THE BROOKINGS INSTITUTION

WEBINAR

# HOW TECHNOLOGY CAN REDUCE GOVERNMENT FRAUD

Washington, D.C.

Wednesday, September 15, 2021

PARTICIPANTS:

**Welcome:**

DARRELL WEST
Vice President and Director, Governance Studies
Brookings

**Panelists:**

MELISSA KOIDE
CEO and Director, FinRegLab

BETH SIMONE NOVECK
Director, The Governance Lab
Chief Innovation Officer, State of New Jersey
Professor, Northeastern University

* * * * *

P R O C E E D I N G S

MR. WEST: Good afternoon. I'm Darrell West, vice president of Governance Studies at the Brookings Institution. And I am pleased to welcome you to our webinar on "How technology can reduce government fraud."

We live in a time of rapid change. AI is being deployed in many different sectors, ranging from healthcare and education, to transportation and e-commerce. Organizations are using algorithms to screen resumes, assign students to schools, and develop new vaccines. But one of the areas where AI and machine learning are being deployed is in the public sector in the financial services area. We're seeing the use of technology to improve government effectiveness and to try to reduce fraud.

I actually have a new Brookings paper out this week entitled, "Using AI and machine learning to reduce government fraud." The paper looks at traditional fraud approaches, and then how AI and machine learning are starting to be used to ferret out various cases of fraud and abuse. So, we look at some of the obstacles to public sector innovation. The paper makes a number of recommendations. I'll just summarize a few quick ones, but then refer you to the paper, which you can read at brookings.edu.

So, one of the recommendations for moving forward is for organizations to be proactive about developing responsible AI. And this involves hiring emphasis, creating internal review boards, and developing mitigation strategies early in product design and deployment before the products actually are put out into the marketplace. We also are encouraging organizations to use evidence-based evaluation to help determine the efficacy of new algorithms and to see how they actually are affecting decision making. It's very important as government agencies start to deploy these new techniques to think about workforce development and giving their employees new skills to use these types of techniques. So, it's important to develop partnerships with higher education, community colleges, technical institutes, online providers, and firms offering personalized learning, not just so people have the skills so that they understand what these applications are doing, what the opportunities are, but also what the risks are. We think organizations should embrace lifelong learning and expand their professional development

programs, both for their technical and their non-technical staff. In the public sector, in particular, it's important to reform government procurement processes that is a barrier to innovation in the government agencies. It's also important to build a culture of innovation within these organizations. Finally, we suggest that organizations should employ pilot projects to launch innovation efforts, and also to figure out safe ways to scale up the pilot projects to the entire organization. And again, if you want more details, you can find the paper online at brookings.edu.

To help us understand these and other issues related to technology innovation, we are delighted to have two distinguished experts with us. Melissa Koide is the CEO and director of the Financial Regulation Lab. There she works on issues related to the financial sector and ways to provide needed oversight. Beth Simone Noveck is the director of The Governance Lab at New York University, and also the chief innovation officer for the State of New Jersey. And she's also a professor teaching at Northeastern University.

For our audience, if you have questions for our panelists, you can email them to events@brookings.edu. That's events@brookings.edu. Or you can tweet using the #AIgovernance. That's #AIgovernance. So, we'll try and get to as many of your questions as possible.

So, I'd like to start with Melissa. So, you work on issues related to financial regulation. And I noted on your website that your organization talks about advancing safe and smart technologies in the financial services area in particular. So, my question for you is just, what are the problems you're seeing in this area? And how can technology help address those issues?

MS. KOIDE: Of course, I'm still making that mistake talking before the mute button is off. Well, Darrell, thank you for having me. It really is a pleasure to be joining this conversation.

I established FinRegLab after leaving the Office of Consumer Policy in the Treasury Department, where I was there for 4-1/2 years. And it was at the point in time where we were both seeing the growing use of new types of data and new technologies in the financial system. And we were looking at those developments with a lot of questions around what are the implications for people, for households, communities, the financial system in the micro sense, but the financial system at large? And ultimately

how does public policy, the laws and regs that we have in place, allow for, or frankly, prevent the benefits that we may get from the use of data and technology, but also make sure that we're guarding against some of the risks that come with new data being used for traditional purposes, whether it's assessing for fraud, which we're going to talk about today, credit risk assessing, or even helping individuals make determinations about their longer term savings and retirement goals?

So, FinRegLab is the organization I stood up to really help the collective financial sector generate some of the empirical analysis that we all collectively need to know and understand the benefits, but also the risks of a particular data or technology use. And so, I think what we've been going through over the past few years is both a collective and across all sectors about a recognition that technology, we're not going to put this genie back in the bottle and we actually really do need to better understand how do these more complex mathematical approaches to evaluating data, how do they work.

