



Featured Topics Newsletters Events Podcasts

SIGN IN

SUBSCRIBE

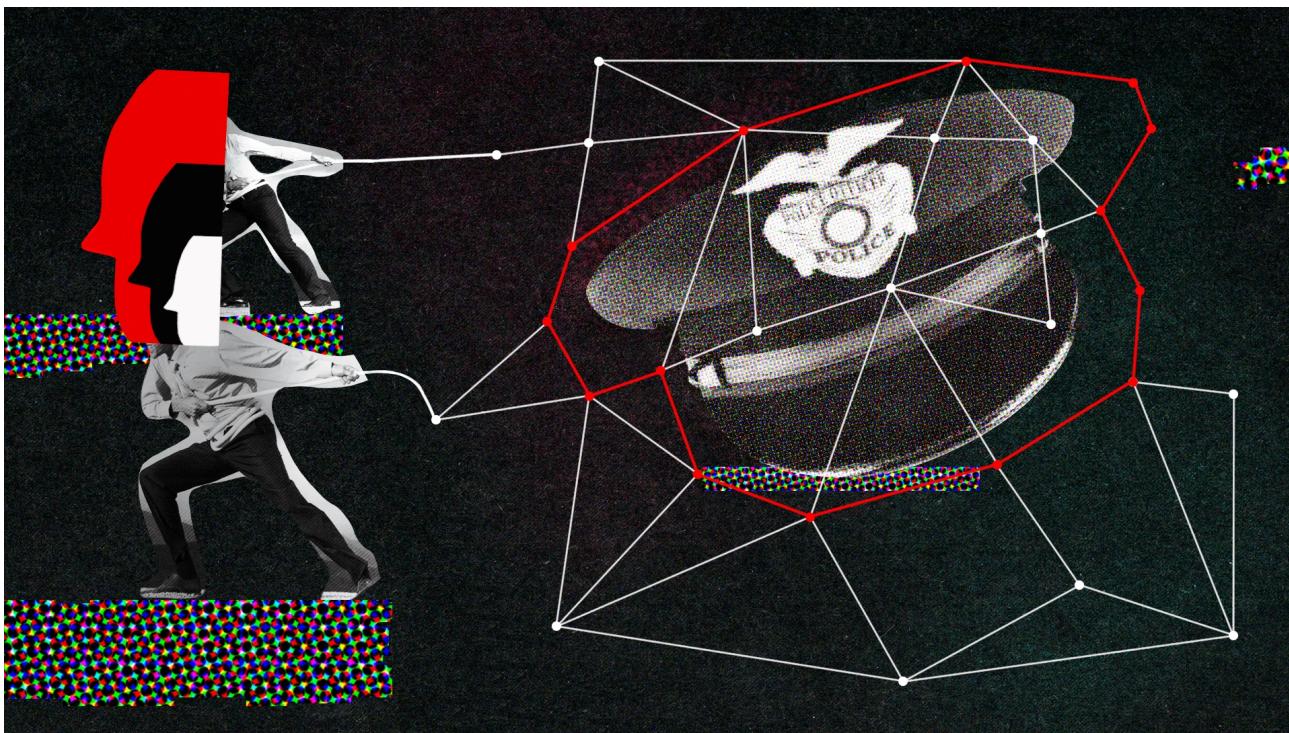
ARTIFICIAL INTELLIGENCE

Predictive policing algorithms are racist. They need to be dismantled.

Lack of transparency and biased training data mean these tools are not fit for purpose. If we can't fix them, we should ditch them.

By Will Douglas Heaven

July 17, 2020



FRANZISKA BARCZYK

Yeshimabeit Milner was in high school the first time she saw kids she knew getting handcuffed and stuffed into police cars. It was February 29, 2008, and the principal of a nearby school in Miami, with a majority Haitian and African-American population, had put one of his students in a chokehold. The next day several dozen kids staged a peaceful demonstration. It didn't go well.

That night, Miami's NBC 6 News at Six kicked off with a segment called "Chaos on Campus." (There's a [clip on YouTube](#).) "Tensions run high at Edison Senior High after a fight for rights ends in a battle with the law," the broadcast said. Cut to blurry phone footage of screaming teenagers: "The chaos you see is an all-out brawl inside the school's cafeteria."

