CSAI Foundation | Frontier Ready Cybersecurity

Frontier Ready Daily

Issue 85 · Member edition

2026-11-19



CSAI FOUNDATION INITIATIVE

Frontier Ready.

Machine-speed agentic cybersecurity

cloud security alliance®

CSAI

© 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.

In this issue

Editorial notice. This issue was published without meeting every Frontier Ready editorial rule. Unmet on publication: item 1 (Five teams found the same nginx HTTP/2 flaws): confidence segment 'CVE-2026-123007 reported 16 November) and its list of five i' does not begin with a vocabulary term. Provenance labels only — vendor-interest disclosure belongs in the 'Vendor interest' field. It is published because a missing issue serves readers worse than a flagged one; treat the affected items with corresponding caution.

The lead item is nginx's two critical HTTP/2 flaws, fixed on 18 November 2026. Five research teams reported them independently, one with an AI harness that flagged the first bug in 90 minutes, and a public proof of concept was out before Ubuntu or Debian had shipped fixed packages. The others are CISA adding a LiteLLM model-gateway flaw to its exploited list after it was used to turn inference keys into admin credentials, Ruby fixing seven memory-safety flaws (four credited to Anthropic with Ada Logics) under GitHub advisory IDs that CVE-keyed scanners will not yet match, telemetry showing that AI agents mostly run on their builders' OAuth grants, and new IETF drafts for agent discovery that leave catalog trust unsettled. We examined GTIG's COLD RELIC cloud-impersonation report (an identity intrusion with no agentic or machine-speed element), the TheMailer AI phishing kit (an actor using models for lure-writing, which our scope excludes), PHP's seven critical fixes and the PostgreSQL/MySQL/Oracle RCE disclosures (no AI-discovery or capacity signal in the reporting), the Danfoss controller zero-days (found by manual review), and the agent-framework, Java4Shell, Puppet and GTIG exploit_agent stories (already covered in recent issues), and left them out.

Five teams found the same nginx HTTP/2 flaws and a public exploit was out before Ubuntu or Debian had patched them

Category: vuln_storm **Significance:** major **Confidence:** VERBATIM (PROVIDER) for the 18 November 2026 fixes (nginx 1.31.9 and 1.30.7, NGINX Plus 37.3.0.1 and 37.0.8.1), the affected range 1.9.5 through 1.31.8 for CVE-2026-123007, the public proof of concept on GitHub, the Ubuntu and Debian trackers listing nginx as still vulnerable on disclosure day, and the absence of reported exploitation, as reported by The Hacker News; SELF-REPORTED (PROVIDER METRIC) for Wiz's disclosure timeline (CVE-2026-123633 reported 21 October and already known to F5 from another team; CVE-2026-123007 reported 16 November) and its list of five independent reporters, and for Xint's statement that its AI tooling flagged the first bug in about 90 minutes with enough detail to build a working exploit in under ten

(publication of F5's own credit list, or an independent reproduction of Xint's run, would confirm these); LINK ONLY — VERIFY AT SOURCE for the exposure estimates of nearly 10 million services and 98.9% of version-identifiable servers, which count versions rather than exploitable configurations; CHARACTERIZATION (CSA) for treating the five-way collision as evidence that the embargo window no longer protects anyone **Vendor interest:** Wiz and Xint, whose accounts supply the discovery timeline and the AI figures, both sell vulnerability-discovery or exposure-management capabilities, so AI-speed framing is commercially favourable to them. **Source:** The Hacker News — 10-Year-Old NGINX HTTP/2 Flaw Lets Unauthenticated Attackers Execute Code on Servers

What changed. F5 fixed two critical nginx HTTP/2 flaws on 18 November 2026. CVE-2026-123007 is an integer underflow in HTTP/2 flow control that dates back to 2015 and is reachable in default configurations with HTTP/2 on. CVE-2026-123633 is a use-after-free that matters when nginx is an HTTP/2 reverse proxy. Both are rated 9.2. According to Wiz's write-up, the flaws were reported independently by Wiz, Ikotas Labs, Metnir, Rapid7 and Xint. Wiz used an AI research harness to find both bugs. Its 21 October report on the first was a duplicate of one F5 already had, and its report on the second arrived two days before disclosure. Xint says its AI tooling flagged the first bug in about 90 minutes and that a working exploit took under ten more. It also chained a Linux kernel SCTP bug to reach root on AWS, GCP and Azure hosts. A proof of concept went up on GitHub at disclosure. That day, Ubuntu's tracker listed nginx as vulnerable from 16.04 through 26.10 and Debian's listed it as vulnerable from bookworm through sid. No exploitation has been reported. The previous known pre-auth RCE in nginx was in 2013.

