FEDERAL TRADE COMMISSION

COMPETITION AND CONSUMER PROTECTION
IN THE 21ST CENTURY

Tuesday, November 13, 2018
9:00 a.m.

Howard University School of Law
2900 Van Ness Street, NW
Washington, D.C. 20008

FEDERAL TRADE COMMISSION

I N D E X

PAGE:

Welcome and Introductory Remarks 3

Introduction to Algorithms, Artificial Intelligence, and Predictive Analytics 6

Opening Address by Michael Kearns 32

Understanding Algorithms, Artificial Intelligence, and Predictive Analytics through Real World Applications 58

Perspectives on Ethics and Common Principles in Algorithms, Artificial Intelligence, and Predictive Analysis 139

Consumer Protection Implications of Algorithms, Artificial Intelligence, and Predictive Analytics 217

WELCOME AND INTRODUCTORY REMARKS

MR. GAVIL: Good morning, everyone. My name is Andy Gavil, and I’m a Professor here at the Howard University School of Law. On behalf of Dean Danielle Holley-Walker, my faculty colleagues, and our students, I’d like to welcome the FTC and all of you to Howard for Hearing Number 7 of the FTC's hearings on Competition and Consumer Protection in the 21st Century. We are very happy to cosponsor today's event, and I want to thank the FTC and the many people at the agency and here at Howard who have worked hard over the past few months to organize these hearings.

As you all know, today's topic is Algorithms, Artificial Intelligence and Predictive Analytics. As is immediately evident from both the list of questions the FTC has posed and the agenda for today and tomorrow's programs, these hearings have been purposefully designed to take a broader and more interdisciplinary perspective than any of the previous ones.

Moving well beyond the usual collection of academic and practicing economists and lawyers, FTC staff have assembled an impressive collection of academics, public servants, technologists, scientists engineers, and industry leaders, but, of course, there's still lots of lawyers and economists.

The goal is to educate the agencies and the broader competition and consumer protection policy community so we can all obtain a better understanding of the technologies that are transforming our economy, as well as our political and social environs. We'll hopefully learn more so we can better understand the business models and practices of our time and so we can differentiate myth from reality, promise from near and long-term prospect.

The ability to take on this kind of prospective study is a hallmark of the FTC and one of its great institutional strengths. It is especially fitting that such a forward-looking approach is being taken here at Howard. Only two years after Howard University was chartered by Congress in 1867, this law school was founded with the aspiration of producing lawyers who would lead the future fight to realize the full promise of the reconstruction amendments to the Constitution of the United States.

Next year, we will celebrate our sesquicentennial, and for that occasion, instead of looking backward, we have selected a theme that looks forward, “The Next 150." As is true for the FTC and for today's hearings, any institution that fails to look forward is bound to fall backward.

In closing, please note that the event is being photographed and webcast and will be posted on the FTC's website, and that by participating all attendees consent to those conditions.

Please also note that our students will be coming and going throughout the day and are available to answer your questions. Please get to know them while you are here and feel free to seek them out if you have any questions or concerns.

Finally, it's my great pleasure to introduce our first presenter. Our scheduled presenter, Michael Kearns, has been slightly delayed, so we're going to start with John Dickerson from the University of Maryland, and hopefully Michael will arrive in time to follow John. Again, welcome, thank you, and enjoy the hearings.

PRESENTATION: INTRODUCTION TO ALGORITHMS,
ARTIFICIAL INTELLIGENCE, AND PREDICTIVE ANALYTICS

MS. GOLDMAN: Hi, I’m Karen Goldman. I'm an attorney in the Office of Policy Planning at the Federal Trade Commission, and I just want to introduce you to John Dickerson, who is an Assistant Professor in the Department of Computer Science at the University of Maryland, College Park. Welcome.

MR. DICKERSON: Thank you, Karen. It’s a pleasure to be here. I am John Dickerson, I'm a, I guess, third-year Assistant Professor at the University of Maryland and right up the street in College Park, and today I'll be talking about an introduction briefly introducing the audience to algorithms, AI, and predictive analytics.

And so for this talk, I'd like to start with a motivational quote which sounds like it was written a long time ago, and that's because it was. So “although machines can perform certain things as well or perhaps better than any of us can, they infallibly fall short in others,” by which means we may deduce that they did not act from knowledge but only from the disposition of their organs.

And this sounds old because it was written a long time ago. It was written by Descartes, who was a philosopher and mathematician in the 1600s. So quite a long time ago, folks were already thinking about what does it mean to think, can we mechanize thought?

Another famous philosopher from the 1600s, Hobbes, states, “reasoning is nothing but reckoning.” So reckoning here is just a reference to mathematics. So reasoning is nothing but mathematics essentially.

And some time passed, 1600s, 1700s, 1800s, until the 1900s, when some breakthroughs occurred in logic and mathematics and philosophy. Folks like Boole, folks like Hilbert, made some breakthroughs in the formalizations of mathematical reasoning. So recall, we think reasoning is nothing but reckoning, and now we can reckon perhaps with mathematics.

So there were some proofs showing that some hard limits -- there are some hard limits to what mathematical reasoning can do, but subject to those limits, folks like Alan Turing came around, Church came around and said there are certain machines -- simple machines -- that for any of these mathematical reasoning problems, subject to these limits, we can create a machine that can do this.

So this is nice, this built on now hundreds of years of philosophy and mathematics, but the general pitch here is that if intelligence can be simulated by mathematical reasoning, that is reasoning is just reckoning, and mathematical reasoning can be simulated by a machine, then can a machine simulate intelligence?

So AI, artificial intelligence, the word was coined by John McCarthy in either 1955 or 1956, depending on how you count, it’s '55 in a proposal, to fund the Dartmouth Summer Research Project on Artificial Intelligence. And you'll hear this called the Dartmouth Conference. This occurred in the summer of 1956.

And there are some fun quotes in there saying basically we can solve artificial intelligence in three months or we can solve artificial intelligence in one generation, but the one I'd like to pull out is that every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it. So even in the 1950s, 1960s, folks were making statements like this.

So a quick spoiler, this hasn't happened yet. We can just shut this down right now. But, progress has been made. So how does that progress occur? Well, this is a cycle of basically R&D

progress that you'll see repeating in the AI world, and this has happened since basically 1956, where some new advance, maybe a new technique, new hardware happens. Fast progress is then made on old, hard problems. So it could be a new mathematical technique, it could be new hardware, GPUs, these graphics processing units, are one of the main drivers in the current sort of fast progress being made on problems that we're seeing now.

But eventually you start to hit road blocks. And at this point, the academic community, the industrial community starts to get pessimistic, this bleeds into the press, and at that point, everyone is pessimistic about progress, funding dries up, progress dries up and so on. We wait until the next large advance.

And so this is the cycle that occurs in most sorts of verticals. It occurs in AI research as well. In AI, though, we call it a cycle of basically AI summers and AI winters. The winters are when funding dries up and nothing happens; the summers are basically what we’re going through right now, where we’re seeing large advances driven by sort of recent hardware and mathematical advances.

So this is a bit pessimistic, this cycle, but like I said, progress has been made. So this has been cycling for arguably maybe six or seven times since the 1950s, but every time we go through this loop, progress is made, new problems are solved, and new problems are encountered.

So what is AI? AI, many definitions, the one I'll use here is the ability to process and act based on information via automation. So we can break this down roughly into four segments. One is perception. I want to be able perceive the world around me. That could be the physical world; that could be the virtual world. I want to be able to learn something about it. So I get some signals about the world, then I learn something about them. Maybe I learn a model.

I want to abstract and generalize that model so that I can use it in other situations. And what do I mean by use? Well, maybe I can reason about this information, I can reason using my model and then act within the world. Again, that could be virtual, that could be physical.

So if I can create this automated system, roughly, I have created what we would call AI. So let's keep moving through this history of AI until we are where we are today. Roughly we can split AI research into some first-wave AI, second-wave AI, and then maybe 2.5 or third-wave, which is where we are right now.

In this first wave, primarily, researchers focused on what is called search. So this is either searching through a potential solution space, some quick examples, chess is a good example here where we had, say, Deep Blue beating Kasparov via a sophisticated algorithm that did search through using domain-specific heuristics, expert knowledge, for instance. Folks who played a lot of chess encoded heuristics into the search algorithm; it would search through the solution space to find, say, the next move to play.

Now, another hallmark of first-wave AI is something called expert systems. And this also relies on basically bringing in a lot of expert knowledge to form some sort of large database of rules, of knowledge, of facts about the world, using some sort of inference engine, typically based on logical reasoning, to make new sort of conclusions based on these facts, and then some sort of action, I/O system to interact with the human. So this is basically the world up until maybe the ‘80s in terms of AI.

Now, there were some large successes here, so one example that I used earlier is this chess champion falling to basically a sophisticated search algorithm. And there are many more. And, in fact, techniques from first-wave AI are still used in practice, but they're decidedly brittle and they really don't have any real learning capability. So they're really sort of a function of just the knowledge that you encode into them.

There's a huge overhead to encoding that knowledge. Right, I have to ask, say, every member of the audience and everyone watching to tell me all the facts that they know about the world and then I have to store that somehow, and that might be brittle and that might not generalizable. They're very, very brittle systems, but they do allow me to do in-depth specific reasoning. Right, if I ask a bunch of experts for facts on a specific vertical, then I can do a lot of fast automated reasoning about just that vertical. So that can be good, but it's very difficult to generalize.

And if you recall back to that earlier slide, we want generalizability, we want abstraction because we want to create some system that's able to encounter new environments and still act in a reasonable way.

So in terms of those four boxes, first-wave AI generally does perception reasonably well in the sense that I have asked all audience members to give me facts, and it can do reasoning and acting reasonably well, but it won't learn and it won't generalize.

Now, there were some transition points in multiple areas of sort of AI research. One of these is something called natural language processing, which says, can I get a computer to ingest, say, raw text or can I get Alexa to ingest signal from your voice and then have it understand that in some sense. So in natural language processing, up until about the late 1980s, most of the rules for doing this sort of translation or understanding were done via hardcoded sort of expert rules.

Around the late ‘80s, probabilistic models started to come into play. Okay, so this is going to sound more like machine learning like folks have maybe heard about in the press. These are models that ingest, in this case, large text corpora and learn patterns in that data.

To look at a different vertical in AI, so autonomous vehicles rely heavily on something called computational vision, which says, hey, I have a video image, can I understand what’s going on in that image. Say I'm a car and I’m driving along, and I have a still image of the road in front of me, can I understand that there's a stop sign and a pedestrian and dog in front of me and so on. So in autonomous vehicles, in the mid 2000s, ARPA ran what they call a Grand Challenge, in fact their first Grand Challenge, which asks, can I create a vehicle that can drive some hundred-plus miles across the desert autonomously?

In 2004, no vehicles completed this task. In fact, I think the longest trip that a vehicle took was something like ten miles. And these vehicles relied heavily on hand-coded rules that say something like, in general, when you're, you know, ten degrees away from the sun and you're driving forward at a particular speed, then a shadow is going to be a shadow instead of a rock with some set of features associated with it. And, again, this is a very brittle system. This is not going to generalize very well.

But then in 2005, five teams completed the entire trip, so 100-plus miles. And this is because they started using these probabilistic models. And, in fact, you can see the general manager for the program, Strat at the time, had a fun quote: “[Vehicles] were scared of their own shadow, hallucinating obstacles when they weren't there.” And this is for those prior systems. And then probabilistic models allowed them to get around this.

So you can see similar transition points throughout all core AI areas, in the late ‘80s, in the ‘90s, up and through basically the mid-2000s. And this happened because of three things. One is computational power increased, and this is the story of basically computation since the ‘40s or ‘50s. This has played a driving role in AI development as well.

Number two, storage costs decreased. I don't have to pay a lot of money to store a lot of data. And, three, everyone in this world now relies on statistical models, maybe with some expert input, but still statistical models.

So this takes us into the second wave of AI, and there's no hard date for this because it happened differently in different verticals in this world. Here, we're relying on this assumption now that we've learned the hard way, multiple times, that encoding all knowledge explicitly does not work. It doesn't scale. It's very brittle and it's very difficult to handle uncertainty.

The new idea is that we should create a general statistical model for a problem domain. We should create a statistical model for natural language or for a type of natural language or for autonomous driving, a type of autonomous driving. What do we do with that model? Well, we feed in data from the real world or maybe simulated data until it looks right. And this is going to be characterized by statistical learning.

So the reason why these models have taken off is because if we input a different data set or, say, set of data sets into these models, we'll learn a different model and then we can deploy that in a different environment. So it's much more generalizable.

Now, some examples. In machine translation, for instance, going back to this natural language that we discussed earlier, we can feed in multilingual text corpora to learn relationships between language. So say we want to translate French to English, one of the early multilingual text corpora came from Canada, where there are rules stating that, say, any government ruling has to appear both in English and French. And so now we have a mapping between English and French documents, we can feed that into a model and we can learn a way to translate between the two systems.

