

# AI technologies, like police facial recognition, discriminate against people of color

August 25 2020, by Jane Bailey, Jacquelyn Burkell and Valerie Steeves



Credit: AI-generated image (disclaimer)

Detroit police wrongfully arrested Robert Julian-Borchak Williams in January 2020 for a shoplifting incident that had taken place two years earlier. Even though Williams had nothing to do with the incident, facial recognition technology used by Michigan State Police "matched" his face with a grainy image obtained from an in-store surveillance video



showing another African American man taking US\$3,800 worth of watches.

Two weeks later, the case was dismissed at the prosecution's request. However, relying on the faulty match, police had already handcuffed and arrested Williams in front of his family, forced him to provide a mug shot, fingerprints and a sample of his DNA, interrogated him and imprisoned him overnight.

Experts suggest that Williams is not alone, and that others have been subjected to similar injustices. The ongoing controversy about police use of Clearview AI certainly underscores the privacy risks posed by facial recognition technology. But it's important to realize that <u>not all of us</u> <u>bear those risks equally</u>.

## **Training racist algorithms**

Facial recognition technology that is <u>trained on and tuned to Caucasian</u> <u>faces</u> systematically misidentifies and mislabels racialized individuals: numerous studies report that facial recognition technology is "<u>flawed and</u> <u>biased</u>, with significantly higher error rates when used against people of <u>color</u>."

This <u>undermines the individuality and humanity of racialized persons</u> who are more likely to be misidentified as criminal. The technology—and the identification errors it makes—reflects and further entrenches long-standing social divisions that are deeply entangled with racism, sexism, homophobia, settler-colonialism and other intersecting oppressions.

## How technology categorizes users



In his game-changing 1993 book, *The Panoptic Sort*, scholar Oscar Gandy warned that "complex technology [that] involves the collection, processing and sharing of information about individuals and groups that is generated through their daily lives ... is used to coordinate and control their access to the goods and services that define life in the modern capitalist economy." Law enforcement uses it to pluck suspects from the general public, and private organizations use it to determine whether we have access to things like <u>banking</u> and <u>employment</u>.

Gandy prophetically warned that, if left unchecked, this form of "cybernetic triage" would exponentially disadvantage members of equality-seeking communities—for example, groups that are racialized or socio-economically disadvantaged—both in terms of what would be allocated to them and how they might come to understand themselves.

Some 25 years later, we're now living with the panoptic sort on steroids. And examples of its negative effects on equality-seeking communities abound, such as the false identification of Williams.

## **Pre-existing bias**

This sorting using algorithms infiltrates the most fundamental aspects of everyday life, occasioning both direct and structural violence in its wake.

The direct violence experienced by Williams is immediately evident in the events surrounding his arrest and detention, and the individual harms he experienced are obvious and can be traced to the actions of police who chose to rely on the technology's "match" to make an arrest. More insidious is the <u>structural violence</u> perpetrated through <u>facial recognition</u> <u>technology</u> and <u>other digital technologies</u> that rate, match, categorize and sort individuals in ways that magnify pre-existing discriminatory patterns.



Structural violence harms are less obvious and less direct, and cause injury to equality-seeking groups through systematic denial to power, resources and opportunity. Simultaneously, it increases direct risk and harm to individual members of those groups.

Predictive policing uses <u>algorithmic processing of historical data to</u> <u>predict when and where new crimes are likely to occur</u>, assigns police resources accordingly and embeds enhanced police surveillance into communities, usually in lower-income and racialized neighborhoods. This increases the chances that any <u>criminal activity</u>—including less serious criminal activity that might otherwise prompt no police response—will be detected and punished, ultimately limiting the life chances of the people who live within that environment.

And the evidence of inequities in other sectors continues to mount. Hundreds of students in the United Kingdom protested on Aug. 16 against the disastrous results of Ofqual, a flawed algorithm the U.K. government used to determine which students would qualify for university. In 2019, Facebook's microtargeting ad service helped dozens of public and private sector employers exclude people from receiving job ads on the basis of age and gender. Research conducted by ProPublica has documented race-based price discrimination for online products. And search engines regularly produce racist and sexist results.

## **Perpetuating oppression**

These outcomes matter because they perpetuate and deepen pre-existing inequalities based on characteristics like race, gender and age. They also matter because they deeply affect how we come to know ourselves and the world around us, sometimes by pre-selecting the information we receive in ways that reinforce stereotypical perceptions. Even technology companies themselves acknowledge the urgency of stopping algorithms from perpetuating discrimination.



To date the success of ad hoc investigations, conducted by the tech companies themselves, has been inconsistent. Occasionally, corporations involved in producing discriminatory systems withdraw them from the market, such as when <u>Clearview AI announced it would no longer offer</u> <u>facial recognition technology in Canada</u>. But often such decisions result from regulatory scrutiny or public outcry only *after* members of equality-seeking communities have already been harmed.

It's time to give our regulatory institutions the tools they need to address the problem. Simple privacy protections that hinge on obtaining individual consent to enable data to be captured and repurposed by companies cannot be separated from the discriminatory outcomes of that use. This is especially true in an era when most of us (including technology companies themselves) cannot fully understand what algorithms do or why they produce specific results.

### Privacy is a human right

Part of the solution entails breaking down the current regulatory silos that treat privacy and human rights as separate issues. Relying on a consent-based data protection model flies in the face of the basic principle that privacy and equality are both human rights that cannot be contracted away.

Even <u>Canada's Digital Charter</u>—the federal government's latest attempt to respond to the shortcomings of the current state of the digital environment—maintains these conceptual distinctions. It treats hate and extremism, control and consent, and strong democracy as separate categories.

To address algorithmic discrimination, we must recognize and frame both privacy and equality as human rights. And we must create an infrastructure that is equally attentive to and expert in both. Without



such efforts, the glossy sheen of math and science will continue to camouflage AI's discriminatory biases, and travesties such as that inflicted on Williams can be expected to multiply.

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

Provided by The Conversation

Citation: AI technologies, like police facial recognition, discriminate against people of color (2020, August 25) retrieved 27 April 2024 from <u>https://techxplore.com/news/2020-08-ai-technologies-police-facial-recognition.html</u>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.