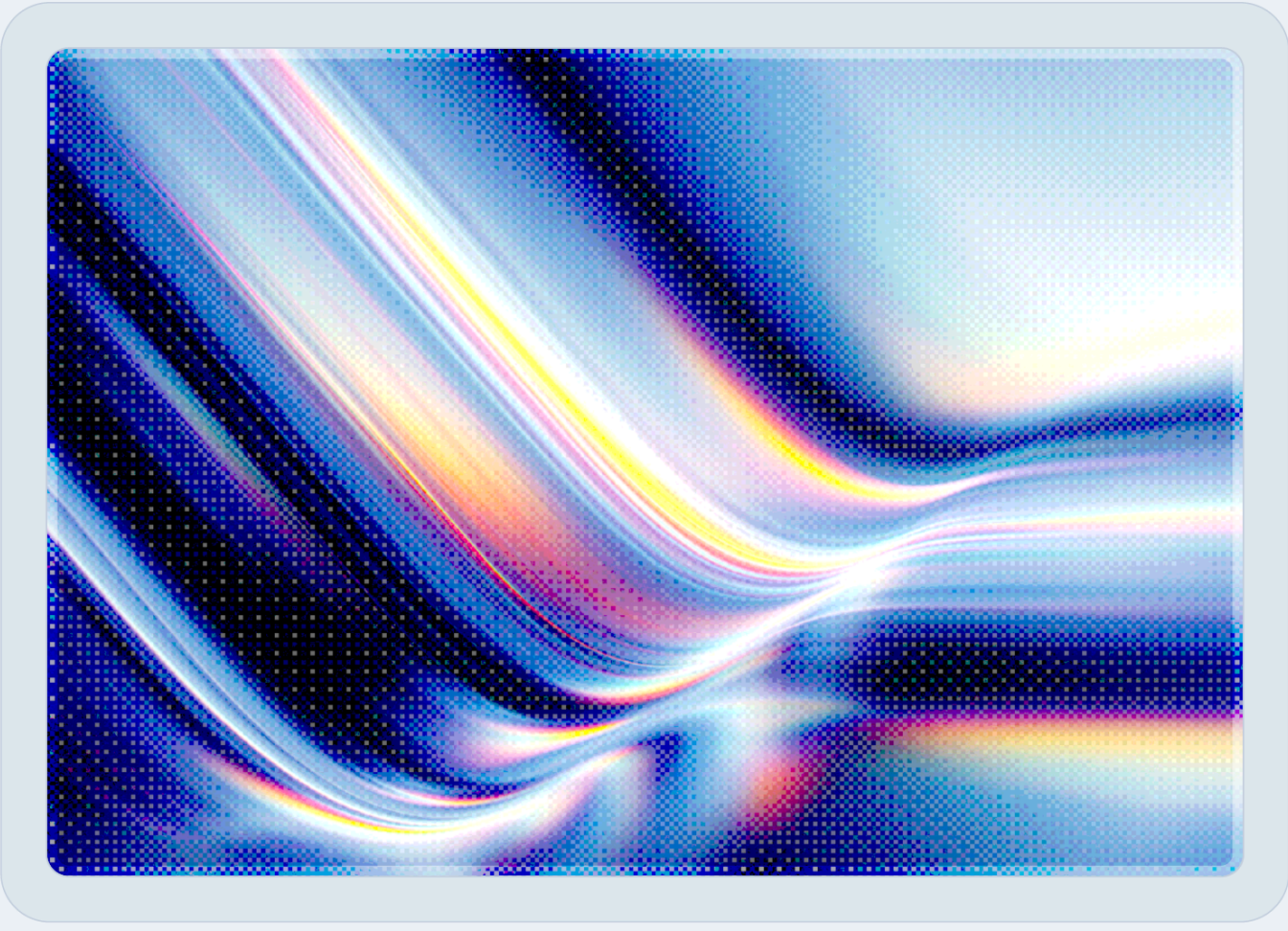
CSAI Foundation | Cloud Security Alliance

Two Attackers, Two Playbooks: Langflow Under Siege

VulnCheck Canaries Reveal Parallel Credential-Theft and Cryptomining Campaigns

2026-09-08

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

© 2026 Cloud Security Alliance. Some rights reserved.

