Stenographic Transcript
Before the

Subcommittee on Cybersecurity

COMMITTEE ON
ARMED SERVICES

UNITED STATES SENATE

TO RECEIVE TESTIMONY ON HARNESSING ARTIFICIAL INTELLIGENCE CYBER CAPABILITIES

Tuesday, March 25, 2025

Washington, D.C.

ALDERSON COURT REPORTING
1029 VERMONT AVE, NW
10TH FLOOR
WASHINGTON, DC 20005
(202) 289-2260
www.aldersonreporting.com

TO RECEIVE TESTIMONY ON HARNESSING ARTIFICIAL INTELLIGENCE CYBER CAPABILITIES

Tuesday, March 25, 2025

U.S. Senate
Subcommittee on Cybersecurity
Committee on Armed Services
Washington, D.C.

The committee met, pursuant to notice, at 3:31 p.m. in Room SR-232A, Russell Senate Office Building, Hon. Mike Rounds, chairman of the subcommittee, presiding.

Committee Members Present: Senators Rounds [presiding], and Rosen.

OPENING STATEMENT OF HON. MIKE ROUNDS, U.S. SENATOR FROM SOUTH DAKOTA

Senator Rounds: Good afternoon, and I'd like to thank our witnesses for appearing today to discuss how artificial intelligence can be utilized to enhance the Department of Defense's cyber capabilities. We have just heard from experts in our closed session from the U.S. Cyber Command, the Defense Advanced Research Projects Agency, or DARPA, and the DOD's Chief Digital and Artificial Intelligence Office. These organizations all play a crucial role in making sure the Department is postured to carry out its national security mission in cyber space.

Recent cyberattacks against U.S. critical infrastructure are a stark reminder of the growing sophistication and persistence of cyber threat actors. To outpace our adversaries in the cyber domain, the Department must rapidly harness the advances of AI technologies. This means that the Department of Defense needs capable partners outside of the Pentagon who are moving at breakneck speed to solve our national security challenges.

This brings us to our hearing topic today; how the Department can leverage AI-enabled capabilities to field exquisite, offensive, and defensive cyber tools, enhance our ability to detect cyber threats, and automate threat mitigation to gain an enduring advantage in cyberspace.

I also look forward to hearing from the witnesses about how the Department can be better equipped to counter enemy AI-enabled cyber capabilities, and leverage AI to enhance our overall war fighting ability in the cyber domain. Our innovators and tech companies are one of our asymmetric advantages in the cyber fight, but the gap is steadily closing.

At the tip of the spear is artificial intelligence. Unfortunately, the Chinese Communist Party understands this all too well. Xi Jinping has spoken about the importance of AI. With the release of DeepSeek earlier this year, it is clear unless we act decisively and soon, China will not be playing catch up. We will.

U.S. advancements in this critical technology are impressive, and we are fortunate to have some of the best innovators in the world. As Silicon Valley and other leading technology developers continue their research and development of AI at the bleeding edge, our job must be to integrate those tools in a secure, but rapid fashion into our cyber capabilities.

I look forward to hearing from our witnesses who all bring unique and firsthand experience about how the Department can speed up its use of AI in the cyber domain. Again, thank you to our witnesses for coming here today.

And before I introduce them, I'll now recognize Ranking

Member Senator Rosen.

STATEMENT OF HON. JACKY ROSEN, U.S. SENATOR FROM NEVADA

Senator Rosen: Well, thank you, Chairman Rounds. And I'd like to begin by welcoming our panel, and thanking you all for joining us. This topic has profound implications for our national security, I would say, for our personal security, for everything in our world to come.

But this is actually my first hearing as ranking member of this subcommittee, and I am really honored to work alongside Chairman Rounds, our colleagues, and each of you on how we can responsibly integrate innovation and the increasing pace of technology including artificial intelligence into our national defense strategy and into the hands of our service members to enhance their speed, their capabilities, and their operating picture. Well, of course, all the time we have to balance the risks and rewards concerns of AI and what it teaches us.

So, with great promise comes great responsibility. We know that our adversaries are developing new AI tools and have the potential to fundamentally shift the nature of warfare. We've began to see how new uses of AI can help our own service members counter such threats and take proactive offensive actions in the moment as well.

However, the rapid pace of AI innovation also raises really important questions about its ethical implications,

its governance, and the security risks it poses as well. We're operating in a new world without guardrails and we need to tread carefully, balancing such caution with the need to create an environment that allows for innovation and agility.

And there are also challenges we must overcome in order to both mitigate the risks of AI and make the most of the opportunities that I know it presents. In particular, we need to further invest in and expand the AI workforce, both at DOD, and across the government, across the private sector. We have to increase it everywhere to harness our full potential. I truly believe this.

As a former computer programmer, systems analyst, myself, I can say from firsthand experience that AI has vastly changed the technology landscape since I began my career. Many of the coding and the programming skills that people like me brought to the table, which form the backbone of what CYBERCOM personnel do every day, in both offensive and defensive operations, can now be supplemented by AI.

And I know it doesn't replace us, that's for sure. But however, this does pose its own set of risks, and it creates a deep need for us to invest in that new kind of cyber workforce that is centered around understanding these AI skills. And we continue to have a cyber and AI skills gap.

And until we meet that challenge of bridging it,

understanding it, being able to see its potential, and at the same time understand how it improves our own potential as human beings, we're going to continue to be at the risk of our adversaries having the upper hand.

So, I look forward to discussing such challenges today and over the course of this Congress. I thank our panel once again for your expertise and contributions to that effort. And I thank you again, Mr. Chairman.

Senator Rounds: Thank you. And it is a pleasure to have you here on the team with us. And this is one of those subcommittees in which it is very bipartisan, and we have focused on this since the creation of this by Senator McCain back in 2017, I believe. And the path forward, I think, has been made better because of the work that we've done in the past on a bipartisan basis to keep everything on the straight and narrow.

