



Cold Data is Taxing Your Budget!
Cut Data-Tax Now.

 komprise ^x


[Learn How](#)

Emergent Tech ▶ **Artificial Intelligence**

Sorry, Dave, I can't code that: AI's prejudice problem

The life-changing effects of algorithmic bias

By [Danny Bradbury](#) 8 May 2017 at 09:01

47  [SHARE](#) ▼



Bureaucrats don't just come in uniforms and peaked caps. They come in 1U racks, too. They'll deny you credit, charge you more for paperclips that someone in another neighbourhood, and maybe even stop you getting out of jail. Algorithms are calling the shots these days, and they may not be as impartial as you thought.

To some, algorithmic bias is a growing problem. Among these people: Jeanna Matthews, associate professor of computer science at Clarkson

University and on the executive committee for the Association for Computing Machinery's US Public Policy Council.

Algorithms are increasingly making decisions that have significant personal ramifications, warns Matthews: "When we're making decisions in regulated areas – should someone be hired, lose their job or get credit," she says. "All of those things have huge impacts on the core pieces of peoples' lives."

In some cases, computers will simply spit out a non-transparent score that determines someone's future, she warns. While the algorithms may serve the decision maker's agenda, what happens if it conflicts with the individual's interests?

Complex analytics systems may simply not take human nuances into account, or weigh some data too heavily, she warns. Consequently, [good teachers lose their jobs](#), while 72 per cent of applicants' resumes are [tossed out by software gatekeepers](#) that may have their own biases.

Others think the problem is overblown. Joshua New is a policy analyst at the Center for Data Innovation, an initiative from the Information Technology and Innovation Foundation, a Washington think tank focused on technology policy. New thinks we shouldn't get too alarmist – it isn't as though human beings weren't biased before, and digitising decision making throws this into the spotlight.

"The very process of creating one of these learning systems that can make decisions for you involves quantifying what were originally subjective factors or human processes that never necessarily had to be quantified and analysed before," he argues.

Still, the examples of algorithmic bias keep coming. Advertising networks [served women](#) fewer instances of ads encouraging high-paying jobs. People with names commonly identified as African-American were [shown more online commercials](#) about arrests and criminal records.

If anything, we're not paying these things enough attention, warns Matthews. "It's incorrectly sold as a solution to a problem when in many cases it's making the problem much worse," she says.

Bias ex machine

Bias makes its way into algorithms in several ways. Firstly, it can be intentional. A company may deliberately craft its code to adjust pricing for people [based on where they live](#), say, or how far they are from a rival firm's store.

More commonly, though, questions of bias arise in situations where the code isn't making explicit decisions. In artificial intelligence applications – particularly machine learning and its cousin, deep learning - computers don't follow an explicit set of sequential instructions to arrive at their results. Rather, they process data statistically, reaching conclusions via processes that are often opaque, and defy analysis.

AI gets biased in several ways. One example is “crowding out” by other algorithmic considerations. Young women [were served](#) fewer advertisements for STEM jobs via social media. It didn't happen because a bunch of guys around a green beize table somewhere deliberately wanted to deny them work in science and tech, but because ad pricing algorithms consider them higher-value targets, and make them more expensive to serve ads to. So algorithms designed purely for cost-effectiveness will generally favour cheaper groups.

Crummy data

Bias can also make its way into the data sets used to train AI algorithms. This is a common trope of brogrammer culture. If a young white guy is collating a data set, then the chances are it's going to reflect the world view of a young white guy.

What does this look like in practice? Consider the experience of MIT grad student Joy Buolamwini, who started playing with other people's facial recognition algorithms. In her [Ted Talk on the subject](#), she recalls watching the algorithm recognise every white face in front of the camera. Her black face went unnoticed, presumably because whoever trained the machine forgot to include a broad range of skin tones in the training set. Discrimination by omission is still discrimination. So she started the Algorithmic Justice League to try and redress the imbalance.

Data bias can be incredibly subtle. Matthews recalls a study where algorithms analysed Google News stories for word embedding, which represents text data as vectors. “If you learn about the world from that set, you learn analogies like ‘[man is to computer programmer as woman is to home maker](#)’ (PDF),” Matthews says. Now train an HR system to analyse CVs based on that data, and find out why Jeffrey can code, and Jenny isn’t allowed to.

Scummy data

Frank Pasquale, professor of law at the University of Maryland’s Francis King Carey School of Law, argues that sometimes malicious players deliberately tinker with datasets, targeting algorithms designed to learn as they go.