Autonomous vehicles. We can feed in videos and tests of successful driving into a model and then learn what scenarios are safe or not safe or maybe put some error bars around what scenarios are safe in general.

Face detection, face recognition. I can feed in many labeled faces of people. Here is where the face is, or here is where the face is and an idea associated with that, to learn what a face looks like or to learn what, say, your face looks like.

So these types of models are very good at perception, and they're very good at learning. Remember, we're training these models, these general models, based on a data set, and if we feed in a different data set, we're going to get a different result, so they’re reasonably good at abstraction and generalization as well, so long as your model is general enough and so long as you have enough data. But there is no reasoning or acting. I've made no statements about, say, when one should turn the car in -- turn the wheel in the autonomous vehicle.

So a quick example model. Remember, these are systems that rely on statistical learning to train probabilistic models that will tell us something about the world. A quick example is a neural network. So these appear a lot in the news now, which is why I've chosen them, but they're not a new idea. Indeed, that 1955 proposal where McCarthy defined AI, used the term AI for the first time, also discusses neural networks. I believe they were called neuron networks at the time. So this is not a new idea.

The general idea of neural networks is that one should pass information into this input layer, which you see on the left side of the screen. So that information could be pixels of an image. That information could be something with audio signal. It will cascade through the network, along basically a series of pipes that go through nodes, and these pipes have, say, different widths that can be controlled by a learning algorithm.

And then the final layer of this network that has information flowing through it will create some sort of guess. In the case of, say, classifying images, here we have cats and dogs, it's going to create, say, a probabilistic model of whether or not an image is a cat or a dog. And that gives you some signal as to how good or bad your statistical model --

in this case a neural network -- is acting.

A very general model, so long as we can feed information into it via that input layer and so long as we can judge the output and so long as we can actually learn, so make the network better, using sophisticated optimization techniques, we can use this for many problems and, indeed, that is what we've seen, so long as we can, again, train these models through repetitive sort of optimization algorithms.

So another sort of buzzword that one sees in the press a lot is a deep neural network. Again, not a new idea. These existed, I think, since the 1980s, and they're just these neural networks that we had on the last slide but with more, quote, unquote, hidden layers. These are the layers in between that input and that output. So I can add more and more of these. I can create more piping -- intricate piping between these different nodes to learn new patterns in the data.

And sometimes, indeed, we can stack many, many, many, many, many more nodes, so order of hundreds of thousands, millions, et cetera. So these are very large models. And, again, this is because we have increased computational power and cheap storage.

That idea for deep networks has existed

since the ‘80s, but we've seen them taking off in the last five to ten years because of advances in hardware, because of a huge increase in the amount of data that exists. So we have large firms collecting data; we have the government collecting data; and we can now store it cheaply, access it quickly, and because, indeed, from the R&D community, there have been much better methods developed for learning basically how to make a good one of these.

They're hugely successful. They're good at detecting anomalies in data, for instance, credit card fraud. They're good at voice recognition. You've seen Alexa, Siri, Google Assistant, et cetera. They're great at machine translation, language generation, game playing. Some recent high-profile success stories such as AlphaGo playing basically expert-level, Go, DeepStack Plane, expert-level Heads-Up Poker.

Self-driving cars are starting to take off. Video search, audio search, finance, et cetera. These are all success stories in part due to deep learning. Not a new idea, driven by advances in hardware and training them.

Nobody understands why they work very well, and this is a common story in AI as well and this is something that we're seeing more and more appearing, which is humans have sat down, they've designed the network structure, they've designed what those nodes and what those connections between the nodes look like. Maybe they're encoding some domain expertise. There are some known heuristics that you can rely on. There's a trial-and-error process, and maybe actually other AI is actually coming in and trying to train these models or structure these models in a better way, but nobody knows when or why they don't work in general.

So they work well in expectation, which is why we see machine translation systems, which is why we see Alexa and Siri in households now, but when they fail, it can be very confusing, it can be reasonably catastrophic, and it can be very hard to explain.

And some recent research pushes funded by the DOD, funded by industry, funded by nonprofits, have started noticing that, hey, an adversary can exploit this kind of behavior. When I have a system I trust most of the time but it can be exploited in very odd ways and I don't understand why or when that happens, then I can wreak some havoc in these systems.

So I'd like to take a step back. So now we've talked about deep learning, we’ve talked about machine learning, and we've talked about AI. And, roughly, AI is this sort of four-pillar approach to perceiving the world, learning about it, building an abstract and general model, and then using that to act and reason. Machine learning is just one way to build these models, where we do not focus on acting and reasoning but we focus on perception, on learning, on abstraction, and on generalization. And deep learning is just a specific form of basically representational learning, so it's a type of machine learning.

Right, so every time you hear deep learning in the news, you can replace it with machine learning mentally. It's just a way to solve a machine learning problem.

So some present-day movements in AI, understanding bias and methods for debiasing. You'll hear about this I think throughout today and tomorrow, many of the topics on this slide. So this is sort of a teaser. Understanding bias and methods for debiasing. So if I feed skewed training data into these systems -- remember, these are statistical models that are trained on data from somewhere in the world. If I feed skewed data into the system, then I'm going to learn something that represents that skewed data. So how do we understand when that happens and can we create systems that still feed in this biased data which might be the only data that exists but spits out a model that is debiased?

As mentioned before, adversarial reasoning in multi-agent systems, learning to act with cooperative actors, learning to act with adversarial actors, so bringing in older fields such as game theory into these new methods for solving those problems. How do I say design -- well, I'll talk about this in a few slides, but how do I design policies as a firm to compete with other, say, firms that are both cooperative and adversarial? Can I do this based on machine learning?

Also mentioned on the previous slide, robustness to noise, robustness to adversarial attacks, both in terms of theoretical robustness and empirical robustness. How do I design automated systems that fail less, that are robust to attacks and that fail more predictably, because obviously these systems will always fail at some point.

And in that vein, explainable AI, there's a lot of money going into this as well because it's very difficult to interpret the results that come out of these systems from time to time, so can we produce human-understandable models that also work well?

And one final move in the AI community has been reinforcement learning. It's a type of machine learning, but it's a type of machine learning that also focuses on learning to act and reason. So now we're getting closer to that initial definition of artificial intelligence. Here we have an agent, maybe physical, maybe virtual, that's going to act within an environment. It's going to receive a reward signal and then maximize total reward. It wants to find the actions to take for any state in the world such that when it takes that action, it is treated well in the future, it receives reward and expectation in the future. And I'll give you some examples of this at the end of the talk.

So here again, again, reinforcement learning, not a new idea, but deep networks have been used extensively here to revolutionize their use and practice. So here we have deep networks that are used to, say, reduce the complexity of representing the environment. Remember, I can't actually write everything down, I don't want to represent every single fact in my computer, so now I'm going to learn some abstraction of the world and then act on that.

So reinforcement learning is taking us

closer to what we want to call AI. We have perception, we have learning. These are just like machine learning. We have abstraction and generalization, again, moving toward that. Again, if we train these models on different data, we get a different trained model, and we're starting to move toward reasoning and acting here.

So in the context of this audience, I thought I would do maybe a quick deep dive into a few uses of AI, particularly in something called market design. So markets provide agents the opportunity to gain from trade. Many markets require structure to operate efficiently. Market design is going to tackle this problem via what's called economic engineering. So I put on my economist hat and I put on my engineer’s hat and I put on my mathematician's hat. I’m wearing three hats at this point, but I can use these hats to design a market, how do I structure the market, how do I constrain the market such that I achieve some sort of efficiency goals.

AI is increasingly helping with the design of these markets. For instance, these automated methods can use data to help designers characterize families of market structures. They can be used obviously for predictive methods that anticipate, say, future supply and demand in electricity markets or finance markets.

One example, as a computer scientist, this is close to my heart because a lot of the money in our world comes from this, is using AI in online advertising. So online advertising markets generally match advertisers with consumers. Many billions of dollars, and this is an increasing market, many, many billions of dollars are being used here, and it's a driving force in the technology sector.

Machine learning models in this space right now are being used to divide customers into very fine-grained and automatically generated segments. So no longer just male/female but something far, far more fine-grained than that. That’s learned automatically. They're being used to set reserve prices and auctions based on user modeling and bidder behavior, again automatically.

They're being used to automatically generate the creatives, that is, the artwork that you see pop up on your screen, to automatically generate those, say without human input, to fit a specific customer’s predicted wants. All automated.

Reinforcement-learning-based tools -- remember, this is that form of machine learning that also focuses a bit on acting -- are being used to help advertisers, for instance, bid automatically on these very fine-grained segments. Remember, now we have, say, millions of segments. How do I bid on that, I can use a machine-learning-based model to do this.

Another example, AI in electricity markets. Here, matching supply and demand is extremely important. It relies heavily on demand forecasting. Machine-learning-based techniques are going to provide very accurate demand forecasting, which leads to very stable market prices and more efficient power usage.

Reinforcement-learning-based techniques -- remember ML plus some form of acting -- are going to allow us to activate or deactivate expensive heterogeneous power sources to maintain that stability. So I can predict better demand, I can predict better demand at particular time points further into the future, and then I can make a plan to boot up or boot down particular power sources such that I maintain market stability, such that I reduce brownouts and so on. Again, automated.

And my final example is AI and kidney allocation. This is close to my heart. I've done a lot of work in this space. So here, kidney exchanges, for instance, are an organized market where patients with end-stage renal disease enter and are able to swap donors -- willing living donors -- to receive new organs.

It's a really interesting paradigm that's been around for, say, 15 years now, and it accounts for something like 10, 11, 12, 13 percent now of all U.S. living donations of kidneys. Hundreds of transplant centers are involved in this organized market, in fact, multiple organized markets.

And, here, AI-based tools are also operating. Now, this isn't fully automated, but they are, for instance, semiautomatically and optimally subject to human value judgments, matching donors to patients, both in the U.S. and also worldwide. Here, I've called out the United Kingdom and the Netherlands, but in many countries. They're providing sensitivity analysis at a level that humans cannot for new policies. And they're learning from data the quality of, say, potential matches in this market.

Now, let's return to some open questions and some recent pushes which will, I guess, trigger good discussion for the rest of today and tomorrow. So one is, how and why does deep learning work? So I've mentioned not a new idea. Neural networks existed since the ‘50s; deep learning existed since the ‘80s. Now we have new hardware and now we have new training techniques, these tend to work very well in expectation, but when they fail, they fail confusingly. Why do they work?

How can we handle incentives of competing agents? All those three market examples that I showed you, firms are obviously going to compete against each other in this space. The government, regulatory agencies have their own incentives as well. Individual participants have their own incentives. How do we handle this, how do we encode other aspects such as fairness, accountability, and transparency into these systems?

How do I ensure that my automated system doesn't marginalize, say, an already marginalized class in the ever sort of increasing hunt for efficiency? How do I even define this? How do I define fairness? This is a classic question in economics that computer scientists are now starting to struggle with as well. How do I implement this in a scalable way, in an understandable way?

Ethical AI, this will be talked about, I believe, later, by folks like Henry Kautz, how do I divide the labor between policymakers, such as those in this audience, who are ethically trained and ethically minded and technically trained, perhaps ethically neutral AI and machine learning experts? So I can implement, say, a very sophisticated system, but I need to then produce some sort of aggregate output that I can pass back to policymakers to ensure that this is reflecting the aggregate human value judgments of those who control the systems. How do I do that? And there are close ties in this sort of exploration to the world of privacy and the world of social norms.

So in general, our end goal is to create these systems that perceive the world, learn from it, create some sort of generalizable model and then inevitably learn to act using that model. We're not quite there yet, but there's a lot of hope in this space. But I'm going to say that maybe this isn't even the biggest problem. The biggest problem is going to be the interplay between these systems and society, ethical issues, societal norms, human value judgments. How do we play between, say, these sort of sophisticated machine-learning-based approaches to what I've shown here on this slide and the rest of the real world? So I'll leave it at that.

MS. GOLDMAN: Thank you very much, Professor

Dickerson, for that excellent introduction to the field and for the questions that will be coming throughout this hearing.

(Applause.)

OPENING ADDRESS

MS. GOLDMAN: As I mentioned, I'm Karen Goldman. I’m with the Office of Policy Planning at the FTC, and now I'd like to introduce our next speaker, Michael Kearns, who is a Professor at the Department of Computer and Information Science at the University of Pennsylvania. Welcome, and we're looking forward to your address.

DR. KEARNS: Okay, thank you. So not only am I late, but I also missed the deadline to give slides last week. So there's two strikes against me already. But hopefully the time I would have spent hacking PowerPoint I put into thinking instead, and so I'm just going to speak informally from notes.

