THE BROOKINGS INSTITUTION
WEBINAR

FORENSIC ALGORITHMS: THE FUTURE OF TECHNOLOGY IN THE US LEGAL SYSTEM

Washington, D.C.
Thursday, May 12, 2022

PARTICIPANTS:

Fireside Chat:

REP. MARK TAKANO (D-Calif.)
Chairman, House Veterans' Affairs Committee

REBECCA WEXLER
Nonresident Fellow
Center for Technology Innovation
The Brookings Institution

Panel:

JULIA ANGWIN, Moderator
Founder
The Markup

REDIET ABEBE
Assistant Professor
Computer Science
University of California, Berkeley

GLENN RODRIGUEZ
Co-Director
Youth Services
Center for Community Alternatives (CAA)

ANDREA ROTH
Professor of Law
University of California
Berkeley School of Law

REBECCA WEXLER
Nonresident Fellow
Center for Technology Innovation
The Brookings Institution

* * * * *

P R O C E E D I N G S

MS. WEXLER: Welcome to the audience. Today, we are going to talk about forensic algorithms and the U.S. criminal legal system. First a note, you only get me on audio. I'm very sorry about that. I had some video tech problems.

But I'm Rebecca Wexler. I'm an Assistant Professor of law at U.C. Berkeley School of Law and I work on data technology in the criminal legal system. I'm super excited to welcome Congressmember, Mark Takano here who you get to see on the screen. Representative Takano has been the U.S. Representative for California's 41st Congressional District since 2013.

He's a member of the Democratic party and in 2012, Representative Takano was the first openly gay person of color to be elected to Congress. Most relevant to us today, he introduced the Justice in Forensic Algorithm Act of 2021, which would do a couple of things. We're going to get to chat about it and hear directly from him.

But basically, would impose some standards on forensic algorithms that are being used to bring criminal convictions and put people in prison and it would also prohibit the use of intellectual property laws to block criminal defense scrutiny of law enforcement technologies.

So before I hand this over to a chat with Representative Takano, I'm going to give a little overview of today's program. We're going to start out with a fireside chat between myself and Representative Takano. And then we're going to open into a panel discussion with some other really exciting guests that I'm thrilled to be able to have here today.

So the basic background for this event is the U.S. criminal legal system is becoming automated at every stage of our system. That there's policing and investigations, bail decisions, evidence that's used to convict people at trial, sentencing decisions and even parole, who gets out of prison. In every one of these stages computer systems are starting to play a role. Artificial intelligence is directing police officers on the beat. Audio sensors are generating gunshot alerts that are then being introduced in court.

We've got forensic analysts using complicated software to analyze DNA, fingerprints, face matching. Risk assessment instruments are helping to determine who is incarcerated and for how long. And there are a lot of questions about how these technologies are affecting the criminal legal system.

So, Representative Takano, thank you so much joining us here today to talk about these issues. And I'm wondering if you can tell us just a little bit about what exactly are these forensic algorithms that are currently being used in our system?

MR. TAKANO: Well, there are three main types of forensic algorithms. There's facial recognition. There's latent prints and there's probabilistic genotyping. And those are the three main types of forensic algorithms used.

Latent print is used for finger print and palm prints. DNA analysis is used by probabilistic genotyping. By attempting to determine whether an evidentiary sample from a crime scene is or not associated with a potential source sample from a suspect. And the results can be compared to a reference profile from one or more persons of interest.

Law enforcement agencies are increasingly using a new type of software to partially automate the analysis and interpretation of evidence in criminal investigations and trials. And they're used to analyze everything from degraded DNA samples and faces in crime scene photos to gunshots and online file sharing.

MS. WEXLER: Fabulous. And I'm just wondering. I mean what are really the stakes of using algorithms in this context? Because I understand software is around us all the time, but, you know, I'm thinking, gee, is the criminal legal system -- this is, you know, we put people in prison. Sometimes, we even take their lives.

MR. TAKANO: Right.

MS. WEXLER: And so, why does it matter so much?

MR. TAKANO: Well, I mean the stakes are pretty obvious. People can lose their freedom. They can be incarcerated and they can put on death row based on the evidence that's introduced to them in court.

And as it's simplified by even trying to get onto this platform. We see fallibility of technology, but platforms are different than the types of "black boxes" that algorithms are often presented as solutions. And so, there's the danger is a kind of presentation of infallibility to judges and juries and to people who use them. Like people who are sheriffs and police detectives and the police departments and law enforcement agencies that purchase these programs.

There needs to be, I think an amble amount of skepticism. And look, my position is not that I'm against these algorithms and against technology but we need to make sure that defendants in a court of law and their attorneys are able to exercise under the constitution due process rights to a fair trial. We must ensure that defendants have --

MS. WEXLER: Let me just --

MR. TAKANO: Go ahead. Go ahead.

