STATEMENT BY
ALEXANDR WANG
FOUNDER AND CHIEF EXECUTIVE OFFICER
SCALE AI

BEFORE THE
HOUSE COMMITTEE ON ENERGY AND COMMERCE

APRIL 9, 2025

Chairman Guthrie, Ranking Member Pallone and Members of the House Committee on Energy and Commerce. Thank you for the opportunity to be here today to discuss the steps that the United States must take to stay ahead of China in the emerging global race for leadership in Artificial Intelligence (AI). Today's hearing comes at a critical moment in this competition, and I greatly appreciate the Committee's bipartisan interest in ensuring—and advancing—our country's continued leadership.

Introduction

My name is Alexandr Wang and I am the Founder and CEO of Scale AI (Scale). Today's hearing is deeply personal for me, as I grew up surrounded by scientists committed to American leadership in cutting edge technology. I was fortunate to be born and raised in Los Alamos, New Mexico where my parents worked as physicists at the National Lab. During my childhood, I saw firsthand how the development of world-leading technology is critical to our national competitiveness.

After high school, I enrolled at MIT and started studying AI. From day one, I knew it would be the most transformative technology of our time, much like nuclear technology had been for my parents' generation. Early on, I learned that progress in AI depends on three foundational elements: compute, algorithms, and data. Compute is the processing power that drives AI systems. Algorithms guide that compute, turning raw data into large language models (LLMs) like Gemini,

ChatGPT, Grok, Llama, and others. Data is AI's oil, gas, wind, solar all wrapped into one. It is the raw material that enables AI to learn, adapt, and improve over time. When I was in school, most of the field focused on compute and algorithms, and almost no one was thinking seriously about data. That's what pushed me to leave MIT after my freshman year and start Scale. In our early days, Scale delivered high-quality data for autonomous vehicle companies. Today, we provide advanced, expert-level frontier data to leading AI labs like Google DeepMind, Meta, OpenAI, and xAI.

Data is unique because it is fundamentally human. Every major AI advancement is grounded in data shaped by human expertise. For AI model training, we operate a global marketplace powered by hundreds of thousands of human experts dedicated to enhancing the quality and intelligence of data for AI models. This marketplace has created significant economic opportunities for participants, who are located in over 9,000 cities and towns in the United States. The marketplace paid out nearly $500 million globally in 2024.

Apart from providing frontier training data to AI labs, Scale is also helping to prepare enterprises and the U.S. Government for an Agentic World. This will be a world where AI agents, autonomous intelligent systems capable of performing a wide range of tasks, go beyond conversations and recommendations to actively assist humans across every industry.

Scale now partners with global enterprises, including media companies like Time, law firms like DLA Piper, and tech companies like Cisco to train proprietary AI models on their vast data ecosystems and provide solutions for building agentic AI applications.

Scale also partners with the United States Government, particularly the Department of Defense, to equip our military with cutting-edge commercial technology for their missions. One such example is a recent prime contract that was awarded to Scale by the Defense Innovation Unit (DIU) for Thunderforge. This flagship AI program will see agentic AI leveraged for US military planning and operations. As part of our work with the United States Government, we operate a St. Louis AI Center, which produces millions of high-quality geospatial data annotations, builds out Generative and Agentic AI applications, and has created hundreds of AI jobs in the greater St. Louis region.

As AI continues to advance, and we shift from chat-centric Large Language Models (LLMs) like ChatGPT to new agentic applications, making sure that we keep humans at the center of all major advancements is imperative. AI should always work for us, not the other way around.

The Chinese Communist Party's AI Master Plan

I saw first hand the beginnings of the Chinese Communist Party's (CCP) AI master plan in 2018 when an investor invited me to visit. I witnessed how advanced their AI capabilities were, and how they were leveraging them in a variety of useful ways, but also to surveil and suppress their population.

Later that same year, China President Xi Jinping declared China's plans to dominate AI by 2030[1] and they are working to do just that. From Scale's perspective, today, China leads on data, the United States leads on compute, and the two are tied on algorithmic development. It's a neck-and-neck race, which is why the policies this Congress promotes will likely determine the outcome.

