

**ARTIFICIAL INTELLIGENCE APPLICATIONS TO
OPERATIONS IN CYBERSPACE**

HEARING

BEFORE THE

SUBCOMMITTEE ON CYBERSECURITY

OF THE

**COMMITTEE ON ARMED SERVICES
UNITED STATES SENATE**

ONE HUNDRED SEVENTEENTH CONGRESS

SECOND SESSION

—————
MAY 3, 2022
—————

Printed for the use of the Committee on Armed Services



Available via: <http://www.govinfo.gov>

—————
U.S. GOVERNMENT PUBLISHING OFFICE

WASHINGTON : 2025

COMMITTEE ON ARMED SERVICES

JACK REED, Rhode Island, *Chairman*

JEANNE SHAHEEN, New Hampshire	JAMES M. INHOFE, Oklahoma
KIRSTEN E. GILLIBRAND, New York	ROGER F. WICKER, Mississippi
RICHARD BLUMENTHAL, Connecticut	DEB FISCHER, Nebraska
MAZIE K. HIRONO, Hawaii	TOM COTTON, Arkansas
TIM Kaine, Virginia	MIKE ROUNDS, South Dakota
ANGUS S. KING, Jr., Maine	JONI ERNST, Iowa
ELIZABETH WARREN, Massachusetts	THOM TILLIS, North Carolina
GARY C. PETERS, Michigan	DAN SULLIVAN, Alaska
JOE MANCHIN III, West Virginia	KEVIN CRAMER, North Dakota
TAMMY DUCKWORTH, Illinois	RICK SCOTT, Florida
JACKY ROSEN, Nevada	MARSHA BLACKBURN, Tennessee
MARK KELLY, Arizona	JOSH HAWLEY, Missouri
	TOMMY TUBERVILLE, Alabama

ELIZABETH L. KING, *Staff Director*
JOHN D. WASON, *Minority Staff Director*

SUBCOMMITTEE ON CYBERSECURITY

JOE MANCHIN III, *Chairman*

KIRSTEN E. GILLIBRAND, New York	MIKE ROUNDS, South Dakota
RICHARD BLUMENTHAL, Connecticut	ROGER F. WICKER, Mississippi
JACKY ROSEN, Nevada	JONI ERNST, Iowa
	MARSHA BLACKBURN, Tennessee

CONTENTS

MAY 3, 2022

	Page
ARTIFICIAL INTELLIGENCE APPLICATIONS TO OPERATIONS IN CYBERSPACE	1
MEMBERS STATEMENTS	
Statement of Senator Joe Manchin	1
Statement of Senator Mike Rounds	2
WITNESS STATEMENTS	
Lohn, Andrew, PhD, Senior Fellow, Center for Security and Emerging Technology, Georgetown University	5
Moore, Andrew, PhD, Vice President and Director of Google Cloud Artificial Intelligence, Google Corporation	9
Horvitz, Eric, PhD, Technical Fellow and Chief Scientific Officer, Microsoft Corporation	15

ARTIFICIAL INTELLIGENCE APPLICATIONS TO OPERATIONS IN CYBERSPACE

TUESDAY, MAY 3, 2022

UNITED STATES SENATE,
SUBCOMMITTEE ON CYBERSECURITY,
COMMITTEE ON ARMED SERVICES.
Washington, DC.

The Subcommittee met, pursuant to notice, at 2:43 p.m. in room SR-232A, Russell Senate Office Building, Senator Joe Manchin (Chairman of the Subcommittee) presiding.

Committee Members present: Senators Manchin, Blumenthal, Rosen, Kelly, Rounds, Ernst, and Blackburn.

OPENING STATEMENT OF SENATOR JOE MANCHIN

Senator MANCHIN. The meeting will come to order.

I want to extend a warm welcome and thanks to our distinguished witnesses today, who have all taken time out of your important duties for your companies and academic institutions to help educate all of us the Cyber Subcommittee of the Senate Armed Services Committee on the application of artificial intelligence and machine learning technology to the critical missions of offensive and defensive operations in cyberspace.

Artificial intelligence and machine learning are extremely technically complex topics so I would highly encourage our witnesses to provide as many real-world examples as they can. What I am saying is bring it down to our level, okay—

[Laughter.]

Senator ROUNDS. All the way to kindergarten?

[Laughter.]

Senator MANCHIN. Might have to—in your answers and simplify technical concepts as much as humanly possible for the benefit of the members and the public that are viewing this hearing.

I cannot overstate our need for AI [artificial intelligence] application in cyberspace operations, and I believe our witnesses' prepared statements will eloquently express your sentiments.

There is a huge shortfall of technically trained cybersecurity personnel across the country in government and industry alike. This shortage is likely to continue to worsen, especially as cyber threats intensify in scope and scale. Keeping up with the demand of capacity in this field will therefore require massive gains in workforce productivity, which, practically speaking, means automation by computers. AI technology can power this automation and productivity growth.

Not to belabor the point but China has four times our population. There is no way we are going to win a competition in manpower, or woman power, or person power that can be dedicated to an important mission. Computer-driven automation powered by superior software innovation is the only option that we have. As Dr. Moore wrote in his prepared statement, with AI the work of 5,000 people can become the equivalent of 50,000 people.

Additionally, AI can discover subtle signals and patterns of malicious cyberattacks in a sea of noise better and faster than humans. AI can also help to automate actions to contain and eradicate cyber penetrations.

Commercial computer-aided intrusion detection technologies that are widely used today already process enormous quantities of data, provide alerts to human analysts of suspicious actions and anonymous events. But these products generate enormous numbers of false positive—false alarms, if you will. So many, in fact, that our analysts are overwhelmed and cannot possibly investigate them all. This is why we fail to find the genuine needles in the haystack, even when they may be noted by our security event management systems. AI, however, will increase the rate of detection of real intrusions while lowering the false alarms.

AI, in short, can enable our cyber forces to achieve scale and speed in defensive cyber operations. The flip side of this is that AI can also tremendously benefit the offensive side of cyber operations. Just as AI algorithms can scan our own networks for vulnerabilities, they can discover vulnerabilities and attack vectors and adversary networks that we can exploit.

Make no mistake. Our adversaries will capitalize on this technology, using AI to power attacks on our networks as well as increasing their ability to detect our intrusions on their networks and to respond quickly. We can use the Russian SolarWinds attack to illustrate the potential danger. The SolarWinds software supply chain operation compromised thousands of networks, but the Russians can only manually exploit a limited number of the targets they infected.

However, the use of AI technology in the future will enable Russia or China to take advantage of every target that they compromise. It would be disastrous if we failed to be ready. Yet, while the Defense Department is developing AI applications for business efficiencies and warfighter support, I fear we are not moving at the necessary speed in cyberspace.

Commercial cybersecurity companies have, for a number of years, been developing and applying AI technology to their products, and the Department of Defense is benefitting from that investment. Microsoft's Defender product is a good example.

A direct DOD [Department of Defense] investment in cyber AI is lagging. I look forward to hearing recommendations from our witnesses on what we could be investing in and where we need to focus our attention.

So I turn now to my friend, Senator Rounds, for his remarks.

STATEMENT OF SENATOR MIKE ROUNDS

Senator ROUNDS. Thank you, Senator Manchin. First I would like to thank our witnesses for appearing at our hearing today.

The topic of today's hearing is one that is of particular interest to me. Over the last few years this subcommittee has witnessed firsthand, at our many hearings and briefings, how dynamic and rapidly evolving the cyberspace domain is. New technologies are emerging all the time, and that is a good thing, but it also poses new challenges. Malicious cyber actors have demonstrated time and time again how quickly they can exploit these new technologies to attack our systems and infrastructures. The Department of Defense must move just as quickly to understand these emerging technologies, both to provide our United States Cyber Command with cutting-edge capabilities for their cyberspace mission and also to defend against these technologies being used against our Nation. I cannot think of a technology that will have a broader impact on cyberspace than the application of artificial intelligence or AI.

I would like to share an excerpt from the final report of the National Security Commission on AI—this is the NSCAI—which captures the landscape nicely, and I will quote:

“AI-enhanced capabilities will be the tools of first resort in a new era of conflict as strategic competitors develop AI concepts and the technologies for military and other malign uses and cheap and commercially available AI applications, ranging from deep fakes to lethal drones, become available to rogue states, terrorists, and criminals. The United States must prepare to defend against these threats by quickly and responsibly adopting AI for national security and defense purposes.”

“Defending against AI-capable adversaries operating at machine speeds without employing AI is an invitation to disaster. Human operators will not be able to keep up with or defense against AI-enabled cyber or disinformation attacks, drone swarms, or missile attacks without the assistance of AI-enabled machines. National security professionals must have access to the world's best technology to protect themselves, perform their missions, and defend us. Put simply, our adversaries are going to use AI against us, so we must use AI to defend against them.”

I look forward to hearing from our witnesses today. But to begin with, I would like each witness to give a short, basic introduction to AI that will help us understanding these technologies better and help us describe these issues to our Senate colleagues so that we can have the policy discussions that need to be completed. Please give us a short overview of the difference between a normal computer program, machine learning, artificial intelligence, and quantum computing.

Now I know that sounds like a crazy thing, but clearly if there is anybody that can do it, I would just ask you to keep down at like our kindergarten or first-grade level.

I would also like to hear from the witnesses on their perspectives of the current state of adoption of AI technologies in industry to defend against AI-capable adversaries. How are your companies leveraging AI today to defend your cyberspace infrastructure? How do you think the Department of Defense needs to leverage AI for their cyberspace missions? I would appreciate your thoughts on the

best ways to leverage AI-enabled cyber defense to protect against AI-enabled cyberattacks.

Thank you again to our witnesses for coming here today. Senator Manchin.

Senator MANCHIN. Thank you, Senator Rounds. Before I begin I want to recognize you three for being here, and I really, really appreciate it. I think it is tremendous. It will be a tremendous hearing here.

We have Dr. Eric Horvitz. He is a Technical Fellow and Chief Scientific Officer for Microsoft. We have Dr. Andrew Lohn, who is the Senior Fellow for Security and Emerging Technology at Georgetown University. We have Dr. Andrew Moore. He is Vice President and Director of Google Cloud Artificial Intelligence at Google.

So we look forward to hearing your updates and we will start, Dr. Horvitz, with you.

Dr. HORVITZ. Thank you. Let me first answer the overview question.

AI systems are programs, just like any other computer software, but they are special in that they are designed to emulate aspects that we would call human intelligence. So what are the capabilities we recognize as intelligence? The ability to perceive, to see and hear; the ability to reason about situations, for example, by considering multiple pieces of information or observations; the ability to make good decisions, even where uncertain; the ability to adapt to learn from experiences and information over time; the power to use and understand language; and other capabilities that are a little bit more nuanced, like the ability to generalize from specifics, to form useful abstractions about the world. So AI scientists write programs to emulate these capabilities of intelligence.

I should say that there has been progress on all those fronts that I just mentioned, all those dimensions of intelligence. But over the last 20 years we have seen an absolute revolution in the learning part. This is the learning part of AI and it is called machine learning. So it is a part of the larger discipline of artificial intelligence. It is one sub-area but it has come to be so important in supercharging the other areas, including computer vision, language abilities, speech recognition, and so on.

Now quantum computing is a very different thing. Quantum computers harness quantum physics to computer, that use behaviors seen on a microscope scale, behaviors discovered by physicists with interesting names like “superposition” and “entanglement,” and to clean up any potential misconception, or a broad one, successes in quantum will not give us general purpose computers. A quantum computer solves special kinds of problems, like factoring large numbers, critical cryptography. So working quantum computers, when they come to be, at scale, will be able to solve extraordinarily hard problems in those areas that they are great for, thus, for example, breaking current cryptographic protections, which makes them of very deep interest for national security.

Senator ROUNDS. [Presiding.] On behalf of the chairman, thank you very much. I appreciate it. Did you have anything else that you wanted to add before we move forward?

Dr. HORVITZ. Well, I can answer your second question. I guess you asked a very broad question about what companies and enterprises are doing to protect themselves right now.

You know, we are building infrastructures, and I would love to see more effort in DOD and other Federal agencies, infrastructures that go from being able to sense across many computers for patterns, being able to collect that data across the world, for example, and across organizations, of course, to employ machine learning on the infrastructure, to build predictive models, and to build filters and detectors.

We have to have a great workforce of professionally trained cybersecurity experts to work with these AI systems, because despite what we think about AI, the big gain is going to be in human AI iteration and collaboration. So we need those teams, no matter how good our AI is, and lastly we need to have a system of pushing out updates quickly, to make patches and to stay in touch with end users.

Senator ROUNDS. Thank you. On behalf of the chairman, and he shall return very quickly. Dr. Lohn?

STATEMENT OF ANDREW LOHN, PhD, SENIOR FELLOW, CENTER FOR SECURITY AND EMERGING TECHNOLOGY, GEORGETOWN UNIVERSITY

Dr. LOHN. Thank you. I would like to start by thanking Chairman Manchin and Ranking Member Rounds and the Members of the Subcommittee. Thank you for inviting me to be here. I am Andrew Lohn from the CyberAI project at the Center for Security and Emerging Technology at Georgetown University. It is an honor to be here.

When we talk about AI, to answer your question, I like to use the Defense Science Board's definition. They say the capability of a computer system to perform tasks that normally require human intelligence. As an example, accounting software used to be AI when tax filing normally required humans, but now it is so common that it is no longer considered AI.

But if AI is about what software can do then machine learning and normal programs are about how that software was made. For normal programs, somebody writes all the logic themselves—if this, then that, many times. For machine learning, nobody sets those if-then rules. The computer determines them after many examples.

Quantum computing is, as Dr. Horvitz said, kind of a different sort of process that touches a little bit on normal computer programs, machine learning, and AI, but is mostly separate.

With that background in hand, I would like to talk about three areas where AI intersects with cybersecurity: one, how AI promises to improve cyber defense; two, how AI may improve offensive cyber operations; and three, how AI is itself vulnerable.

AI for cyber defense is not a new concept. Spam and anti-phishing filter have been protecting users for many years, and AI has long been touted as a tool for companies that hunt for malware or search for intrusions. Some of these techniques have become the foundations of modern cybersecurity. But in general there is a back-and-forth. Whereas an AI learns attacker tactics, the attackers adapt their tactics to evade that AI.

To date, those attacker tactics have not relied much on AI. That is likely because so much has already been automated. A human can direct a computer to find possible targets on a network, then direct it to exploit those targets, then delist the files or folders to encrypt or extract. The human really only has to manage the system while the computers already do most of the work.

That said, there are reasons to automate attack code. In 2015, when Russia first cut power to Ukraine, the hackers had to take over the mouse and manually shut down the grid. By the next year they developed new malware that had more automation.

An attacker may just simply want to operate a machine's speeds. In 2016, DARPA [Defense Advanced Research Projects Agency] hosted the Cyber Grand Challenge, where fully automated systems competed to secure themselves while breaking into each other. These systems relied more on hard-coded rules than machine learning, but they were impressive. The winning system competed against some of the world's top humans the following day, and though it ultimately finished last there were times where it was leading some of these human teams, which is an impressive result in only its first year.

This was the first and last such challenge in the United States, but China was struck by the potential and has hosted at least seven of their own autonomous hacking challenges since. It is unclear how capable their systems are, but it is clear that both China and Russia are working to develop software that can discover vulnerabilities and in some cases run their cyber offenses more autonomously.

AI systems are technological marvels but they too are software with their own vulnerabilities. Most famously, it is easy for an attacker to change just a few pixels in an image to make a detection system to stake objects it is looking for. It is easy to imagine these techniques disguising parts of an invading force or directing drones or coastal defense systems to the wrong targets. It is even easier to envision digital decoys that overwhelm that system. It is not clear how susceptible these systems are in the real world yet, but we may soon find out as countries rush to deploy autonomous military capabilities.

But rather than wait for our systems to be deployed, our adversaries may target the AI supply chain. Our systems are often merely adapted from existing ones that may or may not be trustworthy, and the data used to train or adapt those systems can be compromised too.

