

Data ethicist cautions against overreliance on algorithms

May 26 2022, by Laurel Hamers



Credit: AI-generated image ([disclaimer](#))

Pigeons can quickly be trained to detect cancerous masses on X-ray scans. So can computer algorithms.

But despite the potential efficiencies of outsourcing the task to birds or computers, it's no excuse for getting rid of human radiologists, argues

UO philosopher and data ethicist Ramón Alvarado.

Alvarado studies the way that humans interact with technology. He's particularly attuned to the harms that can come from overreliance on algorithms and machine learning. As automation creeps more and more into people's daily lives, there's a risk that computers devalue human knowledge.

"They're opaque, but we think that because they're doing math, they're better than other knowers," Alvarado said. "The assumption is, the model knows best, and who are you to tell the math they're wrong?"

It's no secret that algorithms built by humans often perpetuate the same biases that went into them. A face-recognition app trained mostly on white faces isn't going to be as accurate on a diverse set of people. Or a resume-ranking tool that awards greater preference to people with Ivy League educations might overlook talented people with more unique but less quantifiable backgrounds.

But Alvarado is interested in a more nuanced question: What if nothing goes wrong, and an algorithm actually is better than a human at a task? Even in these situations, harm can still occur, Alvarado argues in a recent paper published in *Synthese*. It's called "epistemic injustice."

The term was coined by feminist philosopher Miranda Fricker in the 2000s. It's been used to describe benevolent sexism, like men offering assistance to women at the hardware store (a nice gesture) because they presume them to be less competent (a negative motivation). Alvarado has expanded Fricker's framework and applied it to data science.

He points to the impenetrable nature of most modern technology: An algorithm might get the right answer, but we don't know how; that makes it difficult to question the results. Even the scientists who design today's

increasingly sophisticated machine learning algorithms usually can't explain how they work or what the tool is using to reach a decision.

One often-cited study found that a [machine-learning](#) algorithm that correctly distinguished wolves from huskies in photos was not looking at the canines themselves but rather homing in on the presence or absence of snow in the photo background. And since a computer, or a pigeon, can't explain its [thought process](#) the way a human can, letting them take over devalues our own knowledge.

Today, the same sort of algorithm can be used to decide whether or not someone is worthy of an organ transplant or a credit line or a mortgage.

The devaluation of knowledge from relying on such technology can have far-reaching negative consequences. Alvarado cites a high-stakes example: the case of Glenn Rodriguez, a prisoner who was denied parole based on an algorithm that quantified his risk upon release. Despite prison records indicating that he'd been a consistent model for rehabilitation, the algorithm ruled differently.

That produced multiple injustices, Alvarado argues. The first is the algorithm-based decision, which penalized a man who, by all other metrics, had earned parole. But the second, more subtle, injustice is the impenetrable nature of the algorithm itself.

"Opaque technologies are harming [decision-makers](#) themselves, as well as the subjects of decision-making processes, by lowering their status as knowers," Alvarado said. "It's a harm to your dignity because what we know, and what others think we know, is an essential part of how we navigate or are allowed to navigate the world."

Neither Rodriguez, his lawyers, nor even the parole board could access the variables that went into the algorithm that determined his fate, in

order to figure out what was biasing it and challenge its decision. Their own knowledge of Rodriguez's character was overshadowed by an opaque computer program, and their understanding of the computer program was blocked by the corporation that designed the tool. That lack of access is an epistemic injustice.

"In a world with increased decision-making automation, the risks are not just being wronged by an [algorithm](#), but also being left behind as creators and challengers of knowledge," Alvarado said. "As we sit back and enjoy the convenience of these automated systems, we often forget this key aspect of our human experience."

More information: John Symons et al, Epistemic injustice and data science technologies, *Synthese* (2022). [DOI: 10.1007/s11229-022-03631-z](#)

Provided by University of Oregon

Citation: Data ethicist cautions against overreliance on algorithms (2022, May 26) retrieved 13 March 2024 from <https://techxplore.com/news/2022-05-ethicist-cautions-overreliance-algorithms.html>

<p>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.</p>
--