So as she said, I'm on the computer science faculty at the University of Pennsylvania. My main research area is machine learning and related areas. I've been in this area since I was a doctoral student in the 1980s, before machine learning was a thing in society. And so I was just asked to give some introductory framing remarks based on the agenda that I saw, which contains, like, lots of topics that are near and dear to my technical and related interests.

In particular, following on the last speaker, in the last few years, I’ve been thinking quite a bit about ethical and social issues in the use of machine learning and algorithmic decision-making more generally. And I also saw that there are some discussion or a panel about sort of competition and marketplace questions introduced by machine learning. I hope to make some less technical remarks about that because I think that's less scientific to say there but a lot of interesting things to discuss, and also relatedly topics related to consumer protection and abuses by machine learning and AI.

And so what I just want to do with my time is make some informal remarks, provide some personal opinions on these topics based on my own experiences and research, and, you know, hopefully cue things up for the next couple of days for the rest of the speakers.

So as the last speaker mentioned, there has been a lot of discussion not really first in the technical community but first in the mainstream media and society at large, about the many things that can go wrong when applying machine learning and AI and related methods to algorithmic decision-making.

And before I describe -- say a little bit more about what can go wrong and what we might hope to do about it scientifically, I thought I would start by sort of just framing how things can go wrong in the first place. And so one thing you might wonder is, you know, if there is a lending model or a consumer credit-scoring model that exhibits racial bias, for instance, or there’s some data analysis or machine learning methodology that leaks personal, private consumer data, you might -- it's a reasonable thing to wonder whether this happens through active malfeasance. You know, are there evil programmers at tech companies who, you know, put in a line in their code that says if the person’s race is this then do this; if it's some other race, then do something else; or whether they program back doors into their code that permit privacy leaks.

And there’s good news and bad news here. My strong belief, and I think those people who work in the field would say that, no, there is absolutely no such malfeasance going on by evil programmers at technology and other companies. So that's the good news. The bad news is that the truth might actually be a little worse, which is these sort of collateral damage or consequences are actually the natural byproduct of applying the formal, fundamental, scientific principles of machine learning and AI. And I'll say a little bit more about that.

The vast majority of what I think we would call algorithmic decision-making is actually a little bit more specifically almost always driven by machine learning these days. So, in particular, when you think about the algorithms that make things like lending decisions or decide what ads to show you on Facebook or Google, these generally are not what you should think of as hand-coded or programmed algorithms, but rather they're the result of taking data, you know, historical data, whatever that means in a given domain, giving it to an algorithm, and that algorithm, of course, trains a model on the data. And then at the end of the day, it's the model that's actually making the decisions. It's the model that's actually deployed in the field.

And, typically, the algorithm that transforms the data into a model is actually tremendously simple and very principled from a scientific standpoint. So if I had slides, one thing I like to do in forums like this is put up the Wikipedia pseudocode for the so-called back propagation algorithm for neural networks which the previous speaker mentioned. And that pseudocode is literally a simple loop with about ten lines of code in it.

And a real working version of it wouldn't be that much more complicated. And it's doing the most obvious thing you can possibly imagine, which is essentially going through the training data and adjusting the parameters or nods of the model in order to minimize some -- you know, usually accuracy or error-based cost function, okay? So that algorithm is not opaque at all. It's entirely transparent.

Sometimes, you know, when I talk to people who aren’t in the field, they naturally assume that machine learning algorithms -- you know, the code for them might look like something like I imagine the code to a video game like Grand Theft Auto looking, you know, hundreds of thousands of lines of spaghetti code with all these special cases and details, and it's not like that.

So, then, the natural question to ask next is if the complexity doesn't lie in the algorithms themselves then where does the complexity creep in? And, of course, it's from the interaction of the data being processed to produce a model mediated by these very, very simple algorithms, okay? And so the problems arise these days not so much from the algorithms themselves, which, again, are very simple and operating on very basic, kind of well-motivated scientific principles, the problem is really when you work in extremely large complicated model spaces, of which, you know, deep learning is just one and perhaps the most recent example, the sort of space of models has a lot of sharp corners in it, as I might put it, which allow to you kind of optimize the thing that you're trying to optimize like minimizing the error on the data at the expense of other things that you didn't explicitly ask for like fairness or privacy.

And I think if there’s one kind of important adage to understand about machine learning, it’s that basically modern machine learning will not give you for free anything that you don't ask for and specify. And in general, you shouldn't expect it to avoid things that you don't want that you didn't tell it you didn't want. Okay?

And this is, I think, the source of a lot of the kind of violations of social norms and values that we’ve seen by machine learning and AI in recent years. So that's a little bit about what can go wrong. Now let me talk a little bit about -- sorry, that's a little bit about how things can go wrong.

And, so, with that background, I want to talk about, well, what are the different things that can go wrong, and, most importantly, what can we do about them from a kind of scientific standpoint. So, you know, the things that can wrong are things that I’ve mentioned already, which is violations of things like privacy or fairness or interpretability and transparency, or even safety or morality, if you like. You know, the sort of logical extreme of this for those of you who’ve heard of it is, you know, this sort of parlor game or science fiction thought experiment known as the singularity on which AI, you know, sort of -- AI achieves superhuman intelligence to the point that, you know, for lack of a better term, the robots become our overlords.

While that's a fun thing to think about, I don't know many sane people in the machine learning community who actually think that that's our sort of gravest technological risk anytime soon. All you need to do is come and see what AI and machine learning can actually achieve right now and compare it to humans or other biological species and you will be deeply underwhelmed by what we can accomplish so far. But violations of social norms are, like, already with us now today and on a very large scale, whether we are -- whether we know it or not or whether we're measuring them properly or not.

And, you know, I think it's important to say to this audience that I think I and many of my colleagues, you know, we do believe that better laws and better regulations are possible and should be developed. And I'm sure that that's being worked on and is a necessary activity. But I think my opinion is that that will be woefully inefficient in the algorithmic era to actually keep up with the types of violations of social values that we're seeing because it just -- you know, basically human organizations don't scale, and you can't sort of expect to police the sort of violations I'm talking about with sort of regulatory agencies that are pouring over the decisions made by algorithms on a sort of a human time scale and hope to keep up.

So an alternative approach, which I'm a great advocate of and as are a growing number of people who do technical work in these areas is to design better-behaved algorithms in the first place and to actually endogenize various notions of social norms inside of our algorithms and asking that our algorithms -- that the actual code obey some definition of privacy, some definition of fairness, some definition of morality, if you like.

And, of course, this leads immediately to two very difficult questions. The first difficult question is, you know, how do you define these things as the last speaker said. How do you define algorithmic fairness, how do you define algorithmic privacy? And, then, if and when you can come up with such a definition, it’s going to come at some cost, right.

So if I have some notion of fairness in models that are used to provide criminal sentencing guidelines, my asking for fairness from that model by gender or race will come at a cost of accuracy. What I’m saying is like a tautology. If I sort of -- if I ask myself to find the model in some space of models which minimizes the error period, and then I ask to find the model that minimizes the error subject to your favorite definition of fairness, the error can only get worse.

And so in a model like -- let’s say in a setting like criminal sentencing, this means that a cost to accuracy might mean sort of, you know, hard things to think about. It might mean incarcerating more innocent people, or it might mean letting more guilty people go free. So there will be societal and technical costs to imposing these sorts of constraints on our algorithms, but I think my view and the view of many people in the field is that we have to go down the road, we have to decide algorithms that incorporate these values, we have to talk about what the possible definitions for these values are, and we need to study these tradeoffs between the thing that people optimize for in machine learning, which is accuracy, and the tradeoffs to different social norms.

Okay. And so what I want to do with most of my remaining time is just tell you a little bit about the sort of very active research that's going on in the computer science and machine learning and related communities on this scientific agenda, sort of picking definitions for different social values or norms and actually encoding those norms inside of our algorithms and then trying to study what the tradeoffs will be with, you know, things like accuracy and other more standard objectives.

So let me first talk about the work that goes on in the area of privacy in machine learning, and not just in machine learning but more generally in kind of data analysis and data science. And I think it's helpful to say just a little bit about the distinction between what I’m thinking of as privacy and a closely related and complementary area, which is that of security and cryptography.

So security and cryptography, to a first approximation, is a technology about keeping data locked down. It’s about controlling access to data and making sure that people who shouldn't have access to data don't get that access by basically hacking into a system that they shouldn't hack into. And this is largely the domain of security and cryptography, and that’s one notion of privacy. That’s sort of control of your data and making sure it doesn’t get -- you know, it doesn’t get accessed or stolen by people who shouldn't.

Here, I’m talking about something a little bit different and more nuanced but in many ways is equally as pervasive and important as notions of security, which is the fact that, you know, you have all of this data that’s being collected by different companies and agencies and other organizations. And you might worry about what -- not just sort of, you know, how -- who’s accessing that data but what can be inferred about you from that data that isn’t directly in the data itself.

So the kind of thing that I'm concerned about here is that if your medical record is used as part of a study to build a predictive model, let’s say, for some disease based on symptoms, and then that model is published or used in the field, could it be that the use of that model or the publication of the model, perhaps combined with other publicly available data sets, actually reveals a great deal about your own personal medical status and record. Okay?

And, you know, if you go down the road of thinking about possible technical definitions of this type of privacy, I believe that most of you would eventually come to two kind of, I think, important conclusions or desiderata from any sort of privacy definition for machine learning or data science.

One is that, you know, you need to account for the fact that any particular data set that you want to, you know, keep private in some technical sense, will not be the only data set in the world. And, in particular, that data set might be combined with other data sets that you don't know about or didn’t foresee or don’t even exist yet but might exist in the future.

And one consequence of this that I will

state without proof is that this means that any definition of privacy that it involves anonymization is essentially a flawed definition of privacy, right, because anonymization refers to taking the data set that's in front of you and doing things like eradicating personally identifiable information.

But the literature and news is, you know, rife with examples where you anonymized one data set, somebody else anonymized a second data set. Those two data sets were combined and then maybe combined with some publicly available information, and your specific data could be backed out of that. You could be, as we like to say, reidentified, or the data set could be, you know, deanonymized as they say.

And, you know, I think many people feel strongly enough about this assertion that there is sort of a saying in the field, which is, you know, anonymized data isn't, meaning that, you know, whatever you think you did to deidentify individual identity in a data set, that can often be undone through the unforeseen combination of that data set with other data sets.

The other, I think, sort of axiom for any definition of privacy that's important is that in order to have a definition of privacy that still allows to us do anything useful with data, it's important to isolate, you know, the potential harm that comes to somebody as the result of use of their data in some analysis or model-building exercise versus the harm that might come to them just because data analysis reveals some facts about the world.

So, for instance, if you were a smoker in the early 1950s before there was discovered a link between smoking and lung cancer, well, when somebody did data analysis and discovered that there was a strong correlation between lung cancer and smoking and you were a smoker, that fact does harm to you, but it doesn't matter whether your data was used in that analysis or not, right?

Researchers were going to discover this fact whether your particular data was used or not, and a harm was done to you by the fact that suddenly it's revealed that smoking is bad for your health and you were a smoker and everybody knows it. But that harm was not done to you specifically as a result of the data analysis using your data or not. You were basically -- you know, this harm was going to be done to you whether your particular medical record went into those studies or not.

And so there is a very rich definition of data privacy known as differential privacy that was introduced about 15 years ago and has received a great deal of scientific attention in the interim, and now has a very rich theory and a very rich body of algorithms that basically on the one hand meet this sort of very strong notion of data privacy which has to foresee the possibility of triangulation through the combination of multiple data sets on the one hand but still permits sort of powerful use of data.

And so, you know, one kind of pseudo-theorem that I will state to you is that everything that we pretty much know to do today in machine learning we know also know how to do in a differentially private way. And it's just a matter of companies adopting this technology and choosing to, you know, do their machine learning and data analysis in a differentially private way. And we’re actually starting to see large-scale deployments.

Both Google and Apple use differential privacy in meaningful, large-scale ways in some of their services, and maybe more importantly, the 2020 U.S. Census, every single statistic or report that they release as the result of the 2020 Census they are planning to do so under the constraint of differential privacy. And so this is an example, I think, of a very promising kind of case study, right? Of course, people have thought about different definitions of privacy and data analysis and machine learning for a long time. There was a struggle to kind of come up with the right definition. Many of us believe that sort of definitions based on anonymization are fundamentally flawed.

But then a better definition came along around 15 years ago. There’s been a huge amount of research on this particular definition, and, you know, the good news is that in this particular -- for this particular definition of privacy and this particular social norm, it is possible to sort of give these very powerful guarantees at not too great a cost to accuracy or computational efficiency and the like. We can sort of, you know, have the best of both worlds, if you like.

