Langflow RCE CVE-2026-33017: Exploited Within 20 Hours

Unauthenticated Remote Code Execution in AI Pipeline Infrastructure and the Accelerating Weaponization of AI Platform Vulnerabilities

Unofficial AI-assisted Research

Cloud Security Alliance AI Safety Initiative

2026-03-25

Key Takeaways

- CVE-2026-33017, rated CVSS 9.3 Critical, is an unauthenticated remote code execution vulnerability affecting all Langflow versions prior to 1.9.0. Organizations running exposed Langflow instances must upgrade immediately.
- Active exploitation began approximately 20 hours after public disclosure on March 17, 2026 [1][6]. Attackers developed working exploits without relying on any published proof-of-concept code, suggesting that the advisory text and patch diff alone provided sufficient technical basis for exploitation.
- The vulnerability exploits an unsandboxed Python `exec()` call in a publicly accessible API endpoint — the same class of flaw exploited by CVE-2025-3248 — suggesting a pattern of incomplete remediation that addressed specific endpoints rather than the underlying unsafe execution design [2].
- Attackers progressed within 10 hours from automated scanning to targeted credential exfiltration, harvesting OpenAI, Anthropic, and AWS API keys along with database connection strings and application secrets [1].
- A second critical vulnerability, CVE-2026-33309 (CVSS 9.9), was disclosed concurrently and also patched in version 1.9.0. Organizations should treat the upgrade as urgent regardless of internet exposure posture [3].
- This incident illustrates a systemic risk in the AI ecosystem: platforms that bridge AI models to cloud infrastructure represent high-value targets whose compromise can cascade across multiple downstream services.

Background

Langflow is an open-source, low-code visual development platform for building AI agents and Retrieval-Augmented Generation (RAG) pipelines. Originally developed by Logspace, a startup founded in 2020, Langflow was acquired by DataStax in April 2024 [4]. The platform had accumulated more than 145,000 GitHub stars as of March 2026 [1][5], and is used by developers and enterprises to prototype and deploy applications powered by large language models. Its visual workflow model allows users to chain together

LLM providers, vector databases, data loaders, and external APIs without writing extensive custom code, reducing the development time required for teams building applications on models from Anthropic, OpenAI, and others [5].

This breadth of integration is also what makes Langflow a particularly attractive target. A single compromised Langflow instance is not simply a server breach; it is a potential entry point into every connected AI service credential, cloud account, and database the platform has been configured to access. Langflow deployments are often configured to integrate with multiple credential-holding services simultaneously — AI API providers, vector databases, and cloud accounts — concentrating credential risk in a single platform. Where traditional application server compromises might yield access to a single backend, Langflow deployments can aggregate sensitive API keys and credentials across an organization's entire AI stack. The security posture of the AI pipeline layer therefore substantially shapes the exposure surface of the AI supply chain it connects to.

CVE-2026-33017 was reported by security researcher Aviral Srivastava on February 26, 2026, and publicly disclosed on March 17, 2026, alongside the release of Langflow 1.9.0 containing the fix [1]. The vulnerability carries a CVSS base score of 9.3 [1][7], reflecting the combination of network accessibility, lack of authentication requirements, and the severity of impact — full system compromise without credentials.

Security Analysis

Vulnerability Mechanics

The core flaw lies in the `POST /api/v1/build_public_tmp/{flow_id}/flow` endpoint, which is intentionally designed to allow unauthenticated users to interact with publicly shared Langflow flows. While the design intent was reasonable — enabling public demos and shared AI applications — the endpoint accepts attacker-controlled flow data containing arbitrary Python code in node definitions. This code is passed directly to an unsandboxed `exec()` call in the `prepare_global_scope()` function within Langflow's code validation module. No sandboxing, no bytecode restriction, and no input validation stands between the attacker-supplied payload and full code execution on the host [2].

The architectural flaw is not new. CVE-2025-3248, which carried a CVSS score of 9.8 and was added to CISA's Known Exploited Vulnerabilities catalog in May 2025, exploited the same underlying `exec()` pattern through the `/api/v1/validate/code` endpoint [1][2][13]. The patch applied at the time added authentication gates to the validation endpoint but left the `exec()` call itself intact and

unsandboxed. CVE-2026-33017 represents an attacker routing around that gate through a different, still-unauthenticated endpoint. The remediation history of this codebase demonstrates the danger of treating authentication controls as a substitute for secure execution design: so long as `exec()` remains unsandboxed anywhere in an unauthenticated code path, the vulnerability class persists regardless of how many specific instances are addressed.