Today, most of the models, datasets, and tools are provided by trustworthy organizations such as those represented by Dr. Horvitz and Dr. Moore. But China, in particular, is making a push to provide more of these resources. If they succeed, then DOD would face an unwelcomed decision between using the most capable systems or the most trustworthy ones.

I do not wish to overstate the impact of artificial intelligence on cybersecurity nor the severity of the vulnerabilities in AI. I only hope to alert you to the potential that is being developed. Our adversaries are highly capable and grow more emboldened every year, and they have been developing increasingly autonomous attack software. Similarly, although we have seen only a few attacks di-

rectly on AI systems, the potential is no secret. Our adversaries are surely aware of the vulnerabilities, and we should expect attacks as soon as AI systems prove their value on the battlefield.

Thank you.

[The prepared statement of Dr. Andrew Lohn follows:]

PREPARED STATEMENT BY DR. ANDREW LOHN

Chairman Manchin, Ranking Member Rounds and Members of the Subcommittee, thank you for the opportunity to testify before you today. I am Andrew Lohn, Senior Fellow in the CyberAI Project of the Center for Security and Emerging Technology at Georgetown University. It is an honor to be here with Dr. Horvitz and Dr. Moore.

At the CyberAI project, we try to anticipate the impact of artificial intelligence and cybersecurity coming together. In these opening remarks I'd like to touch very briefly on three areas of that intersection: 1) How AI promises to improve cyber defenses, 2) How AI may improve offensive cyber operations, and 3) How AI itself is vulnerable.

Before I begin I would like to make clear that everything I am saying comes from an external vantage point. At CSET, we do not use classified sources and I do not have access to any private corporate data. Much of the cybersecurity world exists behind those closed doors, so there are surely capabilities or incidents that I am not aware of. However, much of it plays out in public, so we can try to extrapolate the future from the past.

CYBER DEFENSE

AI for cyber defense is not a new concept. Spam and anti-phishing filters have been protecting users for many years. And AI has long been touted as a tool for companies that either hunt for malicious software or search for irregular behaviors that could indicate the presence of an intruder. Some of these techniques have become the foundations of modern cybersecurity while others are marketing hype. Sometimes it is difficult to tell the difference. In general, there is a back and forth where once an AI learns attacker tactics, those attackers adapt to evade that AI.

CYBER OFFENSE

To date, those attacker tactics have not relied much on artificial intelligence. That is likely because so much has already been automated that humans only need to manage the attack. A human can select a computer script that scans the victim network and reveals possible targets. The human can then run another script that tries to exploit the vulnerabilities found by the first one. Then another script can enumerate the files and folders to encrypt or extract. The human only has to manage the system while computers already do most of the work.

That said, there are a few reasons to want the attack code to be able to make those decisions by itself. For example, the number of victims may be too large for humans to manage, or the targets may be difficult to communicate with over the internet. In 2015, when Russia first cut power to Ukraine, the hackers took over the mouse and had to manually select components of the grid to shut down. By the next year, they had developed new malware that was programmed with the ability to make some of those decisions without direct human involvement. The second version of that malware that was discovered last month is still being evaluated but appears to follow suit.

In addition to being able to operate where command and control might be difficult, an attacker may simply want to make decisions at machine speeds. In 2016, the year of the second power grid attack on Ukraine, DARPA hosted the Cyber Grand Challenge where fully automated systems competed to secure themselves while breaking into each other. These systems relied more on hardcoded rules than the advanced techniques we think of as AI today, but they showed some signs of promise. The winning automated system competed against some of the world's top human teams the following day. Though it ultimately finished last, there were periods where it outscored some of the human teams, an impressive result in only its first year.

This was the first and last such challenge in the United States, but China was struck by the potential and has hosted at least seven of their own autonomous hack-

ing challenges.¹ It is unclear how capable their systems are, but it is clear that both China and Russia are working to develop software that can discover vulnerabilities and, in some cases, is capable of running their cyber offensives more autonomously.

The threat extends beyond software that can autonomously find and exploit vulnerabilities. The human component is becoming more vulnerable. Humans are usually the weakest point in the security of a system, which is why 36 percent of intrusions involve phishing attacks.² Click rates have been falling for years but recent advances have made AI-generated text nearly as convincing as what humans can write. Combining that writing ability with the vast amounts of personal data on the internet provides a concerning potential for AI to make phishing campaigns even more effective than they already are.

VULNERABILITIES OF AI

Today's AI systems are technological marvels but they too are software complete with vulnerabilities of their own. They share some of the same vulnerabilities of more traditional software, but also introduce some new ones that can be very difficult to fix.

Most famously, it is easy for an attacker to change a few pixels in an image to make a detection system miss objects that it is looking for or to mistake objects in a scene for what the attacker wants them to see. Most strikingly, the attacker's manipulations can be so subtle that humans cannot tell the difference between the original and the doctored images.

It is easy to imagine these techniques being used to disguise parts of an invading force, or to direct autonomous search and destroy drones or coastal defense systems toward the wrong targets. It is even easier to envision digital decoys that overwhelm the system or its human operators. It is not clear yet how susceptible these systems are in the real world rather than just the laboratory setting, but we may find out soon, as many countries have become more keen to deploy autonomous military capabilities.

The United States is among those deploying autonomously capable systems, but our adversaries may not wait to subvert them. There are plenty of opportunities for interference throughout the design process. AI can be very expensive to train, so rather than starting from scratch, a system is often adapted from existing systems that may or may not be trustworthy. And the data used to train or adapt the systems may or may not be trustworthy too. It takes surprisingly few nefarious volunteers or low-paid online workers to corrupt a dataset in ways that give attackers a backdoor to control the model. Today most of these models and datasets are built and hosted by relatively trustworthy organizations such as those represented by Dr. Horvitz and Dr. Moore, but China in particular is making a push to provide more of these resources. If they succeed, then DOD would face an unwelcome decision between using the most capable systems or the most trustworthy ones.

CONCLUSION

I do not wish to overstate the impact of artificial intelligence on cyber security nor the severity of the vulnerabilities in AI. Cyber operations are still human-intensive both on offense and on defense. And there are few openly reported cases outside of a laboratory environment where AI algorithms were attacked directly. I only hope to alert you to the potential that is being developed. Our adversaries are highly capable and grow more emboldened every year. They have been developing increasingly autonomous attack software for years, and we should expect that those preparations will eventually come to fruition. Similarly, although we have seen only a few attacks directly on AI systems, the potential is no secret. Our adversaries are surely aware of the vulnerabilities and we should expect attacks as soon as AI systems prove their value on the battlefield.

Dr. HORVITZ. Senator Rounds? Just to ask courteously, I thought you were asking us to go round robin on your special questions first, but I have a prepared statement as well.

Senator ROUNDS. Oh. That was your question, was it not?

We will go to Dr. Moore and I will come back to you.

Dr. HORVITZ. Thank you very much.

Senator ROUNDS. Dr. Horovitz, I am sorry.

¹Dakota Cary, Robot Hacking Games, 2021.

²Verizon Data Breach Investigations Report, 2021.

Dr. Moore?

**STATEMENT OF ANDREW MOORE, PhD, VICE PRESIDENT AND
DIRECTOR OF GOOGLE CLOUD ARTIFICIAL INTELLIGENCE,
GOOGLE CORPORATION**

Dr. MOORE. Thank you very much, Chairman Manchin and Ranking Member Rounds, and Members of the Committee. My name is Andrew Moore. I am Vice President and General Manager of Google Cloud AI. I most recently served as a Commissioner with Dr. Horvitz on the NSCAI, and I previously served as Dean of Carnegie Mellon University, which I cannot help but mention, won the grand challenge of which you spoke.

[Laughter.]

Dr. MOORE. I really want to thank the committee's support for advancing artificial intelligence.

Chairman Manchin, you have really supported the relationship between National Science Foundation and West Virginia University. I really respect WVU, and I go there frequently. It is a really great asset.

Dr. Rounds, as Ranking Member Rounds, thank you for your support of actually doing AI baselining at the Department of Defense. This really, really matters, so thank you for that. I greatly appreciate all the support you have given to NSCAI's recommendations as well.

My colleagues nicely defined AI. I am going to just leave it simply that AI refers to technologies that can make decisions from billions of possible alternatives in almost real time, and modern AIs do improve themselves as they are doing this.

I want to give you a tangible example because that is what Chairman Manchin asked for. If I am lowly drone trying to attack a U.S. battle fleet—and this is a hypothetical, non-classified example—if I am a lowly drone trying to attack a huge battle fleet you might think I have got no chance because I am so outgunned. But suppose I can search, in the space of a second, over a trillion possible trajectories, misleading directions relative to the sun, deal with all the various possible other tricks, maybe even a flock of seagulls, at the same time. I have got this advantage that I am not fighting against a battle fleet. I am fighting against the worst-case scenario out of a trillion scenarios for that battle fleet. So that is what the power of AI is. It is where we have these supercomputers, so superhuman abilities to search lots of alternatives.

AI powers many of our products, and we are using it to help organize the world's information. For example, AI is used to help you predict the best route in Google Maps. Many of our Google Cloud solutions are used by the Department of Defense. One of my favorite examples is our partnership with the U.S. Navy, where autonomous drones are able to take pictures of corrosion on the sides of warships and quickly and efficiently inspect what is at most danger, what needs servicing as quickly as possible. This not only saves a large amount of repair money but it helps keep us in better readiness than we would otherwise.

There are many other examples of our work with DOD, and I think it is fair to say that all the large what we call hyperscalers,

the big internet companies, are proud of the opportunity to help serve the U.S. Government.

Now I have got to talk about cybersecurity. Cybersecurity, as my colleagues have mentioned, is interesting because everything happens just so fast. Google has a huge network which is being attacked all the time from huge numbers of places, including many state actors, so we have to have everything we can do to secure it.

What we have done is a pattern that I see developing in the DOD. I strongly recommend it. I am going to sort of highlight it now. There are three parts to it. The first one is using AI to defend against attacks, the other two are how we organize the data and people in the Department of Defense.

Using AI to defend against attacks, first, the most obvious one that I have already kind of illustrated is you want to be watching millions of possible attacks, known attacks, every second, looking out for all of them. That is the basic one, and that is where you cannot possibly afford to use humans for that. Things are happening too fast.

The second one, which is interesting, is emerging attacks, people ingeniously coming up with new methods, and AIs are coming up with new methods, so you have to be learning new patterns or detecting whole new kinds of attacks in real time. This is where the full power of adversarial AI comes in.

Finally, while you are doing all of this on your perimeter you have got to be ready for the insider threat. So artificial intelligence is extremely important and it plays a large part in conjunction with the Zero Trust approach that the Department of Defense has brought in. That plays a large part in how to deal with the very real, unfortunately, insider threats, looking to see strange human patterns.

I cannot resist following up on one of Chairman Manchin's comments about we are building these AIs on the other side of building these AIs. New technologies, which I would like to make sure that the government is aware of, are things you will see, for example, in poker-playing robots. One of these championed at Carnegie Mellon University, which are using the work of mathematician John Nash to solve game theory games. The important things about that are AI are aware of the facts that the other person is learning from them at the same time they are taking their actions, and the AI cannot just automatically do the most obvious thing, because it actually has to conceal its activities.

So National Science Foundation is funding this kind of research into very advanced AI, and it is very important that we do not ignore that aspect.

I want to talk about the second part of all of this, which is the data inside the Department of Defense. It is not okay if there are lots of different silos of data. We need, especially in certain major scenarios, we need something to have a full understanding of what is going on, and to do that it is not okay for people to need to pick up a phone call, to phone to ask for help from a different set of sensors or a different database somewhere else.

So the notion of using concepts such as knowledge graphs to join together information from many different sources of data to form a more complete picture, extremely important. For example, I am

extremely supportive of the Joint All-Domain Command and Control, JADC2, which is seeking to do this by allowing information sharing through interfaces and services across all domains.

AI without data these days is pretty worthless, and so the absolute importance of getting through the sort of social or organizational hurdles, for people to share information about threats, is essential.

The final thing I want to quickly mention is humans and machines working together. I know that there are bills which advocate for a cyber reserve unit, for example, and thank you for those. I strongly support that. As it comes in, the people that we are putting on the frontlines with AI need powerful tools designed for humans to work with machines. Many of us in industry are working incredibly hard at the moment to make sure that those tools are usable by folks trained up to become an AI force as easily as possible. So we have put lots of effort into AI platforms which help guide users to quickly be able to respond and work on new and important AI issues as they come up.

Let me be clear about what I mean here. If we get a threat, some major, new attacks surfaces, and we have to get together a whole bunch of people to deal with it, that is done in an hour or so, at the very latest, and you immediately have people with the tools, who know how to use them, to combine the data to build a system against some new threat in ideally less than a day, and within a week or two all you are doing is double-checking the patches and doing postmortems to make sure it never happens again.

The nightmare for me is if, instead, the U.S. Government ever found itself in a position it said, “Hey, this is not really working. We better start a procurement process to find a contractor to bid on solving this thing.” I strongly believe you actually need people in the Armed Services with the capabilities to get on this stuff right away.

So with that I again want to express my appreciation. I have a lot more thoughts on this.

Senator MANCHIN. We are going to have questions for you too, Doctor. We are going to have a lot of questions for you.

Dr. MOORE. Great. So thank you for the opportunity, and I look forward to helping continue work with Congress on this issue.

[The prepared statement of Dr. Andrew Moore follows:]

PREPARED STATEMENT BY DR. ANDREW MOORE

Chairman Manchin, Ranking Member Rounds, and Members of the Committee, thank you for the opportunity to appear before you this morning.

My name is Andrew Moore. I am Vice President and General Manager of Google Cloud Artificial Intelligence (AI). I most recently served as a Commissioner on the National Security Commission on AI (NSCAI) and I currently serve as a task force member on the National AI Research Resource (NAIRR). I previously served as Dean of the Carnegie Mellon University School of Computer Science and have spent my career as a computer scientist specializing in machine learning and robotics. I have also spent time as an advisor to the Department of Defense as a member of Google Cloud’s leadership team.

I appreciate the Committee’s support for advancing AI—thank you Chairman Manchin for your leadership in driving a partnership between the National Science Foundation and West Virginia University to ensure more funding for AI research in last year’s appropriations bill, and thank you Ranking Member Rounds for your continued support of AI baselining at the Department of Defense. And I greatly appreciate the support both of you have provided for the NSCAI and its work. During my

time there, NSCAI submitted strong recommendations to the Committee and the Department of Defense (DOD). In addition to the NSCAI recommendations, it is also worth revisiting the recommendations led by the National Academies and sponsored by the Office of the Director of National Intelligence on the Implications of Artificial Intelligence for Cybersecurity. AI can be an incredible asset but, as with any new technology, can also present new vulnerabilities.

A useful definition of AI is a machine which seems to have human or sometimes superhuman capabilities at a task we might previously have said needs uniquely human intelligence. In recent years some of the biggest advances have come from neural networks, which simulate billions of neural connections in biological nervous systems. The two big technological battles happening in academia and corporations around the world are first, how to scale it up to trillions of connections, and second how to turn really amazing technology demonstrations into practical deployed systems that are actually useful.

AI can refer to any number of technologies involving artificial systems designed to or having the ability to learn. One way Congress has itself described it is as “An artificial system designed to think or act like a human, including cognitive architectures and neural networks.” A neural network is a computation system used to classify and analyze data using a process that mimics the function of the human brain. The data is fed into the first layer of a neural network, with each layer making a decision, then passing that information onto multiple nodes in the next layer. Some modern neural networks have hundreds or thousands of layers, with millions and even billions of parameters—the output of which can do such things as classify an object, or find patterns in data. This means that AI can process more information more quickly than a human: finding patterns and discovering relationships in data that any human would never be able to process on its own given the volume of data being processed. And, AI is not limited by time of day, the need for breaks, or other human encumbrances. In the cloud, AI and machine learning can be “always on,” continuously working on their assigned tasks.

For cybersecurity and in the context of national security, having the upper hand in AI against your adversary is critical. There is a race to see who can get machines to provide as much defense as possible. For example, AI systems are absolutely necessary to automate aspects of cybersecurity. The U.S. remains the leader in AI, but we must ensure we continue to do this at scale.

