

'Encoding the same biases': Artificial intelligence's limitations in coronavirus response

September 7 2020, by Natalie Grover, Horizon Magazine



Artificial intelligence is being used to understand and address coronavirus but the results will only be as unbiased as the information fed into the algorithms. Credit: Daniele Marzocchi/Flickr, licensed under CC BY-NC 2.0

As the coronavirus pandemic endures, the socio-economic implications of race and gender in contracting COVID-19 and dying from it have been laid bare. Artificial intelligence (AI) is playing a key role in the response, but it could also be exacerbating inequalities within our health systems—a critical concern that is dragging the technology's limitations back into the spotlight.

The response to the crisis has in many ways been mediated by data—an explosion of information being used by AI algorithms [to better understand and address COVID-19](#), including tracking the virus' spread and [developing therapeutic interventions](#).

AI, like its human maker, is not immune to bias. The technology—generally designed to digest large volumes of data and make deductions to support decision making—reflects the prejudices of the humans who develop it and feed it information that it uses to spit out outcomes. For example, years ago when Amazon developed an AI tool to help rank job candidates by learning from its past hires, the system mimicked the gender-bias of its makers by [downgrading resumes from women](#).

"We were seeing AI being used extensively before COVID-19, and during COVID-19 you're seeing an increase in the use of some types of tools," noted Meredith Whittaker, a distinguished research scientist at New York University in the US and co-founder of [AI Now](#) Institute, which carries out research examining the social implications of AI.

[Monitoring tools](#) to keep an eye on white collar workers working from home and educational tools that claim to detect whether students are cheating in exams are increasingly growing common. But Whittaker says that most of this technology is untested—and some has been shown to be flawed. However, that hasn't stopped companies from marketing their products as cure-alls for the collateral damage caused by the pandemic,

she adds.

In the US for instance, a compact [medical device](#) called a [pulse oximeter](#), designed to gauge the level of oxygen in the blood, had some coronavirus patients glued to its tiny screens to decide when to go to the hospital, in addition to its use by doctors to aid in [clinical decision making within hospitals](#).

The way the device works, however, is prone to racial bias and was likely calibrated on light skinned users. Back in 2005, a study definitively showed the device 'mostly tended to [overestimate](#) (oxygen) saturation levels by several points' for non-white people.

The problem with the pulse oximeter device has been known for decades and hasn't been fixed by manufacturers, says Whittaker. "But, even so, these tools are being used, they're producing data and that data is going on to shape diagnostic algorithms that are used in health care. And so, you see, even at the level of how our AI systems are constructed, they're encoding the same biases and the same histories of racism and discrimination that are being shown so clearly in the context of COVID-19."

Evidence

Meanwhile, as the body of evidence accumulates that people of [colour are more likely to die from COVID-19 infections](#), that diversity has not necessarily been [reflected in the swathe of clinical trials](#) christened to develop drugs and vaccines—a troubling pattern that has long preceded the pandemic. When it comes to gender diversity, a recent review found that of 927 trials related to COVID-19, more than half explicitly excluded pregnancy, and [pregnant women have been excluded altogether from vaccine trials](#).

The outcomes of products in these clinical trials will not necessarily be representative of the population, notes Catelijne Muller, a member of an EU high-level expert group on AI and co-founder of ALLAI, an organisation dedicated to fostering responsible AI.

"And if you then use those outcomes to feed an AI algorithm for future predictions, those people will also have a disadvantage in these prediction models," she said.

The trouble with use of AI technology in the context of COVID-19 is not different from the issues of bias that plagued the technology before the pandemic: if you feed the technology biased data, it will spout biased outcomes. Indeed, existing large-scale AI systems also reflect the lack of diversity in the environments in which they are built and the people who have built them. These are almost exclusively a handful of technology companies and elite university laboratories – 'spaces that in the West tend to be [extremely white, affluent, technically oriented, and male](#),' according to a 2019 report by the AI Now Institute.

But the technology isn't simply a reflection of its makers—AI also amplifies their biases, says Whittaker.

"One person may have biases, but they don't scale those biases to millions and billions of decisions," she said. "Whereas an AI system can encode human biases and then can distribute those in ways that have a much greater impact."

Complicating matters further, there are automation bias concerns, she adds. "There is a tendency for people to be more trusting of a decision that is made by a computer than they are of the same decision if it were made by a person. So, we need to watch out for the way in which AI systems launder these biases and make them seem rigorous and scientific and may lead to people being less willing to question decisions made by

these systems."

Safe

There is no clear consensus on what will make AI technology responsible and safe en masse, experts say, though researchers are beginning to agree on useful steps such as fairness, interpretability and robustness.

The first step is to ask 'question zero,' according to Muller: what is my problem and how can I solve it? Do I solve it with [artificial intelligence](#) or with something else? If with AI, is this application good enough? Does it harm fundamental rights?

"What we see is that many people think that sometimes AI is sort of a magic wand...and it'll kind of solve everything. But sometimes it doesn't solve anything because it's not fit for the problem. Sometimes it's so invasive that it might solve one problem, but create a large, different problem."

When it comes to using AI in the context of COVID-19—there is an eruption of data, but that data needs to be reliable and be optimised, says Muller.

"Data cannot just be thrown at another algorithm' she said, explaining that algorithms work by finding correlations. "They don't understand what a virus is."

Fairness issues with AI showcase the biases in human decision making, according to Dr. Adrian Weller, programme director for AI at the Alan Turing Institute in the UK. It's wrong to assume that not using algorithms means everything will be just fine, he says.

There is this hope and excitement about these systems because they

operate more consistently and efficiently than humans, but they lack notions of common sense, reasoning and context, where humans are much better, Weller says.

Accountability

Having humans partake more in the [decision-making](#) process is one way to bring accountability to AI applications. But figuring out who that person or persons should be is crucial.

"Simply putting a human somewhere in the process does not guarantee a good decision," said Whittaker. There are issues such as who that human works for and what incentives they're working under which need to be addressed, she says.

"I think we need to really narrow down that broad category of "human" and look at who and to what end."

Human oversight could be incorporated in multiple ways care to ensure transparency and mitigate [bias](#), suggest ALLAI's Muller and colleagues in a [report](#) analysing a [proposal EU regulators](#) are working on to regulate 'high-risk' AI applications such as for use in recruitment, biometric recognition or in the deployment of health.

These include auditing every decision cycle of the AI system, monitoring the operation of the system, having the discretion to decide when and how to use the system in any particular situation, and the opportunity to override a decision made by a system.

For Whittaker, recent developments such as EU regulators' willingness to regulate 'high-risk' applications or community organising in the US leading to bans on facial recognition technology are encouraging.

"I think we need more of the same...to ensure that these systems are auditable, that we can examine them to ensure that they are democratically controlled, and that people have a right to refuse the use of these systems."

Meredith Whittaker and Catelijne Muller will be speaking at a panel to discuss tackling gender and ethnicity biases in artificial intelligence at the [European Research and Innovation Days conference](#) which will take place online from 22-24 September.

Provided by Horizon: The EU Research & Innovation Magazine

Citation: 'Encoding the same biases': Artificial intelligence's limitations in coronavirus response (2020, September 7) retrieved 13 May 2024 from <https://medicalxpress.com/news/2020-09-encoding-biases-artificial-intelligence-limitations.html>

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.