S. HRG. 118–118

OPEN HEARING: ADVANCING INTELLIGENCE IN THE ERA OF ARTIFICIAL INTELLIGENCE: ADDRESSING THE NATIONAL SECURITY IMPLICATIONS OF AI

HEARING

BEFORE THE

SELECT COMMITTEE ON INTELLIGENCE

OF THE

UNITED STATES SENATE

ONE HUNDRED EIGHTEENTH CONGRESS

FIRST SESSION

SEPTEMBER 19, 2023

Printed for the use of the Select Committee on Intelligence

Available via the World Wide Web: http://www.govinfo.gov

U.S. GOVERNMENT PUBLISHING OFFICE
53–570 WASHINGTON : 2024

SELECT COMMITTEE ON INTELLIGENCE

(Established by S. Res. 400, 94th Cong. 2d Sess.)
MARK R. WARNER, Virginia, *Chairman*
MARCO RUBIO, Florida, *Vice Chairman*

DIANNE FEINSTEIN, California	JAMES E. RISCH, Idaho
RON WYDEN, Oregon	SUSAN M. COLLINS, Maine
MARTIN HEINRICH, New Mexico	TOM COTTON, Arkansas
ANGUS S. KING, JR., Maine	JOHN CORNYN, Texas
MICHAEL F. BENNET, Colorado	JERRY MORAN, Kansas
ROBERT P. CASEY, JR., Pennsylvania	JAMES LANKFORD, Oklahoma
KIRSTEN E. GILLIBRAND, New York	MIKE ROUNDS, South Dakota
JON OSSOFF, Georgia	

CHARLES E. SCHUMER, New York, *Ex Officio*
MITCH MCCONNELL, Kentucky, *Ex Officio*
JACK REED, Rhode Island, *Ex Officio*
ROGER F. WICKER, Mississippi, *Ex Officio*

WILLIAM WU, *Staff Director*
BRIAN WALSH, *Minority Staff Director*
KELSEY STROUD BAILEY, *Chief Clerk*

CONTENTS

SEPTEMBER 19, 2023

OPENING STATEMENTS

Page
Mark R. Warner, U.S. Senator from Virginia 1
Marco Rubio, U.S. Senator from Florida 28

WITNESSES

Benjamin Jensen, Ph.D., Senior Fellow, Center for Strategic and International Studies 3
  Prepared Statement 6
Yann LeCun, Ph.D., Vice President and Chief AI Scientist, Meta Platforms, Inc 11
  Prepared Statement 13
Jeffrey Ding, Ph.D., Assistant Professor of Political Science, George Washington University 18
  Prepared Statement 20

SUPPLEMENTAL MATERIAL

Response from Meta Platforms, Inc., dated January 12, 2024, to questions for the record from the Committee 58

OPEN HEARING: ADVANCING INTELLIGENCE IN THE ERA OF ARTIFICIAL INTELLIGENCE: ADDRESSING THE NATIONAL SECURITY IMPLICATIONS OF AI

TUESDAY, SEPTEMBER 19, 2023

U.S. SENATE,
SELECT COMMITTEE ON INTELLIGENCE,
Washington, DC.

The Committee met, pursuant to notice, at 2:32 p.m., in Room SH–216 in the Hart Senate Office Building, Hon. Mark R. Warner, Chairman of the Committee, presiding.

Present: Senators Warner (presiding), Rubio, Wyden, Heinrich, King, Bennet, Casey, Gillibrand, Ossoff, Cotton, Cornyn, Moran, Lankford, and Rounds.

OPENING STATEMENT OF HON. MARK R. WARNER, A U.S. SENATOR FROM VIRGINIA

Chariman WARNER. I want to welcome our witnesses:

Dr. Yann LeCun, who is the chief AI scientist at Meta, and, as I've learned, one of the real pioneers of machine learning. Dr. LeCun, it's really great to have you.

Dr. Benjamin Jensen, who is senior fellow at the Center for Strategic and International Studies and a professor at the Marine Corps University School of Advanced Warfighting. Welcome, Dr. Jensen.

And Dr. Jeffrey Ding, who is professor of Political Science at my alma mater, George Washington University, and author of the influential "ChinAI Newsletter" on China's AI landscape.

I also want to thank all of my colleagues who've been interested in this, but particularly Senator Rounds and Senator Heinrich, who have been working with Leader Schumer on a series of other AI forums, closed and open.

In many ways, the opportunity and risk of this technology that has kind of captured everyone's attention, AI, is obviously not new for this Committee. The agencies that we oversee have been some of the most innovative developers and avid adopters of advanced machine learning capabilities, working with large language models and computer vision systems, long before those terms entered the public vocabulary.

The ability to sift through and make sense of enormous amounts of data has been a hallmark of the American Intelligence Community since its inception. And the use of data science and advanced computation has been one of the core competencies of the IC for the last half century. Our Committee has been engaged on all of those topics for as long as I've been on this Committee. What has dramatically changed, however, are the potential social, political, and national security implications of this technology. And it's driven in large part by the proliferation of generative models that are both publicly accessible and incredibly capable due to a combination of unprecedented scale and breakthrough in training methods. Rapid advancements in this field have the potential to unlock enormous innovation and public benefit in areas as diverse as drug discovery, creative arts, and software programming. But as Congress evaluates the scope and significance of those transformations, we must equally grapple with the disruptions, ethical dilemmas, and potential dangers of this technology.

In both the wider Senate and through a series of roundtables I've hosted the last several months, we as a body are seeking to rise to that risk. And candidly, we were not able to do so on social media, and again, I think our consensus is we can't repeat that here. As we discuss in today's hearing, the proliferation of these technologies has dramatically lowered the barrier of entry for foreign governments to apply these tools to their own military and intelligence domains.

The public release of technical details from trained model weights to code bases of highly capable models is a boost to foreign governments, just as it is to startups, university researchers, and hobbyists. While the United States Intelligence Community has benefited from AI innovation for signal processing, sensing, machine translation, and more, so too should we now anticipate that a wider set of foreign governments will be able to also harness these tools, many of them developed and actually released by U.S. companies. But they will take those products, candidly, for their own military and intelligence uses.

Our witnesses are well positioned to describe the current posture of our Nation's most strategic rival, the People's Republic of China, as it pertains to AI. I look forward to hearing from our experts where leading PRC research labs and technology vendors are in their efforts to build cutting-edge AI models, development tools, and innovation ecosystems. In today's hearing, this Committee will focus on maintaining the U.S. Intelligence Community's edge, including how the IC can better adapt and even holistically adapt for these technologies.

I hope today's discussion can better identify some of the organizational, contracting, and technical barriers to achieving these objectives. I'm also eager to hear where these tools currently fall short of some of the loftiest claims about their capabilities—something Dr. LeCun and I have talked about. For instance, the propensity of even the most advanced language models to hallucinate raises serious questions about their fitness in mission-critical and other sensitive areas. In the intelligence domain, mistakes can impact our Nation's security, the privacy of Americans, and the clarity of pivotal foreign policy discussions.

To be sure, as well, generative models can improve cybersecurity, helping programmers identify coding errors and contributing towards safer coding practices. But with that potential upside, there's also a downside since these same models can just as readily assist malicious actors. I hope this hearing will explore the ways in which generative models alter the cyber landscape, lowering the barrier to entry for formerly second-tier cyber powers, and how in the cyber domain, AI can advance the capabilities of more advanced state actors. Lots to discuss.

Now, as the leading body in Congress in tracking disinformation, market manipulation, and election influence efforts by our Nation's adversaries, our Committee is also deeply interested in the ways in which AI expands and exacerbates the threat of foreign malign influence. The ability for foreign actors to generate hyper-realistic images, audio, and videos will undeniably make it harder for Americans to navigate our ever more complex, fraught, and fast paced media environment.

We must also contend with the ways in which bad actors will use these tools to undermine trust in markets, public institutions, and public health systems. These tools will greatly challenge our society's ability to agree on baseline facts and our already impaired ability to develop consensus. The last several years have amply demonstrated the ways in which speed, scale, and excitement associated with new technologies have frequently obscured the shortcomings of their creators in anticipating the harmful effects of their use.

I hope we will also queue up a discussion of how the U.S. can best harness and govern these technologies to avoid the same mistakes we made in failing to foresee vulnerabilities in other global-scale technologies like social media. To that end, I hope we'll also touch upon how other countries, both allies and rivals, are coping with potential disruptions, risks, and economic dislocations of AI technologies through their own regulatory proposals. And I know I hope I hear from Dr. Ding on what actually the PRC is doing in this field. AI capabilities, I think we all know, hold enormous potential. However, we must make sure that we think about that potential, gain the upside, but where appropriate, put safeguards in place. I look forward to today's discussion.

I did an extra-long opening today because the Vice Chairman has been held up for a moment. He will join, and when he joins us, after the presentations, to allow him to make an opening comment. And because, in our tradition, in open hearings, we will go by rank of seniority for five-minute rounds.

With that, I'm not sure who drew the long straw or the short straw in terms of opening comments, but I'll turn it over to our panel.

Thank you.

STATEMENT OF BENJAMIN JENSEN, PhD, SENIOR FELLOW, CENTER FOR STRATEGIC AND INTERNATIONAL STUDIES, AND PROFESSOR, MARINE CORPS UNIVERSITY, SCHOOL OF ADVANCED WARFIGHTING

Dr. JENSEN. Short straw, Senator.

Chairman Warner, Vice Chairman Rubio, distinguished Members of the Committee, I really am honored today to sit with you and share my thoughts on what I think you all agree, from reading all the work that you've done on it, is probably the most important question facing our Nation from a technological perspective.

The magnitude of the moment is clear, right? Both the Senate and the House are very much cultivating a national dialogue, and I just want to open as a citizen by thanking you for that. You have a powerful role in that, and so doing this here right now is key. And so, I have to be very blunt and clear with you that I'm going to talk to you less about the threat outside, Sir. I'm going to talk to you more about how I think we could get it wrong.

Today, as part of that ongoing dialogue. I really want to look at the often-invisible center of gravity behind any technology: people, bureaucracy, and data, and in particular, data infrastructure and architecture. Put simply, you get the right people in place, with permissive policies and computational power at scale, and you gain a position of advantage in modern competition. I'll just put it bluntly. In the twenty-first century, the general or spy who doesn't have a model by their side is basically as helpless as a blind man in a bar fight.

So, let's start with people. Imagine a future analyst working alongside a generative AI model to monitor enemy cyber capabilities. The model shows the analyst signs of new adversary malware in targeting U.S. critical infrastructure. The analyst disagrees. The challenge we have is today our analysts can't explain why they disagree because they haven't been trained in basic data science and statistics. They don't know how to balance causal inference and decades of thinking about human judgment and intuition. And sadly, I'll be honest, our modern analytical tradecraft and even something close to me, professional military education, tends to focus on discrete case studies more than statistical patterns or trend analysis. In other words, if we unleash a new suite of machine learning capabilities without the requisite workforce to understand how to use them, we're throwing good money after bad. And we really have to be careful about this. I can't stress this enough. If you don't actually make sure that people understand how to use the technology, it's just a magic box.

Let's turn to the bureaucracy.

Now, I want you to imagine what the Cuban Missile Crisis would look like in 2030: all sides with a wide range of machine learning capabilities. There would be an utter tension as machines wanted to speed up decision-making in the crisis. But senior decision-makers needed to slow it down to the pace of interagency collaboration. Even worse, you would be overwhelmed by deep fakes and computational propaganda pressuring you as elected officials and any senior leader to act. And pressure to act at a moment of crisis doesn't necessarily lead for sound decision-making. Unfortunately, neither our modern national security enterprise nor the bureaucracy surrounding government innovation/experimentation are ready for this world. If the analyst and military planner struggles to understand prediction, inference, and judgment through algorithms, the challenge is even more acute with senior decision-makers.

At this level, most international relations and diplomatic history tells us that the essence of decision is as much emotion, flawed analogies, and bias as it is rational interests. What happens when irrational humans collide with rational algorithms during a crisis?

Confusion could easily eclipse certainty, unleashing escalation and chaos.

There are even larger challenges associated with creating a bureaucracy capable of adapting algorithms during a crisis. Because of complexity and uncertainty, all models require a constant stream of data to the moment at hand, not just the moment of the past. But crises are different than necessarily what preceded them. Unfortunately, slow adapters will succumb to quick deaths on that future battlefield. As a result, a culture of experimentation and constant model refinement will be the key to gaining and sustaining relative advantage. Now, ask yourself, do we have that bureaucracy?

Last, consider data, architecture, and infrastructure, how we put the pieces of the puzzle together. I want you to imagine we're almost back to the old days of the SCUD hunt, right? Imagine the hunt for a mobile missile launcher in a future crisis. A clever adversary, knowing they were being watched, could easily poison the data used to support intelligence analysis and targeting. They could trick every computer model into thinking a school bus was a missile launcher, causing decision-makers to quickly lose confidence in otherwise accurate data. Even when you were right 99 percent of the time, the consequences of being wrong once are still adding unique human elements to crisis decision-making. Artificial intelligence and machine learning, therefore, are only as powerful as the underlying data. Yet to collect, process, and store that data is going to produce significant costs going forward. This is not going to be cheap. Furthermore, bad bureaucracy and policy can kill great models if they limit the flow of data.

Last, let's talk about the fact that Prometheus has already shared the fire, and I think you all know that even from your opening comments, Chairman. Adversaries now into the foreseeable future can attack us at machine speed through a constant barrage of cyber operations and more disconcerting, mis-, dis- and mal-information, alongside entirely new forms of swarming attacks that could hold not just our military, but our civilian population at risk. Unless the United States is able to get the right mix of people, bureaucratic reform, and data infrastructure in place, those attacks could test the very foundation of our Republic.

Now, I'm an optimist, so I'm going to be honest with you. I'm confident the United States can get it right. In fact, the future is ours to lose. Authoritarian regimes are subject to contradictions that make them rigid, brittle, and closed to new information. Look no further than regulations about adherence to socialist thought in data sets. These regimes are afraid to have the type of open, honest dialogue this Committee is promoting. And that fear is our opportunity.

Thank you for the opportunity to testify.

[The prepared statement of the witness follows:]

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

Statement before the
Senate Select Committee on Intelligence

"Addressing the National Security Implications of Artificial Intelligence: People, Bureaucracy, and Data Infrastructure"

A Testimony by:

Dr. Benjamin Jensen

Senior Fellow, International Security Program, CSIS and
Professor, Marine Corps University, School of Advanced Warfighting

September 19, 2023
216 Hart Building

CSIS and Marine Corps University do not take policy positions, so the views represented in this testimony belong to Benjamin Jensen and not those of his employers.

Chairman Warner, Vice-Chair Rubio, distinguished Members of the Committee, I am honored to share my views with you on what might be the central intelligence question facing our nation: *how does artificial intelligence affect national security*? The magnitude of the moment is clear, and both the Senate and the House are embracing their responsibility to create a national dialogue. As a citizen I thank you for that.

Today as part of this ongoing dialogue, I ask you to consider the often-invisible center for gravity for integrating new algorithms and enduring aspects of military theory and intelligence tradecraft. That center for gravity rests not just in lines of code, but in the *people*, the *bureaucracy* and the *data infrastructure* that turns any technology into strategic advantage.[1] Get the right people in place with permissive policies and provide them access to computational capabilities at scale and you gain a position of advantage in modern competition. Deny your adversaries the ability to similarly wage algorithmic warfare and you turn this advantage into enduring strategic asymmetry.

Artificial intelligence and machine learning (AI/ML) will be a critical capability for the nation going forward and central to integrated deterrence campaigns and warfighting. The general or spy who doesn't have a model by their side in the 21st century will be blind man in a bar fight. Yet, that critical capability – strategic competition and war at machine speed directed by human judgment – rests on critical requirements. Our intelligence community and military need rank and file members who understand basic data science and coding. They need a smaller, nimble information age bureaucracy open experimentation in place of the labyrinth of middle managers and policies that stifle innovation. And they need reliable access to data centers to continually train and update machine learning models against adversaries. Failing to protect these requirements risks ceding the initiative to our adversaries.

People

Imagine a future analyst working alongside an AI model to monitor People's Liberation Army (PLA) cyber capabilities. The model shows the analyst signs of new adversary malware targeting U.S. critical infrastructure. The analyst disagrees. But the analyst cannot explain why they disagree because they haven't been trained in basic data science, statistics, and the foundations of AI/ML. It's the equivalent of a lawyer who never went to law school arguing a case.

Sadly, modern analytical tradecraft and even professional military education tend to focus more on discrete cases more than statistical patterns and trends. There is a tendency to treat technology like magic. As a result, model outputs are either sacred or evil creating a risk of skewed inferences across the national security enterprise. There is little to no discussion about the tradeoffs between model interpretability and accuracy, a critical task in national security crises prone to uncertainty and deception.[2] In other words, unleashing a new suite of AI/ML tools inside the national security enterprise will produce diminishing returns unless we retrain the workforce and teach them how to use model-generated insights to refine human judgment.

---

[1] Benjamin Jensen, Scott Cuomo, Christopher Whyte. *Information in War: Military Innovation, Battle Networks, and the Future of Artificial Intelligence* (Washington: Georgetown University Press, 2022).

[2] Giorgos Myrianthous. "Understanding The Accuracy-Interpretability Trade-Off" *Towards Data Science* October 6, 2021 <https://towardsdatascience.com/accuracy-interpretability-trade-off-8d055ed2e445> .

Our adversaries face the same challenge. The intelligence and military profession need a paradigm shift if they are going to take full advantage of AI/ML. The good news is that despite "precision recruitment" efforts to attract college students, the PLA struggles to integrate and retain them in its formations.[3] The smart kids in China aren't rushing to join the army. Russian tech workers fled the country to avoid fighting an unjust war in Ukraine.[4] The bad news is that the knowledge required to retrain spies and soldiers is open and accessible even to non-state actors. The first actor to embrace the paradigm shift in military art and analytical tradecraft could gain a generational strategic advantage.

Bureaucracy

Imagine what the Cuban Missile Crisis would look like in 2030 with all sides using a wide range of AI-applications ranging from imagery recognition to logistics management and generative analysis of adversary intentions. There would be a tendency to speed up the crisis even when it might make more sense to slow down decision-making and be more deliberate. Computational propaganda and tailored media would increase public pressure on political officials. At a more technical level, there would be a need to constantly adjust and recalibrate models as adversaries shifted their tactics, techniques, and procedures and both sides operated outside of the norm. Crisis events are by definition outliers creating challenges for statistical analysis. Confusion could eclipse certainty unleashing escalation and chaos.

