ASSESSED INTELLIGENCE

ASSESSED SIGNAL

April 2026

At the Intersection of Policy, Use & Threat Intelligence

The regulatory collision between California's AI procurement Executive Order and the Trump administration's federal preemption framework has created a dual compliance burden, with Colorado's AI Act and the EU AI Act both reaching enforceability within the same operational quarter. The U.S. Intelligence Community's 2026 Annual Threat Assessment designates AI as a defining national security priority for the first time at this prominence, documenting adversarial use in active conflict operations and state-sponsored data extortion campaigns. At RSAC 2026, CrowdStrike reported the average adversary breakout time has collapsed to 29 minutes, AI-enabled attacks increased 89 percent year-over-year, and one Fortune 500 organization discovered more than 600 AI agents running in its environment without any security ownership.

Secure & Responsible Technology.

AssessedIntelligence.com

The Federal-State Collision Is Now Operational

The regulatory environment governing AI has entered a new phase. It is no longer a debate about whether to regulate; it is a collision between two competing governance architectures, neither of which has a clear mandate to prevail.

California is advancing on multiple fronts simultaneously. Executive Order N-5-26, signed by the Governor on March 30, directs state agencies to overhaul procurement standards and requires companies seeking state government contracts to demonstrate how AI systems should address risks such as model bias, civil rights violations, and illegal or exploitative content.[1] This order also establishes California's authority to evaluate federal supply-chain risk designations independently, operating separately from federal procurement frameworks. California acts, and companies will need to adapt to remain in the world's fourth-largest economy.[2] Federal policy continues to emphasize an innovation-first approach while state laws begin taking effect in 2026, and the federal government signals increasing willingness to contest or preempt state approaches through the DOJ AI Litigation Task Force.[3] The legal challenges are anticipated, what remains uncertain is the compliance burden organizations now carry while that litigation proceeds. According to EY's recent global survey, the majority of C-suite leaders identify non-compliance with AI regulations as the most common AI risk they face.[3]

Organizations deploying AI in consequential decisions, including financial services, healthcare, education, employment, housing, and essential government services, must prepare for new high-risk AI requirements entering effect throughout 2026.[4] Colorado's AI Act, postponed from February 1 to June 30, 2026, establishes requirements for developers and deployers of high-risk AI systems, including annual impact assessments, risk management obligations, disclosure requirements, and the duty of reasonable care to protect consumers from algorithmic discrimination.[5] The EU AI Act is fully applicable by August 2, 2026. Two major regulatory deadlines now sit within the same operational quarter.

This is not a moment to wait for legal clarity. Organizations that have not established a governance baseline will manage compliance retroactively, which is the most expensive and least defensible posture. The *ARISE Framework™* addresses this directly through its GOVERN and MANAGE domains, which establish the cross-functional oversight structures and documented system boundaries that regulators on both sides of this conflict require.

Sources:

1. California Governor's Office. Executive Order N-5-26, March 30, 2026. https://codewithvamp.medium.com/owasp-top-10-for-agentic-applications-2026-the-ultimate-guide-to-securing-ai-agents-2459c39e37a9
2. Axios. California cements its role as the national testing ground for AI rules. April 3, 2026. https://codewithvamp.medium.com/owasp-top-10-for-agentic-applications-2026-the-ultimate-guide-to-securing-ai-agents-2459c39e37a9
3. Software Improvement Group. AI legislation in the US: A 2026 overview. January 28, 2026. https://codewithvamp.medium.com/owasp-top-10-for-agentic-applications-2026-the-ultimate-guide-to-securing-ai-agents-2459c39e37a9
4. Wilson Sonsini. 2026 Year in Preview: AI Regulatory Developments for Companies to Watch Out For. January 13, 2026. https://codewithvamp.medium.com/owasp-top-10-for-agentic-applications-2026-the-ultimate-guide-to-securing-ai-agents-2459c39e37a9
5. Brownstein Hyatt Farber Schreck. Colorado's Landmark AI Law Coming Online. March 4, 2026. https://codewithvamp.medium.com/owasp-top-10-for-agentic-applications-2026-the-ultimate-guide-to-securing-ai-agents-2459c39e37a9