So let me say a few words about research in fairness in machine learning and algorithmic decision-making, which is much more recent. It’s a much more nascent field than the study of privacy and machine learning and AI, but we already know a fair amount about it. And one of the things we already know about it is that it's going to be a little bit messier then privacy. So my claim is that if you waded into these literatures and you looked at the work that's gone on in deferential privacy and looked in particular at the definition of differential privacy, you perhaps, like many people, might sort of agree that this is sort of the right definition of privacy.

So we already know that there's not going to be a right definition of fairness in algorithmic decision-making and machine learning. And what do I mean by we know there's not going to be a right definition? So there's already from the last several years several examples, several papers which have results of the following form. They basically say, well, you know, whatever your definition of fairness is, wouldn't we all agree you'd want it to have at least the following three properties. And you kind of look at those three properties and you'd say yes, yes, I would definitely want any definition of fairness to at least meet those three properties and probably other stuff, too.

And, then, of course, the punch line of these papers is, well, guess what, here's a theorem proving to you that you cannot simultaneously achieve all three of those properties in any definition of fairness. Okay, so those of you that are -- have kind of an economics or social choice background might know about, like, arrows and possibility theorems for sort of voting systems. These have a very similar flavor.

And these -- and so this has very concrete consequences. So in particular, a typical notion of fairness in machine learning would ask for the approximate equality of false positive or false negative rates across different groups. So let me give an example. Suppose you're trying -- you know, suppose you’re a mortgage company and you’re trying to build a statistical model that tries to predict from people’s loan applications and credit history whether they will repay a loan if you give it to them or not, okay? A very natural thing to want to do. And you want this model so that when people apply you can make a prediction about whether they’ll repay or not, and then you want to give loans to people that will repay you and not give loans to people that you predict won't repay you.

But because this is a statistical model, you’re going to make mistakes. You’re going to have both false positives and false negatives, right? And we might think of false negatives as really the case that causes harm to the consumer in question, right? False negative being somebody who's creditworthy would have repaid the loan if you didn't give it to them, but you decided to reject them, okay?

So we might think of false negatives as a harm inflicted on a consumer, and false positives are sort of the people that got lucky. So a typical definition of fairness would basically say that, look, you're going to make false -- you're going to make false rejections. We're not going to try to prevent that, but across different racial groups, it cannot be the case that the false rejections rates differ wildly. It cannot be the case that the rate at which you falsely reject qualified African American applicants is three times the rate at which you falsely reject qualified white applicants, okay? So this is a very natural constraint. And these impossibility theorems basically say if you ask for that and you also ask for a quality of false positives, i.e., the people got lucky, and one other related condition, you cannot simultaneously achieve all of these.

So what this means is that we already know that in fairness we're going to have to simultaneously entertain multiple competing definitions of fairness.

And so not only will there be sort of tradeoffs in competition between fairness and accuracy, there is even going to be competition between different notions of fairness. If you optimize for one notion of fairness or constrain for one notion of fairness, you might be damaging or losing on another notion of fairness, okay?

But nevertheless, you know, we know this and we have to proceed anyway, and so there's been a great deal of research in the last several years on algorithmic fairness and on different notions of fairness and what the tradeoffs between that particular notion of fairness and accuracy is. And this is an area where, you know, to again echo something the previous speaker said, when you sort of talk about the potential interfaces between technical people and policymakers and other stakeholders, I think there are very, very good starting points.

So one thing you can do, for instance, is if you pick a particular definition of fairness like approximate equality of false rejections in a lending application, and you have a data set in front of you, a historical data set of people who did and didn't repay loans, you can trace out empirically -- I can give you -- I would have shown this slide if I'd met the deadline -- I can actually show you an empirical tradeoff where on the X axis would be the error -- the predictive error of your model. On the Y axis would be a numerical measure of the extent to which you violated this fairness notion, so 5 percent would mean -- sort of 5 percent unfairness means that let’s say between African Americans and other races there’s as much as a 5 percent disparity in the false rejection rates. And I can just trace out a curve for you that shows you the menu of choices you have.

So you can get the lowest error, but, you know, if you sort of ignore fairness entirely, that will give you the lowest error but it will give you the worst unfairness. At the other extreme, I can demand that the false rejection rates differ by 0 percent across populations. It’s a very strong constraint. And I will get the worst error but the most fairness, and in between you’ll get something in between.

And I think this is the type of, you know, sort of quantification of the tensions that we face as a society in making these kinds of decisions that’s the right at least initial interface between, you know, people like me and people like you for lack of a better term, right, because it sort of really shows starkly the choices that you have available.

Just to say a little more about fairness, most definitions of fairness, like the ones I’ve been discussing, actually only hold at the group or aggregate level. So you’re only making promises to sort of groups of people in a statistical sense, and you’re not making promises to individuals, so, you know, sort of more prosaically, if you are a -- you know, if you’re a person of a particular race that was falsely rejected for a loan, you would have repaid that loan, the consolation that you have in these types of definitions is, like, well, we’re also rejecting people from other races who would have repaid their loans at the same rate that we’re rejecting people from your race, which is sort of cold comfort if you're somebody who was mistreated in this way.

And so a lot of recent research, including some of my own, is it trying to move towards definitions of fairness and studying their algorithmic implications that try to make finer-grained promises? Maybe not all the way down at the individual level, but to much finer-grained groups than just things like, you know, race -- you know, top-level race or gender or the like.

So these are two social values -- privacy and fairness -- on which in relative terms we know quite a bit already scientifically. And I think we’re well on the way to kind of developing both a science and an engineering of designing better algorithms and understanding what the tradeoffs are between accuracy and the various definitions of the social values that we’ve come up with.

My former grad student and colleague, Jen Wortman Vaughan, is giving the keynote tomorrow. And she's done a lot of recent research on interpretability and transparency of machine learning, which is another, of course, important social norm. I think we know a lot less there so far, partly because we just haven't had as much time, but one of the problems with sort of coming up with satisfying definitions of things like transparency and interpretability is that there’s fundamentally an observer in kind of the middle of such a phenomenon, right? So when you talk about interpretability, for instance, of a statistical model, you have to talk about interpretability to whom and what reason and in particular the sort of numeracy of the audience you have in mind will matter greatly, right, if we’re talking about interpretability to people with like a statistics training, that means one thing. If we’re talking interpretability to the average American citizen, well, you know, the average American citizen has not been exposed to linear regression and may find it a little bit bewildering to even talk at all about an abstract mathematical mapping from loan applications to lending decisions.

And so I think much of the research that needs to happen on that topic will have to have like a cognitive and behavioral element to it. You’ll need to do human subject studies with the type of audience that you're interested in and ask them what they think is interpretable to them or whether you can explain models to them and the like.

So I'm almost out of time, but just to sort of quickly touch on a couple of other things that I saw on the agenda, I saw that there was one discussion -- there was one panel title that had a very intriguing name, which was Algorithmic Collusion. And I'm not sure exactly what the context that’s meant there is. But, you know, if your concern is that, you know, we might be entering an era where algorithmic decision-making causes in some implicit or explicit kind of large-scale way collusion between different entities, whether it’s on things like pricing or decision-making and the like, I definitely think this is already happening.

One area in which I'm very familiar with this already is on Wall Street where quantitative trading teams tried to build statistical models to predict the directional movement of stocks and, so to speak, beat the market. And my basic belief there is that there’s a huge amount of implicit sort of collusion going on there, and it's really because, you know, when we all use the same or similar data sets, and when we all use the same or similar algorithms to train our models, then even if we think we're clever and independent and creative, we are going to be strongly correlated just through the data, right?

If we’re trying to predict the same thing and we’re using similar data sets and similar methods, then no matter what else we do -- everything else we’ll do is a second-order effect from the fact that the data itself will correlate us. And so I think that this is an interesting topic on which there is probably interesting scientific things to say but I haven't thought about it yet, and I don't know of a large body of research on it

But I'm out of time, so let me stop there and let the agenda move on.

(Applause.)

MS. GOLDMAN: Well, thank you so much, Professor Kearns, for that great overview and introduction to all the issues that will be covered in this hearing.

DR. KEARNS: Okay, thank you.

MS. GOLDMAN: So now it is 10:15, and we're going to be taking a little break until 10:30, at which time we will be back for the first panel.

(End of Presentation.)

UNDERSTANDING ALGORITHMS, ARTIFICIAL INTELLIGENCE, AND PREDICTIVE ANALYTICS THROUGH REAL WORLD APPLICATIONS

MS. GOLDMAN: So we're now going to begin the first panel on Understanding Algorithms, Artificial Intelligence, and Predictive Analytics through Real World Applications. As I mentioned, in case someone had just come in, I'm Karen Goldman. I'm an attorney advisor in the Office of Policy Planning at the Federal Trade Commission. And this is my comoderator, Dr. Harry Keeling, who is an Associate Professor in the Department of Computer Science here at Howard University.

So we hope that this panel, which will cover applications that are currently in use and on the horizon, will provide a sense of the variety of uses to which these digital tools can be put and highlight that no single application is truly representative of their use.

I just want to mention that anyone in the audience who would like to ask questions of panelists should write their questions on the notecards that are being passed out and will be collected later on with.

With that, I'd like to introduce the distinguished members of this panel. So we have Dana

Rao, who is an Executive Vice President and General Counsel of Adobe. Next, we will have Henry Kautz, who is the Division Director for the Division of Information Intelligence Systems in the Directorate for Computer and Information Science and Engineering at the National Science Foundation.

Then we will have Angela Granger, who is Vice President of Analytics at Experian. And then Melissa McSherry, who is Senior Vice President, Global Head of Data Products at Visa. We have Michael Abramoff, who is the Founder and CEO of IDx, and Professor of Ophthalmology and Visual Sciences at the University of Iowa and also Professor of Engineering and Computer -- of Electrical and Computer Engineering and Biomedical Engineering.

And then we will have Teresa Zayas Caban, who is the Chief Scientist at the Office of the National Coordinator for Health Information Technology.

So with that, Dana, would you like to begin your presentation?

MR. RAO: Thank you.

Thank you. Thanks for being here. So the first thing I wanted to just sort of get out there, I'm a lawyer, and people are like, why are you talking about AI, and I thought I'd put it out there because there are some very distinguished computer scientists on this panel. So I was a actually an electrical engineer undergrad and going to a university, and so when I was at law school, I was going to write a paper, a note for the journal, and this book was on my dad's desk, Understanding Neural Networks. This was back in 1996. So I thought, oh, that would be fun to read, and I read it, and I wrote my paper, which got published, called "Neural Networks -- Here, There and Everywhere," which was a wildly inaccurate characterization of where neural networks were in 1996. So don't come to me for your stock advice, but it was -- it's been a fascinating topic for me, and at Adobe, we're really interested in this topic.

And for us, AI is special because we have this entire business that's focused on helping people be creative. And creativity is a part of the brain that doesn't follow rules. It's unstructured, and traditional software programming is a very structured form of algorithms. It's predictive. You understand the rules, and you understand how to characterize it, and that's actually not a great fit for creatives who tend to break rules.

And so our products have always struggled to bridge that gap between innovation and creativity and the structure form of traditional computer programming. And AI bridges that gap, and it really allows us to create tools that are better for our creative customers.

So when we think about how we look at AI and digital creativity, we're really focused on minimizing the mundane, eliminating those repetitive tasks that everybody has in their day. And so for creative professionals, there's a lot of complexity in the tools and in setting up the camera shots or the video shots that are not actually the highest value add that they have, where they're really trying to get their artistic sense across or fulfill the goal of a marketing campaign as they create content for it, the complexity of adjusting each pixel's luminance or the color or the background or the lighting gets in the way of them actually doing the part of the work that they're getting paid to do. So that's where we're really interested in using AI is it would eliminate those mundane tasks.

And we also at Adobe, we've noticed there's a huge demand for content now, and that's either because there's social media channels and people are posting content all the time on Instagram and Snapchat and Facebook, or on ad campaigns -- digital media advertisement campaigns where -- digital marketing campaigns where you are personalizing content for each consumer. So there’s a huge demand for content, more than ever before, and our creative professionals need to be able to create content at a higher velocity, and that's what AI is helping us do.

So when we think about AI, we think about it in the creativity space in two buckets -- content understanding, computational creativity. And Adobe also has an experience intelligence business. I’m not going to talk about that much today, but just for transparency, we also have this other business that also uses AI in a different way.

Content understanding is really trying to get behind what’s in an image, for example, or a video. So it's easy to look at an image of a cat and say there’s a cat, or there’s a house and just do sort of basic object recognition. What AI allows you to do is provide that insight into the image and add an abstract layer, a conceptual layer above what you typically can do pre-AI so we can understand things like actions and concepts and styles and sentiments, so just abstract concepts that are in your image that the AI can infer from looking at it.

And so we have a couple of demos that we're going to show. We're hopeful they're all going to work correctly. I think this is going -- yeah, it's going. And these -- this deck will be published in the Adobe public policy blog, so anyone who wants to see the full deck and watch the videos through can do that. But we're just going to talk through a few -- a couple seconds of these.