The master plan that I saw pieces of seven years ago is now much clearer. It includes:

- ***A whole of government approach:*** In order to compete more aggressively, and win, the CCP has launched their AI plus initiative[2] to implement a "whole of country" approach, utilizing its government, industry, and military to accelerate efforts to become the global standard for AI. The CCP understands that, if deployed properly, AI can serve as the engine of efficient government and—for the first time in history—China is benchmarking its AI investments against leading technology companies rather than the United States government.
- ***Widening their lead on data:*** When it comes to building out their data pipeline, China has spared no expense. Our analysis of publicly available materials has determined that China spends billions per year on AI ready data alone. The government has also unveiled a new data strategy[3] which includes unlocking all of its vast array of "public" data to train AI systems as

[1] See, https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/
[2]See, https://www.globaltimes.cn/page/202403/1308210.shtml
[3] See, https://babl.ai/china-unveils-comprehensive-plan-to-boost-data-labeling-industry-growth/

well as providing large subsidies for this work through vouchers, tax breaks and other means.[4] In May 2024, the National Data Bureau announced the construction of seven specialized data annotation bases[5]: Chengdu (Sichuan), Shenyang (Liaoning), Hefei (Anhui), Changsha (Hunan), Haikou (Hainan), Baoding (Hebei), and Datong (Shanxi).[6]

- ***Catching up on compute and physical infrastructure:*** When it comes to compute and infrastructure the CCP has continued to advance their domestic capabilities. While export controls have helped – for the time being – maintain the United States lead on compute, China's homegrown chips from companies like Huawei and SMIC are quickly closing the "capability" gap with their foreign competition.[7] In 2025, they're expected to spend at least $110 billion in chip manufacturing[8] and recent studies have shown that between 2018 and 2023, Chinese authors published 50% of the published research papers on foundational chip research compared to the only 22% published by United States' authors.[9]

  The CCP has also long recognized that winning in AI also requires a robust government strategy to ensure the availability of energy and infrastructure.

[4] See, https://english.www.gov.cn/news/202501/14/content_WS67859ba1c6d0868f4e8eeca1.html
[5] See, https://www.globaltimes.cn/page/202404/1309974.shtml
[6] See, https://dub.sh/SySgFXA
[7] See, https://s201.q4cdn.com/141608511/files/doc_financials/2025/q4/177440d5-3b32-4185-8cc8-95500a9dc783.pdf
[8] See, https://www.reuters.com/technology/china-seen-leading-chipmaking-investment-again-2025-semi-group-says-2025-03-26/#:~:text=China%20is%20the%20largest%20consumer,imposed%20by%20the%20U.S.%20government.
[9] See, https://www.nature.com/articles/d41586-025-00666-3

In response to the Stargate announcement[10], the Bank of China announced a 1.4 trillion yuan investment into AI infrastructure[11] and is setting up a National Venture Capital fund specifically focused around technologies like AI.[12]

- ***Develop leading tech, implement it rapidly and export it to the world:*** Deepseek may have sent shockwaves around the world, but it is not the only advanced Chinese model. Since the launch of Deepseek, China has also released or announced multiple additional open-source models like Baidu's Earnie series[13], Moonshot's Kimi[14], 01.ai's Yi[15] and Alibaba's Qwen[16] each rivaling leading United States tech companies' top models.[17]
  As competition in AI intensifies, China is both rapidly implementing and looking to export itsAI systems. The People's Liberation Army issued 81 contracts with Large Language Model companies in the first half of 2024 alone[18] and similar to their Belt and Road initiative[19] and 5G ambitions[20] the

[10] See, https://openai.com/index/announcing-the-stargate-project/
[11] See, https://www.medianama.com/2025/01/223-bank-of-china-announces-1-trillion-yuan-ai-industry-investment/
[12] See, https://www.reuters.com/world/china/china-set-up-national-venture-capital-guidance-fund-state-planner-says-2025-03-06/
[13] See, https://research.baidu.com/Blog/index-view?id=185
[14] See, https://kimi.moonshot.cn/
[15] See, https://github.com/01-ai/Yi
[16] See, https://www.reuters.com/technology/artificial-intelligence/alibaba-prepares-flagship-ai-model-release-soon-april-bloomberg-news-reports-2025-04-01/
[17] See, https://www.cnbc.com/2025/03/18/baidu-shares-jump-10percent-following-release-of-new-open-source-ai-models-.html
[18] See, https://www.scmp.com/tech/tech-trends/article/3267866/chinas-public-sector-accelerates-ai-adoption-2024-zhipu-and-iflytek-emerge-winners
[19] See, https://www.cfr.org/backgrounder/chinas-massive-belt-and-road-initiative
[20] See, https://www.noemamag.com/the-world-china-is-building/

Chinese Government is also working to export their technology to the rest of the world, and is wholly committed to positioning themselves as the global standard.

