FTC FinTech Forum: Artificial Intelligence and Blockchain
March 9, 2017
Panel 1
Transcript

DUANE POZZA: Thanks again, Deirdre for that presentation, which has given us a lot to talk about in terms of the issues. Our panelists list is distinguished and I will be up here for half an hour exhausting their biographies. I'll give the high-level details and refer you to the handout and the website for more details. Sitting to my left is Pam Dixon, the founder and executive of the World Privacy Forum, a US based public interest research group focused on consumer-privacy research. Pam is an author and a researcher in area privacy, whose work includes the Scoring of America, a substantive report on predictive analytics and privacy. Rayid Ghani comes to us from the University of Chicago, where he is the director of the Center for Data Science and Public Policy, and a senior fellow at the Harris School of Public Policy, and the Computation Institute. Rayid focuses on the use of data science, machine learning, and AI in solving large public policy and social challenges.

Deirdre Mulligan is an associate professor in the School of Information at UC Berkeley, and the faculty director of the Berkeley Center for Law and Technology. Her research explores legal and technical means of protecting values and emerging technical systems. Morgan Reed is executive director of ACT, The App Association, which represents more than 5,000 app makers and tech companies in the mobile ecosystem. Reed specializes in a number of policy issues, including security, privacy, and connected health. Ken Schneider is an assistant regional director in the San Francisco office of the Securities and Exchange Commission. His experience includes reviewing the operations of advisers, taking advantage of FinTech innovations, such as marketplace lending, virtual currencies, and automated investing. And Paul Schwartz teaches at UC Berkeley School of Law and is director of the Berkeley Center for Law and Technology. His many publications include the casebook Information Privacy Law and the distilled guide, Privacy Law Fundamentals.

So I want to dive right in. We had asked, in terms of framing the issues, Deirdre to try to focus in particular on one of the consumer implications of AI in a particular kind of product that we're already seeing trying to come on the market, which is these sort of robo-advisors. And I would broaden it to include not just ones to give you investment advice, which implicates the FCC for example, but also ones that are providing sort of money management and personal advice more generally. But in the end, I think that's very helpful in the way we frame these issues that we'll talk about throughout our discussion. But I also wanted to start out by talking about all the other applications that AI has in the financial sphere and beyond. So my first question is, where is AI going to be integrated in ways that consumers can see, and also in ways that will affect consumers but might not be obvious to them. And the first question will go to Morgan.

MORGAN REED: Thank you very much, and thank you for all the distinguished panelist. And Deirdre, thank you very much for the opening presentation. I think it's interesting when we talk about the financial impact of AI. There is this tendency to focus on what we're talking about, robo-advisors, and all these other aspects. The reality of AI is the idea of machine learning. It's always funny because we talk about AI and it's this year's buzzword. We used to call it big data.

We've called it any number of things. It's simply the use of machines to look at patterns and to make determinations about how things work. If I were to quickly say, everybody raise your hand and show a finger, well, I can use math and calculate the average, divide it by the number of people, how many fingers are held up, what was the average number of fingers that everyone in the room held up? Now imagine that I instead try to look at this data and I realize that oh, groups of three or more tend to hold up the same figure; people who sit across from each other have very wide differing fingers that they hold up; and then, I can start to look through that data and make deterministic revelations about the group that I'm dealing with.

And so when I hear people talk about the algorithm, it has that kind of Terminator feeling that it's going to walk in on metal legs. But in reality, algorithms are the same thing we all did in science back in the 80s before we had those kinds of computers. Where we'd pour through spreadsheets and look through for things that sparked our interest or our attention that show groupings of information. And so that is what we're dealing with. Only we're doing it a lot faster with a lot more powerful machines and a lot more data.

And I'll give two kind of benchmark points on where this AI lives right now. I see the gentleman there pulling up his Starbucks coffee and drinking it. Well, interestingly enough, Starbucks pulls down datasets through a deal with AccuWeather to make determinations about long-term staffing predictions for Starbucks coffee. And you may say, what? Well it turns out that it takes longer for Starbucks to make a cold drink than it does for a hot drink. And so by looking at long-term weather patterns and long-term temperature patterns in areas, Starbucks is able to actually offset the idea of getting that call from your manager at the last minute to say, we don't have enough people, come in now, drop what you're doing, and actually make predictive modeling about how many people we need on days where it's a little warmer and people are going to have more cold drinks.

Now that seems like a really small use of AI and insignificant, but if you look at staffing issues, if you look at shipping, if you look at all of the other issues that are involved, there is an enormous amount of data that can be pulled down by Starbucks on an issue like that. I'm going to go to the other extreme. One of the areas that I spend a lot of focus on in AI is called clinical decision support. And that's the idea that the ability to use machine learning to make a recommendation to a physician-- we refer to it as a learned intermediary-- on the treatment course that he might offer a patient. And this is very important, the average physician, sparklingly physician, great record, maybe in their lifetime, she will see 15,000 patients total. 15,000. And with your particular disease, your comorbidity, your genetic markers, your age, and all the other things, in their entire lifetime they may see 500. Maybe they'll read a journal article about your condition, maybe they'll talk with a couple of friends, but their total dataset to make the determination of what medication that you will be recommended will be infinitesimally small.

Now imagine that your physician is not dealing with the 500 patients with your condition they've seen over their lifetime, and instead you were able to bring the power of 10 million people with your same comorbidity, your same age, your same genotype, your same information. And you can determine that guess what, this medication doesn't work as well, or surgery isn't recommended. And that sounds very broad, but I'll give a very specific example. Nearly one in eight women in America will suffer some kind of thyroid problem. We have an epidemic of chronic fatigue syndrome, we have an epidemic of Graves' and Hashimoto Syndrome. And yet, if you talk to the women of America who suffer from this, they will tell you almost uniformly that their endocrinologist says, here's some Levoxyl, have a nice day. And so the web is filled with what amounts to other physicians who have the condition saying, well I tried this diet and it worked; or, I took this supplement and it worked. We have returned to almost the age of witchcraft in trying to determine what to do about thyroid disease in America. And yet, if you look at it, there are probably a whole series of patterns, of behavior, of genotype, of intake of food, of age, of where you live, that might inform people about which diet works better for these conditions, which treatment isn't necessary in this case, what dosage makes the right sense for this set of genotypes?

And so, when we talk about the range of where we have in the financial impact, if you realize that one out of eight American women are going to suffer from thyroid disease, 20 million are currently suffering from some form of thyroid injury, you think of the economic impact of that and what it does to the stock market, and what it does to the company you work for, and in fact, how you can take care of your kids, that's what we see as where we need to head on AI. So in short, I want you to think about those as kind of the bookends of what we're talking about in its power.