So this is a person in the, let me just go back here. Set this up. So this is a person using our stock photography service. And so they wanted to start a creation. And so they wanted to be able to say, I have an ad campaign for Nike, how should I start. And they go to our stock photography site and they just search for things to sort of -- as inspiration for the ad campaign.

And so for example, in this example, this person's going to say, you know, I see this image of this woman with a ribbon jumping. That sort of captures the aesthetic of what I want. And here we go, here we go. And so she -- say they choose this picture, and then what Adobe Stock does, it recommends other pictures that are very similar to this picture. So in this case, she says, okay, I like this, this is a good start for me, and then Adobe Stock at the top does sort of a normal picture recommendation. Here are other pictures of people with ribbons, and that may be what you're looking for.

But in this case, that's not what we want. Like, Nike actually wants this sort of freedom. And so we select the woman jumping, and our AI understands that what we want is actually the action of jump. Like that's what we want out of this picture, not the color, not the ribbon, not the blue sky. We want the action of jump. And so now we actually recommend pictures that are about jumping.

So we can take the concept of that picture and using AI understand, okay, they actually wanted jumping, and so now we can just show them these other pictures.

Now, the next level is we say, okay, well, Nike didn't really want a picture of random people jumping. It was actually supposed to be a family picture. So we take family and we use the jump concept from the first image, so you see how they're stacked on the right, and now you have families jumping. And now the creative professional could say, that's where I want to start, I want to choose one of those pictures and start my campaign from there.

So how do we do it? So what we do is our AI will analyze these -- in this case, an image -- and look for the concepts behind it. So you can see in the middle, there's concepts, and on the right, there are percentages. The percentages are the confidence that our AI is actually accurately predicting what is going on in there. But what you can see is we've analyzed those faces and we've analyzed the context of the picture, and you can see that where you said, oh, there's happiness there, there's love there, there's joy there, we've understood the abstract concept of those pictures. And so you can go, if you're a creative professional, and say I need pictures, my theme is love, you can type in love as a search term, and you're going to get a wide variety of images, but they're going to have this concept in them.

You can also look for families, right? And it will understand that the connection of these three people plus the expression on their face means that they're a family. And you can understand -- and you can search for concepts like family as part of this. And so you can see all the different kinds of categories that you are able to search on using our Adobe AI technology to analyze what is actually going on inside the picture.

We also have a PDF and Acrobat service, you know, and that has lots of text, and we’ve actually run our AI on the text to understand the intelligence behind the words. And we have married that up to images to allow you to do automatic phrasing. And, again, we can do very basic captioning. So you put your photo there, and we can say couple on a bike and that’s object recognition. But then we see AI, then there’s a little slider you can see that’s moving. And you can say I want to see what the AI thinks this is. And it says young couple on a bike, or in this case, it said beautiful peacock, right? So it understands not just the image but also the concepts behind the image. So if you wanted to search for “beautiful,” you’d get that peacock, for example.

So these are the techniques that are being used when we talk about content understanding, the first part of how we looked at AI and creativity. You know, traditional machine learning, it’s traditional deep learning, and when you look at all these things like aesthetics and style and color, as part of the -- we train our AI to understand these concepts, and then we are able to provide these services to our creative professional.

The second piece of what we do is try to make the creative professional day's faster. And that’s what we call computational creativity. And that is trying to help their work flow. How do we help them do those tasks even faster than they used to have to do under traditional software? So here’s an example. Let's say somebody wants -- Macy's wants an ad campaign and they told you to go out and shoot a cityscape at night, and you go out and you spend six months getting this shot. It had the right lighting, the right building, the right angle, and you’re like, all right, I'm great, I’m happy.

And then you turned it in, and Macy's was like, you know what, we’ve changed our mind, we want a different setting. We want it to be the sunset. And so then traditionally, you’d have to go spend another six months reshooting this picture trying to get the lighting correct.

So with our AI, we can automatically segment out the part of the picture that’s of interest to you, and then that's the cityscape. And then we let you import another picture that is of the desired lighting and sky that you want. And with one click, you can now take that lighting and put it in your picture.

So that's probably not 100 percent of what the creative professional wants for their Macy’s campaign, but it's probably 80 percent or 90 percent of what they want, and now they can take this picture and make it into exactly what they want with very little extra effort. So you've just taken six months of extra work, of not exciting work, that was not the fun part of their day. The fun part of their day was setting up that shot to get that image in the first place. And now they can take this and they can go back to Macy's, and if they come back and Macy's says, you know what, we've changed our mind, snowy, blue-sky day, five minutes later, you can just change. And so the AI really helps drive that routine out of your day.

Another example is what we call neural stylization. And so, again, this is the idea that we've been able to understand the style of an image. And so we've trained our AI demonstration on the style of different famous paintings. And so if you have your photograph on the left and you said, I want it to look something like the interpretation of these two different paintings, you can do it. All it does is understand the style of whatever painting you put in, and it's just the style of it. So it's not just copying the colors broadly like you might have expected pre-AI. It understands what the style of the image was and applies it to the image.

So just understanding that concept of -- I think this is going to play. And so this is not just for creative professionals. This is for hobbyists. You can take your own pictures and you can upload whatever artist you want and it’s going to take the style of the artist and apply it to your picture. And it understands that concept.

We can also use AI and we do use AI for our video editing products. So this is a project called Cloak, and this is a normal example where you have -- you shot a scene and then in post-production, you want to get rid of something you don't like. In this case you don't want that couple there. So in using AI, we segment the image and understand who’s in the image and who they are, and we can also fill in the background with copied pixels to make the background look perfect.

So on the left is the original footage, and on the right is post-AI, and it looks like they’ve just vanished, right? And then that used to take months of work to do to edit two people walking out of the video, and now you can do it in minutes.

So as I mentioned, we also have an experience intelligence business. This is the other side of our business. This is a digital marketing business that allows you to target advertisements and allows chief marketing officers to understand what the content in the campaign is doing. So we provide that service and use AI there as well. We use it to help you predict the results of a campaign before you even launch it. We may say this is going to be successful in the northeast, or this is going to be successful in California based on our analysis of customer data from interacting with their website. That's another way we’re using AI at Adobe.

So I think the question is how we get there. How do you actually produce the AI, and I know there’s going to be a lot of people talking about the nuts and bolts of the computer science so I’m not going to spend too much time on this, but this is how Adobe does it. Our AI product called Sensei. And this is the architecture.

And so what we do is -- what we do typical of any neural network, we have the neural network and then we train it with data, and we train it for an outcome. And using this architecture, we're able to create the neural network; we freeze it in place; and we ship it PhotoShop; we ship in Premiere, and that's the result you see as a consumer.

So the principles -- this is my second-to-last slide -- the key principles for training AI that is important to Adobe and just a takeaway for everyone is how do we make this product work as well is we need millions of pieces of data to train it. You need lots of examples of artists; you need lots of examples of images in order to train a neural network to understand the insights that we're able to show you.

So when you think about how do we make AI beneficial, how do we get the rewards of AI, you need access to data, you need access to a lot of data and you need access to a variety of data, and that variety of data will make your neural network accurate. And a variety of data will also eliminate bias.

You can imagine bias when you're looking for images, that is inherent because you may have trained your AI on that particular kind of a person, and if you go searching for a job or an occupation, you're always going to get that person because that's what you trained it with. So the wider variety of data you put into the AI, the more likely it is your results are going to be unbiased.

So thank you for your time. This is our presentation. Creativity in AI, with AI is what Adobe is focused on. It's how we believe AI will help transform the creative professionals for today and tomorrow. Thank you.

MS. GOLDMAN: Thank you so much for that colorful and creative presentation.

So, next, Henry Kautz will begin his presentation.

MR. KAUTZ: Thank you. So I’m going to focus my talk on the work we’ve been doing at NSF to support AI applications for social good. So when we look at a proposal, we have two major criteria. First, we want to advance science or engineering looking at fundamental advances, but we also consider potential for broader positive impacts on society.

Now, the traditional broader impacts that were frequently mentioned in proposals, we’re training graduate students and potential future applications of the result. So someone, say, I'm doing this fundamental research, and maybe someone in the future will come along and do something to benefit society with us.

But, increasingly, we see that the fundamental science and these broader impacts are entwined, that as you work on an application for social good you discover new questions that require fundamental scientific advances. And from those advances, you discover that there are new opportunities.

So AI and broader impacts. So AI methods, taken broadly, that includes machine learning, knowledge representation and reasoning, and what we might call deliberative intelligence, making optimal decisions, are being used by researchers in every discipline that’s funded by NSF. I'm from the computer science, and my particular division funds a lot of the fundamental work in AI, but there’s really no area of NSF now, including the social sciences, where you don't see people talking about AI. And, increasingly, we're partnering with other agencies, funding or taking advantage of work in these fields.

So we’ve seen -- over the last decade, we’ve grown up quite a rich portfolio of what we call cross-cutting programs. So these are interdisciplinary funding opportunities that involve multiple directorates within NSF and sometimes with other agencies. Some of the most important are the Smart and Connected Health Program that we run with NIH. And so there, we are looking at AI research that is a bit more applied than traditional work funded by NSF but is not yet ready for the kinds of actual clinical uses that NIH would fund. So we both put money in there, and then we help bridge the gap between those agencies.

Smart and Connected Communities looks at applications of AI to all kinds of problems facing urban life from pollution, policing, and violence, transportation, other issues. We’ve had a program for several years now called Big Data in science and engineering, and that is to support broad collaborations -- or collaborations that can cover a number of fields. So you might have material scientists together with a computer scientist or, you know, electrical engineer together with the computer scientist or even medical people.

And through that Big Data program, we’ve also funded what are called big data hubs, so the idea that these are a set of universities that act as resources to all of the universities in that region for activities such as helping making connections to government agencies. And through that, we’ve had programs like the Civic Innovation Challenge.

One of our most recent programs that is particularly relevant for broader impacts is one called the Future of Work at the Human Technology Frontier. And it's a very interesting combination of directorates -- computer science, engineering, education, and then our social, behavioral and economic sciences.

So we’re now looking at the future of the workplace and in particular how AI will be impacting that future. So we want to fund both the kind of technology we might see in the future. So, for example, in a recent -- we just completed the first year of the program, and one of the awards was on smart classrooms, so how we might integrate AI as a teacher’s assistant, and not replacing a teacher but assisting a teacher.

But we also will be looking for work where technologists work with social scientists to look at both the positive and the negative consequences. Will AI throw millions of people out of work? That's absolutely an open question. And if you look back at the history of science and technology, you can make quite good arguments either way that AI will lead to permanent unemployment or that AI will lead to new opportunities for employment.

This is another example of the work from this most recent program -- solicitation. So Whole-body Exoskeletons for Advanced Vocational Enhancement. So, here, we’re looking, you know, at something a little bit different than your traditional robotics for manufacturing but augmenting the human worker to give the human worker superhuman strength and endurance, or as I mentioned in classroom teaching, where a system that is monitoring a classroom and noticing when students -- those students who have become apparently disengaged are not working or not attending and can perform such tasks as simply alerting the teacher or engaging in a personalized activity with the student.

So one of our very largest grant programs is called Expeditions in Computing. These are typically $10 million over four to five years. So, here, we’re really looking for research of the highest intellectual merit. All of our reviewing is a system called peer reviewing, where we get unbiased scientific experts from the community to review. And in Expeditions, we have multiple layers of viewing because we really want to get the best of the best.

And in addition, these -- the work we fund should address the nation's greatest needs. So to give just a case study of the synergy between positive broader impacts and scientific merit, I’d like to just mention some of the work going on at the Institute for Computational Sustainability, which is a -- the result of actually two successful Expeditions in Computing that went to a consortium of Cornell, Stanford, and

University of Southern California.

So the problem here is looking at sustainability problems, and by sustainability, we’re looking at environmental sustainability, economic sustainability, resources, social sustainability, very broadly, as complex problems that are really too difficult to solve with human intelligence alone. So we want to employ AI techniques and large amounts of data to solve optimization -- essentially resource optimization problems that are far beyond the kinds of linear optimization that most of the people in this audience would be familiar with.

These are highly nonlinear problems where we must model uncertainty. So we can't -- we just can’t ignore the fact that many -- there are many variables that are unobserved. Okay.

Now, you might think that, well, these are all different problems, but what has been so fascinating by this Expeditions is that problems that seem to be quite different often have very -- have shared technical solutions, okay? So this is a subway map that the research group created. And as we see, each of the tracks of the subway, the six tracks -- the six tracks are scientific themes. So pattern decomposition, crowdsourcing, mechanism design, so social choice theory, and economics, spacio-temporal modeling probabilistic inference, and sequential decision-making. And then each of those tracks is going through the stops, where the stops are the particular application.