I do think, and we may talk about this, we also realized that, you know, data is out there. But we have to get really clear with ourselves around how we're making sure that data privacies are put in place, but that we do it in a way that is allowing for the benefits that can accrue from data use to really be deployed. And so, where we spend our time is quite focused on what are the inclusion implications, the potential inclusion benefits from using data, maybe financial data and maybe non-financial data, for things like, are you who you say you are, which gets to some of the fraud questions we're dealing with. But also where data and these more complex analytics may be instructive for making sure that the up to 50 million people who are insufficiently able to be credit risk assessed are able to be extended prudently credit, which is so important for near-term needs, but also longer term asset building opportunities.

So, I think collectively, we're all realizing that there's a lot that we need to know and better understand about how do we safely deploy AI and machine learning. And I continue to underscore data because that is how these machines ultimately work. And I think we're all digging in, whether it's the consumer advocacy community, the traditional financial sector, some of the new technology firms, obviously, Big Tech, too. I think they're coming to the table and asking how do we make sure that we are setting the rules and setting the governance expectations up, so that we're getting the benefits and,

again, mitigating the risks.

MR. WEST: I love your emphasis on data, because you're absolutely right that in terms of both machine learning and AI algorithms, the algorithms are only as good as the data. And as you very well know, sometimes some of the problems we see in terms of bias come from incomplete or unrepresentative data. And you're certainly correct to highlight the importance of protecting consumer privacy as these new algorithms are (inaudible) --

MS. KOIDE: And balancing as well, depending upon the context, too, right?

MR. WEST: Yes, definitely.

MS. KOIDE: Yeah.

MR. WEST: So, Beth, I want to bring you into the conversation. And Beth has a fascinating new book out from Yale University Press. It's entitled, *Solving Public Problems: A Practical Guide to Fix Our Government and Change Our World.* And I love that title. It's obviously a very ambitious title. So, how can technology help us solve public problems?

MS. NOVECK: (Laughter) How long do we have? (Laughter) Standing on one foot.

And so, let me connect back to some of the technologies that we're talking about, and in the process, say a little bit more about why I think that, in fact, the technologies, particularly the technologies of data, just to connect back to what Melissa is emphasizing, can be so powerful as tools for social change. I think you've detailed wonderfully in this report, which I really commend to everybody to read, because it provides just such a succinct overview of the powerful ways in which new technologies of AI and machine learning, natural language processing can be used to do things like spot fraud, waste, and abuse. And then obviously, it's a tremendous benefit.

And I think one that is under realized and talked about far too little amidst the welter of discussion, very needed discussion about the danger of nontransparent algorithms, about the danger of the abuses of AI, about the dangers of, again, the lack of transparency and control, which we're starting to get into and will in addition. But I think, as Melissa pointed out, it's a balancing act. And we do have to at the same time educate ourselves about the powerful things that AI can do for us, many of the

examples of which you provided in the report. And others of which you can know and experience with in my own work.

So, if you take something like Brazil's Court of Accounts, their primary auditing court, they are using -- like other institutions, they are using AI to audit public purchasing and contracting. And you see comptrollers' offices -- forward-leaning comptrollers' offices and auditing agencies around the world really turning to the use of machine learning, because it helps you frankly just to process huge quantities and volumes of information. And it can make it possible, therefore, to spot fraud, waste, and abuse. But I think also importantly, as we saw in Brazil during COVID, not only decrease corruption, but also speed up the process of contracting. So, it's a very, you know, pro-business strategy, not just anti-business. It can be really used to help make contracting processes with government and other organizations more efficient, as well as more accountable.

So, you see lots of uses of basically this just unbelievable power now to crunch through huge volumes of data to do things like -- you know, you probably received -- I'm looking for a bill that I might have received in the mail today that got to me through the post office thanks to the power of AI that helped the post office to sort and scan that piece of mail and get it to me faster.

So, in the book, *Solving Public Problems,* I do try to talk about the power of these technologies, AI and machine learning, among them big data more generally. But I really try to go into as you talk about in the report, and I am really influenced by a lot of the work that you have done, Darrell, now and of course in the past for which I'm really grateful, is the fact that we do need to upskill people in government, as well as outside about this equation that is balancing. What are the benefits? What are the things that we can do when we use big data to really help us to better understand the problem, but then to use things like the technologies of machine learning and AI to really help us to solve problems?

So, without that understanding of what this toolkit is that's now at our disposal in the 21st century, we really are going to be backwards in terms of our ability to know how we can do things differently, more effectively, and more efficiently, including to do things like spot fraud, waste, and abuse. The ability to know how to do that and how we can actually audit people better, how we can better spot,

you know, the people who are trying to commit fraud, whether it's in food benefits or unemployment benefits, and thereby make the system more efficient, more effective. We can do those things so much better, but we really have to understand what these tools are, and at the same time, understand those risks. Understand how, as you pointed out, Darrell, if we're using machine learning based on bad data or based on incomplete data, based on discriminatory data, we're only going to end up perpetuating bias. We're going to end up creating more problems than we'd solve. If we're throwing too much money into uses of AI to spot fraud, waste, and abuse among those applying for unemployment or food stamps, for example, those who are the most needy, we may be -- you know, it may be a bad use of our technology investments when we should be spending more time on investing in the technologies that we need to actually help deliver better services to people, especially during a pandemic and during rampant unemployment.

