U.S. SENATE COMMITTEE ON
COMMERCE, SCIENCE & TRANSPORTATION
Senator Maria Cantwell, Chair

U.S. Senator Maria Cantwell

Senate Committee on Commerce, Science, and Transportation
Hearing: The Need to Protect Americans' Privacy and the AI Accelerant

July 11, 2024

Opening Statement
VIDEO

The Senate Committee on Commerce, Science, and Transportation will come to order. I want to thank the witnesses for being here today for testimony on the need to protect Americans' privacy and AI as an accelerant to the urgency of passing legislation. I want to welcome Dr. Ryan Calo, University of Washington School of Law and Co-Director of the University of Washington Technology Lab; Ms. Amba Kak, Co-Executive Director of the AI Now Institute in New York; Mr. Udbhav Tiwari, Global Product Policy Director for Mozilla from San Francisco; and Mr. Morgan Reed, President of ACT, the App Association, of Washington D.C. Thank you all for being here on this very, very important hearing.

We are here today to talk about the need to protect Americans' privacy and why AI is an accelerant that increases the need for passing legislation soon.

Americans' privacy is under attack. We are being surveilled…tracked online in the real world, through connected devices. And now, when you add AI, it is like putting fuel on a campfire in the middle of a windstorm.

For example, a Seattle man's car insurance increased by 21% because his Chevy Bolt was collecting details about his driving habits and sharing it with data brokers, who then shared it with his insurance company. The man never knew the car was collecting the data.

Data about our military members, including contact information and health conditions is already available for sale by data brokers for as little as 12 cents. Researchers at Duke University were able to buy such data sets for thousands of active military personnel.

Every year, Americans make millions of calls and text chats to crisis lines seeking help when they are in mental distress. You would expect this information would be kept confidential. But a nonprofit suicide crisis line was sharing data from those conversation with its for-profit affiliates that it was using to train its AI product.

Just this year, the FTC sued a mobile app developer for tracking consumers' precise location through software embedded in a grocery list in a shopping rewards app. The company used this data to sort consumers into precise audience segments. Consumers who used this app to help them remember when to buy peanut butter didn't expect to be profiled and categorized into a precise audience segment like "parents of preschoolers."

These privacy abuses and millions of others that are happening every day are bad enough. But now, AI is an accelerant and the reason why we need to speed up our privacy law.

AI is built on data, lots of it. Tech companies can't get enough to train their AI models -- your shopping habits, your favorite videos, who your kids' friends are – all of that. And we're going to hear testimony today from Professor Calo about how AI gives the capacity to drive sensitive insights about individuals. So, it is not just the data that is being collected. It is the ability to have sensitive insights about individuals in the system.

This, as some people have said, referring to [Dr. Calo's] testimony now, is creating an inference economy that could become very challenging. That is why you also point out, Dr. Calo, that a privacy law helps offset the power of these corporations and why we need to act.

I also want to thank Ms. Kak for her testimony because she is clearly talking about that same corporate power and the unfair and deceptive practices, which we've already known should be given to the FTC as their main authority.

The lack of transparency about what is going on with prompts and the AI synergy is that people are no longer just taking personal data and sending us cookie Ads. They are taking that are putting that actually into prompt information. This is a very challenging situation. And I think your question is, are we going to allow our personal data to train AI models is very important for our hearing today.

We know that they want this data to feed their AI models to make the most amount of money. These incentives are really a race to the bottom where the most privacy protective companies are at a competitive disadvantage.

Researchers project that if current trends continue, companies training large language models may run out of new publicly available, high-quality data to train AI systems as early as 2026.

Without a strong privacy law, when the public data runs out, nothing is stopping them from using our private data. I'm very concerned that the ability to collect vast amounts of personal data about individuals, and create inferences about them quickly at very low cost, can be used in harmful ways, like charging consumers different prices for the same product.

I talked to a young developer in my state and I said what is going on? And he said, well I know one country is using AI to basically give it to their businesses. And I said, well why would they do that? He said, they want to know when a person calls up for a reservation at a restaurant how much income they really have. If they don't really have enough money to buy a bottle of wine, they are giving the reservation to someone else.