Why it reaches you. The component class is the internet-facing reverse proxy and ingress tier, including Kubernetes ingress controllers and the origin servers behind a CDN. Most estates take nginx from the distribution, not from nginx.org, so the effective patch date is the distribution's date, which had not arrived when the exploit code went public. Distribution packages may also carry backported fixes without changing what `nginx -v` reports, so version-based scanning will show false positives and negatives in both directions. A reload does not swap out the vulnerable binary, so you need a full restart. With five teams finding the same bug, an embargo protected no one. A sixth finder with no disclosure intent is a reasonable planning assumption.

What to do. Edge and platform owner: escalate. Inventory every nginx listener with HTTP/2 enabled. Test origins directly rather than through the CDN by checking for `h2` with ALPN. Patch from upstream or the vendor image now where the distribution has not shipped. Where you cannot, set `http2 off` on affected listeners, which blocks CVE-2026-123007 but not the remaining configuration-specific flaws. Confirm completion by restart and binary hash, not by version string. Urgency: escalate.

CSA coverage. CVE-2026-119618: AI-Discovered Apache HTTP/2 RCE and the Distribution Patch Gap; see also The Collapsing Patch Window: Project Zero's Full X and AI-Speed Remediation.

CISA added a LiteLLM gateway flaw to its exploited list after attackers used it to turn inference keys into admin credentials

Category: agentic_surface **Significance:** major **Confidence:** VERBATIM (PROVIDER) for CISA's 18 November 2026 KEV addition of CVE-2026-40217 (CVSS 8.8, code execution through the `/guardrails/test_custom_code` endpoint) and the 2 December federal deadline, as reported by The Hacker News; LINK ONLY – VERIFY AT SOURCE for GTIG's attribution of the exploitation to LLMjacking campaigns that raised limited inference keys to full admin credentials across finance, retail, transport and technology, which reaches us only through secondary reporting; LINK ONLY – VERIFY AT SOURCE for the fixed version, which public advisories give inconsistently (releases after 8 April 2026, or before 1.25.0) and with conflicting statements on whether a master key is needed; CHARACTERIZATION (CSA) for the roughly seven-month interval between the fix and the KEV listing **Source:** The Hacker News – CISA Warns of 4 Exploited Flaws Linked to Ransomware, Cryptomining, and LLMjacking Attacks

What changed. On 18 November 2026 CISA added CVE-2026-40217 in LiteLLM, the open-source model gateway, to its Known Exploited Vulnerabilities catalog, with a federal deadline of 2 December. The flaw lets a caller run arbitrary code through the endpoint that tests custom guardrail code, by escaping that endpoint's hand-built Python sandbox. The Hacker News reports Google Threat Intelligence Group's finding that attackers used it to turn limited-access inference keys into full admin credentials and to persist, as part of LLMjacking campaigns that resell or abuse model access. Public advisories date the fix to around April 2026. If that holds, the gateways being exploited now have gone unpatched for about seven months.

Why it reaches you. The component class is the model gateway: the proxy that holds every provider key, enforces spend limits and, increasingly, runs your guardrails. The vulnerable endpoint is part of the guardrail feature, so the control meant to constrain agent traffic is the way in. Once an attacker holds admin on the gateway, they hold every upstream provider key, they can read or rewrite prompts and tool calls passing through, and they can switch guardrails off. Earlier this week this feed reported Amazon's decoys catching agents hunting exposed gateways for default master keys. This listing shows the same component class being exploited in production estates. Gateways are often deployed by platform or data-science teams outside the vulnerability-management inventory, which is how a seven-month-old fix goes unapplied.

