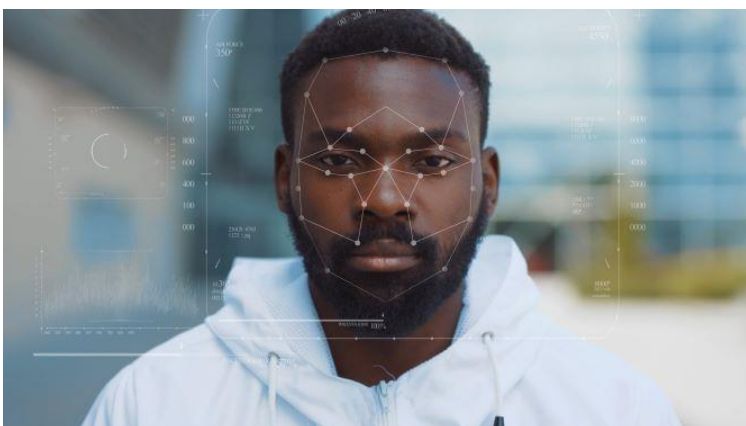
Afterward, then-Dallas Police Chief David Brown called the decision sound. Before the robot attacked, Mr. Johnson had shot five officers dead, wounded nine others and hit two civilians, and negotiations had

stalled. Sending the machine was safer than sending in human officers, Mr. Brown said.

But some robotics researchers were troubled. “Bomb squad” robots are marketed as tools for safely disposing of bombs, not for delivering them to targets. (In 2018, police officers in Dixmont, Maine, ended a shootout in a similar manner.) Their profession had supplied the police with a new form of lethal weapon, and in its first use as such, it had killed a Black man.

“A key facet of the case is the man happened to be African-American,” Ayanna Howard, a robotics researcher at Georgia Tech, and Jason Borenstein, a colleague in the university’s school of public policy, wrote in a 2017 paper titled “The Ugly Truth About Ourselves and Our Robot Creations” in the journal *Science and Engineering Ethics*.

Like almost all police robots in use today, the Dallas device was a straightforward remote-control platform. But more sophisticated robots are being developed in labs around the world, and they will use artificial intelligence to do much more. A robot with algorithms for, say, facial recognition, or predicting people’s actions, or deciding on its own to fire “nonlethal” projectiles is a robot that many researchers find problematic. The reason: Many of today’s algorithms are biased against unlike the white, male, affluent and



able-bodied designers of most computer and robot systems.

While Mr. Johnson's death resulted from a human decision, in the future such a decision might be made by a robot — one created by humans, with their flaws in judgment baked in.

"Given the current tensions arising from police shootings of African-American men from Ferguson to Baton Rouge," Dr. Howard, a leader of the organization Black in Robotics, and Dr. Borenstein wrote, "it is disconcerting that robot peacekeepers, including police and military robots, will, at some point, be given increased freedom to decide whether to take a human life, especially if problems related to bias have not been resolved."

Last summer, hundreds of A.I. and robotics researchers signed statements committing themselves to changing the way their fields work. One statement, from the organization Black in Computing, sounded an alarm that "the technologies we help create to benefit society are also disrupting Black communities through the proliferation of racial profiling." Another manifesto, "No Justice, No Robots," commits its signers to refusing to work with or for law enforcement agencies. Over the past decade, evidence has accumulated that "bias is the original sin of A.I.," Dr. Howard notes in her 2020 audiobook, "Sex, Race and Robots." Facial-recognition systems have been shown to be more accurate in identifying white faces than those of other people. (In January, one such system told the Detroit police that it had matched photos of a suspected thief with the driver's license photo of Robert Julian-Borchak Williams, a Black man with no connection to the crime.)

There are A.I. systems enabling self-driving cars to detect pedestrians — last year Benjamin Wilson of Georgia Tech and his colleagues found that eight such systems

were worse at recognizing people with darker skin tones than paler ones. Joy Buolamwini, the founder of the Algorithmic Justice League and a graduate researcher at the M.I.T. Media Lab, has encountered interactive robots at two different laboratories that failed to detect her. (For her work with such a robot at M.I.T., she wore a white mask in order to be seen.)

The long-term solution for such lapses is "having more folks that look like the United States population at the table when technology is designed," said Chris S. Crawford, a professor at the University of Alabama who works on direct brain-to-robot controls. Algorithms trained mostly on white male faces (by mostly white male developers who don't notice the absence of other kinds of people in the process) are better at recognizing white males than other people.

"I personally was in Silicon Valley when some of these technologies were being developed," he said. More than once, he added, "I would sit down and they would test it on me, and it wouldn't work. And I was like, You know why it's not working, right?"

Robot researchers are typically educated to solve difficult technical problems, not to consider societal questions about who gets to make robots or how the machines affect society. So it was striking that many roboticists signed statements declaring themselves responsible for addressing injustices in the lab and outside it. They committed themselves to actions aimed at making the creation and usage of robots less unjust.

"I think the protests in the street have really made an impact," said Odest Chadwicke Jenkins, a roboticist and A.I. researcher at the University of Michigan. At a conference earlier this year, Dr. Jenkins, who works on robots that can assist and collaborate with people, framed

his talk as an apology to Mr. Williams. Although Dr. Jenkins doesn't work in face-recognition algorithms, he felt responsible for the A.I. field's general failure to make systems that are accurate for everyone. "This summer was different than any other than I've seen before," he said. "Colleagues I know and respect, this was maybe the first time I've heard them talk about systemic racism in these terms. So that has been very heartening." He said he hoped that the conversation would continue and result in action, rather than dissipate with a return to business-as-usual.

