We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information

Business Impact

Inspecting Algorithms for Bias

Courts, banks, and other institutions are using automated data analysis systems to make decisions about your life. Let's not leave it up to the algorithm makers to decide whether they're doing it appropriately.

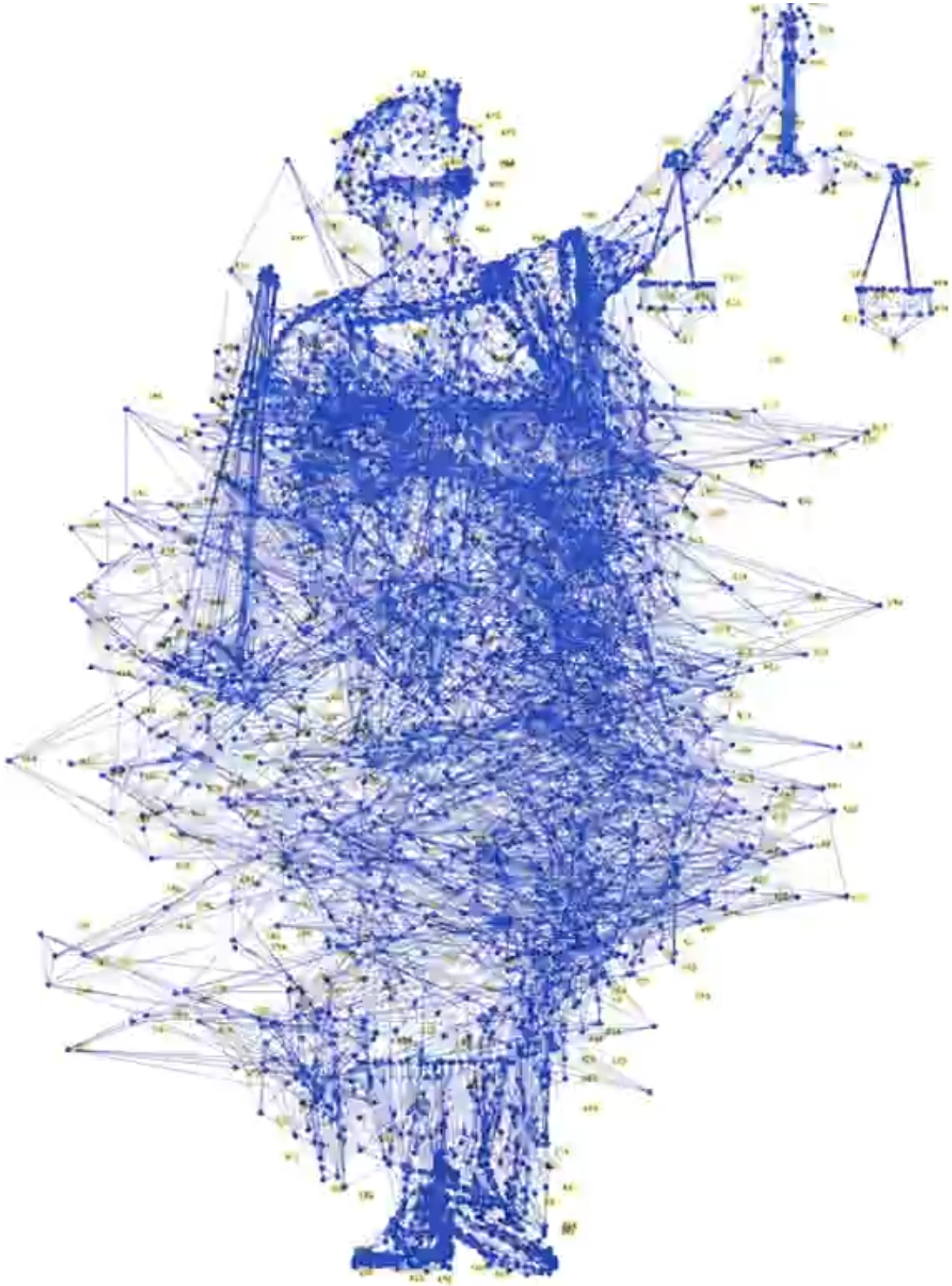
by **Matthias Spielkamp** Jun 12, 2017

You've read **2/3** of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in

We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information



You've read **2/3** of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in

We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information

ProPublica, a Pulitzer Prize–winning nonprofit news organization, had analyzed risk assessment software known as COMPAS. It is being used to forecast which criminals are most likely to reoffend. Guided by such forecasts, judges in courtrooms throughout the United States make decisions about the future of defendants and convicts, determining everything from bail amounts to sentences. When ProPublica compared COMPAS's risk assessments for more than 10,000 people arrested in one Florida county with how often those people actually went on to reoffend, it discovered that the algorithm "correctly predicted recidivism for black and white defendants at roughly the same rate." But when the algorithm was wrong, it was wrong in different ways for blacks and whites. Specifically, "blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend." And COMPAS tended to make the opposite mistake with whites: "They are much more likely than blacks to be labeled lower risk but go on to commit other crimes."

This story is part of our July/August 2017 issue

See the rest of the issue ›

Subscribe ›

Whether it's appropriate to use systems like COMPAS is a question that goes beyond racial bias. The U.S. Supreme Court might soon take up the case of a Wisconsin convict who says his right to due process was violated when the judge who sentenced him consulted COMPAS, because the workings of the system were opaque to the defendant. Potential problems with other automated decision-making (ADM) systems exist outside the justice system, too. On the basis of online personality tests, ADMs are helping to determine whether someone is the right person for a job. Credit-scoring algorithms play an enormous role in whether you get a mortgage, a credit card, or even the most cost-effective cell-phone deals.

