THE BROOKINGS INSTITUTION

FALK AUDITORIUM

LESSONS OF HISTORY, LAW, AND PUBLIC OPINION FOR AI DEVELOPMENT

Washington, D.C.

Thursday, December 12, 2019

PARTICIPANTS:

DARRELL WEST, Moderator
Vice President and Founding Director, Center for Technology Innovation
The Brookings Institution

ALEX ENGLER
David M. Rubenstein Fellow, Governance Studies
The Brookings Institution

MELISSA WHITNEY
Counsel
DLA Piper

BAOBAO ZHANG
Ph.D. Candidate
Yale University

* * * * *

P R O C E E D I N G S

MR. WEST: Good morning. I'm Darrel West, vice president of Governance Studies and director for the Center for Technology Innovation here at the Brookings Institution. And I would like to welcome you to this forum on artificial intelligence.

So AI is not the first technology to concern consumers. You can go back and look at events such as the automobile, television, and robots that have generated skepticism and sometimes outright fear. And those worries have forced both government and business to take action to address public concerns.

So we now face a similar situation in regard to AI. As algorithms are being deployed, people are worried about privacy, human safety, and a lack of transparency. A number of individuals worry that we are moving towards more inequality and higher rates of bias. So we're at the point where we need to think about how to deal with public concerns and what measures should be undertaken in order to protect consumers.

To help us understand these issues, Brookings has launched a new AI paper series. We published almost a dozen papers in recent weeks and there will be many more coming in the next couple of months. It covers topics such as facial recognition, products liability law applied to technology cases, the role of insurance in mitigating AI risks, and how to deal with AI biases in a variety of different sectors. Each of the papers identifies particular AI problems and outlines recommendations for how to deal with those particular issues.

Our goal is to try to come up with fresh ideas for moving forward and to raise awareness about AI issues. And if you're interested, those papers are available at brookings.edu.

So today, we have brought together three of our paper writers to discuss their particular ideas in regard to AI. So I'd like to introduce Melissa Whitney, who is a counsel at DLA Piper, a D.C.-based law firm. Previously, she was at the Federal Judiciary Center where she wrote about AI and technology in regard to the courts. In our paper series she's written a paper on ways to improve access to technical expertise for judges in AI kinds of litigation.

Baobao Zhang is a Ph.D. candidate in Political Science at Yale University and currently is at MIT on a post-doctoral fellowship. She studies public opinion on AI and wrote a report for us on the

lessons from public opinion for how we should think about regulating AI.

Alex Engler is the David Rubenstein fellow in Governance Studies at Brookings and previously was at the University of Chicago. He has written about deep fake videos and the need for greater AI transparency.

At each of your seats you will find an event evaluation form that looks like this in which at the end of this event we would ask you to fill them out and there will be people standing in the back of the room who can collect them. So we're interested in your feedback as we plan additional AI events in the future.

So I want to start with Melissa. So technology policy cases increasingly are ending up in the courts as people start litigating various types of issues. You have written about the need to provide greater technical expertise for judges in AI related cases. So I guess I have a two part question.

What is the nature of the expertise problem in the judiciary and how should we address that issue?

MS. WHITNEY: Okay. So the problem really arises because our traditional framework for handling cases really deals with human errors and human decision-making and human fault. Here, where judges don't understand a technology at issue in a litigation, it makes it incredibly difficult to determine who or what should be responsible when there is an AI-related error or accident. These can have implications because we're setting legal precedence here that arguably could stifle innovation going forward if we get those decisions wrong.

In addition, judges also play a crucial role as gatekeeper. They must be responsible for understanding whether a technology is sufficiently reliable that it can form a basis for making decisions, for making legal decisions. And there you see things, technologies, like facial recognition software or prediction of recidivism that can have real consequences if they're not reliable on defendants' freedom. So that's really the problem, is making sure we get these decisions right, and with judges that aren't necessarily familiar with these technologies.

There are a number of solutions. First, we can educate judges about the methods that they have available to them now. There are technology tutorials in the courts and I've written about that previously. There are technical advisors that judges can appoint to help them with specific issues in a

case they're facing immediately. And in some situations, and increasingly (inaudible) there's still infrequently used, there are court-appointed experts that can serve as expert witnesses and actually present in trial to judges and to juries about the technology at issue.

So first, judges need to become familiar with and comfortable with using these strategies.

Second, we need to encourage additional pilot programs and research into other strategies that may help judges understand these issues. There have been a number of proposals over the years including the idea of technical expert panels or judges, science judges. Basically, science and technology judges handling cases for which they have particular expertise.

To have these strategies, both the ones in existence and to promote the use of pilot programs, third, we need the buy-in from the legal community in general. Judges are understandably hesitant to try to introduce one of these strategies in the courtroom if the parties are not onboard or in some cases if the parties don't propose it themselves. It's part of the adversarial system. We consider sacrosanct a party's ability to decide exactly how they want to go about proving up their cases.

And then fourth and finally, I think most importantly for the group today, we need to really harness the expertise of the AI professionals that we have. They can be leading these educational strategies for judges. They are the ones that are going to service technical experts on panels and hopefully develop the reference guides that judges will use and will turn to and will cite very frequently as they face these new cases with AI-related technologies.