I want to thank all of you once again for coming in and participating in this open session. And we have with us, today, all three of you here. Beginning with Mr. Jim Mitre, Vice President and Director of RAND Global and Emerging risks. Mr. Mitre, welcome. Mr. David Ferris, Global Head of Public Sector, Cohere. Welcome. And Mr. Dan Tadros, Head of Public Sector, Scale AI.

And I understand that the agreement has been made that Mr. Mitre, you will begin today. So, we welcome you for

your opening statement, sir.

STATEMENT OF MR. JIM MITRE, VICE PRESIDENT AND DIRECTOR, RAND GLOBAL AND EMERGING RISKS

Mr. Mitre: Terrific. Chairman Rounds, Ranking Member Rosen, thank you so much for the opportunity to testify today on the national security implications posed by the potential emergence of advanced artificial intelligence, or artificial general intelligence, AGI.

Leading AI companies in the United States, China, and the rest of the world, are in hot pursuit of AGI, which would possess human level or potentially even superhuman level intelligence across a wide variety of cognitive tasks. The pace and potential progress of AGI's emergence, as well as the composition of a post-AGI future, are uncertain and hotly debated. Yet the emergence of AGI is plausible and the consequences so profound that the U.S. national security community should take it seriously and plan for it.

Consider the following. What would the U.S. government do if in the next few years, a leading AI company announced that its forthcoming model had the ability to produce the equivalent of 1 million computer programmers as capable as the top 1 percent of human programmers at the touch of a button. The national security implications are substantial and could cause a significant disruption of the current cyber offense defense balance.

At RAND, we are planning for it. Our work has revealed that AGI presents five hard national security problems. First, AGI might enable a significant first-mover advantage via the sudden emergence of a decisive wonder weapon. For example, a capability so proficient at identifying and exploiting vulnerabilities in enemy cyber defenses, that it provides what might be called a splendid first cyber strike, that completely disables a retaliatory cyber strike. Such a first mover advantage could disrupt the military balance of power in key theaters, create a host of proliferation risks, and accelerate technological race dynamics.

Second, AGI might cause a systemic shift in the instruments of national power that alters the balance of global power. The history of military innovation suggests that being able to adopt a new technology is more consequential than being the first to achieve a specific scientific or technological breakthrough.

As the U.S. allied and rival militaries establish access to AGI and adopted it at scale, it could upend military balances by affecting key building blocks of military competition such as hiders versus finders, precision versus mass, or centralized versus decentralized command, and control. States that are better postured to capitalize on and manage systemic shifts caused by AGI could have greatly expanded influence.

Third, AGI might serve as a malicious mentor that explains and contextualizes the specific steps that non-experts can take to develop dangerous weapons such as violent cyber malware, widening the pool of people capable of creating such threats.

Fourth, AGI might achieve enough autonomy and behave with enough agency to be considered an independent actor on the global stage. Consider an AGI with advanced computer programming abilities that is able to break out of the box and engage with the world across cyberspace. It could possess agency beyond human control, operate autonomously, and make decisions with far reaching consequences.

Fifth, the pursuit of AGI could foster a period of instability as nations and corporations race to achieve dominance in this transformative technology. This competition might lead to heightened tensions reminiscent of the nuclear arms race, such that the quest for superiority risks triggering rather than deterring conflict. Misinterpretations or miscalculations could precipitate preemptive strategies or arms buildups that destabilize global security.

As the U.S. Department of Defense embarks on developing the National Defense Strategy, it will have to grapple with how advanced AI will affect cyber along with all other domains. The five hard problems that EGI presents to

national security can serve as a rubric to evaluate how the strategy addresses the potential emergence of AGI.

Thank you for the opportunity to testify. I welcome your questions.

[The prepared statement of Mr. Mitre follows:]

[SUBCOMMITTEE INSERT]

Senator Rounds: I thank you. Mr. Tadross, unless you folks have agreed on a different. Mr. Tadross.

STATEMENT OF MR. DAN TADROSS, HEAD OF PUBLIC SECTOR, SCALE AI

Mr. Tadross: Chairman Rounds, Ranking Member Rosen, members of the subcommittee, thank you for the opportunity to be here today.

My name is Dan Tadross. I lead Scale AI's public sector business. Every day, my team is singularly focused on how to bring best-in-class AI into the DOD and other agencies. Scale was founded in 2016, and since that time, has powered nearly every AI innovation. Our role in this critical ecosystem provides us a unique opportunity to understand how to build high quality AI systems powered by the world's best data.

Our work is deeply personal to me as I have worked nearly my entire career at the intersection of AI and the government. During my time as an active-duty marine, I had the privilege of helping to stand up the Joint Artificial Intelligence Center, which enabled me to see firsthand the challenges and struggles associated with the DOD's implementation of AI.

This hearing comes at a critical time for the future of AI leadership. And before we discuss what the United States must do to win, it's important to analyze where things stand today.

AI is made up of three main pillars; compute, data, and algorithms. More than one year ago, the United States was clearly ahead on all three. However, today, that is no longer the case. Advancements from China have shown that they've closed the gap. And, today, China is leading on data. We're tied on algorithms, but the United States remains ahead on compute. It's clear that the race is neck and neck.

In order to compete more aggressively, the CCP has implemented a whole-of-country approach to accelerating its pursuit of becoming a global standard for AI from an investment standpoint. And for the first time in history, China is benchmarking AI investment off the leading tech companies and not the United States government.

Last year, China spent at least $1.2 billion on data labeling alone compared to our under $100 million by the United States. And as part of China's AI Plus initiative, the government established seven data labeling centers around the country to mainly support public sector application.