Students told reporters that police hit them with batons, threw them on the floor, and pushed them up against walls. The police claimed they were the ones getting attacked—"with water bottles, soda pops, milk, and so on"—and called for emergency backup. Around 25 students were arrested, and many were charged with multiple crimes, including resisting arrest with violence. Milner remembers watching on TV and seeing kids she'd gone to elementary school with being taken into custody. "It was so crazy," she says.

For Milner, the events of that day and the long-term implications for those arrested were pivotal. Soon after, while still at school, she got involved with data-based activism, documenting fellow students' experiences of racist policing. She is now the

"There's a long history of data being weaponized

against Black communities."

director of [Data for Black Lives](#), a grassroots digital rights organization she cofounded in 2017. What she learned as a teenager pushed her into a life of fighting back against [bias in the criminal justice system](#) and dismantling what she calls the school-to-prison pipeline. "There's a long history of data being weaponized against Black communities," she says.

X

THE SPARK

A weekly newsletter explaining the tech combating the climate crisis.

Enter your email

 Sign up

By signing up, you agree to our [Privacy Policy](#).

Inequality and the misuses of police power don't just [play out on the streets](#) or during school riots. For Milner and other activists, the focus is now on where there is most potential for long-lasting damage: predictive policing tools and the abuse of data by police forces. A number of studies have shown that these tools [perpetuate systemic racism](#), and yet we still know very little about how they work, who is using them, and for what purpose. All of this needs to change before a proper reckoning can take place. Luckily, the tide may be turning.

There are two broad types of predictive policing tool. Location-based algorithms draw on links between places, events, and historical crime rates to predict where and when crimes are more likely to happen—for example, in certain weather conditions or at large sporting events. The tools identify hot spots, and the police plan patrols around these tip-offs. One of the most common, called PredPol, which is used by [dozens of cities in the US](#), breaks locations up into 500-by-500 foot blocks, and updates its predictions throughout the day—a kind of crime weather forecast.

Other tools draw on data about people, such as their age, gender, marital status, history of substance abuse, and criminal record, to predict who has a high chance of being involved in future criminal activity. These person-based tools can be used either by police, to intervene before a crime takes place, or by courts, to determine during pretrial hearings or sentencing whether someone who has been arrested is likely to reoffend. For example, a tool called COMPAS, used in many jurisdictions to help make decisions about pretrial release and sentencing, issues a statistical score between 1 and 10 to quantify how likely a person is to be rearrested if released.

The problem lies with the data the algorithms feed upon. For one thing, predictive algorithms are easily skewed by arrest rates. According to US Department of Justice figures, you are [more than twice as likely to be](#)

arrested if you are Black than if you are white. A Black person is five times as likely to be stopped without just cause as a white person. The mass arrest at Edison Senior High was just one example of a type of disproportionate police response that is not uncommon in Black communities.

The kids Milner watched being arrested were being set up for a lifetime of biased assessment because of that arrest record. But it wasn't just their own lives that were affected that day. The data generated by their arrests would have been fed into algorithms that would disproportionately target all young Black people the algorithms assessed. Though by law the algorithms do not use race as a predictor, other variables, such as socioeconomic background, education, and zip code, act as proxies. Even without explicitly considering race, these tools are racist.

That's why, for many, the very concept of predictive policing itself is the problem. The writer and academic Dorothy Roberts, who studies law and social rights at the University of Pennsylvania, put it well in an online panel discussion in June. "Racism has always been about predicting, about making certain racial groups seem as if they are predisposed to do bad things and therefore justify controlling them," she said.

Risk assessments have been part of the criminal justice system for decades. But police departments and courts have made more use of automated tools in the last few years, for two main reasons. First, budget cuts have led to an efficiency drive. "People are calling to defund the police, but they've already been defunded," says Milner. "Cities have been going broke for years, and they've been replacing cops with algorithms." Exact figures are hard to come by, but predictive tools are thought to be used by police forces or courts in most US states.

The second reason for the increased use of algorithms is the widespread belief that they are more objective than humans: they were first introduced to make decision-making in the criminal justice system more fair. Starting in the 1990s, early automated techniques used rule-based decision trees, but today prediction is done with machine learning.