DEIRDRE MULLIGAN: Health care is always such a compelling example, and there's a paper that came out, Rich Caruana, he's a researcher up at Microsoft, recently, it just shows both the power and highlights some of the issues that I was just talking about on bias. And it was a presentation, the [INAUDIBLE] Machine Learning Conference. He was one of the keynote presentations on one of the sessions. The paper basically was working with health care data, and they were trying to figure out better ways based on detailed records of hospital admissions to get the right people into the hospital with pneumonia. Like who really needed to escalate versus who could be sent home. And they were using some really complicated algorithms, the things that we would kind of associate with deep learning, around that sort of stuff, and they turned out some patterns that seemed to be highly predictive. However, Rich, having a deep appreciation of this "do no harm," really knew that he needed to understand what were the rules that seemed to be getting pulled out from this algorithm. So they looked.

They went through the records to look and see what happened, and it's ironic, but it's really problematic that people with asthma were going to be sent home. Now, anybody who's had asthma, or like me, had a kid who had asthma, a brother with asthma no no no, those are the people that go straight to the top of the list. And that was the problem. That they weren't in the dataset. Because if you come in with your kid with asthma, you go straight into the ICU, the NICU. That's where you go. And so basically, what they ended up doing was using an algorithm that was actually less predictive, but where they could interpret the rules that were being derived. Because in the medical context, the need to make sure that the medical professional retains and being understanding of the algorithmic output, correlation is just not good enough.

MORGAN REED: It's not causation.

DEIRDRE MULLIGAN: So all I just wanted to say is that all of this stuff that Morgan is saying is incredibly true; but you need to be really sensitive to some of the blind spots in our data, which again, had nothing to do with anybody's desire to be biased. It's just that there was a gap in the population represented in the data, which would have had disastrous consequences for treatment of certain patients. I'm certain doctors would have intervened and still sent the people with asthma to the hospital, don't get me wrong, but it shows that you need to really be sensitive in thinking about both the algorithms and the data.

DUANE POZZA: And Pam has a comment, too.

MORGAN REED: Just really quickly, I want to clarify one thing, though. That's why I hate the idea of AI as this thing that's other. Because what you're really talking about is people who didn't go through all the processes that were necessary, and that's how we learn. Algorithms are designed by people. Some of you will probably design algorithms that we use. And so ultimately, I think what you're highlighting is why what we need to do is not just turn it over and say, oh well, the computer made the decision. No, it's actually to say, how do we look at that data? How do we reinterpret it? How are we constantly putting it under the pressure it needs to be so that it reveals more accurate information? So I think you're completely right. And that's why it's very important for us to separate this, AI is an other, and a separate, and to understand that AI is created by people. By programmers with flaws and biases and that's a big part of that.

RAYID GHANI: I don't think it's as simple as saying, algorithms are created by people. Algorithms are initiated by people.

MORGAN REED: Yeah, that's more accurate.

RAYID GHANI: After the data comes in, they adapt, and no programmer can understand at that point. I think that's a big distinction and you should be careful about that.

MORGAN REED: Great point.

RAYID GHANI: I think Pam had a comment.

PAM DIXON: I do. Thank you so much. First, thanks to the FTC for holding this event and I appreciate your invitation and this timely topic. So you asked a very broad question. So I just want to frame things very broadly. Almost all of my work has been on convolutional nueral networks and looking at long tail and high dimensionality and what you do with that and how fairness interacts with that. So that's been my long-term focus in AI. So just zooming back to frame all of this, so a really good way of thinking about AI is to understand this-- the ecosystem has moved. What we were dealing with is no longer what we are dealing with. And AI is not a service; it's the plumbing that's in the new ecosystem.

And the tension points change because the ecosystem has moved, and because the AI is built into the new ecosystem. And as a result, there are new tension points. And a very, very simple way to think about how this works out in practicality is this-- So if you look 25 years ago, you can think of a software being programmed. So you would be a programmer. You would program the software. But the way to think about it in terms of shifting to an AI environment is that the software isn't programmed anymore, it's trained. And when software is trained, there's the crux of the difference-- it's between programming and training. And when it's trained, basically you're hauling a dataset. And datasets can now also have algorithms as factors. So we have to understand we're not dealing with algorithms per se anymore. We're really dealing with networks of information that have capacities that are sometimes surprising, but the bottom line is that when you're building, when you're working with and training these items, it's all about inference. It's all about, should you build this? Can you build this? Can you make this inference?

Once you have an inference, how on earth do you interpret it? What Deirdre was saying in her framing remarks about the issues with interpretability, that's a really important focus. And that's actually something that I'm very focused on. Because if someone comes out with a number-- like a credit score is a great example of this-- you can be defined by that number. But right now, for the credit score in particular, there is all sorts of regulation around what the credit score means, what factors go into the credit score, how it can and will be interpreted, what the range is, and how it will be used. But because the ecosystem has moved, because we've got plumbing here, some of those old ways of doing things are much less feasible. For example, Deirdre mentioned back-end notice of practices. It just doesn't work in this environment. We need more of an interpretation on, should this inference be done in the first place? And that's just for starters. But that's just the broad framing.

DUANE POZZA: So one thing we hear is that AI is going to become sort of ubiquitous in our lives. So if you think about the Internet of things, you think about connected devices, personal assistants, connected cars, the different platforms that can be providing services to consumers, AI will be steadily more integrated into them. Is that what you see happening with AI? And I'll open this up to Deirdre first. So is that what you see happening with AI, and are there any sort of particular problems that arise when you have AI being integrated across different sectors of a consumer's life, basically?

DEIRDRE MULLIGAN: So I do think that AI is going to become prevalent in many consumer devices and some of the specific challenges that arise when we think about different AI systems interacting with one another or questions of composability. So if you are somebody who works on computer security for example, you understand that you can make a product that's completely secure until you connect it to everything else. And then it turns out that the system might have a set of vulnerabilities that are really distinct from the product. And so this question of how we compose things really matters. And when we're thinking about AI, think about the autonomous vehicles that are being developed. And right now those autonomous vehicles-- the algorithms-- I don't know whether or not there's any standardization happening, and Professor Ghani might have some ideas for us. But my understanding is that no. And that's appropriate. People are trying different things, and they have different kinds of data.

If you look at the conditions under which some of the automakers only let their cars operate-- when the roads are dry and when it's a sunny day, et cetera, and others of them are letting the cars operate all the time and in more urban conditions. They have totally different datasets. And as Pam said, if we're thinking about this being some combination between Professor Ghani and Dr. Dixon, that they're initiating and then they are training on different datasets. So those algorithms-- even if the initiation-- the kinds of algorithms they're using, they're training in super different ways. So imagine one of my favorite comics, the breakfast comic-- I don't remember them but-- there's this great comic and it's like the Rawlsian car and the Kantian car and then along comes the Nietzschean truck and it just like, phew! Takes everybody out.