AI as a Declared National Security Priority

The political framing around AI shifted again in March. The Office of the Director of National Intelligence released the 2026 Annual Threat Assessment on March 18, and the placement of AI within that document signals something material. The Intelligence Community describes AI as a "defining technology for the 21st century," notes it is being used in active combat operations, and identifies China as the most capable competitor to the United States in this domain.[1] The assessment treats AI differently than it treats conventional threats. Unlike enduring threats from China, Russia, Iran, and North Korea, AI is positioned less as a discrete actor or capability and more as a cross-cutting force that amplifies each of them.[1] That framing has operational implications for enterprise risk. It means the governance question is not contained within a technology department; it extends to supply chain integrity, data provenance, and organizational exposure to adversarial AI operations.

The Intelligence Community assessed that China aims to displace the United States as the global AI leader by 2030, and that AI adoption at scale poses serious risks, including its documented use in weapons design and in recent conflicts to influence targeting and streamline battlefield decision-making.[2] DNI Gabbard described a China-run data-extortion operation from August 2025 in which an AI tool was used to target international government, healthcare, public health, and emergency services sectors. That operation is not a future scenario. It is a documented past event with documented victims. The 2026 assessment's most consequential shift is its explicit recognition that threats now converge across domains, exploiting organizational boundaries and decision-making processes designed for an earlier era. Cyber operations enable space attacks. Transnational crime funds state weapons programs. Advances in AI itself create risk.[3] Organizations that continue to treat AI governance as a compliance function rather than a security function are building their posture on a premise the Intelligence Community has already rejected.

1. Defense One. US intelligence elevates AI as a top global threat in new report. March 2026. https://www.kiteworks.com/cybersecurity-risk-management/agentic-ai-attack-surface-enterprise-security-2026/
2. Office of the Director of National Intelligence. 2026 Annual Threat Assessment. March 18, 2026. https://www.kiteworks.com/cybersecurity-risk-management/agentic-ai-attack-surface-enterprise-security-2026/
3. The SCIF / National Security Institute. NSI Experts Weigh In: 2026 Annual Threat Assessment. March 2026. https://www.kiteworks.com/cybersecurity-risk-management/agentic-ai-attack-surface-enterprise-security-2026/

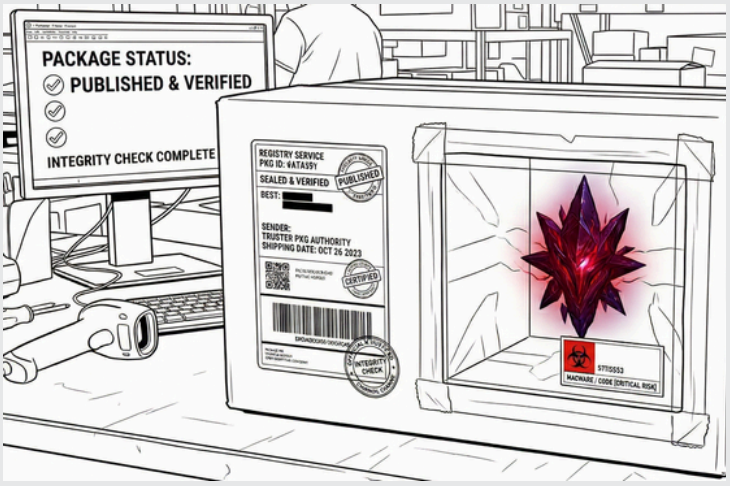
The LiteLLM Compromise: When the AI Stack Becomes the Attack Surface