Beyond data, while the U.S. has been stuck in a research and pilot mindset, the CCP has rapidly increased their investment in fielding AI capabilities. In the first half of 2024 alone, the PLA issued 81 contracts with large language model companies to rapidly grow their capability.

To win, the U.S. needs to unleash our technology to the warfighter at an unprecedented pace.

When it comes to adopting and implementing AI, the DOD has not launched a new AI program in nearly a decade. For the past four years, DOD leadership spent countless hours developing potential use cases for AI, researching and piloting AI systems, and even putting out guidance to stop users from utilizing AI.

We still have time, but the window is closing. If we want to win, we must not only buy into a vision, but it also takes three clear and decisive actions. Number one, is put the right AI foundation in place to start. The DOD lacks the foundation piece, the foundational pieces necessary to build scale and implement widespread AI solutions. This needs to change, and we must put in place the elements necessary to expand the use of AI programs. And this starts with data.

To truly prioritize and execute the strategy, it requires two main aspects; AI-ready data requirements, and enterprise-wide AI data infrastructure. The U.S. government is the world's leading producer of both quantity and diverseness of data. But nearly all that data is going unused. If the U.S. wants to turn our data into an advantage, this must change.

In multiple NDAAs, his committee has directed,

suggested, and tried to require the DOD to prioritize AI-ready data requirements, but it's clear that more must be done. In parallel to implementing the requirement, the Department should also set up enterprise-wide AI data infrastructure.

This commercial best practice ensures that AI programs are developed in the most efficient and cost-effective manner, and leading tech companies have long realized this requirement for effectiveness. And for that reason, China is mirroring this same approach.

Number two, is to shift our mindset to be an implementation-first. If the U.S. is going to win, we must shift into an implementation-first mindset. In order for this to occur, Scale believes that the DOD must set must first set a North Star related to robust AI implementation in no more than five years.

This should focus on agentic applications such as agent warfare, and would provide an ambitious vision and enable tangible multi-year plan to reach it. Scale is actively working on deploying the first instance of this in INDOPACOM and EUCOM through DIU's Thunderforge effort.

Number three, is to ensure our acquisition system no longer slows us down. AI is unique in that it is software, but needs to be maintained like hardware, which presents challenges for the DOD given that it doesn't neatly fit into

a legacy acquisition system. Congress took a strong first step by requiring the DOD to break out AI elements of programs in the future budgets. And it is critical that Congress continues to provide oversight to push the DOD to do so quickly as possible.

In addition to proposals like the FoRGED Act, Scale also believes that we need to continue to look at finding ways to break through the challenges of multi-year budgeting, which is clearly still holding back the DOD's implementation of AI. With these three decisive actions, the DOD will be better positioned to adopt and effectively implement AI solutions.

Thank you again for the opportunity to be here, and I look forward to your questions.

[The prepared statement of Mr. Tadross follows:]

[SUBCOMMITTEE INSERT]

Senator Rounds: Thank you very much, sir. Mr. Ferris.

STATEMENT OF MR. DAVID FERRIS, GLOBAL HEAD OF PUBLIC SECTOR, COHERE

Mr. Ferris: Chairman Rounds, Ranking Member Rosen, distinguished members of the subcommittee, thank you for the opportunity to testify today.

My name is Dave Ferris, and I'm the Head of Global Public Sector at Cohere. I previously served nearly 17 years in the Canadian Armed Forces, including deployments to Afghanistan and Ukraine, and spent the last two years of my career on the U.S. Joint Staff in the Pentagon.

Cohere is a leader in building AI systems designed exclusively for government and enterprise use, prioritizing privacy, security, multilingual capability, and verifiability. Our expertise spans from building foundational AI models, to developing AgentX systems. We focus on operationalizing AI, integrating it into real missions, under real world constraints. We partner with allied governments, agencies, and leading global companies.

Our primary goal is seamless integration, deep customization, and accessible solutions that deliver immediate practical value and confidence. We specialize in private deployments, even air gapped environments where we do not see our customer's data.

Today, I would like to highlight four key topics of

focus gleaned from having worked with high security cyber defense government organizations. The first key topic is how AI can significantly enhance the Department of Defense's mission, particularly in cybersecurity and intelligence.

AI systems can dramatically improve pattern recognition and anomaly detection across vast data sets. They can be invaluable for sorting through and synthesizing huge volumes of multi-source information, and they can help automate a number of crucial tasks to provide early warnings and free humans to focus on making strategic decisions.

Similarly, effective AI adoption requires integrating technology thoughtfully with existing workflows. Human AI teaming is crucial in ensuring AI tools have user-friendly interfaces. It helps build trust and maximizes operational value.

A second key topic is to consider how AI can help fight back against competitor nations and malicious actors that are already employing AI-enabled cyber capabilities. Reports have shown these countries are automating their intrusion attempts using AI to generate deceptive deep fakes, develop more convincing phishing lures, and create information warfare.

To stay ahead of these AI augmented threats, DOD must likewise incorporate AI across its offensive and defensive cyber operations. Large language models provide a unique

ability beyond traditional, rule-based machine learning systems for language understanding and reasoning capabilities that allows for dynamic identification, analysis, and generation of conclusions across a wide range of use cases.

The third key topic is to understand how technical considerations are critical to successful AI deployments in defense. Models should be right-sized for their specific mission. Specialized efficient AI models can often outperform larger general-purpose systems. This enables deployment even on limited hardware such as edge devices like laptops or classified data centers.

Flexible secure deployment architecture is critical. AI systems must be deployable across multiple secure environments and ensure AI sovereignty. Similarly, ensuring models are hardware agnostic and interoperable, so there is no lock into one cloud or one chip provider, is essential to ensuring supply chain and operational security.