AI experts also really need training and communication to be able to translate these concepts to lay audiences. I think that's really critical and that will help not only judges but that's a scale that we need to inform policymakers and legislators as well because ideally we'll have some legal systematic frameworks to deal with the AI-related technologies that are now in cases, rather than having judges have to decide issues de novo and piecemeal across the country.

MR. WEST: So just a quick follow-up question. How much buy-in are we getting from the legal community on some of these reform proposals? And are judges actually starting to avail themselves of the various suggestions you have?

MS. WHITNEY: Judges are starting to avail themselves. Once people have familiarity once with holding a technology tutorial before a judge, they're more willing to recommend it.

Unfortunately, there's not really systematic training or introductions to these strategies, and so attorneys are inherently risk averse suggesting strategies for their clients. And I think until we have more mainstream recognition of these strategies they haven't really picked up the pace like they need to.

MR. WEST: Okay. Thank you.

So Baobao, you have a new paper on the Brookings website on the Public Opinion Lessons for AI Regulations based on a very detailed set of analysis of public opinion data that she has undertaken. So what are the lessons that you have found based on the survey data?

MS. ZHANG: That's a great question.

So as we move away from considering ethical AI in terms of abstract principles, many governments, whether that's national governments or state and local governments, are considering regulation of artificial intelligence. So this is where I think mass politics and public opinion come into play because a lot of these issues might become important to voters. An overwhelming majority of Americans believe that artificial intelligence is a technology that should be carefully managed.

In my survey work with my team based in Oxford, we found that Americans tend to prioritize issues related to cybersecurity, privacy, and digital manipulation, which are all big things that have recently come up in the news.

In the new report that I've written for Brookings, there's three key takeaways. First of all, national level public opinion could sometimes belie opposition by particular subgroups. So in the case of facial recognition software, a majority of the American public actually support law enforcement using facial recognition software, but while 56 percent of the American public supports this use of technology by the police, subgroups, especially those who are marginalized, including African-Americans, they have a much greater opposition to law enforcement using facial recognition software. So what you see is in cities or states where there is more opposition, you've seen moratoriums or bans on police use of facial recognition. And this is where you might see more movement in the future, more at the local or state level where there's organized interest among the public or with civil society groups that oppose some sort of use of the technology.

The second point is that as AI move into the mass politics space, you might see partisan divisions. Right now there's a tech lash in the U.S. It may appear bipartisan on the surface, but actually,

there's a difference between what conservative legislatures and what liberal legislatures prioritize, especially in the regulation of algorithms used by social media companies.

On the left you see much more emphasis on data transparency, privacy, preventing digital manipulation. And on the right the priorities focused on this sort of opposition to this alleged conservative bias on social media. And these partisan divisions could prevent potential federal regulation.

The third lesson is besides the public there are other important interest groups that could push for regulation. In the case that I look at which is a ban on lethal autonomous weapons, the machine learning community has been really active in pushing for this ban. A lot of tech companies, heads of tech companies, as well as machine learning researchers have signed on to this call for a ban that was organized by the Future of Life Institute, and this is where you might see potential again for movement in self-governance, for instance. Engineers at Google have successfully protested against their parent company to pull out of Project Maven. This is a very fast-moving field. As we move towards regulation of the technology, I think a lot of the conflicts in the mass politics space will come into play.

MR. WEST: Okay. Thank you very much.

So Alex, you have a new paper on the problems of deep fake videos. So why don't you start just by telling us about deep fake videos, what are they and why we should be worried about them?

MR. ENGLER: Sure. Thanks, Darrell.

So deep fakes are getting a lot of attention right now, and I think most of us have probably been exposed to them. But think highly realistic videos of real people doing what seem to be real things but are in fact fake and made with modern AI.

Okay. There are examples out there of Mark Zuckerberg and Barack Obama and Donald Trump that are kind of funny. You might want to watch those. They're entertaining. But there are some less enjoyable ones that are possibly coming out.

So I'm going to tell you the story of Ali Bongo in Gabon. So Gabon is a country on the west coast of Central Africa, and Ali Bongo is its dictator president. Last year he got sick and was out of the public eye for weeks. Very few press releases, very few statements. And at some point some reporters were saying he may or may not be dead. He's not a super popular president dictator, and

various people started to claim that Ali Bongo was dead and it was time for a democratic revolution in Gabon.

Okay. Then comes out a video from the government, from the Ali Bongo government showing -- purporting to show Ali Bongo sitting at a desk and giving a speech, giving one of his annual speeches. But the video is a big strange. He's sort of making odd facial expressions. His voice sounds slurred and abnormal. He's not very animated. Maybe it's because he's sick, maybe not. But he's not blinking enough. He's blinking half as much as he should be. Less than half as much as a normal person blinks. He doesn't have his normal affect. The camera is cutting back and forth between two shots, maybe suggesting that it's been tampered with, the video. And that's what the opposition party suggested. This is a deep fake. This is a fake video put out by the government to reassure people that he's alive when, in fact, he may actually be dead. It actually leads not only to an opposition, but to a small coup, a failed coup in the country by military leaders saying that Ali Bongo was dead and it was time for a new government.

At the time -- this, by the way, is last year. This is 2018, for a little bit of a timeline. The underlying technology was invented in 2014. The first deep fake is 2017. 2018 we have a deep fake in a fairly influential situation. We still don't know if it was real or not. So, and that's maybe not so much because it's an unsolvable problem but it's sort of an incentives problem. And I'll explain what I mean by that.

