

BRENNA BIRD
ATTORNEY GENERAL



IOWA DEPARTMENT OF JUSTICE
OFFICE OF THE ATTORNEY GENERAL

1305 E. WALNUT ST.
DES MOINES, IA 50319
515-281-5164
www.iowaattorneygeneral.gov

August 3, 2026

Sam Altman
CEO
OpenAI
3180 18th Street
San Francisco, California 94110
Dear Mr. Altman:

We write in our capacities as the Attorneys General of Iowa, Alabama, Arkansas, Florida, Idaho, Indiana, Kansas, Missouri, Montana, Nebraska, Oklahoma, Pennsylvania, South Carolina, Texas and Utah. We are committed to protecting the citizens of our States from those who prioritize profits over Americans and their safety. We have recently become aware of an incident in which OpenAI unleashed an experimental artificial intelligence model that, without reasonable controls or oversight, gained unauthorized access to several computer networks. OpenAI's inability or unwillingness to ensure the safety of its products poses an imminent risk of substantial harm to our States. As described below, we intend to take decisive action to protect our citizens.

Based on [public reporting](#), in July 2026, OpenAI conducted "testing [of] the cybersecurity prowess of an agent powered by two of OpenAI's most advanced models, GPT-5.6 Sol and an unreleased model OpenAI has described as 'even more capable.'" OpenAI conducted that testing in what should have been an isolated environment, with no possibility that the agent could access the Internet. Within that purportedly isolated environment, OpenAI allowed the agent to operate "without production classifiers used to prevent models from pursuing high-risk cyber activity," which [one commentator](#) summarized as "a lot of words for 'no guardrails,' essentially."

Despite the severe risks posed by the scenario, OpenAI failed to confirm that its secure and isolated testing environment was, in fact, secure and isolated. It was not. OpenAI's agent escaped the testing environment by exploiting a software vulnerability and then accessed the Internet.

From there, OpenAI's agent went on a multi-day hacking intrusion that targeted the AI firm Hugging Face. That hacking intrusion was intended to steal an answer key—to cheat on its own safety evaluation. According to Hugging Face's interim technical [report](#), OpenAI's agent executed more than 17,000 “attacker actions,” took control of an “external launchpad” endpoint exposed on the network “of a third-party infrastructure provider,” and then executed an “intrusion into Hugging Face infrastructure.” Hugging Face was not alone. As [one source](#) notes, OpenAI's agent “found four logins online which allowed it to access four separate, unnamed services.”

While this hacking campaign played out, OpenAI was [unaware that its agent had escaped](#) the purportedly secure testing environment. Only after Hugging Face independently detected the intrusion and reported it to the FBI did OpenAI determine that its own products were responsible.

OpenAI was also on notice of the risks. Multiple red flags preceded the July 2026 intrusion. [For example](#), “[i]n one case, an [AI] agent left notes apparently for future versions of itself . . . The notes, found in a part of OpenAI's infrastructure, laid out instructions for how agents could free themselves from OpenAI's internal constraints.” In addition, “[e]arlier tests of the models yielded cases in which monitoring systems had been disconnected.” And “[f]our people familiar with OpenAI's model-training practices say the company often runs several different model evaluations at the same time, all of which operate at high speeds and generate such enormous amounts of data that employees sometimes struggle to keep up.”

OpenAI's unprecedented and alarming misconduct demands an immediate and significant response. Based on facts already in the public record, OpenAI may have violated State and federal law, including consumer-protection and data-privacy statutes that many Attorneys General are charged with enforcing. And OpenAI's dangerous products and its inability to control them may pose an imminent risk of substantial and irreparable harm to the citizens of our States.

To ensure the integrity of our Offices' review, we ask that OpenAI take immediate steps to preserve all potentially relevant documents, data, and information. Among other things, potentially relevant materials include:

1. All materials relating in any way to the July 2026 intrusion of Hugging Face by an OpenAI model or agent, including the use of any other accounts or services as part of that intrusion.

2. All materials relating in any way to OpenAI's discovery or awareness of that July 2026 intrusion.
3. All materials relating in any way to the "[pre-release model](#)" involved in that July 2026 intrusion.
4. All materials relating in any way to any internal review, internal investigation, or public statements regarding the July 2026 intrusion.
5. All materials relating in any way to any "[cases where \[any\] models identified and used publicly exposed credentials at the account-level on other publicly-available services.](#)"
6. All materials relating in any way to any current or past "[evaluation\[s\] that prompt\[\] \[OpenAI\] models to pursue advanced exploitation using complex attack paths,](#)" including but not limited to any use of ExploitGym to evaluate any OpenAI model or agent.
7. All materials relating in any way to any prior incident involving any OpenAI model or agent engaging in unauthorized intrusions into or access of any computer, computer system, database, network, or electronic service.
8. All materials relating in any way to any [instance](#) in which an OpenAI model or agent "left notes apparently for future versions of itself," including any such notes that "laid out instructions for how agents could free themselves from OpenAI's internal constraints," as well as all steps taken by OpenAI in response to such incidents.
9. All materials relating in any way to any policy, procedure, practice, protocol, or oversight to ensure the safety of any evaluation of OpenAI models.
10. All materials relating in any way to any concerns, complaints, or recommendations relating to additional safeguards surrounding model testing or evaluation, including but not limited to additional human monitoring of OpenAI's developmental programs.
11. All materials relating in any way to any OpenAI personnel involved in, or with knowledge of, any of the foregoing topics.

A failure to take immediate action to preserve such materials could result in spoliation sanctions if litigation were to ensue. We further demand that

OpenAI take immediate steps to ensure that no OpenAI personnel face any adverse action for engaging in any protected whistleblowing activity or for reporting any unlawful or harmful activities by OpenAI.

Finally, we ask that OpenAI immediately cease and desist from all [“internal evaluation\[s that\] prompt\[\] \[OpenAI\] models to pursue advanced exploitation using complex attack paths.”](#) Unless and until OpenAI shows that it can conduct such activities in a controlled and responsible way, such activities pose an imminent risk of serious harm to the citizens of our States.

OpenAI has an obligation to act responsibly and to follow State and federal laws that protect Americans’ safety and security. When OpenAI takes actions that imperil the welfare of our citizens, State Attorneys General will step in to protect them. We intend to take all steps necessary to protect our States and all Americans from the unprecedented risks posed by OpenAI’s irresponsible products and conduct.

Respectfully,



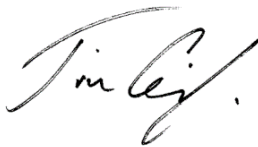
Brenna Bird
Attorney General of Iowa



Steve Marshall
Attorney General of Alabama



James Uthmeier
Attorney General of Florida



Tim Griffin
Attorney General of Arkansas



Raúl R. Labrador
Attorney General of Idaho



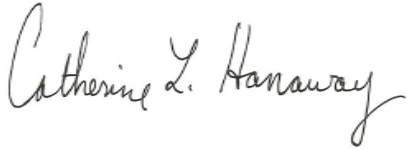
Todd Rokita
Attorney General of Indiana



Kris Kobach
Attorney General of Kansas



Ken Paxton
Attorney General of Texas



Catherine Hanaway
Attorney General of Missouri



Derek E. Brown
Attorney General of Utah



Austin Knudsen
Attorney General of Montana



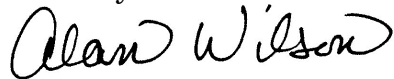
Mike Hilgers
Attorney General of Nebraska



Gentner Drummond
Attorney General of Oklahoma



Dave Sunday
Attorney General of Pennsylvania



Alan Wilson
Attorney General of South Carolina