Collaborative development through public-private partnerships allows for rapid customization of or creation of new AI models to meet specific operational context while protecting sensitive information. The DOD does not need to undertake the costly, time-consuming task of developing every AI model from scratch.

The final key point is to highlight that Congress can

take immediate action to accelerate responsible AI adoption. Congress should modernize procurement processes to allow innovative AI startups easier entry. Procurement should reward innovation, agility, and performance, not just size or past contracts. New legislation should promote interoperability, and open standards to prevent vendor-locking and enable diverse AI solutions to seamlessly integrate into defense ecosystems.

Finally, Congress should support robust internal benchmarking and testing specific defense applications rather than the use of generic academic benchmarks. This would ensure AI reliability and trustworthiness in critical missions.

In conclusion, Cohere is committed to partnering with DOD in Congress ensuring AI tools are secure, effective, and mission-ready. Thank you, and I look forward to your questions.

[The prepared statement of Mr. Ferris follows:]

[SUBCOMMITTEE INSERT]

Senator Rounds: First of all, thank you to all of you, and I appreciated your opening comments. We'll pass this back and forth a little bit with regard to questions and so forth, but we'll try to get to as many as we can in a short period of time.

I wanted to begin, Mr. Mitre. The artificial intelligence is here to stay. It's not going away. You gave us some warning signs out there, but I wanted to hear from you. We can't slow down on the development of AI, or we know that our competitors will clearly outpace us.

Give me your rendition of how we do this without losing facts or losing sight of the facts that there can also be some dangers involved. You've identified a number of the possible dangers, but how are we going to do this and still keep that in mind?

Mr. Mitre: That's a great question, and I welcome it. And I wholeheartedly agree that it's in America's interest to stay at the forefront of the development of generative AI and AI technologies more broadly.

So, the way in which we can address this issue is, first, it's helpful for the U.S. government to really understand what the current state of the technology is, and make sure that folks within the government, particularly those that are working in the national security community, really understand what's happening with the technology.

Because one of the challenges with this technology is that it's not being developed by government, it's being developed by the private sector. So, just understanding what the current state is critical so there aren't technological surprises that come out that shock people in the national security community.

The second thing that government should be doing here is really looking for applications in the national security context. What are the specific use cases that it can be applied? What are potential pathways to wonder weapon or ways in which it could be highly advantageous in a military competition that's critical to do, and that means having the AI in an environment where you've got sufficient compute, where you've got the right networks, et cetera. You can act actively, experiment with it, and get the technology in the hands of the operators to play around with it.

The third thing is preparing for contingencies. There's a wide range of possible things that could happen. A loss of control scenario, for example, areas where there is technological surprise and the Chinese get ahead. What would the U.S. government do in such contingencies? We should think that through in advance and have plans ready to address it.

Senator Rounds: Thank you. Mr. Tadross, this works right into some of the comments that you had made. And I

want to just, number one, I think it would be a statement we would all agree on that continuing resolutions are absolutely not the long-term plan that we need.

If we're going to be able to move forward with the investment in AI that we need, that may very well save a lot of lives in the battlefield. So, I would recognize that up front, and I think you were rather suggesting that a little bit in terms of our failure to keep up with the demands of how quickly AI is developing elsewhere.

You also said something else, though, and I wanted to touch on two items. Number one, you talked about the fact that we have data, which is unused. And I want you to explain that a little bit. And then, second, of all feeding into to what Mr. Mitre talked about, you talked about agentic warfare.

And can you talk a little bit about what that really means for the -- I mean, we've got a lot of folks out here that this may be their first introduction to the coordination of different applications that are directly involved in warfare versus the application of AI in general. So, first of all, data unused, and second of all, agentic warfare.

Mr. Tadross: Of course, Senator, and thank you for the question. So, in terms of data being unused, the approach that I was kind of looking at there is the aspect that,

right, now an enormous amount of information is being collected day to day. But to take kind of a quote from one of the previous Secretaries of the Air Force, "We treat data like exhaust as opposed to something that's really critical to use."

So, as a result, every time that we run an exercise, run a command post exercise in terms of large amounts of chat data is being developed, large amounts of chat data is being traced back and forth, what's happening is at the end of that exercise, all of those hard drives are just being purged or being neglected and goes into storage.

So, those are instances where the interactions between participants of a staff, for example, should be getting captured, and we should be using that to help develop training data to using it to help develop benchmarks against how these algorithms should operate. And then by doing so, are eventual development of agentic solutions can be more in line with what is required by those end users, which I think then brings us into the idea of like agentic warfare.

And really what that means, my interpretation of this, is we're trying to move humans, move to a position from humans are the loop to humans on the loop. So, right now, if a staff at INDOPACOM, or at EUCOM, or any other combatant command needs to make a decision, the process at which they do that hasn't really changed since the advent of the

Napoleonic staff structure. We take the problem, we divide it up, and then what's required is that the commander at the last minute has to synthesize all of those things together and then make an informed decision.

The effort of agentic warfare is to move to the point where much of that low-level staff work can be done by these AI agents through automated methods with human oversight and supervision of the process. It's important to maintain some human oversight of the entire process to ensure that human-context judgment, and the competitive advantage of the U.S. military, which is the fact that we have the most well-trained, well-versed staff and NCOs on the globe.

Senator Rounds: Thank you. And, Mr. Ferris, I've got some questions for you as well, but my first five minutes is up. We will do a second round, but at this point, I'll come back to Senator Rosen.

Senator Rosen: Thank you. You know, I want to talk a little about guardrails and benchmarks. Both, I believe they go hand in hand. And over the last year, discussions between Congress, prior administrations, they've always centered around trying to come up with guardrails to promote responsible AI. You all know what I'm talking about; nobody wants it to become an unchecked technology.