Right now there's a lot of money going into automated deep fake detection. This is the solution that most people in the field are working on. Can you detect a deep fake by the sort of pixels that result from the video, right, using algorithms to detect that it's been tampered with? But that's a probabilistic approach. So you get a percentage of probability. This is 85 percent likely to be fake.

Can you think of a lot of tech companies that are going to want to look at the video of a president who may or may not be dead and say, oh, don't worry, that's a fake video? Our model said it's 80 percent sure. Probably not.

So we have this big incentives problem. The people working on and building the technology that could tell us more certainly if these things are real aren't going to want to do it. So we might want to be thinking about how to arm third parties who are in a better position to make these

judgments with those tools and better tools. We might want to give them the most modern deep fake detection algorithms and maybe not actually spread them publicly so they can be used for ill. We might want to make sure we're training them on it. We might want to give them other tools, like reverse video search. Putting a video into a search engine and finding videos that are highly similar. Maybe the origin of a video that was then used for deep fake. And also, working towards finding groups, maybe government actors, more likely third-party actors in the near future who really feel like it's their role to make some definitive decisions around these deep fakes because right now it's not really clear who that's going to be and we can expect these videos to get very, very, very realistic in the immediate future.

MR. WEST: Okay. Thank you.

So Melissa, in your paper you write that AI is ushering in a new era of quantitative legal decision forecasting. So can you tell us what that is and what that will mean for potential litigants?

MS. WHITNEY: Certainly.

So legal decision forecasting has been around a long time and people have always wanted to try to find variables to predict what the outcome of cases will be. And the classic example there is trying to predict Supreme Court justice votes based on whether a republican or democratic president nominated that justice. Historically, that scholarly work and prediction hasn't fared much better than flipping a coin to predict the outcome of a case, but AI has really revolutionized that space.

So in 2017, some researchers out of Illinois came up with an AI method that they used to predict with 70 percent accuracy all of the Supreme Court case decisions from 1816 through the end of 2015. That's striking. It's incredibly impressive because that legal landscape has changed a lot over 200 years. These AI prediction methods for legal decision forecasting now are hitting a level of reliability where litigants may want to use them to consider whether it's worth bringing certain types of claims or not. Attorneys that take cases on contingency fee arrangements where they do not get paid unless their client is successful may reconsider using these AI-driven methods. And very interesting over the past decade, these AI-driven legal forecasting models have spawned and really funded an entire -- fueled an entire industry and that's third-party litigation funding where essentially hedge funds are deciding whether to invest or to place a cash advance on a litigation in which they are not a party and otherwise wouldn't have involvement in return for a stake in a potential jury award or a global settlement agreement. And those

stakes may be as high as 10 to 30 percent returns. So that's entirely revolutionized the way that's done. And these AI-driven legal decision forecasting methods, they reduce the amount of time it takes to evaluate whether an investment is worthwhile, and it takes some of the risk out as these algorithms get better in terms of accuracy.

MR. WEST: So it sounds like it could have a major impact on the way litigation develops in the future.

So Baobao, you note that there are interesting subgroup differences in attitudes towards AI based on gender, race, and age. And in your opening remarks you mentioned the case of facial recognition software which for the public at-large seems to have some support when used by law enforcement, but that African-Americans in particular are opposed to that.

So can you tell us a little bit more about these differences and what they could mean for future regulation of AI?

MS. ZHANG: Sure. That's a great question.

So I my team's research, we found that while a majority of Americans support the development of AI and related technologies, when you break it down into demographic subgroups you see clear differences. Respondents who have lower levels of education, who are low income, who are female and who have less science or technology training in their background and their education tend to be less supportive of the development of AI. And there are several reasons that we hypothesize. One of them is that there's a lot of conversation about how AI might impact society, whether that's the future of work or in policing, and these more marginalized groups could face disparate impacts that they might be disadvantaged, whether that's in the labor market or that they will be policed more harshly.

And in terms of regulation of AI, as I pointed out, there's heterogeneity, geographic heterogeneity in terms of support for using the software, or opposition to using the software. So in my state of Massachusetts, we're actively considering a moratorium on facial recognition software. And I think that's in part because 79 percent of Massachusetts voters oppose the police using facial recognition software.

MR. WEST: Okay. Actually, if you can hold off just for a minute.

But I have a follow-up question for you just on the gender gap because you mentioned,

because you said that women tend to be a little less supportive of AI applications than men.

Do you have a sense of why that is the case?

MS. ZHANG: That's a very good question. This is a topic of active investigation by my team.

Women tend to be less supportive of technology across the board, whether that's in the use of GMOs or vaccines. So it could be technophobia that just maps onto AI. But there have been other research by labor economists that suggest that women are going to be more negatively impacted by automation than men. So that could play a role in this difference of opinion.

MR. WEST: Okay. You had a quick point of information you wanted to ask?

SPEAKER: Yes. There is a problem with the facial recognition technology as it applies to African-Americans because our faces create light shadows that make it difficult for the technology to pick up the image correctly. That is basically the source of the widespread opposition to -- excuse me. I have a mint in my mouth -- the widespread opposition to the applications of AI.

I was at a conference last week where a research project which had been deemed a failure was described. I didn't hear the reasons -- I didn't come early enough to hear the reasons why they deemed it a failure, what the design problems were. But the research generated an outcome where six football players were identified -- are you familiar with this?