There's been a fairly rich conversation about what sort of ethical priors should the autonomous vehicle algorithms have. I would say a more pressing concern is how are they going to reason together. They can't all be doing different reasoning or things are going to turn out really badly. And that's just one example. As we think about different AI systems, how do they compose? How do they reason together? And how do humans-- who might be the kind of sole input-- like where the place it's getting all the outputs and providing all the inputs, what does that mean for us as we're trying to understand what we should do based on all those AI outputs?

DUANE POZZA: So is it a good thing? What are the benefits to consumers we think of having these kind of devices? AI being implemented in consumer-facing devices and working together? I'll give it to Morgan first.

MORGAN REED: Absolutely. Most of us are tired of user experiences that don't meet our expectations, that alter with every update, that drive us to both distraction-- literally in the case of automobiles-- and to making really bad decisions when it comes to our financial health. We definitely see that improving the way that people interact with the devices that are part of their life is absolutely essential to better engaging with their needs. I think one of the issues that we all have to face on this and I think Deirdre's point about, how do connected cars reason is absolutely critical. The question that we're going to have to figure out is, which reasoning system is best? And we need to figure that out prior to putting millions of cars on the road.

The question that arises with those of us who are working on this area is, where is the safe sandbox to do this testing? How do we do it? In what environments? Because ultimately, to Deirdre's point about someone driving on a wet road, someone driving on dry roads, and the ultimate confusion area is black ice. Does it flex differently? Is it wet? Is it ice? Do I lengthen the time that the tire stops? Do I change the way the braking happens? Do I alter the distance between cars? So we are going to have to work to develop sandboxes so that the impact of those kinds of algorithms and decision-making processes are improved. But yes. Is it in the net good? Absolutely. If the machines do a better job of working to serve our needs, then absolutely.

KEN SCHNEIDER: Can I make a comment on the investment management space? First, before I speak as an FCC employee I have to provide [INAUDIBLE]

MORGAN REED: Nothing you say! [LAUGHTER]

KEN SCHNEIDER: Of course, the views I express are my own and not those of the commission or any of my colleagues on the staff. [INAUDIBLE] So having said that though, I do feel like it's a very good thing for the investment management industry. This whole automation of investment management has really driven down the cost a lot and made it much more accessible to retail investors. Over history, it's been a few big innovations or changes in the investment management industry; like the index fund, ETFs, and I think this is going to be, as far as I can see, another one. Where we have the automation of investment management making it more accessible to the individuals. I think this year I saw a report that there was approximately $100 billion in assets under management that's being advised by automated advisors. A few years ago that was just a few billion. And estimates are that it might be as much as in the trillions by 2020 worldwide. So I think it's growing a lot, and it's having a big impact on investors. And not only is it driving down costs, but also, account minimums are $500 now for some of these robo- advisors versus $500,000 for some other advisors out there that are managing high-net-worth clients. So it's making it more accessible, encouraging people of all different levels of wealth and particularly younger people to start investing early and to get into that habit. [INAUDIBLE]

PAM DIXON: Can I jump in?

DUANE POZZA: Sure.

PAM DIXON: Just really quickly, there are different models, and they all have different functions. So if you look at-- without going into all the geek speak-- there are for example, models that detect patterns. So really good benefits of that, for example, are the fraud-detection systems, where we get annoying phone calls that won't let us buy coffee in an airport. Or, I guess we're in a different space. But fraud detection in AI has been enormously successful. It's driven the fraud rates of, for example, credit cards quite low because it's been quite accurate. There's different kinds of data-mining algorithms which can find disparate items and connect them together. Health care is a very interesting use for that. And also, if the AI model has been fit correctly, it's not overfit and it's not underfit. It's not overgeneralizing or undergeneralizing, you can also get some good accuracy increases. We've seen this in biometrics for example, when you have a convolutional neural network which you will increase accuracy profoundly. And that assists people directly because you don't have false positives, and that can be very helpful. Particularly when it's being used for, for example, governmental identity purposes. So there are a lot of benefits. I will refrain from going into the risks, but there are benefits. And I would say it's not just benefits, it's almost like saying, OK when we move from horses and buggies to cars, what are the benefits? It's just a huge change and it's along those lines.

RAYID GHANI: One thing I think we should realize is, we'll be calling the AI today what they used to call AI 10 years ago. AI 10 years ago was a bunch of rules. You would write thousands of rules and that was AI. And that's embedded in pretty much every system we consume today. Every service we're getting-- simple things like health insurance. A typical high-rate insurance company has about half a million rules inside the system to process your claim. And so AI has always been in there; it's just as soon as it becomes deployed, it's not called AI because it's now real, it's not magic.

So I think the way to think about it isn't whether it's going to help or not, or does it do something well or not, but does it do it better than what we're doing today? And I think that gets to deepest point of values. If we take a problem and we solve it and now it's 90% correct, well what does that mean? Do we only care about overall correctness, or do we care about how does it have disparate impact on different types of people? If it's 10% wrong, is it 10% wrong on everybody equally? Or is it 80% wrong on certain types of people, and 0.01% wrong on the rest of the people? So the average error is 10%, are we happy with that?

And it's sort of ironic. Where we're talking bias in AI, but look at all the problems we're talking about-- self-driving cars, not being able to buy coffee in airports, devices that we're all getting information from, that's a horribly biased view of the world. I mean we're very close to Oakland, and there's some really real problems. Self-driving cars other than Uber being in Oakland is not a problem for them right now. Buying coffee is not a problem. They're dealing with education issues, health-care issues, criminal-justice issues, public-safety issues, policing issues-- I mean the police department has had how many chiefs fired over the last year?

So I feel like we kind of need to think differently about AI. There are consumers of private services that-- we care about because of the venue and FTC, but a lot of consumers of public services today that are getting subjected or will be subjected to decisions made by AI. How is sentencing going to be? And there's been a lot of work over the last few years around the bail process and how do we use AI to improve the bail process? There's work around people who are going into incarceration. They have other needs.

Two-thirds of our people incarcerated have mental-health issues, health issues, substance-abuse disorders. AI is also being used to redirect them to care instead of putting them into jails and prisons as we're doing today. Policing, with the work we're doing around police misconduct and excessive use of force. AI is also being used to detect which officers are likely to go through that and can we intervene early. Same for HIV. People who are not getting into care, who are not getting tested. There are a lot of issues there, and I think the challenge becomes-- people like us don't work on problems like that.