What to do. AI platform owner: escalate. Find every LiteLLM instance, including those teams run themselves, and upgrade it. Block `POST /guardrails/test_custom_code` at the reverse proxy whatever the version. Rotate the gateway master key and every upstream provider key it held, and check provider billing and the gateway's key list for keys you did not issue. Put model gateways in the same patch SLA tier as identity infrastructure. Urgency: escalate.

CSA coverage. Self-Funding Agents: ARTEX Campaigns Harvesting AI-Gateway Keys.

Ruby fixed seven memory-safety flaws, four credited to Anthropic, under advisory IDs that CVE-keyed scanners will not yet match

Category: vuln_storm **Significance:** notable **Confidence:** VERBATIM (PROVIDER) for the 18 November 2026 fixes in Ruby 4.0.10, 3.4.13 and 3.3.14, the per-advisory credits (four to Stephan Lipp of Ada Logics and Anthropic), the maintainers' statement that they are aware of proofs of concept reaching remote code execution, CVE IDs having been requested but not assigned, and Ruby 3.2 and earlier receiving no fix; LINK ONLY – VERIFY AT SOURCE for The Hacker News's count of five flaws with public RCE proofs of concept and its reference to an earlier regex fix release on 23 October; NO PROVIDER CLAIM for the method behind the Anthropic-credited findings, which the advisory does not describe; CHARACTERIZATION (CSA) for the scanner-coverage gap **Source:** Ruby – Security advisories: multiple vulnerabilities in Regexp and String case conversion

What changed. On 18 November 2026 Ruby fixed seven flaws in its regular-expression engine and string case conversion: heap overflows, a double free, an integer overflow and out-of-bounds reads. Four are credited jointly to Ada Logics and Anthropic. The maintainers say they know of proofs of concept that reach remote code execution. CVE identifiers were requested but had not been assigned at release, so the advisories carry only GitHub Security Advisory IDs. Ruby 3.2 and earlier are end-of-life and will not be fixed. The Hacker News reports this is the second regex security release in under four weeks.

Why it reaches you. The component class is any application runtime that runs regex or case conversion on attacker-controlled strings, which includes most Rails applications and much of the CI tooling written in Ruby. The process problem is separate from the bugs. Many vulnerability scanners and SBOM matchers key on CVE IDs, so a fix published only under GHSA IDs may not appear in their findings until CVEs are assigned and propagated, even though exploit code reportedly exists already. Applications still on Ruby 3.2 have no fix at all.

What to do. Application security owner: validate. Confirm that your dependency scanner matches GHSA identifiers for the Ruby runtime and does not wait for CVEs. Upgrade internet-facing Ruby services to the fixed releases. List any remaining Ruby 3.2 and earlier workloads as unpatchable and plan their migration. Urgency: validate.

CSA coverage. When Discovery Outruns Remediation: Open-Source Disclosure Strain.

Most enterprise AI agents run on their builders' own OAuth grants, so audit logs record the employee rather than the agent

Category: security_operating_model **Significance:** notable **Confidence:** SELF-REPORTED (PROVIDER METRIC) for Nudge Security's telemetry figures: a median of seven agents in organisations that run any, about 44% running more than ten, agents spread across 15 building platforms, Anthropic integrations in 74% and OpenAI integrations in 67% of organisations, and 2,927 distinct AI vendors observed (the total number of organisations is not disclosed, and a comparable dataset from a party without a governance product would confirm the figures); CHARACTERIZATION (CSA) for the conclusion that agent actions cannot be separated from human actions in standard identity logs **Vendor interest:** Nudge Security sells SaaS and AI-usage discovery and OAuth governance, which is the remedy its report recommends. **Source:** Help Net Security – AI agents inherit employee access, and nobody's watching how they use it

What changed. On 19 November 2026 Nudge Security published aggregated customer telemetry on AI adoption. Among organisations running at least one AI agent, the median ran seven and about 44% ran more than ten, built on 15 different platforms such as work-management, automation and coding-agent products. The report says agents typically act through the OAuth grant or API key of the employee who built them, or through a shared connection, rather than through an identity of their own. Their actions are therefore logged as the employee's, and their access can outlast that person's change of role. OAuth-connected AI tools move data cloud to cloud, outside the corporate network and away from gateway and DLP inspection, and the grants stay live until someone revokes them.

