

GENDER SHADES

# The influential project that sparked the end of IBM's facial recognition program

June 10, 2020



By **Nicolas Rivero**  
Reporter



REUTERS/LOREN ELLIOTT

FROM OUR OBSESSION

## Machines with Brains



AI is upending companies, industries, and humanity.

On June 8, IBM CEO Arvind Krishna announced the end of his firm's involvement in facial recognition in a letter to US Senators. The company, he wrote, "firmly opposes and will not condone uses of any technology, including facial recognition technology offered by other vendors, for mass surveillance, racial profiling, violations of basic human rights and freedoms."

IBM's decision to abandon facial recognition can be seen as a reaction to protests against racism that have called for police reform in the United States, including law enforcement's use of the flawed technology. But the process began much earlier. In part, it can be traced back to one influential piece of research: the Gender Shades project, from MIT Media Lab's Joy Buolamwini and Microsoft Research's Timnit Gebru.

The computer scientists' 2018 study showed that commercial facial recognition software was significantly less accurate for darker-skinned women than for lighter-skinned men. They found that IBM had the greatest disparity: Results for darker-skinned women were 34.4% less accurate than those for lighter-skinned men.

“What started this conversation was the research by Timnit Gebru, Joy Buolamwini, and Inioluwa Raji on bias in facial recognition,” said Maria De-Arteaga, an incoming assistant professor at UT Austin who studies ethics in AI. “That was the work that started showing the computer science community that this is something that is wrong with this technology.” Others in the field had noted the potential for facial recognition to introduce racial bias. But the 2018 paper was the first to evaluate the biases embedded in commercial products already in use.

In a January 2019 follow-up study, Buolamwini and Raji, who is a tech fellow at NYU's AI Now Institute, showed that the firms studied in the Gender Shades study had each made changes to their facial recognition tools to reduce racial and gender biases.

While many applications of AI are subject to potentially harmful biases, Raji said that facial recognition is especially perilous. “It involves easily accessible, sensitive, identifiable

information about many, many people,” she said. “That makes facial recognition a particularly dangerous technology that has an incredibly high potential for surveillance and can be easily coopted and used by the police in a way that puts communities of color at risk.”

The same month the follow-up paper published, IBM—citing Gebru and Buolamwini’s original study—responded by releasing the “Diversity in Faces” dataset, which contained 1 million images meant to sample a more diverse group of faces. In a press release, the company said it hoped the dataset would “advance the study of fairness and accuracy in facial recognition technology.”

Unfortunately, IBM scraped its million photos from Flickr without asking the photographers or their subjects, as NBC reporter Olivia Solon revealed in March of 2019. Another round of controversy followed. In September, IBM quietly removed the “Detect Faces” tool from its public API, meaning that developers could no longer simply buy access to the company’s facial recognition tools.

IBM told Quartz in an email that it has been gradually rolling back facial recognition for existing clients over the course of months. Yesterday, in the wake of widespread protests over the killing of George Floyd and calls for police abolition, the company publicly announced those changes. It is no longer researching, developing, marketing, or selling facial recognition tools to any client, and is not using the technology itself. It will, however, continue to develop visual detection tools for objects—for example, building agricultural machines that can recognize crops.

In a Medium post following IBM’s announcement, Buolamwini said the company “made a bold move in the right

direction” and called on IBM to follow up by donating \$1 million to organizations fighting discrimination in tech. “I am proud of the work [my colleagues and I] have done together and call on the tech sector and other researchers to do more,” she said in an email.

“The fact that IBM said they would stop research on [facial recognition] will be something that people point to as evidence that there is something unsavory about being in this space now,” said Jeanna Matthews, a professor of computer science at Clarkson University. “I think that will change the conversation.”

“I think it is an important stance to have a company as big as IBM disavow facial recognition,” Raj said. Although other players like Microsoft and Amazon sell their recognition technology to more clients, she said IBM’s move “does influence the public understanding and perception of the dangers of the technology. To have any kind of reinforcement of that message makes it easier for policy makers to push for regulation that will affect the Amazons and the Microsofts.”

*Correction: This post has been updated to correct Jeanna Matthews’ title. She is a professor of computer science at Clarkson University, not an associate professor.*



**Subscribe to the Daily Brief, our morning email with news and insights you need to understand our changing world.**

By providing your email, you agree to the [Quartz Privacy Policy](#).