MS. ZHANG: Yes.

SPEAKER: Six football players were identified as criminals. And I asked a question, how many of those football players were black? And the answer was, guess, all of them.

So that suggests a real problem with the technology. And I just wanted to make sure it's important for people to have that kind of background because these people are out there making policies and they need to know why there are problems.

MR. WEST: Okay. Thank you for that clarification.

Did you have a quick response?

MS. ZHANG: Yes. So this is a really prevalent problem right now in facial recognition software. So a recent study by researchers at MIT found that the leading facial recognition software by tech companies -- so these are commercial software -- misidentified women of color 33 percent of the

time or about a third of time, where the inaccuracy was about one percent for white men. And that's a big problem if you're going to apply it in these high stake situations such as in law enforcement.

MR. ENGLER: Okay. Just a really quick plug.

Joy Buolamwini is one of the researchers most focused on this. She just gave a really incredible speech at the Stanford Human-Centered AI Conference. If you're interested in this topic, you should watch that. It's absolutely phenomenal and really covers the state of the field right now very well.

MR. WEST: Okay. Thanks for that.

So Alex, when you were talking about deep fakes, you suggested several possible courses of action. So automating deep fake detection, reverse video searchers, and there are other possibilities as well. How likely are these to actually deal with the issue, or are there other more fundamental things we should be thinking about?

MR. ENGLER: It's not great. So we have not solved, you know, photo shopped images. Right? You can make a photo shopped image with enough effort that it is good enough to be indistinguishable from a real photo. That is currently possible.

The pace at which the deep learning technologies are advancing is very, very fast. And this is a little confusing if you look at -- there's like a lessons from history in the title. So one lesson from history. You can hear a lot of things that seem to conflict about the history of AI. Some people will say, well, you can really trace it back to the 1940s and '50s and '60s. That's true. But then there is the AI winter at some point and we stopped working on it for a while, and that's true. And well, this modern stuff is all these rehashed older methods and that's also sort of true. Right? It's sort of confusing to parse out. But the key for the last decade is we're in a fundamentally new stage of AI history. And the pace of the increasing -- the quality of the technologies being used for narrow tasks, not for being a human, not for being generally an intelligent -- not for general intelligence but for narrow AI tasks, the pace at which they're getting better is really, really dramatic.

So there is a lot of speculation by people, by the researchers in this field that deep fakes are going to get to a point in the relatively immediate future where they will be indistinguishable from videos otherwise. We might need to think about how we evaluate and how our citizens are prepared to evaluate general information they encounter. And it's not just deep fakes. We could also talk about -- so

I'm reading right now a little bit of about AI transparency in chatbots as well. Right now, half of all Americans don't think that they could identify whether a chatbot was human or not. And that's probably doable but probably but pretty hard. It might take you a little bit to figure it out but they would start to falter if you kept talking to one. In the not so distant future that is going to be impossible. Or maybe if not impossible, it's going to take too long to really be something anyone has time to do. Every single time you interact with an online tool you're like, okay, I'll spend a half hour running a personal touring test to see if this is a real thing or not. Like, it's maybe not literally impossible but it's logistically impossible. We need to start really thinking about how we require disclosure from some of these systems and in that circumstance as well.

MR. WEST: And just on a quick footnote, too, what Alex said. So in the current UK election, you know, all of us are used to thinking about the problem of Russian interference in various elections. What they are finding is the Russians have basically taught everybody else how to do it. And in Great Britain, there are domestic actors that are basically engaged in exactly the same behavior. Deep fake videos, disinformation, social medial dissemination of false narratives and so on. So it's something that as we head into our own elections over the next year we should be worried about.

MR. ENGLER: A quick note on that. It does also mean that domestic legislation can matter. Right? Some people will say, well, it doesn't matter if you pass AI transparency laws because Russia and actors and subversive international foreign influence campaigns don't care. But they are domestic actors, too. Some of them are companies trying to sell you stuff with a chatbot on Instagram, and others are domestic political actors, too. So it is still worth thinking about legal requirements around AI transparency.

MR. WEST: So one more question for the panel. Then we're going to open the floor to any questions or comments from you.

This I'm going to just throw out for all of you, and you can answer as you would like.

What responsibilities should companies have for some of the AI problems that we've been talking about? And are there ways that AI can be part of the solution?

Does anybody want to jump in on that?

MS. ZHANG: I can talk about the first aspect of the question.

I think tech companies should assume more responsibility before they deploy their products. I collaborate with researchers at Open AI, a tech company based in San Francisco. And earlier this year they were going to put out a tax generation software, GDPII, which you can type in a prompt and it will generate text. Obviously, there could be a lot of malicious use for this software. You can more easily generate fake news articles. So before they had released this product they thought about it and said we should do a limited release. So what they've done is they've worked with academics, with other third-party nonprofit organizations, with researchers, to test out the safety of this product to anticipate what are some potential malicious uses. And then they've been gradually launching this product by upping the volume of the training model so that it gets more sophisticated over time. And I think this more careful approach in launching a product is something that other tech companies should look to.