You may download, store, display, view, print, redistribute, and link to this document in its original, unmodified form, provided that attribution to the Cloud Security Alliance is maintained and all trademark and copyright notices remain intact.

This document may not be modified or altered. You may quote portions of the document as permitted by the Fair Use provisions of the United States Copyright Act, provided that attribution is given to the Cloud Security Alliance.

This document may be shared on professional and social media platforms in its original form with attribution.

This document was generated with AI assistance and has not undergone official CSA review and approval processes.

Key Takeaways

VulnCheck's threat research team deployed vulnerable Langflow honeypots, or "canaries," across the open internet and captured two independent, financially motivated attackers running distinct intrusion campaigns against the same AI orchestration platform over roughly the same period in 2026 [1]. Published August 28, 2026, VulnCheck's report — "Same Target, Different Playbooks: Two Attackers, Two Different Paths to Pwning the AI Stack" — documents an operator who entered through CVE-2026-5027, a path-traversal-to-remote-code-execution flaw, and over the ten days between May 20 and May 30, 2026, deployed a Python credential harvester, proxy agents, and a SimpleHelp remote-access tool; the operator did not establish an IRC-based command-and-control channel until June 8, roughly three weeks after initial access [1]. A second, unrelated operator entered through CVE-2025-3248 as early as April 22, 2026, and over five distinct waves through late June built a Monero cryptomining operation, layering in a Chisel SOCKS5 tunnel, disabling the host's audit daemon to blind forensic logging, and ultimately exploiting a second flaw, CVE-2026-0769, to drop additional persistence and pivot toward other targets [1].

Neither campaign has been attributed to a named threat actor or group; VulnCheck characterizes both as "financially motivated" despite pursuing different monetization paths — one harvesting credentials for resale or downstream access, the other mining cryptocurrency directly on compromised infrastructure [1]. The report situates these two intrusions within a much larger pattern: VulnCheck counts twelve distinct Langflow vulnerabilities exploited in the wild during 2026 alone, against a single known exploited Langflow flaw in all prior years combined, with cumulative successful exploitation attempts against CVE-2026-0769, CVE-2025-3248, and CVE-2026-5027 alone — the three CVEs central to this and the prior CSA notes on Langflow — exceeding 15,000 [1][6]. CSA has previously documented three of those individual vulnerabilities — CVE-2026-33017, CVE-2026-5027, and CVE-2026-0768 — as standalone incidents [7][8][9]; this note instead examines what VulnCheck's canary telemetry reveals about attacker tradecraft once code execution is achieved, since the same underlying flaws are now confirmed to support multiple independent operators running qualitatively different post-exploitation playbooks against the same product.

Based on these two campaigns, the practical implication for defenders is that patching a specific CVE addresses only the entry vector observed here, not necessarily the full range of objectives an attacker might pursue once inside. A Langflow host compromised via any of its dozen exploited-in-2026

vulnerabilities should be assumed capable of hosting credential exfiltration, remote-access tooling, cryptomining, or lateral movement infrastructure, and incident response should scope accordingly rather than checking only for the specific payload associated with the CVE that provided initial access.

Background

Langflow is an open-source, Python-based visual builder for large language model pipelines, retrieval-augmented generation applications, and autonomous AI agents, distributed under DataStax's stewardship following IBM's 2025 acquisition of DataStax [10][11]. The project has grown rapidly among developers assembling agentic workflows without extensive custom code, reporting more than 153,000 GitHub stars and roughly 400 contributors as of VulnCheck's report [1]. That popularity, combined with Langflow's practice of storing API keys for LLM providers such as OpenAI and Anthropic, cloud credentials for AWS, and other integration secrets directly on the host running the application, has made it an attractive target: a single compromised instance can yield authenticated access to every downstream service the workflow was configured to call [1][9].

VulnCheck's canary methodology involves deploying real, vulnerable Langflow instances on the public internet and observing the full lifecycle of unsolicited attacker interaction, rather than relying on static honeypot signatures or simulated services [1]. This approach let researchers capture not just the initial exploitation request but the complete sequence of follow-on commands, file drops, and network connections each attacker executed after gaining code execution, producing an unusually direct view of post-compromise tradecraft rather than inferred behavior from scan logs alone. VulnCheck maintains its own Known Exploited Vulnerabilities catalog alongside CISA's KEV list and reported that several of the Langflow CVEs implicated in these campaigns appear on one or both lists, reflecting confirmed real-world exploitation rather than theoretical risk [1].