AI powers all Google’s products. And, importantly, we use AI to monitor our network infrastructure and attempt to predict and detect threats to our network or users. One of AI’s critical uses is finding anomalies in activity that would indicate a new threat vector.

We of course also use it to support users when you search using Google Search. AI enables the most relevant responses to surface. AI is used to help predict the best route for you in Google Maps, detect misspellings or grammar mistakes in Google Docs and more. AI makes our products better by making them work for the user, by understanding and anticipating the user’s preferences and needs. The same AI technology is used at Google to keep our users secure from phishing attacks on email, from malicious actors hacking into documents, and more.

AI also powers a lot of the solutions Google uses to serve the Department of Defense. For example, one of my favorite partnerships between the Department and Google Cloud is with the U.S. Navy, where commercial drones are used by the Navy to take millions of images of the hull of ships and other hard-to-reach parts of ships, and then sends the images to Google Cloud to analyze the images using AI technology. We have trained Google Cloud to recognize any picture of rust corrosion and when spotted, the system alerts a Naval analyst to review and schedule the ship for repair. By leveraging Google Cloud’s native computer-vision capabilities, the team successfully identified “corrosion of interest” in aerial images of vessels, with confidence scores of more than 90 percent and with very few false positives. This was an engineering feat that required complex integration between emerging software and hardware technologies, and has saved the Navy thousands of hours a year in readiness.

There are other examples of our work with the Department I would be happy to share—including using AI imaging to detect cancer, using AI to assist building simulation technology to train Air Force pilots and more.

As I mentioned, a critical use of AI is in cybersecurity solutions. While it is often hard to predict new kinds of attacks and new threats as they are constantly emerging, Google runs one of the largest and most secure networks in the world. Due to its scale and the threats it faces on a daily basis, we have a level of insight and visibility into the world of cyber threats, through all our global platforms, that allows us to assess and develop cutting edge defenses to whole classes of threats, not just particular attacks.

At Google Cloud, we have leveraged this expertise to deliver a new, unified AI experience through our Cloud services which give every data scientist, data analyst, and machine learning (ML) engineer the same tools we use at Google to secure their own networks. Like the Department of Defense, we must be constantly vigilant and ensure Google Cloud's security solutions and updates are informed by vulnerability and threat information as it evolves in real time. Indeed, as we have seen in many recent cyberattacks, some of the most dangerous attacks are those where multiple systems communicate in unforeseen ways to create chaos and wreak havoc. With this in mind, I'd like to offer the following observations and recommendations for how this committee can further support the Department of Defense in its mission using AI capabilities to secure its networks, applications and personnel:

1. *Using AI to defend against attacks.*

As we have learned through recent events, our customers in the public and private sector increasingly understand that they must protect different parts of their network with different applications. There are known threat factors but all organizations must be able to spot new threat vectors that are constantly emerging and recognize that insider threats continue to be a real concern. DOD must stay on top of ensuring they have the right resources. And I'll attempt to illustrate this with how Google thinks about each of these threats:

- a. *First, AI allows for monitoring known threats at a massive scale.*
 - i. Threat hunting and investigation tools are used to look at historical data and determine if exploitation was attempted—or they can be used as vehicles for monitoring active exploitation.
 - ii. *On-demand scanning* of containers (containers are isolated software packages that contain everything the software needs to run).
 - iii. Active scanning that detects Domain Name System (DNS) calls to known malicious sites (the DNS is effectively the “phonebook” of the internet).
 - iv. Tools to detect common exploit attempts.
- b. *Second, AI excels at anomaly detection and emerging threats.*
 - i. Implementing passive detection rules in *Event Threat Detection (ETD)* and *Security Health Analytics*.
 - ii. Tools to detect potential attacks include using *custom reports in Edge API Analytics*. (API stands for Application Programming Interface, which is a software intermediary that allows applications to talk to one another)
 - iii. Tools to create web application firewalls as layered defenses to protect against attacks until all vulnerabilities can be patched.
- c. *Finally, AI can assist in identifying insider threats.* AI is particularly best suited to identify insider threats because it has the capacity to analyze billions of parameters an hour. The need to protect against insider threats is also part of the Administration's push toward agencies embracing a Zero Trust philosophy.

It is worth noting that AI is trained and powered by data and so having accurate, well curated sources of data is key to threat hunting. For example, tools like VirusTotal provide threat context and reputation data to help analyze suspicious files. These tools use live flux samples of data against historical data in order to track evolution of certain threat actors, malware families and automatically generate “indicators of compromise” to protect organizations.

2. *Breaking down data silos to harness the full power of AI.*

Today, data exists in many formats, is provided in real-time streams, and stretches across many different data centers and clouds all over the world. From analytics, to data engineering, to AI/ML, to data-driven applications, the ways in which we leverage and share data continues to expand. Data has moved beyond the data analyst and now impacts every employee, every customer, and every partner. With the dramatic growth in the amount and types of data, workloads, and users, we are at a tipping point where traditional data architectures—even when deployed in the cloud—are unable to unlock their full potential. As a result, the data-to-value gap is growing.

Insights are not just locked in raw data—they're locked in data from many sources and silos—meaning the ability to unify datasets is a prerequisite to applying AI, in a structured and purpose-built manner, to applications. There are many opportunities to ensure the Department can operate different services across different and disparate data networks. For example, Joint All Domain Command and Control (JADC2) is seeking to do just this by allowing information sharing through interfaces and services across all domains. AI can enhance the security of this effort and ensure that the Department is reviewing the data for learnings, anomalies, changes and patterns.

A great example of this is how we are using AI systems for anti-money laundering and countering the financing of international terrorism (“AML/CFT”). Money laundering fuels drug trafficking, human trafficking, and terrorist activities. AI-enabled AML/CFT approaches, on the other hand, can develop a much more sophisticated analytic lens capable of ingesting massive volumes of data, in a more timely way, to detect new patterns and anomalies that might bypass simple, rules-based logic. These engines can be trained to improve accuracy, reduce false-positives, and help perform internal risk assessments and better determine when, amongst millions of legitimate transactions being processed, bad actors are trying to move criminal money. AI can further incorporate more contextual signals and generate more targeted flags for investigators, reducing toil and allowing them to focus on the most serious issues that are identified. AML highlights the opportunities this committee, the Department, and the private sector can focus on as we ensure the United States continues to lead in the development and deployment of artificial intelligence.

At Google Cloud, we have made it a priority to deliver cutting-edge cloud-native capabilities for distributed workloads spanning public cloud, private cloud, and multi cloud environments. Additionally, managing data across disparate locations creates silos and increases both risk and cost—especially when data needs to be moved. Innovations such as data lakes offer the ability to unify data stored across multiple cloud providers without worrying about the underlying storage format or system, which eliminates the need to duplicate or move data, which in turn reduces cost, inefficiencies, and security risks. This approach permits innovation by using multiple vendors, clouds and technologies, but it also increases competition and will likely lower prices for the Department and taxpayers.

But, it is not just about ensuring we have thousands of databases and data tables. The personnel at DOD must have the proper skills and training to capitalize on these insights. If an AI system identifies 27 new threats, we need DOD teams sitting inside the Department to quickly prioritize and address the threats. This is a vastly different way of thinking than the traditional “waterfall approach” which involves slower, deliberate planning and can constrain the more agile type work that is necessary in these scenarios. This is a classic challenge in large bureaucratic organizations. At Google, by the time a threat is discovered, we need to have a patch in place well within 24 hours. In two weeks, we need to have developed a permanent solution, and shortly thereafter, we need to have a post-mortem which describes the event and includes a recap of the timeline, description of user impact, root cause, action items and lessons learned.

3. Capitalizing on data insights through human-machine teaming.

To understand the full opportunities of AI in DOD’s mission, it must also ensure the Department can inject AI into its workflows. Understanding of AI-based tools cannot be limited to those with programming skills only.

To be clear, this is not a procurement issue. Instead, what is needed is leaders to think about whether AI tools within the Department can help solve the challenge. Usually the answer is yes. Then the Department must have the ability for teams to quickly build/adapt/leverage an AI system—in hours or days—to address problems like finding a ship lost at sea or responding to an active threat event. Vetex AI and AI infrastructure provide tools for data scientists to build custom AI for their own problems at scale. Today, AI platforms like ours require nearly 80 percent fewer lines of code to train a model with custom libraries and data scientists can now build and train models 5X faster on Vetex AI than on traditional notebooks.

Human teams, such as those formed by analysts and data scientists, must have a common understanding and opportunity to bring machine capabilities into the mission by building out an end-to-end AI experience where they can extract value from data and use AI out of the box to maximize value at a moment’s notice. Imagine for a moment that there were different types of databases across the department that track shipping container movements around the world. Then imagine that another database holds information about the contents of each container and yet another that can analyze components or materials used in individual products inside the containers. Brought together, an AI system then identifies that there is a particular metal alloy used in each of the products that all appear to be heading to the same country in different pots. Cross-linking and joining data in this manner allows for constant pattern detection for unexpected defensive concerns and can help analysts identify emerging trends from data across different departments in new and novel ways.

This is especially important as our adversaries will continue to look for gaps in systems—including AI systems—that may be exploited in both simple and complex ways. The term “adversarial AI” may be known to you already but it is an increasing area of research. As I mentioned earlier, the most dangerous attacks are those where multiple systems communicate in unforeseen ways to create chaos and wreak havoc. AI is further enabling these kinds of attacks, but it can also help defend against them.

In the last several years, researchers at Carnegie Mellon proved that AI can act in super-human ways. This was recently demonstrated in a straightforward game of poker. Operating on incomplete information and against multiple parties, the system beat leading professionals by bluffing and misleading human adversaries. This is an indication of more to come. The poker demonstration offers valuable insights into the future of cyberdefense and warfare: our adversaries will continue to understand new and novel ways to leverage AI to mislead and attack.

As you can see, from poker games, to thwarting money laundering, to protecting networks from cyberattacks, to spotting corrosion on the hull of Navy ships, AI can be used to spot patterns and anomalies generally faster and with more precision than humans. AI technology can help the Department scale its analysis of these patterns and anomalies for threats and learnings. I urge the Department to embrace AI, particularly in its efforts to secure its networks.

Let me conclude by recognizing the importance of the work of this subcommittee, and its efforts to ensure the United States remains a leader in AI and cybersecurity, given the increasingly complex landscape. With AI, the work of 5,000 people can become the equivalent of 50,000.

My hope is that the Department will continue to make the right investments in training, technology, and management that will facilitate more experimentation, prototyping, and execution that will be necessary. It is also critical that the Department continues to make comprehensive technology investments—in cloud migration, data set curation, API management, network connectivity to increase operational effectiveness and deliver proven innovation.

We all have a role to play to prevent and detect threats online. Being transparent with governments, customers, and government entities when it comes to cyberattacks is one of our key principles and is critically important when responding to incidents at scale. I suggest this committee continue to encourage the use of modern, cloud-based technologies to improve long-term security, based on investments in defense-in-depth. Diversity in the ecosystem, especially with cloud-based solutions, reduces overall risk and fosters and improves resilience against attacks. In addition, products and services that enable portability and interoperability foster resiliency.

Thank you for the opportunity to speak with you today. I look forward to continuing to work with Congress on these important issues, and I’m happy to answer any questions you might have.

Senator MANCHIN. [Presiding.] Thank you, sir. Thank you.

Dr. Horvitz, I am sorry we misinterpreted. I thought that is where Mike was coming.

Dr. HORVITZ. Yeah, so did I.

Senator ROUNDS. What were you thinking?

Senator MANCHIN. His intro was so profound that I thought, well, here we go.

**STATEMENT OF ERIC HORVITZ, PhD, TECHNICAL FELLOW
AND CHIEF SCIENTIFIC OFFICER, MICROSOFT CORPORATION**

Dr. HORVITZ. So Chairman Manchin, Ranking Member Rounds, and Members of the Subcommittee, thanks for inviting us today to testify on this important topic. I am Eric Horvitz. I currently serve as the Chief Scientific Officer of Microsoft.

AI researchers and engineers work to automate tasks that are typically associated, as I mentioned earlier, with human cognition, such as perception, pattern recognition, prediction, reasoning, and learning. We are seeing developments in AI now at a pace we could not have predicted just a few years ago.

I will focus my remarks today on three areas that lie at the intersection of AI and cybersecurity” number one, advancing our cybersecurity with AI; number two, malicious uses of AI to power cyberattacks; and three, an interesting area evolving quickly, attacks on AI systems themselves.

First, using AI in cyber defense. It is an exciting area and it is being used today to detect attacks and respond to attacks in real time, at scales that would be nearly impossible with manual techniques. These methods can recognize patterns of activities associated with attacks, they can adapt to new attacks, and detect attacks never seen before by identifying subtle similarities and signals that adversaries try hard to hide.

AI methods help cybersecurity teams to scale their efforts, which is critically important when there is a global deficit of nearly three million cybersecurity professionals and when cybersecurity job opportunities are projected to grow 33 percent over the next decade.

Second, AI-powered cyberattacks, that is using AI on the offense, is an important area of concern. To date, there is scarce information on the active use of AI in cyberattacks. It is expected, though, that AI technologies will be used to scale cyberattacks and increase their efficacy, and the power of offensive AI, we will call it, has been demonstrated by red teams and a growing community of researchers. So given the pace of AI, we have to prepare ourselves.

Offensive AI spans several areas. Researchers have demonstrated the ability to efficiently guess passwords, to attack industrial control systems, to create malware that can evade detection.

Another form of attack uses AI methods for social engineering. This is aimed at the soft, human side of cybersecurity. The work includes impressive formal demonstrations that show how AI can be used to ultra-personalize phishing attacks on individuals, generating content that compels people, even security experts, to click on links that emit malware.

Finally, another rising concern is attacks on AI systems themselves, what we call—and you will hear this over the years—adversarial AI. These attacks use AI techniques to disrupt the operation of target AI systems or gaining access to their data or processes.

Here is an example about how AI attackers have used AI techniques to fool AI systems, causing the system to fail dramatically. In stunning demonstrations, researchers can make a stop sign look like a yield sign by injecting patterns of dots too fine to be seen by human eyes, into an image. The stop signs look the same but they look differently to the AI system.

The same kind of thing has been done with stealthy audio signals embedded in voice commands, where a speech recognition system hears the commands that the attacker wishes to execute, not what the owner says or hears.

Other types of attacks include methods that steal secrets about the operation of the AI system or the proprietary data that was used to train the system. In another attack, adversaries poisoned the AI systems by injecting erroneous or biased training data into the system.

So to conclude I will highlight five recommendations for you to consider.

One, we need to invest in core R&D [research and development] on harnessing AI to push ahead on the frontier of defense and to better understand offenses that will be on the horizon. This includes red-teaming. This is imagining what adversaries can do and developing strategies to protect our systems in advance.

We need to incentivize the creation of cross-sector partnerships to promote sharing and collaboration around data, experiences, best practices, and research.

Three, we need to ensure that AI systems are designed with awareness and best understandings about handling these special adversarial attacks.

Four, we need to develop training programs to educate cybersecurity and AI workforce teams on the special security vulnerabilities of AI systems and their components.

And finally, we need to ensure that DOD and Federal AI agency systems are developed in a secure manner across the lifecycle of these projects to protect the data, protect the executables, and the programs.

Thank you again for your leadership on this important topic and for giving me the opportunity to testify today. I look forward to hearing your questions.

[The prepared statement of Dr. Horvitz follows:]

PREPARED STATEMENT BY ERIC HORVITZ

Chairman Manchin, Ranking Member Rounds, and Members of the Subcommittee, thank you for the opportunity to share insights about the impact of artificial intelligence (AI) on cybersecurity. I applaud the Subcommittee for its foresight and leadership in holding a hearing on this critically important topic. Microsoft is committed to working collaboratively with you to help ensure new advances in AI and cybersecurity benefit our country and society more broadly.