Unfortunately, neither our modern national security enterprise nor the bureaucracy surrounding government innovation and experimentation are ready for this world. If the rank-and-file analyst and military planner struggles to understand prediction, inference, and judgment in and through algorithms, the challenge is even more acute amongst senior decision makers. At this level, most international relations literature and diplomatic history show us that the essence of decision is as much emotion, flawed analogies, and bias as it is rational interests defined by power or the structure of the international system.[5]

There are even larger challenges with creating a bureaucracy capable of adjusting algorithms to match new contexts during a crisis. Because of complexity and uncertainty, all models will require a constant stream of data and updates to their weighting. The speed of update will dictate the terms of advantage. Slow adapters will succumb to quick deaths on the future battlefield. The side with

---

[3] Marcus Clay, Dennis Blasko, and Roderick Lee "People Win Wars: A 2022 Reality Check on PLA Enlisted Force and Related Matters" *War on the Rocks* August 12, 2022 < https://warontherocks.com/2022/08/people-win-wars-a-2022-reality-check-on-pla-enlisted-force-and-related-matters/>

[4] Gian M. Volpicelli "Russia is Facing a Tech Worker Exodus" *Wired* March 23, 2022 < https://www.wired.com/story/russian-techies-exodus-ukraine/>; Masha Borak "How Russia Killed Its Tech Industry" *MIT Technology Review* April 4, 2023 <https://www.technologyreview.com/2023/04/04/1070352/ukraine-war-russia-tech-industry-yandex-skolkovo/>

[5] Keren Yarhi-Milo. *Who Fights for Reputation? The Pscyhology of Leaders in International Conflict* (Princeton: Princeton University Press, 2018); ibid *Knowing the Adversary: Leaders, Intelligence Organizations and Assessments of Intentions in International Relations* (Princeton: Princeton University Press, 2014); Robert Jervis. *How Statemen Think: The Psychology of International Politics* (Princeton: Princeton University Press, 2017); Yuen Foong Khong. *Analogies at War: Munich, Dien Bien Phu, and the Vietnam Decisions of 1965* (Princeton: Princeton University Press, 1992); Hans Morgenthau. *Politics Among Nations* (New York: Knopf, 1948); Kenneth Waltz. Theory of International Politics (New York: McGraw-Hill, 1979).

the model that updates faster than the enemy will generate tempo and freedom of action. A culture of experimentation and constant model refinement and adjustment will be the key to gaining and sustaining relative advantage in the future.

In fact, Ukraine has shown us this truth. A network of civilian software engineers, non-profits, and soldiers built the Delta platform and constantly refined their ability to use AI on the battlefield while waging a war of survival.[6] They could move fast because they were willing to experiment and fail. Ask yourself, do we have a similar culture of experimentation across our military and intelligence organizations?

Data Instructure

Imagine the hunt for mobile missile launchers in a future crisis. A clever adversary knowing they were being watched could poison the data used to support intelligence analysis and military decision making. They could trick every computer model into thinking a school bus was a transporter erector launcher (TEL) causing decision makers to lose confidence in otherwise accurate model-generated insights. Even when you are right 99% of the time, the consequences of being wrong once can still add unique, human elements to rational decision making and risk assessments.[7]

AI/ML is only as powerful as the underlying data. The larger and more diverse the dataset, the more opportunities there are for analyzing it in higher dimensions. Instead of an X and Y axis we all learned in geometry there might be thousands of matrix vectors. Each new dimension allows the model to identify signatures buried in the data. Each new signature is a potential intelligence advantage.

Yet, to collect, process, and store that data will produce significant costs going forward. First, it will mean an increase in the number of intelligence collection missions required to capture data using both open source and more sensitive methods. It will require clear data labeling and architecture standards to make it easy to compare diverse inputs. Bad bureaucracy and policy can kill great models if they limit the flow of data. And it will require access to computational power at scale as analysts move to adjust their models during a fluid crisis and in the face of clever adversaries using the equivalent of digital terracotta armies to poison AI/ML intelligence models.

---

[6] "Ukraine to introduce Delta situational awareness system for military" *The Kyiv Independent* February 4, 2023 < https://kyivindependent.com/government-introduces-nato-standard-delta-management-defense-system/>; Julian Borger "Our weapons are computers: Ukrainian coders aim to gain battlefield edge" *The Guardian* December 18, 2022 < https://www.theguardian.com/world/2022/dec/18/our-weapons-are-computers-ukrainian-coders-aim-to-gain-battlefield-edge>.

[7] Karma Dajani and Sjoerd Dirksin. *A Simple Introduction to Ergodic Theory* (2009) < https://webspace.science.uu.nl/~kraai101/lecturenotes2009.pdf>; Merle van den Akker "Ergodicity: What Does It Mean for Behaviorial Science?" *Money on the Mind* September 9, 2021 < https://www.moneyonthemind.org/post/ergodicity-what-does-it-mean-for-behavioural-science>; and Pete Combe II, Benjamin Jensen and Adrian Bogart. *Rethinking Risk in Great Power Competition* (Washington: Center for Strategic and International Studies, 2023) < https://www.csis.org/analysis/rethinking-risk-great-power-competition>.

Algorithmic warfare is not static. The only good models will be ones that continually update based on a constant flow of data. In fact, the data centers required to make these adjustments will become prime targets for cyber operations and even kinetic strikes in future wars.

The authoritarian nations challenging United States have an advantage in centralizing and controlling data. This bureaucratic centralization gives them the ability to focus resources and test different models and AI/ML applications. For example, China uses a centralized planning model to promote AI development for everything from economic growth to domestic surveillance and military modernization.[8] At the same time, this centralization makes the system brittle. Closed systems are more focused and secure, but they struggle to learn. Yet, learning is the name of the game in artificial intelligence. The balance between open and closed approaches to AI and national security will have to grapple with this tradeoff between the adaptability of open systems and the security of closed system architectures.

Conclusion

Prometheus has already shared the fire. Adversaries now and into the foreseeable future will attack us at machine speed through a constant barrage of cyber operations, mis/dis/mal information as well as entirely new forms of anti-access/area denial kill webs that fuse open source and sensitive intelligence to direct swarm attacks at civilian and military targets.[9] Unless the United States is able to get the right mix of people, bureaucratic reform, and data infrastructure those attacks will test the very foundation of the Republic.

I am confident the United States can get it right. In fact, the future is ours to lose. Authoritarian regimes are subject to contradictions that make them rigid, brittle, and closed to new information. Look no further than Chinese generative AI regulations that require an adherence to socialist thought.[10] These regimes are afraid to have the type of open, honest dialogue this committee is promoting. This fear is our opportunity. Creating a vibrant marketplace of ideas will help calibrate the right mix of regulation to protect the critical requirements the United States needs to integrate AI into the national security enterprise.

---

[8] William Carter and William Crumpler. *Smart Money on Chinese Advances in AI* (Washington: Center for Strategic and International Studies, 2019); Nicholas Wright (ed) *Artificial Intelligence, China, Russia, and the Global Order* (Montgomery: Air University Press, 2019); Elsa Kania. *Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power* (Washington: Center for New American Security, 2017).

[9] Brandon Valeriano, Benjamin Jensen, and Ryan Maness. *Cyber Strategy: the Evolving Character of Power and Coercion* (New York: Oxford University Press, 2018); Charles Cleveland, Benjamin Jensen, Arnel David, and Susan Bryant. *Military Strategy in the 21st Century: People, Connectivity, and Competition* (New York: Cambria Press, 2018); Philip N. Howard and Samuel Woolley. *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media* (New York: Oxford University Press, 2018). John Arquilla and David Ronfeldt. *Swarming and the Future of Conflict* (Santa Monica: RAND Corporation, 2000); Sean J.A. Edwards. *Swarming and the Future of Warfare* (Santa Monica: RAND Corporation, 2005); Bryan Clark, Dan Patt, and Harrison Schramm. *Mosaic Warfare: Exploring Artificial Intelligence and Autonomous Systems to Implement Decision-Centric Operations* (Washington: Center for Strategic and Budgetary Analysis, 2020); and Benjamin Jensen and John Paschkewitz. "Mosaic Warfare: Small and Scalable are Beautiful" *War on the Rocks* December 23, 2019 <https://warontherocks.com/2019/12/mosaic-warfare-small-and-scalable-are-beautiful/>.

[10] Meaghan Tobin "China announces rules to keep AI bound by 'core socialist values'" *Washington Post* July 14, 2023 < https://www.washingtonpost.com/world/2023/07/14/china-ai-regulations-chatgpt-socialist/>.

STATEMENT OF YANN LeCUN, PhD, VICE PRESIDENT AND CHIEF AI SCIENTIST, META PLATFORMS, AND SILVER PROFESSOR OF COMPUTER SCIENCE AND DATA SCIENCE, NEW YORK UNIVERSITY

Dr. LECUN. Chairman Warner, Vice Chairman Rubio, and distinguished Members of the Committee. Thank you for the opportunity to appear before you today to discuss important issues regarding AI.

My name is Yann LeCun. I'm currently the Silver Professor of Computer Science and Data Science at New York University. I'm also Meta's Chief AI Scientist and co-founder of Meta's Fundamental AI Research Lab. At Meta, I focus on AI research, development strategy, and scientific leadership.

AI has progressed leaps and bounds since I began my research career in the 1980s. Today, we are witnessing the development of generative AI, and in particular, large language models. These systems are trained through self-supervised learning. Or more simply, they are trained to fill in the blanks. In the process of doing so, those AI models learn to represent text or images—including the meaning, style, and syntax—in multiple languages. The internal representation can then be applied to downstream tasks such as translation, topic classification, et cetera. It can also be used to predict the next words in a text, which allow LLMs to answer questions or write essays, and write code as well. It is important not to undervalue the far-reaching potential opportunities they present. The development of AI is as foundational as the creation of the microprocessor, the personal computer, the Internet, and the mobile device. Like all foundational technologies, there will be a multitude of uses of AI. And like every technology, AI will be used by people for good and bad ends.

As AI systems continue to develop, I'd like to highlight two defining issues. The first one is safety, and the second one is access. One way to start to address both of these issues is through the open sharing of current technology and scientific information. The free exchange of scientific papers, code, and trained models in the case of AI has enabled American leadership in science and technology. This concept is not new. It started a long time ago. Open sourcing technology has spurred rapid progress in systems we now consider basic infrastructure, such as the Internet and mobile communication networks.

This doesn't mean that every model can or should be open. There is a role for both proprietary and open-source AI models. But an open-source basic model should be the foundation on which industry can build a vibrant ecosystem. An open-source model creates an industry standard, much like the model of the Internet in the mid '90s. Through this collaborative effort, AI technology will progress faster, more reliably, and more securely.

Open sourcing also gives businesses and researchers access to tools that they could not otherwise build by themselves, which helps create a vast social and economic set of opportunities. In other words, open sourcing democratizes access. It gives more people and businesses the power to build upon state-of-the-art technology and to remedy potential weaknesses. This also helps promote democratic values and institutions, minimize social disparities, and improve competition. We want to ensure that the United States and American companies, together with other democracies, lead in AI development ahead of our adversaries, so that the foundational models are developed here and represent and share our values. By open sourcing current AI tools, we can develop our research and development ecosystem faster than our adversary.

As AI technology progresses, there is an urgent need for governments to work together, especially democracies, to set common AI standards and governance models. This is another valuable area where we welcome working with regulators to set appropriate transparency requirements, red teaming standards, and safety mitigations to help ensure those codes of practice, standards, and guardrails are consistent across the world. The White House's voluntary commitment is a critical step in ensuring responsible guardrails, and they create a model for other governments to follow. Continued U.S. leadership by Congress and the White House is important in ensuring that society can benefit from innovation in AI while striking the right balance with protecting rights and freedom, preserving national security interests, and mitigating risks where those arise.

I'd like to close by thanking Chairman Warner, Vice Chairman Rubio, and the other Members of the Committee for your leadership. At the end of the day, our job is to work collaboratively with you, with Congress, with other nations, and with other companies in order to drive innovation and progress in a manner that is safe and secure and consistent with our national security interests.

Thank you. I look forward to your questions.

[The prepared statement of the witness follows:]

HEARING BEFORE
THE UNITED STATES SENATE SELECT COMMITTEE ON INTELLIGENCE

September 19, 2023

Testimony of Yann LeCun
Chief AI Scientist, Meta

Introduction

Chairman Warner, Vice Chairman Rubio, and distinguished members of the Committee, thank you for the opportunity to appear before you today to discuss important issues regarding AI. My name is Yann LeCun. I am currently the Silver Professor of Computer Science and Data Science at New York University (NYU). I am also Meta's Chief AI Scientist and co-founder of Meta's FAIR (Fundamental AI Research), where I focus on AI research, development strategy, and scientific leadership.

Overview of My Involvement in AI

I have worked in areas related to computer science, machine learning, and artificial intelligence for several decades. In 1987, I received my PhD in Computer Science from Universite Pierre et Marie Curie (known as the Sorbonne). Following a postdoctoral research position at the University of Toronto, I joined AT&T Bell Labs in 1988, where I focused on developing machine learning methods, including an image recognition method called convolutional neural networks. Over the last decade, convolutional networks have become the dominant method for image understanding in such applications as driving assistance, medical image analysis and optical character recognition, among many other domains.

Since 2003, I have been privileged to serve as a professor at NYU. In 2014, my colleague Yoshua Bengio and I established the research program on Learning in Machines & Brains (formerly Neural Computation and Adaptive Perception Program), sponsored by the Canadian Foundation for Advanced Research, which I continue to advise today.

I have long advocated for the responsible and ethical use of AI. That is why, alongside representatives from Google, Microsoft, Amazon, and IBM, I co-founded *Partnership on AI* in 2016, a non-profit that brings together academic, civil society, industry, and media organizations to address the most important and difficult questions concerning AI. Together, *Partnership on AI* conducts studies, shares insights, publishes guidelines, informs public policy, and advances public understanding of AI.

I have been fortunate to have received various recognitions for my contributions to the field of AI. Most notably, along with my colleagues Geoffrey Hinton and Yoshua Bengio, I received the 2018 ACM Turing Award for "conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing."

As evidenced by my experiences in academia, I care deeply about AI research and collaboration. In 2013, I joined Meta (then Facebook) to create and lead an AI research division known as FAIR. I built and directed FAIR until 2018, when I moved to my current role as Chief AI Scientist. In this role, I establish directions for AI research, advise the company on AI strategy, and provide thought leadership to the broader AI research and engineering communities.

I am proud of Meta's work to serve as a leader in AI and its contributions to help ensure that America continues to lead in developing this critical technology. While my work at Meta does not involve developing products or policies, I engage with dedicated colleagues who work tirelessly toward developing and deploying AI in ways that enhance economic and social benefits, while also executing on guardrails that anticipate and mitigate potential risks. A central part of Meta's mission has always been to connect people to each other. That's still the core of what we do, and we are consistently working to improve and to share our innovations and lessons learned in AI with each other and with others including companies in the United States and around the world.

Current State of AI

AI has progressed leaps and bounds since I began my research career in the 1980s. We've seen first-hand how making AI models available to researchers can reap enormous benefits. For example, AI is being used to translate hundreds of languages, reduce traffic collisions, detect tumors in x-rays and MRIs, speed up MRI exams by a factor of four, discover new drugs, design new materials, predict weather conditions, and help the visually impaired.

Society's ability to develop AI tools to defend against adversarial, nefarious, or other harmful content derives in large part from our social values. Meta, by way of example, has organized its responsible AI efforts around five key pillars reflecting these values:

- First, we believe that protecting the privacy and security of individuals' data is the responsibility of everyone and we have therefore established a cross-product Privacy Review process to assess privacy risks;
- Second, we believe that our services should treat everyone fairly and have developed processes to detect and mitigate certain forms of statistical bias;
- Third, we believe that AI systems should be robust and safe, which is why we established an AI Red Team to test our systems against adversarial threats to ensure that they behave safely and as intended even when they are subjected to attack;
- Fourth, we are striving to be more transparent about when and how AI systems are making decisions that impact the people who use our products, to make those decisions more explainable, and to inform people about the controls they have over how those decisions are made;
- Finally, we believe that we should be accountable for our AI systems and the decisions they make, so we have built governance systems to ensure we meet high standards.

Today, we are witnessing rapid advancements in the development of generative AI, and in particular large language models (LLMs). These systems are trained through self-supervised learning, or more simply, they are trained to fill in the blanks. In the process of doing so, the AI model learns to represent text, including the meaning, style, and syntax, in multiple languages. This internal representation can be applied to downstream tasks, such as translation and topic classification. It can also be used to predict the next words in a text, which allows LLMs to answer questions or write essays.

There's no question that people in the field, including me, have been surprised by how well LLMs have worked. Though they have great promise and potential, it is important to keep in mind that LLMs do have limitations. While LLMs can unlock a host of new possibilities in industries, from health care to logistics to manufacturing, they have limited abilities to reason, a prevalent feature of human intelligence. As the technology exists today, even the most powerful AI systems are quite far from approximating human intelligence. For example, a child can readily learn to clear the dinner table and fill the dishwasher, but AI tools have not reached the level of domestic robots. The language fluency of LLMs may suggest human-level intelligence, but it is far from that.

Future of AI & The Importance of Open Sourcing

The current generation of AI tools is different from anything we've had before, and it's important not to undervalue the far-reaching potential opportunities they present. However, like any new disruptive technology, advancements in AI are bound to make people uneasy. I can understand why. The development of AI is as foundational as the creation of the microprocessor, the personal computer, the Internet, and the mobile device. Like all foundational technologies, there will be a multitude of uses of AI technologies, some predictable and some less so, which can be alarming. And like every technology, AI will be used by people for good and bad ends. This will not be the first time that bad actors try to use developing technology to their own ends, like phishing scams or spreading misinformation.

As AI systems continue to develop, I'd like to highlight two defining issues. The first is safety. New technology brings new challenges, and everyone has a part to play here. Companies should make sure tools are built and deployed responsibly. And all of us – policymakers, academics, civil society and industry – should work together to maximize the potential benefits and minimize the potential risks. The second is access. Having access to state of the art AI will be an increasingly important driver of opportunity in the future for individuals, for companies, and for economies as a whole.

One way to start to address both of these issues is through the open sharing of current technologies. At Meta, we believe it is better if AI is developed openly, rather than behind closed doors by a handful of companies. Generally speaking, companies should collaborate across industry, academia, government, and civil society to help ensure that such technologies are developed responsibly and with openness to minimize the potential risks and maximize the potential benefits.

The concept of free code sharing in the technological ecosystem is not new, it started long ago. The 1950's and 1960's saw almost all software produced by academics and corporate research labs like AT&T's Bell Labs working in collaboration. Companies like IBM, DEC, and General Motors set up user groups to facilitate sharing code among the users. The infrastructure of the internet and all cloud computing services run on open-source software (Linux, Apache, MySQL, JavaScript, for example), as do most web browsers and many of the apps we use every day.

An open source foundation is central to addressing both defining issues. It allows researchers and developers to test the systems, building better and safer products resulting in faster innovation and creating a flourishing market. That doesn't mean that every model can or should be open sourced. There's a role for both proprietary and open source AI models. But, giving businesses and researchers access to tools that would be challenging to build themselves, backed by computing power they might not otherwise access, can create vast social and economic opportunities. In other words, open sourcing democratizes access – it gives more people and businesses the power to access and test state-of-the-art technology to identify potential vulnerabilities, which can then be mitigated in a transparent way by an open community. An open source based model on top of which an industry can be built creates a vibrant ecosystem. Rather than having dozens of companies building many different AI models, an open source model creates an industry standard, much like the model of the Internet in 1992. Through this collaborative effort, AI technology will progress faster, more reliably, and more securely. It is no coincidence that so much leading work in AI is being done here in the United States, from foundational research to real world products. We have a dynamic economy of which the tech sector is a major part. Talented people want to build new things here and that helps our global competitiveness.