Exploitation requires a single crafted HTTP POST request to the vulnerable endpoint. No prior authentication, session tokens, or knowledge of specific flow identifiers beyond the default is required to achieve remote code execution as the Langflow service user.

Exploitation Timeline and Attacker Behavior

Active exploitation was first observed by Sysdig's threat research team approximately 20 hours after the advisory publication, beginning at 16:04 UTC on March 18, 2026 [1][8]. The attack campaign progressed through three observable phases that reveal an increasingly sophisticated threat posture.

The initial phase consisted of automated scanning using Nuclei-based payload templates, identifiable by the HTTP header `Cookie: client_id=nuclei-scanner`. These scans executed basic system commands (`id`) and exfiltrated results through interactsh callback infrastructure, confirming exploitability and establishing that the executing context was `uid=1000(langflow)`. Twelve unique interactsh subdomains across `.oast.live`, `.oast.me`, `.oast.pro`, and `.oast.fun` were observed as callback destinations [1].

The second phase, beginning several hours after initial scanning, involved custom Python-requests scripts performing active reconnaissance: directory enumeration, system fingerprinting, and attempted delivery of a stage-two dropper payload via curl from an attacker-controlled staging server. The transition from templated scanning to bespoke tooling suggests that at least one actor independently analyzed the advisory and developed custom exploitation capability without relying on shared public tooling.

The third phase, involving targeted data harvesting, produced the most directly actionable attacker outcomes. Attackers dumped full environment variable sets, extracted `.env` files, and exfiltrated database connection strings and API credentials for OpenAI, Anthropic, and AWS services [1]. This phase began within approximately 30 hours of public disclosure. Six unique source IP addresses were observed across the campaign, all registered to commercial VPS providers including DigitalOcean, Contabo, and AEZA GROUP in Frankfurt, Singapore, Paris, and the Netherlands — infrastructure consistent with opportunistic campaigns that prioritize operational security through commercial cloud anonymization.

No named threat actor or group has been publicly attributed to these attacks as of this writing. The observed tooling, infrastructure, and targeting patterns are consistent with financially motivated opportunistic actors rather than a sophisticated nation-state campaign, though the value of the credentials being harvested — particularly AI model API keys — warrants continued monitoring for secondary exploitation.

Concurrent Vulnerability Disclosure

CVE-2026-33309, disclosed on March 24, 2026, with a CVSS score of 9.9, represents a distinct but architecturally related vulnerability in Langflow's file upload functionality [3]. An authenticated attacker can submit a crafted multipart upload request with a path-traversal filename to the `POST /api/v2/files/` endpoint, writing arbitrary files anywhere on the host filesystem and thereby achieving remote code execution. This vulnerability affects versions 1.2.0 through 1.8.1, overlapping completely with CVE-2026-33017's affected range, and is also resolved by upgrading to version 1.9.0. Its root cause — inadequate input sanitization in `LocalStorageService` — represents a bypass of a prior patch for CVE-2025-68478, further reinforcing the pattern of incomplete remediation in this codebase.

Additionally, CVE-2026-27966 (hardcoded RCE in Langflow's CSV Agent component) and CVE-2026-21445 (broken authentication in Langflow monitor APIs) were disclosed in the same period [3]. The concurrent nature of these disclosures suggests coordinated research activity targeting the Langflow codebase as a whole, and organizations should expect continued vulnerability discoveries in AI pipeline platforms as security research attention increases.

Systemic Risk to the AI Supply Chain

The broader significance of CVE-2026-33017 extends beyond any single compromised instance. Langflow deployments serve as aggregation points for AI infrastructure credentials, often storing API keys for multiple foundation model providers alongside database access credentials and cloud account configurations. A single exploitation event can therefore yield simultaneous access to an organization's AI services budget and model access, cloud storage and compute resources, database contents that may include sensitive business or user data, and any downstream pipelines or automated agents that use the exfiltrated credentials.

This aggregation risk is a direct consequence of how AI pipeline platforms are architected and used. The convenience that makes tools like Langflow valuable — centralizing connections to many external services — creates a concentration of credential risk that security teams must explicitly account for. The

attack pattern observed in CVE-2026-33017, where credential exfiltration begins within hours of the first access, suggests that financially motivated actors are well aware of this value and have optimized their post-exploitation workflows accordingly.