MS. WHITNEY: From a legal perspective, your company's value has always been in the proprietary elements of their software. And now increasingly, the value is placed on the data that informs you of these algorithms. So there is kind of a tension there in whether to protect and whether kind of to preserve the proprietary nature of that data or whether to open things up to collaborate with other companies and to provide a judge when you have these technologies eventually come before a court, to provide transparency, to kind of shine a light on what is now a black box. There's a tension there. There's always been a tension there between educating the public and between protecting your intellectual property rights.

I think now where the stakes are so high and we have these advancements that Allison has talked about where public literacy and judicial literacy has not kept pace. I think companies are beginning to recognize that we have to find a way to agree to educate judges about what we have before we engage in the advocacy in front of a court. And so I think you will see parties, you will see attorneys on both sides agree to hold technology tutorials and to agree to at least fundamental principles that a judge needs to understand to be able to decide a case effectively. And I think the companies have a responsibility to consider these approaches, to using these strategies in courts, and to sharing enough that we make sure decisions are made in a sound way and not just in a way where our side has the best outcome today but faces real consequences with legal precedents set tomorrow.

MR. WEST: Alex, your thoughts on corporate responsibility?

MR. ENGLER: Let's see. The loss side (phonetic) of this is really interesting, too.

So there's some discussion over putting more onus on digital platforms for the content that ends up there. Currently, they're protected from this from a legal perspective by section 230 of the Communications Decency Act. And that's like a really touchy subject to get into. That being said, a really important question if you were to roll back that protection and hold digital platforms to a higher standard is could you do it in a universal way that made any sense? The answer is probably no, which means the courts are going to have to make a whole bunch of decisions. Just one example, if you said to Twitter you have to start labeling automated accounts, right, if you have a strong reason to believe something is automated, you need to tell us or you need to make it public.

What exactly is the fair and reasonable standard for them to do that is a pretty hard question. They wouldn't be able to get all the bots but there is some research that says they could label more than they're currently, you know, sometimes they're taking them down. Taking them down or labeling. And that would certainly end up in the courts. And so certainly, we may be approaching a time where we're going to hold them to a higher standard. It can't be a set universal standard because universal standards don't make sense for every company in every circumstance and it's going to depend on the courts to work out some of it.

MR. WEST: And there's an interesting angle. You mentioned this issue of the current legal protections for digital platforms and just the difficulty of holding them accountable for what takes place on their platform. In two new treaties, basically, the U.S. Government has agreed to extend those current legal protections and, one, part of the NAFTA Agreement. The current section 230 protections are part of that treaty. And then also the new treaty with Japan also includes that. There's some sentiment in Congress to start to revisit that. Congress already has carved out an exemption to those protections in the case of human trafficking and there are some legislators who would like to extend that to some other areas as well.

Why don't we open the floor to questions and comments from you? So if you have a question, raise your hand.

Right here in the front row is a gentleman with a question. And if you can just give us

your name and organization.

MR. CHECCO: Larry Checco.

MR. WEST: I know this is an emotional topic for you.

MR. CHECCO: Yes, it is. I get all chocked up.

I recently had the opportunity to ask a former FCC chairman what the difference between 4G and 5G is. You know what can be better than 4G? And he said that, and I hope I'm getting this right because I'm not a technocrat, he said that 4G is basically pushing data between two points. Like in a cellphone kind of situation. I'm calling you, you're call me. It's pushback and it happens fast. With 5G, he said the difference is that it's technology that's being coordinated out there. It's no longer point to point. The information is being coordinated. Think of a non-driver car. That's all -- and the thing that scares me because now at the beginning of this you mentioned a couple of technologies and maybe we shouldn't worry because, but those technologies were like a car, point to point. The problem with AI is it's so pervasive and it's just -- and we talk about Congress. It was embarrassing to watch Congress ask, what's his name? Mark Zuckerberg questions. I mean, it sounded like they were coming from the Jurassic period. They can't keep up with this stuff.

So I guess my question is, where are we going? And the fake stuff drives me crazy. I mean, you know, fact and fiction is just so blurred right now. I mean, what kind of world are we entering into?

MR. WEST: A risky world.

Our panelists, any reactions to that?

MR. ENGLER: I'll take the Congress question side of it real quick.

MS. WHITNEY: Address the fact from fiction.

MR. ENGLER: Yeah. Just I think that there is a growing recognition that the technical capacity of Congress is not currently where it needs to be. I will give a lot of credit to the organizations working on this, the Tech Congress Fellowship is fantastic, as is the American Academy for the Advancement of Science has a science and technology fellowship that places people with really high levels of technological expertise in the halls of Congress. That is the majority of the technical capacity there but there is a conversation about bringing back something like the Office of Technology

Assessment to inform Congress and also the GAO is doing some great work with a new group called STA, which is their science and technology advisory group, essentially.

So there is a recognition of that and hopefully maybe some signs that something is going to change. I know that was only one part of your question.

MR. CHECCO: What scares me is the anti-science environment (inaudible). That's a little bit -- talk about being scary.

MR. WEST: And one of the ironies is the Office of Technology Assessment was killed by Newt Gingrich, 1995, just as the Internet was coming into being. So talk about bad timing from the standpoint of going out of the technology business and offering technology advice when we actually start to innovate.

Baobao, you had a comment, too?