So in each application you had domain experts. So let's say there's one there on landscape scale conversation and rural communities. That included, you know, people who knew a lot about that topic and had been studying and working with communities in Ecuador, but it made use there of temporal modeling, probabilistic inference, and sequential decision-making. So you see it's quite a variety here -- flight call detection, and I'll mention that again, wind and solar forecasting, all the way over to microbial fuel cells.

Now, but one thing I should point out is AI covers many things. There's sometimes a tendency because of the great success of what are called artificial neural networks to say that that is AI. And as we just saw from the previous speaker, artificial neural networks are wonderful when you're dealing with patterns, doing pattern recognition, and essentially trying to emulate those parts of intelligence that don't involve essentially logical thinking but are more based on pattern recognition and intuition, the kinds of problems we don't think about when we solve them -- recognizing your friend’s face, right? We don't think consciously about it.

By and large, the work in this particular set of projects, though, involves what we may call your Type 2 intelligence, your deliberative rational intelligence where you have many alternatives to consider. In fact, there is such a large number of alternatives, you can't simply enumerate them all one after the other. You have to have very clever ways of essentially searching through an enormous, sometimes infinite space of possibilities and narrowing in on those points that are near optimal.

So just going down a little bit deeper, the problem of data -- of decomposition in big data. So this is -- so a core technical problem. You have some kind of very complex signal, and you want to reduce it to something simpler, right, to a small -- the one measurement or a small number of measurements. So this is also called dimensionality reduction, source separation, sometimes called segmentation. But it makes use of a body of algorithms that have come up in computer science, electrical engineering, and particularly more and more in work in AI.

So we had a -- there were a series of projects, one on detecting gunshots. And you can imagine security applications in a city. Another one detecting elephant calls. So you can put out audio monitors in the jungle and use that to conduct a census of elephants, right, based on their calls. That same work was then used to detect birdcalls of actually birds in flight for a project with the School of Ornithology at Cornell. And perhaps, surprisingly, is with very few changes, that same algorithm was used in a project on crystal phase mapping, which is in material discovery, so a problem where you’re coming up with a mix of new materials, you hope they have some property, and you’re analyzing the results of shooting x-rays at those new materials.

Another example -- my last example here -- is dealing with hydropower in the Amazon Basin. So there are a great potential for getting more hydropower from the Amazon Basin. And, in fact 170 dams have already been built or under construction, about 300 dams are planned or proposed.

Now, there’s obviously a big problem here. If all of these damns are built, not only will there be quite a lot of devastation to wildlife, but they will become less effective because one damn is going to affect the water flow to another damn. So you have to look at this as a multi-objective optimization problem to balance off energy, fisheries, transportation, and navigation. Obviously, as you put in more dams, you’re going to make river transportation more expensive, and finally looking at the long-term effects, how will all these damns affect the natural flow of sediment and nutrients and how that affect farming. So this becomes a multi-objective optimization problem.

And then the goal is to look at the tradeoffs between these different factors and have a new algorithm that can present, well, here is the possible best tradeoffs. There’s no single best tradeoff, but you can look at that any solutions that don't fall along this line are provably worse, so they're worse in some respect and no better in any other respect. So this tremendously reduces this sort of infinite space of the number of dams and the placement of dams to one that now can be decided by humans. Yeah, that’s showing where they’re the dams.

And interesting that this same effort has led to startups. For example, ATLAS AI, that is basically a for-profit AI for social good company. This also received funding from the Rockefeller

Foundation, looking at providing -- helping developing nations be more sustainable in their agricultural practices. Networks of CompSustNet, a larger network that includes this group of these three universities with others to address these problems.

And with that, I’ll conclude. Thank you.

MS. GOLDMAN: And thank you so much for showing us the diverse portfolio that NSF is supporting.

And, now, Angela Granger will begin her presentation.

MS. GRANGER: Thanks. Sorry, it’s a little tight up here, so we thought that would be the better route to get around.

I lead analytics for Experian, and one of those areas that I'm responsible for is credit scoring product development, and for those of you that don't know, Experian is a global leader in consumer and business credit reporting and marketing services. We support clients in over 80 countries, and we have approximately 17,000 people in 37 different countries.

We believe it's our responsibility to assist lenders in managing consumer credit risk and empowering consumers to understand and responsibly use credit in their financial lives. We’re committed to being the consumers’ credit bureau, and I thank you guys for having me here today.

To set the context for today, there’s a lot of different areas of application for credit scoring, so we’re going to -- I’m going to specifically talk to scores used to assess eligibility for credit where adverse action may be taken. The example was used a couple of times earlier today specifically of application of credit for an example where you could be approved or declined, your application for credit. That would be the credit scoring context we're talking about today.

Benefits of AI or machine learning, for both lenders and consumers in our industry, are ultimately better lending decisions. If you have greater insights into the data that you're using, better accuracy in the scores, you’re going to have better decisions being made.

And, secondarily, financial inclusion. Where we’re really finding the power of AI and machine learning techniques is our ability to evaluate new data sources more quickly and incorporate that new data into credit scores, thus broadening the access for credit for people who maybe have thin credit or are new to credit and don’t have a credit file with us today.

Where we like to start is with the data. If you think about predictive modeling, and any kind of modeling for that matter, it's important to understand the data that’s feeding into the model. For us, we talk about traditional credit data. And when you think about traditional credit data, what we refer to is what you typically find on the core credit databases at the major credit reporting agencies. And this includes information around what we call trade lines or account-level information where you get access to a consumer’s payment history on a certain type of account, their outstanding balances, that sort of thing.

We also have information on inquiries that are made into the credit bureau for applications for credit for example. And we have public record information, particularly on bankruptcies. We also maintain some additional information that you might think of as being part of a credit application, such as income and employment.

We also like to talk about alternative credit data. So this goes by many terms. In our industry, when we say “alternative credit data,” we really mean data that is not on that core credit database that I talked about a minute ago. So types of alternative credit data that aren't reported to the core credit database today can include rental payments, asset ownership, alternative financing such as payday loans, short-term loans, rent-to-own-type loans.

There's additional public record information out there that's not on the core credit database. And, most recently, we've incorporated consumer permission data.

Both alternative data and traditional credit data have been found to be very predictive of a consumer's creditworthiness. And, particularly, the alternative data comes into play in those cases of thin file and no-hit-type consumers that I mentioned a minute ago.

The Fair Credit Reporting Act regulates the collection, dissemination, and use of consumer credit information, and so all data used in credit scores are what we would call FCRA-complaint. What does that mean? That means the data needs to be accurate, so the credit reporting agencies must do their best to ensure their data is accurate. The data is disclosable, so consumers can see that information. Consumers can get one free credit report every 12 months, and they can see their credit information if they're denied credit as an example.

The data furnishers also play a role in the process. They have to confirm information when disputes happen, and they're held to certain turnaround times as well as part of the dispute process. And, lastly, we were set up pretty nice earlier around fairness. Fairness is another part of the FCRA. So scores are -- they cannot discriminate based on these different ECOA factors such as gender, marital status, race, and religion.

So for about 30 years, we've been using scores kind of in their current form, which means they're using this core credit information that I talked about earlier. And so between that and our experience over time, we've come up with things that are generally acceptable in our space, data that complies with those FCRA rules that I mentioned earlier, proven payment information, rental data, account transactions from your demand deposit accounts are generally deemed acceptable. Generally not acceptable are things like social media data, you know, who your Facebook friends are sort of things, and any data that could discriminate in decisions or that could be discriminatory, I should say.

Under consideration right now, we're looking at education level, again, something to help us in that new-to-credit space. Think of students graduating from universities and having that information available so that they can more easily get credit and join the credit ecosystem.

So one of the things about our industry is not only is the data itself, which we just went through, regulated but the scores or the models are regulated as well. There’s regulatory guidelines around accuracy and fairness that have been put out by the OCC. Those documents or those guidelines, I should say, are pretty extensive. They cover the model development process, they cover model use, they cover model monitoring, when to redevelop. And they create quite an extensive amount of documentation.

And in order to meet these model governance guidelines, many of our clients -- so think of, you know, big banks, big lenders -- have had to create entire staffs just to take on this model governance requirements.

We talked about the controls around discrimination which lead to the need for transparency. And then in the FCRA, we are also required to provide your top four reasons for your score being what it is as well. And so the need for transparency, or what we call explainability in scores, is very high.

Some key considerations when developing credit scores to meet all these needs, I won't go through all of these in particular, but they really cover the full life cycle. We talked about, at one of the earlier sessions, generalization. So our models need to essentially replicate. They can't just work really well on the training sample. They have to work well in production. If you think about credit scores in use today -- think about mortgage in particular -- the credit scores being used there are about 20 years old, right? So these models need to continue to replicate and still rank-order consumers in terms of their creditworthiness.

Today, models have an average shelf life of about three years, so we're looking at AI to help us get models to market faster. Some research that we did, we tested several different techniques around machine learning. I won't go into each of them. You can see that here. But suffice it to say the gradient boosting models are the ones for credit scoring and credit risk in particular that seem to be rising to the top.

When we let the machine run by itself, these are the type of results we get. We see anywhere between a 5 percent to 10 percent lift depending on the situation. This is a more generic sample for auto and bank card, so we see about a 5 percent lift if you were to do the math here. But our clients report anywhere up to a 15 percent lift as they start to really look at specific portfolios or specific lenders.

This, however, is when you just let the machine run itself and you don't take into consideration some of those things we talked about earlier.

We do something that we call model refinement, and this is where you have to go in and ensure your model is working as expected, that you can explain what’s happening. You want to make sure that a credit score doesn't go down if a consumer has made some impact to their credit such as paying off some of their debt or lowering their utilization. And if you don't do this refinement and you don’t understand what’s happening under the covers, that can happen.

So when you go in and you refine the model through the requirements that we talked about before, you’ll see that the lift in performance from the -- in this case, extreme gradient boosting methodology, is lessened. So in our particular example, the lift went from 5 percent to 2 percent. In other examples, we’ve see that 15 percent or 10 percent lift come down to 5 to 8 percent, right? So on average, we’re seeing about a 5 percent lift in accuracy from applying some of these techniques outside of our traditional regression methods.

This is just another example of addressing overfitting, which tends to be a problem with some of these new methodologies that aren't -- haven’t been used in practice as long. What you tend to do if you throw all of the data into the machine and let it do its work, we have over 2,000 attributes, variables, characteristics that we will throw into a model, and it will use almost all of them if it can, right.

And that tends to overfit and the model doesn't generalize. And so you do have to go in and manually intervene and not let the machine do all the work.

Some of the advantages for AI in credit scoring go beyond just the modeling. You know, I mentioned a 5 percent improvement, and I'm sure you guys are all sitting there, going, whoo, 5 percent, 5 percent, right? But in the credit risk world and creditworthiness world, we have very predictive models today. And so a 5 percent improvement is actually a big improvement. The data that we use in the models is very accurate, and so we get very good models. So 5 percent improvement is significant, but we're looking to use machine learning and AI methodologies across the model development life cycle and not just in the model development itself.

Probably most importantly to take away from today is in credit scoring. Credit scores are static models. So most of us when we think of AI think of realtime updating, self-learning type models. Those are not in use in our industry today. These are static models. So while we're looking at these additional techniques outside of regression, we're still talking about static models. I mentioned the turnaround time or the shelf life of a model is about three years right now. With these new techniques, that's going to come down, but we have to have the ability to go back in time and replicate our models.

So, lastly, there's some future policy regarding credit scoring that we wanted to make sure you were aware of. Today, unlike what people think, your telephone bill, utility payments are not reported to the credit bureau. Those are very powerful predictors just like other payment methods of future payment behavior and so of creditworthiness. And there’s been several studies that show that today.

And so with that, I would like to thank you for giving me this opportunity and hopefully this gave you a quick glimpse into the status of AI and how it’s being applied in credit scoring. Thank you.

(Applause.)

MS. GOLDMAN: And thank you, Angela, for that very interesting presentation on credit scoring and bringing in the related legal and policy issues.

So, next, Melissa McSherry will begin her presentation.

MS. MCSHERRY: Thank you very much, and thank you so much for having me today. I work with Visa. Visa is the world’s largest payment network, and what that means is basically when you use a Visa card your -- the merchant where you use the Visa card basically calls their bank and says can I authorize this transaction. And then Visa connects the merchant’s bank with your bank, who says yes or no, that's a good transaction to authorize. And then that signal goes back to the merchant, and all of that happens if everything goes according to plan. All of that happens almost instantaneously.

In that -- in that context, Visa is very -- we work very, very hard to make sure that the transactions that are going through are legitimate transactions or not fraudulent transactions. I think fraud worldwide today is something like $600 billion, so it's a lot of money, and we want to make sure that we do as much as we can to help banks prevent any of those fraudulent transactions from going through while still making sure that all of the good transactions go through. Basically, when you are actually the one using your card, if you try to use it, that it actually works.

So what I’m going to talk about today is one way in which Visa is using AI, specifically computer vision, to help us do that work of looking after and preventing fraud on the Visa system.

