Lowenstein Sandler

Client Alert

Data Privacy, Security, Safety & Risk Management

November 18, 2025

Anthropic Reports First Known AI-Orchestrated Cyber Espionage Campaign: Raising Stakes for Data Security

By Amy S. Mushahwar, Cairon S. Overton, and Tricia Y. Wagner CIPP/US, CISSP, CISA

On November 13, AI company Anthropic reported that its Threat Intelligence team had disrupted a state-sponsored Chinese threat actor conducting what is believed to be the first largely autonomous AI-orchestrated cyber espionage campaign. The threat actor used Claude Code with autonomous agentic orchestration to execute the majority of the intrusion life cycle—reconnaissance, exploitation, credential harvesting, lateral movement, and data exfiltration—across multiple global sectors. Claude Code is a developer-focused variant of the Claude large language model created by Anthropic, designed to function as an advanced, agentic coding assistant that automates complex software development.

The key technical aspects of the campaign reveal that the AI agent autonomously carried out approximately 80 percent to 90 percent of all operational tasks. Human operators were responsible for selecting targets and giving strategic approvals, while the AI handled the majority of the tactical actions. Rather than leveraging zero-day vulnerabilities, the attack primarily utilized widely available open-source commodity tools, which were scaled through AI orchestration. Additionally, the AI was able to generate operational documentation, network mappings, and internal notes as well as manage task handoffs between phases. While largely AI-generated, some of the AI's outputs required human validation due to occasional inaccuracies.

Legal and regulatory implications

Data and privacy: This reported use of autonomous intrusion techniques significantly compresses the time frame for detection and notification, making it more challenging to assess the scope and impact of a breach. Agentic AI systems review payloads, identify high-value information, and categorize and sort data, thereby reducing the resources and time required to understand and exploit compromised data.

Supply chain: Organizations that utilize or expose AI models through their vendors may face new legal obligations concerning transparency, prevention of misuse, and auditability. These factors increase the importance of robust contract management and due diligence for third-party AI services.

State-sponsored activity: Although this reported incident is believed to be a state-sponsored group, the deployment of agentic AI lowers the resource threshold for executing global, simultaneous attacks by more-pedestrian threat actors. This technology enables smaller criminal groups to conduct large-scale campaigns, and successful attribution could invoke national security concerns, sanctions, or critical infrastructure protections.

AI governance: Organizations must update their governance programs to address the risks associated with model misuse and adversarial AI. This includes enhancing monitoring and response strategies specific to AI-driven threats.

Business risk implications

The adoption of these advanced intrusion techniques significantly lowers the barriers for smaller and less-resourced threat groups to conduct sophisticated espionage operations. This democratization of capability means that organizations may face a broader range of adversaries, each capable of executing complex attacks that were previously limited to highly skilled or well-funded actors.

As a result, organizations face increased exposure of critical assets, including intellectual property, trade secrets, and research and development initiatives. The heightened risk to these valuable resources can have far-reaching consequences for business competitiveness and long-term innovation.

Incident response procedures become more challenging and forensic investigation timelines are extended due to the complexity introduced by these advanced threats. The ability to quickly identify, contain, and remediate incidents is hindered, potentially leading to greater damage and prolonged recovery times.

Furthermore, these evolving risks may influence cyber insurance underwriting practices, as insurers reassess the likelihood and impact of successful attacks. This, in turn, can affect how risk is reported and managed at the board level, necessitating more comprehensive strategies and oversight to address emerging threats.

The campaign targeted a wide array of high-value entities, including major technology corporations, government agencies, financial institutions, and chemical manufacturing companies across multiple nations. The scale and diversity of chosen targets proves that no sector is safe, and any organization with valuable intellectual property or sensitive data may be targeted for autonomous attacks.

Business risk implications

Short-term actions (first 30 days)

1. Conduct an AI attack tabletop exercise: Simulate an AI-driven attack using this incident as a scenario. This exercise will help teams identify gaps in current preparedness and response strategies related to advanced, autonomous threats.
2. Review breach response plans for high-automation events: Evaluate and update existing breach response plans to ensure they address incidents involving highly automated and agentic AI techniques. Focus on detection, containment, and notification procedures.
3. Assess third-party AI risk and update vendor obligations: Examine potential risks introduced through third-party AI services and revise vendor contracts to reflect updated obligations regarding transparency, prevention of misuse, and auditability.
4. Strengthen AI governance and misuse detection: Enhance oversight structures and implement or improve mechanisms for detecting the misuse of AI systems within the organization.
5. Prepare communications for AI-driven breach notification scenarios: Develop and refine communication templates and protocols to ensure prompt and accurate notifications in the event of an AI-enabled breach.

Mid- and long-term actions (3-12 months)

1. Integrate AI-enhanced detection tools: Adopt and deploy advanced detection solutions that leverage AI to identify and mitigate sophisticated threats more effectively.
2. Upgrade red team exercises to assume AI-enabled adversaries: Modify red team scenarios to include adversaries utilizing agentic AI, ensuring that organizational defenses are tested against the latest threat models.
3. Reevaluate cyber insurance coverage for AI-orchestrated attacks: Review and, if necessary, update cyber insurance policies to account for risks and liabilities related to AI-driven attack campaigns.
4. Update board briefings and investor communications: Ensure that regular reports to the board and communications with investors include insights on emerging AI-related risks and response measures.
5. Increase participation in threat-sharing networks: Engage more actively in industry and cross-sector threat intelligence sharing to stay informed about evolving AI threat landscapes and best practices.

Key takeaways

- This is the first documented agentic AI-orchestrated espionage operation. In previous attacks, AI provided advice on how to implement an attack and humans implemented it. Here, humans advised and AI implemented the attack.
- The reported campaign's speed, scale, and autonomy challenge traditional legal and technical response frameworks.
- Organizations must enhance AI governance, vendor oversight, and detection capabilities.

- Lowenstein Sandler's Data360 approach continues to deliver dual-threat legal and technical insights to help clients to embed privacy, security, and safety controls directly into business operations and product development lifecycles.

For guidance or additional information on how to test and strengthen your organization's data security infrastructure, please contact the authors of this article.

Contacts

Please contact the listed attorneys for further information on the matters discussed herein.

AMY S. MUSHAHWAR
Partner
Chair, Data Privacy, Security, Safety & Risk Management
T: 202.753.3825
amushahwar@lowenstein.com

CAIRON S. OVERTON
Counsel
T: 202.753.3747
coverton@lowenstein.com

TRICIA Y. WAGNER CIPP/US, CISSP, CISA
Counsel
T: 202.753.3658
twagner@lowenstein.com