On March 24, 2026, two versions of the Python library LiteLLM were published to the official PyPI repository containing malicious code. Versions 1.82.7 and 1.82.8 were not typosquatted packages and were not submitted through unofficial channels. They were published using legitimate maintainer credentials obtained through a prior compromise of Trivy, an open-source security scanner embedded in LiteLLM's own CI/CD pipeline. The packages were available for approximately three hours before PyPI quarantined them. LiteLLM processes roughly 3.4 million downloads per day and is present in an estimated 36 percent of cloud environments, according to Wiz research. That download volume, combined with the three-hour exposure window, created a blast radius that security teams are still scoping.[1,2,3]

The attack was attributed to TeamPCP, a threat group previously responsible for compromising Aqua Security's Trivy scanner and Checkmarx KICS. The Trivy compromise is the documented root cause of the LiteLLM incident: attackers obtained the LiteLLM maintainer's PyPI publishing token from the GitHub Actions runner environment during the Trivy build process, then used that token to publish malicious packages directly to PyPI, bypassing the project's standard GitHub release workflow entirely. Neither version 1.82.7 nor 1.82.8 has a corresponding git tag in the official repository. The two versions used distinct injection techniques. Version 1.82.7 embedded a base64-encoded payload inside the proxy server source file, executing on any import of the proxy module. Version 1.82.8 escalated this by adding a .pth file named litellm_init.pth to the Python site-packages directory; Python's site module processes .pth files automatically on every interpreter startup, meaning the payload executed on any python, pip, or pytest invocation in the affected environment, with no import statement required.[2,4,5]

The payload operated as a three-stage attack. The first stage conducted a comprehensive credential harvest targeting environment variables, SSH keys, cloud credentials from AWS, GCP, and Azure, Kubernetes configuration files and service account tokens, CI/CD secrets, Docker configurations, database credentials, and cryptocurrency wallets. The collected data was encrypted using AES-256-CBC with the key further protected by an embedded RSA public key, then exfiltrated to attacker-controlled infrastructure at models.litellm.cloud. The second stage attempted lateral movement across Kubernetes clusters by deploying privileged pods to every available node and extracting cluster secrets. The third stage installed a persistent systemd backdoor that polled external attacker infrastructure at checkmarx.zone every 50 minutes for additional second-stage payloads, deliberately abusing the Checkmarx brand name to evade DNS allowlists. BleepingComputer reported that the estimated number of data exfiltrations reached approximately 500,000, with many duplicates across affected environments.[4,5,6]

The organizational implications of this incident extend well beyond LiteLLM users. Downstream projects including DSPy, MLflow, OpenHands, CrewAI, and Arize Phoenix filed security pull requests to pin away from the affected versions on the same day. The attack was not discovered by any security tooling; it was identified by a researcher who noticed anomalous behavior while testing a Cursor MCP plugin that pulled LiteLLM in as a transitive dependency. That discovery path is the governance gap the incident exposes: organizations that cannot answer which AI libraries are running in their environments, which versions are installed, and which CI/CD jobs installed them during a specific time window, cannot determine their exposure. Standard hash verification provides no protection in this scenario because the malicious content was published using legitimate credentials; the hash matches what PyPI advertised, and the package passes all standard integrity checks.[1,4,7]



The *ARISE Framework™* addresses this directly through its IDENTIFY and PROTECT domains. Organizations must maintain a complete AI System Inventory that includes third-party libraries, their versions, their installation paths, and the credentials and data they can access in the environments where they run. The LiteLLM incident demonstrates that an AI library operating as an LLM gateway occupies a uniquely privileged position: it sits between the developer environment and every major model provider endpoint, with access to API keys, cloud tokens, and Kubernetes credentials as a functional requirement. We assess with high confidence that TeamPCP will continue targeting AI infrastructure components for exactly this reason; the credentials available in an AI development environment represent a high-value, concentrated target that is frequently under-governed relative to its access level. Organizations that have not yet inventoried their AI dependency graph and associated credential exposure must treat that gap as an unacceptable risk posture, not a deferred compliance task.