So you might be asking what do puppy dogs and blueberry muffins have to do with preventing fraud. And I put this up just to sort of illustrate both the challenges and the opportunity in computer vision. So all of you could look at these pictures and very easily discern what’s a blueberry muffin and what’s a puppy dog. But using the techniques that were available up until, you know, call it 2012, 2013, this was actually a pretty hard problem for most computers to solve. They would get it right about 75 percent of the time.

And in I think it was 2013 -- there’s a competition that is run every year. And new techniques, specifically things called convolutional neural networks, started coming into play and started dramatically improving the ability of computers to correctly differentiate the muffin from the dog. And so we're now at the point where these techniques can generally differentiate not just muffins and dogs but can differentiate different images about 97 percent of time as opposed to 75 percent of the time, which is really quite good.

If you think about human beings -- although if you’re sitting there concentrating, you know, you would always be accurate since most people don't concentrate all the time and they do sometimes make careless errors, human beings run at about 95 percent of the time, right, when you give them a lot of images. So this ability to look at a picture and so to say this picture looks like this one, and this other picture looks like this other one, this is one of the applications of AI that has dramatically improved.

And so now I’m going to talk a little bit about how we use that application of that computer vision application of AI in the context of fraud. So just a couple of terms before we get started with this particular example. First of all, what is a fraud score? Like I said, whenever you use a card, Visa basically attaches a score to the transaction that goes to your bank that says how likely is it that we think that this is actually you using your card versus someone who’s trying to commit fraud using your card. We provide that information to the bank so the bank can make a decision about whether or not they want to authorize the transaction.

And as you can imagine, we process a lot of transactions, right? So that first thing we do in every transaction is we attach a score from zero to 99. But then if we look across all of the transactions, we can actually say, for instance, all of the transactions in an hour, how many of them were at, like, the highest score, got a score of 99? How many of them were at the lowest score, got a score of zero. And it's helpful to us to look at the percentage of scores that are in each of those bands.

And the reason why is if you -- if we're running along and 1 percent of the population is getting the highest score, that 99, and it's nice and steady and then all of a sudden like 10 percent of the population is getting a 99, that means that probably one of two things is happening. Either there’s a giant fraud attack, and there are fraudsters that are trying to, in a very coordinated way, steal a lot money, and this does happen sometimes, right, in which case we need to intervene. And we typically intervene by calling the banks that this is happening to.

Or there is something wrong with our models or system or how we’re processing things. And, again, that's a situation in which we need to intervene and we need to make sure that everything is actually working as we expect. So not only do we look at the fraud scores, we also look at the distribution of those scores.

And so the next page, this is just -- this is a made-up example, but I think it sort of illustrates what’s going on. So you can imagine that this is a graph looking at the percentage of transactions in a particular score band. And in this particular case, I just did it over days, and it goes up and down, and it goes up and down because, for instance, the kinds of -- the mix of transactions that you see on like a Friday night can be pretty different than the mix of transactions you see on a Sunday morning. And so the mix of transactions in a particular score band can go up and down.

Now, if you look at this, it's pretty easy, again like the puppy dogs and the muffins. It's pretty easy to see that at the end there’s something that looks a little bit different, right, that doesn't -- that pattern doesn't look like all of the other patterns that came before it.

And this is, again, pretty easy for everyone in the audience to see that that’s different, but it's actually kind of hard for the tools that we had prior to those computer vision tools to pick this up, like you can’t -- like a traditional control chart, it’s hard to write a rule that will get this because the actual numbers are sort of -- they’re inside the range of the historical range, they’re going up, they’re going down. They’re not -- it's just -- it’s hard to write the rules. But, again, it's easy to see it using computer vision tools.

And so what the computer vision tools let us do is basically do what a person would do in terms of looking at this and seeing a pattern that's different. But the computer vision tools let us do that every hour of every day. I mean, the computer doesn't get tired and people do, like, they need to go do something else other than look at charts all day.

It lets us look at hundreds of metrics, not just one, right? And if you think about this, this is a pretty simple chart that I put up here, right? This is basically one-dimensional, right? We sort of look at the scores, versus one-dimension. And so it's easy to see the variation. If I had put a chart up here that had multiple dimensions, like we were varying a couple things at the same time, that very quickly gets really hard, even for people, to see the differences. But, again, the computer vision techniques that we've been talking about can pick those variations up pretty quickly and can identify those. So we can not only monitor what's going on versus one dimension, we can monitor what's going on versus multiple dimensions, and it makes our monitoring that much better and that much faster.

So just a quick explanation of how we've applied this in our particular situation. Basically, we built a model that looks at the distribution of each of those score bands that we just talked about, so, you know, for instance, scores of 10 to 19, right, so it does this for each score band. And it looks at those distributions for a five-hour period over each of the last 120 days. Right, so this is lots of data that’s coming in. Think of the computer as looking at a chart, an hourly chart over the 120 days.

And from that, it forms an expectation of what the current five-hour period is going to be, right? Is the score -- is the distribution going to be going up and then down? Is it going to be going down -- you know, down and then up? Is it going to be going, you know, one direction -- it forms an expectation. And, then, and this is the part that relates back to the puppy dogs and the muffins, it looks at the actual picture and it compares it to its expectation that it created based on the last 120 days, right?

And so on the top row, we see on the right is sort of what we would expect, right, for this time period from the data that’s come in over the last 120 days. And what we see on the left is what actually came in. In those two pictures, the computer would say, yep, those two things -- they look similar, they're both blueberry muffins or they’re both puppy dogs, right?

But on the lower band, what we see is the expectation for the particular time period that we're looking at is just that the scores will be going up during the time period. But what we actually see is that they're going up and then coming back down. And the computer at that point says, no, no, no, these do not look like they’re the same. This is not -- something is not matching here.

And that, then, causes the system to generate an alert and say, hey, a person, a human being, needs to go look at this, right? It might be that it’s fine. It might be that it’s just, I don’t know, Black Friday, right, and so all kinds of things are different. Or it might be that there is an actual problem and we need to get engaged and figure out what the problem is, and we need to figure that out promptly.

So in this particular case, what’s going on is the computer is basically taking a lot of work that might have been kind of boring and tedious for the people and doing the boring and tedious part and then just pulling out the things that are interesting and require human intervention so that human can then go and figure out what we actually need to do differently.

One thing I just want to call out about this particular example is, you know, so Visa is using a lot of different AI techniques across a lot of different places in our system. These particular

can you replicate that or what alternatives should there be?

MR. CALO: Fred is a deep expert on notice and choice, one of the leading experts on notice and choice in America. But I will hazard something which is that what is interesting about artificial intelligence, at least when we come to embody it in an agent, which is something that somebody asked about, is that it can be awfully contextual and dynamic.

So I think that we ought to be encouraging -- you know, the possibility of having a conversation with Alexa about Amazon’s privacy practices is, I think, quite exciting, you know, and the idea -- maybe you are anti, but the idea being that you could ask specific questions rather than have some stupid thing that was like really long and you are never going to read it. But you could say, hey, Alexa, can Amazon turn on you remotely to listen in on a conversation, and then get an answer about that. I think that is actually pretty powerful, personally.

MR. GILLULA: I am not anti, I just think maybe only the people on this panel would find it super exciting to have a conversation about Alexa about --

(Laughter.)

MR. GILLULA: Which is not to say it would be me, I agree. I just do not think the vast majority of consumers would get a ton out of it.

MR. CALO: I mean, I think it is critical when you are thinking about emerging technology generally not just think focus on what is loss, but what new affordances might be there or what you might gain. I think that these things are quite powerful. I think we are getting to a place where natural conversations are becoming more viable and I think that we should therefore -- I mean, if you think about it, notice and choice, we have been operating under basically Gutenberg technology all this time, right?

We just publish a long thing whether it is a digital or a print, just a bunch of words on a page. Yet, you know, here we have companies that are doing these amazing things about organizing information and gauging you and so on. Anyway, I think there is a lot of innovation that could be occurring with notice. And part of it would be to contextualize and actually answer questions about this consumer instead of just having something that no one reads.

MR. CATE: I would echo everything Ryan said. I would just like to make two points. One is we put in the record a paper that I did with some colleagues at the Center for Information Policy Leadership about AI, how it is used today and some of the issues it raises, and one of the things we talk about in there is the way AI is already being used to enhance privacy protections, not just to make them more easily understood or explainable, but to actually activate them. So in other words, you can identify somebody's privacy preferences as they start expressing them and then you can start predicting what they will be so that you offer them the default they are more likely to care about. Rather than the default that you want, you try to give them the default that they want.

I would say just, in general, though, back to the original question on notice and choice. As I said earlier, we have relied on this largely because we have not known what else to rely on for 50 years now, with not a lot to show for it. And so I think we should recognize that notice should be used and choice only where there is something meaningful to tell the individual and only where there is something they can do about it. So I think it is terrific when my iPhone says, did you know this app is using your contacts, do you want to permit that? That is meaningful notice and I can do something. I can say yes or no, I can alter it.

But making my doctor add another paragraph to the 65 paragraphs of the HIPAA notice saying, by the way, your scans are going to be read by AI and, by the way, you have no choice about that whatsoever because it is far more accurate than humans, I am not sure that is overly meaningful. I think we have to be very contextual with notice because the effect when we do not is that we just teach people to ignore all of it. We get people in the habit of knowing that notice is meaningless and so they do not read it, whereas if we would use notice when there actually is something worth telling them and something they can do about it, we might resurrect notice as a meaningful data protection tool.

Now, having said that, I am not disagreeing with Irene. The law requires, both in Europe and in some industries in the United States, notice and choice, it is just bad law. In other words, it is causing people to ignore these notices by providing them when you cannot do anything about them and nobody cares.

So one of the things we often talk about at universities is, you know, a teachable moment. You know, you can only teach someone when they are interested in learning. Similarly, you can only give meaningful notice when there is something that is going to cause them to care about it. And that cannot be they woke up in the morning or they went to a doctor’s office. It might very well be where there is an event, there is something happening, there has been some effect on them, there is some reason that they would care, and then using the tools that Ryan was talking about would be fabulous to really make notice meaningful and interactive.

MS. LIU: There is always a conflict within companies with product design when you are trying to design products that is easy to use and that is easy to understand. When you are throwing in all sorts of consents and notices, it can make it really difficult. And so there is often a conflict between the lawyers and the product design teams about how can we make it look beautiful without all your verbiage. So that is something that we struggle with.

And I completely agree with Fred that meaningful consent is ultimately more beneficial to society and to consumers for how their information is being used and how the company is using it versus just providing our lengthy privacy policies that most companies have.

MS. GEORGE: And as a corollary to that, does the notion of opt-out work in an AI context and does that vary based on I think the stage of the product life cycle, be that data collection, you know, product design when it is rolled out to market and being used or other instances?

MS. LIU: Jeremy and I were talking about this earlier. So from a GDPR standpoint, you do have a right to erase your data. So there is an obligation for companies to be able to remove that data. And depending on how you configure that information, it can be difficult. So that is something that you need to think about from the beginning in the design phase to ensure that companies, especially with the California Privacy Act as well, it is important to design these products in such a way that there is an opt-out notion.

To opt out of AI, typically if a company -- if someone wants to opt out of AI completely, that is like let’s say if I am using Netflix and I want to opt out of using the choices, the different types of videos or shows that they are showing to me, it is basically opting out of using Netflix completely. So you have to think about, like, are you trying to opt out of the product or are you trying to opt out of the database use as well? So there are different ways of viewing opt out, and I think Jeremy can probably talk more about the technological ways of opting out.

MR. GILLULA: Yeah, there has been some recent papers that show that for neural networks you can actually reconstruct what the training data was if you are given sufficient time and access to -- and able to run test data through the neural network, which basically means that if I am a service and I used your data to train my neural network, I cannot remove your -- the fact that you are -- the imprint your data has left on my neural network basically without retraining it from scratch and retraining it, once again, without your data. So it is technically -- is it technically possible? Yes. Is that potentially a huge burden on the company? Potentially, yes.

Then there is the other question of, how much benefit do you get from having your imprint removed from whatever model was generated? Because it does take quite a bit of effort to reconstruct all of the training data, and so that is in the unlikely but feasible attack. So there we do have to get sort of into this balancing act a little bit.

MS. WORTHMAN: Another question from the

audience. In cases where autonomous systems result in consumer harm, who should be held liable and to what degree?

MR. GILLULA: Just send the robots to jail.

(Laughter.)

MR. CALO: Well, I mean, I think that is a genuine puzzle. I mean, so you have -- in criminal law and in tort law, we generally require that you do something either on purpose or that you -- a reasonable person would be able to foresee the category of harm that occurred, right? And so when you, for example, have a bot, which this really happened, that is supposed to buy things randomly on the web and buys methamphetamine and the police come and say, you know, you bought methamphetamine, and you say, no, no, it was the bot, right?

