[Provocations](#)

Our Neophobic, Conservative AI Overlords Want Everything to Stay the Same

By [Cory Doctorow](#)01/01/2020

For the [Provocations](#) series, in conjunction with UCI's "[The Future of the Future: The Ethics and Implications of AI](#)" conference.

I've been a technology activist for decades now, and I've read innumerable profound and enduring critiques of technology. In recent years, though, artificial intelligence has come under more fire than most developing trends. The pronouncements, hype, and foolishness surrounding it have risen to heights that stand out even by the outlandish standards of tech absurdity. Like me, you've probably encountered some of the better, smarter critiques along with all the silliness and insanity. Some of the greats are Cathy O'Neil's outstanding 2016 book *Weapons of Math Destruction*, and the excellent research reports from the nonprofit AI Now institute, and also Patrick Ball's spectacular papers published through the essential and dreadfully under-resourced Human Rights Data Analysis Group.

But of all these wonderful, smart, sharp analyses, none has left as enduring an impression as Molly Sauter's odd and lyrical 2017 essay "Instant Recall," published in the online magazine *Real Life*.

Sauter's insight in that essay: machine learning is fundamentally *conservative*, and it hates change. If you start a text message to your partner with "Hey darling," the next time you start typing a message to them, "Hey" will beget an autosuggestion of "darling" as the next word, even if this time you are announcing a break-up. If you type a word or phrase you've never typed before, autosuggest will prompt you with the statistically most common next phrase from all users (I made a small internet storm in July 2018 when I documented autocomplete's suggestion in my message to the family babysitter, which paired "Can you sit" with "on my face and").

This conservativeness permeates every system of algorithmic inference: search for a refrigerator or a pair of shoes and they will follow you around the web as machine learning systems "re-target" you while you move from place to place, even after you've bought the fridge or the shoes. Spend some time researching white nationalism or flat earth conspiracies and all your YouTube recommendations will try to reinforce your "interest." Follow a person on Twitter and you will be inundated with similar people to follow. Machine learning can produce very good accounts of correlation ("this person has that person's address in their address-book and most of the time that means these people are friends") but not causation (which is why Facebook constantly suggests that survivors of stalking follow their tormentors who, naturally, have their targets' addresses in their address books).

Nor is machine learning likely to produce a reliable method of inferring intention: it's a bedrock of anthropology that intention is unknowable without dialogue. As Cliff Geertz points out in his seminal 1973 essay, "Thick Description," you cannot distinguish a "wink" (which means something) from a "twitch" (a meaningless reflex) without asking the person you're observing which one it was.

Ultimately, machine learning is about finding things that are similar to things the machine learning system can already model. Machine learning systems are good at identifying cars that are similar to the cars they already know about. They're also good at identifying faces that are similar to the faces they know about, which is why faces that are white and male are more reliably recognized by these systems — the systems are trained by the people who made them and the people in their circles.

This is what makes machine learning so toxic. If you ask an ML system to predict who the police should arrest, it will suggest that they go and arrest people similar to the ones they've been arresting all along. As the Human Rights Data Analysis Group's Patrick Ball puts it, "A predictive policing system doesn't predict crime, it predicts policing."

But there's a difference between police rounding up the usual suspects on their own, and police doing so because an algorithm told them to: empiricism-washing makes bias seem objective because bias has been quantified. As Congresswoman Alexandria Ocasio-Cortez discovered in 2019 when she gave a Martin Luther King Day speech that described "algorithms" as racially

biased, there is a substantial fraction of people who find this idea risible on its face, because they believe that "math can't be racist."

Empiricism-washing is the top ideological dirty trick of technocrats everywhere: they assert that the data "doesn't lie," and thus all policy prescriptions based on data can be divorced from "politics" and relegated to the realm of "evidence." This sleight of hand pretends that data can tell you what a society wants or needs — when really, data (and its analysis or manipulation) helps you to get what you want.

Think of the UK's recreational drug reclassification exercise under former "drugs czar," the eminent psycho-pharmacologist David Nutt. Asked to assign a risk category to each drug with a recreational use, Nutt convened an expert panel to rank each drug based on how dangerous it was to its users, their families, and society as a whole. Then Nutt went to Parliament and said, "You tell me what your priorities are — whether you're more interested in protecting users, families, or society — and I'll tell you where the drugs go." Empiricism gave Nutt the tools to explain how to categorize drugs based on policy priorities, but those priorities were matters of human judgment, not empirical truth.

Data analysis is as old as censuses of the tax collectors of antiquity — it's as old as the Book of Numbers! — and it is unquestionably useful. But the idea that we should "treasure what we measure" and a reliance on unaccountable black boxes to tell us what we want and how to get it, has delivered to us automated systems of reaction and retreat in the guise of rationalism and progress. The question of what the technology does is important, but far more important is who it is doing it *for* and who it is doing it *to*.

Cory Doctorow is a science fiction author, activist and journalist — the co-editor of Boing Boing and the author of many books, most recently Radicalized *and* Walkaway*, science fiction for adults,* In Real Life*, a graphic novel,* Information Doesn't Want to Be Free*, a book about earning a living in the Internet age, and* Homeland*, a YA sequel to* Little Brother. *His next book is* Poesy the Monster Slayer*, a picture book for young readers.*