The two vulnerabilities central to this report both trace to weaknesses that CSA has previously analyzed in isolation. CVE-2026-5027 is a path-traversal flaw in Langflow's file-upload endpoint that allows an unauthenticated attacker to write arbitrary files to the filesystem, achieving remote code execution because Langflow's default configuration issues valid session tokens without requiring credentials [8]. CVE-2025-3248 is an earlier, distinct unauthenticated RCE that Langflow's maintainers patched in version 1.3.0; independent vulnerability research attributes the fix to removing the `/api/v1/validate/code` endpoint's reliance on Python's `exec()` function [5]. CVE-2026-0769, the flaw the cryptomining operator used mid-campaign to extend persistence, is a separate eval-injection vulnerability in the `eval_custom_component_code` function that likewise permits unauthenticated remote code execution through unsanitized user input passed to Python's code-execution path [4][12]. That three functionally similar but independently numbered vulnerabilities –

spanning file upload, code validation, and custom component evaluation — all reduce to the same class of unauthenticated-input-reaches-code-execution defect suggests a recurring pattern in how user-supplied data has been handled across Langflow's subsystems, though the two-campaign, three-CVE sample examined here is not sufficient to characterize the project's broader engineering practices.

Security Analysis

Attacker One: Compressed Credential Extraction

VulnCheck's canary first recorded Attacker One's activity on May 20, 2026, entering through CVE-2026-5027 [1]. Rather than lingering for reconnaissance, this operator moved through a tight, largely automated sequence: deploying a Python-based credential harvester alongside proxy agent software and an installation of SimpleHelp, a legitimate remote-monitoring-and-management tool frequently repurposed by intrusion operators for persistent remote access [1]. The credential harvester itself executed in under 21 seconds once triggered, consistent with a pre-scripted tool rather than manual, interactive exploration of the host [1]. Persistence was established through a cron entry — `0 * * * * /usr/bin/3WA72N.sh &` — that re-executed an attacker-controlled shell script every hour, ensuring the harvester or its supporting tools would survive a reboot or accidental process termination [1]. Harvested material was exfiltrated to an external endpoint at `23.234.98[.]182:9999`, and the operator additionally stood up an IRC-based command-and-control channel at `185.117.74[.]172:6667`, a comparatively old-fashioned C2 protocol that nonetheless remains effective against monitoring tools tuned primarily for HTTP- and DNS-based exfiltration [1]. By May 30, ten days after initial access, the harvester, proxy agents, RAT, and cron persistence were fully deployed; the operator did not stand up the IRC-based command-and-control channel until June 8, nineteen days after initial access. Even accounting for that longer full timeline, the pace at which the core credential-harvesting toolkit was assembled and triggered indicates either a highly automated toolkit or an operator working from a well-rehearsed playbook rather than improvising against each new target [1].

Attacker Two: Patient, Multi-Wave Cryptomining Expansion

The second operator's campaign unfolded on a markedly different timescale, beginning even earlier — April 22, 2026 — and progressing through five distinct waves over roughly two months [1]. Initial access came through CVE-2025-3248, notable because that vulnerability had already been patched in Langflow 1.3.0 more than a year before this campaign began, meaning the operator was working exclusively against unpatched, internet-exposed instances that had never applied a fix available since

early 2025 [1][5]. VulnCheck's wave-by-wave reconstruction shows the operator deploying a proxy agent in the second wave (May 19), installing a Chisel SOCKS5 tunnel for pivoting and traffic relay in the third wave (May 25), and launching an XMR cryptocurrency miner in the fourth wave (May 29) — each step separated by days to weeks rather than executed as a single automated burst [1]. This unhurried cadence suggests an operator managing a portfolio of compromised hosts rather than optimizing for speed on any single target, consistent with cryptomining operations that value sustained, low-visibility resource theft over rapid one-time payout.