MS. WEXLER: Yeah. Sorry to interrupt you, Representative. I just wanted to emphasize that one of the things you're saying. There might be errors in these tools that we wouldn't even know about, right? That the people using them wouldn't even know about. And you're telling me, well, one of the ways to fix that is if criminal defense counsel who we rely on to scrutinize the government's evidence could really probe these tools. I think that's what you're saying, and I just wanted to emphasize that.

MR. TAKANO: That was where I was going to go. What defendants need to have access to is source code and other information necessary. The underlying data on which that source code operates that data need to be sort of validated and shown not to have biases embedded in it.

We need to be able to challenge the evidence that is being brought against them. That has been processed, analyzed or interpreted by using the forensic algorithm. And basically, they need to know how this algorithm works. And but what stands in the way is often the people who create these algorithms use their copyright protections, the proprietary rights they have to tell judges no, you can’t do that. And so --

MS. WEXLER: Let me jump in and just push on that point. So there’s a variety of intellectual property protections we have in the legal system, right? And you're right some people have asserted copyrights. And some people have asserted trade secret laws as well to block scrutiny of these forensic algorithms.

And so, for those of in the audience who might not be as familiar with the USIP system. Trade secrets are a form of intellectual property that allows companies to protect confidential information or information that’s secret from being stolen by business competitors. So the whole idea of trade secret law is, well, you know, let me have some legal protection for this information. So if I share it with a business negotiator or if I share it with a subcontractor, they can’t just go around and steal it. The point of trade secret law is to protect businesses from competitors stealing their information.

Representative Takano, that doesn’t seem like anything that’s happened in the criminal justice system. Who would be trying to steal this information among criminal defense counsel? How could this be going on here?

MR. TAKANO: Well, I mean if I am someone with -- if I put a lot of investment and money into creating these programs, I can see where there would be apprehension if people have access to this. But let’s also remember that some defendants may not have the tools or the resources to even have given the source code. And even if given access to the algorithm, they may not have the ability to analyze it themselves.

And if you do have capacity, you might outsource it to somebody. And the outsourcing may be a point of contention. You know, who you outsource it to maybe the vulnerability to you as the proprietor. But and there are probably ways to mitigate that as well, but there's --

MS. WEXLER: Absolutely. And there are. Sorry to interrupt you. I don't have my video on so it's a little hard, but I wanted to just point out, yeah. Exactly as you're saying. There are way to mitigate that risk and in civil legal disputes. Disputes about contract violation or even trade secret misappropriation itself.

Companies are regularly disclosing this information under something called a protective order. And so, we could easily have the companies disclose the information to defense counsel or expert witnesses under a protective order.

But let me get to your bill actually because we only have a couple of more minutes. And I was wondering if you could tell us what is the Justice and Forensic Algorithms Act? What does it accomplish? How did you learn about it yourself and why did you decide to do something about it?

MR. TAKANO: Okay. Well, I'll try to answer that really quick. How I learned about it myself was I had a tech Congress fellow in my office, Emily Paul. Emily worked for me two or three years ago. Brilliant person. She read an article by you, Dr. Wexler, Professor Wexler in one of the law journals about this very issue.

And I green lighted her curiosity and said, you know, let's do something about this issue that Professor Wexler has raised. And it shows you, Professor Wexler, writing in law journals, you know, often comes to good things.

MS. WEXLER: I'm so glad to hear it. That's great.

MR. TAKANO: Well, there are two parts of this bill. And let's discuss them in turn. One thing that this bill does is it stops forensic algorithm vendors from using trade secret law to block criminal defense counsel from scrutinizing these tools.

And why is this so urgent? The trade secret privileges of software developers should never trump the due process rights of defendants in the criminal justice system. That's a bedrock principle that we have in our criminal justice system.

MS. WEXLER: Well said.

MR. TAKANO: We also -- but I want to get to the other thing. I know time is short. I can get into this in more deeply, but the other thing the bill does is it -- knowing that not everybody has the resources to test the algorithms. My will would mandate that NIST would create standards for testing computational forensics software including testing for disparate impact on the basis of race, ethnicity and social economic status, gender and other demographic features.

It also requires a NIST to institute a program to perform testing and mandates the federal law enforcement only use forensic algorithms that have been tested through this program. Because right now, we have this --

MS. WEXLER: Can I just ask you -- yes.

MR. TAKANO: Go ahead. Because we have right now, we have --

MS. WEXLER: It seems obvious that we should have this, right?

MR. TAKANO: I think this is obvious because we've got judges and jurisdictions all over the map across the country. And judges and juries and lawyers, frankly, have varying degrees of understanding of this technology.

And we do need the -- I think, the resources of the federal government through NIST to be able to establish these standards. And to say, this is a standard. The standards are established standards by which all forensic algorithms need to meet in order to be considered reliable and admissible as evidence in a court of law.

MS. WEXLER: But, Representative Takano, this sounds so obvious that we would need this. Are you telling me, and I think you are, and I think you're right. That there is no current federal standard to control the quality of forensic algorithms that we're using to put people in prison or even to take their lives?