CLAY BANKS VIA UNSPLASH

Yet increasing evidence suggests that human prejudices have been baked into these tools because the machine-learning models are trained on biased police data. Far from avoiding racism, they may simply be better at hiding it. Many critics now view these tools as a form of tech-washing, where a veneer of objectivity covers mechanisms that perpetuate inequities in society.

“It’s really just in the past few years that people’s views of these tools have shifted from being something that might alleviate bias to something that might entrench it,” says Alice Xiang, a lawyer and data scientist who leads research into fairness, transparency and accountability at the Partnership on AI. These biases have been compounded since the first generation of prediction tools appeared 20 or 30 years ago. “We took bad data in the first place, and then we used tools to make it worse,” says Katy Weathington, who studies algorithmic bias at the University of Colorado Boulder. “It’s just been a self-reinforcing loop over and over again.”



The Accessibility issue

Subscribe for full access the Accessibility Issue and learn how technology can build toward a more inclusive future.

SUBSCRIBE NOW

Things might be getting worse. In the wake of the protests about police bias after the death of George Floyd at the hands of a police officer in Minneapolis, some police departments are doubling down on their use of predictive tools. A month ago, New York Police Department commissioner

Dermot Shea sent a letter to his officers. “In the current climate, we have to fight crime differently,” he wrote. “We will do it with less street-stops—perhaps exposing you to less danger and liability—while better utilizing data, intelligence, and all the technology at our disposal ... That means for the NYPD’s part, we’ll redouble our precision-policing efforts.”

Police like the idea of tools that give them a heads-up and allow them to intervene early because they think it keeps crime rates down, says Rashida Richardson, director of policy research at the AI Now Institute. But in practice, their use can feel like harassment. Researchers have found that some police departments give officers “most wanted” lists of people the tool identifies as high risk. This first came to light when people in Chicago reported that police had been knocking on their doors and telling them they were being watched. In other states, says Richardson, police were warning people on the lists that they were at high risk of being involved in gang-related crime and asking them to take actions to avoid this. If they were later arrested for any type of crime, prosecutors used the prior warning to seek higher charges. “It’s almost like a digital form of entrapment, where you give people some vague information and then hold it against them,” she says.

Similarly, studies—including one commissioned by the UK government’s Centre for Data Ethics and Innovation last year—suggest that identifying certain areas as hot spots primes officers to expect trouble when on patrol, making them more likely to stop or arrest people there because of prejudice rather than need.

“It’s almost like a digital form of entrapment.”

Another problem with the algorithms is that many were trained on white populations outside the US, partly because criminal records are hard to get hold of across different US jurisdictions. Static 99, a tool designed to predict recidivism among sex offenders, was trained in Canada, where only around 3% of the population is Black compared with 12% in the US. Several other tools used in the US were developed in Europe, where 2% of the population is Black. Because of the differences in socioeconomic conditions between countries and populations, the tools are likely to be less accurate in places where they were not trained. Moreover, some pretrial algorithms trained many years ago still use predictors that are out of date. For example, some still predict that a defendant who doesn’t have a landline phone is less likely to show up in court.

But do these tools work, even if imperfectly? It depends what you mean by “work.” In general it is practically impossible to disentangle the use of predictive policing tools from other factors that affect crime or incarceration rates. Still, a handful of small studies have drawn limited conclusions. Some show signs that courts’ use of risk assessment tools has had a minor positive impact. A 2016 study of a machine-learning tool used in Pennsylvania to inform parole decisions found no evidence that it jeopardized public safety

(that is, it correctly identified high-risk individuals who shouldn't be paroled) and some evidence that it identified nonviolent people who could be safely released.

Another study, in 2018, looked at a tool used by the courts in Kentucky and found that although risk scores were being interpreted inconsistently between counties, which led to discrepancies in who was and was not released, the tool would have slightly reduced incarceration rates if it had been used properly. And the American Civil Liberties Union reports that an assessment tool adopted as part of the 2017 New Jersey Criminal Justice Reform Act led to a 20% decline in the number of people jailed while awaiting trial.



Rashida Richardson is director of policy research at the AI Now Institute. She previously led work on the legal issues around privacy and surveillance at the American Civil Liberties Union.