Or in another instance, where a company made a bot that was arguably hacked into, but at least was subverted by trolls that wound up denying the holocaust which is not lawful in some jurisdictions where -- that had access to this bot. You know, you would be sort of hard pressed to bring a criminal case to it. And certainly in many categories where -- something happens where the system just behaves in a way that was not anticipated, you do not have what is called proximate causation for purpose of bringing a tort lawsuit, which is what I teach.

And that is not a great place to be because you wind up in a situation where you have victims, but not perpetrators. And I do not know how much that would really matter to FTC enforcement, specifically, because I think you could get around it just by saying, look, you created these conditions that were deceptive or unreasonable and these unexpected things happened, but something was going to go wrong. But I think it is pretty serious in tort and criminal law. I think it is hard.

MS. WORTHMAN: We have also had a question from the audience about retail price discrimination at the individual consumer level and what is the material harm to the consumer in price discrimination and maybe price discrimination can be sort of whether or not it is advertising different things, not on a prohibited basis under ECOA, but just because you are using a different type of computer, because you are purchasing tickets on your mobile rather than on a laptop. What is the harm, what is the cost-benefit analysis in that particular instance?

MR. CATE: So this is a place where actually notice would be quite useful. This would be much more, in other words, to say if you visit this website, we are going to use pricing based on information about where you are coming from, the computer you are using, whatever because it would then empower you to say, well, I am going to go have my friend check and see what the price is to see if I can get a better price. In other words, that would be actionable notice, you could really do it. And by the way, having to disclose it would probably slow people down -- companies down actually wanting to do that.

I mean, remember, we have discriminated on price for forever, I mean, for generations. Every time you fly, there could not be -- there is more discrimination for all sorts of reasons, how long you are willing to stay, what nights you will stay, and so forth. We discriminate based on zip code, we discriminate based on all sorts of other information that have been imperfect. Now, we are going to be able to discriminate better. I mean, we are going to have both better technology and better data and the two together are going to make much more precise discrimination. You know exactly what I will pay. eBay knows exactly what I will pay for something because it has watched me pay that for years.

So this is actually a place where you could say, first of all, we need to figure out is that a harm. Is it something we are going to say is unfair? Is it something that we are going to say causes injury? And if not, maybe disclosure is sufficient. To say, look, we are not willing to say we are going to prohibit it, but we are going to say you get notice. So now, you can figure out if you want to try to come back at the system the other way. They are doing it to you, can you do it to them?

But this is why we have to remember, again, it is going to be very contextual and it is not something new. It is not something AI is going to create. AI is going to make it better in the sense of potentially more precise or more tailored.

MR. CALO: I will give you my two favorite examples of price discrimination after -- I mean, and by favorite, I do not mean I like them. One of them was a couple of years ago a marketing firm was using this tool to figure out when women felt worse about themselves and they labeled these “prime vulnerability moments.” And they suggested that perhaps you should advertise or charge people more during those moments, you know what I mean? That strikes me as not a very good use of price discrimination.

Another one of my favorites, although they claim they never did this, is when Uber experimented with figuring out whether you would be more willing to pay surge prices when your battery was low on your phone because maybe you would get stranded there. Lovely, also. They say they have never done this and I believe them about that.

The issue is not price discrimination. The issue is taking advantage of people, which happens, it happens a lot. And, yes, from an economic perspective, better information is better. Maybe we would worry at one level about all the social surplus going to the firm. You know, they know your reservation price. There is no windfall for you because they charge you more if you would be willing to pay more, so they get social surplus. We have seemed to have moved away from the original understanding of how consumer protection worked, which was that it was immoral for firms to extricate all of the social -- we seem to have moved away from that model, and that is fine.

But I think it is the advantage-taking that I really would worry about, and that is the kind of thing I want there to be hard questions asked about.

MS. LOPEZ-GALDOS: Yeah. What I think is that the questions we are addressing here, like from the liability question and the answer from a total perspective to this question right now is that there are no new issues. Discrimination, price discrimination has existed forever. It does not matter whether a machine makes the decision or not, the debate is the debate. We should analyze whether we still -- whether price discrimination, for example, is procompetitive or not or on the other side whether consumers are being harmed or not, which approach we want to take. But it is a debate that we should be having and we have been having even without machines.

So I think we just need to continue talking about these things, but I do not think it makes a difference whether a machine makes a decision or a human being makes that decision.

MR. GILLULA: So actually, I want to disagree. There is something fundamentally different. And if you lump in AI and big data and predictive analytics altogether, then I agree there is nothing new separate on AI. But a major difference is that now there is a -- you are making a decision based on a tremendous amount of data that has been collected as opposed to just like, say, one data point that you happen to notice or one data point you got like the zip code or how many nights you want to stay for the flight or something like that, something that is very clear.

Now, you can potentially make price discrimination decisions based on what websites the person was visiting. Were they looking at budget travel websites versus high-end travel websites? And then there is the question of what happens if -- how were the price discrimination decisions made if you do not have any data on the person? And do they suffer a penalty for preserving their privacy?

If I use a tracker blocker app on my phone and I go to your website and I try to buy a plane ticket and you do not have any history, am I automatically categorized as I have to pay the highest price or not as a punishment for not giving you data about what level I might be willing to pay? So I think that is a difference as opposed to say, you know, what we have been doing for generations. It is not different versus what we have been doing for the last 10 or 15 years.

MS. WORTHMAN: Following up on that, though, is there -- even though these are problems that we have faced before, are there any particular harms that are new based on price discrimination from AI or that is a result of AI? Any new types of harms or is this just the same thing that we have seen before?

MR. CALO: Well, I think there is a huge difference. I think that -- again, I do think you have to group together a bunch of different phenomenon. It is not AI particularly. But, you know, look, for a long, long time, companies have noticed that -- and not just companies like mom and pop shops, everybody has noticed, that there are just cognitive limitations that we all have, right? We just have these limitations to our rationality and that is why everything costs $9.99, right? I mean, obviously, okay?

There is a set of cognitive limitations that behavioral law and economists, Ariely, Kahneman, and so on have -- Christy Jolls at Yale -- have been surfacing over a long period of time. And these are things like prospect theory and status quo bias. And sometimes the FTC actually intervenes and says, you seem to be using status quo bias here with these rebates. We are going to intervene because it does not seem to be fair and you do not seem to understand what is going on.

The issue is that even with all these behavioral economists thinking about how we have cognitive limitations, the list of cognitive limitations is about 45 long, okay?. What artificial intelligence permits you to do because it is so good at pattern matching is to model what rational consumer behavior would look like in a particular environment and then look for deviations that are particular to you, even if they are explicable. Turns out when you are watching “Buffy the Vampire Slayer” on Tuesday night, you are going to pay more for ice cream. I know I am. But the point of the matter is that there will be situations that are very, very specific to you and perhaps not even have a theory behind them.

But what it allows is the mass production of bias. That is what it allows. It allows these systems to figure out where you are specifically susceptible. And, indeed, we see early signs of this already. I mean, you heard earlier a presentation about how Netflix is showing different people posters for shows based on guesses about their demographics or their qualities. You know, that is part of the phenomenon that in the literature is referred to as persuasion profiling, the idea that not just that you be matched to your interests, but that the messages to you to sell you things would be matched to your unique vulnerability.

So, for example, for some reason in your life you are really worried about scarcity, well, that advertisement will say, “while supplies last,” right? And this is the kind of move that marketers are making and it is only possible because of the way that we are mediated by digital technology and we have these intense analytic capabilities and, respectfully, I think that is am enormous distinction from what has come before.

MS. LOPEZ-GALDOS: So obviously, before, we did not have self-driving cars and now, apparently, we are going to have self-driving cars. So we are going to see new things happening. Now, a self-driving car might just cross over a person. What I was trying -- the point that I was trying to make is that the thought process of analyzing the problems and analyzing who is at fault, what was the causality, I mean, the thought process is the same. The same that exists without human beings is just applied to the new setting.

I think the theories and thought process should remain -- we should not think in the abstract. We should think like we have a lot of analysis in tort law, for example, and we want to say who is responsible, who is not. In a self-driving car, there is software, hardware, there are apps, there might be somebody inside the car that was doing something as well. And what I mean is that in the thought -- when we are analyzing who is at fault and who is liable for crossing over two people, the thought process of, for example, causality should be the same as without AI. That is an example -- for example, of the point that I was trying to make.

MS. GEORGE: So I am going to ask one final question and then I think we are going to wrap up. It is going to be a compound question. Because I like that.

So are there ways in which the FTC should expand or rethink the notions of unfairness and deception when it comes to AI and what educational role should the FTC play with these new technologies, both for consumers and businesses?

Marianela, do you want to start?

MS. LOPEZ-GALDOS: I think it is a very good final question. I think the FTC is doing a great job in putting together these hearings, as I said in the beginning. I think AI is just a machine learning -- it is at a nascent moment. I think it is very important to keep having a dialogue with businesses, with the community, with the consumers, with experts, and see where we are going to and see whether there is anything that needs to be refined, for example, of existing laws or not.

But what is very, very important is not to think in the abstract of AI. We talk about AI as if -- you know, at this moment, there are marvelous things that can be done. I think there is a lot of potential, but I really think that before stepping and regulating or saying, oh, this is going to be a disaster, everything is going to be mass-biased, et cetera, we really need to understand where we stand, what engineers can do, what companies are working on.

I think companies, at least the ones that CCIA represents, are willing to cooperate with the authorities, are willing to engage in adopting principles. And I think having an open and frank dialogue about what is going on is key to make sure we get the right approach. So society can really profit from AI.

MS. GEORGE: Irene, you just want to continue down the line?

MS. LIU: Sure. Again from the beginning, I feel that the FTC framework and the existing laws are sufficient and the fact that it is broad enough that it can capture AI, I think that is great. I think FTC has withstood the test of time because it is broad.

But at the same time, I do think -- I agree with Marianela that it is important for the FTC to continue talking to the industry, also with other regulators and academics to make sure that they are abreast of this nascent technology.

There is also movements across the globe, it is not just a U.S. phenomenon, but just globally. Again, there is a recent universal guideline for AI that was launched in 2018 by a number of data protection officers recently. The World Economic Forum is working on this issue. Regulators in Europe, China, have taken a deep interest in AI and so there is a lot of cross-country developments within AI as well that the FTC can also engage in to make sure that it stays ahead in terms of the policy developments around the world so that we are not hindering innovation, but fostering it as well. So from that perspective, I think the FTC Act is moving in the right direction with these types of hearings as well.

From an education standpoint, the FTC can also play a role in educating consumers to understand what is AI. Again, because it is a new technology, people hear about it. We talk about it all the time in Silicon Valley, but it may not be known to the rest of the country. So just educating people about what chatbots are, what it means when you are choosing Netflix on a Tuesday night and watching "Buffy the Vampire Slayer," what the impact might be. It might be that your ice cream prices might go up or it may be that your Netflix fee might go up if you are a more avid watcher than others.

So just understanding the impact of the data would be helpful to consumers and also encouraging companies to implement AI not just to exploit data, but to think about it holistically is really important and encouraging companies to do that from that framework of advancing society versus exploiting the data is something that FTC can take on, too.

MR. GILLULA: So I am actually going to answer the question in reverse order. In terms of consumer education, I think to accomplish that mission, the FTC needs a much more robust staff of technologists. They have only somewhat recently started having technologists on staff. I feel like the FTC should have as many technologists as lawyers at this point. And, obviously, that is not where we are.

I also realize that is not in the FTC's ability to change. So if you are a Congressman or a

Congresswoman sitting in the audience, this is my plea to you is increase the funding for technologists at the FTC because those technologists can help with explaining AI and what to expect in a consumer standpoint to consumers. They can also help explain it to the lawyers at the FTC when they are doing enforcement actions or they are doing investigations. They can help explain it to policymakers. So I think there is a real need for a really robust technical staff there.

In terms of whether or not the FTC Act sufficiently captures everything that we might worry about with regards to AI, I still worry a little bit about the fact that -- I mean, I guess there are two parts. One is whether -- I mean, at least -- and, again, you got the only nonlawyer I think on the panel talking. The FTC Act -- when you are talking about harms and unfair and deceptive, you are talking about what is the cost-benefit analysis. And I worry a little bit that when we are talking about privacy, in particular -- so, again, this comes back to rolling AI and big data and predictive analytics into the same thing.

But when you are talking about privacy, what may be good for society is not necessarily good for

CERTIFICATE OF REPORTER

I, Linda Metcalf, do hereby certify that the foregoing proceedings were digitally recorded by me and reduced to typewriting under my supervision; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were transcribed; that I am not a relative or employee of any attorney or counsel employed by the parties hereto, not financially or otherwise interested in the outcome in the action.

s/Linda Metcalf
LINDA METCALF, CER
Court Reporter