The current administration has raised concerns that guardrails might inhibit innovation. I believe we need both

effective guardrails and benchmarks because the benchmarks, just as if your child goes to school, they're the test to show if they're learning and going in the direction that you're expecting them to go. That's what's going to keep that circle in check.

So, I'm going to have questions for all three of you, but I'll start -- they're similar, but I'm going to start with you, Mr. Mitre. How should we develop guidelines, or the guardrails, and benchmarks in ways that mitigate risk without stifling innovation?

And I might also add, I'm actually going to ask all three of you this. How do we develop, for those of us sitting in this seat with all of you, a common policy language that is both nimble, but provides the availability for us to do effective oversight?

Mr. Mitre: Thank you, Senator. So, I wholeheartedly agree that it's important for us to understand what these models are capable of doing, right? They're developed, and they're released into the world with no user manual. It's not entirely clear what applications they'll be able to perform or how capable they'll be at doing that.

So, benchmarks are crucial, particularly in a national security context. It's helpful to understand what might the latest generation model be able to do in terms of offensive cyber defensive, cyber capabilities in terms of potentially

informing non-experts on how they go about designing a bioweapon that could be highly transmissible and lethal, et cetera. So, the real focus that is warranted is on developing benchmarks to really just evaluate and understand what the risks are.

Separate question in terms of what should government do about those risks if they emerge, and should regulations or something along those lines be appropriate in that regard? I defer to government for specific thoughts on that. What we're trying to do is just understand at first pass what are some of the risks here and make sure that people are well informed on that point.

Senator Rosen: Thank you. And I'm going to just go down. Mr. Tadross, the same thing. Developing the guardrails. The benchmarks tell us one thing, the guardrails tell us another. I guess I'll make it all the same question. We are going to struggle. We have to put this down in some way on paper that allows us to be nimble and provide that ability to do the oversight we need to.

And so, if you have thoughts about how we develop this common language that we can all speak from or start from, I think is really critical, so.

Mr. Tadross: Absolutely. So, the way that our company kind of looks at this, at least as it relates to guardrails in the implementation of AI in the Department of Defense, is

to really look at it from a perspective of people, process, and technology. And that while the technology needs to have guardrails by itself in terms of like its responses when it will trigger a refusal, or when it may not, there still needs to be the other two portions of this triangle.

So, people need to be trained on how to best leverage the capability. And then, the process needs to be adapted. Because if we just bolt AI onto an existing process, then the advantages are somewhat lost. So, the doctrine and training of the individuals needs to adapt at the same time as the technology has fielded.

And this goes back to my position about implementation. The only way to do this is to experiment in low-risk environments and to iterate very quickly. short of that, I'm afraid that the concern about trying to write out the full answer at the beginning of the test is probably unlikely. So, you need to be able to learn from doing and be able to build off of that.

As it relates to benchmarks, this is an area where our company's done quite a bit of interesting work. So, we have a paper that we've published showing that most of these large language models and AI systems will essentially cheat off of existing benchmarks. They've seen them, they understand the rules of the test, and as a result, they will score abnormally high.

The approach that we've taken in partnership with organizations like CSIS and the CDAO is to build custom benchmarks that are focused on the domain at which it actually matters to test. So, we've built these custom benchmarks. The algorithms have never seen them, they've never been incorporated in their training data. And as a result, you can have a little bit more faith in the performance of those algorithms.

Senator Rosen: Thank you. Mr. Ferris?

Mr. Ferris: Thank you, Senator. I echo the sentiment of my colleague on the panel here. I think public benchmarks can often be gamed. I'll start from the perspective of benchmarks because I think it's relevant to what my colleague was saying. They don't typically show the performance in real-world context. So, we would --

Senator Rosen: Is using the word "audit" better than benchmark?

Mr. Ferris: Well, no, I think we would say creating custom benchmarks.

Senator Rosen: Just like right-sizing your model.

Mr. Ferris: Yeah, exactly. Okay. And, you know, to kind of take that down one step further, we work very closely with our customers from beginning to end in order to ensure that we're right-sizing that model, developing the benchmarks. But that also includes some human evaluations

because that human AI interface is obviously imperative as we're moving down this.

And with respect to guardrails, you know, there's this healthy tension between accountability and agility, I would say, in this environment. And so right now, we obviously would suggest that we want to lean into the agility. We want to take an adoption mindset, but can't, you know, sacrifice really the security reliability and verifiability.

So, you know, ensuring that you have clear visualization into the data lineage, ensuring that you have a good understanding of how those safety measures have been built into the model during its development and deployment, I think, is imperative.

Senator Rosen: Well, I think because you say you want to lean in to -- oops, I'm going over my time. I'm sorry. Can I finish the thought? Lean into the agility, but if you don't keep humans, if you don't keep someone else in the loop, people's lives are on the line. And it's still a computer just analyzing data. And so, at that execution point, you have to consider leaning into agility. But at what execution points do we allow for a better decision? And I'll let it go to my -- maybe that's a philosophical question.

Senator Rounds: Well, look here, and I'm going to lead into this a little bit, too. And I'm going to start with

Mr. Ferris. We talked about right-sizing systems. And kind of along the same line here, I'm going to compare that because I'm not sure if I'm thinking the same thing that you're proposing.

But loitering, munitions as an example, we have clear evidence that in the Nagorno-Karabakh War between Azerbaijan and Armenia, loitering munitions were used. They were able to, as you know, basically unmanned aerial vehicles, they moved into a particular kill box, identified targets that were there. And then without a human in the loop, they were able to identify the types of systems that were there, whether it was a tank and an armored personnel carrier, a command center, a radar station aircraft, and so forth.