If China leads on Generative AI and Agentic AI, there could be serious consequences for the United States. President Trump rightly called the proliferation of Chinese AI a "wake up call."[21] Which begs the question, how should the United States respond?

The United States' Four Pillars to Win—Dominate, Unleash, Innovate and Promote

If the United States wants to beat China, we must not only match but exceed the CCP's intensity on AI. President Trump has already made this aspiration clear, stating: "It is the policy of the United States to sustain and enhance America's global AI dominance in order to promote human flourishing, economic competitiveness, and national security".[22] Global AI dominance is not about trying to level the playing field by mimicking China's Authoritarian way of government, it's about the Administration and Congress working together to take bold steps. Scale believes the path forward must include decisive action on four main pillars—Dominate, Unleash, Innovate and Promote.

Pillar One: Dominate – As evident by China's investments, the country that wins on data will almost certainly win the AI race and this has been made abundantly clear by nearly every recent AI advancement being rooted in a data advantage.

[21] See, https://www.nbcnews.com/tech/innovation/trump-china-deepseek-ai-wake-call-rcna189526
[22] See, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/

Over the past few years, we've seen Congress and Administrations prioritize the onshoring of chip manufacturing as well as renewed focus on domestic energy production. This should help us maintain our lead in chips and algorithms, but it won't help the United States with data.

The United States Government is already one of, if not, the largest producer of data quantity and diversity in the world, but today nearly all of that data is unavailable to improve our AI systems in contrast to the Chinese approach. If we want to win on AI, we must turn our data asset into an advantage. To achieve Data Dominance, we need a whole of government approach, which should:

- ***Establish a National AI Data Reserve:*** The National AI Data Reserve would serve as a centralized data hub for all of the government's AI programs to leverage and should include all relevant government data. This would allow for the data to be easily shared between agencies and be leveraged for widespread AI adoption. The Department of Defense is currently working towards its own version of this, but if the United States wants to lead, this must be government-wide.
- ***Make all government data AI-ready:*** AI is only as good as the data that it is trained on, and without AI-ready data, no use case will be successful and we will never be able to truly leverage the technology.
- ***Stand up government-wide AI data infrastructure:*** Once the data is made AI-ready, this commercial best practice makes it so that each Agency can

tap into the Data Reserve and efficiently scale AI programs in the most cost effective manner.

- ***Invest in AI ready data as a national priority:*** The government must invest both in continued investment to keep feeding data into that National AI Data Reserve and to incentivize our private sector to create an asymmetric data advantage. The CCP has recognized this, and Congress should look at all tools at their disposal, whether that is tax credits or stronger private-public partnerships, to incentivize Data Dominance.

Pillar Two: Unleash - In order to win, the US needs to unleash our technology. In Scale's direct work with other Governments, such as Qatar[23] and our observations of developments in Singapore, the United Kingdom[24] and elsewhere, other governments around the world are actively implementing AI across their public sectors much faster than governments at all levels in the United States.

To date, no Federal Agency has truly made progress implementing AI in any meaningful way, despite identifying over 1,700 use cases[25], and developing many research reports and pilots. Fortunately, we still have time to catch up, but the window is closing. This starts with the National AI Data Reserve laying the foundation for applications like AI agents and agentic capabilities.

For private enterprises, AI Agents have already proven to provide strong results for things like coding and customer service applications. Identifying and implementing agentic applications for the government should be no different. If

[23] See, https://scale.com/blog/mcit-scale
[24] See, https://committees.parliament.uk/work/8367/use-of-artificial-intelligence-in-government/
[25] See, https://fedscoop.com/federal-government-discloses-more-than-1700-ai-use-cases/

done correctly, it could revolutionize government efficiency, allowing humans to shift from manual decision making to a role focused on providing oversight of AI systems. This will free up public employees to think more strategically and ensure our democratic values are reflected in the systems our governments develop. Agentic Government will speed efforts to reduce the regulatory backlog, streamline permitting processes, and provide the Department of Defense asymmetric advantages over adversaries on topics like planning and OSINT collection.