COURTESY OF AI NOW

Advocates of such tools say that algorithms can be more fair than human decision makers, or at least make unfairness explicit. In many cases, especially at pretrial bail hearings, judges are expected to rush through many dozens of cases in a short time. In one study of pretrial hearings in Cook County, Illinois, researchers found that judges spent an average of just 30 seconds considering each case.

In such conditions, it is reasonable to assume that judges are making snap decisions driven at least in part by their personal biases. Melissa Hamilton at the University of Surrey in the UK, who studies legal issues around risk assessment tools, is critical of their use in practice but believes they can do a better job than people in principle. “The alternative is a human decision maker’s black-box brain,” she says.

But there is an obvious problem. The arrest data used to train predictive tools does not give an accurate picture of criminal activity. Arrest data is used because it is what police departments record. But arrests do not necessarily lead to convictions. “We’re trying to measure people committing crimes, but all we have is data on arrests,” says Xiang.

"We're trying to measure people committing crimes, but all we have is data on arrests."

What's more, arrest data encodes patterns of racist policing behavior. As a result, they're more likely to predict a high potential for crime in minority neighborhoods or among minority people. Even when arrest and crime data match up, there are a myriad of socioeconomic reasons why certain populations and certain neighborhoods have higher historical crime rates than others. Feeding this data into predictive tools allows the past to shape the future.



The Accessibility issue

Subscribe for full access the Accessibility Issue and learn how technology can build toward a more inclusive future.

[SUBSCRIBE NOW](#)

Some tools also use data on where a call to police has been made, which is an even weaker reflection of actual crime patterns than arrest data, and one even more warped by racist motivations. Consider the case of Amy Cooper, who called the police simply because a Black bird-watcher, Christian Cooper, asked her to put her dog on a leash in New York's Central Park.

"Just because there's a call that a crime occurred doesn't mean a crime actually occurred," says Richardson. "If the call becomes a data point to justify dispatching police to a specific neighborhood, or even to target a specific individual, you get a feedback loop where data-driven technologies legitimize discriminatory policing."

As more critics argue that these tools are not fit for purpose, there are calls for a kind of algorithmic affirmative action, in which the bias in the data is counterbalanced in some way. One way to do this for risk assessment algorithms, in theory, would be to use differential risk thresholds—three arrests for a Black person could indicate the same level of risk as, say, two arrests for a white person.

This was one of the approaches examined in a study published in May by Jennifer Skeem, who studies public policy at the University of California, Berkeley, and Christopher Lowenkamp, a social science analyst at the Administrative Office of the US Courts in Washington, DC. The pair looked at three different options for removing the bias in algorithms that had assessed the risk of recidivism for around 68,000 participants, half white and half Black. They found that the best balance between races was achieved when algorithms took race explicitly into account—which existing tools are legally forbidden from doing—and assigned Black people a higher threshold than whites for being deemed high risk.

Of course, this idea is pretty controversial. It means essentially manipulating the data in order to forgive some proportion of crimes because of the perpetrator's race, says Xiang: "That is something that makes people very uncomfortable." The idea of holding members of different groups to different standards goes against many people's sense of fairness, even if it's done in a way that's supposed to address historical injustice. (You can try out this

trade-off for yourself in our [interactive story on algorithmic bias in the criminal legal system](#), which lets you experiment with a simplified version of the COMPAS tool.)

At any rate, the US legal system is not ready to have such a discussion. “The legal profession has been way behind the ball on these risk assessment tools,” says Hamilton. In the last few years she has been giving training courses to lawyers and found that defense attorneys are often not even aware that their clients are being assessed in this way. “If you’re not aware of it, you’re not going to be challenging it,” she says.

The lack of awareness can be blamed on the murkiness of the overall picture: law enforcement has been so tight-lipped about how it uses these technologies that it’s very hard for anyone to assess how well they work. Even when information is available, it is hard to link any one system to any one outcome. And the few detailed studies that have been done focus on specific tools and draw conclusions that may not apply to other systems or jurisdictions.

It is not even clear what tools are being used and who is using them. “We don’t know how many police departments have used, or are currently using, predictive policing,” says Richardson.

For example, the fact that police in New Orleans were using a predictive tool developed by secretive data-mining firm Palantir came to light only after an [investigation by The Verge](#). And public records show that the [New York Police Department has paid \\$2.5 million to Palantir](#) but isn’t saying what for.