And so the AI developers, because they're consumers of all the services we talked about, they have an intuition of what could go wrong, what could be good. The car might hit the truck because there was a reflection coming in. So we understand those issues, but we don't understand the other issues. Most of us are not consumers of criminal-justice services, of public-health services. So it's important for us to think about how we balance the issues that our consumers are facing from private services versus issues that consumers are facing from public services, because those issues are very different. Because the risks are much much, much higher, but the benefits can also be higher. So I don't know what people think of it

MORGAN REED: I guess I can tell a positive story, because as you point out, in fact, I would say that AI and the use of remote-patient monitoring is actually having a profound impact in communities that are in the biggest strain. The University of Mississippi Medical Center is the leading academic medical center on remote-patient monitoring. You'd say, Mississippi? Seriously? They're the leader on telemedicine? Absolutely. Mississippi has a huge population of people with obesity and type 2 diabetes. Many in that particular community are terrible about coming in until we're at the amputational stage.

So what they have looked at is, how do we do more to interact earlier? Especially with older men in rural parts of Mississippi who are culturally not particularly happy to go to a doctor's office?

Culturally that's not something that's seen as a positive impact. And they're two hours away. They have a health care professional shortage area. And so how do we actually use AI to reach out to those people? And right now, the leader of telemedicine is the University of Mississippi Medical Center. And so it is being done, It is being looked at, and it is being targeted, exactly as you say, to try to take out some of the human bias that has existed in physicians and medical professionals who for a while, unfortunately, that prejudiced them against treating some of those communities in a way that was more appropriate. So hopefully AI can actually break some of those biased value chains because unfortunately, as we all know, we're all biased.

RAYID GHANI: I think it has the potential to break them.

MORGAN REED: And if we do it right.

RAYID GHANI: It's expensive to break it. Because the cheapest way to build AI is to take existing data and pull some part of it out and test on that. So any false positive numbers, any accuracy numbers-- Pam you were talking about how we can do better-- that's conditioned on being correct on the data you collected. So if you're dealing with any type of fraud problems, those things become really tricky because if you've only investigated 5% of any sort of audit in cases of any sort, 95% were not investigated. So when the system predicts something there, you're only evaluating your system on the 5% to 1% that have. So that requires you to go collect additional data. That requires you to go on randomized experiments. And that's one scary to people, that's expensive, and most of the times, it doesn't really happen today. So I think we have to make sure we invest the extra money that's required and teach people who are using them to evaluate them in a way that's actually correct as opposed to using the historical data that exists.

PAM DIXON: I agree with that completely. And I think that's a profound problem today because we have a lot of really dirty datasets out there that are for sale for really cheap money. And by dirty, I mean that they are really inaccurate.

MORGAN REED: Wanna buy some data? [LAUGHTER]

PAM DIXON: But it's true. OK, I won't nag on data brokers, but the bottom line is-- If you look at actuaries and their job, they're purchasing a lot of retail data and working with a number of insurance clients to figure out, OK, so how do we set rates based on behavior? And people's online shopping habits have been the basis of some algorithms that were of some models that were of identifying risks. And how accurate is that data? And what kind of decisions are impacting people based on this? Was the model fit correctly? I think these are all very important questions. And this is not the benefit, this becomes part of the risks. And there are a lot of financial models that are also incredibly risky. If we have time, I'd rather talk about summarized credit statistics because that's another really problematic model. But I agree with you just wholeheartedly on that.

DUANE POZZA: Before we get into talking about the biases-- which I want to talk about a little bit more, along with how to be correct for them-- before we get away from the benefits, in between buying coffee and large-scale societal problem seems to be something we've see in personal financial management; which, for millennials and the younger generation, is both a personal problem and a sort of broader issue. And some of the sort of FinTech innovations that are in AI seem to be aimed at trying to improve financial management in part by this promise of AI processes making better decisions than people would normally make. Do we think this is an actual potential benefit of personal financial management? Do we think that's going to expand? And is that good for consumers? And I'll ask Paul this question first.

PAUL SCHWARTZ: OK, thank you. I think that a way to think about AI is that there are going to be winners and losers. And we may decide, as Morgan has said, that it's going to be a net benefit to society. But I think we also have to think about the people who are going to be the potential losers. So the big benefit is that there can be a democratization of access to financial services, which we've heard about. I think the potential challenge then is the choices that you make available to different groups and the biases that might get baked into that. And so that's something that there has to be transparency about. The other thing I think to think about is in terms of what Deirdre has said about hand-offs. And I think you are also going to see a series of handoffs between what happens in the AI box and then the law. And so I'll take you quickly through three areas in which I think those handoffs between AI-- whether FinTech or other-- and then the legal system. So we think about the financial silo. The big issue is going to be [INAUDIBLE] fiduciary responsibility. So then the question will be, it seems to me, in programming the AI.

Rayid has said the machine learns. After a while, the machine is kind of on its own doing its thing. You kind of set it off into the universe and it goes where it goes. And so what does it mean to have a fiduciary responsibility in designing these kinds of systems and checking on it? Now let's switch to the torts of silo. So in torts, we've talked about connected cars. And Deirdre's mentioned the kind of Nietzsche truck that takes people out and kills them. Or the other way to think about it-- is the updated version of the trolley problem, which you may remember from college or law school or graduate school. Where you have to pull the switch; and does the trolley kill two people or three people? And so what does it mean then into the tort silo to make a reasonable decision about that and to program it in a reasonable way? And by the way, there will be litigation about all these issues.

Finally, the privacy silo. And the privacy silo, the way we've proceeded for decades now has been through fair-information practices or fair information principles. And so several of them are really going to be very contested now. And so the question is, what does it mean to have in this new world of AI? Second, what does it mean to have consent? At what point in the process do you consent to what? And to whom? How is that collected? And then the final two, three, and four are fair-information principles that are particularly important in Europe.

And so that's the notion of data minimalization. Which means the system, at least in kind of the classic European approach to fair information principles, this system should only collect the minimal of data to make the decision to which you have agreed or consented to. What does that even mean? Big data means we collect everything; we look for interesting correlations. And then finally, in classic EU data protection law going back to French law in the 1970s, we don't like automatic decision making, surprisingly enough. And so automatic decision making is forbidden, and there has to be a human at the end of the process. And so if we're going to scale up our systems to be compatible with EU regulation, what does that mean to have a human approve an automatic decision? So I think, just in summary, in all these silos-- financial, tort law, privacy, AI raises immense challenges and then if we think about the benefits, well, we're going to have winners and losers at the end when the legal-system handoff is made.

DUANE POZZA: So in terms of the potential bias issues that we've been talking about, what do we think the answer is to solving them or making them more transparent or dealing with them?

PAM DIXON: Certainly one of the things that needs to be done is to really investigate the long tail or the noisy data points. So if AI has been fit correctly and trained properly, there's going to be a that occurs. And it's going to smooth out the data. You know all the different aspects of the data, but there's always going to be someone, an individual in certain datasets who's noisy. So if you have a beautiful Gaussian curve or clump of data and there's someone over here-- in financial terms you could think of it as credit scoring. Here's your average credit score and then here's a noisy data point-- maybe the person had cancer and had to go bankrupt. Their credit score will be tanked and there's a human behind that noisy data point and the reason that they didn't fit in the curve. And we've got to think about how we handle the noisy data points in this world.