So, there's a lot of balancing questions that we can get into. But I think the broader point here and the lesson we truly learned during COVID is that the technologies of big data, the technologies that enable us to collaborate with people, the technologies that enable us to communicate across the distances, the broader toolkit that we now have available to us in the 21st century, AI is the most important among those tools. It can help us to solve problems more effectively. And we have to really deepen our understanding of what that broader toolkit is.

MR. WEST: Yeah, those are all great points. And those of you who want more details, check out Beth's book, *Solving Public Problems,* because she has some great recommendations on ways public agencies can do a better job.

The good news from the AI standpoint is there actually are a number of government agencies already starting to deploy AI and machine learning to combat fraud. In the paper, I talked about the Securities and Exchange Commission, the IRS, the Department of Treasury, CMS. There are lots of the issues in the healthcare area that need that tracking, and agencies are starting to use these tools.

But we also know that there are well-known problems of public sector innovation that limit or restrict the ability of government agencies to use technology in these innovative ways. And this

includes issues such as organizational siloing, not sharing data, cultural barriers to change within the agency, procurement system issues, and as each of you pointed out, an inadequately trained workforce.

So, a question for each of you. And we'll start with Melissa. How can we overcome these obstacles? You know, how can we deal with the organizational issues, the cultural issues, and the structural issues that sometimes restrict change in the public sector? Melissa, we'll start with you.

MS. KOIDE: Sure. So, when I was sitting in the Treasury Department in the Obama administration, I think many people realized how convinced the president in that administration was in terms of bringing in technologies and literally placing them within different departments and different offices throughout the federal government. I think that was a really helpful start to even enabling government employees, career civil servants to get exposed to some of these technologies, who were coming in saying, "What's the public policy mission that we're after? Let me introduce some different approaches for how we might thinking about getting after some of those challenges." I think -- at the same time, I think government is learning also from the private sector. You know, there are many agencies that have contractual relationships with the private sector for doing the business of government. And government is learning as it is engaging with those private entities who themselves are deploying data in new ways and using more and more complex, more sophisticated analytics for providing those government functions on behalf of government. And so, I think that helps.

But I also want to caution, and I'll tout some of the work that we're doing. I was a -- I was an econ major. And it is an evolution in terms of -- and it's -- there's an effort required to really understand how does -- how do different algorithms work, and how are they different from the ways that many of us were trained especially as economists with linear and logit regression approaches to evaluating quantitatively questions that we're grappling with. And so, there is a -- I think we have to be realistic with ourselves. There is a learning curve.

There was a lot that especially the federal government and importantly -- I mean, I spend a lot of time on the financial regulatory side, but then we're all going to have to be learning about, how does a math work? How much can we understand the math? And I think really importantly -- and Beth,

you alluded to this, there is -- there are still these very important and open questions around transparency and explainability, and even evaluating which public policy questions or applications we're trying to solve for to evaluate how much explainability or transparency do we think we really need as we're deploying more complex, you know, black box types of models and their use.

So, I think, you know, the government is making a headway. There's no question, I think, especially among the federal financial regulators. They are hiring up. They are staffing up. They are engaging in things that have been happening for many years, data policies, tech sprints. But there is going to be a learning curve for most of us who didn't grow up as data scientist or computer scientist. So, I'm happy to save a little bit more as we get further into the conversation about that.

MR. WEST: Well, Beth, I'd love to get your views on how we can overcome these barriers in the public sector. Like Melissa, you worked in the Obama administration, so you have the experience in the public sector before you returned to academia. How can we overcome some of these obstacles?

MS. NOVECK: Absolutely. And I have the privilege now also to serve in the public sector at the state level, which is where the action is at right now, especially when it comes to frontline delivery of services during COVID. So, it's been a great honor to do public service work at various levels of government.

So, where do I begin? (Laughter) Melissa has already pointed to a number of the challenges and approaches to them. But let me focus on something that's really kind of I've zeroed in on and that you emphasized strongly in the report and talked about now in your opening remarks, which is training. So, you know, that awareness building, Melissa, just to pick up where you left off, on the fact that none of us were trained as data scientists. Even if we took stats, if you -- like me twice in college, I think to take --

MS. KOIDE: I think I need to take it again still. (Laughter)

MS. NOVECK: I definitely wish I had paid more attention. Let me --

MR. WEST: Maybe that's true for all of us. (Laughter)

MS. NOVECK: I definitely wish I'd paid more attention. And yes, I definitely need to take it again, given where the world has gone in the -- so, even if we took stats in college, we -- there's a lot of development. It may have -- it did teach us the basics of how to formulate a hypothesis, which in the end is the most important thing we need to know how to do. But it's important to recognize that the skill, first of all, of how to define a problem and define a hypothesis, what is the thing that we're actually trying to solve for, is actually something that's not innate. It's something that we have to learn.