But because they had that capability, they could then choose which weapon system based upon which drone was there in the area and at an appropriate time attack each of them. Is that the type of -- can you talk about, is that what you mean when you say right-sizing in terms of having the capability for that particular mission set? Or share with me what you mean by that.

Mr. Ferris: Yeah, thank you, Senator. In that context, I think when we talk about right-sizing the model, we're talking about making sure we're bringing the appropriate solution to the use case. So, to use your example, we would be looking at, you know, how the models

are used to analyze all that multi-source information that's coming into the system and from various sources, but also potentially from different sensors and systems.

I think what's important is that we would suggest that by analyzing, using artificial intelligence to analyze all of that data, it allows you to elevate the level at which a human can make that decision. We would still suggest that the human AI interface is important, and that should be maintained during these types of operations. But really what AI allows you to do is to elevate that decision and make it closer to when it needs to be taken, potentially.

Senator Rounds: I'm going to -- you're following right into what my next question was going to be, and that is with regard to -- and I'm going to run this all the way down the line again, but I want to talk a little bit about humans on the loop, and humans over the loop, and defining each of them, if you would, in terms of where we're at today and where we're going to be tomorrow.

And I'm going to talk about it in both offensive and defensive capabilities. And the example that I would use that if you could build upon, is we have systems right now that for defensive capabilities, we arm them, but once they've been armed, they can automate to protect our platforms.

And that means if you have incoming missiles,

particularly if you're talking, you know, less than a minute to respond, to be able to identify a missile incoming, such as what we've seen in the Red Sea region with regard to Houthis attacking our systems.

But to be able to identify it, identify the type of weapon system necessary to take it out, and then to be able to execute and then to have backups along with it, how far along are we, and what will AI do with regard to having that whether there's a human directly in the loop of making that decision, or on the loop having armed it, or over the top of the loop, not engage at all.

I'd like your thoughts, then I'm going to ask our other two members here as well for their thoughts.

Mr. Ferris: Yeah. Thank you, Senator. So, obviously, I would say that, currently, we're supporting or we're seeing AI deployed in an environment with humans in the loop, as you described, and on the loop where there's some oversight. But certainly, I don't think we're yet at that over the loop where they're elevated outside of the analysis and execution of the mission set, if you will. But, certainly, as agentic AI becomes more advanced, and the models improve, and become more precise, and relevant, which is happening at an incredible pace, I would say we'd be able to see some of that.

But again, our position at Cohere would be that we want

to work -- we would develop -- because we deploy models, you know, with our customers in their environments, we would suggest that that integration on the front end with the customer and with our partners having that partnership in development, deployment, and then, you know, ultimately the decisions in how those guardrails are put in place. I think that's important on the front end of really understanding where in that loop it's necessary to have the human placed.

Senator Rounds: Mr. Tadross?

Mr. Tadross: The way that I would kind of look at this is for human in the loop. What you're sacrificing is speed over the oversight required to ensure that you're rendering it. In those cases, I think on or over the loop, it really comes down to the use case and the speed at which you have to make the decision.

So, if the use case is such in a defensive manner, similar to like a CIWS or an Aegis Cruiser, which if certain triggers are hit, you default to the machine's knowledge because the speed at which things are changing is so great that you can no longer support the decision-making process.

I think what it comes down to with that's a heuristic-based system where it's like very clear triggers to be able to implement that same type of approach with AI would require a certain amount of evaluation of those systems.

So, going back to the benchmarking question from

earlier, it would also require having a data infrastructure layer in place to be able to retrain those models effectively when the environment changes significantly. And as a result of doing that, you can ensure that this rapid iteration of retraining, and testing, and evaluation can occur that would still provide the commander the opportunity to make that informed decision about if the staff needs to be in on or over the loop.

Senator Rounds: Thank you. Mr. Mitre? And I apologize, am I saying your name correctly? Is it Mitter?

Mr. Mitre: Mitre.

Senator Rounds: Mitre.

Mr. Mitre: Mitter is fine, too, though. We get it all the time. Not a problem.

Senator Rounds: Thank you.

Mr. Mitre: Yeah, no worries, Senator. On this point, I think fundamentally what the Department of Defense is looking for are weapons systems and military systems more broadly that are effective. And so, the question is, what is effective in a particular use case in particular context?

Now, certainly as the technology progresses, there are more opportunities to use it in different ways, and along with that can come greater dependence on the technology. And with greater dependence, you potentially open up new vulnerabilities and new risks associated with that. So,

it's incredibly important to understand what are ways in which it could go sideways.

What are some of the vulnerabilities there? When you're integrating in a broader weapon system where it might act in ways that are inconsistent with human intentions, and do you have the right safeguards put in place to guard against those cases? Are there kill switches that might be necessary? Are there ways in which you're dealing with a model that's breaking out of the box and engaging more with the cyber world? Are you able to cut it off from certain applications if you need to?

I think it's helpful for the Department to think through the wide range of potential applications here, and then make sure that it's thought through how you ensure effectiveness despite different ways in which the model could react in a particular context.

Senator Rounds: Thank you. Senator Rosen.

Senator Rosen: I want to talk about energy limitations, but I'm not going to ask this as a question. I'm just going to make this as a general statement, philosophically. Because if we move to no humans in the loop, why not just create a grand video game and save lives? Because at the end of the day, if it's the AI making the choice, there's still people on the ground. All of us. Not just men and women in the military, but the rest of us that

live in the world that the computer may or may not really care too much about.

So, it's a bigger philosophical question as we move forward. Not expecting it to be answered here, but in a way, we have to be sure that we think about that because for every action these computers might take to each other, theirs versus ours, the fallout happens to us living here on earth. That's all I'm going to say. But we got to speak about living here on earth.

We got AI energy limitations. You know, a lot of data centers in Nevada. Let me tell you, there's an increasing demand for energy. They just gobble it up. And it's a hardware problem, software problem. And it's largely based of course, on the current architectures that we have.