And that also means that we're going to have to figure out how to interpret the inferences that come out of these models. It's really important. And Paul, to your comments, I really like your set up of that and I think that you're really right. And one of the things that I've been contemplating-- though by no means have arrived at a conclusion, but thinking about-- perhaps notice really turns into interpretability. I'm not sure about that, but I'm thinking about it and seeing how that works out. I might be really wrong, but I'm just thinking about it because something has to change. And Paul's really right about that.

And in terms of historic bias, we have to really look at the datasets that we're using. And this should be done on the front end before the network is trained. If we're training with historical bias we're going to have historical bias somewhere in the inference. So we've got to really watch that and that's tough, and it is expensive, and I think we're really at the beginning of learning that we have to be careful of these kinds of datasets. Even in the financial world, for example salary histories can be incredibly difficult for, for example, women.

RAYID GHANI: I think AI is the tool we want to use. So I imagine you're going to have these adversarial systems in place where you've got people producing different AI systems to make certain decisions. And sometimes they'll be black boxes because a vendor will build a tool, and it will spit out protections and some unsuspecting service organization will use that and do something with it. And we can argue whether that's right or wrong and whether it should be regulated or not, but on the other side needs to be another AI system that's looking at those predictions and recommendations and decisions and auditing them. And we should have a list of the things we care about in this audit.

So if we're saying we're going to take the top 20% least risky people and give them a certain service-- next 20, next 20, next 20, the audits should say, who are the people? I don't care what inputs you used, I don't care what algorithm you used, who are the 20%? How are they different from the rest of the 80%? And we do that today anyway, but now we can start making it more complicated and say, can we start having more complex audit systems that are then given to policymakers to decide, are you OK with this bias? It may be intentional bias, it may not be, but we do want to make sure that these people get the services they need. And so we want to disproportionately put them in that bucket.

The second thing we want to do with AI is try to correct for that bias. And that's tricky because right now a lot of these algorithms, when they're trained, as Pam was saying, they're trained to maximize generalization performance. [INAUDIBLE] And right now we have [INAUDIBLE] that one thing in there is, how well does it generalize to new data? Now, it might be biased data, but that's a separate conversation. So there's some work we're doing we're looking at. But what if we add another term to it? What if we also say, you want to do well on future data, but you also want to be unbiased? And that definition of unbiased again is subjective, but now it puts that definition right up front. Where the policymakers have to say, what bias is OK? What do I mean by bias? Is it disproportionately having people in that bucket or is it disproportionately not having people? Is it false positive or is it false negative? What type of people? How much [INAUDIBLE] Putting these values up front and almost codifying them is really uncomfortable for people. To codify these values.

If you're running a political campaign you're going to put those values on GitHub. Here is my code, and I'm going to govern using these values. And the algorithm uses them and it is transparent. The bottom line for me is that AI is a really useful tool in auditing these systems and correcting for these systems. And there's some early work going on. There's people looking at these tradoffs. If you make it less biased, does it also hurt accuracy of the system? And sometimes it does, and sometimes it doesn't. Because there are 10 equally good models you could pick and right now you're picking the one that's slightly better, but there's no constraint. I'll give the constraint up, I want it to be unbiased. You might pick the one that's 0.01% worse, but it's 50% better on the bias side. And I think right now, we're not thinking that way. But there's some work by Cynthia Dwork at Microsoft Research, she's moving over to Harvard, she's doing some really good work in this area. There is some other work around this tradeoff. I think that's an area that's starting up and that's something we have to encourage and almost help people define what these policies should be so that they have a way of seeding that as opposed to coming up with random policies for debiasing.

DEIRDRE MULLIGAN: So can I just add, I raised three different kinds of biases. There's intentional bias. And there we need to figure out ways in which we can continue to engage in this sort of policing in the marketplace that we've always relied on. People who want to intentionally disadvantage certain portions of the population, they can do it with their policies and processes, they can do it with math. And we need to figure out how to update our regulatory tools. And part of that is, what sort of datasets do you need to test against? What do you have to disclose? Not just about your model but about how you're weighting different variables. What do we need to know about the data you're putting in? And I completely agree with Professor Ghani, here-- the ability to use AI to police AI in a way that is fully independent. We think about software independence. That we can't rely on the machine to police itself. We need to have a completely independent piece of software to police the software. Humans are not going to be so good at that. So that's intent.

How do we deal with the systematic biases that might come in either through data or through the design of the system I think is a much more complicated problem, but one that because we're here at Berkeley, we can think about it. Part of it starts with how we educate people. And if you look at data science programs right now, there are like three of them that actually have classes that require people to think about ethics and legal implications of their data, of how they work with data. And I think that there is an enormous need. In the White House, one of the last reports that came out on AI said, look, we need educational opportunities because most of the people who are designing systems, they're actually not intending to do bad things. Or when they're running data through their system, they're actually looking to optimize for something. They want to do good. And the question is, how do we empower them so they understand their blind spots, they have good tools for cleaning data, all those sorts of things.

But in order for the things that are happening in academia to then translate in a meaningful way into the marketplace, we need incentives. And right now, again, there is a lack of incentives for people to do anything other than go for the 0.001% if we can get it. Because you can always say, well you know, it was the best model. And the ability to say, well it might have been best overall, but we actually care about the disproportionate effect you're having on a section of the population requires us to have a hard conversation, which Professor Ghani was pointing out about what does it mean to be fair. And fairness is a really complicated philosophical question. We have different ideas of fairness in the law, but we can design these systems to be fair. We need a political consensus about what fair is, and I think the Federal Trade Commission has a lot at stake around that political definition. And then the final sorts of biases are the ones that arise from complexity. And there, I think we need to think about who's in the best position to address it.

And I talked about this work by Michael Chance and [INAUDIBLE] and [INAUDIBLE] data, looking at the advertising outputs of Google's ad platform and we can look and see that men and women are getting different ads. And in certain areas of the market, we care about that. If you're advertising financial products or homes, we care about that. I'm also deeply certain that was not Google's intent. But the question is, who's in the best position to understand how it's happening and to potentially do something about it? I would suggest Google, yet we're not going to hold them liable in that particular instance. But we have to think about values in these complex systems in a way that help those sorts of actors who are in a good place to intervene do so.

RAYID GHANI: By the way, if you're interested in these types of issues, there's a workshop that happens every year called FATML-- Fairness Accountability and Transparency in Machine Learning. And it's gone from 20 people the first year, and last year was the fourth year and there were a few hundred people. So look at that. There are lots of good presentations and papers that go into details of things that we're talking about here.