And especially in government, we tend to start with the solution. We tend to start with that thing that our agency has always done, whether it's a grant program or a particular kind of service that we deliver. Maybe we're typically in the vein of thinking about a policy solution, like a legislative or a regulatory solution. And, you know, we're not always pausing, especially now that we have such a broad range of tools at our disposal in ways of doing things, to really learn how to define a problem.

And then furthermore, to learn how do I define a problem with data, how do I then figure out what data sets I can use to answer my question, where do I find that data, how do I access it, how if I need to actually get the help of other people to gather it, so to crowdsource that data, or to evaluate the quality of that data to bring in other people to help me, none of those are intuitive skills. And even -- and it's not necessary, let me say, for us to be, you know, fully trained data scientists. But we do need to have enough awareness of what it means to solve problems with data and to use data to enable us to then reach out to collaborators and universities. And in an industry of whom there's a growing number, we still have to have some baseline understanding. And that training, sadly, is missing.

There are many, many, many, many more courses now offered in universities in data science. But at least in schools of public administration, colleagues of mine at Cornell School of Labor Relations did a piece. Maybe it's a year or two old now, so maybe there's been some change. But the study that they did said that basically in the top 25 schools of public administration, public management, and public policy, none of them had a requirement around data science. And if you look at training at the federal level and at the state level, I am required to take training in cybersecurity, in sexual harassment, in ethics. At the federal level, you have to do some also workplace safety and constitutional basics. None

of that is teaching us -- we have no requirement that you learn data science. It doesn't mean that, again, you can't go out and do that, but in a world in which in corporate America people are scrambling to upskill for the future of work, which includes a knowledge of how to think in a data analytical way and how to use data, to my mind, it's a huge gap that we are not training people in the public sector to define problems and to define problems using data and to understand how we can solve them using machine learning and AI.

So, these questions of, how do we spot fraud, waste, and abuse, you know, and how do I do that requires some training. And I think that's the lowest hanging fruit where we can start to remedy some of these gaps and deficits. The much more thorny and difficult questions about when in that balancing act we should use AI, we should not use AI, we should allow a black box, we should not allow a black box, those are difficult case by case questions. You know, I think you're absolutely right that there are plenty of cases though in which AI is just helping us to sort things and count things and manage a lot of content in ways that just makes it more efficient that are fairly straightforward. But we're never going to get to even those basic benefits if we don't invest more in talent and in upskilling. And that's not going to happen just with hiring a couple of innovation fellows or a couple of data scientists, as great as those things are. And I love those things because they allowed me to work in the federal government. We need, as they say, humans in the loop when it comes to the oversight of AI. And that means more humans who understand what AI is and what big data is, and that means training.

MR. WEST: Yes, I agree 100 percent with that. And I'll just add a quick footnote, which is we need the training not just for the technical people, which is what government agencies often emphasize in order to, you know, develop data skills and figure out how to analyze all the information, but the non-technical people, the people who are not writing the code, not writing the algorithms, they need to understand how to think about it in a broader sense.

And one of the issues I raised in the paper is just this whole idea of how you develop responsible AI that adheres to ethical values and broader societal values. So, Melissa, I'm just curious, how do you think we can develop more responsible AI? What should public sector agencies be thinking

about? How can we get both the technical people, as well as the non-technical government employees, to think about these broader considerations?

MS. KOIDE: Well, you emphasized it in your report. And I think it is something we all need to keep browbeating ourselves into realizing it is making sure that we have diversity of viewpoints at the sort of build -- not even just at the build stage, but at the -- what are we trying to apply the technology to. And then making sure that we have that diversity of perspectives and voices and consideration as we then go to actually deploy these more complex technologies, and making decisions about which data, you know, we want to be bringing into these decisions. I think that's an important place to start. I'm going to use this to be very concrete in terms of I think a consideration for us, it's where we spend a lot of our time. It's thinking about the public policy implications of using AI and machine learning in financial services.

I think one of the things that is maybe quite helpful as we contemplate questions around deploying AI in many different sectors, including government use, is looking at where we have existing laws that are forcing the deployment of AI to be done in ways that answer core fairness and inclusion and transparency ways. Now, to be fair, not every AI application requires all of that. But now, let me be much more concrete. In credit underwriting, which is where we're doing some research, in particular looking at AI and ML algorithms of different types, we actually have laws on the books that require that if you are a lender, and you are going to be making loans to consumers -- customers, you need to make sure that you can explain to the customer why a certain outcome was reached, a credit decision on the price that was offered. You need to, from a prudential regulatory standpoint, know that the models you're deploying are going to be stable and robust, even as the data underneath may be shifting. You also need to make sure that the models that you're using are going to generate fair outcomes. And we have a disparate impact analysis for evaluating that question. Sorry, cookies, we're making the rounds. I'm actually in the office today. (Laughter) Hence the wave.