MR. TAKANO: There is no current standard. And I think, Professor Wexler, you cited a case in New York State where there were contradictory conclusions by these forensic algorithms on the same case.

MS. WEXLER: Absolutely true. Yes, there have been cases where there was the same evidence analyzed by two different algorithms has produced different results. So that's really concerning.

And I just want to flag how important Congressmember Takano's contribution is here because it is here because it is true as he says. There is no current federal standard controlling the use of these algorithms. My colleague, Andrea Roth, is going to join us in the panel was pointing out to me last night.

Well, gee, we do have a standard for algorithms in breath test devices. So the devices they use to convict people of drunk driving. And I want to say, you can't see me on the screen, but I am a white, female and, you know, a relatively privileged in society. These are the algorithms used to convict people who look like me. Guess what? We've got standards for those algorithms.

The rest of the algorithms primarily used to convict people of color, people from marginalized and underprivileged backgrounds, no standards. That is a serious problem. So,

Representative Takano, I'm so grateful to you for having put this bill into Congress. And I just want to commend you for your work on this.

MR. TAKANO: Well, I want to say -- I want to commend too. I want to commend you for the scholarship and for, you know, being the point of inspiration for Emily Paul, my tech Congress fellow. And I want to especially make sure I acknowledge Emily because, frankly, members of Congress need resources such as Emily. I was fortunate enough to have a tech fellow.

They often choose to go to the Senate instead of House members, but we would all be better off if all members of Congress had access to really skilled and technologically and scientifically competent advisors such as Emily. So I really want to give an applause too.

MS. WEXLER: Fantastic. Tech Congress fellows, we should send more of them from the academy to Congress and great, Representative Takano. Hosted Emily Paul.

Thank you so much, Republican Takano, for joining us and for introducing the Justice in Forensic Algorithm Act of 2021. And I am now delighted to introduce our panel.

Julia Angwin, Andrea Roth, Glenn Rodriguez, Rediet Abebe, I'm excited to do a further conversation with these experts on these issues. Thank you, Representative Tanako, so much.

MS. ANGWIN: Thank you for such a great conversation. I'm just waiting for more panelists to arrive. Okay. Great. So we're here. Thank you very much. This has been a great discussion. Thank you, Representative Takano, for sharing with us what you're trying to do about forensic algorithms.

We have a great panel here of people to talk about what is happening right now in the criminal justice system with algorithms. And so, I want to pick up where you guys left off and discuss things that are happening right now.

I'm going to start with just a brief introduction of myself. Why am I here? And then I'm going to hand off to each one of the panelists to talk about their perspective, and then we'll have a discussion and we'll open it up for questions for the last 10 minutes.

So I'm Julia Angwin. I'm an investigative journalist and the Founder of a nonprofit newsroom called The Markup. And in 2016, I did an investigation of an algorithm called Compass that is used throughout the U.S. to predict whether defendants are likely to commit a future crime within two years of when they were assessed by this algorithm.

And it is used at all stages of the criminal justice system. So I wrote about it. It is used in pre-trial assessments. It's used in sentencing in some areas and it's used parole decisions. And I, as a journalist, was just really surprised when I learned that it hadn't been independent assessed. That basically, when I looked at it, there were different jurisdictions using it had written their white papers basically saying like it works.

And I thought, you know, that seems suspicious. So I went and obtained through the Freedom of Information Act request thousands of scores of people who were arrested in Broward County, Florida and then did the hard work of assessing whether those predictions were correct.

And it took almost a year of four people's work, but we did find that the algorithm was not particularly accurate. It was only about 60 percent accurate and I would just like to say that I, personally, if I was 60 percent accurate in my job, I would be fired, right? So like that was not a standard that like I've lived by.

And then when it was wrong, it was wrong in really different ways, in racially biased ways. So basically, the algorithm was very likely to give white defendants a low-risk score even if they were actually did go onto commit a future crime and be arrested.

And so, it was twice as likely to give a white defendant an incorrect low risk score as a black defendant. And it was twice as likely to give a black defendant a high-risk score when they were actually did not get arrested in the next two years as a white defendant. So basically, the rates were biased in different directions.

So white defendants got like an unfair pass and black defendants got an unfair penalty. And so that study of Compass is still being used today to discuss sort of what algorithmic bias can look like. And what the human stakes of it are. And it has prompted a lot of conversations about how to retune these algorithms. And honestly, I think I would like to say that there should probably be more conversations about whether there should be such algorithms.

But that is why I'm here to discuss and lead this discussion around the use of algorithms. Because we know that it's not just theoretical. It's happening right now. And so, the person who can really speak to it most directly is our panelist, Glenn Rodriguez, who himself was assessed by the Compass algorithm and actually managed to successfully dispute it, which is I think extremely rare. Possibly the only case.