My perspective is grounded in my experiences working across industry, academia, scientific agencies, and government. As Microsoft's Chief Scientific Officer, I provide leadership and perspectives on scientific advances and trends at the frontiers of our understandings, and on issues and opportunities rising at the intersection of technology, people, and society. I have been pursuing and managing research on principles and applications of AI technologies for several decades, starting with my doctoral work at Stanford University. I served as a Commissioner on the National Security Commission on AI (NSCAI), was president of the Association for the Advancement of Artificial Intelligence (AAAI), chaired the Section on Computing, Information, and Communication of the American Association for the Advancement of Science (AAAS). I am a member of the National Academy of Engineering (NAE) and the American Academy of Arts and Sciences. I currently serve on the President's Council of Advisors on Science and Technology (PCAST) and on the Computer Science and Telecommunications Board (CSTB) of the National Academies of Sciences.

I will cover in my testimony four key areas of attention at the intersection of AI and cybersecurity that warrant deeper understanding and thoughtful action:

- Advancing cybersecurity with AI
- Uses of AI to power cyberattacks
- Vulnerabilities of AI systems to attacks
- Uses of AI in malign information operations

Before covering these topics, I will provide brief updates on the cybersecurity landscape and on recent progress in AI. I'll conclude my testimony with reflections about directions.

1. Cybersecurity's changing landscape

Attacks on computing systems and infrastructure continue to grow in complexity, speed, frequency, and scale. We have seen new attack techniques and the exploitation of new attack surfaces aimed at disrupting critical infrastructure and access-

ing confidential data.¹ In 2021 alone, the Microsoft 365 Defender suite, supported by AI techniques, blocked more than 9.6 billion malware threats, 35.7 billion phishing and malicious emails, and 25.6 billion attempts to hijack customer accounts targeting both enterprise and consumer devices.^{2 3} Multiple independent reports have characterized the nature and status of different forms of cyberattack.⁴ As detailed in Microsoft’s recent Digital Defense Report,⁵ cyber criminals and nation-state actors continue to adapt their techniques to exploit new vulnerabilities and counter cyber defenses.

To help mitigate these concerning trends, the U.S. Government has taken significant steps forward to secure our cyber ecosystem. Congress enacted several recommendations that came out of the Cyberspace Solarium Commission, such as creating the Office of the National Cyber Director and enacting cyber incident reporting legislation. Almost a year ago, the Administration issued Executive Order (E.O.) 14028, Improving the Nation’s Cybersecurity, which directs agencies to develop and implement a variety of initiatives to raise the bar on cybersecurity across areas, such as supply chain security, and requiring agencies to adopt a zero-trust model. Microsoft has worked diligently to meet deadlines specified in the E.O. on cybersecurity and we support these efforts to encourage a cohesive response to evolving cyber threats.

We expect to face continuing efforts by creative and tireless state and non-state actors who will attempt to attack computing systems with the latest available technologies. We need to continue to work proactively and reactively to address threats and to note changes in systems, technologies, and patterns of usage. On the latter, cybersecurity challenges have been exacerbated by the increasing fluidity between online work and personal activities as daily routines have become more intertwined.⁶ The large-scale shift to a paradigm of hybrid work coming with the COVID-19 pandemic has moved workers further away from traditional, controlled environments. Cybersecurity solutions must enable people to work productively and securely across various devices from a variety of non-traditional locations.

2. *Advancements in Artificial Intelligence*

Artificial intelligence is an area of computer science focused on developing principles and mechanisms to solve tasks that are typically associated with human cognition, such as perception, reasoning, language, and learning. Numerous milestones have been achieved in AI theory and applications over the 67 years since the phrase “artificial intelligence” was first used in a funding proposal that laid out a surprisingly modern vision for the field.⁷

Particularly stunning progress has been made over the last decade, spanning advances in machine vision (e.g., object recognition), natural language understanding, speech recognition, automated diagnosis, reasoning, robotics, and machine learning—procedures for learning from data. Many impressive gains across sub disciplines of AI are attributed to a machine learning methodology named deep neural networks (DNNs). DNNs have delivered unprecedented accuracy when fueled by large amounts of data and computational resources.

Breakthroughs in accuracy include performances that exceed human baselines for a number of specific benchmarks, including sets of skills across vision and language subtasks. While AI scientists remain mystified by the powers of human intellect, the rate of progress has surprised even seasoned experts.

Jumps in core AI capabilities have led to impressive demonstrations and real-world applications, including systems designed to advise decision makers, generate textual and visual content, and to provide new forms of automation, such as the control of autonomous and semi-autonomous vehicles.

AI technologies can be harnessed to inject new efficiencies and efficacies into existing work flows and processes. The methods also can be used to introduce fundamentally new approaches to standing challenges. When deployed in a responsible and insightful manner, AI technologies can enhance the quality of the lives of our

¹ <https://www.microsoft.com/security/blog/2021/12/15/the-final-report-on-nobeliums-unprecedented-nation-state-attack/>

² <https://news.microsoft.com/wp-content/uploads/prod/sites/626/2022/02/Cyber-Signals-E-1-218.pdf>, page 3

³ <https://www.microsoft.com/en-us/research/group/m365-defender-research/>

⁴ 2018-Webroot-Threat-Report_US-ONLINE.pdf

⁵ Microsoft Digital Defense Report, October 2021

⁶ <https://www.microsoft.com/security/blog/2021/05/12/securing-a-new-world-of-hybrid-work-what-to-know-and-what-to-do/>

⁷ J. McCarthy, J., M.L. Minsky, N. Rochester, N., C.E. Shannon, C.E. A Proposal for the Dartmouth Summer Project on Artificial Intelligence, Dartmouth University, May 1955. <http://www.formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

citizenry and add to the vibrancy of our Nation and world. For example, AI technologies show great promise in enhancing healthcare via providing physicians with assistance on diagnostic challenges, guidance on optimizing therapies, and inferences about the structure and interaction of proteins that lead to new medications.

AI advances have important implications for the Department of Defense, our intelligence community, and our national security more broadly. Like any technology, the rising capabilities of AI are available to friends and foes alike. Thus, in addition to harnessing AI for making valuable contributions to people and society, we must continue to work to understand and address the possibilities that the technologies can be used by malevolent actors and adversaries to disrupt, interfere, and destroy. AI has important implications for cybersecurity as the technologies can provide both new powers for defending against cyberattacks and new capabilities to adversaries.

3. *Advancing Cybersecurity with AI*

The value of harnessing AI in cybersecurity applications is becoming increasingly clear. Amongst many capabilities, AI technologies can provide automated interpretation of signals generated during attacks, effective threat incident prioritization, and adaptive responses to address the speed and scale of adversarial actions. The methods show great promise for swiftly analyzing and correlating patterns across billions of data points to track down a wide variety of cyber threats of the order of seconds. Additionally, AI can continually learn and adapt to new attack patterns—drawing insights from past observations to detect similar attacks that occur in the future.

3.1 *Assisting and Complementing Workforce*

The power of automation and large-scale detection, prioritization, and response made possible by AI technologies can not only relieve the burden on cybersecurity professionals but also help with the growing workforce gap. On the challenges to current cyber workforce: the U.S. Bureau of Labor Statistics estimates cybersecurity job opportunities will grow 33 percent from 2020 to 2030—more than six times the national average.⁸ However, the number of people entering the field is not keeping pace. There is a global shortage of 2.72 million cybersecurity professionals, according to the 2021 (ISC)2 Cybersecurity Workforce Study released in October 2021.⁹

Organizations that prioritize cybersecurity run security operations teams 24/7. Still, there are often far more alerts to analyze than there are analysts to triage them, resulting in missed alerts that evolve into breaches. Trend Micro released a survey in May 2021 of security operations center decision makers that showed that 51 percent feel their team is overwhelmed with the overall volume of alerts, 55 percent are not confident in their ability to efficiently prioritize and respond to alerts, and that 27 percent of their time is spent dealing with false positives.¹⁰

AI technologies enable defenders to effectively scale their protection capabilities, orchestrate and automate time-consuming, repetitive, and complicated response actions. These methods can enable cybersecurity teams to handle large volumes of classical threats in more relevant time frames with less human intervention and better results. Such support with scaling on the essentials can free cybersecurity professionals to focus and prioritize on those attacks that require specialized expertise, critical thinking, and creative problem solving. However, additional attention should also be given to general cybersecurity training, security awareness, secure development lifecycle practices, and simulated training modules, including using AI to run intelligent and personalized simulations.

3.2 *AI at Multiple Stages of Security*

Today, AI methods are being harnessed across all stages of security including prevention, detection, investigation and remediation, discovery and classification, threat intelligence, and security training and simulations. I will discuss each of these applications in turn.

Prevention. Prevention encompasses efforts to reduce the vulnerability of software to attack, including user identities and data, computing system endpoints, and cloud applications. AI methods are currently used in commercially available technologies to detect and block both known and previously unknown threats before they can cause harm. In 2021, AV-Test Institute observed over 125 million new malware

⁸ <https://www.bls.gov/ooh/computer-and-information-technology/information-security-analysts.htm>

⁹ <https://www.isc2.org/News-and-Events/Press-Room/Posts/2021/10/26/ISC2-Cybersecurity-Workforce-Study-Sheds-New-Light-on-Global-Talent-Demand>

¹⁰ <https://newsroom.trendmicro.com/2021-05-25-70-Of-SOC-Teams-Emotionally-Overwhelmed-By-Security-Alert-Volume>

threats.¹¹ The ability of machine learning techniques to generalize from past patterns to catch new malware variants is key to being able to protect users at scale.

As an example, last year Microsoft 365 Defender successfully blocked a file that would later be confirmed as a variant of the GoldMax malware. Defender had never seen the new variant of GoldMax. The malware was caught and blocked leveraging the power of an AI pattern recognizer working together with a technology known as “fuzzy hashing”—a means for taking a fingerprint of malware.¹² It is important to note that GoldMax is malware that persists on networks, feigning to be a “scheduled task” by impersonating the activities of systems management software. Such hiding out as a scheduled task is part of the tools, tactics, and procedures of NOBELIUM, the Russian state actor behind the attacks against SolarWinds in December 2020 and which the United States Government and others have identified as being part of Russia’s foreign intelligence service known as the SVR.

In other work, we have found that AI methods can improve our ability to detect sophisticated phishing attacks. Phishing attacks center on social engineering, where an attacker creates a fake web page or sends a fraudulent message designed to trick a person into revealing sensitive data to the attacker or to deploy malicious software on the victim’s device, such as ransom ware. To help protect people from harmful URLs, AI pattern recognizers have been deployed in browsers and other applications as part of their security services. AI methods can improve detection while lowering false positive rates, which can frustrate end users.¹³

Detection. Detection involves identifying and alerting suspicious behaviors as they happen. The goal is to quickly respond to attacks, including identifying the scale and scope of an attack, closing the attacker’s entry, and remediating footholds that the attacker may have established. The key challenge with detecting suspicious activity is to find the right balance between providing enough coverage via seeking high rates of accurate security alerts versus false alarms. AI methods are being leveraged in detection to (1) triage attention to alerts about potential attacks, (2) identify multiple attempts at breaches over time that are part of larger and lengthier attack campaigns, (3) detecting fingerprints of the activities of malware as it operates within a computer or on a network, (4) identifying the flow of malware through an organization,¹⁴ and (5) guiding automated approaches to mitigation when a response needs to be fast to stop an attack from propagating. For example, an automated system can shut down network connectivity and contain a device if a sequence of alerts is detected that is known to be associated with ransomware activity like the way a bank might decline a credit card transaction that appears fraudulent.

There are several technologies available today to help detect attacks. I will use Microsoft 365 Defender capabilities as an example. A set of neural network models are used to detect a potential attack underway by fusing multiple signals about activities within a computing system, including processes being started and stopped, files being changed and renamed, and suspicious network communication.¹⁵ ¹⁶ In addition, probabilistic algorithms are used to detect high likelihood of “lateral movement” on a network.¹⁷ Lateral movement refers to malware, such as ransomware, moving from machine to machine as it infects an organization. The goal is to detect signals of concerning patterns of spread and to shut down the infection by isolating potentially infected machines and alerting security experts to investigate. As numerous legitimate operations can appear like lateral movement of malware, simplistic approaches can have high false-positive rates. AI systems can help to raise the rate of capture and block these spreading infections, while reducing false positives.¹⁸

As a recent example, in March 2022, Microsoft leveraged its AI models to identify an attack attributed to a Russian actor that Microsoft tracks as Iridium, also referred to as Sandworm. The United States Government has attributed Iridium activity to a group allegedly based at GRU Unit 74455 of the Main Directorate of the General Staff of the Armed Forces of the Russian Federation. The actor deployed

¹¹ <https://www.av-test.org/en/statistics/malware/>

¹² <https://www.microsoft.com/security/blog/2021/07/27/combining-through-the-fuzz-using-fuzzy-hashing-and-deep-learning-to-counter-malware-detection-evasion-techniques>

¹³ <https://www.microsoft.com/en-us/research/publication/urltran-improving-phishing-url-detection-using-transformers/>

¹⁴ <https://dl.acm.org/doi/10.1145/3471621.3471858>

¹⁵ <https://www.microsoft.com/security/blog/2020/07/23/seeing-the-big-picture-deep-learning-based-fusion-of-behavior-signals-for-threat-detection/>

¹⁶ <https://www.microsoft.com/security/blog/2020/08/27/stopping-active-directory-attacks-and-other-post-exploitation-behavior-with-amsi-and-machine-learning/>

¹⁷ <https://www.microsoft.com/security/blog/2019/12/18/data-science-for-cybersecurity-a-probabilistic-time-series-model-for-detecting-rdp-inbound-brute-force-attacks/>

¹⁸ <https://www.microsoft.com/security/blog/2020/06/10/the-science-behind-microsoft-threat-protection-attack-modeling-for-finding-and-stopping-evasive-ransomware/>

wiper malware at a Ukrainian shipping company based in Lviv. Wiper malware erases data and programs on the computers that it infects. The first documented encounter of this malware was on a system running Microsoft Defender with Cloud Protection enabled. The ensemble of machine learning models in Defender, combined with signals across client and cloud, allowed Microsoft to block this malware at first sight.

Investigation and remediation. Investigation and remediation are methods used following a breach to provide customers with a holistic understanding of the security incident, including the extent of the breach, which devices and data were impacted, how the attack propagated through the customer environment, and to seek attribution for the threat.¹⁹ Gathering and doing synthesis from telemetry sources is tedious. Efforts to date include multiple tools to collect telemetry from within and across organizations. The use of AI for investigation and remediation is a promising and open area of research.^{20 21}

Threat intelligence. Threat intelligence enables security researchers to stay on top of the current threat landscape by tracking active malicious actors, at times deliberately engaging with them and studying their behavior. Today, Microsoft actively tracks 40+ active nation-state actors and 140+ threat groups across 20 countries.²² AI methods help to identify and tag entities from multiple feeds and intelligence sharing across agencies. AI models show promise with their ability to learn and make inferences about high-level relationships and interactions by identifying similarities across different campaigns for enhancing threat attribution.^{24 25}

Recommendations: Advance development and application of AI methods to defend against cyberattacks.

- Follow best practices in cybersecurity hygiene, including implementation of core protections such as multifactor authentication. Bolster security teams, regularly test backups and update patches, test incident response plans, and limit internet access to networks that do not require internet connectivity.
- Invest in training and education to strengthen the U.S. workforce in cybersecurity, including education and training programs on cybersecurity for both traditional and AI systems.
- Invest in R&D on harnessing machine learning, reasoning, and automation to detect, respond, and protect every step of the cyberattack kill chain.
- Incentivize the creation of cross-sector partnerships to catalyze sharing and collaboration around cybersecurity experiences, datasets, best practices, and research.
- Develop cybersecurity-specific benchmarks and leaderboards specific to validate research and accelerate learnings.