Dr. Jenkins was one of the lead organizers and writers of one of the summer manifestoes, produced by Black in Computing. Signed by nearly 200 Black scientists in computing and more than 400 allies (either Black scholars in other fields or non-Black people working in related areas), the document describes Black scholars' personal experience of "the structural and institutional racism and bias that is integrated into society, professional networks, expert communities and industries."

The statement calls for reforms, including ending the harassment of Black students by campus police officers, and addressing the fact that Black people get constant reminders that others don't think they belong. (Dr. Jenkins, an associate director of the Michigan Robotics Institute, said the most common question he hears on campus is, "Are you on the football team?") All the nonwhite, non-male researchers interviewed for this article recalled such moments. In her book, Dr. Howard recalls walking into a room to lead a meeting about navigational A.I. for a Mars rover and being told she was in the wrong place because secretaries were working down the hall.

The open letter is linked to a page of specific action items. The items range

from not placing all the work of "diversity" on the shoulders of minority researchers to ensuring that at least 13 percent of funds spent by organizations and universities go to Black-owned businesses to tying metrics of racial equity to evaluations and promotions. It also asks readers to support organizations dedicated to advancing people of color in computing and A.I., including Black in Engineering, Data for Black Lives, Black Girls Code, Black Boys Code and Black in A.I.

As the Black in Computing open letter addressed how robots and A.I. are made, another manifesto appeared around the same time, focusing on how robots are used by society. Entitled "No Justice, No Robots," the open letter pledges its signers to keep robots and robot research away from law enforcement agencies. Because many such agencies "have actively demonstrated brutality and racism toward our communities," the statement says, "we cannot in good faith trust these police forces with the types of robotic technologies we are responsible for researching and developing."

Last summer, distressed by police officers' treatment of protesters in Denver, two Colorado roboticists — Tom Williams, of the Colorado School of Mines and Kerstin Haring, of the University of Denver — started drafting "No Justice, No Robots." So far, 104 people have signed on, including leading researchers at Yale and M.I.T., and younger scientists at institutions around the country.

"The question is: Do we as roboticists want to make it easier for the police to do what they're doing now?" Dr. Williams asked. "I live in Denver, and this summer during protests I saw police tear-gassing people a few blocks away from me. The combination of seeing police brutality on the news and then seeing it in Denver was the catalyst."

Dr. Williams is not opposed to working

with government authorities. He has conducted research for the Army, Navy and Air Force, on subjects like whether humans would accept instructions and corrections from robots. (His studies have found that they would.) The military, he said, is a part of every modern state, while American policing has its origins in racist institutions, such as slave patrols — “problematic origins that continue to infuse the way policing is performed,” he said in an email.

“No Justice, No Robots” proved controversial in the small world of robotics labs, since some researchers felt that it wasn’t socially responsible to shun contact with the police.

“I was dismayed by it,” said Cindy Bethel, director of the Social, Therapeutic and Robotic Systems Lab at Mississippi State University. “It’s such a blanket statement,” she said. “I think it’s naïve and not well-informed.” Dr. Bethel has worked with local and state police forces on robot projects for a decade, she said, because she thinks robots can make police work safer for both officers and civilians. One robot that Dr. Bethel is developing with her local police department is equipped with night-vision cameras, that would allow officers to scope out a room before they enter it. “Everyone is safer when there isn’t the element of surprise, when police have time to think,” she said. Adhering to the declaration would prohibit researchers from working on robots that conduct search-and-rescue operations, or in the new field of “social robotics.” One of Dr. Bethel’s research projects is developing technology that would use small, humanlike robots to interview children who have been abused, sexually assaulted, trafficked or otherwise traumatized. In one of her recent studies, 250 children and adolescents who were interviewed about bullying were often

willing to confide information in a robot that they would not disclose to an adult. Having an investigator “drive” a robot in another room thus could yield less painful, more informative interviews of child survivors, said Dr. Bethel, who is a trained forensic interviewer.

“You have to understand the problem space before you can talk about robotics and police work,” she said. “They’re making a lot of generalizations without a lot of information.”

Dr. Crawford is among the signers of both “No Justice, No Robots” and the Black in Computing open letter. “And you know, anytime something like this happens, or awareness is made, especially in the community that I function in, I try to make sure that I support it,” he said.

Dr. Jenkins declined to sign the “No Justice” statement. “I thought it was worth consideration,” he said. “But in the end, I thought the bigger issue is, really, representation in the room — in the research lab, in the classroom, and the development team, the executive board.” Ethics discussions should be rooted in that first fundamental civil-rights question, he said.

Dr. Howard has not signed either statement. She reiterated her point that biased algorithms are the result, in part, of the skewed demographic — white, male, able-bodied — that designs and tests the software.

“If external people who have ethical values aren’t working with these law enforcement entities, then who is?” she said. “When you say ‘no,’ others are going to say ‘yes.’ It’s not good if there’s no one in the room to say, ‘Um, I don’t believe the robot should kill.’”