And so, I want to kick it to Glenn first to tell us about his experience.

MR. RODRIGUEZ: Good morning, everyone. My name is Glenn Rodriguez. I currently serve as Co-Director of Youth Services at the Center for Community Alternatives and as Julia mentioned I was directly impacted by the algorithm utilized under Compass in New York State, which is utilized in the context of parole decisions.

So I was denied parole at my first appearance despite the fact that I had actually earned a reduction of my sentence due to significant programmatic accomplishments as well as behavioral adjustment. I was denied parole because of Compass. And that for me just didn't necessarily go well.

Obviously, I had a certain expectation. I had worked very hard convincing that panel that I was no longer the 16-year-old who committed the offense for which I was incarcerated. And this is 27 years later, by the way. So, you know, over the course of the last 15 of that sentence. I had worked very hard to convince that panel. I obtained college degrees. I became a certified dog trainer.

I did pretty much every program that was in every facility that I was housed. I pretty much took part in every program. I volunteered for approximately 10 years to counsel at risk youth. And none of that was enough to kind of go against Compass.

And I think that in large part that is because Compass obviously, you know, gives one perspective. But folks are concerned about potentially losing their jobs, right? So in this case, in New York State, the way it is set up. Compass is supposed to inform parole on its decision. And often times, what happens is parole commissioners are reluctant to kind of go against what Compass is dictating pretty much.

So regardless of what I had done in terms of accomplishments, Compass said that I was a high risk because of my prison misconduct. By the way, which dated over a decade back. Yes, early on during the course of my incarceration, I was involved in a number of incidents. This was during the ‘90s. Mind you, I am appearing before a parole panel in 2016. And I’m being penalized. And I was eventually denied parole because of actions, because of disciplinary actions that were -- you know, I was the subject of disciplinary actions in the ‘90s.

This is precisely one of the reasons why this algorithm was actually implemented in New York State. The rationale for implementing it was obviously to reduce human bias. And not only that but also to kind of refocus the parole from like the nature of the crime. And kind of focusing on past events and kind of looking at the individual in terms of what they have accomplished and who they are at the moment. Not necessarily focusing on who they were 20 years ago.

So in my case, I was denied parole. Obviously, I did not agree with that decision. And so, I sought out the advice of counsel. I met with a number of folks in the prison. I became pretty much a researcher. I started collecting folks, their results and comparing them to mine.

And ultimately, what I was able to do -- and by the way, Rebecca who was also on this panel and to whom I am very gratefully for. She actually wrote an article where she featured my story. I was able to reverse engineer some of what was happening, right? So I was able to accumulate over 100 Compass results and kind of get a sense as to what was happening and how the questions were weighed.

And it appeared at least from my findings that one of the suggestive question was more heavily weighted than all the questions leading up to it under the disciplinary history section of the Compass. And it was the question that said, does this person appear to have notable disciplinary issues?

I don't know what someone appears to -- I don't know what someone looks like who appears to have notable disciplinary issues. So I personally don’t think that person should even be there because the way it’s worded doesn’t necessarily make sense. But that question appears to be more heavily weighted then all the questions leading up to it which are all based on numbers and stats.

So how many misbehavior reports has this person had in the past two years? All of the questions preceding it focused on the past two years. And for me in my case, I did not have any instances of misbehavior. I had nothing but accomplishments. And regardless of all of that, I was ultimately denied parole on the basis of prison misconduct.

So I am very thankful to Representative Takano for the introduction of the Justice and Forensic Algorithms Act because I believe that there should be greater transparency as it relates to these predictive software because, you know, we look at this and we don't necessarily always get the impact that it’s having. But I was personally impacted and if it were up to Compass, I would still be in prison today.

And here I am five years later. I’ve been employed since two weeks out of prison. I’ve received five promotions. I’m currently sitting in a senior leadership position at my organization overseeing the provision of afterschool programs, recreational programs for young people in secure detention facilities, working closely with ACS, working closely with docks, with upstate.

And so, none of this would have been possible had it been for Compass. And so, definitely I appreciate the Brookings Institution for shining the light on this issue because this is something that we should be talking about. So thank you.

MS. ANGWIN: Thank you, Glenn. That is just exactly the kind of horrifying story that is what prompted me to want to investigate this algorithm.

And I have to say that I also find it really surprising that I wrote this investigation in 2016. You, you know, found this error that same time. And this algorithm is still in use and unchanged. And that's something that just is surprising to me.

Next, I want to move to Andrea. She'll introduce herself and explain her work on the right to cross examine algorithms.

MS. ROTH: Hi. Thank you so much to Julia and Brookings and Representative Takano for taking on this issue and everybody on this conference. I haven't met Glenn before. I know of your story from others, but I just wanted to say how inspiring it is that you took all these traumatic things that have happened to you and are working to help other people not be subject to it. It's very inspiring.

