

OpenAI asks governments to require monitoring, shutdown controls and incident reports

Source copy kept by wheresthe.ai for the story "OpenAI asks governments to require monitoring, shutdown controls and incident reports" (<https://wheresthe.ai/d/openai-cyber-defenders-blueprint-2026-10/>).

Source: <https://openai.com/index/empowering-cyber-defenders/>

Text saved 22 Oct 2026 (the page text through a reader, because the site refused a direct fetch).

SHA-256 of the saved text: 25521a2f456525c626fe099e6b979a8e29772fc3bd28f64cb0ea1e95d934370d

AI is changing the speed, scale, and economics of cybersecurity. It can help defenders find and fix vulnerabilities, detect intrusions, and respond to attacks—but those same capabilities can enable malicious actors and create new risks when AI systems act beyond their intended boundaries. We are in a defender's window: a limited opportunity to put advanced AI into defenders' hands and strengthen protections before advanced cyber offensive capabilities become widely available.

Today, we're publishing a blueprint for empowering cyber defenders, building on our earlier cybersecurity action plan and frontier safety blueprint([opens in a new window](#)). Our experience with misalignment and security incidents involving our own models underscores the need for stronger safeguards and accountability across the industry. Drawing on those lessons, the blueprint puts forward concrete proposals for governments and industry in three areas:

* **Strengthening safety and security for advanced AI:** Establish clearer responsibilities for securing frontier models and agents, detecting dangerous behavior, and maintaining reliable shutdown controls for operations under each organization's control.

* **Modernizing cybersecurity for the intelligence age:** Equip defenders with advanced tools, training, and technical support; address longstanding cybersecurity weaknesses; and invest in securing essential services and replacing outdated systems.

* **Building resilient and adaptive public institutions:** Strengthen government evaluations and independent technical assessments, establish consistent incident reporting requirements, and improve coordinated response and recovery.

AI labs have a responsibility to address the risks their systems create and help strengthen society's defenses. Governments must also have the expertise, resources, and authority to protect the public as capabilities advance. We encourage policymakers to act now—establishing safeguards, learning from experience, and strengthening requirements as the technology evolves.