GETTY

Most tools are licensed to police departments by a ragtag mix of small firms, state authorities, and researchers. Some are proprietary systems; some aren’t. They all work in slightly different ways. On the basis of the tools’ outputs, researchers re-create as well as they can what they believe is going

on.

THE DOWNLOAD

Sign up for your daily dose of what's up in emerging technology.

Enter your email

 Signup

By signing up, you agree to our [Privacy Policy](#).

Hamid Khan, an activist who fought for years to get the Los Angeles police to drop a predictive tool called PredPol, demanded an audit of the tool by the police department's inspector general. [According to Khan](#), in March 2019 the inspector general said that the task was impossible because the tool was so complicated.

In the UK, Hamilton tried to look into a tool called OASys, which—like COMPAS—is commonly used in pretrial hearings, sentencing, and parole. The company that makes OASys does its own audits and has not released much information about how it works, says Hamilton. She has repeatedly tried to get information from the developers, but they stopped responding to her requests. She says, “I think they looked up my studies and decided: Nope.”

The familiar refrain from companies that make these tools is that they cannot share information because it would be giving up trade secrets or confidential information about people the tools have assessed.

All this means that only a handful have been studied in any detail, though some information is available about a few of them. Static 99 was developed by a group of data scientists who shared details about its algorithms. Public Safety Assessment, one of the most common pretrial risk assessment tools in the US, was originally developed by Arnold Ventures, a private organization, but it turned out to be easier to convince jurisdictions to adopt it if some details about how it worked were revealed, says Hamilton. Still, the makers of both tools have refused to release the data sets they used for training, which would be needed to fully understand how they work.

Not only is there little insight into the mechanisms inside these tools, but critics say police departments and courts are not doing enough to make sure they buy tools that function as expected. For the NYPD, buying a risk assessment tool is subject to the same regulations as buying a snow plow, says Milner.

Buying a risk assessment tool is subject to the same regulations as buying a snow plow.

“Police are able to go full speed into buying tech

without knowing what they're using, not investing time to ensure that it can be used safely," says Richardson. "And then there's no ongoing audit or analysis to determine if it's even working."

Efforts to change this have faced resistance. Last month New York City passed the Public Oversight of Surveillance Technology (POST) Act, which requires the NYPD to list all its surveillance technologies and describe how they affect the city's residents. The NYPD is the biggest police force in the US, and proponents of the bill hope that the disclosure will also shed light on what tech other police departments in the country are using. But getting this far was hard. Richardson, who did advocacy work on the bill, had been watching it sit in limbo since 2017, until widespread calls for policing reform in the last few months tipped the balance of opinion.

It was frustration at trying to find basic information about digital policing practices in New York that led Richardson to work on the bill. Police had resisted when she and her colleagues wanted to learn more about the NYPD's use of surveillance tools. Freedom of Information Act requests and litigation by the New York Civil Liberties Union weren't working. In 2015, with the help of city council member Daniel Garodnik, they proposed legislation that would force the issue.

"We experienced significant backlash from the NYPD, including a nasty PR campaign suggesting that the bill was giving the map of the city to terrorists," says Richardson. "There was no support from the mayor and a hostile city council."

With its ethical problems and lack of transparency, the current state of predictive policing is a mess. But what can be done about it? Xiang and Hamilton think algorithmic tools have the potential to be fairer than humans, as long as everybody involved in developing and using them is fully aware of their limitations and deliberately works to make them fair.

But this challenge is not merely a technical one. A reckoning is needed about what to do about bias in the data, because that is there to stay. "It carries with it the scars of generations of policing," says Weathington.

And what it means to have a fair algorithm is not something computer scientists can answer, says Xiang. "It's not really something anyone can answer. It's asking what a fair criminal justice system would look like. Even if you're a lawyer, even if you are an ethicist, you cannot provide one firm answer to that."

"These are fundamental questions that are not going to be solvable in the sense that a mathematical problem can be solvable," she adds.

Hamilton agrees. Civil rights groups have a hard choice to make, she says: "If you're against risk assessment, more minorities are probably going to remain locked up. If you accept risk assessment, you're kind of complicit with

these tools. For example, the NYPD has been involved in dozens of lawsuits concerning years of biased policing. “I don’t understand how you can be actively dealing with settlement negotiations concerning racially biased practices and still think that data resulting from those practices is okay to use,” she says.