MS. ZHANG: Yes. I have a comment about our current fears around technology and with deep fakes and fake news. I have a somewhat critical perspective. My colleagues who work in academic research, they've tracked Internet users and found that leading up to the 2016 election, exposure to fake news wasn't as prevalent as people who are scared about fake news seems to predict. It's, I think, less than 10 percent of Facebook users were exposed to fake news and it tends to be a small segment of the population who are consuming most of the fake news. So while it is something that's coming online and might become more prevalent, we should think about it in terms of the entire media landscape.

I think what's driving our fears about fake news and about deep fakes, et cetera, is this fundamental rise in the level of distrust in society. The public doesn't trust the traditional news media anymore. The public doesn't trust government officials. And I think this is not a problem that just can be solved by more technology when it's a more deeply rooted problem, especially in American society where we have pretty extreme polarization and high levels of distrust of institutions.

MR. WEST: Okay. Other questions?

Right here is a gentleman with a question. Yeah. Right there.

MR. COLEMAN: Richard Coleman. I retired from what is now called Customs and Border Protection. At one time it was Bureau of Customs and then it became the Customs Service. And

I was hired to answer congressional complaints, among which were many involving people getting stuck because they had a common name and our information systems couldn't discriminate between them on the basis of names. So from my perspective, one of the advantages of facial recognition would be to assure the government who you're not, not falsely accuse you of who you are. So there's a benefit to that.

I know that CBP just announced that they're not going to go -- they're not going to keep facial recognition at least at this time. Is there any kind of idea of what the -- an accuracy of an acceptable level of accuracy would be so that people would be reassured? I know for white men 99 percent sounds good. If you're not a white man, 67 percent doesn't sound too good. Is there a target that's been discussed in the industry as far as you know?

MS. WHITNEY: So I know that we're revising some of these older technologies and forensic techniques, and people are evaluating what is the validity here? What is the reliability of these methods? It's important before we use them that we make sure that we are not engaging in technologies that because we come up with a number at the end of the day, they look like they're a reliable basis. And there are so many examples in the past where people have come into a court and before a judge said, you know, I'm 99 percent confident that this is -- that I have correctly identified the defendant as the perpetrator. So I understand the need to avoid that.

I know that there are really responsible groups, and even the Deparmtent of Justice Office of Legal Policy that are evaluating these technologies and are trying to come up with a way to characterize the accuracy and reliability that it can apply to all citizens, all humans. And we'll see. I think that's a huge, broader societal question but I think first we need to just be able to characterize that correctly so that when it goes before a court and we try to use it for the first time, judges can evaluate its reliability generally and we don't end up with these errors like we've had in the past with other technologies where there are a lot of people who were falsely convicted using what we thought were really valid and scientific techniques that prove not to be.

MR. ENGLER: I'll that there are some standards, too. NIST, the National Institute of Science and Technology has standards around facial recognition, and by all accounts they're getting better for everyone. For white men and for women and for people of color, and it's even possible that

they could all hit that very, very high range. What that still won't tell you is how effective they are when they're applied to grainy, CCTV cameras sitting outside of a supermarket or something. Right? And so, in low light at night with a person's face partially covered.

So even though there are standards and we are improving on those standards, I think there is certainly an insufficient amount of knowledge known about how, when we apply these broadly again, you know, police body cameras is another example, where you can see the quality of the algorithmic side of it not necessarily being the only factor in them being a trustworthy technology to be used in a courtroom.

MS. ZHANG: Can I just briefly add to that? While I think increasing accuracy is one target, on the other -- at the co-consideration is that what are we using these technology for? There may be other values that we want to emphasize. For instance, privacy. So some police departments have partnered with the Amazon Ring service, their home security camera systems, to track people who walk by your house. For some voters that's unacceptable. Why should my walking by my neighbor's house, why should the police have access to this information? So I think beside the point about increasing accuracy, there are some other values we might want to consider when passing legislation.

MR. WEST: And on that particular issue there's a question of data retention. Like, with all these cameras and with facial recognition being deployed to use them, how long should the information be stored? Like, you know, there are some cases where we might think there's a legitimate interest in collecting people's images, for example, in the case of an imminent threat, airport security and so on, but once the threat is past, which could be the next day, do we still need to retain those images?

Right here there's a question.

MR. SAKOCHI: Yeah, hi. Damian Sakochi. I'm with Georgetown's School of Business.

First, thanks, Brookings, for focusing so many great papers on this important topic, and I really appreciate it.

The thing that brought me today was really the history that you were bring up, and I was wondering if any of the panelists could talk about the importance of transparency or maybe look to historical examples because AI is complicated and somewhat opaque. And in all of your areas it seems to me that trust is a key thing for bringing these communities. And (inaudible) are looking for good

examples of where in previous historical, technical initiatives that there was transparency. We were able to get there. The genomics issue, I think we didn't quite do it right. Maybe we have now. But maybe you could just point me towards some good examples you've come across. Thank you.

MR. WEST: Any suggestions for our panelists?

MS. ZHANG: Recently I wrote a paper that has a literature review on the trust issue with regard to other technological applications, and GMO was one of the things I focused on. And according to past survey research, it seems that distrust of scientists, distrust of agricultural companies tend to be correlated with opposition to GMO. And it seems the opposition is much stronger in Europe. That could explain why they have tougher legislation against GMO foods. So that's an example.