4. AI-powered cyberattacks

While AI is improving our ability to detect cybersecurity threats, organizations and consumers will face new challenges as cybersecurity attacks increase in sophistication. To date, adversaries have commonly employed software tools in a manual manner to reach their objectives. They have been successful in exfiltrating sensitive data about American citizens, interfering with elections, and distributing propaganda on social media without the sophisticated use of AI technologies.^{26 27 28} While there is scarce information to date on the active use of AI in cyberattacks, it is widely accepted that AI technologies can be used to scale cyberattacks via various forms of probing and automation. Multiple research and gaming efforts within

¹⁹ <https://www.microsoft.com/security/blog/2021/12/02/structured-threat-hunting-one-way-microsoft-threat-experts-prioritizes-customer-defense/>

²⁰ <https://www.microsoft.com/security/blog/2020/07/09/inside-microsoft-threat-protection-correlating-and-consolidating-attacks-into-incidents/>

²¹ <https://www.microsoft.com/security/blog/2020/07/29/inside-microsoft-threat-protection-solving-cross-domain-security-incidents-through-the-power-of-correlation-analytics/>

²² <https://www.microsoft.com/security/blog/2022/02/03/cyber-signals-defending-against-cyber-threats-with-the-latest-research-insights-and-trends/>

²³ <https://www.microsoft.com/security/blog/2021/05/12/securing-a-new-world-of-hybrid-work-what-to-know-and-what-to-do/>

²⁴ <https://www.microsoft.com/security/blog/2021/04/01/automating-threat-actor-tracking-understanding-attacker-behavior-for-intelligence-and-contextual-alerting/>

²⁵ <https://dl.acm.org/doi/pdf/10.1145/3448016.3452745>

²⁶ Cybersecurity Incidents (opm.gov)

²⁷ Russian Interference in 2016 U.S. Elections-FBI

²⁸ Characterizing networks of propaganda on twitter: a case study

cybersecurity communities have demonstrated the power using AI methods to attack computing systems. This area of work is referred to as^{29 30} offensive AI.

4.1 Approaches to offensive AI

Offensive AI methods will likely be taken up as tools of the trade for powering and scaling cyberattacks. We must prepare ourselves for adversaries who will exploit AI methods to increase the coverage of attacks, the speed of attacks, and the likelihood of successful outcomes. We expect that uses of AI in cyberattacks will start with sophisticated actors but will rapidly expand to the broader ecosystem via increasing levels of cooperation and commercialization of their tools.³¹

Basic automation. Just as defenders use AI to automate their processes, so too can adversaries introduce efficiencies and efficacies for their own benefit. Automating attacks using basic pre-programmed logic is not new in cybersecurity. Many malware and ransomware variants over the last five years have used relatively simple sets of logical rules to recognize and adapt to operating environments. For example, it appears that attacking software has checked time zones to adapt to local working hours and customized behavior in a variety of ways to avoid detection or take tailored actions to adapt to the target computing environment.^{32 33} On another front, automated bots have begun to proliferate on social media platforms.³⁴ These are all rudimentary forms of AI that encode and harness an attacker's expert knowledge. However, substantial improvements in AI technology make plausible malicious software that is much more adaptive, stealthy, and intrusive.³⁵

Authentication-based attacks. AI methods can be employed in authentication-based attacks, where, for example, recently developed AI methods can be used to generate synthetic voiceprints to gain access through an authentication system. Compelling demonstrations of voice impersonations to fool an authentication system were presented during the Capture the Flag (CTF) cybersecurity competition at the 2018 DEF CON meeting.³⁶

AI-powered social engineering. Human perception and psychology are weak links in cyber-defense. AI can be used to exploit this persistent vulnerability. We have seen the rise of uses of AI for social engineering, aiming the power of machine learning at influencing the actions of people to perform tasks that are not in their interest. As an example, AI methods can be used to generate ultra-personalized phishing attacks capable of fooling even the most security conscious users. A striking 2018 study demonstrated how AI methods could be used to significantly raise the probability that end users would click on malevolent links in social media posts. The AI system learned from publicly available data including online profiles, connections, content of posts, and online activity of targeted individuals. Machine-learning was used to optimize the timing and content of messages with a goal of maximizing click through rates—with significant results.³⁷ A 2021 study demonstrated that the language of emails could be crafted automatically with large-scale neural language models and that the AI-generated messages were more successful than the human-written messages by a significant margin.³⁸ In a related direction, Microsoft has tracked groups that use AI to craft convincing but fake social media profiles as lures.

4.2 AI-powered cyberattacks on the frontier

The need to prepare for more sophisticated offensive AI was highlighted in presentations at a National Academies of Sciences workshop on offensive AI that I organized in 2019. The workshop, sponsored by the Office of the Director of National

²⁹ <https://arxiv.org/pdf/2106.15764.pdf>

³⁰ B. Buchanan, J. Bansemmer, D. Cary, et al., Automating Cyber Attacks: Hype and Reality, Center for Security and Emerging Technology, November 2020. <https://cset.georgetown.edu/wp-content/uploads/CSET-Automating-Cyber-Attacks.pdf>

³¹ How cyberattacks are changing according to new Microsoft Digital Defense Report

³² Intelligence, FireEye Threat. "HAMMERTOSS: Stealthy tactics define a Russian cyber threat group." FireEye, Milpitas, CA (2015).

³³ Virtualization/Sandbox Evasion, Technique T1497 - Enterprise / MITRE ATT&CK®

³⁴ <https://www.jmir.org/2021/5/e26933/>

³⁵ See for example, see documentation of Deep Exploit, tools and demonstration showing the use of reinforcement learning to drive cyberattacks: https://github.com/13o-bbr-bbq/machine_learning_security/tree/master/DeepExploit ³⁶ <https://www.defcon.org/>

³⁶ <https://www.defcon.org/>

³⁷ J. Seymour and P. Tully, Generative Models for Spear Phishing Posts on Social Media, 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 2017. <https://arxiv.org/abs/1802.05196>

³⁸ <https://www.wired.com/story/ai-phishing-emails/amp>

Intelligence, led to a report available from the Academies.³⁹ The report includes discussion of the applications of AI methods across the cyber kill-chain, including the use of AI methods in social engineering, discovery of vulnerabilities, exploiting development and targeting, and malware adaptation, as well as in methods and tools that can be used to target vulnerabilities in AI-enabled systems, such as autonomous systems and controls used in civilian and military applications.

The cybersecurity research community has demonstrated the power of AI and other sophisticated computational methods in cyberattacks. Adversaries can harness AI to efficiently guess passwords, to attack industrial control systems without raising suspicions, and to create malware that evades detection or prevents inspection⁴⁰ AI-enabled bots can also automate network attacks and make it difficult to extinguish the attacker's command and control channels.⁴⁶ In another direction, a competitor demonstrated at a DARPA Cyber Grand Challenge exercise in 2016⁴⁷ how machine learning could be used to learn how to generate "chaff" traffic, decoy patterns of online activity that resemble the distribution of events seen in real attacks for distraction and cover-up of actual attack strategies.⁴⁸

It is safe to assume that AI will improve the success, impact, and scope of the full breadth of threats present today. AI will also introduce new challenges, including special cyber vulnerabilities introduced with general uses of AI components and applications, which create new apertures for adversaries to exploit.

RECOMMENDATIONS: PREPARE FOR MALICIOUS USES OF AI TO PERFORM CYBERATTACKS

- Raise DOD and other Federal agency awareness of the threat of AI-powered cyberattacks and directions with defenses against them, including detecting and thwarting new forms of automation and scaling.
- DOD should deeply engage with the cybersecurity community, participate in R&D and competitions on AI-enhanced cyberattacks and continue to learn from frontier advances, findings, and proposed mitigations.
- Increase R&D funding for exploring challenges and opportunities at the convergence of AI and cybersecurity. Consider the establishment of federally funded R&D centers of excellence in cybersecurity. Execute on the NSCAI recommendation to invest in DARPA to facilitate greater research on AI-enabled cyber defenses.⁴⁹
- Formalize and make more efficient cross-sector networks for sharing updates on evolving technologies, data, attack vectors, and attacks.

5. *Special vulnerabilities of AI systems*

The power and growing reliance on AI generates a perfect storm for a new type of cyber-vulnerability: *attacks targeted directly at AI systems and components*. With attention focused on developing and integrating AI capabilities into applications and workflows, the security of AI systems themselves is often overlooked. However, adversaries see the rise of new AI attack surfaces growing in diversity and ubiquity and will no doubt be pursuing vulnerabilities. Attacks on AI systems can come in

³⁹ Implications of Artificial Intelligence for Cybersecurity: A Workshop, National Academy of Sciences, 2019. <https://www.nationalacademies.org/our-work/implications-of-artificial-intelligence-for-cybersecurity-a-workshop>

⁴⁰ Hey, My Malware Knows Physics! Attacking PLCs with Physical Model Aware Rootkit—NDSS Symposium (ndss-symposium.org)

⁴¹ B. Hitaj, P. Gasti, G. Ateniese, F. Perez-Cruz, PassGAN: A Deep Learning Approach for Password Guessing, NeurIPS 2018 Workshop on Security in Machine Learning (SecML'18), December 2018. <https://github.com/secml2018/secml2018.github.io/raw/master/PASSGAN-SECML2018.pdf>

⁴² S. Datta, DeepObfusCode: Source Code Obfuscation through Sequence-to-Sequence Networks In: Arai, K. (eds) Intelligent Computing. Lecture Notes in Networks and Systems, vol 284. Springer, Cham. <https://doi.org/10.1007/978-3-030-80126-7-45>, July 2021.

⁴³ J. Li, L. Zhou, H. Li, L. Yan and H. Zhu, "Dynamic Traffic Feature Camouflaging via Generative Adversarial Networks," 2019 IEEE Conference on Communications and Network Security (CNS), 2019, pp. 268–276, doi: 10.1109/CNS.2019.8802772. <https://ieeexplore.ieee.org/abstract/document/8802772>

⁴⁴ C. Novo, R. Morla, Flow-Based Detection and Proxy-Based Evasion of Encrypted Malware C2 Traffic, Proceedings of the 13th ACM Workshop on Artificial Intelligence and Security 2020, <https://doi.org/10.1145/3411508.3421379>.

⁴⁵ D. Han et al., "Evaluating and Improving Adversarial Robustness of Machine Learning-Based Network Intrusion Detectors," in IEEE Journal on Selected Areas in Communications, vol. 39, no. 8, pp. 2632–2647, Aug. 2021, <https://ieeexplore.ieee.org/abstract/document/9448103>

⁴⁶ A botnet-based command and control approach relying on swarm intelligence-ScienceDirect <https://www.darpa.mil/program/cyber-grand-challenge>

⁴⁸ 48 R. Rivest, Chaffing and Winnowing: Confidentiality Without Encryption." CryptoBytes, 4(1):12–17, <https://pdfs.semanticscholar.org/aaf3/7e0afa43f5b6168074dae2bc0e695a9d1d1b.pdf>

⁴⁹ <https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf>. page 279.

the form of *traditional vulnerabilities, via basic manipulations and probes*, and via a new, troubling category: *adversarial AI*.

5.1 Attacks on AI Supply Chains

AI systems can be attacked via targeting traditional security weaknesses and software flaws, including attacks on the supply chain of AI systems, where malevolent actors gain access and manipulate insecure AI code and data. As an example, in 2021, a popular software platform used to build neural networks was found to have 201 traditional security vulnerabilities, such as memory corruption and code execution.⁵⁰ Researchers have demonstrated how adversaries could use existing cyberattack toolkits to attack core infrastructure of the software running AI systems.⁵¹ Multiple components of AI systems in the supply chain of AI systems can be modified or corrupted via traditional cyberattacks. As an example, data sets used to train AI systems are rarely under version control in the same way that source code is. Researchers from NYU found that most AI frameworks downloaded from a popular algorithm repository do not check the integrity of AI models, in contrast to the standards of practice with traditional software, where cryptographic verification of executables/libraries has been standard practice for well over a decade.⁵²

5.2 Adversarial AI

Adversarial AI or adversarial machine learning methods harness more sophisticated AI techniques to attack AI systems. Several classes of adversarial AI have been identified, including *adversarial examples*, the use of basic policies or more sophisticated machine learning methods to fool AI systems with inputs that cause the systems to fail to function properly. A second type of attack is called *data poisoning*, where data used to train AI systems are “poisoned” with streams of data that inject erroneous or biased training data into data sets, changing the behavior or degrading the performance of AI systems.⁵³ A third type of attack, called *model stealing*, seeks to learn details about the underlying AI model used in an AI system.⁵⁴ A fourth category of attack, called *model inversion*, seeks to reconstruct the underlying private data that is used to train the target system.⁵⁵

With adversarial examples, basic manipulations or more sophisticated application of AI methods are used to generate inputs that are custom-tailored to cause failures in targeted AI systems. Goals of these attacks include disruptive failures of automated message classifiers, perceptions of machine vision systems, and recognitions of the words in utterances by speech recognition systems.

As an example of basic manipulations of inputs, a group, alleged to be within the Chinese Government, attempted to amplify propaganda on Uyghurs by bypassing Twitter’s anti-spam algorithm via appending random characters at the end of tweets.⁵⁶ The approach was viewed as an attempt to mislead the algorithm into thinking each tweet was unique and legitimate. In another example, researchers from Skylight appended benign code from a gaming database to Wannacry ransomware to cause the machine-learning-based antivirus filter to classify the modified ransomware as benign.⁵⁷ In related work on the fragility of AI systems, researchers showed that simply rotating a scan of a skin lesion confuses a computer recognition system to classify the image as malignant.⁵⁸

In uses of AI to generate adversarial examples, researchers have demonstrated stunning examples of failures. In one approach, adversarial methods are used to inject patterns of pixels into images to change what an AI system sees. While the changes with AI inferences are dramatic, *the changes to the original images are not detectable by humans*. Sample demonstrations include the modification of a photo

⁵⁰ <https://www.cvedetails.com/product/53738/Google-Tensorflow.html>

⁵¹ Xiao, Qixue, et al. “Security risks in deep learning implementations.” 2018 IEEE Security and privacy workshops (SPW). IEEE, 2018.

⁵² Gu, Tianyu, Brendan Dolan-Gavitt, and Siddharth Garg. “Badnets: Identifying vulnerabilities in the machine learning model supply chain.” arXiv preprint arXiv:1708.06733 (2017).

⁵³ Jagielski, Matthew, et al. “Manipulating machine learning: Poisoning attacks and countermeasures for regression learning.” 2018 IEEE Symposium on Security and Privacy (SP). IEEE, 2018.

⁵⁴ Yu, Honggang, et al. “CloudLeak: Large-Scale Deep Learning Models Stealing Through Adversarial Examples.” NDSS. 2020.

⁵⁵ Ziqi Yang, Ee-Chien Chang, Zhenkai Liang, Adversarial Neural Network Inversion via Auxiliary Knowledge Alignment, 2019.

⁵⁶ <https://www.nytimes.com/interactive/2021/06/22/technology/xinjiang-uyghurs-china-propaganda.html>

⁵⁷ <https://skylightcyber.com/2019/07/18/cylance-i-kill-you/>

⁵⁸ Finlayson, Samuel G., et al. “Adversarial attacks on medical machine learning.” Science 363.6433 (2019): 1287–1289.

of a panda leading an AI system to misclassify the panda as a gibbon and changes to a stop sign to misclassify it as a yield sign.⁵⁹ ⁶⁰ Similar demonstrations have been done in the realm of speech recognition, with the injection of hidden acoustical patterns in speech that changes what a listening system hears.⁶¹ Attacks leading to such misclassifications and malfunctions can be extremely costly, particularly in high-stakes domains like defense, transportation, healthcare, and industrial processes.

Challenges of adversarial AI and a set of recommendations are called out in the final report of the National Security Commission on AI (NSCAI).⁶² I chaired the lines of effort on directions with developing and fielding trustworthy, responsible, and ethical AI applications, leading to chapters 7 and 8 of the report and the appendix on NSCAI's recommendations on key considerations for fielding AI systems that align with democratic values, civil liberties, and human rights.⁶³ ⁶⁴ ⁶⁵ Chapter 7 of the report covers rising concerns with adversarial AI, including the assessment that, "*The threat is not hypothetical: adversarial attacks are happening and already impacting commercial ML systems.*" In support of this statement, over the last five years, the Microsoft cybersecurity team has seen an uptick in adversarial AI attacks.⁶⁶ I believe the trend will continue.