And so, those laws that we have in place are then forcing the use of these models to -- the model builders to go through the process of making sure that we can answer those kinds of questions. And so, when we think about what are sort of the processes that we want to build around the

decision to use AI and, frankly, explainable AI from the start or post hoc approaches to explaining it, I think governance and defining what are those must-haves that you need on the outside of what the AI may generate are really important steps to making sure that you're going to be doing this in a way that you're going to feel like it's ethical and it's well considered.

MR. WEST: Yes, it definitely makes a lot of sense. Beth, your thoughts on ways we can develop responsible AI.

MS. NOVECK: Well, let me -- I think, obviously, there is a huge movement now among universities, to begin with, to start by, again, thinking about training people and dive -- which has the center for responsible AI, there have been a number with centers of excellence around AI, which have started. I'm going to Northeastern University now where I've started a job, which has a new center for experiential AI, as they put it, because it's really designed to be hands-on in terms of answering these questions not in a theoretical way, but in a very practical way, so that we can think about these balancing questions and decisions. And again, this question comes back to training people also in government.

So, to continue on that theme, which we started before, I will just put in a little plug for something that we have been doing in New Jersey and that we're expanding now and that we have been doing both for public servants and more broadly. So, there's not just a book called *Solving Public Problems,* but there's a course called Solving Public Problems, which is available for free at that title, solvingpublicproblems.org. That's aimed at a global social innovators. It builds off of work we did in New Jersey, which is at innovation -- sorry, skills.innovation.nj.gov, which is where we got our start, which was building a really early version one of a training in digital and data skills. We framed it as innovation skills, but it's really, again, thought about as problem solving skills for people in government. But it starts with the digital basics, which just explains -- you know, answers those questions that people are afraid to ask about, what is big data? What is AI? What is machine learning? And how do you use it in government?

This was a very, again, very early stage. We had no budget to do this, originally. So, it's just me in front of the camera, but providing a free course to the 70,000 public servants that we have. And now, thankfully, we have both legislative appropriation and philanthropic funding to allow us to

expand this to multiple states, and to grow this we hope to other areas of -- to other parts of governments like -- other governments -- I'm getting tongue-twisted here -- like Singapore, Germany, and Canada, which have already made investments again and upskilling people in AI and in specifically in these questions of responsible AI, so that you can have educated humans providing that oversight. And that's going to mean, you know, real changes in terms of just as you said, Darrell, like it can't just be about teaching the data scientists. The people who are in charge of buying this stuff, the procurement people have to understand what they're buying. The lawyers who review these agreements and contracts have to understand it.

You know, I'll give you a quick example and a shoutout here. I am also the chair of New Jersey's Future of Work Task Force. And in that role, we started looking at questions of the use of AI and machine learning in the workforce context to decrease costs and create efficiencies for employers in interviewing and selecting.

So, if anybody watching has a college kid or a recent graduate, you probably know something about -- you know, your kid has probably been subjected to an entry level interview, where they get interviewed by video, not by a human. And that video is used to select them or not select them for their internship for their job. It has huge cost savings and benefits for employers, but it runs real risks when as it has been divulged, the training data sets that are used to train that machine learning is the data set of existing employees. So, if a company only has White employees, and you are a person of color interviewing, or you are a person with an accent, or you are a person who doesn't look like or sound like the people who are ready at that company, you are going to get poorly scored in terms of the review by that piece of machine learning.

So, when looking at that issue and trying to examine that, you know, trying to look at that from the perspective of the state and understand what's the conversation we want to have with employers about educating them, for example, a lot of all big employers use this kind of software already. But as you start to talk about small and medium-sized businesses, you know, what's the -- or with large businesses, what's the conversation that we can have with them to educate them about the ethical risks,

as well as then what are the legal issues here, all of that long example to say, “The lawyers don’t necessarily understand or know what the machine learning software is or how it works.”

So, luckily, I had a wonderful intern named Shyamala Ramakrishna, who eventually became a full-time employee. She’s now back in law school, thankfully. And I say thankfully, because we need people like her who understand what the tools could do, can write about them, can then educate the lawyers, who can then help with the review process, as well as the other officials, to help people just understand what does that tech do. And so, we want people like that, who are bilinguals, who will be able to speak law and data, who can speak contracting and machine learning, who cut across fields. So, yes, it can’t be limited to just the data scientists. We need to expand our education, so that we can have humans in the loop in providing oversight over this AI.