DUANE POZZA: I want to talk just a little bit about, and ask Ken a little bit about the FTC guidance that just came out last week, I think the week before.

KEN SCHNEIDER: I want to briefly mention to Rayid, I'm glad you're talking about using AI to audit other AI systems because one of the issues we have is that it's hard to do that. And we're very concerned of course with fairness and things like intentional bias, and it's very challenging for us to do an audit on that sort of thing. So we kind of rely on technology to audit technology often. Right now there's the Association of Certified Fraud Examiners. We either had or have an upcoming day devoted to regulatory technology, which is going to look at all these issues and others that Pam discussed about identifying fraud and things like that. Because when we have these departments of technical people involved in creating algorithms and machine-learning-based algorithms, who's overseeing that? We could have someone in the compliance department, maybe with legal background, but very little technology background. And that can sometimes create problems, whether it's intentional bias or even errors that aren't identified or are identified and may be covered up in some way.

You may remember, we sued Axa Rosenberg, a quantitative investment advisor, six or seven years ago. He found an error in the model and did not disclose it to clients, or actually maybe misled clients, about what the standard error was, and he would end up paying clients over $200 million. So these things add up. And this highlights the need to have interaction between compliance departments, regulators, and the people coming up with these algorithms. And to your question about the guidance that came up, if you have any questions, we did come up with some guidance to both the industry and to investment advisors engaged in the robo-advisors space, as well as to investors, things they want to keep an eye out for if you're considering investing in a robo-advisor. Things you might want to consider, even as an investor-- is this something that is right for you? And if it is, find out what is their personal interaction, what is the algorithm being overwritten by a person, and if you feel comfortable with that, and so on.

DUANE POZZA: Part of the discussion centers around certain kinds of disclosures that need to be made. Do you think those kind of disclosures are effective to a consumer trying to determine whether or not to use an AI-enabled device?

KEN SCHNEIDER: I think they are. I think they're important. One of the first problems is, how are these disclosures made and are consumers even provided with these disclosures in a clear and meaningful manner? The technology-driven firms a lot of times are on a website somewhere. You scroll down about 40 feet down to the bottom, and you click off that you read everything. We have a little bit higher level of disclosure requirement required of investment advisors. So we want to make sure that information is disclosed to clients. So that's one of our concerns right from the very beginning.

MORGAN REED: Well I'm glad you came back to that because I didn't want to lose track of Pam's very excellent point that was built off of Paul's about how we handle disclosure. And one of the most important things you were saying, Pam, about noise, one of the problems that we face is there are solutions for disclosure and data minimization that Paul talked about. But those very same techniques of whether it's differential privacy or anonymization or others, remove the actual really important noise when it comes to health. So if you think about the information that I might need for behavioral targeted advertising, I probably can scrub that data. I can dump it. But if I don't have that outlier factor, that noise that you talked about that put you up here that I don't recognize, that can be a cancer cluster. If I can find five people similar-- what do they eat? What did they do? How did they do it? What are their exercises? So we have to understand that we're going to have to develop wildly different disclosure and education regarding the kinds of information that we're collecting. How we use it, how long we keep it, and what we do with it.

And I want to give credit. Pam and I spent and Deirdre spent about a year working on short-form privacy notices, and we came out at the end doing some excellent user testing with some folks. And we learned just how wrong we were about how people, all of us privacy experts in a room saying, well this is obvious, let's put this in! And then when you actually handed the device to a user they were like, I don't understand what this means. So I think Ken's point about the 40 page-- how do we handle the 40 page? But Pam and I also learned it's really hard to get the short form right, too. And so we have three kind of disparate buckets here. How do we do data scrubbing, differential privacy limitations, and retain the data in a way that's useful to develop the engines we need?

But then, how do we also retain the data to have the impact on people's health and wellness that they absolutely need to make better decisions and move out of 19th century medicine? So we're going to have to have different regimes about how we report, how we handle, how we protect, how we segregate that data. And that's going to take a lot of sandboxing and work within the industry. And I think Deirdre, a really important point is ultimately, it's going to be the people with the primary incentives who are going to have to be brought to the table and help to make some of those decisions.

RAYID GHANI: I think a third thing, we're talking about the three buckets, right, is there are strong privacy protections. correct services being provided in an unbiased-- and I think you might have to get to a point where you just have to pick two.

MORGAN REED: Be a triangle.

RAYID GHANI: Yeah, you have to pick two because if you want an optimized data you cannot figure out if you're being biased against certain types of people. If you want to be biased, not unbiased and correct, you might need individual data. So it might just be that you have to pick two and you decide as the policy, which two are you comfortable with in different types of services in different industries for different types of problems.

KEN SCHNEIDER: And I think there's a balancing act between fiduciary duty and privacy as well. You're often talking about what is the most information to gather from a client to make sure that you're fulfilling your fiduciary duty. Do you just need to gather four points of data or ask four questions? Is that enough? 12? 20? How much information do you gather? You gather a lot of information and a lot of data to fulfill your fiduciary duty. Now you have logins and passwords of all the other financial accounts, other information on that client, what are you doing with that data? Where is it being collected? And how is it being protected? That's a lot of information to have at your disposal or someone else's disposal who isn't entitled to it.

PAM DIXON: If we look at privacy as a subset of autonomy, which I do, some of these issues fall into a little bit of a different pattern. As opposed to thinking of privacy as hiding information or withdrawing it from public view, I think we've got to shift to a different view where privacy is a subset of autonomy and can grow to something larger. I do think that the tradeoffs are profoundly different, to Professor Ghani's point. I don't think we get to have all the same things that we used to. I think we may get more of some things. I just think that the ground is going to be quite different. I really like the idea of audits. I think that works. Notice is profoundly difficult. And I really like the idea. We're not going to have one giant silver bullet policy, like the Command and Control regulations that came down in the 1960s and '70s. It's just not going to happen that way.

I think we're going to have many, many layers, many different points, so Morgan is discussing that he needs the noisy data points. That means that you're going to have more data in the clear, but there are some instances where you want to really have a smooth fit and reduce the data points. So I think we're going to end up with a lot of different models applied to different sectors in different ways with probably best practices, let's hope. Please. Especially regarding inferences and should this inference even be made in the first place? But something to think is where does all of this go? Where are we going with this? And we're trying to make things better. I think everyone would agree. It's just that getting there, the path is not quite set. And that's, of course, what's very exciting about all of this.

DUANE POZZA: Do you think that consumer expectations are changing or will change to the extent that AI becomes more ubiquitous in things like devices or cars? About the kind of data that's being collected and the purposes for what it is being collected? Or is it too soon to tell?