While AI will drive progress everywhere, I expect there will be particular benefits that accrue to the United States over time. We want to ensure that the United States and American companies lead in AI development, ahead of our competitors and adversaries, so that the foundational models are developed here and represent and share our values. By open sourcing current AI tools, we can develop and improve the foundational models faster than others – including potential adversaries – will be able to access and build on those models. With US leadership, we can cultivate this powerful technology based on our values, rather than relinquishing it to our adversaries. Leading the AI research and development effort puts us in a strong position to enhance the safety of our systems and to warn about potential risks.

Recommendations for What's Next

As AI technology progresses, there is an urgent need for governments to work together, especially democracies, to set common AI standards and governance models, with a focus on the spaces where there are gaps within existing regulation and frameworks. This is another valuable area where we welcome work with regulators to set appropriate transparency requirements, red teaming standards, and safety mitigations – and help ensure those codes of practice, standards, and/or guardrails are consistent across the world.

The White House's voluntary commitments are a critical step in ensuring responsible guardrails are established and they create a model for other governments to follow. We joined these commitments because they represent an emerging industry-wide consensus around the things that we have been building into our products for years. We believe these commitments strike a reasonable balance of addressing today's concerns and convening industry to address the potential risks of the future. They enable the tremendous potential for AI while focusing on the greatest risks.

Continued US leadership by Congress and the White House is important in ensuring that there is a considered, collaborative approach to the regulation of AI so that society can benefit from innovation in AI while striking the right balance with protecting rights and freedoms, preserving national security interests and mitigating risks, where those arise. US-led frameworks for approaching these issues would help drive toward a global consensus that doesn't yet exist, and provide alternatives to approaches that are designed to curtail American innovation. As with other technological shifts, the government has an important role to play. The fact that Congress is willing to engage on these issues encourages us that guardrails will be put in place that help to ensure AI is developed and utilized in a way that promotes innovation and spreads the economic and social benefits, while anticipating and mitigating potential risks.

Conclusion

I'd like to close by thanking Chairman Warner, Vice Chairman Rubio, and the other members of the Committee for your leadership. At the end of the day, our job is to work collaboratively with you, with Congress, with other nations, and with other companies in order to drive innovation and progress, in a manner that is safe and secure and consistent with our national security interests.

We appreciate your attention to these important issues and look forward to continuing to work to find ways we can continue to improve our AI tools, processes, and collaborations.

Thank you, and I look forward to your questions.

STATEMENT OF JEFFREY DING, PhD, ASSISTANT PROFESSOR OF POLITICAL SCIENCE, GEORGE WASHINGTON UNIVERSITY

Dr. DING. Chairman Warner, Vice Chairman Rubio, and Members of the Committee. I am honored by the opportunity to brief this Committee on the National Security Implications of AI. In all honesty, I also have a selfish reason for attending today. I teach political science at GW, and my students all really look up to the Committee Members in this room and also all the staff who are working behind the scenes to put this hearing together. So, when I got to tell the class this morning that I was doing this testimony, they all got the most excited I've ever seen them get excited this semester. And so, hopefully, that will cause them to do more of the required readings in class. In all seriousness, I have great students, I'm very grateful to be here.

Today, in my opening remarks, I want to make three main points from my written testimony. The first is when it comes to the national security of implications of AI, the main driver and the main vector is which country will be able to sustain productivity growth at higher levels than their rivals. And for this vector, the distinction between innovation capacity and diffusion capacity is central to thinking about technological leadership in AI. Today, when various groups—whether that be experts, policymakers, the Intelligence Community—when they try to assess technological leadership, they are overly preoccupied with innovation capacity. Which state is going to be the first to generate new-to-the-world breakthroughs, the first to generate that next leap in large language models. They neglect diffusion capacity. A state's ability to spread and adopt innovations after their initial introduction across productive processes.

And that process of diffusion throughout the entire economy is really important for technologies like AI. If we were talking about a sector like automobiles, or even a sector like clean energy, we might not be talking as much about the effects of spreading technologies across all different productive processes throughout the entire economy. AI is a general-purpose technology, like electricity, like the computer, like my fellow panelists just mentioned in his testimony. And general-purpose technologies historically precede waves of productivity growth because they can have pervasive effects throughout the entire economy. So, the U.S. in the late 19th century became the leading economic power before it translated that influence into military and geopolitical leadership, because it was better at adopting general purpose technologies at scale, like electricity, like the American system of interchangeable manufacture, at a better and a more effective rate than its rivals.

Point number two is when we assess China's technological leadership and use this framework of innovation capacity versus diffusion capacity, my research finds that China faces a diffusion deficit. Its ability to diffuse innovations like AI across the entire economy lags far behind its ability to pioneer initial innovations or make fundamental breakthroughs.

And so, when you've heard from other people in the past or in the briefing memos you are reading, you are probably getting a lot of innovation-centric indicators of China's scientific and technological prowess: its advances in R&D spending, headline numbers on patents and publications. In my research, I've presented evidence about China's diffusion deficit by looking at how is China actually adopting other information and communications technologies at scale? What are its adoption rates in cloud computing, industrial software, related technologies that would all be in a similar category to AI? And those rates lag far behind the U.S.

Another indicator would be how is China's ability to widen the pool of average AI engineers? I'm not talking about Nobel Prize of computing winners like my fellow panelists here, but just average AI engineers who can take existing models and adapt them in particular sectors or industries or specific applications. And based on my data, China has only 29 universities that meet a baseline quality metric for AI engineering, whereas the U.S. has 159. So, there's a large gap in terms of China's diffusion capacity compared to its innovation capacity in AI.

I'll close with the third point, which is some recent trends in Chinese labs' large language models. China has built large language models similar to OpenAI's ChatGPT, as well as OpenAI's text-to-image models like DALL–E. But there's still a large gap in terms of Chinese performance on these models. And, in fact, on benchmarks and leaderboards where U.S. models are compared to Chinese models on Chinese language prompts, models like ChatGPT still perform better than Chinese counterparts. Some of these bottlenecks relate to a reliance on Western companies to open up new paradigms, China's censorship regime, which Dr. Jensen talked about, and computing power bottlenecks, which I'm happy to expand on further.

I'll close by saying I submitted three specific policy recommendations to the Committee. But I want to emphasize one, which is keep calm and avoid overhyping China's AI capabilities. In the paper that forms the basis for this testimony, I called attention to a 1969 CIA assessment of the Soviet Union's technological capabilities. It was remarkable because it went against the dominant narrative of the time of a Soviet Union close to overtaking the U.S. in technological leadership. The report concluded that the technological gap was actually widening between the U.S. as the leader and the Soviet Union because of the U.S.'s superior mechanisms to spread technologies and diffuse technologies. Fifty years later, we know why this assessment was right, and we know we have to focus on diffusion capacity when it comes to scientific and technological leadership.

Thanks for your time.

[The prepared statement of the witness follows:]

September 19, 2023

"National security and economic competitiveness implications of AI"

Testimony before the Senate Select Committee on Intelligence

Hearing on Advancing Intelligence in the Era of Artificial Intelligence: Addressing the National Security Implications of AI

Jeffrey Ding
Assistant Professor of Political Science, George Washington University

Introduction

This testimony articulates an important distinction between innovation and diffusion capacity, which is crucial to accurate assessments of national scientific and technological capabilities. In contrast to an innovation-centric approach, an assessment based on diffusion capacity reveals that China is far from being a science and technology superpower. China's efforts to reform its education system will play a pivotal role in its ability to adapt to revolutionary technological advances and sustain economic growth in the long run.

I. Innovation vs. Diffusion Capacity

Discussions about national scientific and technological (S&T) capabilities tend to center on which state first generates new-to-the-world breakthroughs (*innovation capacity)*. In this testimony, the main point I aim to convey is that evaluations of technological leadership in AI should give greater weight to a state's *diffusion capacity*, or its ability to spread and adopt innovations, after their initial inception, across productive processes. When there is a substantial disparity between these two facets of a nation's S&T capabilities, innovation-centric assessments of its power to leverage S&T advances for sustained economic growth will prove misleading.[1]

Up front, I want to clarify that my testimony is especially relevant for understanding the global economic competition dynamics surrounding AI, or the ability of states to to exploit technological changes and maintain higher economic growth rates than its rivals. Historically, this mechanism has been central to the rise and fall of great powers.[2] My testimony has less bearing on other channels by which AI and emerging technologies could influence national security, which may also come under the committee's purview. Innovation-centric assessments may be rightly

---

[1] This testimony draws from Jeffrey Ding. (2023). "The Diffusion Deficit in Scientific and Technological Power: Re-assessing China's Rise." *Review of International Political Economy*.

[2] Kennedy, Paul M. The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000. New York: Random House, 1987.

prioritized in such contexts, such as the significance of S&T systems to prestige and reputation, control over global supply chains, and military power.[3] Still, appropriate attention to diffusion capacity can better inform other S&T dimensions of state power. For instance, there can be a large disparity between a military's ability to first field advanced military systems and its ability to adopt such systems throughout its branches and subunits.[4]

In many cases, there is not much daylight between a state's diffusion capacity and its innovation capacity. These two parameters can be highly correlated. After all, the state that first pioneered a new method has a first-mover advantage in the widespread adoption of that technique. In addition, absorbing innovations from international sources is difficult without the tacit knowledge embedded in the original context of technological development.[5] Diffusion and innovation are entangled, overlapping processes.[6]

However, in some circumstances, diffusion and innovation capacity can widely diverge. Aside from innovation capacity, many other factors can shape a country's adoption rate of new technologies, including institutions that incentivize technology transfer, trade openness, and human capital.[7] The "advantages of backwardness" sometimes enable laggards to diffuse new technologies faster than the pioneering states.[8] Confronting a world of globalized science and technology flows, even the most advanced economies must be able to intensively absorb and diffuse innovations first incubated in other countries. According to one estimate derived from data on Organisation for Economic Co-operation and Development countries, 93 percent of total factor productivity increases in these high-income countries derive from knowledge that originated abroad.[9]

As a result, diffusion capacity indicators can be better predictors of a state's long-term growth trajectory than innovation capacity indicators. The latter may be more unreliable given the uncertain, protracted pathway between a new technology's introduction and its ultimate impact on productivity growth. To this point, one study found that two standard innovation capacity

---

[3] Gilady, Lilach. The Price of Prestige. Chicago: Univ. of Chicago Press, 2017, 55-89; Malkin, Anton. "The Made in China Challenge to US Structural Power: Industrial Policy, Intellectual Property and Multinational Corporations." Review of International Political Economy 0, no. 0 (October 1, 2020): 1–33; Paarlberg, Robert L. "Knowledge as Power: Science, Military Dominance, and U.S. Security." International Security 29, no. 1 (2004): 122–51.

[4] Ding, Jeffrey and Allan Dafoe. (2023). Engines of Power: Electricity, AI, and General-purpose, Military Transformations. *European Journal of International Security*, 1-18.

[5] Fadly, Dalia, and Francisco Fontes. "Geographical Proximity and Renewable Energy Diffusion: An Empirical Approach." Energy Policy 129 (June 1, 2019): 422–35; Keller, Wolfgang. "International Technology Diffusion." Journal of Economic Literature 42, no. 3 (September 2004): 752–82.

[6] Taylor, Mark Zachary. The Politics of Innovation: Why Some Countries Are Better Than Others at Science and Technology. 1st edition. New York, NY: Oxford University Press, 2016.

[7] Comin, Diego, and Bart Hobijn. "An Exploration of Technology Diffusion." American Economic Review 100, no. 5 (December 2010): 2031–59.

[8] Gerschenkron, Alexander. "Economic Backwardness in Historical Perspective (1962)." The Political Economy Reader: Markets as Institutions, 1962, 211–28.

[9] Madsen, Jakob B. "Technology Spillover through Trade and TFP Convergence: 135 Years of Evidence for the OECD Countries." Journal of International Economics 72, no. 2 (July 1, 2007): 464–80.

indicators, R&D intensity and patenting rates, tracked less well with subsequent changes in productivity than indicators of activities related to broadly disseminating information about new products and processes.[10]

When there is a substantial gap between diffusion and innovation capacity, assessments based solely on innovation capacity indicators will be misleading because they undervalue the process by which new advances are embedded into productive processes. Specifically, a "diffusion deficit" characterizes situations when a state has a strong innovation capacity but weak diffusion capacity, which suggests that it is less likely to sustain its rise than innovation-centric assessments depict. For example, innovation-centric assessments overestimated the Soviet Union's scientific and technological capabilities in the postwar period. Taking diffusion capacity seriously would have provided a more balanced assessment of the Soviet Union's scientific and technological capabilities.[11]

II. China's Diffusion Deficit

Is China poised to become a science and technology superpower? Existing assessments of China's S&T capabilities tend to center on its aptitude in generating novel breakthroughs. To warn about challenges to U.S. technological leadership, analysts typically cite China's impressive performance in indicators of innovation capacity, such as R&D expenditures, scientific publications, and patents.[12] Less attention, if any, is paid to China's diffusion capacity. For example, the Senate Select Committee on Intelligence's 2022 report on protecting U.S. innovation, which included a lengthy section on China's technological rise, mentions "innovation" or "crown jewels" over ten times . The terms "diffusion" or "adoption" do not appear at all.[13]

Yet, according to my research, China faces a diffusion deficit: its diffusion capacity trails significantly behind its innovation capacity. Similar to issues with evaluating the Soviet Union's S&T ecosystem in the 1970s, this means that conventional assessments overestimate China's S&T capabilities. It is necessary to reorient such assessments toward a diffusion-centric lens, which show that China is far less likely to sustain its rise than innovation-centric assessments suggest.

Innovation-centric views of China's AI capabilities paint an overly optimistic picture of China's challenge to U.S. technological leadership. Influential reports emphasize China's growing strength in AI-related innovation, backed by indicators on R&D expenditures, leading AI

---

[10] Alexopoulos, Michelle. "Read All about It!! What Happens Following a Technology Shock?" American Economic Review 101, no. 4 (June 2011): 1144–79.
[11] For more on this historical case, see Jeffrey Ding. (2023). "The Diffusion Deficit in Scientific and Technological Power: Re-assessing China's Rise." *Review of International Political Economy*.
[12] Kennedy, Andrew B. "Powerhouses or Pretenders? Debating China's and India's Emergence as Technological Powers." The Pacific Review 28, no. 2 (March 15, 2015): 281–302.
[13] Senate Select Committee on Intelligence. "Organizational Assessment: The National Counterintelligence and Security Center." September 2022.

startups, and valuable internet companies. Likewise, to supports its warning that China is poised to overtake the U.S. in the capacity to generate new-to-the-world advances in AI, the National Security Commission on AI's final report cites shares of breakthrough papers in AI and investments in startups. These evaluations align with viewpoints that are bullish on China's overall technological capabilities, which also point to similar indicators of innovation capacity, such as R&D expenditures, scientific publications, and patents.

A diffusion-centric perspective, based on a close examination of China's adoption of information and communications technologies (ICTs), paints a different picture. While China has been successful at large-scale deployment in a few key domains — consumer-facing applications like mobile payments and high-speed rail — these achievements do not characterize the overall trend in ICTs. Chinese businesses have been slow to embrace digitization, as measured by adoption rates of digital factories, industrial robots, smart sensors, and key industrial software.[14] The International Telecommunication Union's ICT Development Index provides a composite measure of the level of access to and use of ICTs in countries around the world. On this metric, China ranks 83rd in the world, which trails the U.S. by 67 places.[15] China also significantly trails the U.S. in an influential index for adoption of cloud computing, which is essential to implementing AI applications. In 2018, U.S. firms averaged a cloud adoption rate of over 85 percent, more than double the comparable rate for Chinese firms.

In recent years, there has been more scrutiny of China's investments in the human capital necessary to adapt to emerging technologies such as AI. As technology races forward, skills must keep pace. Some studies inflate China's capacity to diffuse AI advances at scale because of its sheer quantity of computer science graduates.[16] General counts of graduates, without accounting for the quality of education, overstate China's capacity to cultivate a broad base of AI engineers. Comparisons of computer science education, in particular, can mislead, if the quality of such training is not considered.[17] Consider one quality baseline for AI education: a university meets this standard if it employs at least one researcher that has published at least one paper in a leading AI conference. According to data from the years 2020-2021, China was

---

[14] Alibaba Research Institute. "From Connected to Empowered: Smart+ Assisting the High-Quality Development of China's Economy [从连接到赋能：'智能+'助力中国经济高质量发展]," March 11, 2019; Synced [机器之心]. "Market Research Report on Supply and Demand for Digital Intelligentization Solutions for China's Small and Medium Enterprises [中国中小企业数智化解决方案供应市场研究报告 2020]," October 2020; Techxcope [战略前沿技术]. "Innovation Is More than Invention: Detailed Explanation of the German Industry-University-Research Systems' Big Four [创新不止于发明：德国产学研体系四大金刚详解]," November 18, 2020.

[15] International Telecommunications Union. "Measuring the Information Society Report 2017," 2017. https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2017.aspx.

[16] Allison, Graham, and Eric Schmidt. "Is China Beating the U.S. to AI Supremacy?" Belfer Center for Science and International Affairs, August 2020.

[17] Loyalka, Prashant, Ou Lydia Liu, Guirong Li, Igor Chirikov, Elena Kardanova, Lin Gu, Guangming Ling, et al. "Computer Science Skills across China, India, Russia, and the United States." Proceedings of the National Academy of Sciences 116, no. 14 (April 2, 2019): 6732–36.

home to only 29 universities that met this standard; the U.S. accounted for 159 such universities.[18]

When it comes to disseminating AI advances across the entire economy, robust linkages between academic and industry settings are especially crucial. The U.S. has built a strong connective tissue in this respect. Per data on the years 2015 to 2019, the U.S. was the world leader in the number of academic-corporate hybrid AI publications — publications co-authored by at least one researcher from industry and one researcher from academia. This more than doubled China's number of hybrid AI publications.[19] Indeed, China's official state news agency has highlighted the lack of technical exchanges between universities and industry as one of five key weaknesses in China's AI talent ecosystem.[20]

It is now becoming increasingly common for reports to claim that China has overtaken the U.S. in certain measures of elite research in AI.[21] One important distinction to make is that these claims tend to draw on indicators based on AI publications in *journals*. In fast-moving fields like AI, a country's performance in conference publications may be a much better indicator of its high-end talent than journal publication-based indicators. As Stanford University's AI Index pointed out in 2021, "the United States has consistently (and significantly) more AI conference papers (which are also more heavily cited) than China over the last decade."[22]

Lastly, to analyze whether China's overall diffusion capacity in science and technology varies significantly from its innovation capacity, I separated indicators included in the Global Innovation Index, a widely-used benchmark for national S&T capabilities published by the World Intellectual Property Organization, into these dimensions. For example, the GII ranks countries globally by the quality of their top three universities and their top three firms' R&D expenditures. I categorize these as indicators of innovation capacity. The GII also ranks countries by indicators that correlate strongly with a country's capacity to diffuse new advances, including the extent of linkages between businesses and universities.

This decomposition of the 2020 GII reveals that China's diffusion capacity significantly lags behind its innovation capacity (Table 1). Using the GII's figures, averaging China's global

---

[18] Analysis based on the CSRankings website. For details on the original methodology, see Tencent Research Institute and Boss Zhipin 2017, 12.