“It’s worth highlighting on its own, especially with regards to the Holocaust results that have been so controversial over the past few months at Google,” he says, referring to hate groups that [gamed search results](#) to get their own sites on the radar.

Let’s not forget Microsoft’s Tay, either, a digital AI teen who was [corrupted by online trolls](#).

The other way that algorithms can be biased is by being accurate. We’re a racist, sexist society. Use a data set with as much veracity as possible, and you may reflect that real world too well, surfacing biases that already exist, but shouldn’t.

That can lead to some serious problems. Some people may not get out of jail because a computer decides not to let them.

The investigative journalism organization ProPublica [explored](#) COMPAS, a machine learning system produced by Northpointe Inc, which courts in the US are using to predict recidivism rates in felons. Judges take the software’s musings into account when deciding whether to parole people or not.

The software tended to predict higher recidivism rates along racial lines, said the ProPublica investigation. It alleged that the software was terrible

at predicting who would commit violent crimes, getting only 20 per cent of its risk assessments right. Unfortunately, it also tended to think that black people would be more likely to re-offend.

Northpointe's researchers were busy beard-stroking and couldn't talk to us, the company's PR told *The Reg*, but they pointed us to a [response](#) to the Propublica story defending its software. In summary, it argued that risk scores varied between groups because real-world crime rates did too.

ProPublica published a [riposte](#). Its bottom line: sure, black defendants are more likely to be arrested for new crimes, but Northpointe's scores were still error prone, it said. This meant that the most likely error for a black person was to be incorrectly classified as a higher risk, whereas whites were more likely to be incorrectly classified as lower risk. You can't just look at top-level numbers when individual rights and freedoms are at stake, ProPublica argued.

This isn't the only allegation of bias in the law enforcement and justice system. Predictive policing system PredPol has also been [accused](#) of amplifying racially-biased policing.

Data transparency

It's difficult to truly assess the bias in algorithms without understanding how they reach their decisions. One option is to make algorithmic decisions as transparent as possible, by exposing how they do what they do.

Transparent, as in simply publishing the algorithm? New warns against taking this too far. Making the algorithms or training sets open source might be commercially sensitive, he warns. In any case: "Opening up the hood to the public won't solve anything. People won't know what's going on."

Matthews agrees that we're going to need new skill sets to cope with algorithmic bias. She calls for a kind of algorithmic watchdog, with "investigative data scientists." We can see the Hollywood trailer now.

Who would govern it, though? In the US, the FTC would be one obvious pick. It handles privacy complaints, and tends to settle them on an ad hoc basis. Its chief technologist Ashkan Soltani considers “algorithmic accountability” a [hot-button issue](#).

With a new administration that [downplays civil rights in favour of the right to do business](#), though, it might take some heavy lifting for it to get on the federal radar any time soon.

Things look more positive in Europe, where policy makers are on top of the transparency issue. The General Data Protection Regulation (GDPR), the Europe-wide privacy legislation that replaces the Data Protection Directive, has language devoted to it.

Recital 71 in the Regulation, which we like to call the “computer says no” clause, says that people have a right not to blindly accept decisions with legal effects made entirely by computers with no human intervention. They have a right to contest the decision and express their point of view to a human being. The Recital also calls for transparency of processing.

Some machine learning algorithms, such as decision trees, can show you how they reached their results. Others are less tractable. Neural networks are excellent at narrow AI tasks, but can't articulate how they did it.

No really, what were you thinking?

Some researchers are exploring the idea of “explainer” technology that will help show how an algorithm reached a decision. The University of Washington's [LIME project \(PDF\)](#) is an example. Researchers at the University of Utah have also [proposed](#) a system to find bias in algorithms and remedy them, by blurring demographic attributes in the data.

There's another problem muddying the waters, though: what *is* bias? Understanding how to code fairness suggests that we can define it, and that creates ethical concerns.

“If your AI talks about the most beautiful people in the world, but most of them are white, then that seems to not really reflect the nature of the

world's population," Pasquale says. "That disparate impact measure may be the only way to measure the fairness of some of these systems."

The problem is that many algorithmic outputs are more nuanced and three dimensional than deciding the results of a beauty contest. Deciding who gets credit can depend on many factors. How can you determine what's fair to consider, and how much weight it should have? And who gets to decide those rules?

