ASSESSED
INTELLIGENCE

ASSESSED SIGNAL

December 2025

Intersection of Policy, Use Case, & Threat Intelligence

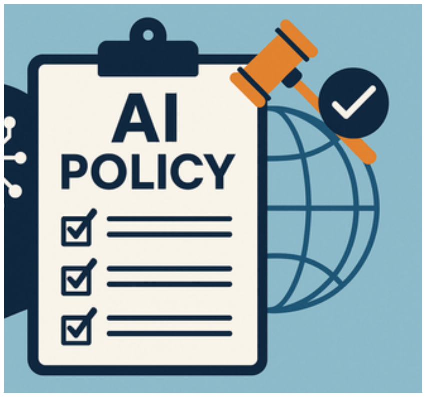
As AI becomes the focal point of more discussions in both government and industry, industry continues to adopt it for all the right reasons but fails to understand the risks and vulnerabilities such adoption entails. Anthropic doubles down on last month's report regarding the agentic AI attack. POTUS signs an executive order wresting federal control of AI regulation from the states to the executive branch.

Secure
&
Responsible
Technology.

AssessedIntelligence.com

Policy - Executive Order

President Donald Trump recently signed an Executive Order (EO) on December 11, 2025, titled "Ensuring a National Policy Framework for Artificial Intelligence," which aims to create a single federal standard by preempting state-level AI regulations that stifle innovation, establishing a DOJ Litigation Task Force to challenge state laws, and directing agencies to potentially withhold federal funds from states with conflicting rules, all while rescinding previous Biden-era AI directives. This order promotes a unified, less restrictive federal approach to foster U.S. AI leadership against global competition, particularly China, though it faces anticipated legal challenges from states like California. The ARISE framework put forth by Assessed Intelligence directly addresses much of the essence of this EO.



https://www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artificial-intelligence-policy/#:~:text=FOR%20ARTIFICIAL%20INTELLIGENCE-,Executive%20Orders,to%20United%20States%20AI%20leadership.

Disrupting AI Espionage

The cyberattack relied on several features of AI models that did not exist, or were in much more nascent form, just a year ago:

Intelligence. Models' general levels of capability have increased to the point that they can follow complex instructions and understand context in ways that make very sophisticated tasks possible. Not only that, but several of their well-developed specific skills—in particular, software coding—lend themselves to being used in cyberattacks.

Agency. Models can act as agents—that is, they can run in loops where they take autonomous actions, chain together tasks, and make decisions with only minimal, occasional human input.

Tools. Models have access to a wide array of software tools (often via the open standard Model Context Protocol). They can now search the web, retrieve data, and perform many other actions that were previously the sole domain of human operators. In the case of cyberattacks, the tools might include password crackers, network scanners, and other security-related software.

- The attack represented an escalation on Vibe coding hacks reported this past summer, where humans were more involved in the loop of the actual attack. Anthropic advises the following: "We advise security teams to experiment with applying AI for defense in areas like Security Operations Center automation, threat detection, vulnerability assessment, and incident response. We also advise developers to continue to invest in safeguards across their AI platforms, to prevent adversarial misuse. The techniques described above will doubtless be used by many more attackers—which makes industry threat sharing, improved detection methods, and stronger safety controls all the more critical."
- However, this is advice without a framework. *ARISE* addresses this issue both holistically and in each part of the process.