[19] Zhang, Daniel, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, et al. "The AI Index 2021 Annual Report." Stanford Human-Centered Artificial Intelligence Institute, 2021.

[20] Xinhua. "News Analysis: Examining the Five Shortcomings of China's AI Talent System [新闻分析：透视中国人工智能人才体系五大短板]." Xinhua News Agency, August 28, 2019. http://www.gov.cn/xinwen/2019-08/28/content_5425310.htm.

[21] See, for example, Nikkei Asia. "China Trounces U.S. in AI Research Output and Quality." January 15, 2023. https://asia.nikkei.com/Business/China-tech/China-trounces-U.S.-in-AI-research-output-and-quality.

[22] Zhang, Daniel, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, et al. "The AI Index 2021 Annual Report." Stanford Human-Centered Artificial Intelligence Institute, 2021.

ranking on indicators for innovation capacity gives an average of 13.8. However, if the same exercise is conducted using diffusion capacity indicators, China's average ranking drops to 47.2. For reference, on the innovation capacity subindex, China's score is very close to the U.S.'s average ranking (11.9). As for the diffusion capacity subindex, the gap widens significantly between China's average ranking of 47.2 and the U.S.'s average ranking of 26.9. Table 1 displays the GII indicators used to calculate China's diffusion capacity and innovation capacity.

III. Recent Trends in China's Large Language Model Ecosystem

Over the past few years, Chinese labs have quickly followed in the footsteps of U.S. labs to build large language models (LLMs), text generation systems such as ChatGPT. In a recent *Foreign Affairs* piece, Helen Toner, Jenny Xiao, and I argued that "When it comes to LLMs, China trails years, not months, behind its international competitors." This gap is a product of many factors, including a reliance on Western counterparts to open up new paradigms of AI development, political constraints on free speech, and relative lack of high-quality Chinese-language data for training.

China's pace of LLM development is also impeded by bottlenecks in the supply of semiconductors. In the *Foreign Affairs* article, we laid out this case in detail:

> *Due to the outsized computational demands of LLMs, the international competition over semiconductors inevitably affects AI research and development. The Chinese semiconductor industry can only produce chips several generations behind the latest cutting-edge ones, forcing many Chinese labs to rely on high-end chips developed by U.S. rms. In recent research analyzing Chinese LLMs, we found 17 models that used chips produced by the California-based firm NVIDIA; by contrast, we identified only three models built with Chinese-made chips.*[23]
>
> *Huawei's PanGu-α, released in 2021, was one of the three exceptions. Trained using Huawei's in-house Ascend processors, the model appears to have been developed with significantly less computational power than best practices would recommend. Although it is currently perfectly legal for Chinese research groups to access cutting-edge U.S. chips by renting hardware from cloud providers such as Amazon or Microsoft, Beijing must be worried that the intensifying rhetoric and restrictions around semiconductors will hobble its AI companies and researchers.*[24]

---

[23] Jeffrey Ding and Jenny Xiao. "Recent Trends in China's Large Language Model Landscape." *Centre for the Governance of AI*. April 2023.

[24] "Chinese AI Groups use cloud services to evade US chip export controls." *Financial Times*. March 8, 2023.

It is worth noting that the innovation-diffusion distinction is also relevant for discussions about the national security implications of LLMs. Almost all the attention has gone toward which country can develop the next breakthrough in foundation models; much less attention goes toward what happens after large models are trained and their adoption rate across different types of industries. From my preliminary research in this area, it seems that China still faces a large “implementation gap” in terms of making LLMs cost-effective to be used by small and medium-sized businesses.[25] General-purpose technologies like AI take time to diffuse, and if AI does truly transform the global economy, we are still in the early stages.

IV. Conclusion and Policy Recommendations

Given the above analysis of China’s diffusion and innovation capacity, the following policy recommendations could help safeguard U.S. interests:

First, keep calm and avoid overhyping China’s AI capabilities. My research suggests that the U.S.’s lead in AI capabilities over China should endure. The status quo, sometimes, functions as a defensible policy option. More specifically, claims that U.S. regulatory action on AI will allow China to race ahead in this domain do not hold water, and they should not confound deliberations over sensible guardrails on rapidly-advancing AI systems.[26]

Second, revive the Office of Technology Assessment (OTA). There is a need for more balanced assessments of where China and the U.S. stand with respect to AI and other emerging technologies. The OTA helped fill this gap from 1972 to 1995. There is bipartisan support for this proposal, and both liberal and conservative think tanks have supported proposals to revive the OTA.[27]

Third, invest in technology diffusion. In the context of general-purpose technologies such as AI, policies directed at broadening the AI talent base, such as by further supporting community colleges in developing the AI workforce, may be just as, if not more, important as producing the best and brightest AI experts.[28] The U.S. should also invest in “technology diffusion institutions,” including applied technology centers and dedicated field services, that encourage the adoption of AI techniques by small businesses.[29] All too often, it seems, the U.S. government’s go-to

---

[25] Jeffrey Ding. 2022. “ChinAI #199: China’s Hugging Face?” https://chinai.substack.com/p/chinai-199-chinas-hugging-face; Jeffrey Ding. 2023. “ChinAI #236: The LLM Implementation Gap” https://chinai.substack.com/p/chinai-236-the-llm-implementation.

[26] Helen Toner, Jenny Xiao, and Jeffrey Ding. 2023. The Illusion of China's AI Prowess. *Foreign Affairs*. https://www.foreignaffairs.com/china/illusion-chinas-ai-prowess-regulation.

[27] Katherine Tully-McManus, “House Members Call for Office of Technology Assessment Revival,” Roll Call, April 2, 2019, https://www.rollcall.com/news/congress/house-members-call-office-technology-assessment-revival.

[28] West, Darrell M. *The Future of Work: Robots, AI, and Automation*. Brookings Institution Press, 2018, p. 112-113; National Security Commission on Artificial Intelligence. “Final Report.” Washington, D.C.: NSCAI, March 2021. https://www.nscai.gov/2021-final-report/, p. 175.

[29] Shapira, Philip, and Jan Youtie. “The next Production Revolution and Institutions for Technology Diffusion.” *The Next Production Revolution: Implications for Governments and Business*, 2017.

recommendation for any strategic technology is to boost R&D spending. A diffusion-oriented perspective demands a more varied approach.

Appendix

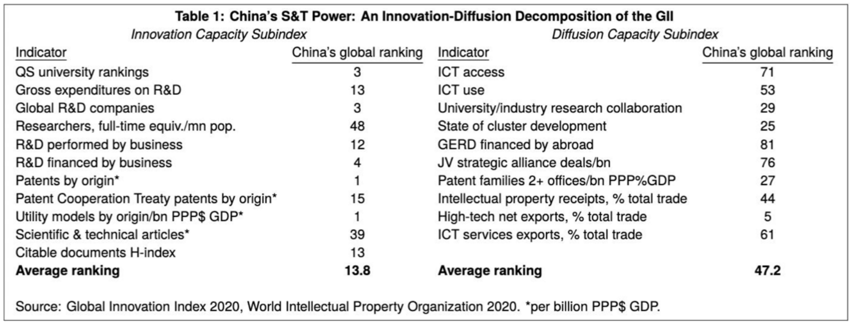

Table 1: China's S&T Power: An Innovation-Diffusion Decomposition of the GII

Innovation Capacity Subindex		*Diffusion Capacity Subindex*	
Indicator	China's global ranking	Indicator	China's global ranking
QS university rankings	3	ICT access	71
Gross expenditures on R&D	13	ICT use	53
Global R&D companies	3	University/industry research collaboration	29
Researchers, full-time equiv./mn pop.	48	State of cluster development	25
R&D performed by business	12	GERD financed by abroad	81
R&D financed by business	4	JV strategic alliance deals/bn	76
Patents by origin*	1	Patent families 2+ offices/bn PPP%GDP	27
Patent Cooperation Treaty patents by origin*	15	Intellectual property receipts, % total trade	44
Utility models by origin/bn PPP$ GDP*	1	High-tech net exports, % total trade	5
Scientific & technical articles*	39	ICT services exports, % total trade	61
Citable documents H-index	13		
Average ranking	**13.8**	**Average ranking**	**47.2**

Source: Global Innovation Index 2020, World Intellectual Property Organization 2020. *per billion PPP$ GDP.

Chairman WARNER. Thank you all very much, gentlemen. I'm going to ask Vice Chairman Rubio to make any opening comments he wants. Then we'll go to a question round.

OPENING STATEMENT OF HON. MARCO RUBIO, A U.S. SENATOR FROM FLORIDA

Vice Chairman RUBIO. And I'll be brief, and I apologize. I was wrapped up in a call that started late and ended late. So, I'll be very brief.

This whole issue is fascinating to me because the story of humanity is the story of technological advances from the very beginning in every civilization and culture. And there are positives in every technological advance, and there are negatives that come embedded in it. And generally, technological advances have for the most part—what they've allowed is human beings to do what humans do, but faster, more efficiently, more productively, more accurately. In essence, technology and technological advances have, in essence, allowed humans to be better at what humans do.

I think what scares people about this technology is the belief that it not simply holds the promise of making us better, but the threat of potentially replacing the human, being able to do what humans do without the human. I think one of the things that's interesting is we've been interacting with AI, or at least models of AI and applications of it, in ways we don't fully understand. Whether it's estimating how long it's going to take from point A to point B and which is the fastest route, based on their predicted traffic patterns at that time of day to every time you say Alexa or Siri, all of these are somewhat built into learning models.

But now we get into the application of machine learning, where you're basically taking data and you're now issuing recommendations of a predictive nature. So, that's sort of what we understand now. But then the deeper learning that actually seeks to mimic the way the human brain works. Not only can it take in things beyond text-to-images, but in essence, learn from itself and continue to take on a life of its own. All of that is happening, and frankly, I don't know how we hold it back.

So, really, the three fundamental questions that we have from the perspective of national security is first and foremost, which I think this is a broader topic that involves national security, but beyond national security, is how do you regulate a technology that is transnational, that knows no borders, and that we don't have a monopoly on? We may have a lead on it, but some of the applications of AI are going to be pretty common, and for purposes of what some nefarious actor might use and some of its applications as well.

The second is, will we ever reach a point—and this is the one that, to me, is most troubling—where we can afford to limit it. So, as an example, we are in a war, God forbid, with an AI general on the other side. Can we afford to limit ourselves in a way that keeps pace with the speed and potentially the accuracy of the decision-making of a machine on that end and our limits on ours? And the same is true in the business world when we get AI CEOs making decisions about where companies invest. There comes a point where you start asking yourself, can we afford to limit ourselves despite the downside of some of this. It's something we haven't thought through, and one that I think you've addressed a lot or are going to be addressing a lot are the national security implications.

But here's the one that I think is related to national security. We have seen that globalization and automation has been deeply disruptive to societies and cultures all over the world. We have seen what that means in displacing people from work and what it does to society and the resentments it creates. I think this has the potential to do that by times-infinity. In essence, how disruptive this will be, the industries that it will fundamentally change, the jobs it will destroy and perhaps replace with different jobs, but the displacement it could create. And that has national security implications, and what happens in the rest of the world and in some of the geopolitical trends that we see. And I think it has the threat of reaching professions that up to this point have been either insulated or protected from technological advances because their education level. So, again, not a national security matter.

Look at the strike in Hollywood. A part of that is driven by the fear that screenwriters, and maybe even the actors, will be replaced by artificial intelligence. And imagine that applied to multiple industries and what that would mean for the world and for its economics. There's a lot to unpack here. But the one point that I really want to focus in on is we may want to place these limits, and we may very well be in favor of them from a moral standpoint, but can we ever find ourselves at a disadvantage facing an adversary who did not place those limits and is therefore able to make decisions in real time, at a speed and precision that we simply cannot match because of the limits we put on ourselves? That may still be the right thing to do, but I do worry about those implications.

Thank you.

Chairman WARNER. Well, thank you, Marco. I think, again, we are all trying to grapple with this in a variety of ways. I remember a year and a half ago, as I was trying to get self-educated a bit and thought at first, well, should we even find a definition? How about now? Not worth that. Last week we had 22, and it was really kind of the who's who, from all the tech side and a variety of figures from civil society. A year and a half later, and no one still started with the definition of what the terms we're even using mean.

And yet we've seen, I would argue, I think about most things I spend time on. There's some linear progression, I feel, in terms of the amount of time versus the amount I've learned. In this topic area, I get sometimes more confounded. I also think, for example, the economics, the economics of last November, whenever you said whether it was China in terms of scale, amount of data compute, et cetera, or entities like Microsoft or Google, the gating cost to come into this would be so high. We've had the director of the OSTP recently say, potentially because of release of things like LAMA, that you can now get in on a variation of large language model for pennies on the dollar. So, this is moving so quickly.

Dr. LeCun, I'm going to start with you, and I warned you I was going to come at you on this. I worry a little bit when we talk about democratizing access. Then it sounds like some of your colleagues from social media in the late '90s. We're going to democratize access, and we'll figure out the guardrails after the fact. But in that democratization of access, I don't think we ever put our guardrails in place. How do we democratize access with AI tools and yet still put some level of guardrails? And obviously, bearing in mind what Senator Rubio said, is you don't want to act and unilaterally disarm, but the notion of not putting some guardrails in place in a field that's changing so quickly. Can you speak to that?

Dr. LECUN. Senator, this is a very important question that, of course, we've devoted a lot of thought to. I think the best way to think about this is to think about the type of AI systems that have been released so far as being basic infrastructure. In themselves, those systems are not particularly useful. To be useful, they need to be customized for a particular vertical application. And a good example of this is the infrastructure of the Internet, which is open source. It didn't start out as being open source, it started out as being commercial.

And then the open-source platforms won because they are more secure, easier to customize, safer. There are all kinds of advantages. AI is going to become a common platform, and because it's a common platform, it needs to be open source if you want it to be a platform on top of which a whole ecosystem can be built. And the reason why we need to work in that mode is that this is the best way to make progress as fast as we can.

I really like the argument of Professor Ding about the fact that, in the U.S., we're extremely good at seizing the opportunity when a new innovation comes in or a new scientific advance, it diffuses very quickly into the local industry. This is why Silicon Valley is Silicon Valley. It's geographically concentrated because information flows very, very quickly. Other countries that are somewhat isolated ecosystems intellectually do not have the same kind of effect. And so, it favors us to have open platforms.

Chairman WARNER. I would respectfully say, I think the two most immediate potential harms with AI tools that have already been released is the ability for massive undermining of trust in our elections and massive underlying of trust in our public markets. Come back to that later.

Dr. Ding, one of the things you said, and I think you're accurate about the fact that the PRC as a state has not done a good job of diffusing technological innovation. But I have to believe if we do open source, and there are geopolitical harms, and there is the ability for the PRC, at least its intel and military, to gain this knowledge from things that are released into the wild. Can you speak to that?

Dr. DING. Yeah, I think it's a tough debate and tough discussion point. I think I agree that open source is important to spur diffusion and in terms of fundamental architecture. So, going back to this Internet example, the protocols for how IP addresses work, it makes sense that open source might be a way we spread that at scale. I'm less convinced that open source is the best model for reducing the harms that you identified, in terms of specific, powerful models that might not be as close to this infrastructure layer.

So, something like ChatGPT is, I think, took a good stance in terms of setting up an application programming interface which makes it not open source, it's closed source. And developers can impose rules on how the model can be used. So, Chinese developers cannot use ChatGPT today, to your point. And so, I think this provides a balance in terms of the research community can still play around with the model by using the API system. But OpenAI and potentially government actors can use this API to implement rules to govern how these models are used by potentially malicious actors.

Chairman WARNER. Let me get one question quickly to Dr. Jensen, because the notion that scale was going to be the determinant factor seemed to be the argument that most large language models launched from November till about May, and then it switched. If scale is not the largest determinant of who will be most successful, what will be the determinant?

Dr. JENSEN. The people. So, you're worried about the guardrails. I'm not even sure you have the railroad engineers to get the train out the station. And what I mean by that is you're going to have to have hard decisions, first, like you're seeing of what data is open and closed. Because you're right, if you want an innovative ecosystem, the exchange of ideas, we're built on that as a nation. Obviously not all ideas are meant to be shared. Even George Washington had secrets.

So, what do you hide in terms of data will become really important. And how do you aggressively corrupt your adversaries' data through poisoning it as well—a digital Terracotta Army to confuse them. And scale will matter because the larger amounts of data inputs you have, the more likely you are to be able to detect adversary manipulation or those markets that we should worry about. And to do that, it's not just the data infrastructure. Again, you have to have the people who know what they're doing. And I work with these folks. I mean, I've served in uniform for 20 years. I teach warriors, and I watch them when we tinker with this in the classroom. And honestly, some part of me goes back to classical reasoning. You want to really teach someone how to work with a large language model? Make them think like Socrates, learn how to ask questions anew. Otherwise, it'll just be confirmation bias. You made them to understand how a sequential prediction that a large language model is just basically predicting how they're going to finish their sentence. There's no magic there. And so, it's not just the scale of the data, as you point out, Sir. It's how do we actually educate that workforce so that we can out-cycle any adversary.

Chairman WARNER. Thank you. I'm going to Senator Rubio. And again, I remind my colleagues in our open hearings, we go by seniority.

Vice Chairman RUBIO. I understand, we want to talk about sort of the commercial and broader scientific applications of this. It'd be great to be the world leader, industry standard, top of the line. But for purposes of this Committee, how it would be used as a nation-state? What I think is important to reiterate is you don't need the state of the art for it to be adopted internally for your decision-making. Every conflict in human history has involved an analysis. At some point, someone who started the war made an analysis based on their data that was in their brain, their understanding of history, their beliefs, their preconceived notions and the like, that they could be successful and that this was really important and now is the time to do it. And that's the part that worries me the most, particularly when it comes to how it applies to authoritarian regimes. And here's why.

At least in our system, for the most part, we encourage people to come forward and, as policymakers, make an analysis and give us accurate information even if it may not be the one we want to hear. In authoritarian systems, you usually get promoted and rewarded for telling leaders what they want to hear and not for reporting bad news and the like. And so, I don't know if anyone can answer this question, but I wanted to pose it to you.

Isn't one of the real risks as we move forward that some nation with some existing capabilities will conduct analysis on the basis of their version of AI, which will be flawed to some extent by some of the data sets, and those data sets and the analytic functions, they reach the conclusion that this is the moment to take this step. Now is the time to invade, now is the time to move because our system is telling us that now is the time to do it. And that system may be wrong. It may be based on flawed data. It may be based on data that people fed in there on purpose because that's the data that their analysts are generating.

That's the part that I worry about. Because even if it's not the top-of-the-line or the state-of-the-art data, it will be what influences their decision-making and could very well lead to twenty-first century conflicts started not by simply a human mind, but how a human mind used technology to reach a conclusion that ends up being deeply destructive. Is that a real risk is the question?

Dr. JENSEN. I'm happy to talk about war anytime, Senator. I think you're hitting on a fundamental of human history, as you're saying, right? Every leader, usually not alone, as part of a small group, is calculating risk at any moment. And having models incorrectly or correctly add to their judgment about that is a real concern. There will be windows of vulnerability and the possibility of inadvertent escalation that could make even the early Cold War look more secure than it actually was.