All of this is possible today, but in order to implement it, Congress and the Administration must establish an ambitious, coherent, and integrated AI strategy. To do this, Washington should:

- ***Require each Agency to stand up at least one flagship Agentic Government program****:* Due to the over 1,700 identified use cases, agencies clearly understand where AI can be most effective, but are still stuck in a research and pilot mindset. This must change and Congress should require that every agency stand up an agency-wide agentic program no later than one year from now.

Pillar Three: Innovate – Scale has long supported the position that the right regulatory framework is one that maximizes innovation, but still creates proper guardrails. For the United States to continue to be the most attractive place for leading tech companies to develop the next generation of cutting edge AI systems, our government must create the best environment for them to do so.

To ensure that the United States is set up to be just that, Congress and the Administration should:

- ***Confirm a sector-specific and use-case based regulatory framework:*** This approach rightly governs the outputs and mirrors how we have always governed technology. For example, a laptop can be used for a number of different purposes including enabling research and connectivity, but it can also be used for hacking. However, the government does not regulate the laptop, it regulates the malicious use of the laptop. This is the same approach the US should adopt for AI.
- ***Ask agencies to identify regulatory gaps and fill them:*** To adequately implement this approach the United States must confirm that our regulatory system is sufficient for AI. To do this, a thorough understanding of where gaps may be in the regulatory system is critical and Scale strongly supports Congress working with the Administration to conduct a comprehensive regulatory gap analysis which would highlight if one exists. If shown to exist, actions must be taken to provide regulatory consistency.
- ***Clarify one single federal AI governance standard:*** In 2024, there were more than 700 different AI bills introduced across nearly every state[26] and it is anticipated that this number will grow in 2025. Beyond the contents of the bills themselves, we have seen inconsistent definitions of key aspects of the AI ecosystem used by each state as well as mechanisms to prove regulatory compliance.

[26] See, https://www.bsa.org/news-events/news/2025-state-ai-wave-building-after-700-bills-in-2024

- ***Put policies in place to let the AI workforce thrive in America:*** Innovation is enabled by people, and much like coding in the 1990s, acquiring the skills necessary to train, fine tune, and evaluate AI systems will be vital to future innovation. These skill sets did not exist a decade ago, but are already proving to be important drivers of economic activity. If our nation wants to build the highest quality AI systems in the United States, they must be trained by people in the United States. Congress and the Administration should confirm that all relevant policies are flexible enough to enable this new kind of work to thrive right here at home.

Pillar Four: Promote - Effective tech diplomacy relies heavily on the access and adoption of US technology by our allies. Over the past few years, the United States has simply not prioritized the tech promotion aspect nearly enough. In the coming year, countries around the world, which we refer to as "AI geopolitical swing states", will be forced to effectively choose between whether they want to deploy Western or Chinese technology.[27]

This is not new and the global roll out of 5G technology offers a stark reminder of what happens if the United States does not step up to lead on new technologies. In the early days of 5G, governments around the world had to choose whether to harness U.S. technology or Chinese technology. This happened because the United States did not actively promote our technology aggressively enough. Now, China's global influence on 5G is much stronger than the United States'.

[27] See, https://www.csis.org/analysis/ai-diffusion-framework-securing-us-ai-leadership-while-preempting-strategic-drift

But tech diplomacy is not a level playing field due to the CCP's willingness to heavily subsidize their technology exports, so the United States needs to look at creative ways that we can win. To do this, Congress and the Administration should:

- ***Resource and empower the National Institute of Standards and Technology (NIST):*** NIST is the world's most well respected science organization. It needs more resources and leverage to be able to complete all relevant measurement science, such as standards and frameworks for AI.
- ***Export NIST's measurement science to the world:*** The Global Network of AI Safety Institutes is made up of 10 countries[28] with the United States currently serving as Chair. This is a ready-made entity to export our measurement science and position it as the global standard. To date, China has not been allowed entry into the body and the United States needs to make sure that remains the case.

Conclusion

America led in the Industrial Revolution, the Space Race, and the Internet Age. AI is the next frontier, and I'm confident that we will lead again. Thank you again for the opportunity to be here today to discuss with you the ways in which the entirety of the AI ecosystem must work together to win. I look forward to your questions.

[28] See, https://time.com/7178133/international-network-ai-safety-institutes-convening-gina-raimondo-national-security/