Why it reaches you. The component class is the SaaS identity layer: the identity provider, OAuth consent and the audit trail of each SaaS application. An incident responder working from that trail cannot tell whether a human or an agent acted. Joiner-mover-leaver processes revoke the person, not the agents the person built. This is the governance side of what this feed reported from Palo Alto Networks' survey on 17 November, where 53% of engineers logged agent activity under their own identities. Two independent datasets now point the same way.

What to do. Identity governance owner: validate. Pull OAuth grants to AI applications from your identity provider and core SaaS tenants. For each agent-building platform in use, require agents to run on a dedicated, named non-human identity with an owner of record. Add a step to role-change and offboarding that transfers or revokes agents the person built. Urgency: validate.

CSA coverage. New – no prior CSA coverage.

IETF drafts for agent discovery bind a DNS key to a single endpoint but leave the trust of the catalog above it unsettled

Category: agentic_surface **Significance:** context **Confidence:** VERBATIM (PROVIDER) for the three Agentic Resource Discovery drafts (draft-bu-dawn-ard-architecture-00, draft-jimenez-dawn-ard-initial-contact-00, draft-bu-dawn-ard-discovery-00) announced on 3 November 2026 and discussed on the DAWN list on 18–19 November, and for the proposed interop test; CHARACTERIZATION (CSA) for reading the exchange as an open gap between catalog integrity and endpoint authenticity **Source:** IETF DAWN mailing list — New drafts: Agentic Resource Discovery (ARD) - architecture, initial contact, and discovery

What changed. Three Internet-Drafts for Agentic Resource Discovery (ARD) are under discussion in the IETF's DAWN group. ARD is a protocol through which an AI client asks a discovery service what agents, MCP servers, skills, plugins or APIs exist for a task. In the 19 November thread, a contributor to the related AID DNS approach agreed that the changing catalog should stay out of DNS. Under AID v2 the DNS-published key covers only one live endpoint, which a client can check against the name it asked for. He proposed an interop test that deliberately includes a stale catalog entry, to see how clients report it. Nothing in the thread describes a security model for the catalog itself. Separately, the ARD project site says there will be many discovery services, each applying its own trust and ranking policy.

Why it reaches you. The component class is the agent's tool-discovery path, the step before an agent picks which MCP server or skill to call. Today that choice mostly comes from configuration files and marketplaces, where this feed and CSA have documented poisoning. Standardised discovery will move the choice to services your agents query at runtime. Whether a client can tell a stale or substituted catalog entry from a legitimate one is being settled now, not after deployment.

What to do. Agent platform architect: monitor. If you evaluate or build agent discovery, require that any catalog entry be verifiable against an endpoint-bound key before invocation, and that clients surface stale entries rather than silently using them. Until then, restrict agents to an allowlisted, organisation-run catalog. Urgency: monitor.

CSA coverage. Poisoned Skills: AI Agent Marketplace Supply Chain Attacks.

Watchlist

- **OpenAI reward-hacking postmortem — downstream response** — No change. *(opened 2026-08-27)*

- **VM/hypervisor containment hardening for cyber-capable agents** – No change. *(opened 2026-08-27)*
- **Claude Code Auto Mode prompt-injection ASR discrepancy** – No change. *(opened 2026-08-27)*
- **AI defensive-triage guardrail evasion** – No change. *(opened 2026-08-31)*
- **AI account session hijacking at scale** – No change. *(opened 2026-08-31)*

Opened this issue

- **Self-hosted model-gateway patch lag** – Watching for GTIG's primary report on LiteLLM exploitation, a definitive fixed version for CVE-2026-40217, exposure counts for internet-facing gateways, and further KEV listings for model-gateway or guardrail components. *(opened 2026-11-19)*
- **nginx HTTP/2 distribution patch gap** – Watching for Ubuntu and Debian fixed packages for CVE-2026-123007 and CVE-2026-123633, the interval from upstream release to distribution release, and any first report of in-the-wild exploitation. *(opened 2026-11-19)*