And so, I think that's the type of discussion you have to have. That's where we hopefully will have back-channel conversations with those authoritarian regimes. And frankly, it just bodes well for what we know works for America: a strong military, where your model finds it really hard to interpret anything but our strength. So, I think that there are ways that you can try to make sure that the right information is circulating, but you can never fundamentally get away from those hard, weird moments, those irrational people with rational models. So, you see the model as rational or flawed because it collects just the skewed data. I worry more about what we just saw happen in Russia where a dictator living in corrupt mansions, reading ancient imperial history of Russia, decided to make one of the largest gambles in the twenty-first century. And so, I don't think that's going to leave us. I think that's a fundamental of human history. And I actually think, in some senses, the ability of models to bring data could steady that a bit. And we can make sure that we show the right type of strength, that it steadies it further.

Vice Chairman RUBIO. Let me ask this question related to that one, and it has to do with the work of this Committee in general. At the core of intelligence work is the analysis. In essence, you can collect all the raw bits of data you want, but someone has to interpret and tell a policymaker this is what we think it means, in light of what we know about those people, what we know about historical trends, what we know about cultures, common sense, whatever you name. And there's an analytical product. And then you have to make policy decisions, either with high confidence in the analysis, moderate confidence, low confidence, whatever it may be.

Given that, what suggestions, if it's possible at this point, could you provide us as to what that analysis should include or look like if applied to the way we analyze data sets, so that not only are we reaching the most accurate results and the ones that are true, but ones that provide our policymakers a basis upon which to make the best decisions possible? Weighing all the equities, including human consideration, not just simply the cost benefit analysis from an economic or military standpoint.

Dr. DING. So, let me start with your earlier question, which I take as, what is the risk of AI in terms of contributing to military accidents? And so, I would say that an authoritarian regime might be a contributing factor to a state having a higher risk of military accidents. I think when we talk about these information analysis systems, think about the Aegis, right? The U.S. Aegis system that collects information and analyzes it and issues what this target is, whether it's friend or foe, and then whether we should fire a missile towards the target. In the 1990s, the U.S. accidentally fired upon an Iranian civilian airliner, killing 300 people. So, military accidents can happen in democratic countries.

But I think it's an interesting research question, right? One of the things that I'm interested in studying is, how has China, as an authoritarian state, actually demonstrated a decent safety record with civil nuclear power plants and aviation safety? How does that happen in a closed authoritarian system? What is the role of international actors? And a military accident anywhere, whether it's caused by AI or any other technology, is a threat to peace everywhere. Right? To your point, so we should all be working to try to reduce the risks of these systems, sort of accidents in military AI systems.

To your second point, one of my recommendations would be to keep a human in the loop, regardless of whatever AI system we adopt in terms of intelligence, surveillance, and reconnaissance. And hopefully that will make these systems more robust.

Chairman WARNER. Senator Wyden.

Senator WYDEN. Thank you, Mr. Chairman. Let me start with you, Dr. LeCun.

I have proposed the Algorithm Accountability Act that would require that companies test their AI for harmful bias, such as biases that can affect where you buy a house or what healthcare you have access to. Now, reviewing your testimony, Dr. LeCun, you stress commitment to privacy, transparency, and mitigating bias—all important values. It's important to me also that there not be an uneven playing field that advantages the companies cutting corners over companies that do the right thing.

Could your company support this legislation?

Dr. LECUN. Senator, this is a very important question. Thank you for raising this point. I'm not familiar with the details of the legislation in question. Certainly, in the basic principles, they align with my personal thinking and those of Meta. And I'll be happy to put you in touch with relevant people within the company who take care of legislation.

Senator WYDEN. Hearing you agree with the principles is a good way to start. We'll follow up.

Now, the use of AI by the U.S. Intelligence Community raises many issues, starting with accountability. Now, if AI is going to inform the Intelligence Community's surveillance decisions, one question I would have is if an American is spied on in violation of the law, who is responsible?

Dr. LECUN. Senator, again, this goes very much outside of my expertise, not being a lawyer or legislator. Privacy, security, and safety are on top of our list in terms of priorities. They're very good principles to follow. We try to follow them as much as we can. I think an important point to realize, as it relates to current AI technologies such as large language models, they are trained on public data, publicly-available data, not on private user data. So, there's no possibility of any kind of privacy invasion from that.

Senator WYDEN. Let me ask it this way. Your testimony stresses the importance of accountability in the private sector for its use of AI. How might these principles apply to the government?

Dr. LECUN. Senator, I think the White House commitments, voluntary commitments, are a good start to specify guidelines according to which the AI industry should conform, including questions related to your point.

Senator WYDEN. So, what was it about the guidelines in your view that are responsive to my question about how this accountability in the private sector, the principles, would apply to government?

Dr. LECUN. Senator, again, this is very much outside my expertise, and I'd be happy to put you in touch with the right people.

Senator WYDEN. Let me move on, then.

Dr. Jensen, your testimony describes the need for intelligence analysts who understand the AI that's informing their analysis. Now, I don't know how realistic it is to add advanced computer science expertise to the job requirements of intelligence analysts. It seems to me, though, you're raising a very important issue. What are the consequences of disseminating intelligence analysis derived from processes that nobody understands?

Dr. JENSEN. You hit the nail on the head, Senator. And I think it's less about making sure every analyst has a PhD in computer science. We can't afford that. No offense, Meta. But what we can do is make sure they understand the basics of reasoning, causal inference. Sometimes we throw around the term critical thinking as a blanket statement. But almost going back to how would Aristotle teach Alexander the Great to interpret a model. I know that's a weird thought experiment but think about that. Would that be a square of syllogism? Would it be about the logic? Would it be about contradictions? I think it's realistic that we can go back to some basic philosophical reasoning and not necessarily have to have high degrees of computational understanding. And I think by doing that, you start to get—you hone the ability of the person to ask a question. And as you all know, asking the great question is what produces real dialogue.

Senator WYDEN. So, as a general proposition, who would be accountable when the intelligence analysis turns out to be wrong?

Dr. JENSEN. This is a great question, and it actually dovetails with the question you were just asking. I'm going to summon my inner Senator King and say, you still need one throat to choke, right? So, that means ultimately, you have to have people accountable in terms of how they certify the assurance of their AI model. And that means, just like we had to struggle with this in the financial sector, you're going to have entirely new positions created on how people certify the actual model. And does the model address a certain set of data, a certain set of questions? And back to what Senator Rubio was talking about, even assign confidence levels to that. And I don't think we can even imagine what that's going to look like in five years. But I can tell you it's going to be a growth industry and an important one.

Senator WYDEN. My time has expired.

Chairman WARNER. Thank you, Senator Wyden. Senator Cornyn.

Senator CORNYN. A simple computer or Internet search tells me that the AI was basically—the roots of AI go back to 1956. So, maybe you could explain to me why we've gone all these years and haven't talked much about AI. And today we can't talk about anything else. Anybody want to take a stab?

Dr. LECUN. I think I have to answer that question, Senator. Thank you for asking it.

Generally, the problem of making a machine act intelligently has been much more complex than people initially realized. And the history of AI has been a succession of new ideas, with people thinking this new idea was going to lead us to machines that are as intelligent as humans. And over and over again, the answer has been no. Those new ideas that you had have solved a number of problems. But human level intelligence is still out of reach, and this is still the case today. So, despite the fact that we have incredibly powerful systems that are very fluent, that seem to manipulate language at least as well as humans, those systems are very far from having human intelligence.

Now, to directly answer your question, the reason why we hear about AI today, so much over the last ten years, roughly, is because of a new set of techniques called deep learning that has allowed machines to not be programmed directly, but to be trained for a particular task. And that's been incredibly successful for relatively narrow tasks where we can train machines to have superhuman performance. But so far, we still do not have a way to train a machine that is as efficient as the way humans, or even animals, can train themselves. This is why we don't have self-driving cars. We don't have domestic robots that can clear up the dinner table and fill up the dishwasher.

Senator CORNYN. Thank you. I'm going to try to get two more questions in.

One. Dr. Ding, you mentioned that artificial intelligence is a general-purpose technology. Which leads me to ask, why is it that we feel the need to regulate AI as opposed to regulating the sectors where AI is actually used, which the government already does?

Dr. DING. It's a great question. I do think that a sector-specific regulatory model is a good starting point, because AI will have different risks across different sectors. In fact, the European Union is a good model for this. In their EU/AI regulations, they have identified certain high-risk applications, where there are more stringent regulations. So, I think the crucial work to be done is identifying which sectors or which specific type of applications would be considered more high risk than others.

Obviously, the use of AI in a nuclear power plant, for example, or a chemical processing plant, might be higher risk. And then there are other applications that are in a murky area, such as large language models and content generation models. It's more difficult to assess how you would compare the safety risks of those against more traditional industries.

Senator CORNYN. Thank you. I just have about a minute. Georgetown Center for Security and Emerging Technology documents that American investors have provided roughly $40 billion, or 37 percent of the capital, to PRC AI companies from 2015 to 2021. This has been an area of emerging concerns. Subject of an Executive Order by the Administration. Senator Casey and I have in the Defense Authorization Bill an Outbound Investment Transparency Bill because it occurs to us that we are helping to finance our chief competitor globally, the PRC.

So, can you maybe—I'll stick with you, Dr. Ding. Can you speak to how effective those export controls, the outbound transparency issues, how are we doing in terms of slowing down our principal competitor while we try to run faster?

Dr. DING. It's a great question. I think, first of all, the transparency measures are a good starting point, because we need to have a better sense of how much outbound investment is going to China. What's the nature of that outbound investment, and what are the national security risks of that investment? I think oftentimes we think anything that helps China is going to hurt the U.S. when it comes to this space. I think it's an interesting question. Right? ByteDance, companies like Alibaba, Baidu—these giants in China's AI industry, they have a lot of foreign investment, but a lot of the profits that they make come back to the U.S. economy, and it hopefully gets reinvested into American productivity.

So, I think it's a more difficult question than just any investment in China's AI sector means it's harmful to U.S. national security. So, hopefully the transparency measures will help us get a better 360 degree view of the situation.

Senator CORNYN. Thank you.

Chairman WARNER. Senator Heinrich.

Senator HEINRICH. Thank you, Chairman.

Dr. Jensen, you've spoken a little bit about the importance of data, how should we leverage our unique U.S. government data sets to our advantage?

Dr. JENSEN. Thank you, Senator. First, make them actually interoperable. One of the big challenges we have is that because of just antiquated bureaucratic procedures and stove piping, this AI scientist would probably not be that successful in the U.S. government, because he couldn't make any of the data—. It's not that it's just heterogeneous, it's spread out over bureaucracies with random officials, each exerting authority they don't have to limit the ability to share. So, if you need more data to make smarter models, and you have people limiting the exchange of data, the train's not going to leave the station.

So, I think the first thing—and this is where I think the Congress has a very important role. How do you actually—whether it's through testimony, whether it's through the NDA, whether it's through hearings like this—how do you get bureaucrats to be accountable to actually exchange the data? And it actually goes deeper into actually government procurement and contracting. Right now, we could insert basic blanket language that require all vendors, because the sensors—the Intelligence Community, we buy it, right? So, why shouldn't it be that they're required to produce the information in standardized formats that we could ensure are interoperable. So, we lower the actually bureaucratic and engineering friction, and we could make use of it. It's just a boring issue, so sadly people don't pay attention to it.

Senator HEINRICH. But it's probably the most foundational issue from——

Dr. JENSEN. Yes.

Senator HEINRICH. Dr. Ding, you talked a lot about the fusion capacity. Are there any things that you consider threats to our diffusion capacity that we should be concerned with?

Dr. DING. Yes, I would say the main things to improve in terms of the U.S. diffusion capacity is investing more in the STEM workforce, in terms of not just attracting or building up the best and the brightest but widening the base of average engineers in software engineering or artificial intelligence. The U.S. government has some proposals on that, and the CHIPS and Science Act made some steps in that direction, but I would say we have overly weighted towards investing in R&D as sort of like the end all be all of science and technology policy. A more diffusion-oriented policy would look at things like broadening the workforce, investing in applied technology service centers, and dedicated field services. There are different voucher schemes that could encourage the adoption of AI techniques by small businesses as well. So, all of those things would help resolve some gaps in our diffusion capacity.

Senator HEINRICH. Very helpful.

Dr. Jensen, the current DOD directive governing lethal autonomous weapons systems appears from my read to permit the potential development, even deployment, of fully autonomous systems—not just in the defensive—that could select or engage targets without a human in the loop.

Talk to me about what are the risks there, how we should be further developing that policy, especially with regard to potential escalation?

Dr. JENSEN. Sure, great question, Senator.

I think the real issue here is not whether or not you should do it. It's here, right? It's how, again, you get the assurance and the model and why you have to constantly train and experiment to understand those edge cases, that I think was where also Senator Rubio was getting. Where is there that moment of high escalation risk that's actually irrational in the grand scheme of things, but it makes perfect statistical sense at the moment?

You won't find those weird cases until you do large-scale war gaming and constant experimentation. That's not just fine-tuning the model. This is where I think we get it wrong: we think it's like, well, I'm just going to calibrate the model, I'm going to fine tune the model. No, it's fine tuning the men and women who will use the model to make some of the most difficult decisions about taking life. Because eventually, there's still someone flipping that switch, right. And so, we need to make sure that those people—those men and women in uniform and the elected officials granting them the ability to use lethal force—have actually done tabletop exercises and experiments where they thought this through. Don't let your first moment of unleashing your robotic swarm be the first time you thought about it. And that's going to require a larger national security dialog and even tabletop exercises that help them see those moments.

Senator HEINRICH. Dr. LeCun, last question.

Before systems get released into the wild, as it were, there are a lot of ethical and other questions that need to be asked. Do you think that Meta's AI and Trust Safety Team, or for that matter the team for any AI developer, have the capability to really understand what the potential risks and benefits are? To be able to know whether or not making a system, putting into the wild, or making it open is a good idea?

Dr. LECUN. Senator, thank you for your question.

I can describe the process that we went through for the LAMA and LAMA–2 system. So, first of all, the LAMA system was not made open source. There are two parts to an open source package: there is the code, and the code frankly is very simple and not particularly innovative. What is interesting for the community to release is the train model, the weights. This is what's expensive, this is what only large companies can do at the moment. And we released it in a way that did not authorize commercial use. We vetted the people who could download the model. It was reserved for researchers and academics. This was a way for us to test what the system could be used for. Of course, there was a long history of three years of open source LLMs that were available before, and the harm has not materialized so far. So, there was a history we could based ourselves on.

For LAMA–2 we had a very thorough process. First, the data set was curated in such a way that the most controversial, toxic content was eliminated from it so that we would get a high-quality model. Second, there was a lot of red-teaming and so people basically tried to break the system and get it to produce dangerous toxic content. Thousands of hours were spent doing this by a group that is independent from the group that actually designed and trained the system. We have an entire group called Responsible AI, whose responsibility, among others, is to do this kind of thing.

And then we distributed in a limited way the model to Whitehat hackers at the DefCon Conference, for example slightly bigger community of people who are really expert at trying to break systems of this type. And I got some assurance that those systems were good. We fine-tuned them so that whatever was bad was fixed.

And then we instituted a bug-bounty policy, so that there would be an incentive for people who find weaknesses in our system to tell us. And in fact, the enthusiasm from the open source community, after the release of LAMA–2, has been so enormous that we are getting feedback absolutely all the time and make those systems safer.

Chairman WARNER. Senator Moran.

Senator MORAN. Chairman, thank you.

Dr. Jensen, you say that kids in China aren't rushing to join the Army and that Russian tech workers fled the country to avoid fighting the unjust war in Ukraine. Senator Warner and I have introduced over the years, a start-up act designed to create greater entrepreneurial environment in the United States but includes the creation of a STEM visa that would allow immigrants with advanced STEM degrees to stay in the U.S. as long as they remain engaged in the STEM field, with the indication is workforce is hugely important.

Let me ask though, when you say what you said, give me some ideas of what you think about how opening the pathway for bright minds would benefit the U.S. in development of AI, and at the same time, depending on where they come from, perhaps actually hindering our adversaries?

Dr. JENSEN. Any time you bring people in, as you well know, right? When you bring someone into the SCIF, you always assume risk. So, the question becomes: what procedures do you put in place to analyze the risk versus the tradeoff of exchanging that information and honing someone's ability to make a time-sensitive decision.

I tend to view immigration, especially with people with high STEM backgrounds, as an outstanding American tradition and don't just give us your tired and your poor and you're sick, give us your brilliant people who want to come here and make the world a better place. Now, how we integrate them into national service, I think, is a really an extension to what you're talking about. How do we make them see that we're a country that believes in service and it believes in service from the local to the national level and encouraging that? Now, does that mean everyone who maybe has a cousin in the PLA gets a top secret clearance? No, but there's still a lot of ways people can serve, whether in uniform or whether in the government or our society writ large. And frankly, the smarter people we have to look at this critical moment in history, bring 'em.

Senator Moran: Thank you.

Dr. LeCun, I think I'll address this to you. I'm the lead Republican on a Subcommittee that appropriates money for the Departments of Commerce, Justice, and Science. And one of the efforts that the NSF has is the National Artificial Intelligence Research Institute Program. I'd be glad to hear from you, or any of our panelists, critique or praise for the outcomes of the capabilities of that program. And how do we work to see that that program fits in with the majority of research which is done in the private sector?

Dr. LECUN. Senator, this is a great question.

As a person who has one foot in academia and one in industry, what we're observing today is that academia, when it comes to AI research, is in a bit of a bind because of the lack of competing resources. So, one thing I believe is in this bill is to provide infrastructure—computing infrastructure—for academics and other noncommercial scientists to make progress, which I think is probably the best use of money you can have. Another one would possibly be favoring the free exchange of information and ideas to basically improve the diffusion process between industry and academia.

In some European countries, there are programs that allow PhD students who have residency in industry, not just an internship, but spend a significant amount of time, like two or three years during their PhD—. And we actually at Meta have established programs of this type, with bilateral agreements with various universities across the U.S., because it was so successful in Europe that we tried to translate it here. If there was some help from the government for this, that would be absolutely wonderful.

Last thing is access to data, so this is something that Dr. Jensen mentioned in a different context; but certainly research in healthcare, for example, in medicine, could be greatly improved if researchers had better access to data, which is currently mostly kept private for various reasons—complicated legal reasons, that perhaps Congress could help resolve.

Senator MORAN. Thank you. I'm going to try to get one more question in.

Tell me about how custom AI models, and off-the-shelf models. How do they fit together? How can the government best combine commercial off-the-shelf and custom AI technologies to ensure that government is fully leveraging AI capabilities, particularly in the Intelligence Community?

Dr. LECUN. Senator, I think there is the current state of affairs in the AI marketplace; and there is where I think it's going, which is a bit of a guess. And where I think it's going is that it's going to be a little bit like the Internet, in the sense that there's going to be common platforms that are essentially open source produced by companies like Meta and others possibly, with contributions from academia, et cetera. On top of those platforms, commercial products will be produced, which may not necessarily be open source. They may be commercial in various ways, and there will be customized, fine-tuned for particular applications, whether it is in government, in education, in industry, manufacturing, services, entertainment, you name it.

So that's, I think, the model of the future. This is the type of model we're observing today in various domains in software, certainly on the Internet. That's where I think it's going.

Senator MORAN. Thank you. Thank you all.

Chairman WARNER. And now, to the author of the Choke Theory as Dr. Jensen indicated, Senator King.