PAUL SCHWARTZ: So I think that consumer expectations assumes that there's a "there" there, and so this is sort of the Louis CK principle, where he has a whole routine about how he's on the plane, and one second ago he didn't realize there was such a thing as Wi-Fi and the plane. And then the next second, he's bitterly disappointed because it doesn't work as well as he would like. So I think our consumer expectations are set by all kinds of things; including advertising, including marketplace players setting it and shaping it. Ultimately, I think there is a collective question about values and what we think is important. I don't think we could kind of avoid that question, and so Ken was talking about trillions of dollars going into the financial marketplace, potentially with greater the democratization of access to financial service projects, which is wonderful. But then I immediately start praying that the FCC stays in the ring and makes sure that certain kinds of financial practices remain forbidden. And so I think as well, if we think about some of these other silos that I've described, we do have a great need for regulation. Collective setting of standards, and they're not always pushing things back to saying, notice, consent, and we're done.

PAM DIXON: A great example of that is the difference between traditional credit scorers and something called summarized credit statistics. So with the traditional credit score, we have Fair Credit Reporting Act rights. We get to see our credit scores, we know exactly what factors went into that score, we know how to interpret it. But with summarized credit statistics, it is not regulated by the Fair Credit Reporting Act. You get a credit score all right, but it's not reported to you. You don't have those rights. Here's the loophole, the Fair Credit Reporting Act applies to an individual, whereas summarized credit statistics are a very, very complex system with thousands of factors in a really big, giant neural network. And the thing is, they're applied to your ZIP+4, so it's a group. It's like a tiny grouping of households. Very, very small. Like 3, 4 households in some cases.

So you get this sort of little mini collective kind of credit score that pulls you out of the Fair Credit Reporting Act. So you see a lot of actors just purchasing the summarized credit statistics and using them in a variety of machine learning. And it stands in as a proxy for the regulated credit score. And this is the kind of thing that we have to look at really carefully. Because the summarized credit statistics are based on your zip codes or your geography. That introduces just loads and loads of really difficult bias in some case. It could also use factors that would be prohibited otherwise; related to your marital status, related to your age, et cetera, et cetera. So these are exactly the kind of issues we have to work through.

DEIRDRE MULLIGAN: And they present those issues about non-distributed group profiles. There may be people who actually don't fit that because it's not individualized, and certainly the bias that might arise from that sort of treatment.

RAYID GHANI: A different tradeoff that comes up is, you've got a different triangle happening here which is, you want transparency and interpretability. Transparency is not the same as interpretability. You can be transparent, but completely uninterpretable. But you also want the systems to not be gameable. Especially the fraud protection things. And so if it's completely interpretable and transparent, it's also gameable. And so if you make it opaque then-- like 20 years ago, there was a rule of if you have a less than $1 transaction at a gas station, stop your credit card. And that was a very interpretive, a very transparent rule; but it can also be gamed. So you start making bigger transactions, and now you've gamed the system.

I don't have an answer, but I think one open area is, can you develop systems that still satisfy as much of the transparency and interpretability and make it really harder to game it so that gaming it will mean doing the right thing? If the factors that go into your credit score are good things, you can do more of them-- you use less of your credit, and you make more payments on time. If do the right thing, that's the only way to game it is by doing the right thing. I don't know have an answer how you do it, but I think that's something that we have to think about. How do we create incentives for these types of problems?

DEIRDRE MULLIGAN: Well it depends on who has control over the inputs. And in credit scores, I don't have complete control over the inputs. I can behave in ways that produce certain sorts of inputs, but I think that the problem you're posing is true of systems where the actors who are providing the inputs are able to alter them kind of at will. But where the data is generated in other ways or comes from other sources, it becomes less of a problem. But I do I do agree with you generally.

RAYID GHANI: I think it happens a lot. And so one problem that I deal with is policing and these early-warning systems for police officers. And today, the systems are completely useless, but also completely transparent. If you have 3 complaints against you in a 90-day period, raises a flag. If you have six uses of force in 180 day period, it raises a flag. Transparent, interpretable, useless.

DUANE POZZA: I do want to play cruise director a little bit and move us along and talk, before we run out of time, a little bit about something that I'm not sure quite what the shorthand is, but it's sort of the handoff question, or the question of autonomy or responsibility. Where you go from the human is making a decision to giving control over to an AI process? And starting with the basic question that Paul and Deirdre raised about, how is this disclosed to a consumer? How does a consumer understand when they're giving over control to an automated machine process?

PAM DIXON: Right now, I don't think that that happens very frequently. Unless it's over a ripple investor sort of thing. I think it's very rare.

KEN SCHNEIDER: I guess as long as it's in a space where it's being disclosed. We want to make sure it's very clear to the consumer that they're giving up decision making to a machine. When would that machine's decision-making authority cut be overwritten? And when do they have the opportunity to speak to a person? To ask questions? And also, kind of along the lines of fiduciary duty, how is that advisor dealing with anomalies or problems with answers the client is giving into that system? Will someone call them up? Or will a box pop up on the website asking more questions to figure out why you're saying one thing and this seems to contradict it? Understanding how they deal with these sorts of things should be clear and disclosed to the consumer as well.

PAUL SCHWARTZ: So a couple of ways to think about this. One is that there should be a clear moment if you start talking to a machine. And you should tell the person. And why? And I think this then goes back to that French law from 1974, which is the notion that we think humans will be more responsible. That they will care more. And that I guess in the French law it's just kind of creepy if a machine is making your decision. So one approach would be to say to the person, you're now interacting with a financial adviser that's a bot. There's just a machine there. The second answer is it's not really going to matter at the end of the day. And this goes back to research that Joseph Weizenbaum was doing at MIT in the '70s, Sherry Turkle has written about that. Which is that humans have a scary ability to start interacting with the machine.

So the Weizenbaum system in the '70s was called ELIZA. And it was just a computer and it would mimic psycho-therapeutical interactions. And the psychotherapist would just kind of answer what you said-- I'm worried about my brother. Why are you worried about your brother? And so it's about 100 lines of code and it wasn't even a kind of human-looking like robot. And Weizenbaum, to his surprise, found that the graduate students who had programmed it would spend hours interacting with it. And then he found--and it's just a computer-- and then he came in one day and his secretary was sitting there and she actually didn't want him to see what he she was typing in, even though it was clearly a computer and she knew that he and his graduate students had written it. So then the kind of troubling thing or problematic thing is, we can tell people that and it may not matter ultimately to humans.

DEIRDRE MULLIGAN: I think our governance model, at some level, has to try to help people understand that it's never going to be fully a machine and it's never going to be fully human. They're all going to be these hybrid systems.

MORGAN REED: [LAUGHS] That's a movie.