And what I'm sort of an evidence criminal procedure scholar. And so, I study how we scrutinize testimony and how we scrutinize factual claims that are made to put somebody in prison or to find somebody guilty. And so, the thing about algorithms is again going to Julie's question is that you can't cross examine them. And you can't put them under oath. And you can't physically confront them.

But I just wanted to, I guess say three things because I know one thing that always looms over these discussions is, yes, but in general I know that, you know, some algorithms might only be performing at 60 percent or whatever. But in general, aren't these improvements over humans? You know, isn't this generally an improvement?

So as long as we think that they're generally accurate, they've got to be better than what we have. So let's not be too precious about insisting that they're perfect. Don't let the perfect be the enemy of the good. That's what I hear a lot.

So I just wanted to back up and say a couple of things. One is that, you know, with humans we subject them not just -- we don't just say, okay, this expert in this white lab coat, they're pretty good. Let's just let them say anything and you don't get to cross examine them. You don't get to talk to them before trial. You don't get to find out why they're saying what they're saying. Just trust us.

We don't do that with human witnesses. We subject them to a lot of adversarial scrutiny even after a judge has deemed them reliable enough for the jury to hear them. And so, that's what confrontation is. Above and beyond just, you know, a stamp of approval from a judge. Our system, for better or worse, has adversarialism built into it so that really you've got somebody with literally skin in the game.

You know, Glenn shouldn't have to be out there reverse engineering this even though he did. You know, we have people whose job it is to be devil's advocate and to think, you know, what could have gone wrong here that could have led to somebody being falsely accused or in prison for longer than they should be.

And so, that's what adversarialism gets you. Now, that's our system for better or worse. So if we're going to have that system, we've got to subject factual claims from machines and algorithms to that same level of adversarial scrutiny.

And the other thing about humans is that we have a general sense of their limits and animals. We have a general sense of their limits. So when an eye witness says something or if somebody with, you know, seven perjury convictions makes a claim, we at least have some framework of having lived in the world about how to assess the probative value of that evidence.

So doubling down on something that Representative Tanako said in the chat with Rebecca. You know, it's the presentation of infallibility. It's the fact that we don't know the limits of the claims that are being made by this software. So the jury, the fact finder, the parole judge, you know, doesn't know what this score means in the same way that they might be able to at least put something in context that's more human rendered.

So that's the, you know, view from 30,000 feet. And then if I could just take 15 more seconds to say so what's wonderful about this bill is that, you know, especially for algorithms used in post-conviction or pre-trial, they aren't subject to the right of confrontation. So if we're going to increase the amount of adversarial scrutiny of algorithms like Compass, it's got to be through statutes and not the constitution for the most part.

So, you know, so many others including, you know, the distinguished computer scientists on this panel can say much more than I can about what adversarial scrutiny looks like in this context. But, you know, independent validation, you know, industry standard, software testing unfettered access to the programs, possibly corroboration by more than one algorithm if it's going to be the primary evidence of guilt or liability. These are things that, you know, we need to pass by statutes or court rule, I think. So let me stop there because I've spoken too much, but thanks.

MS. ANGWIN: Thank you, Andrea. I have to say I hear that all the time. People say about my reporting. Oh, whatever. You wrote that. But like judges are really bias.

And you know, it's interesting because when I was looking in the history. It used to be that they would bring in psychologists to testify in court and see if somebody was likely -- particular in violence -- to see if somebody was likely to be violent in the future.

And the Compass algorithm, by the way, had a special violence score, which was only 10 percent accurate. Which is like completely insane, but that’s what our data showed. And interestingly, psychologists were sort of taken out of the business of testifying about future violence of offenders because they were only 50 percent accurate.

So that was the history. And then we put in this machines, right? And I think that that speaks to like the human faith in machines, right? Like we're just, oh, they're so smart. They're like robot gods, right? And so, I want to use that as an introduction to our computer scientist, Rediet Abebe, who’s going to tell us are they robot gods?

MS. ABEBE: Well, okay. Hi, everyone. I am Rediet Abebe. I am an assistant professor of computer science here at UC, Berkeley. And I’m also one of the cofounders of Black in the Eye which is an organization that focuses on increasing representation and inclusion and economy of black people in black communities in the field of artificial intelligence.

And I guess I’m here to be the Debbie Downer and say that algorithms don’t know things that we don't necessarily -- that we don't know. So there are sort of two lines of work that I’m doing that are directly related to several of the things that have been mentioned.

So the first one is this concern that you all have brought up around, well, humans are bias, right? And so, if it’s going to be an improvement over humans than why can’t we use these algorithms? That’s for your people in prison. That’s for your people, you know, denied bail and so on and so forth, right?

And this issue of like evaluation, right? You know, what standard are we holding machine learning algorithms against? And what do we assume they're capable of is something that is less examined than you would expect on the sort of machine learning AI side of the research community. And in particular this human versus algorithms comparison in how it is discussed by the machine learning AI research community is maybe the thing that frustrates me most, right?