5.3 Efforts to Mitigate Adversarial AI

Pursuit of resistant systems. Computer science R&D has been underway on methods for making AI systems more resistant to adversarial machine learning attacks. One area of work centers on raising the level of robustness of systems to attacks with adversarial inputs as described above.⁶⁷ ⁶⁸ Approaches include special training procedures to include adversarial examples, validation of inputs to identify specific properties that can reveal signs of an attack and making changes to the overall approach to building models, and modifying the objective functions used in optimization procedures used to create the models so that more robust models are created. While the latter techniques and research directions behind them are promising, the challenges of adversarial examples persist, per the large space of inputs to machine learning procedures. Thus, it is important to continue to invest in R&D on adversarial AI, to perform ongoing studies with red-teaming exercises, and to remain vigilant.

5.4 Tracking, Awareness, and Resources

Front-line awareness. Despite the opportunities that adversarial AI methods will provide to state and non-state actors for manipulating and disrupting critical AI systems and rising evidence of real-world attacks with adversarial AI, the idea of protecting AI systems from these attacks has been largely an afterthought. There is an urgency to be aware and to be ready to respond to adversarial AI threats, especially those used in critical areas such as defense. A Microsoft survey of 28 organizations in 2020 showed, despite the rise in attacks on AI systems, companies are still unaware of these kinds of intentional failures to AI systems and are massively underinvested in tools and processes to secure AI systems.⁶⁷ Ryan Fedasiuk, a noted researcher at Georgetown's Center for Security of Emerging Technology specializing in China's AI operations, notes that Chinese military officers have explicitly called out that the United States defenses are susceptible to data poisoning, and even so

⁵⁹I.J. Goodfellow, J. Shlens, C. Szegedy, Explaining and Harnessing Adversarial Examples, ICLR 2015. <https://arxiv.org/pdf/1412.6572.pdf>

⁶⁰N. Papernot, P. McDaniel, I. Goodfellow, et al., Practical Black-Box Attacks against Machine Learning, ASIA CCS '17, April 2017. <https://dl.acm.org/doi/pdf/10.1145/3052973.3053009>

⁶¹M. Alzantot, B. Balaji, M. Srivastava, Did you hear that? Adversarial Examples Against Automatic Speech Recognition, Conference on Neural Information Processing Systems, December 2017. <https://arxiv.org/pdf/1801.00554.pdf>

⁶²<https://www.nscai.gov/>
⁶³"Upholding Democratic Values: Privacy, Civil Liberties, and Civil Rights in Uses of AI for National Security," Chapter 8, Report of the National Security Commission on AI, March 2021. <https://reports.nscai.gov/final-report/chapter-8/>

⁶⁴"Establishing Justified Confidence in AI Systems," Chapter 8, Report of the National Security Commission on AI, March 2021. <https://reports.nscai.gov/final-report/chapter-7/>

⁶⁵E. Horvitz J. Young, R.G. Elluru, C. Howell, Key Considerations for the Responsible Development and Fielding of Artificial Intelligence, National Security Commission on AI, April 2021. <https://arxiv.org/ftp/arxiv/papers/2108/2108.12289.pdf>

⁶⁶Kumar, Ram Shankar Siva, et al. Adversarial machine learning-industry perspectives. 2020 IEEE Security and Privacy Workshops (SPW). IEEE, 2020.

⁶⁷<https://cacm.acm.org/magazines/2018/7/229030-making-machine-learning-robust-against-adversarial-inputs/fulltext>

⁶⁸A. Madry, A. Makelov, L. Schmidt, et al. Towards deep learning models resistant to adversarial attacks, ICLR 2018. <https://arxiv.org/pdf/1706.06083.pdf>

far as calling data integrity as “the Achilles’ heel” of the U.S. joint all-domain command and control strategy.⁶⁹

Resources and Engagement. Microsoft, along with MITRE and 16 other organizations created the Adversarial ML Threat Matrix to catalog threats to AI systems.⁷⁰ The content includes documentation of case studies where attacks have been made on commercial AI systems. For engineers and policymakers, Microsoft, in collaboration with Berkman Klein Center at Harvard University, released a taxonomy of machine learning failure modes.⁷¹ For security professionals, Microsoft has open-sourced Counterfit, its own tool for assessing the posture of AI systems.⁷² For the broader community of cybersecurity practitioners interested in AI and security, Microsoft hosts the annual Machine Learning Evasion Competition as a venue to exercise their muscle in attacking and securing AI systems.⁷³ Within the Federal Government, the DOD has listed safety and security of AI systems in its core AI principles.⁷⁴ And there is encouraging activity by NIST on an AI Risk Assessment Framework to address multiple dimensions of AI systems, including robustness and security.⁷⁵

RECOMMENDATIONS: RAISE AWARENESS AND ADDRESS VULNERABILITIES OF AI SYSTEMS

- Secure engineering supply chains for Federal AI systems, including use of state-of-the-art integrity checking for data, executables, libraries, and platforms used to construct AI systems; ensure that a security development lifecycle approach is in place for sensitive code and data.
- Require security reviews of AI engineering projects at DOD and other Federal AI agencies.
- Bring AI development and cybersecurity teams together to establish best practices and review programs.
- Raise DOD awareness of challenges of adversarial AI and consider the vulnerabilities of AI systems and components.
- Pursue the use of robust machine learning algorithms to bolster resilience of systems in the face of adversarial examples.
- Develop training programs to raise awareness of cybersecurity and AI engineering workforce on security vulnerabilities of AI systems and components, risk of attacks with adversarial AI methods, and means for reducing risks.
- Invest in R&D on trustworthy, robust, and secure AI systems.

6. AI in Malign Information Operations

Advances in machine learning and graphics have boosted the abilities of state and non-state actors to fabricate and distribute high-fidelity audiovisual content, referred to as synthetic media and deepfakes. AI technologies for generating deepfakes can now fabricate content that is indistinguishable from real-world people, scenes, and events, threatening national security. Advances that could only be found with the walls of computer science laboratories or in demonstrations that surprised attendees at academic AI conferences several years ago are now widely available in tools that create audio and audiovisual content that can be used to drive disinformation campaigns.

6.1 Challenges of Synthetic Media

Advances in the capabilities of generative AI methods to synthesize a variety of signals, including high-fidelity audiovisual imagery, have significance for cybersecurity. When personalized, the use of AI to generate deepfakes can raise the effectiveness of social-engineering operations (discussed above) in persuading end-users to provide adversaries with access to systems and information.

On a larger scale, the generative power of AI methods and synthetic media have important implications for defense and national security. The methods can be used by adversaries to generate believable statements from world leaders and commanders, to fabricate persuasive false-flag operations, and to generate fake news events. A recent demonstration includes the multiple examples of manipulated and more sophisticated deepfakes that have come to the fore over the course of the Rus-

⁶⁹ <https://breakingdefense.com/2021/11/china-invests-in-artificial-intelligence-to-counter-us-joint-warfighting-concept-records/>

⁷⁰ <https://atlas.mitre.org/>

⁷¹ <https://docs.microsoft.com/en-us/security/engineering/failure-modes-in-machine-learning>

⁷² <https://github.com/Azure/counterfit/>

⁷³ <https://mlsec.io/>

⁷⁴ <https://www.defense.gov/News/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>

⁷⁵ <https://www.nist.gov/itl/ai-risk-management-framework>

sian attack on Ukraine. This includes a video of President Volodymyr Zelenskyy appearing to call for surrender.⁷⁶

The proliferation of synthetic media has had another concerning effect: malevolent actors have labeled real events as “fake,” taking advantage of new forms of deniability coming with the loss of credibility in the deepfake era. Video and photo evidence, such as imagery of atrocities, are being called fake. Known as the “liar’s dividend,” the proliferation of synthetic media emboldens people to claim real media as “fake,” and creates plausible deniability for their actions.⁷⁷

We can expect synthetic media and its deployment to continue grow in sophistication over time, including the persuasive interleaving of deepfakes with unfolding events in the world and real-time synthesis of deepfakes. Real-time generations could be employed to create compelling, interactive imposters (e.g., appearing in teleconferences and guided by a human controller) that appear to have natural head pose, facial expressions, and utterances. Looking further out, we may have to face the challenge of synthetic fabrications of people that can engage autonomously in persuasive real-time conversations over audio and visual channels.

6.2 Direction: Digital Content Provenance

A promising approach to countering the threat of synthetic media can be found in a recent advance, named digital content provenance technology. Digital content provenance leverages cryptography and database technologies to certify *the source and history of edits* (the provenance) of any digital media. This can provide “glass-to-glass” certification of content, from the photons hitting the light-sensitive surfaces of cameras to the light emitted from the pixels of displays, for secure workloads. We pursued an early vision and technical methods for enabling end-to-end tamper-proof certification of media provenance in a cross-team effort at Microsoft.^{78 79} The aspirational project was motivated by our assessment that, in the long-term, neither humans nor AI methods would be able to reliably distinguish fact from AI-generated fictions—and that we must prepare with urgency for the expected trajectory of increasingly realistic and persuasive deepfakes.

After taking the vision to reality with technical details and the implementation of prototype technologies for certifying the provenance of audiovisual content, we worked to build and contribute to cross-industry partnerships, including Project Origin, the Content Authenticity Initiative (CAI), and the Coalition for Content Provenance and Authenticity (C2PA), a multistakeholder coalition of industry and civil society organizations.^{80 81 82 83} In January 2022, C2PA released a specification of a standard that enables the interoperability of digital content provenance systems.^{84 85} Commercial production tools are now becoming available in accordance with the C2PA standard that enable authors and broadcasters to assure viewers about the originating source and history of edits to photo and audiovisual media.

The final report of the NSCAI recommends that digital content provenance technologies should be pursued to mitigate the rising challenge of synthetic media. In Congress, the bipartisan *Deepfake Task Force Act* (S. 2559) proposes the establishment of the National Deepfake and Digital Provenance Task Force.⁸⁶ Microsoft and its media provenance collaborators encourage Congress to move forward with standing-up a task force to help identify and address the challenges of synthetic media and we would welcome the opportunity to provide assistance and input into the work.

⁷⁶ See: <https://www.youtube.com/watch?v=X17yrEV5sl4>

⁷⁷ *The Liar’s Dividend: The Impact of Deepfakes and Fake News on Politician Support and Trust in Media* / GVU Center (gatech.edu)

⁷⁸ P. England, H.S. Malvar, E. Horvitz, et al. AMP: Authentication of Media via Provenance, ACM Multimedia Systems 2021. <https://dl.acm.org/doi/abs/10.1145/3458305.3459599>

⁷⁹ E. Horvitz, A promising step forward on disinformation, Microsoft on the Issues, February 2021. <https://blogs.microsoft.com/on-the-issues/2021/02/22/deepfakes-disinformation-c2pa-origin-cai/>

⁸⁰ Project Origin, <https://www.originproject.info/about>

⁸¹ J. Aythora, et al. Multi-stakeholder Media Provenance Management to Counter Synthetic Media Risks in News Publishing, International Broadcasting Convention 2020 (IBC 2020), Amsterdam, NL 2020 <https://www.ibt.org/download?ac=14528>

⁸² Content Authenticity Initiative, <https://contentauthenticity.org/>

⁸³ Coalition for Content Provenance and Authenticity (C2PA), <https://c2pa.org/>

⁸⁴ C2PA Releases Specification of World’s First Industry Standard for Content Provenance, Coalition for Content Provenance and Authenticity, January 26, 2022, <https://c2pa.org/post/release-1-pr/>

⁸⁵ <https://erichorvitz.com/A-Milestone-Reached-Content-Provenance.htm>

⁸⁶ Deepfake Task Force Act, S. 2559, 117th Congress, <https://www.congress.gov/bill/117th-congress/senate-bill/2559/text>

RECOMMENDATIONS: DEFEND AGAINST MALIGN INFORMATION OPERATIONS

- Enact the Deepfake Task Force Act.
- Promote uses of digital media provenance for news and communications in defense and civilian settings.
- Adopt pipelines and standards for certifying digital content provenance of signals, communications, and news at DOD and other Federal agencies, prioritized by risk and disruptiveness of fabricated content.
- Review potential disruptions that malign information campaigns could have on DOD planning, decision making, and coordination based on manipulative uses of sophisticated fabrications of audiovisual and other signals, spanning traditional Signals Intelligence (SIGINT) pipelines, real-time defense communications, and public news and media.
- Invest in R&D on methods aimed at detection, attribution, and disruption of AI-enabled malign information campaigns.

Summary

I have covered in my testimony status, trends, examples, and directions ahead with rising opportunities and challenges at the intersection of AI and cybersecurity. AI technologies will continue to be critically important for enhancing cybersecurity in military and civilian applications. AI methods are already qualitatively changing the game in cyber defense. Technical advances in AI have helped in numerous ways, spanning our core abilities to prevent, detect, and respond to attacks—including attacks that have never been seen before. AI innovations are amplifying and extending the capabilities of security teams across the country.

On the other side, state and non-state actors are beginning to leverage AI in numerous ways. They will draw new powers from fast-paced advances in AI and will continue to add new tools to their armamentarium. We need to double down with our attention and investments on threats and opportunities at the convergence of AI and cybersecurity. Significant investments in workforce training, monitoring, engineering, and core R&D will be needed to understand, develop, and operationalize defenses for the breadth of risks we can expect with AI-powered cyberattacks. The threats include new kinds of attacks, including those aimed squarely at AI systems. The DOD, federal and state agencies, and the Nation need to stay vigilant and stay ahead of malevolent adversaries. This will take more investment and commitment to fundamental research and engineering on AI and cybersecurity, and in building and nurturing our cybersecurity workforce so our teams can be more effective today—and well-prepared for the future.

Thank you for the opportunity to testify. I look forward to answering your questions.

Senator MANCHIN. First of all, thank you all so much.

We are going to do rounds of 7 minutes. Being it is just the three of us, I think we will not—

Senator ROSEN. My favorite subcommittee.

Senator MANCHIN. I know it is. I can tell. I mean, Jackie—

Senator ROSEN. You are talking my language.

Senator MANCHIN. Let me tell you one thing. She is ready to—she might take more than 7. It will be all right with me. But she is ready to go.

Senator ROSEN. I have got all the questions.

Senator MANCHIN. I want to thank all three of you.

I am going to start with simply an overview. We have been hearing an awful lot about artificial intelligence and machine learning. Are they one in the same? That is one thing. You can maybe answer very quickly.

I really want to know, and Mike and I both serve on Armed Services—this is a subcommittee of Armed Services that all three of us serve on—where are we in the pecking order of what is going on in this unbelievable world that you are explaining to us? Are we behind? Are we in the hunt? Are we on the cutting edge? What more can we do besides, we know, investing? But we want to invest in the right places to get the best results.

So is the private sector, are you moving us to a position to where—I will use the whole SpaceX program, what they have been able to do in the private sector for the defense of our country and the amount of money we have saved because of the efficiency of the private sector? Can that be duplicated here, in artificial intelligence and machine learning, better invested in? Because we are contracting, as the Federal Government, for our defense programs, with SpaceX, putting different types of articles that we need in space, as you know.

So with that, we can start, and we will start, Dr. Moore, if you can, and keep them fairly concise, if you can, in your answer, because everyone has an awful lot of interesting questions.

Dr. MOORE. Thank you, Chair Manchin. Yes, I will be concise. Artificial intelligence without machine learning gave us things like Deep Blue, where the American IBM computer Deep Blue beat the Russian chess master Kasparov, Gary Kasparov, back in the 1990s. We were all so happy about that in the AI world.

But these systems did not adapt over time, and so that is why machine learning, in the early 2000s, has come in and made AI much more powerful than it was in the days of Deep Blue.

Senator MANCHIN. So basically it has been integrated into one? It is all one, AI and machine learning is now integrated as one?

Dr. MOORE. That is right. In the old days you could have AI without machine learning. These days you always want AI with machine learning.

Senator MANCHIN. And on the other, real quickly, on the other, where do we rank? Just give me a ranking. You do not have to name countries, but are we behind in the hunt or are we on the cutting edge?

Dr. MOORE. We are ahead. We are losing ground. I am most worried about our structures. Bringing in massive scale, super-human automation means changing organizational structures and change management. That is what I believe companies are really quite good at.