And really, the best systems are those that provide that safeguard of human check. You know, Canada’s employment insurance sickness program uses optical character recognition and natural language processing to review doctors’ notes and to adapt abuses of their sick leave program. So, that sounds great, unless the machine learning works poorly. And essentially, you know, disadvantages, people deprive some of their benefits, which is why that system has a human in the loop review. The machine learning essentially offers up things for humans to look at, and then the humans make the ultimate decision. So -- and so, we need those humans to be educated, so that they can provide proper accountability and oversight for the technology and know what technologies to buy in the first place.

MR. WEST: Beth, I have to say, it sounds like you have about five different jobs. I’m really impressed at your ability to multitask --

MS. KOIDE: I am, too.

MR. WEST: -- for all of these activities.

MS. NOVECK: Let’s be clear, I don’t get paid for enough of them. (Laughter) If anybody gets the wrong idea here. (Laughter)

MR. WEST: There are none?

MS. NOVECK: They’re all nonprofit or public sector.

MR. WEST: Yes. So, I have one last question for you, and then we're going to move to some questions from our audience. And just to remind people, you can submit questions for our panelists by emailing us at events@brookings.edu. So, each of you at various points have mentioned this issue of workforce development, the importance of upskilling the federal workforce. I emphasized that theme in my new paper as well. Each of you have talked about the importance of that. I just want to push a little further in terms of, how do we do this? Who pays for it? And when and how often do people need to upgrade their job skills? Melissa, we'll start with you now.

MS. KOIDE: Yeah, that's -- it's a hard question. And I don't know if this is going to be specially satisfying answer. I have three teenagers -- three teenage sons. And I worry a little bit frankly -- and I might be slightly too myopic here -- that we're underprepared in a really big way. And it's to Beth's point earlier, we're going to be -- as we see the next generation coming up that we're potentially overtraining on the like how do you code and undertraining on the how are you a critical thinker. And I'm even thinking about my own kids a little bit in this way. But there's a lot of motivation to like code I want to be doing. And I think that's true. I think we also all recognize we have real challenges with the education system generally on sufficient critical thinking, critical analysis. And so, I think spending more time making sure that we've got people who are coming up with the ability to analyze, to ask questions, to feel confident too, frankly, in pushing back, like that's going to be one of the most important aspects of the future workforce. And they're the really sort of smart and safe use of deployment of AI in the end.

I'll let Beth jump in. I'd love to talk about data, though, if we get a minute, because I do think data is -- again, the AI is the sexy stuff, but it really is like, what data are you using? And I think just making sure that we underscore that in some ways would be good for the audience to hear about.

MR. WEST: Our next question actually is going to get into this data.

MS. KOIDE: Okay. Thank you.

MR. WEST: It's just a little to come back to on that.

MS. KOIDE: Yeah.

MR. WEST: But before we do that, Beth, your thoughts on workforce development and

how we actually do this?

MS. NOVECK: Well, I'm going to be very brief, because I want to hear you talk about data too. (Laughter)

And just to agree -- just to double down and agree with what you said which is why I like to use the term solving public problems, because I think we are overinvesting in the coding and in the purely technical aspects of things. I mentioned Singapore before, but that -- but let me come back to them for a moment to say that's a place where the public sector -- public -- their civil service training college has doubled down on making a big bet in saying we need to teach every public servant to learn how to code. I frankly think that's the wrong bet. I do think that making a big bet and saying in the 21st century we think there's a set of skills that people ought to know is the right thing. And I'd rather make a wrong bet than no bet, which is where we're at in this country. But I think it's the wrong bet, because teaching people how to think critically, how to define a problem, how to formulate a hypothesis, and above all how to think about doing that in collaboration with the communities whom we are serving is the most important part. Because as we think about using machine learning, whether it's for fraud, waste, and abuse, we need to be then in touch with the businesses who this is going to affect.

Or, you know, the -- if we're talking about risk recidivism algorithms, the incarcerated and formerly incarcerated, there are people's lives and livelihoods who are ultimately at the end of this. And so, we can't just be thinking about AI and machine learning simply as a tool of taxpayer accountability without thinking about the people in this ecosystem. And so, that means training people also in how to have those conversations and how to collaborate with others in making these decisions, again, in a very practical, efficient way of recognizing. At the end, we need to buy a piece of software to allow us to scan the mail or scan doctors' notes or sort a bunch of comments or a bunch of information, we can do this efficiently. And so, we need to think about very practical forms of training and upskilling that address how do we use these tools in government. So, not in a theoretical way, not simply in terms of questions of moral philosophy, which are great and important, but we really need to address what does this mean day in day out in terms of buying tools and using them in the public sector where we're under tremendous

time constraints.

MR. WEST: Okay. So, we are starting to get some questions from the audience. And Melissa, you're in luck. Our first question actually deals with data, your favorite topic. So, the specific question this person wants to ask is, how can government shift to more data-driven decision making? And then there are a couple of follow-ups. Where can data guide the goals of government agencies? How can data be used as a metric to determine if the goals actually happened then? And finally, how can data help us make decisions that end up reducing government fraud?