Recommendations

Immediate Actions

Organizations running Langflow should treat the following steps as urgent:

- **Upgrade to Langflow 1.9.0 immediately.** This is the only complete remediation for both CVE-2026-33017 and CVE-2026-33309. No available workaround provides complete equivalent protection; network-level blocking of the vulnerable endpoint reduces exposure but does not address the underlying vulnerability class and should not be treated as a substitute for upgrading.
- **Assume compromise if running a pre-1.9.0 instance that was internet-accessible.** Rotate all API keys for any AI services (OpenAI, Anthropic, Cohere, and others) and cloud providers (AWS, GCP, Azure) that were configured in the Langflow environment. Assume that any credentials stored in `.env` files or environment variables have been exfiltrated.
- **Rotate database credentials** for any database connections configured in Langflow flows, regardless of whether evidence of access has been found. Credential exfiltration — a well-documented characteristic of credential theft attacks — does not always produce immediate observable indicators.
- **Audit network exposure.** Any Langflow instance directly accessible from the internet without network-level controls should be isolated pending upgrade. If the upgrade cannot be applied immediately, block or rate-limit the `/api/v1/build_public_tmp` endpoint at the reverse proxy or WAF layer as a temporary measure; this does not constitute a complete fix.

Short-Term Mitigations

Following the immediate response, organizations should implement controls that reduce the risk of recurrence across AI pipeline infrastructure.

Langflow instances, and AI pipeline platforms generally, should not be directly internet-accessible unless there is an explicit operational requirement. Where public access is required, deployment behind a WAF or API gateway with request inspection for suspicious payloads, rate limiting, and anomaly detection provides meaningful defense-in-depth. Authentication should be required at the perimeter even for endpoints that the application itself treats as public.

Runtime monitoring appropriate for AI pipeline infrastructure should include alerts for unusual process spawning (particularly shell command execution initiated by the Langflow process), outbound connections to interactsh or similar callback infrastructure, reads of credential files or environment variable dumps, and curl or wget execution from within the application process tree. Falco rules or equivalent runtime security tooling can detect these patterns when configured with accurate baselines for the pipeline platform's expected behavior, with false-positive rates that depend significantly on tuning and environment characterization.

For organizations that have not yet done so, credential management for AI pipeline environments should be centralized in a secrets management system (such as HashiCorp Vault, AWS Secrets Manager, or equivalent) rather than stored in `.env` files or application-level configuration. This limits the impact of a future credential exfiltration event by reducing dwell time between exfiltration and detection and enabling rapid rotation without application reconfiguration.

Strategic Considerations

CVE-2026-33017 is not an isolated incident but a manifestation of a structural risk category that organizations deploying AI pipeline infrastructure must address deliberately. The pattern of a critical unauthenticated RCE in a widely-deployed AI development platform, weaponized within hours, exploiting the same code path that was previously patched, and yielding a harvest of credentials that touches the AI supply chain, is likely to recur as the ecosystem matures — the combination of immature security engineering practices, high-value credential aggregation, and rapid researcher attention makes AI pipeline platforms a persistent target category.

Security teams should establish formal asset inventory processes for AI pipeline tools and frameworks, applying the same vulnerability management discipline to Langflow, LlamaIndex, Haystack, Flowise, and equivalent platforms that they apply to web application frameworks and middleware. These tools are now production infrastructure, and should be managed accordingly.

Organizations that rely on foundation model APIs should implement API key scoping and rotation policies that limit the blast radius of a credential compromise. API keys used by pipeline platforms should be scoped to the minimum necessary permissions, should carry configured usage limits that would flag

anomalous consumption, and should be rotated on a schedule independent of any known compromise events.

Finally, the supply chain dimension of this vulnerability class warrants ongoing attention. AI pipeline platforms, by design, sit between developers and the model providers, datastores, and cloud services that power production AI applications. Securing this layer requires treating it with the same rigor applied to identity infrastructure and secrets management — because in practice, it controls access to both.

CSA Resource Alignment

The vulnerabilities described in this research note map directly to threat categories and control domains addressed by multiple CSA frameworks.

The MAESTRO framework for agentic AI threat modeling specifically addresses the risk surface of AI pipeline orchestration layers, including the trust boundaries between orchestrators, tool integrations, and external services [9]. CVE-2026-33017 is a concrete instantiation of the orchestrator compromise scenario that MAESTRO treats as a primary risk vector: an attacker who controls the orchestration layer gains access to the credentials and effective permissions of all downstream integrations configured in that instance. Organizations building on agentic AI platforms should conduct MAESTRO-informed threat modeling of their pipeline topology before deployment.