Senator MANCHIN. You all can do it better than we can do it in the government, is what you are saying, and we can contract out in a very secure situation, like we do with some of our defense. Okay.

Dr. MOORE. Perhaps, yes.

Senator MANCHIN. Dr. Lohn?

Dr. LOHN. Thank you, Senator Manchin. I would like to concur with Dr. Moore that AI is like a broader umbrella that has machine learning within it as a component. Now I understand the confusion because those two terms have become almost synonymous because almost all of the AI that we talk about today is machine learning, but in the past there were other techniques that were not machine learning, so right now they are basically the same thing. And it may be that machine learning will not be the same as AI in the future, but right now they are basically the same thing, and now machine learning is a small subset of AI.

With respect to are we ahead or behind—

Senator MANCHIN. Can you evaluate what is going on? I am sure you all have interaction with your colleagues around the world, in different countries, whether they are adversaries or allies. The sci-

entific world seems to cross over pretty—I wish we could do as well as you all do in that arena.

Dr. LOHN. Yes.

Senator MANCHIN. How do you evaluate it?

Dr. LOHN. I have tried to study this directly, and U.S. is ahead. China has been gaining. We still have an innovation lead, I am confident to say, and we also have companies like those represented here that give us a huge leg up.

What I would like to point out, from a DOD perspective, is that the DOD has an opportunity to step ahead of industry in the adversarial context. A lot of the time my co-panelists here are developing products that do not have a natural adversary trying to mess with them, but the DOD does. And so that is a place where we really need to focus a little bit further on what is somebody going to do to subvert our systems as we deploy them.

Senator MANCHIN. Dr. Horvitz?

Dr. HORVITZ. First let me say that the people in the other fields of AI love machine learning but they have all existed side by side since 1956, when the first proposal was written about using the phrase “AI” for the first time. Machine learning has gained but it is simply—well, I should not say “simply” because it is important—a part of AI. It is not separate. It is one of the important disciplines within AI. That is the way AI researchers view machine learning.

Now it has grown up to be a very big discipline because almost every other discipline leverages the advances in that field, which are moving very quickly.

The U.S. is leading in science at the core principles and creative applications, from my point of view. That said, these days technical advances spread around the world like lightning. So at the scientific frontiers of IC [integrated circuit] scientists really keeping pace with one another around the world, there are issues around who has the right resources to do the computation that is needed, because these models are getting bigger and bigger and they are showing with getting bigger, that we do not see any leveling off just yet. You need tremendous amounts of compute for that kind of thing. There are probably two places in the United States that can compute like that and a couple in China right now.

So thinking about the resource constraints, especially on academic researchers, to push on the research is a very, very important direction.

The private sector is kind of like SpaceX in some ways. Microsoft, for example, is building platforms and tools, and it is working with customers in the Federal Government as well as in civil society and the private sector to understand what it takes to field these applications and technologies.

The one place that I worry about Federal applications in DOD is integrating in these scientific achievements into real-world workflows. I think the devil is in the details there. It gets into lots of engineering, human AI, human factors and human AI collaborative approaches. We need to get our hands dirty and work hard and then share ideas and insights across the sectors.

Senator MANCHIN. Thank you all so much. And then just one final one. I will say, respectfully, all three of you are working with Federal Government and with the Department of Defense and

being able to harden, basically making sure that we are not going to be hacked or the information we have is being protected. I would assume you all have done that, and we will talk about that more too. But I just wanted to make sure about that.

Senator Rounds.

Senator ROUNDS. Thank you, Mr. Chairman. Look, first of all, let me just say thank you very much for taking the time to come in and visit with us today. I think part of the challenge we have here is trying to explain and to express to other members here in the Senate just how serious the threats are but also how great the opportunities are, and the recognition that AI is not something that is 10 years away. It is here, has been here, and it is embedded in a lot of the things that we do right now.

Dr. Moore, I direct this question to you, due to your experience as the Dean at Carnegie Mellon University, but welcome all panelists to respond.

According to an article dated April 13, 2021, in The New York Times, a majority of the AI engineers working in the United States are from China and studied in China. I understand that some of the best programs in AI are at universities in China and they are graduating students at record rates. How can we replicate the same types of success at U.S. universities, especially in places like South Dakota, where we have Dakota State University and others that really do have experience in cyber but they want to continue and grow it? How do we take the next steps to really develop that capability here?

Dr. MOORE. Thank you. A very important question, and I think there is some good news, that for us in the cloud sector the democratization of AI, so that we can have large groups of students learning about it all throughout the United States, has been a major part of our roadmaps. It actually does not work to anyone's interest in the United States for it to only be this small group of like 100 PhDs each year who come out with these skills.

So we are all in the commercial sector working on making it faster and faster and easier for folks to get up the training so that they can use AI usefully in their own jobs. I see it as being incredibly important for the work that we are doing with things like reserve programs and information technology or Cyber Reserve Corps for us to be taking those programs, to train people up using these democratized AI tools.

Senator ROUNDS. Thank you. Dr. Lohn?

Dr. LOHN. Thank you. I would like to maybe make two points, is that AI and cybersecurity are both getting easier to learn. When I started, not that long ago, it was very difficult. You had to go through a lot of math and build things out all from scratch. But now there are many tools and many learning resources available. And so I think that we have an opportunity to pull people through our industry giants but also to bring people through armed services, in the enlisted ranks as well as the officer corps, I think we can push for the development there and create these opportunities for servicemembers to have those skills while they are in service and then also to take them elsewhere.

Senator ROUNDS. Thank you. Dr. Horvitz.

Dr. HORVITZ. First let me say that I am proud that this country is still the world's talent magnet. We have built our country on that and it is fabulous we continue to act in that way and to serve in that role.

That said, we can do a lot better with educating our folks. Community college programs are really fabulous and they can use investment, fabulous faculty, and tools from industry and academia. There is a great deal we could do all the interesting skilling programs that are post-graduate skilling programs, online coursework we can invest in. The tools are becoming more usable and many companies are providing beautiful self-help, self-learning programs to use the tools.

I would like to say that we have new applications of AI even. For example, Microsoft has in private preview a project called Copilot that helps developers learn to code, gain insights about coding, and also having an AI coding companion. We are seeing it in private preview how much this is helping coders right now move ahead and become better as a team with the AI system.

So I think that I am optimistic, but I think we can do better.

Senator ROUNDS. Thank you. Just a question. With regard to the Department of Defense, if you were to grade the Department of Defense in terms of their ability so far and where we are at with regard to the application of AI in multiple application opportunities, what grade would you give the Department of Defense in their implementation and utilization of AI today?

Dr. HORVITZ. Can I just say that I would give most of this country a D, maybe a C minus, given the potential of what can be done. I think about health care and how AI is a sleeping giant for health care, whether it be VA system or other venues.

Senator ROUNDS. Is it fair to say we could find cures for cancer within 5 years if we would fully implement AI?

Dr. HORVITZ. Well, let me just say that advances like AlphaFold and RoseTTAFold are really helping us jump forward in the understanding, for example, of sale of machinery. So I am optimistic. I cannot give you a time that we will understand cancer one day, as a running computer program.

But let me back up a bit and talk a little bit about the possibilities for the Department of Defense. We often think about AI, even in your opening comments, which were fabulous, as on the battlefield, as kinetics. But DOD is a huge operation, in peacetime and in war. The logistics, planning, predictive models, employment, back to health care, the VA system all can benefit greatly by even basic applications of machine learning, predictions, diagnoses, and planning.

So I do not want to call out the DOD as failing when I see them doing fabulous work and really working to get on board quickly and doing some of the most enthusiastic and energetic catch-up right now of any organization. But this whole country can do better.

Senator ROUNDS. I enjoy it when you say the basic application of machine learning. Dr. Lohn?

Dr. LOHN. I am not quite as pessimistic as Dr. Horvitz, although he certainly has reason to be. I hesitate to give a letter grade but I would not put it quite as low as a D. I think that, as you mentioned at the end of your answer, that they have been doing a great

job of catch-up. They have been very enthusiastic within the DOD to adopt and develop technologies and have been trying things and fielding them quickly.

I would like to point out also they have a difficult situation as compared to many other people trying to field AI because of the adversarial and permissive environment that they are trying to do it in.

Senator ROUNDS. Thank you. Dr. Moore, I am out of time but do you want to try to give me a quick shot on it?

Dr. MOORE. I will give you a super-quick answer. The way that we are structured with such brilliant individuals within the U.S. military who are willing to try new things is fantastic. But I am really, really worried if I do not see a concerted effort but instead just lots of talk.

I was very encouraged by the creation of the new Chief Data and Analytics Officer under Deputy Secretary Hicks. I wish that person great success. This is how we are going to succeed is by having a centralized effort to put an artificial intelligence strategy across the whole DOD.

What I worry about, frankly, and what I would be really worried about for this individual is whether they are going to get enough support from the government and from the center of DOD to actually make changes that are needed, because you cannot just magic AI on top of existing systems. You have to think about how you are going to change operations. So please give support to your central AI leaders.

Senator ROUNDS. Thank you. Thank you, Mr. Chairman.

Senator MANCHIN. Thank you, Senator. Senator Rosen.

Senator ROSEN. Well I have been so excited to sit here and listen to all of this because as a former coder I started in the 1970s, 1980s, and 1990s, I wrote a lot of if-then code, so I think it is a good thing that we have moved a little bit forward.

To Senator Rounds, how do we get people going? We have got to start K-12 as early as possible, like my Building Blocks of STEM [Science, Technology, Engineering and Mathematics] Act that was passed into law. You have got to start the pipeline as early as you can to excite people about these jobs.

And Dr. Moore, all of you, thank you for mentioning my Cyber Ready Reserve Act, my Cyber Ready Workforce. How do we surge up the resources from public-private partnerships like we do with our other military reserves? And, of course, we started the Junior ROTC [Reserve Officers' Training Corps]. We are giving them a STEM track as well, so young kids in high school can see themselves doing this and serving in the military.

So I appreciate that, and I do think our challenge is to be sure that we bring these very complex ideas down to something tangible that people can really understand, because they are very, very complex and it is important that we all have a platform, a shared platform, to talk about them in the same way. And that is our challenge today.

But I want to talk just a little bit about international partnerships, because we do have to maintain our technological edge. We have to advance our competitiveness in relation to China and others, and we must act—well, we have to act yesterday. I mean, time

is moving. And so as the National Security Commission on Artificial Intelligence pointed out we have to leverage all of this.

I did join Senators Rubio, Cantwell, and Blackburn in introducing the United States-Israel AI Center Act, and that is bipartisan legislation to create that artificial intelligence collaboration between the United States and Israel, and Israel is an emerging hub for these technologies.

Dr. Horvitz, can you talk about how we can work with our international partners, because this does not happen in a vacuum? You mentioned silos across DOD or private-public and other countries. We know that this quantum computing, these complex problems are best when data is not siloed.

Dr. HORVITZ. In the National Security Commission on AI we focused a bit of our time on opportunities for international coordination among allies and like-minded nations, including sharing technologies, data, both in research and engineering as well as for operations. Lots to be said about that and I am very excited about the possibilities there.

This particular interest, for example, in some of the work that is going on in companies as well as was pointed out on the National Security Commission also on the JAIC, the Joint AI Center in DOD, on responsible development and fielding of AI technologies, fielding technologies that are resonant with the United States' democratic values and principles. It turns out that AI can act in different ways in the world. Bias can be unexplainable. Its use can be a challenge to civil liberties. And the U.S. can be a leader among nations in thinking through how do we actually field these technologies in a way that resonates and is in accordance with our approach to democracy, human rights, rule of law?

Senator ROSEN. Thank you. I want to continue to build on that, so for Dr. Horvitz and then Dr. Moore, you both served on these commissions. And the National Security Commission on AI called for a \$20 million increase to DARPA for AI-enabled cyber defenses. So I know how AI can be applied to detect malware and pattern recognition. Can you talk about how that really works? So right now we see the conflict in Ukraine with Russia. We are bracing ourselves for shields up, as CISA [Cybersecurity and Infrastructure Security Agency] is telling us, for cyberattacks. So can you just try to explain to everybody here a little bit how that pattern recognition works?

Dr. HORVITZ. I can jump in on a recent situation in Ukraine.

Senator ROSEN. Thank you.

Dr. HORVITZ. Microsoft detected, with a neural net model, a piece of malware that was related to a known piece of malware, attributed to a group that we refer to as Iridium—it is also called Sandworm by other teams—based in Russia, that put on machines in Ukraine software called wiper software, that wipes the drives clean.

We detected this and immediately dispatched patch and alerts to the Ukraine to protect their systems. And interestingly, what we are seeing in Ukraine—we just fielded a report a week and a half ago on what we are picking up from our signals in Ukraine—interesting signs of where the world is going with hybrid warfare, with coordinated attacks, kinetics plus cyber, that are not just associ-

ated in time but they are planful, where there will be an announcement about dissatisfaction with disinformation, machines being locked out in a broadcasting station in Kyiv, and then missiles hitting that station. Hybrid warfare, planful and deliberate. We have to look out for that and begin to plan for it.

Senator ROSEN. And so that, of course, goes to the workforce because you need people, not just coders, not just engineers, you need a really robust workforce in every area of the network to do that—oh, I have just about a minute—so that goes to the cyber workforce shortage. We really have to do a lot. It is a huge spectrum. Most people do not understand. They see your PhDs and they wonder what are the 2-year degree or certificate or apprenticeship jobs.

So can you talk about the jobs, the 600,000 jobs that are open in cybersecurity now, the kinds of things that somebody who is looking for a new job now, or maybe somebody coming out of high school even, can go and begin to get into this field at that level? Maybe you could speak to that.

Dr. MOORE. Absolutely. If a student at a community college starts to just learn Python or one of the sorts of basic languages of data science, and then starts to play around with data analysis on projects like that, immediately they are going to find that consulting companies, the big internet companies, and startups are going to be really interested in their skills. And having that applied experience, just downloading from some of the cloud networks, simple AI systems, where you can get up and running in a matter of hours in writing your own machine learning recognition system for computer vision or something.

So I want to see Python taught, followed by a data science class taught, and at that point that person is already very well distinguished for joining an organization which will train them further.

Senator ROSEN. Thank you. I think that really is our task, to try to help everybody understand. Six hundred thousand jobs open. Over 3,500 in my state, just in cybersecurity. What does that mean, because I want to plug people into the way that they can do that. So we will speak offline and maybe some good ways—

Dr. HORVITZ. Senator, just to make a comment. About a year and a half ago we opened up LinkedIn courseware to the world, including really rich sets of classes on cybersecurity, promoted by the (ISC) group, the cybersecurity professional organization, and saw I think nearly three million engagements with the courseware.

So let's think through how we can creatively use our platforms to bring people into the fold and get on the path to becoming cybersecurity professionals.

Senator ROSEN. I want people to see that these jobs are for them, not for somebody else. They can all do them. Thank you.

Senator MANCHIN. Thank you, Senator. Senator Kelly.

Senator KELLY. I see 7 minutes on the clock. Is this a new thing we are doing?

Senator MANCHIN. If more people come in it will not be.

Senator KELLY. Doctor, Doctor, Doctor, thank you all for joining us.

Dr. Lohn, in 2020, you contributed to a RAND study on the military application of artificial intelligence in which it was stated, and this is a quote, "There is also growing interest in the potential for

machines that can find and patch vulnerabilities in friendly systems or find and attack vulnerabilities in enemy systems. But these applications still cannot perform these tasks at the level of experienced humans.” And Dr. Horvitz mentioned dispatching patches and alerts to Ukraine. I imagine that was done with people.

So understanding that this technology is constantly evolving and maturing, are we any closer to leveraging AI to assess and either patch or exploit vulnerabilities in friendly or enemy cyber systems?

Dr. LOHN. We are somewhat closer. Certainly the technology continues to progress and there are new research papers. I think that there is opportunity for us to advance at a faster rate with appropriate funding. As I discussed earlier, we have gone away from the Cyber Grand Challenge model and our adversaries have adopted it, and I think we might consider whether we would want to push to accelerate these technologies faster.

Senator KELLY. What is appropriate funding?

Dr. LOHN. Appropriate funding? I am not sure. I would say in the tens of millions of dollars would let us continue the Cyber Grand Challenge effort.