MR. ENGLER: So I can offer you a vision of what would be nice to have for AI specifically. The example I think about a lot is the news recently about the optimum algorithm which was helping hospital systems decide how much care patients were getting. Their paper came out this year. I think their discovery was sometime last year. Essentially, they found that there is a large amount of bias against African-Americans to how much care they were receiving based on the algorithm. This was being used by Optimum, which is owned by United, so it was like a lot of -- in the tens of millions of people who were receiving a degree of care based on, at least in part based on the system, and systems like it were being used much more broadly. It could have been over $100 million people. It's hard to tell. But it was a lot.

A really important question that we could be asking is what transparency, what either, you know, information given directly to the person being affected by the algorithm or what public transparency that is required of public disclosure, or what disclosure past regulators could have prevented that kind of thing. I think you could look at Optum and say, oh, well, now that they got, you know, dinged in the news for this biased algorithm, everyone else will clearly fix their algorithms and then there will be no more algorithm bias problems in health, you know, in provisioning health care.

That's probably pretty unlikely. And I think we should be thinking about what types of -- maybe a typology looking at the use case of the algorithm and the type of underlying methodologies paired with some sort of disclosure that we think is sufficient. And that's probably a really important and complicated question across many different fields. Think health insurance or maybe also hiring,

mortgage applications, right? The number of things they touch is very broad, but some combination of this transparency fits well with this highly influential use case is something I think we should be working towards.

MR. WEST: And on the history part of your question I would recommend my colleague, Tom Wheeler's book. It's entitled from Gutenberg to Google. So literally history over several centuries. And the short version is he says almost every new technology, certainly any of the broad-based technologies inspired a lot of fear at the time, and the way that countries deal with it is through policy, legal, and regulatory means. The automobile came in. We started to have accidents. People started to die, so people got interested in highway safety and started to impose requirements. So he argues that history offers a lot of lessons, and basically, we need to kind of reconsider our rather laissez-faire attitude towards this sector and basically kind of think about what smart regulations would mean there.

SPEAKER: (Inaudible) has a great book, too, The Technology Trap.

MR. WEST: Yeah, that's right.

This gentleman -- or you had a question. Yeah.

SPEAKER: Thank you very much for your talk. It was great.

I'm a physician. I'm practicing oncology in Southern California and --

MR. WEST: You came a long way for this forum.

SPEAKER: Yes, I did.

So I have a question, one for patient safety. Now the hospitals are using EMRS and there's a lot of push to use radiology using AI and deep learning for radiology. What are the -- what's the legislation that we have now in place, if any? And what are things that hospitals can do or physicians can do to protect patient security and information in the future? And I'll open it up to any of the speakers.

MR. ENGLER: Do you have it?

MS. WHITNEY: In terms of protecting patient security and patient records, there is such a huge conflict there because in theory, the more that we connect sources of records for a single patient, the better the care we can provide. The more data we have from different sources, and even private companies that used to hoard data now are recognizing the need to share because of how powerful it can be in delivering care. So there's a conflict there as we connect these databases. We're going to be able

to identify individuals and there are real risks there will be discrimination or that it will have too much confidence and biased data and make bad calls there. At the same time there's so much potential to improve human care. That's a hard balance and I don't know really how we're going to make that. But I know there are good people that are raising these issues and these questions and trying to come up with a framework for what we're going to do there. I think, you know, to Alex's point, transparency to let people know this is happening will help a lot.

MR. WEST: I mean, I'd say one of the better examples in this area is the National Cancer Institute. There are a lot of medical databases that are now starting to be integrated and so it raises the question of patient privacy. And what they are doing to deal with that is to basically create secure databases that researchers can query on a very restricted basis. So it's not a Cambridge analytical situation where you're just simply dumping the data out and people can use it however they want and abuse it for other types of purposes, like you are limited in how you can use the information. You don't actually get the data. You can query it and get answers to your questions based on your research needs. So that's one platform-based way to try and reconcile, encouraging innovation on the one hand to try and get the benefits of AI and integrated databases while still protecting patient privacy.

In the back there's a gentleman with a question. There's a microphone coming up right behind you.

MR. KROPF: Thank you. John Kropf with Northrop Grumman.

It will be a question directed at Ms. Whitney, which is you started your talk mentioning case law, and I wonder if there are any good decisions out there that you can point to or examples where you think judges got it right in terms of examining AI and whether or not they're citing subject matter experts in the field that you think are at least a step in the right direction. Thank you.

MS. WHITNEY: There is such a range of background experiences for judges in AI. I will try to think of a good case for you, a recent case where these ideal principles were followed and the best experts reached an outcome.

I will say because there is such a concern or hesitancy because these are high stakes, the early decisions for the legal precedence and there are high stakes there. A lot of companies have actually chosen to settle. So you'll see the autonomous vehicle would-be cases. You know, when they

have hit the court, they've ultimately resulted in settlement early on because these are tricky issues, and I think companies in part are hoping for a framework to handle them uniformly as compared to piecemeal in the courts. But I'll think about it and try to come up with a great case for you.

MR. WEST: Over here along the wall there was a question.

MS. PASKA: Hi, Cindy Paska from the Council of Scientific Society Presidents.

This sort of follows on Mr. Kropf's question a bit. When we talk about the question of decisions being made that are right and then we talk about what society looks at as far as the decisions are made and the legal aspects, where do we bring the ethical issues in and what kind of focus is being put on the ethical issues when you're looking at the AI development and the decision-making process?

