The New York Times | http://nyti.ms/28YaKg7

SundayReview | OPINION

Artificial Intelligence's White Guy Problem

By KATE CRAWFORD   JUNE 25, 2016

ACCORDING to some prominent voices in the tech world, artificial intelligence presents a looming existential threat to humanity: Warnings by luminaries like Elon Musk and Nick Bostrom about "the singularity" — when machines become smarter than humans — have attracted millions of dollars and spawned a multitude of conferences.

But this hand-wringing is a distraction from the very real problems with artificial intelligence today, which may already be exacerbating inequality in the workplace, at home and in our legal and judicial systems. Sexism, racism and other forms of discrimination are being built into the machine-learning algorithms that underlie the technology behind many "intelligent" systems that shape how we are categorized and advertised to.

Take a small example from last year: Users discovered that Google's photo app, which applies automatic labels to pictures in digital photo albums, was classifying images of black people as gorillas. Google apologized; it was unintentional.

But similar errors have emerged in Nikon's camera software, which misread images of Asian people as blinking, and in Hewlett-Packard's web camera software, which had difficulty recognizing people with dark skin tones.

This is fundamentally a data problem. Algorithms learn by being fed certain images, often chosen by engineers, and the system builds a model of the world based

on those images. If a system is trained on photos of people who are overwhelmingly white, it will have a harder time recognizing nonwhite faces.

A very serious example was revealed in an investigation published last month by ProPublica. It found that widely used software that assessed the risk of recidivism in criminals was twice as likely to mistakenly flag black defendants as being at a higher risk of committing future crimes. It was also twice as likely to incorrectly flag white defendants as low risk.

The reason those predictions are so skewed is still unknown, because the company responsible for these algorithms keeps its formulas secret — it's proprietary information. Judges do rely on machine-driven risk assessments in different ways — some may even discount them entirely — but there is little they can do to understand the logic behind them.

Police departments across the United States are also deploying data-driven risk-assessment tools in "predictive policing" crime prevention efforts. In many cities, including New York, Los Angeles, Chicago and Miami, software analyses of large sets of historical crime data are used to forecast where crime hot spots are most likely to emerge; the police are then directed to those areas.

At the very least, this software risks perpetuating an already vicious cycle, in which the police increase their presence in the same places they are already policing (or overpolicing), thus ensuring that more arrests come from those areas. In the United States, this could result in more surveillance in traditionally poorer, nonwhite neighborhoods, while wealthy, whiter neighborhoods are scrutinized even less. Predictive programs are only as good as the data they are trained on, and that data has a complex history.

Histories of discrimination can live on in digital platforms, and if they go unquestioned, they become part of the logic of everyday algorithmic systems. Another scandal emerged recently when it was revealed that Amazon's same-day delivery service was unavailable for ZIP codes in predominantly black neighborhoods. The areas overlooked were remarkably similar to those affected by mortgage redlining in the mid-20th century. Amazon promised to redress the gaps, but it reminds us how systemic inequality can haunt machine intelligence.

And then there's gender discrimination. Last July, computer scientists at Carnegie Mellon University found that women were less likely than men to be shown ads on Google for highly paid jobs. The complexity of how search engines show ads to internet users makes it hard to say why this happened — whether the advertisers preferred showing the ads to men, or the outcome was an unintended consequence of the algorithms involved.

Regardless, algorithmic flaws aren't easily discoverable: How would a woman know to apply for a job she never saw advertised? How might a black community learn that it were being overpoliced by software?

We need to be vigilant about how we design and train these machine-learning systems, or we will see ingrained forms of bias built into the artificial intelligence of the future.

Like all technologies before it, artificial intelligence will reflect the values of its creators. So inclusivity matters — from who designs it to who sits on the company boards and which ethical perspectives are included. Otherwise, we risk constructing machine intelligence that mirrors a narrow and privileged vision of society, with its old, familiar biases and stereotypes.

Sign Up for the Opinion Today Newsletter

Every weekday, get thought-provoking commentary from Op-Ed columnists, the Times editorial board and contributing writers from around the world.

Enter your email addr

☑ Receive occasional updates and special offers for The New York Times's products and services.

I'm not a robot

reCAPTCHA
Privacy - Terms

See Sample
Privacy Policy

If we look at how systems can be discriminatory now, we will be much better placed to design fairer artificial intelligence. But that requires far more accountability from the tech community. Governments and public institutions can do their part as well: As they invest in predictive technologies, they need to commit to fairness and due process.

While machine-learning technology can offer unexpected insights and new forms of convenience, we must address the current implications for communities that have less power, for those who aren't dominant in elite Silicon Valley circles.

Currently the loudest voices debating the potential dangers of superintelligence are affluent white men, and, perhaps for them, the biggest threat *is* the rise of an artificially intelligent apex predator.

But for those who already face marginalization or bias, the threats are here.

Kate Crawford is a principal researcher at Microsoft and co-chairwoman of a White House symposium on society and A.I.

Follow The New York Times Opinion section on Facebook and Twitter (@NYTOpinion), and sign up for the Opinion Today newsletter.

A version of this op-ed appears in print on June 26, 2016, on page SR11 of the New York edition with the headline: A.I.'s White Guy Problem.