MS. KOIDE: Wow, that's great. Those are great questions. At a broad point, I just want to call out, I think the events over the past year and a half, the death of George Floyd, that many people recognized the inequalities that have existed and the systemic inequalities that have existed in lots of our foundational systems in education, the financial system. In many ways, it derived from our history but are also very much reflected in the data that we use and the data that we don't use. And I -- there's a lot of anxiety about how AI might perpetuate bias inequality and could.

And I just -- I want to say for the audience at large, I'll come back to the question that was raised, the attention and care to what data we aren't using can be using, should be using to make sure, you said it, Darrell, we have sufficient representation of everybody who were trying to make sure are served, whatever the used case is, that is critical. Again, AI is sexy, but I think much of this comes back to the data that you're plugging into those machines.

On how we sort of enable the government to be a data-driven function of our economy and provide the critical public mission services that it does in a data-driven way, some of the things that I think about when I was sitting in the Treasury Department and we were contemplating, things like -- for instance, we built a pre-retirement savings account aimed at the half of Americans who do not have employer-based retirement systems. Some of the challenges that we had because those are lower income families, was how do we actually reach those families to get this product out? It was the myRA account out in front of those households in the first place. Well, who's got that kind of data? Well, the IRS does, of course.

But the things that you run into an IRS data is precious and needs to be protected. But it is also so valuable for data-driven policy purposes is realizing that we are going to be grappling with ultimately what are legislative changes that are going to have to happen, so that the data that -- will be so advantageous for deployment for a data-driven government system or going to have to potentially be re-contemplated. Whether some of the constraints that preclude different agencies from using different data sources, whether it's something like trying to make sure that Americans are aware of something like this myRA product, whether it's making sure that we're doing a better job of avoiding fraud when we are, you know, dispersing, whether it's stimulus checks or other benefits, there are -- and we ran into this -- and it's been a long enough time that I could talk about this, I think. There are real challenges both from sort of policy perspectives within different agencies, because of how data sources were created. But ultimately, importantly too, we have legal guardrails around how data can be used. But yet things have evolved since when those laws may have been put in place. And I think it is going to require ultimately for better use of data and more sophisticated math for public policy purposes to come back and realize we've got to spend some time thinking about those policies and those laws that are in place that govern the data and their deployment.

Beth, I mean, I suspect you saw some of this too. Not to put you on the spot, but -- and you're probably seeing it at the state level as well, I would imagine. And those are most from my --

MR. WEST: And so, I was going to put Beth on the spot with the very same question, because I know she has written about data-driven decision making. So, Beth, your thoughts on how government can shift to that kind of approach?

MS. NOVECK: Oh, well, there's -- again, this is one -- this is such a huge topic with a lot to say here. But what's -- in addition to the things that we've already talked about, I think there is much more investment that we need to make in integrated data systems as a starting point.

So, we do a lot in terms of data collection for compliance purposes, which is not very useful for purposes in solving -- for the purpose of solving problems. So, we're collecting a lot of data that just sort of goes into the drawer, goes into the black box that doesn't get used. The need to really collect

data in a way that we can actually integrate it with other datasets. So, by that, I mean, you know, where we have IRS data that tells us things about income and outcomes for people, we want to be able to better connect that to education data, testing data, health data, and other data sources, so that it's at those intersections. And that's where also the tools of AI and machine learning are really very helpful for doing that kind of data integration.

But being able to do that, we have to do more to fund those integrated data systems, both inside government and outside government. There are wonderful initiatives in New Jersey. We have something called NJEEDS, the New Jersey Educational Employment Data System, which is one of those integrated data warehouses. It's a collaboration between multiple state agencies and the state's university at Rutgers to make integrated data available. You have some agencies, again, funding longitudinal data systems, but often not integrated data systems. You then have things outside of government, like UPenn's AISP project, Administrative -- I'm going to botch the -- I'll do the acronyms, but I'm going to botch what they stand for. The AISP project, Chapin Hall at the University of Chicago. Julia Lane runs something called the Coleridge Initiative. It's a collaboration between NYU, University of Maryland -- and again, I'm going to forget the third partner there. But these are efforts that provide both cloud-based infrastructure for storing data and integrating it, as well as training programs to help people better make use of that data.