1. Wiz. LiteLLM TeamPCP Supply Chain Attack: Malicious PyPI Packages. March 2026. https://www.wiz.io/blog/threes-a-crowd-teampcp-trojanizes-litellm-in-continuation-of-campaign
2. Snyk. How a Poisoned Security Scanner Became the Key to Backdooring LiteLLM. March 2026. https://snyk.io/blog/poisoned-security-scanner-backdooring-litellm/
3. Datadog Security Labs. LiteLLM and Telnyx compromised on PyPI: Tracing the TeamPCP supply chain campaign. March 2026. https://securitylabs.datadoghq.com/articles/litellm-compromised-pypi-teampcp-supply-chain-campaign/
4. BleepingComputer. Popular LiteLLM PyPI package backdoored to steal credentials, auth tokens. March 2026. https://www.bleepingcomputer.com/news/security/popular-litellm-pypi-package-compromised-in-teampcp-supply-chain-attack/
5. Trend Micro. Your AI Gateway Was a Backdoor: Inside the LiteLLM Supply Chain Compromise. March 26, 2026. https://www.trendmicro.com/en_us/research/26/c/inside-litellm-supply-chain-compromise.html
6. Sonatype. Compromised LiteLLM PyPI Package Delivers Multi-Stage Credential Stealer. March 2026. https://www.sonatype.com/blog/compromised-litellm-pypi-package-delivers-multi-stage-credential-stealer
7. LiteLLM. Security Update: Suspected Supply Chain Incident. March 2026. https://docs.litellm.ai/blog/security-update-march-2026

The Agentic Attack Surface

From Theory to Documented Pattern

RSAC 2026 closed last week in San Francisco. Forty-four thousand practitioners attended. The loudest recurring concern was not ransomware or nation-state intrusion. The dominant theme was AI agent security: autonomous agents and AI tools spinning up inside organizations, accessing sensitive credentials, and operating entirely outside any security team's visibility.[1] Forrester's Q4 2025 AI Pulse Survey found that 50 percent of organizations are currently piloting agentic AI, while 24 percent have agents in production. The concrete illustration of this gap arrived early in the RSAC Innovation Sandbox: a Fortune 500 customer uncovered more than 600 AI agents it did not know existed. However, this was not a surprise; this is now the expectation.[2]

Microsoft's analysis presented at RSAC identified AI appearing across the entire attack lifecycle. In reconnaissance, AI accelerates infrastructure discovery and persona development, compressing the time between target selection and first contact. In resource development, AI generates forged documents, polished social engineering narratives, and attack infrastructure at scale.[3] The CrowdStrike 2026 Global Threat Report, published in February, quantifies the operational consequence: the average adversary breakout time has fallen to 29 minutes, down from 48 minutes in 2024, with the fastest recorded breakout occurring in 27 seconds. AI-enabled adversary operations increased 89 percent year-over-year.[4] Model Context Protocol creates a concrete and currently under-addressed attack surface: prompt injection through tool outputs, server-side request forgery at the agent-to-resource boundary, and authorization gaps where agents inherit credentials without scoped delegation. This follows the same adoption pattern as containers, APIs, and open-source dependencies: organic uptake at engineering speed, with the security reckoning arriving once adoption is effectively irreversible.[5] RSAC's coverage of MCP and agent-to-agent protocol security as distinct risk categories was largely absent from the main floor. The industry continues to run a familiar playbook.

In most default logging configurations, agent-initiated activity looks identical to human-initiated activity in security logs. CrowdStrike CTO Elia Zaitsev described the detection requirement directly at RSAC: distinguishing the two requires walking the process tree at the endpoint level.[6] Most organizations do not have that visibility in place. The ARISE Framework™ addresses this through its IDENTIFY and DETECT domains, which require a complete AI System Inventory and runtime observability of agent actions before those agents reach production.