For Milner, the key to bringing about change is to involve the people most affected. In 2008, after watching those kids she knew get arrested, Milner joined an organization that surveyed around 600 young people about their experiences with arrests and police brutality in schools, and then turned what she learned into a comic book. Young people around the country used the comic book to start doing similar work where they lived.



Yeshimabeit Milner is co-founder and director of Data for Black Lives, a grassroots collective of activists and computer scientists using data to reform the criminal justice system.

COURTESY OF DATA FOR BLACK LIVES

Today her organization, Data for Black Lives, coordinates around 4,000 software engineers, mathematicians, and activists in universities and community hubs. Risk assessment tools are not the only way the misuse of data perpetuates systemic racism, but it’s one very much in their sights. “We’re not going to stop every single private company from developing risk assessment tools, but we can change the culture and educate people, give them ways to push back,” says Milner. In Atlanta they are training people who have spent time in jail to do data science, so that they can play a part in reforming the technologies used by the criminal justice system.

In the meantime, Milner, Weathington, Richardson, and others think police should stop using flawed predictive tools until there’s an agreed-on way to make them more fair.

Most people would agree that society should have a way to decide who is a danger to others. But replacing a prejudiced human cop or judge with algorithms that merely conceal those same prejudices is not the answer. If there is even a chance they perpetuate racist practices, they should be pulled.

As advocates for change have found, however, it takes long years to make a difference, with resistance at every step. It is no coincidence that both Khan and Richardson saw progress after weeks of nationwide outrage at police brutality. “The recent uprisings definitely worked in our favor,” says Richardson. But it also took five years of constant pressure from her and fellow advocates. Khan, too, had been campaigning against predictive policing in the LAPD for years.

POPULAR

Welcome to the new surreal. H generated video is changing fil

Will Douglas Heaven

Modernizing the automotive in Creating a seamless customer

MIT Technology Review Insig

A brain implant changed her lif removed against her will.

Jessica Hamzelou

Longevity enthusiasts want to own independent state. They’re Rhode Island.

Jessica Hamzelou

That pressure needs to continue, even after the marches have stopped. “Eliminating bias is not a technical solution,” says Milner. “It takes deeper and, honestly, less sexy and more costly policy change.” 

by Will Douglas Heaven

DEEP DIVE

ARTIFICIAL INTELLIGENCE



Welcome to the new surreal. How AI-generated video is changing film.

Exclusive: Watch the world premiere of the AI-generated short film *The Frost*.

By Will Douglas Heaven



Modernizing the automotive industry: Creating a seamless customer experience

As digital modernization drives more intelligent vehicles, traditional OEMs are using advanced technologies to keep pace with business needs while balancing governance

By MIT Technology Review Insights

Google DeepMind's game-playing AI just found another way to make code faster

The AI-generated algorithms are already being used by millions of developers.

By Will Douglas Heaven

Eric Schmidt: This is how AI will transform the way science gets done

Science is about to become much more exciting—and that will affect us all, argues Google's former CEO.

By Eric Schmidt

STAY CONNECTED

Illustration by Rose Wong

Get the latest updates from MIT Technology Review

Discover special offers, top stories, upcoming events, and more.

Enter your email

 →

[Privacy Policy](#)

The latest iteration of a legacy

Founded at the Massachusetts Institute of Technology in 1899, MIT Technology Review is a world-renowned, independent media company whose insight, analysis, reviews, interviews and live events explain the newest technologies and their commercial, social and political impact.

Advertise with MIT Technology Review

Elevate your brand to the forefront of conversation around emerging technologies that are radically transforming business. From event sponsorships to custom content to visually arresting video storytelling, advertising with MIT Technology Review creates opportunities for your brand to resonate with an unmatched audience of technology

and business elite.

[READ ABOUT OUR HISTORY](#)

[ADVERTISE WITH US](#)

[About us](#)

[Help & FAQ](#)

[Careers](#)

[My subscription](#)

[Custom content](#)

[Editorial guidelines](#)

[Advertise with us](#)

[Privacy policy](#)

[International Editions](#)

[Terms of Service](#)

[Republishing](#)

[Write for us](#)

[MIT News](#)

[Contact us](#)

© 2023 MIT Technology Review