On June 10, this operator's use of Langflow diverges most sharply from Attacker One's: rather than exfiltrating and departing, the operator disabled the host's `auditd` daemon, removing a primary source of Linux system-call logging that incident responders rely on for forensic reconstruction [1]. Thirteen days later, in the campaign's fifth wave on June 23, the operator exploited CVE-2026-0769, a second and independent Langflow vulnerability, to drop additional persistence mechanisms including files named `.sysd`, `.cache-sysd`, and `.watchdog.sh` [1]. This is a well-documented masquerading technique (see MITRE ATT&CK T1036) for blending malicious processes into a Linux host's normal service inventory, reducing the likelihood that a defender scanning process lists or cron tables notices the addition. On June 25, VulnCheck recorded the campaign's final observed action: the operator conducting reconnaissance against `216.78.235.34` via SSH, indicating an intent to pivot beyond the Langflow host itself toward other reachable systems [1]. Auditd suppression combined with a second independent CVE used weeks later for persistence suggests an operator capable of returning to a compromised host and expanding their foothold over months, rather than merely executing a single opportunistic exploit-and-mine sequence.

Two Campaigns, One Shared Root Cause

VulnCheck's side-by-side documentation of these campaigns is notable less for either individual attacker's tradecraft than for what it demonstrates about Langflow as a target class. Both operators achieved their respective objectives — credential theft in one case, cryptomining infrastructure in the other — through the same fundamental weakness: an unauthenticated code-execution primitive reachable over the network, differing only in which Langflow subsystem exposed it (file upload for Attacker One, code validation and later custom component evaluation for Attacker Two) [1]. Neither operator required a sophisticated exploit chain or privilege escalation; Langflow's documented default of issuing valid session tokens without authentication meant a single crafted request was sufficient in each case to reach full remote code execution [1][8]. That two operators with entirely different monetization strategies converged independently on Langflow as a viable target, using different CVEs but the same class of underlying defect, suggests the platform's exposure to this defect class is not confined to any one patched-and-closed vulnerability, though confirming it is systemic across the broader codebase would require a fuller audit than the three CVEs discussed here. VulnCheck's broader count of twelve

Langflow CVEs exploited in the wild during 2026 — a twelve-fold increase over the single known exploited flaw in all prior years combined — reinforces that this is an ongoing pattern rather than a closed chapter, a trend independently corroborated by security reporting on the platform's most recent exploited CVE and by parallel credential-probing campaigns researchers observed against other application frameworks during the same window [2][3][6].

The report also does not attribute either campaign to a named threat actor, nation-state group, or known cybercrime organization, describing both simply as financially motivated [1]. That absence of attribution is itself informative: it suggests these are opportunistic operators scanning for exposed, vulnerable instances rather than actors specifically targeting particular organizations, meaning any internet-facing Langflow deployment running an affected version is a plausible target regardless of the operating organization's size, sector, or profile.

Recommendations

Immediate Actions

Organizations running self-hosted Langflow should confirm their deployed version against the patched releases for CVE-2026-5027, CVE-2025-3248, and CVE-2026-0769, treating any instance running an affected version and reachable from the public internet as compromised until proven otherwise. Because both campaigns documented here established persistence — a recurring cron job in one case, systemd-mimicking files in the other — version confirmation alone is insufficient; incident responders should also inventory cron tables, `/etc/systemd` and adjacent directories, and running processes for unfamiliar entries, and check outbound network connections against the IOCs VulnCheck published, including the exfiltration endpoint `23.234.98[.]182:9999` and IRC C2 host `185.117.74[.]172:6667` [1]. Any Langflow instance with signs of compromise should have every credential it held rotated immediately, including LLM provider API keys, cloud access keys, and Langflow's own secret key, since either campaign's operator would have had unrestricted code execution sufficient to access all of them.

Short-Term Mitigations

Because Attacker Two's cryptomining campaign specifically disabled `auditd` to blind forensic visibility, organizations should verify that host-based logging and monitoring agents are configured to alert on their own tampering or termination, not merely to log attacker activity that assumes uninterrupted logging. Network egress filtering should restrict Langflow hosts from establishing arbitrary outbound connections, since both campaigns depended on reaching external infrastructure — an IRC C2

server in one case, a SOCKS5 tunnel endpoint in the other — that egress controls scoped to only the LLM and cloud provider APIs a given workflow legitimately requires would have disrupted. Organizations that do not need Langflow reachable from the public internet should place it behind a VPN or authenticated reverse proxy rather than exposing it directly, closing off the opportunistic scanning that both operators in this report appear to have relied on to find targets.