Senator KELLY. And if we were to do that, how does this whole world look in, let’s say, a decade from now?

Dr. LOHN. A decade from now is difficult to say, of course. But what I would say is that the patching of the vulnerabilities is one aspect that is very important, but we already today have a lot of our patches known before we disclose that this vulnerability exists.

The real big push that we need to make on is incorporating the patches. It is a challenge for a lot of companies to take a patch that exists and put it into their systems, knowing that it might break their systems, they might encounter downtimes.

And so these technologies that are developing vulnerabilities, are developing the patches, are making progress. Where we need to put more progress is in deploying those patches. If we do not progress in the deployment of the patches we could actually end up in a more dangerous situation, where the world is flooded with vulnerabilities, and even though we know how to patch them we have not been able to slip them into our code to make the protection.

Senator KELLY. How about the other side of this, which is the exploitation of our enemies systems?

Dr. LOHN. The exploitation of our enemy systems is kind of on that same bend. As we exist today, you can spread these exploits very quickly. The way it works is somebody finds a vulnerability, and then they will develop some attack code for that vulnerability, and then they can post it on the internet or into offensive hacking toolkits. And it just downloads automatically into your toolkit and now you can push a button and go sometimes. That can happen very, very quickly.

And so I think there is actually more opportunity for us to make progress on the defensive side, where we are slow today. I think the offensive side is already relatively quick. And so we have some opportunities to advance there but I would really like to focus on the defensive side. I think that is where the biggest gains are to be made.

Senator KELLY. And Dr. Horvitz or Dr. Moore, where do you see us in about 10 years on this run?

Dr. HORVITZ. One comment is I see tremendous opportunity to automate. When I say that, that does not mean workforce issues go away. I think we need people to be shifting over to doing more intensive, creative work in this space, and we will have plenty of that need arising.

One of the problems with automation right now is false positives. More accurate AI systems that can do better at reducing false positives and false negatives, which will come with more training data over time, will be helpful. Also the whole idea of coming up with strategies, for example, like I will accept, in this setting, higher false positives for shutdown that will be frustrating to protect me in this situation that I am in right now, sort of context-sensitive control of thresholds on automation.

To date, when it comes to an important alert, the AI is helping humans triage through thousands of alerts coming in. I think that will get better and better as we get better and better AI systems.

Senator KELLY. How far are we away from—go ahead, Dr. Moore.

Dr. MOORE. I just wanted to add, it is not going to get automated to the extent that we will need fewer cyber warriors on the U.S. side. You will get hopefully a larger workforce using vastly more powerful tools. So one person does the work of 10,000 people in 2022, but it will still have to be quite an army of humans.

Senator KELLY. How far away are we from having an artificial intelligence system being able to write really powerful code to exploit vulnerabilities with little input, like just giving some AI code, like a set of requirements, we want you to do this. You know, here are the requirements and just hit a button and the code is written.

Dr. HORVITZ. Let me say that the concern with using Copilot, which I mentioned earlier, a system that uses a large-scale, what is called a language model chain on large amounts of code to look at prompts of code being written and writing code for you, can generate all sorts of interesting offense cybersecurity as well as cyber offense and cyber defense code. The study we did of Copilot, pre-general availability, was to make the system safer in that regard.

So to answer your question, automated code-writing systems, given prompts and constraints, are surprisingly real these days. How should we field tools to the general public, how they should be used, different questions?

Senator KELLY. Thank you.

Senator MANCHIN. Thank you. I have just got a couple of quick questions. Do you want another round? We are going to a real quick 5-minute round. So I will just start with this one.

When you look and see the superiority that we do have, or the advancements that you think that we may be, how did the Colonial Pipeline happen, that we were not able to detect that? How are we not able to send a very strong signal—and Russia seems to be prolific. I mean, they just made a business out of this whole hacking and hostage-taking, if you will, for profit. And the other countries that have joined. You know, I am understanding that our country is more hacked than any other country in the world, on a minute-by-minute basis.

How can we not be able to stop that and be able to send a signal strong, or shut some of these rogue actors down? Whoever wants to start?

Dr. HORVITZ. Go ahead, Andrew.

Dr. MOORE. Not all of our own computer systems are created equally, so it is extremely important—

Senator MANCHIN. What now? I am sorry. I did not—

Dr. MOORE. Not all of our U.S. computer systems are created equally. We have a legacy of many systems developed over the last 20, 30, 40 years which have existed with some serious security holes, and it is very hard to manage systems built on on-prem large legacy systems of perhaps some computers from 15 years ago, some from 10, some from 5.

So the more sort of continued modernization of software, whereby software is run on very boringly sensible, secure, small pieces of infrastructure, this is the approach that clouds have adopted, means that is much safer for securing infrastructure than if you are having to remember to deal with hundreds, or actually tens of thousands of different old models and operating systems from the distant past.

One of the reasons I was so attracted to the cloud is because of this extra layer of standardization you get from just using modern, constantly patched systems instead of legacy bits of hardware.

Dr. HORVITZ. So I am going to jump into technology for a second and raise the prospect that colleagues have discussed over the last maybe 4 or 5 years, which is whether there should be new international laws and norms and practices regarding attack of civilian infrastructure—hospitals, pipelines, energy. One of the efforts has been called “digital Geneva Convention.” Let’s thing about that, think through that. Do we need new kinds of conventions and new kinds of laws and practices, internationally?

Dr. LOHN. And I will add on just a little bit. I would like to accentuate that not all computer systems are created equally and some of these ones that are legacy are very difficult to patch, and it might not be easy for us to make those adjustments. So we might need to have more protections on the outside and we might need to have higher standards for what we expect of a company to protect themselves, and we might need to communicate which things are unacceptable for other countries to do to us.

Senator MANCHIN. You would think that, like our grid system, you know, that could be absolutely a tremendous, tremendous challenge for all of us but also a horrible situation if they shut it down. And we have different carriers, different transmission in different parts of the country. I do not even know if they are interconnected. I do not know if they are talking to each other. I really do not know.

Do you know, first of all, if that is being done, and if it is not being done, should it be done? Food supply. The food chains, our basic infrastructure, our water, just the things that we depend on, take for granted every day. I would think that if we are not secure, if they were able to get to Colonial Pipeline and almost shut down tremendous flow of our transportation mode, that would have given them—

Dr. HORVITZ. Yeah. Let me play red team for a bit and imagine the future. And Mr. Kelly is not with us right now but to further answer his question, we can imagine AI technologies being used adversarially to think through not just a single Colonial Pipeline but a multi-pronged attack, a hybrid attack—going back to my comments about Ukraine, what we saw there—that look across multiple systems and sequences of attack and use the AI technology to optimize the plan and to carry it out.

I think we need to start thinking through—this is called red-teaming—in a creative way to prepare for those kinds of futures, to be proactive, to disrupt them before they happen. It is going to take a lot of work.

Dr. LOHN. And with just the last couple of seconds I would like to say that our grid operators took note, in 2015 and 2016, when Russia shut down the grid, but that it still scares me.

Senator MANCHIN. Senator Rounds?

Senator ROUNDS. Thank you, Mr. Chairman. I would agree with you. I think one of the nice things about it right now is that we have multiple grids out there, and they can take one but they would have to basically take multiples in order to get the entire country. But grid by grid, yeah, they are vulnerable.

I am just curious. The NSCAI Commission, of which two of you were members, in your final report you stated that the expanding application of existing AI cyber capabilities will make cyberattacks more precise and tailored, further accelerate and automate cyber warfare, enable stealthier and more persistent cyber weapons, and make cyber campaigns more effective on a larger scale.

I would like to hear your perspectives with regard to the threat assessment today, where we are today, with regard to AI-enabled cyberattacks on the DODIN [Department of Defense Information Network] and on the individual businesses within the United States? Where are we at today?

Dr. LOHN. As I mentioned in my comments, there is scarce evidence of adversaries using advanced AI methods for attacks these days, but most everybody believes that the demonstrations that we have seen, for example, in cybersecurity competitions, team-on-team, have led to lots of learnings. And we know that one of the DARPA Grand Challenge competitions in cybersecurity, which had this gaming going on, was picked up by China, who took quite a bit of interest that we did that and has been holding more of those kinds of competitions and looking at their results than the United States.

Dr. MOORE. Yeah, if I could add, if you look at where folks like myself and Dr. Horvitz are deploying engineers, even within an artificial intelligence group, which you might think is a bunch of mathematicians, a large fraction of all the work is on security, so perhaps these novice engineers who we were talking about earlier who are building AI systems, built on platforms with security guarantees underneath the platforms.

The word “platform” is an incredibly boring word to use. It makes people think of really boring computer science. But it is really important, the notion that a few places, places with resources like Google, are able to put huge amounts of effort into making these Lego blocks to build information systems where we

have had the opportunity to put in every single piece of security, which hundreds of thousands of human engineer years of thought have gone into.

So although I love startups, mom-and-pop shops for all kinds of areas, I would like to see the Department of Defense, as it is building its systems it needs to build them not on my cloud, necessarily, but on a secure cloud, not to try to do it as sort of on legacy bits of hardware. It is really, really important. The government needs secure cloud.

Senator ROUNDS. Dr. Lohn?

Dr. LOHN. I will just add a little bit along the lines of Dr. Moore, is that in addition to the tools and resources being provided by the tech companies that are represented here, there is a lot being done in the open-source community as well. And people will build a model or release a dataset or create some tool and then that is downloaded and used by these relative novices—not you—novices that he was referring to, and those may or may not have the same sort of security that we are expecting from our tech companies. There is an opportunity to help fund them, to do the hygiene and clean up their code as well.

Senator ROUNDS. And one last thought that I have to ask, and that his, when we talk about AI and we are looking at the power it takes, are the existing platforms that are out there, are the existing hardware systems, is the AI dependent on the capability, the power of the computing capability of the actual hardware itself, to an extreme basis, or is it being able to utilize an existing power source or computing capability to a greater extent by using the AI concept?

Dr. MOORE. The good news is there are two lines working, fully supporting each other. Hardware miniaturization is working extremely effectively at the moment, but the software folks are also figuring out new ways to take advantage of all the bits of technology. So that is an area where everything is advancing. And if I told you what was happening today it would be different from 3 months ago.

Dr. HORVITZ. To build the largest models, as we call them, that are showing some of these interesting emergent properties right now, where there is a great deal of interest, it is taking specialized hardware, and a lot of it, and a lot of energy.

Senator ROUNDS. Anything else?

Dr. LOHN. I would just like to add that the ability to keep on that trajectory is starting to look less promising because it requires so much.

Senator ROUNDS. Thank you. Thank you, Mr. Chairman.

Senator MANCHIN. Senator Blackburn.

Senator BLACKBURN. Thank you, Mr. Chairman. I appreciate that.

Let me stay with that AI, because there should be some practical applications that come forward. One of the things that has been of concern to me, as we have done our combatant command hearings, is looking at human capital and the workforce and retaining individuals that can solve some of these complex issues and problems, address these problem sets. So when you look at the utilization of AI you should be able to push forward with problem-solving in the

absence of individuals, by having the brainpower that is there to distill what you are hearing.

Dr. Horvitz, I think it would come to you. Talk to me a little bit about how you are using this, the distillation from AI, to help solve some of these problems of malign activity, business processes. And I would like to hear that from each of you, because that is how we are going to stay in the game when it comes to great power competition.

Dr. HORVITZ. And when you say malign, can you clarify what you mean?

Senator BLACKBURN. Adverse bad actors, trying to do bad things to us—

Dr. HORVITZ. Oh, in the world.

Senator BLACKBURN.—in order to thwart some of our positive activity, carry out malign influence campaigns, things of that nature.

Dr. HORVITZ. I see. Well, as I mentioned in my written testimony, one of the concerns with the rise of power AI technologies is the ability to generate content, for AI systems to generate deep fakes, for example. And we are going to be in a place where humans nor AI will be able to detect and discriminate a deep fake from a real scene, a real event in the world. And so we need technologies for that, and we described at least one technology called digital content provenance, which, in some ways the way I like to describe it is glass-to-glass, can you cryptography to certify this is non-AI technology, dealing with an AI outcome or capability, which is deep fakes, to certify that every time hitting this camera surface is represented by a pixel on display, and no one has changed anything, and you can actually track all the edge in between. So we can imagine working on that. That is an interesting front.

More generally, there is opportunity to study large datasets, and I think in our NSCAI report we talked about this idea of having new kinds of centers that would think through, collect data and do research and R&D on malign information campaigns, their source, how they spread and diffuse, how we might address them ideally.

Senator BLACKBURN. Okay. Dr. Lohn?

Dr. LOHN. Yes. I would like to expand just a little bit on Dr. Horvitz's discussion. Not only is there technology for creating fake images but it can create fake text, and that text can be very convincing. We did a study that found that it could convince people, American population, to oppose Chinese sanctions or to support or oppose the withdrawal from Afghanistan, either way.

But what I would like to kind of point out is that the dichotomy between the amount of skills required. So to build these models that can generate that text requires many, many geniuses, but to use it, not so much. All you have to do is type a couple of words, hit stop, go run, and then it fills out the rest. There is no real programming expertise required.

And so we need really smart people to build some of these technologies, but to use them, to build companies out of them or to defend ourselves, or the adversaries to come after us sometimes requires very little expertise. And that it both an opportunity and a threat.

Senator BLACKBURN. That is why—and I appreciate the mention of our civilian cyber force, which would help with that early re-

sponse, have people there that are able to utilize some of these technologies when we do not have individuals, enough people to do the work that we need to do. We can kind of bring them in an as-needed basis. I think that is a good and positive step, and I appreciate you all mentioning that in the opening.

Dr. Moore?

Dr. MOORE. Thank you. Your question is very on point, and thank you for bringing it up. This notion that folks can actually poison our own systems was kind of science fiction-y 5 years ago but it has happened to me, and I have been on the front lines of dealing with this, and attacks against Google systems. So, as you can imagine, that is now a major aspect of defense.

One thing I would like to mention is we at Google Cloud have partnered with the Defense Innovation Unit to stand up their secure cloud management solution, to be ready for these second-, third-, and fourth-level attacks, where everyone is looking above and beyond what each other are doing. It is absolutely the place where the battle is being fought at the moment.

Senator BLACKBURN. Okay. Thank you all for that.

Dr. Horvitz, Microsoft, what have they learned from, I think it is the Hafnium Project. Could you talk to me just a little bit about what the lessons learned are from that and then how you plan to use that information.

Dr. HORVITZ. The main lesson for the world is on-prem is not as secure as cloud. On-prem requires having your own machines. It might seem like I have my data and it is protected here but the amount of updates that are required to keep up with old software, for example, especially in small and medium-sized businesses that do not have IT teams, for example, it is challenging.

We recommend, for the top-notch security, move to the cloud and let the big tech companies take their best resources and ongoing surveillance and cybersecurity software, let them do the work for the businesses. That was the main finding, from my point of view.

Senator BLACKBURN. Okay. Dr. Moore, I see you shaking your head. Anything to add to that?

Dr. MOORE. [Inaudible. Presumably "no".]

Senator BLACKBURN. Okay. Well, thank you all. I know my time has expired, but to your answer I think the prevailing and unanswered question for the 21st century is who owns the virtual you, which is you and your presence online, and being able to distill some of this information and be able to decide what is real, what is fake, what is a misrepresentation is one that we are going to have to continue to work through.

Thank you all for your time.

Senator MANCHIN. Thank you, Senator.

Let me just again thank all of the witnesses. Thank you all for being here and sharing with us your knowledge and forecasts and what we need to do and how we need to all work together. I tell you, we are mostly committed to that. Artificial intelligence development and the applications to national security and our everyday lives has the potential, really, to revolutionize our lives, and we understand that, and most importantly, our society. But Congress and the Federal Government must be prepared to prioritize—and I

have heard it loud and clear—prioritize the necessary investments now.

So I know Senator Rounds and I share the priority and I look forward to working together on implementing what we have learned today and continuing to work with you all.

With that the meeting is adjourned.

[Whereupon, at 4:05 p.m., the Subcommittee adjourned.]