Because there's just a lot of things that it assumes. So one example that I like to use is if I gave you two numbers, two very large numbers. And I said multiple them, it would take you a while to multiple them, right? It's just multiplying, you know, very large numbers. It's hard. I could use a calculator, multiple it really quickly. Does that mean that this calculator is a better mathematician than you are? No, right? There's more to being a mathematician than being able to multiple really quickly.

These are things that we automate because we can, right? There's not really much insight into being able to multiple things really quickly. And so, there's this sort of assumption that just because algorithms are able to do something very narrowly specified, very quickly that that makes them also very good at looking at a broader set of things, right?

So for instance, you know, you're mentioning this question of, you know, is this person likely to, you know, return court? Whatever it is that you're trying to predict, right? That's to the point in a very specific way. You're going to formulate that in a machine learning algorithm in a very specific way. That loses a lot of context, right? And that assumes a lot of things.

So for instance, you saying things like there were questions around disciplinary behavior, right? Well, we know this in the criminal legal system. We know this in education. We know this in many different context. Who's constantly disciplined, right?

If you have a black student and a white student, you know, in middle school, they do the exact same thing and the white student is going to be given a second chance but the black student is not. Technically, you know, that is a fact that student was disciplined and that's going to be on their record, but does that mean that you're actually treating the situation fairly? You're not, right? We know this to be true.

And so, there’s that aspect of it. This problem formulation and the assumptions that we're making around what each of these datapoints even if they're not -- they don’t seem, you know, quote subjective, right? Even if their numbers and we're like this the counts. This is the number of disciplinary kind of behavior that we’ve kind of seen in this person, right? Even when it’s just a number that you're counting, there is a lot that’s encoded in that.

The other thing I want to also emphasize is that even if, you know, these algorithms were, quote, less bias, right? And it’s not like our criminal legal system as run by sort of humans is just either, but at least in the sort of the latter situation, you have multiple different judges making different decisions, right? Maybe they're quote, less bias or more bias in this algorithm, but at least you have different people making maybe different decisions and potentially making different mistakes.

You replace that by a “less bias algorithm” then what happens? Now, you're going to be kind of penalizing the exact same people because it’s just a single algorithm penalizing. You know, making errors on the exact same type of population.

And the second thing that happens is also that algorithms are just faster than humans, right? And so, if you have a criminal legal system that is bottlenecked by how quickly you're able to process people, right? You remove that bottleneck by moving the humans away and replacing it by algorithms. And now, I guess it becomes a matter of how quickly are we able to build prisons, right?

And so, there’s so many different dimensions to this human versus sort of algorithm comparison that are not mentioned. And there’s also this issue that, you know, I don't fault my community necessarily for this, right?

We get excited about the things that we build. They do things that we get super pumped about them. Or like we're able to translate things really well. And we can detect this and that really well. And, you know, there’s a little bit of cherry picking that we do and what we present to the public, right? And from that we just extrapolate so much, right? We say, I’m able to, you know, these machines are able to do this thing well. It must be that they understand language. These machines are able to kind of detect this thing well. It must be that they can like do this or that, right?

And so, there’s this sort of unquestioned assumption or infrequently questioned assumption that, you know, ultimately everything can just be automated and it can be automated and done better than humans.

And I find that to be concerning, right? Because there is sort of a gap in terms of understanding between sort of folks who are building these machinery algorithms and those who are not necessarily in the weeds of what’s going on. And may not necessarily understand sort of all the different ways that they can fail.

So that’s kind of one-way line of work. And the second one is directly related to Representative Tanako’s act and also what has been mentioned, which is this ability to be able to scrutinize the software. And especially when they're being used as evidence. This is not something, you know, I am in this area. I kind of pay attention. I had not realized the extent to which we were actually using output as evidence.

And we're not even able to cross examine. It’s like a witness standing, you know, in front of you and saying things. Potentially lying and you're not even able to cross examine them. That’s was shocking to me when I learned this. I learned this from Rebecca who, you know, I’ve been collaborating closely with over the past two years and has been, you know, giving me crash courses about, you know, what was going on, and that was really shocking to me.

And so, we spent, Rebecca and I along with our collaborators, we spent quite a bit of time trying to think about what would it look like for the machine learning and AI community to build something that can help with this scrutinizing process? And I want to mention one thing, which is -- so we've mentioned probably six types of software in this panel so far.

And we've mentioned how these types of software could have different biases by gender, by race, by all these things and these are incredibly important. But something I also want to emphasize is that, you know, each case is different. And what's applicable to each case is also very different, right?

And so, it is good to have these -- being able to do like an evaluation, just a broader evaluation by race and by gender and so on. But there's also going to be things that show up case to case, right?

It might be the case that, you know, you're in an situation where you have a sample from a crime scene that includes multiple people's DNA and you're trying to decide, you know, if someone if someone's DNA is included or not. And you're using a software that hasn't been validated for cases where the sample includes, you know, five or more people's DNA. Maybe that is the case, right?