1. TWiT.TV. AI Agents Are the New Security Perimeter: What RSAC 2026 Revealed. March 2026. https://twit.tv/posts/inside-twit/ai-agents-are-new-security-perimeter-what-rsac-2026-revealed
2. Forrester. RSAC Innovation Sandbox 2026: Two Sides of AI on Display. April 4, 2026. https://www.forrester.com/blogs/rsac-innovation-sandbox-2026-two-sides-of-ai-on-display/
3. Microsoft Security Blog. Threat actor abuse of AI accelerates from tool to cyberattack surface. April 2, 2026. https://www.microsoft.com/en-us/security/blog/2026/04/02/threat-actor-abuse-of-ai-accelerates-from-tool-to-cyberattack-surface/
4. CrowdStrike. 2026 Global Threat Report. February 24, 2026. https://www.crowdstrike.com/en-us/blog/crowdstrike-2026-global-threat-report-findings/
5. Futurum Group. RSAC 2026: The AI 'Tragedy of the Commons' and the Future of Agentic Security. April 2026. https://futurumgroup.com/insights/rsac-2026-the-ai-tragedy-of-the-commons-and-the-future-of-agentic-security/
6. VentureBeat. CrowdStrike, Cisco and Palo Alto Networks all shipped agentic SOC tools at RSAC 2026. March 31, 2026. https://venturebeat.com/security/rsac-2026-agentic-soc-agent-telemetry-security-gap

The Pilot-to-Production Gap Is Now the Central Problem

Enterprise AI adoption is not stalled; it is stratified. Organizations primarily have pilots, and very few have production-scale deployments with measurable governance. That gap is now becoming the defining condition of the market, and its consequences extend beyond just efficiency. Ungoverned high-risk AI is a liability, not an operational capability. Analysis of more than 300 enterprise AI sessions at the 2026 Data Innovation Summit found that the challenge is not building intelligence; it is building systems that can support and sustain it. Governance, risk, and compliance are no longer treated as optional considerations; they are the design constraints. The organizations that have moved past this constraint share a common thread: when senior leadership actively shapes AI governance, the organization can achieve significantly greater business value than those that delegate governance work to technical teams alone.[1]

Stanford University's Digital Economy Lab, in its April 2026 Enterprise AI Playbook, reviewed 51 documented enterprise deployments and found that the same use case and the same AI models produced vastly different deployment timelines across organizations. The differentiating factor was never the model. It was always organizational readiness, existing process infrastructure, and the clarity of governance ownership.[2] Organizations acquiring AI capabilities without establishing the governance layer first are not behind on technology; they are behind on the prerequisite conditions for technology to function effectively. A 2025 MIT NANDA initiative study, cited in the Stanford report, concluded that 95 percent of generative AI pilot programs fail to produce measurable financial impact. The failures stem not from model quality but from poor workflow integration and misaligned organizational incentives.[2] Deloitte's *2026 State of AI* in the Enterprise report documents agentic AI deployments operating in production across financial services, manufacturing, and logistics, handling workflows, customer transactions, and product development optimization.[3] These are not experimental environments; they are production operations in regulated industries, operating without the necessary behavioral baselines, data provenance controls, and identity management structures that regulatory and operational liability requires.

The ARISE Framework™ VALIDATE domain addresses this directly. Business logic must remain sound even if the AI agent is removed. That standard, applied before production deployment, is the difference between an AI initiative that scales and one that becomes an incident.

1. Hyperight. What 300+ Enterprise AI Use Cases on the Data Innovation Summit Reveal About 2026. April 2, 2026. https://hyperight.com/enterprise-ai-operationalization-2026/
2. Stanford Digital Economy Lab. The Enterprise AI Playbook: Lessons from 51 Successful Deployments. April 2026. https://digitaleconomy.stanford.edu/publication/enterprise-ai-playbook/
3. Deloitte. State of AI in the Enterprise 2026. 2026. https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-ai-in-the-enterprise.html

ASSESSED INTELLIGENCE

Secure & Responsible Technology.

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

SALES@ASSESSEDINTELLIGENCE.COM
EMEASALES@ASSESSEDINTELLIGENCE.COM