"You can say you want them to be as effective as possible, statistically speaking for society, for decision makers, and there's a good reason to want that," muses Matthews. "But you can also say that it's important to protect the individuals. To protect the outliers. To protect the exceptions." It is okay for a deep learning algorithm to refuse someone credit or a job because their lifestyle doesn't fit the norm in the statistical model, even if closer scrutiny suggests the person isn't a credit risk? It depends whose needs the algorithm is meant to serve – that of the individual or that of the decision maker.

This isn't as esoteric as it might sound. Outliers take all forms, and in some cases they may suffer from little more than invisibility. That's a phenomenon, and it's growing, points out New.

"It's a term called 'data poverty'. It's a logical extension to the digital divide," he warns.

Some communities don't have as much data about them as others. Hispanics in the US typically have ["thin" credit ratings](#). Are you in a low-income household and therefore [statistically less likely](#) to have internet access, and make a digital footprint? Don't have a smart home for people to gather data on? "It means you're not represented in the data set," says New.

With algorithms increasingly making key decisions about our lives, it's important not only to be properly represented in the data they're considering, but to understand how they're reaching their conclusions. More important still, is the ability to inquire should you fall through the

algorithmic net. Software may be eating the world, but that doesn't mean it gets to chew you up and spit you out. ®

Sponsored: [Your Guide to Becoming Truly Data-Driven with Unrivalled Data Analytics Performance](#)

Tips and corrections

47 Comments

 **Sign up to our Newsletter** - Get IT in your inbox daily

MORE [Algorithm](#) [Machine Learning](#)





SIGN UP TO OUR WEEKLY NEWSLETTER

[SIGN UP HERE](#)

You Might Also Like

Recommended by  **Outbrain** | 



Sponsored

Potsdam: If You Aren't Doing This on
Honey

Sponsored

[Pics] Husband Introduces Wife To The Wild
Ice Pop

Sponsored

Taylor 214ce Deluxe - Sunburst with
SWEETWATER

Guy is booted out of IT amid outsourcing,

This is the final straw, evil Microsoft.

Sponsored

All Star Billionaires: Bill Maher is One of
Investing.com

Sponsored

[Pics] Inside Dwayne 'The Rock'
VT

IBM HR made me lie to US govt, says axed

Whitepapers



Rethinking Your Infrastructure for Enterprise AI

This white paper discusses these challenges and looks at how IBM proposes to help businesses overcome them.



Using IBM Cloud Object Storage with Komprise Data Management to Curb NAS expansion

Using Komprise data management software with IBM Cloud Object Storage



2019 Ponemon Report: The Value of Threat Intelligence from Anomali

The Ponemon Institute Surveyed Over 1,000 Security Professionals In the United States and the United Kingdom.



Application Monitoring in the Enterprise

What's the big deal? Why is enterprise monitoring so hard?

More from The Register



Have you always wanted an algorithm that can search like Bing? Well, if you change your mind, one's on GitHub now

Make your app answer all the easy questions, like, 'Where can I download Chrome?'



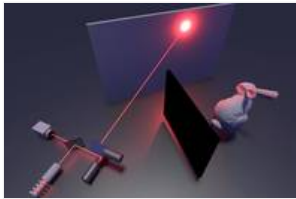
AI solves Rubik's Cube in 1.2 seconds (that's three times slower than a non-AI algorithm)

Even so, human nerds are left in the dust and this neural net can be used for other tricks



Europe seeks company to monitor Google's algorithm in €10m deal

Follows €2.4bn mega fine



New algorithm could help self-driving cars scout out hidden objects

If they are going v-e-r-y slowly



Amazing new algorithm makes fusion power slightly less incredibly inefficient

ANALYSIS Cool, er, hot stuff but no breakthrough (yet)



Machines taught how to 'smell' by new algorithm. How will they cope with shower-dodging nerds?

SIGSTINK: User body odor fault

Sponsored links

[Get The Register's Headlines in your inbox daily - quick signup!](#)

SIGN UP TO OUR WEEKLY NEWSLETTER

SIGN UP HERE

[ABOUT US](#)



[MORE CONTENT](#)



[SITUATION PUBLISHING](#)





The Register - Independent news and views for the tech community. Part of Situation Publishing

SIGN UP TO OUR NEWSLETTERS

Join our daily or weekly newsletters, subscribe to a specific section or set [News alerts](#)

SUBSCRIBE >



Biting the hand that feeds IT © 1998–2019

[Cookies](#)

[Privacy](#)

[Ts&Cs](#)