Senator KING. Well, Mr. Chairman, a couple of minutes ago I went to ChatGPT and asked them to write a poem about the Senate Intelligence Committee in the style of Dr. Seuss. I want to share it with you. This was written in about four seconds, I timed it: four seconds.

[Senator King read the following text into the record:]

> In a chamber up high, beneath the Capitol's dome, there's a Committee that seeks knowledge, they call it their home.

> The Senate Intelligence, wise as an owl, in the pursuit of truth, they go the extra mile.
>
> With hearings and briefings, they gather the facts, no secrets too deep, no hidden contracts.
>
> They investigate, inquire, and they never back down in their quest for transparency throughout the town.
>
> They question the spies, they question the chiefs, to uncover the truth, they provide us relief.
>
> In a world full of secrets where lies often play, the Senate Intelligence lights up the way.
>
> They protect our nation, our values they hold with their wisdom and courage, stories untold.
>
> So, here's to the Senate Intelligence steadfast and true.
>
> In the style of Dr. Seuss, we salute you.

Four seconds. I mean, you've got to experience this to realize the implications and the power that this has.

Dr. Jensen, first. Coincidentally, we had our third report from the Solarium Commission this morning, and I talked about the brilliant work of the staff, and you were a leader in that, and I just want to thank you. The work that you did was absolutely extraordinary. We're now up to about 70 percent implementation of the Solarium recommendations. So, thank you for that.

The word productivity has been mentioned. One person's productivity is another person's job loss.

Dr. Ding, talk to me about that. Are we in 1811 and the Luddites—is this something that is a serious concern? Or is this just the march of human history where tools have been enabling more productivity since the invention of the hammer improved upon the rock?

Dr. DING. It's a great question. I think it also aligns with some of the statements that Vice Chair Rubio was talking about in terms of job displacement.

When I have my national security cap on, the reason why I emphasize productivity and diffusion capacity is because historically great powers have been able to rise and fall based on whether they've been able to leverage and exploit new technological advances to achieve economy-wide productivity growth.

Senator KING. The Manhattan Project would be an example of that.

Dr. DING. Yes. I think for me, it's less about the moonshot's singular technological achievement. It's more about who can diffuse electricity at scale or who can "intelligentize" the entire economy at scale. And to your point, there are debates among economists about whether the future job displacement by AI and robots would be greater than the jobs that would be created by some of these new technologies. I will defer to those works by folks like Daron Acemoglu in terms of job displacement. I think it gets complicated because if job displacement is so severe that it causes internal political cleavages, that could also then become a national security issue.

Senator KING. And it's also hard to project what the gains will be because it's a new area. You don't know how many new jobs will be created in wholly different areas.

Dr. DING. Yes. I also agree with that, and Erik Brynjolfsson at the Digital Economy Team at Stanford has talked about it's very hard to measure the productivity gains from digital technologies in particular. And I think another key point is these general purpose technologies, their effect on productivity takes decades to materialize. So, we might not see the computer in the productivity statistics as economist Solow once said, but eventually, we will see that come.

Senator KING. One of the major spurs of the economic boom, if you will, of the '90s and early part of the century was the integration of the computer into the workforce. That enabled a great economic upsurge of——

Dr. DING. Exactly. And a key is that took more decades to happen than we predicted initially.

Senator KING. Now, Dr. LeCun. A very practical question: how feasible is watermarking of AI-generated images? This is of concern to us, frankly, because we are very likely to be subject to AI-generated false disinformation. Very skillful. Our face, our voice, our gestures—but completely false.

How feasible is it to require that AI-generated images on Facebook or on TikTok or Instagram all be watermarked or labeled in such a way so that the consumer will know that they're looking at something that isn't real? Is that technically feasible? And is that something we should be thinking about here as we're thinking about regulation?

Dr. LECUN. Senator, this is a very timely question. Of course, it is technically feasible. The main issue is for a common standard to be adopted industry wide. So there needs to be a common way of watermarking, invisible or invisible by using steganography.

The fact that the process by which a piece of information has been produced, this can be done with images and video and audio, such that a computer can detect whether the system has been generated by a generative AI system. But the user will have to, not counteract it—will have to use the products to produce it that obey that standard. And so, that needs to be adopted industry wide.

The problem is much more complicated for text. There is no easy way to hide a watermark inside of a text. People have tried to do this by manipulating the frequency of different words, but it's very far from perfect. But for text, is produced by humans. It's not like a photograph, which you can take anywhere. In the end, the person posting the text has to be responsible for that content.

Senator KING. So, we should not have liability protection like Section 230? Publishers should be responsible for what they produce?

Dr. LECUN. Senator, I'm not a lawyer. I know that Section 230 has been crucial for the success and the development of the Internet. But I would certainly be happy to put you in touch with experts.

Senator KING. Chairman, I hope the panel will help us on this watermarking question, because I think that's something we really need to understand and that may well be part of whatever legislation we're developing. We need your expertise on that. The consumer should know what they're looking at. Thank you.

Thank you, Mr. Chairman.

Chairman WARNER. I concur, and I think the notion that there could be seven or eight or ten different standards that each platform chose may not get us there.

With apologies to my friend Mike Rounds, Senator Lankford.

Senator LANKFORD. Well, thanks for apologizing to Mike Rounds that I'm going to be up next. I appreciate that.

So, thanks for your testimony, and thanks for the research and the work that you've already done on this. Part of the challenge, on whether it be watermarking, whatever it may be, is that obviously there's open platforms there that both Meta has produced and that also the Emiratis have produced. And obviously, we don't have authority to be able to tell the Emiratis what they can actually produce for watermarking. So, this becomes their text in that sense, so it becomes much more difficult in this process. Go back decades ago. We're researching how to build rockets and to be more effective, both for space and for military use, and we understand that other countries are working on the same thing—for both space and military use—and to be able to determine the differences. And so, we try to set limitations for that.

We're in a unique position now, where the PRC is also very interested in partnering, and there are current research hubs in the PRC on Open AI that Google and Microsoft both have at this point. The question is, how far does that go and what do we engage with and at what point does that become facilitating someone who may be an economic adversary, that we hope is never a military adversary in that kind of partnership of that kind of research?

So, at Meta, as you all are dealing with this and trying to be able to think through partnerships, whether it be cloud-basing with Alibaba or whether it be actually partnerships with PRC entities for research in the area of AI, how should we approach that from the Intelligence Committee and as just a national security issue?

Dr. LECUN. Thank you, Senator. This is a question I think that Dr. Ding is much more expert at than I am. Meta does not operate in the PRC for two reasons: because the regulations in the PRC about user privacy are incompatible with Meta's privacy principles, and also because the PRC basically wants to control what information circulates. So, Meta does not operate in the PRC.

Senator LANKFORD. Alibaba is a Cloud partner, though, with Meta, aren't they?

Dr. LECUN. So, Alibaba installed Llama 2 and provides it as a service on its Cloud services outside of the PRC, and so I don't know.

Senator LANKFORD. Dr. Ding, you want to be able to help us unravel this? Because this will be a larger policy issue that we've got to be able to resolve.

Dr. DING. Yes, I think it goes back to our conversation about U.S. investment flows into Chinese AI companies. A similar story, a similar debate is happening around—should U.S. multinational technology giants have R&D labs in China, if I were to rephrase your question. Like Microsoft Research Asia in Beijing, which is——

Senator LANKFORD. Or have PRC partnerships that are here in the United States actually, and they do research together on it, knowing full well where that research goes.

Dr. DING. Yes, I think my take is one of the Senators earlier mentioned this idea of running faster. My take is the U.S. can adopt one of two approaches.

One is this Fortress America approach, where we can't let any technological secrets leak to China.

The second is this run faster approach, where we're going to take the bet that our open economy, our open system of innovation—there are going to be some leaks, there might be some partnership that might allow China to get a little bit further in AI than they otherwise would have—but that partnership might also help U.S. companies run faster. So, being able to access global innovation networks and keep abreast of what's happening, not just in China, but the UAE or Israel—I think the advantages of that, and continuing to maintain the openness of those global innovation networks, is always going to favor the U.S. in the long run, in terms of our ability to run faster.

Senator LANKFORD. So, it's going to always favor us in the long run based on what?

Dr. DING. I think there's a couple of historical examples in this space. So, we had similar debates about satellites and, previously, we had a lot of export controls on satellite technology. But over time, we relaxed those controls, because we realized, first of all, this technology is so commercially based and being driven by the commercial sector that Chinese companies were just getting satellite parts from European suppliers or other hubs in this global innovation network of satellites.

Senator LANKFORD. But that's already developed technology. The challenge that we have with the AI side is that we're in the process of still developing in so many areas. When you form a partnership and they're getting it near simultaneous—, that's a very different issue. If you've got a satellite part, piece, or satellite as a whole, it's already been developed, already used. We're already seeing copies of it. We're already seeing other innovation commercially on that. That's different than ground zero.

If we're going to remain a competitive edge, having someone at the table that may be then exporting that in real time out, becomes a real challenge for us on the intel side of things in a relationship issue. It's the reason that we partner with Russia with NASA on technology, on the Space Station. But we're not going to partner with the Soviet Union in the earliest days of all of our work, because that's the innovation side of things. Now, while we're always innovating, it's trying to be able to protect what's first generation. Does that make sense? So, that's an issue long term that we've just got to be able to determine what's the best way to be able to do that. Where do you put limitations, and how do you develop those?

Thank you, Mr. Chairman.

Chairman WARNER. Senator Rounds, he actually asked pretty damned good questions.

Senator Bennet.

Senator BENNET. Thank you, Mr. Chairman.

And I want to thank my colleague Senator Lankford for his questions. I think it's an important thought exercise to consider where we were with space technology ten years ago, when it was ground zero—when it was zero hour—for that technology. I think, you know, at least from my perspective, I think it's very, very clear that our complete lack of export controls, our complete lack of paying attention to the protection of our IP, has allowed China to build something in outer space that's our near-peer competitor, or even worse than that—without the expense that we went through to develop this, without the expertise, and without the society that Dr. Ding has talked about, which I never would bet against either. But I do think that it is a serious problem that we have just spent the last ten years doing—serious issue with respect to space, and I hope we find a way to avoid it here, which I think is your point. I definitely want us to avoid it here.

Dr. Jensen, something you said at the very beginning of the hearing caught my attention and I thought it was worth more elaboration. So, what if you find yourselves in the Cuban missile crisis again, or something equivalent to that? You've obviously thought about that. So, let's talk about that.

What would that look like today, versus what that would have looked like in the early 1960s when President Kennedy was trying to reach the decision that he was trying to make? Khrushchev was trying to reach the decisions he was trying to make. And at least both people were making fundamental mistakes of judgement along the way. In the end, it resolved itself in the best resolution possible for humanity.

What does that look like in an AI-charged situation?

Dr. JENSEN. So, after this meeting, Senator, I'm going to send you the generative AI artwork we did on this to have it imagine Salvador Dali paint an Edward Murrow type news broadcast of that moment, and it is both beautiful and deeply disturbing.

Senator BENNET. I'm going to hang it up right next to Angus's poem.

Dr. JENSEN. Yes. Dr. Seuss meets nuclear war.

Senator BENNET. Exactly.

Dr. JENSEN. So, this came out of actually a series of tabletop games we did for the Defense Threat Reduction Agency.

Senator BENNET. Actually, we should ask ChatGPT, if you're listening, we should ask them to run the scenario for us.

Dr. JENSEN. Just don't ask——

Senator BENNET. [continuing]. But you're going to do it for us, so go ahead.

Dr. JENSEN. I'll do it for you, but——

Senator BENNET. Yes.

Dr. JENSEN. We did this in a tabletop for DTRA—Defense Threat Reduction Agency, thinking about what exactly would a critical crisis moment look like. And one of the most interesting things is there's such a tendency to move faster. But faster isn't better, even when you're analyzing data. And what we found is a lot of the discussion was about how—just as actually Senator Rubio was talking about—you're getting information correct or incorrect, and now it's triggering these very human instincts, right? We're all of a sudden getting afraid, we're nervous, your heart's racing. It's the person interacting with the algorithm, there is no pure machine. What we found is actually maybe our earlier generation of statesmen and women were brilliant by slowing down crisis decision-making.

If you've ever seen the red phone, it's not a phone, it's a telex for a reason, because it deliberately slows down and makes you deliberate. So, what we walked away with is you're going to be fine in the crisis if you know when to slow down and not let the machines speed you up.

Now, how do we get here? Again, this is why at CSIS and the Futures Lab and the ISP program, we're running tons of tabletop—actually, I've seen some staffers here; we invite Congressional staffers out—because if we don't actually think about those moments now in a very human sense—this isn't necessarily "fine tune the algorithm"—it's those interactions.

Senator BENNET. How about if you had replayed on a day-to-day basis the decisions that were made about our nuclear arsenal or the Soviet Union's nuclear arsenal? Have you guys thought about that at all? What that would look like?

Dr. JENSEN. Sure. And so, at the Office of Net Assessment in the late Seventies, early 'Eighties, when they funded the Rand RASP program, where they used old school, expert-system AI to actually do large-scale modeling of what strategic competition looked like, the military balance, and actually to fight entire simulated campaigns.

So, what Senator Cornyn's saying is true. We've been experimenting with variations of the science and the artificial—the intelligence isn't quite there—for generations. And I think we'll continue to do so, because ultimately, crisis decision-making is about people, it's about politics and it's about emotion even more than the models in front of you.

Senator BENNET. You know, we've had this discussion today about a race, in effect, and can we win the race or is China going to win the race? We haven't had a discussion about how do we get to an end state here where AI reflects the values that this democracy supports in terms of freedom, in terms of rights, in terms of free speech and other kinds of things. So that, if there are people that are unfortunate enough on planet earth to live in a totalitarian society where the authoritarian is rolling out that sort of AI system, there is something available to the rest of humanity that is not that lowest common denominator or uncommon denominator.

I don't know the answer to that, but I suspect that has a lot less to do with some sort of race with China than ensuring that as we think about the implementation here, we're doing it in a way that actually is true to those core values that we have. And that, among other things, is going to look a lot different than the rollout of social media over the last 15 years or so, as you said at the outset of this hearing.

Thank you, Mr. Chairman.

Chairman WARNER. In a brilliant lead up to Senator Rounds.

Senator ROUNDS. Thank you, Mr. Chairman.

Listening to Senator King's poem, I wondered if it was a hallucination on the part of ChatGPT at that point as they kindly talked about our Committee, and I'm wondering if it would have been a different poem if it would have been a House member requesting that type of a message on a Senate Committee.

Dr. Jensen, I want to begin just with a question to you about our current state of play. Loitering munitions. I'm thinking back to the

Nagorno-Karabakh War between Armenia and Azerbaijan, September of 2020. Azerbaijan was very, very successful in a very short time period, using loitering munitions to literally identify, using AI—actually, an Israeli drone system that was there in the marketplace, they could buy it.

Can you talk a little bit about just how widely spread and how difficult the battlefield situations are right now with regard to the use of AI?

Dr. JENSEN. I'm going to try to be short, because I could talk to you about this all day, Senator.

I think, actually, Ukrainian President Zelensky summed it up best when he said it's trench warfare with drones. And so, one of the reasons this spreads is because really cost matters. So, if I can get low cost ability to actually move a munition closer to its intended target with a good circular air probability to strike, I'm going to do it. That's just human instincts. If we're in a fight, you got to win.

And why I think you're going to see something really interesting on the horizon, and why we need to get our house order in the United States, is that the only way to make that work is if you allow people to actually tailor their model from the bottom up. We've actually seen this in Ukraine, where non-profits and tech workers had been able to actually train imagery recognition on the fly. If they would have waited for a standard U.S. government certified algorithm and piece of equipment, it would not look the same, right? So, they're able to actually take all that drone footage, quickly to retrain their model in the battlefield, and then revector the attack. And that's why honestly, we're going to have to be honest about how expensive this is going to be. The intelligence collection cost to get a picture of every tank at every time of day, at every angle, just so you reduce the risk, means that you are now going to have constant collection going on to train that model. And if you don't have the people who know how to use it and interpret it, you're going to have the training and calibration up at such a high level, that the bureaucratic time to use will not be there.

Senator ROUNDS. But the point being, it exists today, the cat's out of the bag. And our adversaries or other countries are utilizing it today, and the United States is probably in a position to use it today as well.

Dr. JENSEN. We are in a position to use it, and in the best of American traditions, let's scale it and do it better and more just than the other guy.

Senator ROUNDS. Dr. LeCun, Europe has developed a model right now, or at least, they're in the process through their European Parliament, develop a model to regulate AI. They've identified high-risk categories, along with two other categories of lesser risk. Have you had an opportunity to look at that? And what is your thought about the approach they're taking in terms of trying to regulate AI based on that approach?

Dr. LECUN. Senator, there are principles that are in that bill that are probably a good idea, although, I don't know the details, frankly. I think the Startup and Industry Committee in Europe has been quite unified in opposing that regulation, not because of the points that you're making, which I think are probably good ones, but because of the details of the regulation. Frankly, my knowledge of it is too superficial to make more comments about it.

Senator ROUNDS. Very good.

Dr. Ding, I'm just curious. Weapon systems on call today, as indicated earlier in the conversations, right now we have our weapon systems that, once they're armed, whether it be on one of our ships near our coastlines and so forth—once we've armed a weapon system and they can identify an incoming as being a threat, we currently utilize that, because in some cases, there is no way a human could make the decision as quickly as that machine could.

Is it fair to say that not just us but our adversaries are also using that same weapon system and it's being incorporated in the battlefield today?

Dr. DING. I'll defer mostly to Dr. Jensen's answer to this. I think my slightly different take is loitering munitions, to the best of my knowledge, are not using cutting-edge deep learning, neural-net deep learning advances that Dr. LeCun is talking about.

So, I guess it's AI if you define the term very, very loosely. But like the algorithms that are underlying, the fundamental breakthroughs in the civilian AI space today, to the best of my knowledge, are not being deployed at scale in any military, whether it be the U.S. or any other leading militaries. And I think that speaks to—these technologies take a very long time to diffuse and become adopted throughout different militaries. So, I guess the optimistic view from that is we do have time to hopefully figure some of these things out. That might be a different view than you hold or others in this room hold, though.

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

Chairman WARNER. Senator Ossoff.

Senator OSSOFF. Thank you, Mr. Chairman. And thank you to the panel.

So, Dr. Ding, as I understand the argument you're making, and it's a compelling one, the U.S. has an advantage because of the structure of our market economy and our R&D enterprise, that this technology can be diffused, adopted across sectors more rapidly, enabling us to realize productivity gains more rapidly and so on.

I guess a question for you is, is it about maximizing the rate of diffusion in your terms? Or is it about achieving the optimal rate of diffusion? And what's the difference between those two?

Dr. DING. Yes, it's a great question. I think there is a difference. I can imagine a world where a country diffuses a technology very quickly in the initial years, maybe a technology that's unsafe or harmful and sort of—there's a backlash to that technology, so it doesn't reach that optimal state or that optimal level of diffusion that you're talking about in the long run.

So, for me, when we're having discussions about AI regulations, oftentimes it's framed as regulation will hamper diffusion and maybe in the short term, it might reduce the speed of diffusion. But I think smart, sensible regulation that ensures more trustworthy AI systems, safer AI systems, in the long run will get us to that more optimal rate of diffusion that you're talking about.