DEIRDRE MULLIGAN: And the question is how do we have governance mechanisms that make sure that the values that we care about that people understand, what the rules are. They understand how the service works. They understand to the extent that the service is biased. Your broker makes a commission off this, so does your bot broker. Or, the service that's providing this is affiliated with this industry, and therefore, you're going to see more of their offers. Whatever it is, whether it's being operated through a machine or through a human investor over the phone, those things have to be disclosed to consumers in a way that is meaningful, which is I think is a challenge regardless. I do think one of the things that's really lovely about code is that even if the algorithm is learning as it goes, at some point there is a set of decisions made about what data was going in and what the rules were, even if it learned over time.

And in that way, it actually provides for a different kind of regulatory oversight, then. It's much harder to document-- well why did you go out of band and give that person credit? And it's one of the things we've seen in the financial sector. While the French and the Europeans are concerned about fully automated decision making. A lot of times we're concerned when somebody overrides the output of the machine. Because there's an assumption that the human is doing it in a way that might not be fair. It may be specifically to give white people a better deal, and so some bias happens. And the question is kind of, how we can best manage it. And there are some real advantages to technology that we can read the code to some extent, or we can at least interrogate it and understand how it behaves.

RAYID GHANI: I think the thing you brought up I really like is the question of, why did you do this. And most of the algorithms that are used today cannot answer that question in a way that a human can understand it. They'll tell you why, but the why would be uninterpretable-- well the values for these variables were these and I multiplied them by these variables, and I added them up and the number was greater than 0.7, and so I said yes. OK, well, why did you multiply them together that way? Well, you told me to get data and combine them together in a way that fits that data, so I did that. So that why is not a satisfying why. And I think the handoff-- I wonder if it's a handoff or if it's a conversation. It's really an interaction that you're having where, why did you do this? Well I did it because-- oh but that data is really biased. Oh, I didn't know that. OK, let me change my weights.

And here's a different decision. Does that make more sense? Like, oh yeah, that makes much more sense. So I wonder if we think about the handoff as the developers and the policy makers having a conversation, and then a separate handoff which is the consumer having the conversation and possibly correcting. Like I remember a few years ago, I can't remember if it was Axiom or Experian, where you could see your own data that they were using through the website and possibly even try to correct some of that. And I wonder if that interaction can become much more of a norm in the AI systems, because right now they're kind of one-shot systems.

DEIRDRE MULLIGAN: Yeah the participation, like we can't treat this sort of technology as a procurement decision. That it has fiduciary implications, it has fairness implications. And so the level of involvement by the non-data scientists has to be pretty high if we want to make sure that it ends up doing what we want. It's too easy for people to say, oh god, those engineers messed up and I'm like well, why was it their job?

KEN SCHNEIDER: Do we have greater expectations on an algorithm or an AI system? Because you go talk to your human advisor you don't ask to see transparency into his decisions, why'd he decide that, what happened; but with an algorithm, we want to be able to look into it and see why it did what it did and understand what's going into it.

RAYID GHANI: Yeah I think the bar is much higher. If a human makes a mistake, oh it's just a human but if the computer makes a mistake, oh the computer made a mistake!

MORGAN REED: But just to really finish, I think that gets to the key in a weird way if we do it right, and it will take time and it will happen gradually, we can actually beat out more of the bias than we have in humans because the human you sit across from might have had a bad sandwich from last night, might have fought with his wife, might have not had a happy day talking to their mom on the phone, any number of things that could have an impact. And so if you were to ask your financial planner to give you insight and transparency to why they were grumpy and didn't give you the good advice or told you to sell, it's what you said. And so what we're really looking for is, how do we sandbox this, and how do we start tearing into those questions in a way that allows us to identify those things where we, as you say, we added the 0.007. And we have to ask ourselves, well why? Was it the fact that the program had a bad sandwich that day? But ultimately, we are going to have a chance to tear into some of the biases that we don't even know exist in our own minds and then start moving to correct them.

DEIRDRE MULLIGAN: If they're in our minds, if there is systemic biases that are in historic patterns of data collection, because these things are trained, we will replicate them.

RAYID GHANI: We'll replicate them, but we will know them.

DEIRDRE MULLIGAN: Hopefully.

RAYID GHANI: We'll replicate them, and then we'll see-- that's the choice we have to make as a society, whether we're OK with that. So I think it is hard.

DUANE POZZA: Last question, and then we'll take a break. Which is, is there a role for, or what is the role for self-regulation in this area? Does it make sense when you're talking about AI as a technology being implemented across lots of different sectors? But what is the role of self-regulatory efforts here?

PAM DIXON: Just to jump in quickly, I'm going to tie this back to notifying people about this. If we look at this as the plumbing in a whole new changed ecosystem, we have to understand that this actually is not about self-regulation, it's about really determining what best practices are in this new environment and then figuring that out first. I think we have a lot of work to do right there. I have been concerned for a long time now about what we do about the traditional FIPPs, Fair Information Practice Principles, and particularly in regards to discussing with people who are impacted by the inferences. How you do that in this environment. I just don't have the answer, but it is completely unreasonable to think that we are going to be able to have privacy policies at the ends of our entire environment and have that be meaningful. And notice needs to really be transmuted into something different, and more, and better. So we've got to really reconstruct that. I like the idea of an audit. Maybe audit or something along those lines, maybe notice throughout the system. Anyhow, the bottom line is, we have a whole ecosystem here, and I think that the idea of self-regulation is a bit early, I'd like to see best practices and ethical data use first.

DEIRDRE MULLIGAN: I think the FTC, the FCC, I assume that you have folks who are able to talk to data scientists and able to talk to people who design algorithms, but you're never going to employ as many of them as either this institution here or industry. And there's a need for people who are close to the problem, and who have the right education and training, to have the right incentives. And I think that they're the people who can help figure out how to do what we want. And so the question is, how do you enlist them in this values work. And I think there's a lot to be done, but at the end of the day, yes, best practices, standards, techniques, are going to need to come from the communities that actually have the expertise.

KEN SCHNEIDER: There really needs to be a lot of communication between the technical people and the legal side of things. We rely on firms to not have a self-regulatory organization necessarily, but self-regulate themselves to some extent and have policies and procedures to address their risks that are inherent in their firm. And we oversee it. Like you said, we don't have a lot people with that degree of technical background to go out on every exam who were exposed to that, just a limited number. So we rely on the firm to some extent. And we need to make sure they have adequate procedures to address them.

MORGAN REED: I don't think I could have actually said it better than Pam did. We're going to have to look at sector by sector, industry best practices before we get to the stage where we actually have industry self-reg. But I think it's absolutely critical that we do it in a way that the outcome makes our lives better. And whether that means healthier, that we get to wage equalization for gender, or any of the other issues that we know about, we need to figure out how do we use this technology to improve the outcomes that I think we all want; and I as an employer, want to see in the marketplace. So Pam, I couldn't agree with you more on that.

DUANE POZZA: Great. And on that note, we're out of time. Thanks again to all of our panelists for a very informative discussion.