Maybe you don't even know the number of contributors to that sample. Does that software work in that particular case, right? And so, you could be in a case where you're like I want to know if it's been tested in this very particular setting because that is a setting that applies to my client. Let's say that I'm a defense lawyer and that's the case that I want to evaluate.

And so, what we've done, Rebecca and I and our collaborators, is to provide a framework that uses notions from the robust machine learning community, which is sort of a community that's been doing, you know, a whole set of different things with this notion of adversarial scrutiny, right? Because there is something that's matching here which is that you're saying, here's this thing. It's outputting some evidence and I need to be able to probe it and see that it works in this particular setting that I'm interested in.

And so, this is work that was just recently accepted to the ACM Conference on Fairness, Accountability and Transparency. It's going to be posted publicly in a few days. And it provides a framework that allows defense counsel to be able to scrutinize evidentiary statistical software that's put in front of them for the particular cases that they're working. And they get to say, I want to be able to test this for people, you know, 45 and over in cases where there's at least five people's DNA included in the sample. And, you know, you name it in these kind of geographic settings, whatever it is.

And to be able to say, it could be the case that the software is like 95 percent accurate, right? But maybe in my particular case it's not. Maybe it's 60 percent accurate in the case that you're using it. And that should be also shown to the jury.

And the last thing that I want to say is that I think that there are structural issues in machine learning AI especially on the academic side that I think are really exacerbating this, right? It's only really in 2016 that the community started having a much more open and active discussion around these issues of fairness and accountability in machine learning algorithms in large part because of, Julia, your work.

And that was actually around the time that I started grad school. It was really exciting for me to see like the -- you know, the year before and the year after these conversations were really different. And to me, it was also a huge relief, right? Because I obviously like, you know, my field and the techniques that we are able to build and provide. But as a black woman, as, you know, as someone who has the background that I have. For a while I felt like I couldn't bring my concerns to the table because the community had kind of considered it to be sort of out of scope, right?

If it's, you know, now you're concerned about how these things are going to be used then, you know, that's a social question. It's not really a technical question, right?

And so, in some sense that shift has been something that I've welcomed and I think that it is something that we should continue to discuss. These structural issues that govern the kinds of tools that we build and also the kinds of scrutinizing and evaluating that we do of the systems that we build.

MS. ANGWIN: Thank you. I love that you come up with a framework for this disparate analysis because, you know, one of the things that I sort of think about a lot with the analysis we did of Compass was that we didn't really get to everything.

There was a whole bunch of gender disparities as well. So women who scored medium risk on Compass were actually the same as like a low risk man. And so, the gender disparity is really bad. And that was something we didn't get a chance to address in the story because it had actually been reported. And so, it was sort of known and yet at the same time it hasn't gotten the same scrutiny. And there are so many populations that could be tested, right? And so like the idea of building this framework for it is such a good tool.

And, Rebecca, it's delightful to see your face now. I want to ask you a kind of provocative question, if you don't mind?

MS. WEXLER: Go for it.

MS. ANGWIN: You talk a lot about like the issue of trade secrets around these algorithms and Representative Takano's bill asks for the algorithms themselves to be disclosed. But as you know this is really controversial, right?

The companies really fight hard to protect trade secrets. All companies do. And I want to just sort of ask provocative questions. So the night before I published the Compass analysis, I had sent my analysis to the company in North Point, which I think it has changed its name to something else.

And, you know, asked them -- we had been engaged in weeks and months back and forth where I had showed them my analysis and I wanted to make sure I wasn't getting anything wrong, right? You know, we brought to statisticians, criminologists but it was a significant finding and I wanted their input.

And they were like, look, we obviously didn't intend for any bias in this. This is meant to debias. And they wanted to convince me that it wasn't bias. So they showed me. They were like you can't show this to anyone, but we're going to show you the algorithms so you can see that it's not bias, right? So they send me --

MS. WEXLER: You got into the inner circle.

MS. ANGWIN: So I got to see the algorithm, okay? And linear equation with some constants and some variables. And I studied math in college and I'm from a math family. So I'm familiar with mathematical constructs. And like, I would challenge anyone to look at that algorithm and say it's going to have disparate impact and outcome, right?

The reality is to understand that algorithm, you needed to see the outputs. Because actually the inputs were also known, right? People had seen the questionnaires. They knew that it was -- it had been papers. Lots of legal scholars have written the kind of questions that they ask in these risk scores are going to lead to bias outcomes.

And so, the question I had is do you really need to see trade secrets to do this type of analysis? Or could you just do disparate impact analysis?

MS. WEXLER: Great question. And I am excited to answer it. And before I do I just want to take a moment to introduce myself again because I didn't have a camera before. But I'm Rebecca Wexler, I'm an Assistant Professor at U.C. Berkeley Law and also a nonresident fellow at the Brookings Institution and I work on data technology and the criminal legal system.