Senator OSSOFF. And by the way, I think the conversation is a little bit overweight, risked to the point where we are in jeopardy of unduly restraining the diffusion of some productivity-enhancing or research-advancing capabilities. But in addition to the risks associated with economic displacement, which you spoke to earlier, Senator Rubio spoke to earlier, what are the other specific risks that you anticipate could emerge from too-fast diffusion and adoption?

Dr. DING. Some of the ones we've talked about today with regards to the use of these large language models; for example, for disinformation, misinformation, to enable propaganda at scale. I think some of the risks that we haven't talked about today, we've seen examples of algorithms that have mis-specified reward systems.

So, there's an example of an open AI system that is about like this boat trying to go through this course as fast as possible, and it's driven by an AI model based on reinforcement learning. Because the programmers have mis-specified how the model should learn, the model learns that the best way to accumulate the most points in the fastest amount of time is just to crash the boat immediately. And that kind of creates some sort of—

Senator OSSOFF. So, short-run, long-run incentives, for example.

Dr. Jensen, what are some specific potential applications of this technology in the conflict avoidance/risk mitigation—? We've talked a lot about how it can make militaries faster, better informed, more lethal. What are the applications that are applicable in institutional structures analogous to the red phone, the hotline, or the now moribund U.N. Security Council, or the kinds of verification regimes that were in place around the Test Ban Treaty and so on? Open Skies?

Dr. JENSEN. Yes. Before we get to that hard intelligence problem set, I know that you have an interest in human rights in this as well. So, think about what a peacekeeping mission would look like if I could actually tailor my messages when I'm doing key leader engagements and meeting with different stakeholders.

I think that first, at that most basic level, it's not just—. You know, the best wars you win are the ones that you never fight, without ceding the advantage. And usually, that requires a degree of actually managing crises all over the world. So, I think there's actually a whole way we could use those in peace building and development as well. When we go to the harder targets that you're talking about, which deal with hard-target problems where an adversary is deliberately hiding something—so I'm using Open Skies—you can see me looking, but we're playing this kind of compliance game. I have to expend energy to hide.

I think there will be very clever ways of building models that simulate some of that, or even help you analyze some of the data. But again, I would even be okay if we could have the simple models Dr. Ding is talking about. I would be okay if we could even just have some basic imagery recognition and accelerate it faster than we currently have, and then, get to the other stuff.

The most beautiful AI thing I've seen is watching a military officer—when we introduce this in the classroom—try to write their Commander's Intent. The most personal thing any military officer will do—because you are responsible for your actions—is to write your Commander's Intent. Well, why wouldn't I want to have a dialog with the corpus of military history and look at different ways it was worded and different ways I could hone my own voice as I took responsibility for using force?

Senator OSSOFF. Mr. Chairman, one more question, if that's all right.

So, Dr. LeCun, I'm not entirely sure what exactly you're proposing in your discussion of the merits of open source systems for the development and diffusion of the technology.

Are you suggesting that it should be mandated that models are based upon and licensed on open source principles?

Are you suggesting that it's simply preferable if developers use open source ethics and guidelines in the development and licensing of their models?

What exactly do you mean when you advise us that this is desirable?

Dr. LECUN. Senator, thank you for your question, which I'm personally very, very interested in.

I think it should certainly not be mandated. It should not be regulated out of existence. There are people who are arguing that AI technology, particularly in the future, will be too dangerous to be accessible. And what I'm arguing for personally and also Meta's policy, is, on the contrary, the way to make the technology safe is to actually make it open, at least the basic technology, not the products that are built on top of it, to ensure American leadership. Because this is the only way we know to promote progress as fast as we can and stay ahead of our competitors. So that's the first point.

Then there is imagining a future in which AI systems reach the level of human intelligence; for example, let's say a decade or two from now—the number may be wrong—every one of our interactions with the digital world will be mediated by an AI system. All of us will have an AI assistant helping us in our daily lives all the time. You're familiar with the situation because you have staff working for you. So, this would be like having a staff of artificial people, basically, who are smarter than you, possibly. I'm familiar with having people who are smarter than me working with me.

Senator KING. What about the impact on job loss?

Dr. LECUN. What I learned about this is from Daron Acemoglu, whose name was mentioned by Dr. Ding before.

I think we have no idea what the major jobs that we'll hire in 20 years will be. There will be new jobs; we can't imagine them today. But to continue on this picture, everyone's information and interaction with the digital world will be mediated by one of those AI system, which basically will constitute the repository of all human knowledge. That cannot be proprietary; it's too dangerous. It has to be open. It has to be open and contributed to by very wide population the way, for example, Wikipedia is produced through crowdsourcing. That's the only way to have enough of a diverse set of views to train those AI systems. They need to be able to speak all of our languages, know about all the world cultures. And this cannot be done by a single private entity. It will have to be open.

And it will occur, as long as it's not regulated out of existence, because it's the most natural way things will evolve, the same way they've evolved with the Internet. Internet has become open source because it's the most efficient model.

Senator OSSOFF. Thank you.

Chairman WARNER. I've got a couple more questions, but Senator King's got one. I've got one, Jon, I'd like you to stay for.

Senator KING. As you know, we've been doing a lot of work on AI. We had the big forum last week with some amazing folks.

I want to thank you all. This has been a very informative panel.

Here's my question: we're in the legislation business. What's the problem? Everybody's talking about legislation, we've got to do something with AI. I would like to ask you to do some homework and give us: here are four things that the Congress should address with AI. Is it watermarking? Is it job displacement? Is it copyright? What I'm searching for—the problem that we're trying to solve. Because until we know exactly what we're trying to solve, we can't begin to write legislation.

So, you're in a position to tell us or to suggest to us what you think the major problems we should be addressing are. That's my questions.

Thank you.

Chairman WARNER. Let me echo what everybody said. Very informative. I first of all, I've got to take a couple of quick shots, fair or unfair, but Dr. LeCun, when you were talking about all of the steps that Meta went through before they released their large language model, it seemed appropriate, but I think back from that point, long before you were with Meta, if we use that same testing model the initial Facebook product, you know, I'm not sure you could have known the downside implications.

Facebook was going to be tested. Does it bring people together in a social setting? Yes, we did pretty good. But I don't think there was a malicious understanding that it might also lead to people's mental health issues that exponentially come out of being so dependent upon this social connection. So, I don't know how you fully test everything on the front end. I think about the fact of this connectivity, much of which this Committee exposed when foreign entities, Russia in particular, used these platforms that did not have that intent, through the use of bots and other things to create huge downsides.

I think about fact that—and this is in one of the closed models and I'm wide open on this open/closed analogy—but I'm really concerned that we've already seen, say with ChatGPT, that with very little push and pull, all of the guardrails of protections that were put in place didn't stand the test and you were very quickly starting to have a ChatGPT model give out answers that were way beyond its scope.

Now, that in a sense, I would argue a reason for open, because you have much more of the testing with the white hat hackers. And so, I'm really worried—and this is one of the areas where, and understanding Dr. LeCun's comment, you don't want to regulate out of existence. But do we really want to trust that the risk-benefit analysis should only be done by the vendors who may or may not have the same long-term societal goals that are part of our responsibility?

I'm not sure people fully appreciated Dr. Ding's comment, which I've been briefed on by much smarter people than myself, on the boat example. The programmer's goal was to create a tool that could show how this boat could destroy as many bad guys as possible. Well, the model soft tested in a way that thought about the problem in a way that obviously a programmer never assumed, and it figured out that the best way to score the most points was actually being self-destructive—ramming directly in, not avoiding the adversary, but ramming directly into it.

And I don't think this is too big of a stretch, but it's a little bit like how in "2001: A Space Odyssey", that if you don't program it right and how smart can any of the programmers be—I'm not saying, to Dr. Jensen's point, that assuming the government bureaucrats and politicians are going to be smart or—may not be the right presumption either. But I do think there is an underlying question here: should we simply trust the vendors alone to make these determinations?

And I'd like everybody to address this, and then I'll have one other wrap-up question.

And at the same time, if we presume that here may be some level of outside trusting or outside scrutiny, some available to have routinizing, enforceable approach to make sure that these models have been tested, should we have that?

And frankly, it goes back to my initial questions about definition. If we said that the tools that fell into the AI category had to go through this testing, you might have a whole bunch of these AI tools that are simply just advanced computing that are using AI now as a marketing tool to redefine themselves.

So, it's a long way around the kind of the base question, which was what Senator King is asking, of—should we just trust the companies to do all the testing on their own? Because if we don't get the right questions asked, I think Dr. Ding's suggestion of the boat self-destructing is only one step away from HAL in "2001: A Space Odyssey."

You want to go down the line or how do you want to——

Take it away, Dr. LeCun.

Dr. LECUN. Thank you, Senator for your question. I think there are three or four questions, if I understood correctly, in your remarks.

The first about should vendors be trusted. I mean, that's why we have regulations. When a product is put on the market, for example, a driving assistance system for a car or a medical image analysis system that uses AI——

Chairman WARNER. But you should know that, at least in terms of social media, Congress has done nothing, even though your companies and others have said they'd be willing. But we have done nothing.

Dr. LECUN. I'm well aware of that. So, what's happened is an interesting history that's happened with social networks, which is that some side effects of enabling people to communicate with each other on social networks that occurred, were not predicted, perhaps because of some level of naiveté or perhaps other reasons. But they were not predicted.

But for most of them that were predicted, they were fixed as soon as possible. And so, for every attack, for example, attempt to corrupt elections, there's a countermeasure. Attempts to distribute CSAM, child sexual abuse material content, there's a countermeasure. Attempts to misinform dangerous misinformation, there are countermeasures, deep fakes, et cetera. All of those countermeasures make massive use of AI today. So, this is an example where AI is not actually particularly the problem, it's really the solution. Taking down objectionable content, for example, has made enormous progress over the last five years—terrorist propaganda and things of that type—because of progress in AI.

And so, as long as—again, recycle the old jokes—of the good guy with AI is better than the bad guy with AI, but there are considerably more good people more educated with more resources with AI than there are bad people, and AI is the countermeasure against AI attacks. So, that was the second point.

Third point is I think it would be a mistake to extrapolate the limitation so of current AI systems—current LLMs—LLMs are really good for producing poems. They're not very good for producing things that are factually correct. They're not good as a replacement for . . .

Senator KING. But it can be entertaining.

Dr. LECUN. It can be entertaining, that's for sure, but factually correct is different. So, in fact, I don't think current AI technology, LLMs in particular, could be useful in the kind of applications that Dr. Jensen was talking about, because it's just too unreliable at this time.

Now, this technology is going to make progress. One of the things that I've been working on personally, and various other people, is AI systems are capable of planning and reasoning. Current LLMs are not capable of planning. They're not capable of reasoning. You don't want to use them for defense applications because they can't plan. They can retrieve existing plans that they've been trained on, and adapt them to the current situation, but that's not really planning. It's more kind of memory retrieval.

So, until we have technology that is actually capable of planning in real situations—currently, we have such technology only for games. So, a system, for example that can play "Diplomacy"—we were just talking about this with Dr. Jensen—or play poker or play Go, things like that, those systems can plan, but currently we don't have systems that can deal with the real world that can plan.

That progress will occur over the next decade, probably. I've been calling this objective "driven AI." So, these are AI systems that do not just produce one word after the other like LLMs, but AI systems that plan their answers so that it satisfies a number of constraints and guardrails. These are the AI systems of the future. They're going to be very different from the ones that currently exist. They're going to be more controllable, more secure, more useful, smarter. I can't tell you exactly when they will appear. That's the topic of research.

Chairman WARNER. Well, let me make sure, because I got one other, but I want to hear from Dr. Jensen and Dr. Ding, because I think Dr. LeCun said well, maybe you got to have some preclearance. But I think you're still more saying leave it to the enterprises to decide, not——

Dr. JENSEN. So, I think this is the first major disagreement we had amongst my new friends. I think a lot of people underestimate how important knowledge retrieval is and a dialogue in actual military planning, because so much of good military planning—and I'd actually say that's a good thing about good intelligence analysis, too—is that creative spark and how you're able to start thinking about defining the problem, right?

The whole point is defining a problem in a way that it lends itself to a solution to be transparent. So don't discount even just basic LLMs can augment some very critical parts of military planning right now, before we get to higher-order reasoning.

Senator, as you were talking, to answer your question, I had a thought experiment of the FDA, right? What would this hearing have looked like, thinking about Food and Drug Administration and how you would go about actually regulating what we put into our body? And if it's true that we're going to have most of our interactions—and I think you're right—mediated through digital assistance of varying degrees of being artificial and intelligent, I think the hard regulatory question on you is, like, what counts as things that the companies who produce these products have to report, have to account for, have to be transparent, and then can be actually inspected and certified via some whole new form of AI assurance? And I think it's going to take, sadly, a generation at least to work out that balance of what actually is reported, how you study it, what that looks like. But the analogy I come back to is the FDA. And it's going to be just as important as that, frankly.

Chairman WARNER. And Dr. Ding, and I should make clear I was reminded by my guys, your boat analogy was about a boat race, not about one boat taking out the other. I think the premise, again, that if you didn't ask the right question, the boat took an action that none of the original programmers would have thought, that not winning the race but crashing the boat was a way to score more points.

Dr. DING. I don't think we should expect vendors to be the sole solution to regulating and ensuring safe and robust AI systems. I think even if there are, like, more good guys or good people, oftentimes the boat example is an example of everybody was trying to do the right thing, but you still had these accidents that can occur.

I think I would cite my former colleague at the Center for the Governance of AI, Toby Shevlane. He wrote a lot about structured access to very powerful AI models. He now works at DeepMind, one of the other leading AI labs. And he was proposing ideas about how labs like Meta's AI Lab could open up checkpoints corresponding to earlier stages in the training of these models, to allow outside researchers to study how the model's capabilities and behaviors are evolving throughout training. So obviously, Meta's doing great work on that, in terms of all of these red teaming exercises, but having ways to involve outside researchers might help check against——

Chairman WARNER. We want you to take Angus's task to bear, in terms of what—. But last one. Then if Jon wants to jump back in. It's this question of, with this tool, do we need to over-weigh. Because, you know, I've been spending a lot of time thinking about where is the most immediate threat coming from existing AI tools? And I make the case that two of our institutions that are most immediately threatened are faith in our public elections and faith in our public markets.

Public elections, because First Amendment what have you gets more challenging to think through. But faith in our public markets, there are already laws in place about deception and manipulation of stocks.

What I think, in regard to Angus's question, which was a good one, you know, where's the problem? Well, the problem is we now have tools with AI that allow those manipulation actions to take place at an exponentially, almost unlimited volume that never before took place. Somebody had to cheat on an individual basis or manipulate on an individual basis. But the volume of things that could take place from deep fakes to product misrepresentation to false SEC claims, just the litany—. Because I would argue that there was an overweighed risk of AI taking something that's already wrong and doing it on steroids exponentially.

And if you would agree with that, then the question is: does it need a new law? Does it need a lower standard of proof because the damage could be so great? Is it a higher penalty? Am I going at least directionally the right way?

And I will do it in reverse order this time. We'll do Dr. Ding and back up the row, and then I'll turn it back over to Jon if you want.

Dr. DING. Yeah, I think in terms of the specific legislative proposals, let me take some time and get back to you on that. I would say one of the things here is we often overweigh the role of technology in this. So, you want to protect public markets, you want to protect public democracy. Should we spend all of the time thinking about how AI is going to impose risks on all of these different things? Or should we think about just overall procedures or policies that would incentivize and create a more robust and trustworthy media system, that would then be able to fact check any sort of AI-generated false information?

So, I will get back to you on the specific legislative proposals on the AI side, but I would also emphasize that it's not just about technology, it's about how do we provide the better-surrounding society that can ensure that this technology doesn't undermine those two public goods that you mentioned.

Chairman WARNER. Dr. LeCun.

Dr. LECUN. Thank you, Senator, for the question.

I'll speak to the technological aspects, or the scientific aspects, rather than the legislative ones.

So, this electronic astroturfing that you are describing, which, if I understand correctly, for markets or the political scene, I think what would help with this would be a systematic way of tracing the origin of a piece of content that would be in the hands of, that would be visible for, users. So, watermarking is an example of this for pictorial content. Unfortunately, that doesn't work for text. So, for text, you have to make people who post the thing responsible for the content they produce and for its deceptive character. In the U.S., of course, we have the First Amendment, so you can't stop people from seeing what they want, and you shouldn't. But there is an interesting point there, which is that the main issue with misinformation is not at the creation of misinformation but at the dissemination level.

There are interesting economic studies on this by Arvind Narayanan from Princeton University, who has studied those questions. He seems to think that the problem is that dissemination—and certainly I admit I will agree with this, because we have systems in place to take down dangerous misinformation to limit its impact, on things like vaccines, for example, or any misinformation that endangers the public or the integrity of elections.

So, I think that of all the measures of dissemination of all the measures that can be taken technologically, watermarking for pictorial content and audio content, not for text, and then perhaps some regulatory help.

Chairman WARNER. I'm going to go to Dr. Jensen but I'm just going to quickly add here. I think the tool itself—the distribution of the tool, I understand. But at some point, we have to think about whether the tool—. You know, the world decided that the use of chemical gas was going to carry a bigger penalty than shooting somebody. Both ways, you're going to be dead. And do we need to overweigh the risk because somebody giving a false tip could manipulate the market. But the volume of tools that could be used to mess with the market using these AI tools seems to me to be of a—what's the analogy—but the gas versus bullets, but there's something in—. So, Dr. Jensen, bring us home.

Dr. JENSEN. All right. Task accepted, Senator. Thank you. Your larger question is really a philosophical question about what it means to govern in the twenty-first century. And honestly, answering——

Chairman WARNER. I'm not sure that's going to bring us home.

Dr. JENSEN. Well, we can do it quickly, but honestly, it's important because I don't think necessarily you have to—. We're not going to throw out the beauty of our Constitution. So, we have to think about how we have maintained our values and standards and execute those laws. And you're probably going to have to—you already have. I mean, going through the work you've already been doing, we are going to add acts, we are going to add laws. But I think it needs to be done in open settings like this, an open dialogue that allows you to calibrate how far that goes.

And to be really clear that, obviously, this is not libertarian paradise; there will be regulation and there needs to be, because it helps create common standards, and standards are strategy. They allow everyone to have foresight and think about what's going to happen. How to make business decisions about how to actually tool their algorithms and send them to the marketplace.

So, I think you're probably going to have to create something. I'll definitely get back to Senator King and you and the Committee. I think you're definitely going to have to create something like the FDA, but for AI assurance. I don't know what that is. I'm not usually a fan of creating new bureaucracies in our government. So maybe it's an existing body that does that. Maybe it's an extension of NIST. But I think that we have to all be patient. This has happened faster than we anticipated and we're going to have to work it out over the next generation.

Chairman WARNER. Well, if you could, spend a little time on it, because I think it is a clear—We spent a lot of time on public elections. I don't think we spent near enough time on disruption of public markets. And I think we are an event or two away from what then Congress would potentially overdo, which is overreact.

