NVIDIA Lends Support to Washington's Efforts to Ensure AI Safety

NVIDIA joins leaders from the White House, Congress and tech industry to discuss AI standards and best practices.

Author: Ned Finkle

In an event at the White House today, NVIDIA announced support for voluntary commitments that the Biden Administration developed to ensure advanced AI systems are safe, secure and trustworthy.

The news came the same day NVIDIA's chief scientist, Bill Dally, testified before a U.S. Senate subcommittee seeking input on potential legislation covering generative AI . Separately, NVIDIA founder and CEO Jensen Huang will join other industry leaders in a closed-door meeting on AI Wednesday with the full Senate.

Seven companies including Adobe, IBM, Palantir and Salesforce joined NVIDIA in supporting the eight agreements the Biden-Harris administration released in July with support from Amazon, Anthropic, Google, Inflection, Meta, Microsoft and OpenAI.

The commitments are designed to advance common standards and best practices to ensure the safety of generative AI systems until regulations are in place, the White House said. They include:

Testing the safety and capabilities of AI products before they're deployed,

Safeguarding AI models against cyber and insider threats, and

Using AI to help meet society's greatest challenges, from cancer to climate change.

In his testimony, Dally told the Senate subcommittee that government and industry should balance encouraging innovation in AI with ensuring models are deployed responsibly.

The subcommittee's hearing, "Oversight of AI: Rules for Artificial Intelligence," is among actions from policymakers around the world trying to identify and address potential risks of generative AI.

Earlier this year, the subcommittee heard testimonies from leaders of Anthropic, IBM and OpenAI, as well as academics such as Yoshua Bengio, a University of Montreal professor considered one of the godfathers of AI.

Dally, who leads a global team of more than 300 at NVIDIA Research, shared the witness table on Tuesday with Brad Smith, Microsoft's president and vice chair. Dally's testimony briefly encapsulated NVIDIA's unique role in the evolution of AI over the last two decades.

He described how NVIDIA invented the GPU in 1999 as a graphics processing unit, then fit it for a broader role in parallel processing in 2006 with the CUDA programming software. Over time, developers across diverse scientific and technical computing fields found this new form of accelerated computing could significantly advance their work.

Along the way, researchers discovered GPUs also were a natural fit for AI's neural networks, because they require massive parallel processing.

In 2012, the AlexNet model, trained on two NVIDIA GPUs, demonstrated human-like capabilities in image recognition. That result helped spark a decade of rapid advances using GPUs, leading to ChatGPT and other generative AI models used by hundreds of millions worldwide.

Today, accelerated computing and generative AI are showing the potential to transform industries, address global challenges and profoundly benefit society, said Dally, who chaired Stanford University's

computer science department before joining NVIDIA.

In written testimony, Dally provided examples of how AI is empowering professionals to do their jobs better than they might have imagined in fields as diverse as business, healthcare and climate science.

Like any technology, AI products and services have risks and are subject to existing laws and regulations that aim to mitigate those risks.

Industry also has a role to play in deploying AI responsibly. Developers set limits for AI models when they train them and define their outputs.

Dally noted that NVIDIA released in April NeMo Guardrails , open-source software developers can use to guide generative AI applications in producing accurate, appropriate and secure text responses. He said that NVIDIA also maintains internal risk-management guidelines for AI models.

Making sure that new and exceptionally large AI models are accurate and safe is a natural role for regulators, Dally suggested.

He said that these “frontier” models are being developed at a gigantic scale. They exceed the capabilities of ChatGPT and other existing models that have already been well-explored by developers and users.

Dally urged the subcommittee to balance thoughtful regulation with the need to encourage innovation in an AI developer community that includes thousands of startups, researchers and enterprises worldwide. AI tools should be widely available to ensure a level playing field, he said.

During questioning, Senator Amy Klobuchar asked Dally why NVIDIA announced in March it’s working with Getty Images.

“At NVIDIA, we believe in respecting people’s intellectual property rights,” Dally replied. “We partnered with Getty to train large language models with a service called Picasso , so people who provided the original content got remunerated.”

In closing, Dally reaffirmed NVIDIA’s dedication to innovating generative AI and accelerated computing in ways that serve the best interests of all.

Original URL: https://blogs.nvidia.com/blog/2023/09/12/ai-safety-washington/