Like I said, Nevada's dry weather and our vast open spaces that we have really become a national leader in data storage centers. Our companies are constantly innovating, but we know that the growing use of all this is going to create great energy burdens on our commercial, our government data centers.

And so, I guess we'll go this way. We'll start with Mr. Ferris. How do we address this challenge? Do you see it as a barrier to more widespread DOD and government adoption? And what research, what should we be investing in to try to maybe reduce that that great energy suck as it's

going to take everything it can, right?

Mr. Ferris: Yeah. Thank you, Senator. So, Cohere, this is actually fundamental to our company. We build custom models designed to be efficient and deployable in the environment that our clients and customers are working in. So, in pursuit of that efficiency, a couple of things. One, we're chip agnostic and cloud agnostic. So, that means we've had to focus on building our models in somewhat of a resource-constrained environment. So, we've built --

Senator Rosen: What if you put it on tanks? You've got heat, you can't -- you have to be sure that they adapt in heat environments and they're going to generate energy, right?

Mr. Ferris: Absolutely, Senators. But we've built some of these models to be deployed on as small as two GPUs or even, you know, we're pushing towards edge deployments in laptops. So, being able to bring down that energy cost, but also the infrastructure as a whole. And then, even it has implications, broadly speaking, into the supply chain as well.

Senator Rosen: Thermodynamics. Thank you. What can we do about all the energy we need to do all of this and then make it portable?

Mr. Tadross: Yes, ma'am. So, the way that I kind of look at this is as these technologies start to be fielded,

there's always an interest in the Department of Defense in order to be able to operate in a disconnected environment.

So, what that requirement's going to come along with is fine tune smaller models that can interact together, which is similar to the approach that we're taking with INDOPACOM and EUCOM for agentic warfare. So, what this really results in is a lower power requirement because back at home station, while we've been doing the development and training, we're able to tune these models. You've been using very specific data sets. So, individual models are very good at a specific thing. They've been tested and evaluated, and then the interaction between those models is what can be fielded at the edge. So, that minimizes the energy requirements as these things begin to get fielded and proliferated.

Senator Rosen: Thank you. Mr. Mitre?

Mr. Mitre: The only thing I'll add is that it's important to think about the entire tech stack to include power. Not just the data layer and compute layer. And then, the models itself and certain applications.

So, you're right to think holistically. The power is a big part of that. And certainly, there are ways to find smaller, more efficient models that you could deploy abroad along the lines of what the other panelists said. And it's worth the Department looking at that aggressively.

Senator Rosen: Thank you.

Senator Rounds: Same question for all of you now. You all work with the Department of Defense probably in different ways, but my question is, what can the Department of Defense do with regard to either policy acquisition policies the way that they treat contractors? What can they do to enhance their ability to take advantage of the private sector's capabilities that they're not doing today? Mr. Ferris.

Mr. Ferris: Thank you, Senator. The first thing we'd say is we believe that the Department needs to have an adoption mindset. We've seen a really good shift. You know, the software acquisition pathway and the use of other transaction authorities from an acquisition perspective. There are some really great strides in acquisition.

I would offer using existing mechanisms. I'm an advocate for the simple acquisition threshold being, you know, either a provision similar to what we have currently. The simple acquisition threshold is $250,000 for, you know, contracting officer can buy anything under that without a competitive process.

There's a provision for contingency operations or cyber defense and CBRN defense, where that simple acquisition threshold is raised because of urgent operational requirements. And I think similarly, we could have an

approach in procurement where for artificial intelligence, urgent operational requirements, perhaps the simple acquisition threshold could be a provision for that.

And what that would do is it would shift the burden away from, you know, the DIUs, and DARPAs, and organizations like that that are well versed in using OTAs and allow contracting officers and project managers at like much lower levels in the department to execute and acquire these types of capabilities.

Senator Rounds: Mr. Tadross?

Mr. Tadross: Thank you, Senator. So, when I think about making it easier to acquire this technology, I tend to actually go back to the AI infrastructure standpoint. The reason for that is it actually opens the barrier, reduces the barrier of entry of companies to come in. If they're able to operate off of a central data repository, then that that company's pathway to being able to create relevant technology for the Department of Defense is considerably easier than one of the legacies that have been in that space for a while and may have troves of data that they've saved over 20 years of conflict.

Senator Rounds: Thank you. Mr. Mitre?

Mr. Mitre: I agree with the panelists on everything that relates to narrow AI or AI that exists today. What I think is principally lacking from the Department's approach

to the issue is anticipating where AI might be in a couple of years' time, and really working closely with the technologists that are at the forefront of developing generative AI and frontier AI models to get their head around what that world might look like.

So, there's a lot of attention, rightfully put towards maintaining our lead in the development of technology itself to better promote its development, to better protect our lead through expert controls, and AI security, and things of that nature. But how well does the Department really understand what capabilities it may unearth in the next 2, 3, 4, 5 years, I don't know, and what that means for the future character of warfare. That's crucially important, especially as the Department now embarks on developing a new defense strategy.

Senator Rounds: One last question for all of you, and you don't have to spend a lot of time on this. But is there a place somewhere, a safe space, so to speak, where industry and DOD can actually interface and ask questions of one another, offer ideas, offer products, and so forth that is ongoing? Or is it a case-by-case basis?

In other words, if industry has a particular product that they think would be great in its application within DOD, do they know where to go to get it? And DOD on the other hand, do they have a place where they can go and ask

the questions about what do you have that can help us fix this problem? Does that exist today? Don't everybody speak at once?

[Laughter.]