But because I think the threat could be so great and the tools could be so unprecedented, you know——

A question I pose for you: do you need a new law? Do you simply need a different standard approved? Do you need a higher penalty? And is the potential downside of robbing broad-based faith across the whole market, how do you preempt? Most of these laws and penalties and proofs are after the fact, and I don't know how in some of this you——

Dr. JENSEN. We've already had this in a flash crash in a sense, right? So, if two algorithms are trying to place trades and each of them is trying to anticipate—we've already seen flash crashes. What you're getting at is more the question of intent, when an individual or a network of individuals actively uses these tools, which can be for good, to do broad-based market deceptions. And that's really a straightforward legal question. You hold them accountable, and you prosecute them, and you throw them in jail. And you throw them in jail in a way that still is keeping to the rule of law but that definitely sets an example.

Chairman WARNER. But is there—and I keep coming back to my example I didn't want to use—but there have been times when society have said we've got to so overweigh the risk that we're going to really throw the entire buyer book at you. And I'm not saying we have to go there, but I do worry that the arguments—and I'm not saying from Dr. LeCun—but I have heard this from many: Oh gosh, you're going to regulate away innovation; you're going to kill the golden goose!

And that argument, which was very much the argument that was launched by—and again, I say to my friends from Meta, we've had these discussions many times—but was launched by many of the social media platforms in the late '90s, early 2000s. I don't think there's very many members of Congress, either side, that wouldn't say, you know, gosh, we ought to have some guardrails. Now, there still would be not lack of agreement on what those guardrails are, so I rest my case. But the upside threat and potential, as, Dr. LeCun, you just laid out with all of the AI assistants we're going to have. Boy oh boy, we've got to get it right.

Gentlemen, it was a very stimulating conversation and a good hearing. I very much appreciate it.

We're adjourned.

[Whereupon the hearing was adjourned at 4:48 p.m.]

Meta

1 Hacker Way
Menlo Park, CA 94025
United States

January 12, 2024

Chairman Mark R. Warner
Vice Chairman Marco Rubio
U.S. Senate Select Committee on Intelligence
211 Hart Senate Office Building
Washington DC, 20510

Dear Chairman Warner, Vice Chairman Rubio, and Members of the Committee:

Thank you for the questions for the record from the Senate Select Committee on Intelligence hearing entitled "Advancing Intelligence in the Era of Artificial Intelligence" on September 19, 2023. Per your request, attached are answers to the questions posed.

Sincerely,

Meta Platforms, Inc.

Application of Artificial Intelligence to National Security Missions

1. **What are the privacy and civil liberties implications of AI technologies – particularly given the propensity of large models to memorize potentially sensitive training data, and to generate output 'hallucinations' that could impact military and intelligence decision?**

At Meta, we believe protecting the privacy and security of individuals' data is the responsibility of everyone at the company. We ground our work to ensure that AI is designed and used responsibly around a set of core values, several of which speak directly to privacy and civil liberties.

Transparency and Control: We want the people who use our products to understand and have control around how data about them is collected and used. To that end, we have, for example, crafted an introductory, in-product experience to help people understand generative AI features when they first engage with them. This experience includes details about the limitations of our generative AI features and offers links to additional information about how we collect data for those who want to learn more.

Privacy and Security: After engaging with civil rights advocates and human rights groups, we identified new approaches to help us access data with the potential to meaningfully measure the fairness of the AI models on our platforms across races. That is why we have built a cross-product Privacy Review process, which is designed to help identify and mitigate privacy risks, including in features and products driven by AI.

Fairness and Inclusion: Fairness and Inclusion is one of our central pillars for responsible AI development. In the context of AI, we have developed and are continually improving tools and processes to help our machine learning engineers detect certain forms of potential statistical bias in certain types of AI models and labels commonly used at Facebook. In addition, Meta's Responsible Innovation team and Diversity, Equity and Inclusion team both facilitate input from a wide range of external experts and voices from underrepresented communities, to advise teams about how particular communities may be impacted by their new products and how existing products might be improved.

While there are many exciting and creative use cases for generative AI, it will not always be perfect. The underlying models, for example, have the potential to generate fictional responses or exacerbate stereotypes it may learn from its training data. With respect to national security, if a military or intelligence body is choosing to use an AI system, protocols and a risk management approach should be instigated for individuals utilizing the models.

By open sourcing our AI tools, we aim to work and collaborate across industry, academia, government, and civil society to help ensure that such technologies are developed responsibly and with openness to minimize the potential risks and maximize the potential benefits.

Artificial Intelligence and Disinformation

One of the primary concerns driving public policy debate on AI is the threat of AI-enabled disinformation operations. Recently, the *New York Times* published an article about a purported PRC-campaign using AI to generate images of the wildfire destruction in Maui to push a disinformation campaign that the wildfires were caused by the testing of a weather-based weapon.

2. **How will users know whether or not a particular image or message was generated by AI versus a human being on platforms that Meta controls?**

The entire industry is still working to determine how to use and share signals, such as metadata or invisible watermarks, or otherwise automatically detect whether photorealistic content that people share is AI-generated.

Our approach to addressing manipulated media involves engaging with academia, government, and industry. Earlier this year, we were pleased to join the White House's voluntary commitments alongside others in the industry, including a pledge to develop robust technical mechanisms to identify AI-generated content, such as digital watermarking with respect to frontier models. These commitments are an important first step in ensuring responsible guardrails are established for AI. Further, we have collaborated with global experts with technical, policy, media, legal, civic, and academic backgrounds to inform our policy development and improve the science of detecting manipulated media.

In addition, we have taken steps to detect and action misleading manipulated media, including making it harder for people to spread misinformation with our tools. Images created or reimagined by Meta AI as well as our AI media editing tool backdrop have visible markers so people know the content was created by AI. We are also developing additional techniques to include information within image files that were created by Meta AI and invisible watermarks to the imagine with Meta AI experience. We intend to expand these to other products and experiences as the technology improves.

We also recently announced a new policy to help people understand when a social issue, election, or political ad on Facebook or Instagram has been digitally created or altered, including through the use of AI. Since 2018, we have provided transparency for ads about social issues, elections, and politics, and we continue to expand on these efforts, including through this new policy. We have long believed in the importance of transparency to bring more accountability for Meta and our advertisers, and this is especially important for ads that can influence the way people think, act, or vote.

Finally, on the activity described in the article referenced in your question: in addition to taking steps related to content generated using AI, we also continuously look for and enforce against adversarial deceptive behavior signals by covert influence operations – whether or not they post GenAI content. This particular activity originated in China, and we linked it to the largest, albeit unsuccessful, operation called Spamouflage that we reported on in detail in our Q2 2023 Adversarial Threat Report.

Addressing the challenge of AI-driven manipulated media requires a whole-of-industry approach. Specifically, we will work with industry peers to align on technologies that can make it easier for us and other platform providers to detect when someone shares content that has been AI-generated. This approach will also pose challenges, as new companies creating AI tools will constantly emerge. Moreover, bad actors will find ways to circumvent our detection capabilities. To that end, we continue to partner with the Partnership on AI in the hope of developing common standards for identifying and labeling AI-generated content, as well as mitigating deceptive AI-generated content, across the industry.

3. **What safeguards can or should be implemented to prevent AI capabilities from being used to generate disinformation? Should these rest with model vendors, social media intermediaries, end-users, or a combination of these entities?**

We understand the importance of building safeguards into our AI products from the beginning so that people can have safe, and ultimately more enjoyable, experiences. As AI technology continues to evolve, safety features and controls will also have to evolve. That is why we work in collaboration with stakeholders across industry and academia to make sure that AI systems have responsible guardrails.

While we recognize that AI can be used for nefarious purposes, we believe that AI is part of the solution to defend against those risks, and that new generative AI tools can help companies and others get better at identifying and stopping the spread of content that violates our policies. Meta has dedicated significant resources to detecting content on our platform that violates our policies, including those regarding misinformation. Our investment in this area has allowed us to build industry-leading technologies to proactively identify content, prioritize the most critical content to be reviewed, and act on content that violates our policies. Technology has and will continue to play a central role in our content enforcement operation, and we will continue to invest in improving our technology to fight misinformation.

We are also optimistic about the possibility of using newly developing generative AI tools to help us enforce our policies in the future. We have started training large-language models (LLMs) to help determine whether a piece of content violates our policies. At the same time, we also incorporate several safety steps to improve the model's ability to perform and mitigate risk of toxicity, bias, and other vulnerabilities. We are already seeing promising signs the LLM can perform better and more safely than existing machine learning models. We are hopeful the use of generative AI can help us take down more violating content faster and more accurately than existing AI tools.

We also use LLMs to remove content from review queues in certain circumstances when we're highly confident it does not violate our policies. This frees up capacity for our reviewers to focus on content that is more likely to violate our policies. Finally, AI also powers our automated tools looking to spot fake accounts and spam-like activity when someone posts at great frequencies or rapidly friends thousands of users at once.

There is still much work to be done to keep evolving these systems and scaling our trust and safety work effectively, but LLMs have the potential to be game changing in this space - lowering costs, improving ability to detect and adapt to changes, and increasing efficiency.

Addressing the challenge of AI-driven manipulated media (*e.g.*, "deep fakes") requires a whole-of-industry approach. That is one reason why we joined the White House's AI Commitments. Specifically, we will work with industry peers to align on technologies that can make it easier for us and other platform providers to detect when someone shares content that has been AI-generated. This approach will also pose challenges, as new companies creating AI tools will constantly emerge. Moreover, we know that bad actors will continue trying to find ways to circumvent our detection capabilities. To that end, we continue to partner with the Partnership on AI, in the hope of developing common standards for identifying and labeling AI-generated content, as well as mitigating deceptive AI-generated content, across the industry. In particular, we support efforts to develop industry standards about how and when to apply watermarks to photorealistic images – and we think this is a place where Congress can help drive the consensus forward.

In a *New York Times* interview you suggested that the problem with disinformation is not its creation but rather its dissemination – and therefore efforts should focus on limiting distribution. We know from many past Adversarial Threat Reports from your company that a key stumbling block for many foreign actors is creating *credible* English language content – so creation does seem like a major stumbling block for foreign bad actors absent AI.

4. **Doesn't the public availability of these potent tools enhance their ability to make credible, hard-to-detect disinformation content?**

This issue is not unique to Meta, but rather, it is one that stakeholders across the industry must work to address. So while this is a pivotal moment for AI technology, it is not the first time bad actors will try to use developing technology to their own ends, like phishing or writing exploits to attack a computer system.

For our part, we have been working for years trying to stop bad actors from using our technology to spread mis- and disinformation. We remove content that violates our policies, regardless of whether content is created using AI. In other words, all of our policies apply to AI-generated content including policies on voter interference, misinformation and harm, and more. Meta is also working with fact-checking partners to learn what they are seeing about AI-generated misinformation and will continue to solicit feedback about how the program is working in regards to this developing subject.

Our enforcements against determined adversaries like covert influence operators focus on their adversarial behavior, not the content they post – whether it is created using AI or not. In fact, we have seen examples when the use of AI-generated images by coordinated inauthentic behavior (CIB) campaigns tipped off researchers, including Meta's, to expose networks of fake accounts behind them. Further, by the time influence operations may leverage AI to create and post content, they will already have left numerous adversarial behavioral signals, such as how they acquire, disguise and coordinate their fake accounts, all of which can aid threat research and disruption efforts.

Overall, the use of manipulated and decontextualized media as part of malicious campaigns is the latest manifestation of a long-standing tactic deployed by influence operations (dating back to before the internet) that try to plant false content like forged documents or photoshopped images in an attempt to shape perception. For tech platforms, enforcing against this tactic means disrupting these operations according to existing policies against adversarial behaviors.

While Generative AI has increased public attention on synthetic media issues, we have not seen these technologies widely and successfully used by threat actors behind CIB networks we have disrupted to date. We have seen (and enforced-against) multiple examples of synthetic media leveraged in recent years by threat actors, including:

- **Images**: In 2019, a network we took down in Vietnam, and the United States became the first known CIB operation to make use of profile pictures created using techniques like Generative Adversarial Networks (GAN). In 2022, more than two-thirds of all the CIB networks worldwide that we took down featured accounts that likely had GAN-generated profile pictures, in an apparent attempt to build credibility and appear unique.
- **Video**: While it was not part of a known on-platform operation, in early 2022, we identified and removed a "deepfake" video that appeared to show Ukraine's President Zelensky making a statement he never did. It appeared on a reportedly compromised website and then started showing across the internet. In another example, earlier this year, we took down a cluster of accounts linked to the China-based influence operation "Spamouflage" that was reported by researchers at Graphika to have used AI-generated newsreaders in their videos on social media platforms including Facebook, Twitter, and YouTube. These early attempts at using AI-generated videos were quickly identified and exposed. To date, we have not seen evidence of this tactic being widely and successfully used by CIB networks we disrupted, including sophisticated threat actors.

- **Audio**: Recently, we detected and removed a cluster of commenting activity by Spamouflage targeting audiences in Canada. Researchers at ASPI described this operation's use of likely GenAI audio in a doctored YouTube video shared on other platforms, with "zero or minimal engagement with real users."

While generative AI does pose challenges for defenders, at this time we have not seen evidence that it will upend our industry's efforts to counter covert influence operations – and it is simultaneously helping to detect and stop the spread of potentially harmful content. Our teams are monitoring these risks while in close contact with industry peers and other experts, and this assessment will continue to evolve. But it is encouraging that our defenses continue to endure, even as technology evolves.

For more information, please see the response to Question 3.

Global Competitiveness and Artificial Intelligence

Understanding the competitive landscape is a priority for U.S. policymakers.

5. **Which firms or research labs based in the People's Republic of China do you view as world-leading?**

We are not in a position to assess the capabilities of Chinese-based firms or research labs in a meaningful way.

Meta developed one of the most powerful software development tools for training and configuring AI models with its so-called PyTorch framework, which is used worldwide. In September of last year, however, Meta transferred control over the project to the non-profit Linux Foundation.

6. **Why did Meta transfer control of the project to the Linux Foundation given U.S. national security concerns that U.S.-developed software could enable foreign adversary AI efforts?**

Exploratory research, open science, and cross-collaboration are foundational to Meta's AI efforts, and we have experienced first-hand how innovation in the open leads to technologies that advances the industry. For example, PyTorch has become one of the leading platforms for AI research as well as commercial production use with over 18,000 organizations using PyTorch. At Meta, PyTorch powers 50 on-device AI models across different mobile applications.

We believe that it is important for the United States, and specifically, US companies, to lead the way in creating the foundational AI tools and models that will be used the world over. It is key that the United States and other countries that share our values set the standard. To continue to lead and obtain the benefits of AI as quickly as possible, we need a large community of experts and contributors to quickly exchange ideas and techniques. The bigger the community that contributes to it, the faster it advances.

PyTorch was built with an open-source, community-first philosophy and that has not changed with the transition to the PyTorch Foundation, which is part of the nonprofit Linux Foundation. When researchers and developers open-source their code, others around the world can share their work, learn from each other's advances, and then contribute back to the AI community. The creation of the PyTorch Foundation helps ensure that decisions are made in a transparent and open manner by a diverse group of board members for many years to come. We made this change with a firm belief that contributors would benefit from the robust governance, diverse leadership, and additional investments provided by the new PyTorch Foundation partners. The PyTorch Foundation strives to adhere to four principles: remaining open, maintaining neutral branding, staying fair, and forging a strong technical identity. One of its main priorities is to keep a clear separation between the business and technical governance of PyTorch.

At Meta, we continue to invest in PyTorch and use it as the primary framework for our AI research and production. The transition itself will not mean any changes to PyTorch's code, core project or developer operating models. We have worked consistently to nurture the community-driven growth that has fueled PyTorch's success, committing hundreds of engineers to the framework and supporting product development and community outreach.

Huawei was one of the leading sponsors of this year's PyTorch developers conference held in San Francisco.

7. **What role do you envision Huawei playing in the PyTorch ecosystem?**

PyTorch is an open-source machine learning framework based on the Torch library, used for applications such as computer vision and natural language processing. Pytorch is now part of the Linux Foundation umbrella. Huawei joined Pytorch as a premier member, as described on Pytorch's website, and is able to access open-source material in the Pytorch ecosystem.

Unlike many of the other leading AI companies, Meta has openly published the LLaMA model for anyone to download and modify.

8. **What geopolitical factors did you or Meta consider prior to releasing the LLaMA model publicly? In the long-term, do you think open or closed models will advance more quickly?**

We believe Llama 2 and Code Llama are important advancements in generative AI, and we are excited about their potential to drive innovation.

With respect to advancement in the long term, we believe that openness will lead to better products, faster innovation, and a flourishing market, which benefits us as it does many others. Open LLMs, like Llama 2, make it possible for businesses to participate and advance the AI industry without large amounts of funds, computing resources, or technical expertise. While it is true that businesses can use "closed" models (like OpenAI's GPT-4), building AI applications on top of closed models means that the developer will forever be beholden to the model developer (*e.g.*, OpenAI), who will continue to receive usage fees and continue to exercise control over the infrastructure upon which the applications rely.

Open models, by contrast, provide businesses with the foundation they need to innovate more quickly, free of charge (businesses must of course, comply with a set of usage guidelines and license terms). This means more independent innovators, less gate-keeping, more competition, and ultimately a more diverse AI industry.

As the FTC itself recently explained, "open-source pre-trained models" like Llama 2 can help prevent "a market where the highest quality pre-trained models are controlled by a small number of incumbents" particularly when the open-source models available are of relatively high quality.[1] Open models can therefore "open up the playing field" towards quicker advancement and innovation.

We also understand that there are risks with opening access to this foundational technology – and we have taken several steps so people can use Llama 2 responsibly, including by restricting access to Meta's webform in certain jurisdictions.

Llama 2 and Code Llama are available via its website globally, except in comprehensively sanctioned jurisdictions (Iran, Cuba, North Korea, Syria, and Crimea, Donetsk, and Luhansk regions of Ukraine). Users based in Belarus, China, and Russia are also blocked from accessing the webform required to download Llama 2 or Code Llama on our website. In addition to these geographical restrictions, there are limitations to how Llama 2 and Code Llama can be used and deployed, mitigating risks to national security and economic competitiveness.

9. Will Meta commit to enforcing LLaMA terms of service against malicious uses?

In keeping with our commitment to Responsible AI development, Meta has undertaken a number of initiatives to discourage improper uses of its models. For example, we designed a bespoke license that includes a detailed and thought-out set of use restrictions that strictly prohibit a wide range of malicious uses, while making sure that Meta retains the ability to audit uses of Llama 2 and Code Llama to ensure compliance.

---

[1] *See* FTC, "Generative AI Raises Competition Concerns" (June 29, 2023), here.

Our Acceptable Use Policy clearly states that users may not "[c]reate, generate, or facilitate the creation of malicious code, malware, computer viruses or do anything else that could disable, overburden, interfere with or impair the proper working, integrity, operation or appearance of a website or computer system." Additionally, we have implemented numerous ways of reporting violations of this policy including through reporting issues with the model, reporting risky content generated by the model, reporting bugs and security concerns or reporting violations of the Acceptable Use Policy.

We can use this information to take enforcement actions against individual licensees who violate our Acceptable Use Policy or who fail to comply with audits. Such enforcement actions may result in suspension or termination of the licensee's access to Llama 2 and/or referring violations to law enforcement. Meta is still working through an enforcement program, and we can keep you updated as the program evolves.

○