The CSA AI Controls Matrix (AICM) provides controls directly applicable to this incident [10]. AI-RM-07 (AI Supply Chain Risk Management) addresses the risk of vulnerabilities in AI development tooling and pipeline components. AI-IS-03 (AI Infrastructure Security) covers the requirement for runtime isolation and input validation in AI platform components. AI-DG-04 (Data Governance for AI Systems) is relevant to the credential and sensitive data exfiltration observed in the exploitation campaign. For AI-specific deployments, the AICM provides more targeted controls than the Cloud Controls Matrix (CCM) and should be the primary reference framework for organizations assessing their AI infrastructure posture, supplemented by CCM where cloud-general controls apply.

The CSA Agentic AI Red Teaming Guide, published in 2025, provides structured methodologies for testing the security of agentic and pipeline AI systems, including supply chain and dependency attack categories that directly encompass the risk demonstrated by this vulnerability [9]. Organizations that have not subjected their AI pipeline deployments to red team assessment should treat this incident as a prompt to initiate one.

CSA's AI Organizational Responsibilities publications address governance requirements for AI tooling, including vendor risk management obligations relevant to organizations that have deployed Langflow or equivalent platforms in production without formal security assessment [11]. The incident demonstrates that AI development tooling has outpaced organizational governance processes in many deployments, a gap that CSA's AI organizational responsibility framework is designed to help close.

For organizations seeking to demonstrate their AI security posture to external stakeholders, the CSA STAR for AI program provides a structured self-assessment and certification pathway through the AI Consensus Assessments Initiative Questionnaire (AI-CAIQ) [12]. Completing a STAR for AI Level 1 assessment would surface the control gaps most relevant to the risks illustrated by this incident.

References

[1] Sysdig Threat Research Team, "CVE-2026-33017: How Attackers Compromised Langflow AI Pipelines in 20 Hours," Sysdig Blog, March 2026. https://www.sysdig.com/blog/cve-2026-33017-how-attackers-compromised-langflow-ai-pipelines-in-20-hours

[2] Barrack AI Security Research, "Langflow Got Hacked Twice Through the Same exec() Call," Barrack AI Blog, March 2026. https://blog.barrack.ai/langflow-exec-rce-cve-2026-33017/

[3] GitLab Advisory Database, "CVE-2026-33309: Langflow Arbitrary File Write Leading to RCE," GitLab Advisories, March 24, 2026. https://advisories.gitlab.com/pkg/pypi/langflow/CVE-2026-33309/

[4] Kyle Wiggers, "DataStax Acquires Logspace, the Startup Behind the Langflow Low-Code Tool for Building RAG-Based Chatbots," TechCrunch, April 4, 2024. https://techcrunch.com/2024/04/04/datastax-acquires-logspace-the-startup-behind-the-langflow-low-code-tool-for-building-rag-based-chatbots/

[5] Langflow Project, "Langflow — Official Site," DataStax, 2026. https://www.langflow.org/

[6] The Hacker News, "Critical Langflow Flaw CVE-2026-33017 Triggers Attacks Within 20 Hours of Disclosure," The Hacker News, March 2026. https://thehackernews.com/2026/03/critical-langflow-flaw-cve-2026-33017.html

[7] Wiz Research, "CVE-2026-33017 — Impact, Exploitability, and Mitigation Steps," Wiz Vulnerability Database, March 2026. https://www.wiz.io/vulnerability-database/cve/cve-2026-33017

[8] SecurityWeek, "Critical Langflow Vulnerability Exploited Hours After Public Disclosure," SecurityWeek, March 2026. https://www.securityweek.com/critical-langflow-vulnerability-exploited-hours-after-public-disclosure/

[9] Cloud Security Alliance, "Agentic AI Red Teaming Guide," CSA AI Organizational Responsibilities Working Group, 2025.

[10] Cloud Security Alliance, "AI Controls Matrix (AICM)," CSA AI Controls Framework Working Group, 2025.

[11] Cloud Security Alliance, "AI Organizational Responsibilities — Core Security Responsibilities," CSA AI Safety Initiative, 2025.

[12] Cloud Security Alliance, "STAR for AI Level 1 Submission Guide," CSA STAR Program, 2025.

[13] CISA Known Exploited Vulnerabilities Catalog, "CVE-2025-3248: Langflow Missing Authentication Vulnerability," CISA, May 5, 2025. https://www.cisa.gov/known-exploited-vulnerabilities-catalog