

Congress of the United States

Washington, DC 20515

October 13, 2026

The Honorable Walter “Jay” Clayton III
Director
Office of the Director of National Intelligence
1500 Tysons McLean Drive
McLean, VA 22102

The Honorable Michael Kratsios
Director
Office of Science and Technology Policy
1650 Pennsylvania Avenue
Washington, DC 20502

Dear Director Clayton and Director Kratsios:

We write to urge a coordinated federal response to the growing cybersecurity threats posed by artificial intelligence.

Recent AI security incidents should serve as a warning. Advanced AI agents have escaped testing environments, bypassed safeguards, and gained unauthorized access to real-world systems owned by government agencies and private organizations. AI companies and security researchers are now examining tens of thousands of incidents involving models from OpenAI and Anthropic escaping sandboxes, bypassing safeguards, evading monitoring, and taking other unexpected actions.

These incidents come as AI systems are becoming dramatically more capable at discovering software vulnerabilities, writing and executing code, using digital tools, and operating autonomously.”

We therefore urge the federal government to bring its full capacities to bear on this challenge, building a coordinated national response around standards, testing, auditing, and defensive AI.

First, strengthen federal AI safety and security standards. The National Institute of Standards and Technology (NIST), including its Center for Advancing Innovation and Standards for Super Intelligence (CAISSI), should accelerate the development of technically grounded and measurable standards and risk-management frameworks for advanced AI safety and security. These standards should address the risks posed by advanced AI capabilities, including the potential misuse by malicious actors to conduct sophisticated cyberattacks, automate attacks at scale, and deploy autonomous agents that threaten external systems and critical infrastructure.

The standards should also address the security of the environments in which advanced AI systems are developed, trained, tested, and used internally, including protections against agents escaping testing environments, gaining unauthorized access to external systems, or taking actions beyond their intended permissions. NIST should also work with industry to establish best practices with regard to fundamental aspects of AI development, including standards related to secure testing environments and third-party data sharing and testing. Standards and risk-management frameworks should evolve as capabilities evolve and draw on expertise from across government, industry, academia, and the technical community.

The Administration should also make sure CAISSI has additional resources and the technical expertise necessary to keep pace with rapidly advancing systems, including personnel with deep backgrounds in AI, cybersecurity, model evaluation, and national security. You should look to the private sector to contribute resources and expertise to support CAISSI's efforts in this critical time.

Second, expand independent testing capabilities. The federal government should build the capacity to independently test advanced AI models for serious national-security risks, including sophisticated cyber capabilities, biological and chemical threats, risks to critical infrastructure, the ability to evade safeguards, and increasingly autonomous behavior. CAISSI, the Department of Energy (DOE) National Laboratories, and other appropriate federal facilities can provide secure environments for these evaluations. The federal government capacity should be complemented by universities, independent research organizations, and qualified third-party evaluators including private sector entities, creating a broader testing ecosystem capable of keeping pace with increasingly powerful models.

A mature standards and testing environment can provide the foundation for a broader independent auditing ecosystem. Clear benchmarks and established testing methods will give auditors a common basis for assessing whether testing was conducted rigorously, results were accurately reported, vulnerabilities were addressed, and safeguards are working as intended. As standards and testing mature, independent auditing can provide an additional layer of verification and confidence in the overall system.

Third, move aggressively on defensive AI. The same advances that are making AI-enabled cyberattacks more capable can also strengthen America's cyber defenses. The Defense Advanced Research Projects Agency's (DARPA) AI Cyber Challenge has demonstrated the potential for AI to help identify and patch software vulnerabilities. The federal government should build on that progress by working with industry, universities, state and local government, and critical-infrastructure operators to develop and validate defensive AI tools, including open-source tools where possible. These tools can help detect AI-enabled threats, strengthen and patch vulnerable systems, and help defenders respond more quickly.

NIST, DARPA, DOE and the National Laboratories, the National Science Foundation (NSF), National Security Agency (NSA), and other appropriate agencies can help establish standards for these defensive tools, provide secure testing and validation environments, and expand access to the computing resources needed to develop and use them. Proven defensive AI capabilities should ultimately be available to critical-infrastructure operators, manufacturers, small businesses, state and local governments, and others that increasingly face AI-enabled cyber threats. The federal government should work with leading AI companies to make these capabilities available at scale, not just to a small number of agencies, companies, or governments that already have substantial cybersecurity resources. The goal should be to extend advanced defensive capabilities to these organizations that otherwise lack the expertise, computing resources, or access to protect themselves.

The recent incidents make clear that the security infrastructure surrounding advanced AI must advance as quickly as the technology itself. The United States has an opportunity to lead both sides of this challenge: building the standards, testing, and auditing infrastructure needed for increasingly powerful AI, while harnessing those same capabilities to make American networks harder to attack.

The Task Force for Super Intelligence and OSTP are uniquely positioned to bring these efforts together. As Congress continues to develop legislative solutions, we urge you to make them a central part of the Administration's AI and cybersecurity strategy now.

We look forward to working with you.

Sincerely,



Maria Cantwell
United States Senator

Ranking Member
Committee on Commerce, Science and Transportation



M. Michael Rounds
United States Senator

Chair
Subcommittee on Cybersecurity, Senate Committee on Armed Services

Co-Chair
Senate Artificial Intelligence Caucus