So, the notion is that discriminatory practices can already exist with just a little amount of data for consumers.

AI in the wrong hands is also a weapon. Deepfake phone scams are already plaguing my state. Scammers used AI to clone voices to defraud consumers by posing as a loved one in need of money. These systems can re-create a person's voice in just minutes, taking the familiar grandparent scam and putting it on steroids.

More alarming, earlier this month, the Director of National Intelligence reported that Russian influence actors are planning to covertly use social media to subvert our elections. The ODNI called AI "a maligned influence accelerant," saying that it was being used to more convincingly tailor a particular video and other content ahead of the November election.

Just two days ago, the DOJ reported that it dismantled a Russian bot farm intended to sow discord in the United States. Using AI, the Russians created scores of fictitious user profiles on X, generated posts, and then used other bots to repost, like, and comment on the posts – further amplifying the original fake posts. This was possible at tremendous scale given AI. I'm not saying that misinformation might not have existed before, and may not have been placed in a chat group, but now with the use of bots and AI as an accelerant, that information could be more broadly distributed very, very quickly.

Privacy is not a partisan issue. According to Pew Research, the majority of Americans across the political spectrum support regulation. I believe our most important private data should not be bought or sold without our approval. And tech companies should make sure they implement these laws and help stop this kind of interference.

The legislation that Representative McMorris Rogers and I have worked on does just that.

And I want to say… that Senator Blackburn and I will be introducing [legislation] called the COPIED Act, which provides much-needed transparency around AI-generated content. The COPIED Act will also put creators, including local journalists, artists, and musicians, back in control of their content with a watermark process that I think is very much needed.

First Q&A

[AUDIO] [VIDEO]

Witnesses:

Ryan Calo, Professor, University of Washington Law; Co-Director, UW Tech Policy Lab

Amba Kak, Co-Executive Director, AI Now Institute

Udbhav Tiwari, Director, Global Product Policy, Mozilla

Morgan Reed, President, ACT | The App Association

Sen. Cantwell: I'd like to go to a couple of charts here if we could. One, you know, when we first started working on the privacy law, I'm going to direct this to you, Professor Calo. But what got me going was the transfer of wealth to online advertising. I don't think people really quite understood how much the television, newsprint, radio, magazine, the entire advertising revenue shift went online.

Now we're just talking about the internet - could you bring that a little closer, please? Just up a little closer. So, we are now at 68% - I don't know if people can see that, but we're now at 68% of all spending. Two-thirds of all spending of advertising has now taken place online with data and information. So that trend is just going to keep continuing.

Now, you and you and I have had many conversations about the effect of that on the news media, that having community voices, you know, our community in Seattle, King 5 or the Seattle Times couldn't exist, if it had misinformation, it just wouldn't exist. But in the online world you could have misinformation, there was no corrective force for that. But all the revenue has now gone to the online world.

And the second chart describes I think a little bit about your testimony that I want to ask a question about, and that is the amount of information that is now being derived about you that AI [has] this capacity to derive sensitive insights.

So that trend that I just described, where two-thirds of all advertising revenue, I mean, somebody said, data is like the new oil, right? It's just where everybody's going to go and make the money. So that's a lot of money already in that shift over that, in those years that I mentioned on the chart. But now you're saying they're going to take that information and they're going to derive sensitive information about us. Ms. Kak said it's the way your voice sounds. You've described it as, you know, various features.

So could you tell me how protecting us against that in the AI model, why that's so important. And, you know, I just want to point out, we're very proud of what the Allen Institute is doing on AI. We think we're the leaders in AI applications. We're very proud of that in healthcare, farming, energy.

You know, we have an agreement today between the United States and Canada in principle on the Columbia River Treaty. I think water AI will be a big issue of the future, how do you manage your natural resources to the most effective possible use? So, we're all big on the AI applications in the Pacific Northwest.

But we're very worried about the capacity to derive sensitive insights and then, as you mentioned, an insurance company, or somebody using that information against you. Could you expound on that, please?

Ryan Calo: Absolutely. I was talking to my sister who's on the board of the Breast Cancer Alliance about my testimony and she said, “You know, Ryan, just make sure that people know how important it is for AI to be able to spot patterns and medical records to ensure that people get better treatment, for example, for breast cancer.” And I agree with that.