And as Julia said, one of the main areas I work on is intellectual property and trade secret barriers to criminal defense scrutiny. So -- all right. I'm back.

Technology. All the things my wonderful co-panelists have said today could not happen if companies could use trade secret law to block cross examination, to block access to executables or outputs or inputs as Julia was able to use, to block somebody like Glenn Rodriguez from accessing the survey or results of his Compass assessment, to block the adversarial scrutiny that Rediet Abebe is talking about.

This is wrong. Intellectual property should not block due process. Trade secrets are designed to prevent stealing in business competitions. But the criminal legal system and defense due process is not a business competition.

So Julia is saying, well, what if you don't really need a particular piece of information? Maybe it’s the algorithm. Maybe it’s the source code? Maybe it’s the user manual? This is another thing that companies have asserted trade secrets to block.

And I’m going to say that’s a totally different question. Whether we're going to evaluate what the defense needs in any particular instance. It’s case by case. The problem is that trade secrets are not a good reason to withhold information, period. That’s not what intellectual property is for. And to the extent that these developers actually do have a business interest in some level of confidentiality, we can protect that interest with a protective order, which is what they do in civil cases.

So at a minimum somebody facing incarceration or death, should get the same or better access to information than somebody that’s defending in a contract dispute. So that’s my answer to you. The conversation often goes and I hear this challenge. Oh, Rebecca, what are you complaining about? Why do people need access to that information?

And I’m going to say why are you asserting intellectual property law to block due process? So that’s my answer.

MS. ANGWIN: Thank you. That was a good answer. So we don't have that much time. But I want to end with like a very provocative mini lightning round because the question I have as a journalist writing about this is like, okay. We can debate about making an algorithm open or we can debate about whether it's fair or not.

But like a very important question is should we be using predictions of the future in incarceration decisions at all, right? Given that we know the future is very unpredictable in all cases, right?

Did anyone predict we were going to be -- there was going to be a war in Ukraine this year? No, right? So like I wonder if we could just do a lightning round where I would just like to know is there even a value to using this type of predictions? I'm going to start with you, Glenn.

MR. RODRIGUEZ: I don't see the value in it. I don't think they can accurately reflect what any individual will do in the future so.

MS. ANGWIN: Thank you. Andrea?

MS. ROTH: I don't think risk of dangerousness should be a basis for pre-trial detention, et cetera. But I do think that risk of flight is going to be a legitimate legal basis for detention no matter what. And if algorithms can help us with that determination, I'm open to their use. But that's, you know, that's a big if.

MS. ANGWIN: Rediet?

MS. ABEBE: I'm also going to say no. I think that right now as it stands a lot of the predictions that we are performing in the criminal legal system are known to have a lot of issues. And I think with the ability to better evaluate and scrutinize them, we'll probably find even more issues.

And so, I think that this is not something that we know we can predict well. And it's people's lives in danger. And I think once these things have been embedded within the criminal legal system, especially ones, you know, you've invested a lot of money and resources.

And you've learned how to use them. And you fired a bunch of people who would have done it anyway. It's going to be much, much harder to remove them, right?

So I think right now we just don't have enough good evidence to be able to use them in the way that they're being used.

MS. ANGWIN: If we don't have enough evidence to be able to use them as evidence.

MS. ABEBE: Yeah, exactly. Exactly.

MS. ANGWIN: Rebecca?

MS. WEXLER: Companies are asserting trade secret law to block scrutiny all kinds of algorithms. So it's not just predictable. It's investigative algorithms. It's malware. It's face matching, fingerprint matching, voice matching, bullet analysis, audio analysis, gate matching, handwriting analysis. As Rediet had said, we don't have enough conversation happening right now about the algorithmic output that's actually being used as evidence of guilt in trial.

And those are all happening in addition to the predictive systems that you're talking about, Julia. So I'm going to say regardless of the answer of whether we should be using predictive systems. We should not have trade secret law block scrutiny of any of these technologies.

MS. ANGWIN: And maybe I'll end it with just like a question about if we're going to use like you said, Rebecca, all of these systems that are, you know, the DNA matching. We haven't gotten into, right, is very questionable. Some of the systems don't have a lot of evidence behind them.

And so, the question is are we going to train our judges to be statistically literate? Or how are we going to address this idea of bringing probabilities into the courtroom with due process? So this could be a longer discussion that we have for this panel, but I want to leave the audience with that question.

And thank everybody for a fabulous discussion. And thank you to the audience for bearing with us through our technological difficulties. Rebecca, you overcame them all delightfully so thank you so much.

* * * * *

CERTIFICATE OF NOTARY PUBLIC

I, Carleton J. Anderson, III do hereby certify that the forgoing electronic file when originally transmitted was reduced to text at my direction; that said transcript is a true record of the proceedings therein referenced; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were taken; and, furthermore, that I am neither a relative or employee of any attorney or counsel employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2024