Things reviewed

"Machine Bias"
ProPublica, May 23, 2016

"COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity"
Northpointe, July 8, 2016

"Technical Response to Northpointe"
ProPublica, July 29, 2016

"False Positives, False Negatives, and False Analyses: A Rejoinder to 'Machine Bias'"
Anthony Flores, Christopher Lowenkamp, and Kristin Bechtel
August 10, 2016

It's not necessarily a bad idea to use risk assessment systems like COMPAS. In many cases, ADM systems can increase fairness. Human decision making is at times so incoherent that it needs oversight to bring it in line with our standards of justice. As one specifically unsettling study showed, parole boards were more likely to free convicts if the judges had just had a meal break. This

You've read **2/3** of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in

We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information

clear even to the systems' designers, it's not necessarily possible to determine which algorithms are biased and which ones are not. And even when the answer seems clear, as in ProPublica's findings on COMPAS, the truth is sometimes more complicated.

Lawmakers, the courts, and an informed public should decide what we want algorithms to prioritize.

What should we do to get a better handle on ADMs? Democratic societies need more oversight over such systems than they have now. AlgorithmWatch, a Berlin-based nonprofit advocacy organization that I cofounded with a computer scientist, a legal philosopher, and a fellow journalist, aims to help people understand the effects of such systems. "The fact that most ADM procedures are black boxes to the people affected by them is not a law of nature. It must end," we assert in our manifesto. Still, our take on the issue is different from many critics'—because our fear is that the technology could be demonized undeservedly. What's important is that societies, and not only algorithm makers, make the value judgments that go into ADMs.

Measures of fairness

COMPAS determines its risk scores from answers to a questionnaire that explores a defendant's criminal history and attitudes about crime. Does this produce biased results?

After ProPublica's investigation, Northpointe, the company that developed COMPAS, disputed the story, arguing that the journalists misinterpreted the data. So did three criminal-justice researchers, including one from a justice-reform organization. Who's right—the reporters or the researchers? Krishna Gummadi, head of the Networked Systems Research Group at the Max Planck Institute for Software Systems in Saarbrücken, Germany, offers a surprising answer: they all are.

Gummadi, who has extensively researched fairness in algorithms, says ProPublica's and Northpointe's results don't contradict each other. They differ because they use different measures of fairness.

If used properly, criminal-justice algorithms offer "the chance of a generation, and perhaps a lifetime, to reform

You've read 2/3 of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in

We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information

One problem with this approach is that it tends to increase the number of false positives: people who will be unjustly classified as likely reoffenders. The dial can be adjusted to deliver as few false positives as possible, but that tends to create more false negatives: likely reoffenders who slip through and get a more lenient treatment than warranted.

Raising the incidence of true positives or lowering the false positives are both ways to improve a statistical measure known as positive predictive value, or PPV. That is the percentage of all positives that are true.

As Gummadi points out, ProPublica compared false positive rates and false negative rates for blacks and whites and found them to be skewed in favor of whites. Northpointe, in contrast, compared the PPVs for different races and found them to be similar. In part because the recidivism rates for blacks and whites do in fact differ, it is mathematically likely that the positive predictive values for people in each group will be similar while the rates of false negatives are not.

One thing this tells us is that the broader society—lawmakers, the courts, an informed public—should decide what we want such algorithms to prioritize. Are we primarily interested in taking as few chances as possible that someone will skip bail or reoffend? What trade-offs should we make to ensure justice and lower the massive social costs of imprisonment?

No matter which way the dials are set, any algorithm will have biases—it is, after all, making a prediction based on generalized statistics, not on someone's individual situation. But we can still use such systems to guide decisions that are wiser and fairer than the ones humans tend to make on their own.

The controversy surrounding the New York Police Department's stop-and-frisk practices helps to show why. Between January 2004 and June 2012, New York City police conducted 4.4 million stops under a program that allowed officers to temporarily detain, question, and search people on the street for weapons and other contraband. But in fact, "88 percent of the 4.4 million stops resulted in no further action—meaning a vast majority of those stopped were doing nothing wrong," the *New York Times* said in an editorial decrying the practice. What's more: "In about 83 percent of cases, the person stopped was black or Hispanic, even though the two groups accounted for just over half the population." This example of human bias, illuminated through data analysis, is a reminder that ADM systems could play a positive role in criminal

You've read **2/3** of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in

We use cookies to offer you a better browsing experience, analyze site traffic, personalize content, and serve targeted advertisements. Read about how we use cookies by clicking "Cookie Information." If you continue to use this site, you consent to our use of cookies.

› Cookie Information

But if we accept that algorithms might make life fairer if they are well designed, how can we know whether they are so designed?

Democratic societies should be working now to determine how much transparency they expect from ADM systems. Do we need new regulations of the software to ensure it can be properly inspected? Lawmakers, judges, and the public should have a say in which measures of fairness get prioritized by algorithms. But if the algorithms don't actually reflect these value judgments, who will be held accountable?

These are the hard questions we need to answer if we expect to benefit from advances in algorithmic technology.

Matthias Spielkamp is executive director of AlgorithmWatch, an advocacy group that analyzes the risks and opportunities of automated decision making.

Share Link

Tagged Artificial Intelligence

Author Matthias Spielkamp

Popular

You've read 2/3 of your free monthly feature stories. Subscribe for unlimited access.

Subscribe now Sign in