I'm also proud of all the work we're doing at University of Washington and Allen. The problem is that the ability to derive sensitive insights is being used in ways that disadvantage consumers and they're not able to figure out what's going on and fight back right? For example…

Sen. Cantwell: Thereby driving up costs?

Ryan Calo: For example, right? I mean, we know why everything costs $9.99 right? It's because your brain thinks of it as being a little bit further away from $10 than it really is. But the future we're headed to, and even the present is a situation where you're charged exactly as much as you're willing to pay in the moment.

Say, I'm trying to make dinner for my kids, and I'm just, you know, trying desperately to find a movie for them to stream that they both can agree on, right? You know, if Amazon can figure that out, or Apple can figure that out, they can charge me more in the moment when I'm flustered and frustrated, because they can tell.

If that sounds far-fetched Senator, Uber once experimented with whether or not people would be more willing to pay surge pricing when their batteries were really low on their phone because they'd be desperate to catch a ride. Amazon itself has gotten into trouble for beginning to charge returning customers more because they know that they have you in their ecosystem.

This is the world of using AI to extract consumer surplus. And it's not a good world and it's one that data minimization could help address.

Closing Q&A

[AUDIO] [VIDEO]

Sen. Cantwell: I'd actually like to follow up on that I was going to anyway, so it was a good lead. And I want to thank Senator Thune for his leadership because he not just on data security, but on privacy and now on your proposal as it relates to AI, very good concepts there that we should continue to work on and hopefully we'll get to a markup soon.

But this notion, this data that came out by the Department of Justice that it dismantled a Russian bot farm intended to sow discord in the United States, and that the AI the Russians created scores of fictitious profiles, and then generated these posts.

And so, I'm trying to – you were just talking about this ecosystem that's created, right? And now, here we have bots who are just exploding with the information because we've given them so much data. You can say it came from the social media platform that they collected it and then then that information got scraped, and then the information got to the bots, and then the bots put it on this accelerant and a bad actor can use it against us.

So why is this so important to now have a tool to fight against this? Because it's the bot system is out of control, but the AI accelerant on the bot system makes it an imperative.

Ryan Calo: I can only agree with you, Senator. Obviously, misinformation and propaganda are not new, but the ability of adversaries of all kinds domestic and foreign, to create plausible looking, and very damaging misinformation campaigns has become quite acute.

I mean, I'll just use one example. As you know, the Center for an Informed Public studies, misinformation and disinformation. One example is that someone created a deep fake, that was not real, fictitious, that gave the appearance that there had been a bomb go off at the Pentagon, okay. Which was so concerning to people that it actually caused a dip in the stock market until people figured out that it wasn't real. The ability to create a seemingly real catastrophic event is very, very dangerous.

But you're talking to something even more basic, which is that AI makes it possible to generate much, much more disinformation, and have it appear different from one another. Different media, different phrasing, and everything else. It's deeply concerning.

I think there are ways in which a privacy law could help, actually, but the problem of disinformation misinformation probably is broader still.

Sen. Cantwell: So what do we do about the bots from a regulatory perspective?

Ryan Calo: Yeah that's hard. I mean, so, states, like California have a Bot Disclosure Act requires that if you're operating a bot, in certain ways, commercial and electioneering, that you have to identify yourself as fake. The problem, of course, is that

Russian disinformers are not going to comply with our laws. And so I think part of the response has to be political and economic.

And that's one of the main reasons that the federal government needs to get involved, because it's not something the states can address. States can't find global consensus around sanctioning bad acting around information, you know? But I think that a response placing responsibility on the platforms to do as much as possible, since they have control over their own platforms to identify and disincentivize this kind of misinformation that's automated is also really key.

Sen. Cantwell: Thank you. Well, I think that concludes our hearing. I know it's a very busy morning for everybody. The record will remain open for two weeks, and we asked members to submit their questions for the record by July 18.

I want to thank the witnesses, all of you, very informative panel. We appreciate you answering these questions and in helping us move forward on important privacy and AI legislation.