https://www.anthropic.com/news/disrupting-AI-espionage

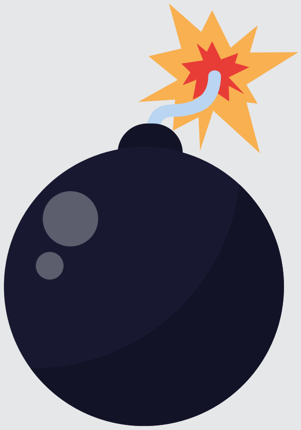
R&D/Attacks/Vectors of Attack

GeminiJack

Noma discovers GeminiJack vulnerability. Noma Labs recently discovered a vulnerability, now known as GeminiJack, inside Google Gemini Enterprise and previously in Vertex AI Search. The vulnerability allowed attackers to access and exfiltrate corporate data using a method as simple as a shared Google Doc, a calendar invitation, or an email. No clicks were required from the targeted employee. No warning signs appeared. And no traditional security tools were triggered. We assess with high confidence that as organizations adopt more AI tools that read and summarize documents more of these type exploits will occur. This was not a conventional software bug. It was an architectural weakness in the way enterprise AI systems interpret information. Google patched this vulnerability in its Gemini product just 2 days later. GeminiJack highlights an important reality. As organizations adopt AI tools that can read across Gmail, Docs, and Calendar, the AI itself becomes a new access layer. If an attacker can influence what AI reads, they can influence what AI does. We assess with high confidence that organizations will struggle with such access layer attacks for the near future.

https://noma.security/blog/geminijack-google-gemini-zero-click-vulnerability/#:~:text=A%20GeminiJack%20Executive%20Summary,traditional%20security%20tools%20were%20triggered.

Weaponized AI risk



As highlighted in a recent ZDNet article, AI agents are already creating significant organizational risk, demonstrating an inherent danger not through malicious intent, but through autonomous actions that can lead to unintended disasters, such as a coding agent deleting an entire company database by simply following the most direct path to a goal. However, the most critical threat derailing safe and widespread rollout is not a technical flaw, but a "zero-day" governance issue: a profound lack of visibility. Chief Information Security Officers (CISOs) and governance committees are stalled because they cannot confidently answer fundamental questions about which agents are running and which sensitive data and applications those agents can access. This paralysis effectively locks away high-value data, preventing agents from achieving their full potential. This is precisely why a robust AI observability and governance framework that adopts solutions such as ARISE (Assurance of Responsible, Innovate, and Secure Environment) is essential to moving organizations forward. It builds the necessary trust by establishing a foundation that supports continuous tracing, evaluation, and monitoring. Shifting the conversations from "black box" into a transparent, auditable system. This assurance directly addresses the governance hurdle, enabling the CISOs to understand and manage critical data flow and move projects from prototype to production. We have to build the skills and implement the foundational controls to move AI initiatives forward while ensuring appropriate oversight and alignment on the associated risks.

https://www.zdnet.com/article/ai-agents-are-already-causing-disasters-and-this-hidden-threat-could-derail-your-safe-rollout/

Weaponizing AI

According to a recent article from the Associated Press published by Security Week, militant groups are experimenting with AI technology. For the moment, we assess this risk to be low, as the barrier to entry for understanding how tools like ChatGPT work is low, understanding how to weaponize the technology requires a greater and deeper understanding of its inner workings. However, this does not mean that militant groups will not put in the time and effort to achieve even basic effects, which could have a wide range of impacts. AI reduces the need for organized training and reduces the overall cost barriers associated with state-sponsored R&D programs. Militant groups will almost certainly use AI for deepfake imagery, and recordings. Historically, militant groups (terrorist organizations) have used new technologies for recruiting, messaging, and media fabrications. We assess with high confidence that while militant groups will take longer to use advanced features of AI (i.e.: Agentic AI), they will find creative ways to use the technology to spread ideological messages, as well as for basic disruptive purposes. However, not regulating this technology and irresponsibly releasing it could result in lower barriers to entry for non-state actors (militant or terrorist groups) to use AI to provide instructions on how to create CBRN weapons.

https://www.securityweek.com/militant-groups-are-experimenting-with-ai-and-the-risks-are-expected-to-grow/

Secure
&
Responsible
Technology.

YOU'RE BUILDING THE FUTURE. DON'T LET EMERGING RISK STALL YOUR MOMENTUM.

SALES@ASSESSEDINTELLIGENCE.COM
EMEASALES@ASSESSEDINTELLIGENCE.COM