Strategic Considerations

VulnCheck's side-by-side documentation of two independent, differently motivated campaigns exploiting the same product illustrates a broader point for organizations evaluating AI orchestration platforms: a vendor's exposure to opportunistic exploitation should be assessed at the level of the product's overall security engineering practices, not vulnerability by vulnerability. A platform that has produced twelve exploited-in-the-wild CVEs within a single year, spanning file handling, code validation, and component evaluation subsystems, warrants heightened scrutiny before being granted access to production-grade credentials, regardless of whether any specific outstanding CVE currently applies to an organization's deployed version. Security teams should require evidence that AI development and orchestration tooling receives the same adversarial testing and patch-cadence discipline applied to customer-facing production systems, and should assume that any internet-exposed instance may already host multiple concurrent, unrelated attacker footholds rather than a single identifiable compromise.

CSA Resource Alignment

CSA has published targeted analysis of three prior Langflow vulnerabilities directly relevant to the campaigns VulnCheck documents in this report.

Langflow Path Traversal: Unauthenticated RCE Actively Exploited: This CSA research note covers the exact vulnerability Attacker One used for initial access, detailing the path-traversal-to-arbitrary-file-write mechanics and Langflow's unauthenticated default session behavior that made single-request exploitation possible [8]. Its remediation guidance on network segmentation and version verification applies directly to the credential-harvesting campaign analyzed here.

Langflow RCE CVE-2026-33017: Exploited Within 20 Hours: Though not one of the two CVEs central to VulnCheck's report, this earlier CSA note documents the same unauthenticated-RCE-to-credential-theft pattern in a different Langflow endpoint, reinforcing that the weakness VulnCheck's dual-campaign analysis surfaces has recurred across multiple, independently numbered vulnerabilities throughout 2026 [7].

Langflow Zero-Day Exploited to Harvest AI and Cloud Credentials: This CSA note on CVE-2026-0768 documents the twelve-CVE exploitation pattern this report also references and provides detailed recommendations on credential rotation and secrets-management practices that apply without modification to hosts compromised via either campaign described here [9].

AI Controls Matrix (AICM) v1.1: As the superset successor to the Cloud Controls Matrix, AICM's vulnerability management and application security control domains provide the framework organizations should use to evaluate whether AI orchestration tooling like Langflow receives adversarial testing commensurate with the credentials it is granted, while its identity and access management domain supports the network egress and credential-scoping controls recommended in this note [13].

References

1. Jonathan Reiter, Guillermo Menjivar, and Patrick Garrity. "Same Target, Different Playbooks: Two Attackers, Two Different Paths to Pwning the AI Stack." VulnCheck, August 28, 2026.
2. Bill Toulas. "Critical Langflow flaw exploited to steal OpenAI and AWS keys." BleepingComputer, September 2026.
3. Ravie Lakshmanan. "Attackers Exploit Critical Langflow and Rails Flaws in Credential-Probing and C2 Activity." The Hacker News, September 2026.
4. GitHub Advisory Database. "Langflow eval_custom_component_code Eval Injection Remote Code Execution Vulnerability (CVE-2026-0769)." GitHub, 2026.
5. GitHub Security Advisories. "Langflow Unauth RCE – GHSA-rvqx-wpfh-mfx7 (CVE-2025-3248)." GitHub, 2025.
6. Forkast News. "Langflow's 12th Exploited CVE Confirms AI Frameworks Are Now Credential Harvesting Infrastructure." Forkast News, September 2026.
7. Cloud Security Alliance AI Safety Initiative. "Langflow RCE CVE-2026-33017: Exploited Within 20 Hours." CSA Labs, 2026.
8. Cloud Security Alliance AI Safety Initiative. "Langflow Path Traversal: Unauthenticated RCE Actively Exploited." CSA Labs, 2026.
9. Cloud Security Alliance AI Safety Initiative. "Langflow Zero-Day Exploited to Harvest AI and Cloud Credentials." CSA Labs, September 4, 2026.
10. IBM Newsroom. "IBM to Acquire DataStax, Deepening watsonx Capabilities and Addressing Generative AI Data Needs for the Enterprise." IBM, February 25, 2025.
11. Langflow. "Big news for Langflow!." Langflow Blog, 2025.
12. Mondoo. "CVE-2026-0769 Vulnerability Intelligence." Mondoo, 2026.
13. Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." CSA, 2026.

© 2026 Cloud Security Alliance. AI-assisted Rapid Research.