




EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

NSTM-4

April 23, 2026

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: MICHAEL J. KRATSIOS 
ASSISTANT TO THE PRESIDENT FOR SCIENCE AND TECHNOLOGY
DIRECTOR, OFFICE OF SCIENCE AND TECHNOLOGY POLICY

SUBJECT: Adversarial Distillation of American AI Models

The United States leads the world in artificial intelligence (AI) technologies. That lead reflects decades of foundational research, bold entrepreneurial risk-taking, and hundreds of billions of dollars in annual private investment. American AI leadership drives economic growth, strengthens national security, and advances the frontiers of science, medicine, and human knowledge. The breakthroughs emerging from American industry raise living standards, expand opportunity, and improve lives around the world.

However, the United States government has information indicating that foreign entities, principally based in China, are engaged in deliberate, industrial-scale campaigns to distill U.S. frontier AI systems. Leveraging tens of thousands of proxy accounts to evade detection and using jailbreaking techniques to expose proprietary information, these coordinated campaigns systematically extract capabilities from American AI models, exploiting American expertise and innovation.

Models developed from surreptitious, unauthorized distillation campaigns like this do not replicate the full performance of the original. They do, however, enable foreign actors to release products that appear to perform comparably on select benchmarks at a fraction of the cost. These distillation campaigns also allow those actors to deliberately strip security protocols from the resulting models and undo mechanisms that ensure those AI models are ideologically neutral and truth-seeking.

The United States is committed to the free and fair development of AI technologies across a competitive ecosystem, from leading frontier models to highly-tuned applied systems, and from open-source frameworks to open-weight models. AI distillation, when legitimately used to produce smaller, lighter-weight models from more advanced systems, is a vital part of that ecosystem. Industrial distillation activities that aim to systematically undermine American research and development and access proprietary information, however, are unacceptable.

To address this threat, the Trump Administration will:

1. Share information with U.S. AI companies concerning attempts by foreign actors to conduct unauthorized, industrial-scale distillation, including the tactics employed and actors involved.
2. Enable the private sector to better coordinate against such attacks.
3. Work together with private industry to develop best practices to identify, mitigate, and remediate industrial-scale distillation activities and build strong defenses against such activities.
4. Explore a range of measures to hold foreign actors accountable for industrial-scale distillation campaigns.

There is nothing innovative about systematically extracting and copying the innovations of American industry, and there is nothing open about supposedly open models that are derived from acts of malicious exploitation.

As methods to detect and mitigate industrial-scale distillation grow more sophisticated, foreign entities who build their AI capabilities on such fragile foundations should have little confidence in the integrity and reliability of the models they produce.

Consistent with America's AI Action Plan, the United States will continue to foster a vibrant open-source ecosystem built on firm foundations, support American industry in making frontier AI broadly accessible to users worldwide, and safeguard the free and fair market competition that enables the broad and beneficial diffusion of these technologies.