MR. WEST: I mean, I think this is part of the tech lash that has been developing which Baobao mentioned that if you look over time, Americans have always been pretty pro-technology in recent years. Because of all the examples that we know people are starting to reassess that. It's not that they've turned against technology but they clearly have a lot of concerns related to privacy, human safety, transparency, explainability, and so on. So what a lot of people are trying to do now on the ethical aspects is to, one, think about what are the ethical principles that we need to safeguard as AI gets deployed, and there are a number of universities, companies, and NGOs that have developed a set of ethical guidelines. There are typically six to eight common ones -- fairness, protecting its bias, maintaining human safety, explainability, transparency, and so on. The difficult part is how to implement and operationalize those types of things. The OECD developed its set of guidelines. They have now developed, I think it's called the AI Observatory, which is an organization committed to sharing best practices across countries. So they're trying to figure out how to implement these ethical principles in AI and then share that information with organizations across national borders.

So I think those are the types of things we need to be thinking about. We certainly have to address the ethical aspects because, you know, some issues are very specific to AI but there are broader considerations, like human agency. Like, are we still going to be in control of this technology or is it at some point going to advance beyond our capability to control?

SPEAKER: There's a bipartisan issue that plays into that as well and who defines what is right and what is ethical (inaudible).

MR. WEST: Yeah, absolutely. And we need to have that conversation right now so that people are comfortable.

I think we have time for one more question. There's a gentleman in the very back who has a question.

MR. COUCHETTE: Yes, my name is Roger Couchette. I work with private equity in the technology sector.

The question I have sort of turns the clock back a little bit, a little piece of history as well as a question I think that's still relevant today. In the mid-'90s when I was heading up IBM policy for Internet and related stuff, we hosted the chairman and most of the members of the Federal Trade Commission at IBM Research for a day of private briefings on what is this new thing, the Internet? How does it work and what are the issues that it raises? And one of the topics we discussed is relevant to the artificial intelligence, which is privacy. And at that time we set forth a framework. And by the way, before anyone overreacts to this little bit of history, please note that I recently wrote an op-ed for The Hill identifying the four major blunders that we made in the mid-'90s about what the Internet would look like and what the issues it would raise. So if you want to know what we did wrong, check out that article.

But what we told them, and I think what precipitated the framework that lasted for a decade was, listen, privacy is simply too subjective, and I remember clearly the example we used. When Mrs. McGillicuddy goes into her butcher and he says, do you want the same porkchops you always get, Mrs. McGillicuddy, she's happy and pleased. He remembers that I like pork chops, and it was a positive reaction. If the butcher tells some door-to-door vendor that Mrs. McGillicuddy buys pork chops and he tries to sell her things for her pork chop preparation, she might be mildly offended that the butcher told somebody else what her personal preferences were. If the butcher takes out an ad in the newspaper and says Mrs. McGillicuddy buys pork chops every week, she would probably be very offended, tell the butcher off, or never go back to that butcher again.

The point of this example which I think was persuasive to the Federal Trade Commission in the mid-'90s was that the subject matter is simply too subjective to be appropriate for flat out government regulation. That there's a continuum between, oh, isn't that nice, they remember me, all the way to what right do they have to tell anybody what I'm doing and where you are in that continuum

depends upon the person, the day of the day, the day of the week, the season, and a million other factors. And therefore, are not appropriate for flag government regulation which can't take into account subjective differences. This was privacy on the Internet but it set a framework for how to look at it which I think reoccurs today in artificial intelligence. And in fact, the two are deeply intertwined. So I guess I'd be interested to see how you react to the philosophical argument that prevailed in the mid-'90s that the subject matter is simply too subjective to be appropriate to a black and white kind of distinctions that are necessary for law. Thank you.

MR. WEST: Okay. That's a great closing comment.

Any reactions from our panelists?

MS. WHITNEY: I mean, I think that our personal information has become such a commodity today. We regulate every industry that trades in that. And so I think we have to find a way to come with some -- these are going to be difficult conversations as Darrell mentioned, but we have to come up with a way to regulate that market and our personal information as well.

MR. ENGLER: Yeah, I'm not a privacy expert really but you actually I feel like in the analogy sort of touched on standards that kind of make sense; right? That if you give someone information willingly, they're allowed to use it but then there could be restrictions on the use if you start passing it off to third parties which is one of the big parties we're having now. It's third-party data brokers who don't have an individual responsibility to the person that they got the data from doing things that we may not want them to be doing. I do think there's space in there for sensible restrictions on how data is used, especially once you start passing it off into companies that you have no direct relationship with whatsoever. And we've certainly seen some bad actors in that space.

MS. ZHANG: That's a very good point. I think the privacy issue is one of many issues that could lead to potential conflict in this space. So there are currently at least 84 sets of AI ethics principles out there. And there are a lot of tensions between some of these principles.

Speaking of privacy, you might want to increase the accuracy of an algorithm, but that might mean collecting more personal data. So these things are in conflict. And how we resolve these conflicts is something that I'm hoping that we can decide democratically.

MR. WEST: I would support that principle right there. That's a great way to close.

Hopefully, we can do that.

Please fill out your event evaluation form. There will be people at the back collecting this, and I want to thank Melissa, Baobao, and Alex for sharing your thoughts. And thank you very much for coming.

(Applause)

* * * * *

CERTIFICATE OF NOTARY PUBLIC

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

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2020