Mr. Mitre: Not in a structured and systematic way, right? I think it happens in ad hoc cases here and there, but not in a coherent approach to really have a tight public-private partnership, if you will, to really understand where are we in the development of AI technologies relative to key competitors, like the Chinese, in particular, what are things that we need to be doing to make sure that America maintains that lead. DeepSeek is a great example here where surprises like that can come out and people wonder, well, what does that mean in terms of where we are?

I don't think we have that kind of environment to enable that constant flow of communication, especially when a cleared environment where you can have more sensitive conversations with key experts in terms of what's happening with this technology and what the U.S. government needs to be doing in partnership with the private sector to maintain America's lead.

Senator Rounds: Thank you. Any other thoughts?

Mr. Tadross: Yes, Senator. So, I think the closest that I've seen of that existing is Project Maven where the

efforts behind that was to bring technology into the Department of Defense in a very aggressive manner. And because they took that approach and because you had a single program that was well-funded, well organized, and manned by the right individuals, what you end up with was a situation in which they were seeking to find as many technology experts as they could bring them and figure out ways to get them into the Department to satisfy a mission requirement that was set forth.

Senator Rounds: Thank you. Mr. Ferris, anything?

Mr. Ferris: I'll just add that, you know, echo that it is very ad hoc and unstructured. However, I think that's precisely why actually, you know, people like us end up staying in these types of companies and working in them for as long as we do because it's important to know those pathways, know those venues in which these conversations do unfold, and how to get after, you know, getting in front of the government customer as quickly and rapidly as possible, especially when you do think you have something that can support the mission. So, it's a little bit at this point, it's experience for some of us where we can find that opening and get in front of the Department.

Senator Rounds: Thank you. Senator Rosen.

Senator Rosen: I have one last question. I think for those of you who don't know, Maven means "know it all" in

Yiddish, I should say. We should have the Maven marketplace. How about that? There you go. That maybe that solves what you need.

And anyway, but what I want to talk about and just finish up with, we can't do any of this without building our AI workforce. And that is something that Congress can help invest and promote. And we can only go as far as we are willing to invest in all of that. And it's just so very important.

So, for all of you, as we just finish up in our last few minutes, the workforce issues that you see in adoption of AI, what do we need to do to grow? Well, coders, engineers? All of the things that we have to do to build out this robust workforce? Because these are the kinds of things that Congress does work on and does fund. What advice would you give to us?

No one starts in the center. We started on the ends. We'll start with you. And I think it's a good way that's something that is in our wheelhouse and work on that Maven marketplace. Will you? There you go. I'm going to trademark that name. You heard it here first.

Mr. Tadross: Absolutely, Senator. So, I can say that I'm actually very, very proud of the work that we're doing in St. Louis. So, in this case, what we're doing is we're taking individuals that would normally not participate in

the national defense and give them an opportunity to support data development and AI development in the St. Louis community.

So, in some cases, what we've done is taken individuals off the fry line, train them on how to look at electro optical imagery, gotten them to the point, through training, that they are then able to look at synthetic aperture radar, get them to the point where they have a clearance, and then even elevate them even further so that they're able to pass certain imagery tests.

Senator Rosen: So, like community college certificate programs to bring people just into the workforce, or would you say even things like that, right?

Mr. Tadross: Yes, ma'am. And give them an opportunity to kind of participate in that national defense. This is an area where like Scale believes very strongly in. Kind of elevating this workforce in order to support the needs of the national defense in this space.

Senator Rosen: Yeah. Perfect. Mr. Ferris?

Mr. Ferris: Thank you, Senator. I agree. I mean, I think what we would say, we try to partner with -- you know, it's a public private partnership. That's extremely important. And workforce development is critical as part of the body of work that the Department and really the government needs to undertake to achieve the advancement in

AI that we're hoping for.

But at within the company, we do partner with educational institutions and within the community, and we're searching for ways to continue to grow that workforce. I do think it's a collaborative process that we need to take with the government and work in concert on it because, you know, from a Cohere perspective, we want to be -- you know, in terms of our deployment and how we work with our customers, it's really early on. So, we want to make sure that we're contributing to the workforce development in a way that's meaningful for the Department as time goes on.

Senator Rosen: Mr. Mitre?

Mr. Mitre: This is not exactly my area of expertise, but in my experience, there's no more compelling reason to go work in government than for the mission. So, emphasizing that is the key ability to attract top technical talent, I think is crucial, as is giving them opportunities to develop their skills.

And that requires actually having the right compute infrastructure and networking analytic tools available so that they can grow and develop their skillset while in government. That's often a challenge to bring together, but there's a broader point than just the technical talent, the AI talent skillset here as well.

Given advances in AI, it's going to impact all elements

of the workforce. And so, what we're seeing in the private sector right now, by way of analogy, is those companies that are better leveraging AI or outcompeting companies that don't have it.

And so, I think that's likely what we could see in the military context, do those militaries that are fully embracing and applying it across a range of applications are going to be at a significant advantage relative to those militaries that aren't. And so, I would think a little bit more holistically on the workforce dynamics here.

Senator Rosen: Thank you. Appreciate it.

Senator Rounds: Well, with that, let me take the opportunity to thank all three of our presenters here today; Mr. Jim Mitre, Vice-President and Director, RAND Global and Emerging Risks. Mr. David Ferris, Global Head of Public Sector, Cohere. And Mr. Dan Tadross, Head of Public Sector, Scale AI. We thank you for participating in this open discussion today that's been very, very helpful.

And my thanks also to my Vice-Chair, Senator Rosen, for participating today as well. We appreciate that. And unless you have any closing comments, I thank you for being here. Thank you for your work, and look forward to continuing to work with you and the ideas you have.

And with that, this subcommittee hearing of the Cybersecurity Subcommittee is now closed.

[Whereupon, at 4:29 p.m., the hearing was adjourned.]