So -- but government needs to do much more. And philanthropy can help to fill some of those gaps, where government is slow to step in to fund just the availability of data and the ability to integrate that data, so that it's more useful. We also need to be able to do more again to open up data that can be open -- that can be made fully open. And here I should distinguish between that I wasn't clear at the outset, there is data that can be published as open data, and we have a lot more of that that's available in machine readable form. But then there's data about things like, as we were talking about, income, education, health records, private data, what we might call administrative data, that needs to be provided on secure infrastructure with limited access that's made available to that. And the ability to actually, you know, provide both of those infrastructure for that, the governance, as Melissa talked about,

for how to access that data appropriately, all of that is really essential preconditions for the ability to then layer on AI and machine learning tools that can help us make sense of all of that data. But without that infrastructure, which you both talked about for data, as well as in the training that has to be layered on top of that, and then the governance that allows us to make very clear who can actually have access to that data, you know, we're not going to get very far. So, these things really sort of go in tandem, go hand and glove together.

MR. WEST: Okay. Thank you. So, Mafi (phonetic) from our audience has a question for each of you as well. This person wants to know, do you -- do each of you believe a global data governance framework is necessary to establish rules to data access? And this person specifically wants to know about the need for a global data governance framework in regard both to fraud reduction and bias issues.

Melissa, do you want to start with that?

MS. KOIDE: Yeah. It's a really -- it's obviously a really important question. It's a challenging question, because I can articulate some of the benefits of having a global consistency in terms of global standards consumer privacy. I mean, it is the -- it's essentially the interoperability right, and importantly, seeing the benefits in terms of data flows.

I'll think about this in the context of correspondent banking and money flows, you know, through large depository institutions who provide correspondent banking capabilities. Somebody is trying to send money to, you know, somebody in Sudan. There I know your customer requirements that are employed, if we had greater consistency in terms of standards around that data. There could be less friction and, frankly, therefore, then less cost associated with making those kinds of remittance services possible at a better price. You know, and I definitely have heard the arguments in terms of large corporations and the challenges they face because of the different governing. I mean, we're seeing this now in the U.S. with some of the state data privacy laws emerging.

On the other hand, I think this is just the hard work that we all need to go through, which is thinking about what are we really trying to solve for with very use case specific purposes. You know,

our financial system I'm biased, but it is very -- it's obviously very integral to thriving economies. And the certainty of controlling risk -- really bad risks all the way to fraud cost is a really important aspect of maintaining a stable and thriving financial system. Data is a critical aspect of that, and inclusion is an important priority for many of us. If we are aware that different types of data may be really instructive for satisfying know your customer requirements, and it may be non-traditional data -- I mean, right now, we rely on name, address, and government-issued ID, other types of data may be really instructive for verifying your identity -- we may want more flexibility in terms of the data that is being used for satisfying that know your customer requirement. That may not be the same case in terms of a commerce activity and whether or not we would have sort of the same data access constraints or not that we have in the financial services sector.

So, I do -- you know, I just think we're at a point where we have to be thinking about, do we need sector-specific laws around consumer data access, consumer data privacy, and what might they look like? And then how do they potentially harmonize with -- I think everybody would agree we need better consumer data privacy at large, but how are we going to normalize and make sure that we've got systems that allow for the benefits of data to accrue in the ways that they can for purposes that are important?

MR. WEST: Okay. Beth, we're going to give you the last word on this. So, your thoughts on the need for a global data governance framework?

MS. NOVECK: Well, let me go in a slightly different direction, just to complement what Melissa said and to just say I think that there are specific sectors that are making interesting progress on this issue. I'm thinking about things like the open contracting partnership, for example, where they have been doing work around the world with governments to try to set a standard for opening up and managing data that is about contracting.

So, coming back to our sort of core issue of fraud, waste, and abuse, I think the idea of really having -- they have found and I think successfully demonstrated that having a global flexible standard that can allow us to do better machine analysis of government contracts and then also have

comparative analysis across jurisdictions is a really powerful and successful example -- or burgeoning successful example. It's building on, you know, other sector-specific work.

Just yesterday, I got an email from NIST about efforts that they're making to develop a global governance -- data governance framework for research data. So, I think that they are all focused on university-specific research. Oh, Melissa, I'm going to --

MS. KOIDE: Right. We need this. (Laughter)

MS. NOVECK: I've been looking for people to invite, so consider yourself --

MS. KOIDE: Good.

MS. NOVECK: Consider yourself invited. (Laughter) So, I think we're seeing a lot of sectors where there has been increasing success and moving in this direction, but we have to be able to answer all of the questions that Melissa so artfully raised.

MR. WEST: Well, Melissa and Beth, I want to thank you very much for sharing your views. Each of you have added important insights to these questions about how technology can further a public sector innovation and both what the opportunities are as well as what the risks.

And for those of you in the audience who would like to learn more about this, you can read our work at brookings.edu. You can also check out our TechTank blog also located at Brookings, and our TechTank podcast, which looks at technology policy issues in a greater detail. So, thank you very much for tuning in.

And Beth and Melissa, thank you for taking the time to join us. I really appreciate getting your viewpoints. Thank you very much.

MS. KOIDE: Thank you.

MS. NOVECK: Thanks for having us.

* * * * *

CERTIFICATE OF NOTARY PUBLIC

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

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2024