Source dossier

The pin held, the checkout moved

Edition 2026-09-21 · narrative: supply chain controls that state intent without verifying it · compiled 2026-09-21

Claims index

Claim	Statement	Status	Sources
CLAIM-2026-09-21-01	AIR Security published the Plugin4Shell research on 17 September 2026; the researchers are Or Nevo, Dor Granat and Niv Hoffman.	verified	airsecurity2026pluginshe, theregister2026ai
CLAIM-2026-09-21-02	Claude Code, OpenAI Codex, GitHub Copilot and Gemini CLI all install marketplace plugins by pinned commit SHA, and none of the four verified that the checkout landed on the pinned commit.	verified	airsecurity2026pluginshe, helpnetsecu2026zeroclick
CLAIM-2026-09-21-03	Git resolves a name as a ref before treating it as an object id, so a branch named with the 40-hex pinned SHA is checked out instead of the pinned commit, with only a "refname is ambiguous" warning.	verified	airsecurity2026pluginshe
CLAIM-2026-09-21-04	The attack is zero-click because marketplaces bump their pinned SHA and agents re-run the checkout in the background without user action.	verified	airsecurity2026pluginshe, helpnetsecu2026zeroclick
CLAIM-2026-09-21-05	Claude Code is fixed in version 2.1.179 and OpenAI Codex in version 0.146.0.	verified	airsecurity2026pluginshe, helpnetsecu2026zeroclick
CLAIM-2026-09-21-06	GitHub Copilot has shipped no fix, and Google deprecated Gemini CLI rather than patch it, directing users to Antigravity CLI.	verified	helpnetsecu2026zeroclick, theregister2026ai
CLAIM-2026-09-21-07	AIR Security discovered the vulnerability in May 2026 with a working proof of concept and disclosed it to all four vendors in June 2026.	verified	airsecurity2026pluginshe, helpnetsecu2026zeroclick
CLAIM-2026-09-21-08	GitHub rejects a 40-hex branch name outright, while Bitbucket and self-hosted git servers allow it, and Anthropic's documentation lists Bitbucket and self-hosted git as valid marketplace backends.	verified	airsecurity2026pluginshe, theregister2026ai
CLAIM-2026-09-21-09	The researchers' recommended check is to compare the commit actually in the working tree against the pinned SHA and abort on mismatch, using git rev-parse HEAD.	verified	airsecurity2026pluginshe
CLAIM-2026-09-21-10	Germany's NIS2 implementation act (NIS2UmsuCG), which amends the BSIG, entered into force on 6 December 2025 and covers approximately 29,500 entities.	verified	nisdirectiv2026germany, secjur2026nis, covingtonbu2026germany

Claim	Statement	Status	Sources
CLAIM-2026-09-21-11	Section 30 BSIG lists supply chain security, including security-related aspects of relationships with direct suppliers, among the risk-management measures an in-scope entity must implement.	verified	covingtonbu2026germany
CLAIM-2026-09-21-12	Section 38 BSIG obliges management bodies to implement and monitor those measures, requires them to participate regularly in training, and makes them liable for damages caused through fault.	verified	covingtonbu2026germany, secjur2026nis

Only claims marked **verified** may carry a post. **single-source** may appear as hedged context; **disputed** must not be published at all.

Plugin4Shell — Zero Click RCE Vulnerability found in top 4 most popular coding agents

PRIMARY

Publisher	AIR Security
Published	2026-09-17
Retrieved	2026-09-21
Cite key	airsecurity2026pluginshe
URL	https://www.air.security/blog-posts/plugin4shell

Assessment. The researchers' own write-up and the only source that states the git mechanism precisely. Primary for the vulnerability, the disclosure timeline and the recommended rev-parse check. Vendor patch status here is the researchers' verification, not a vendor statement.

Plugin4Shell

AIR Security, 17 September 2026. Or Nevo, Dor Granat, Niv Hoffman.

The mechanism

Coding agents install marketplace plugins by cloning the plugin repository and checking out a pinned commit SHA. The pin is the supply-chain control: the marketplace reviews a commit, records its hash, and the agent is supposed to get exactly that code.

"when a name is both a valid ref and an object id, git prefers the ref and only prints a `refname is ambiguous` warning - so it never matters whether the pinned commit is still present, the branch is what gets checked out."

An attacker who controls the plugin repository creates a branch whose name is the 40-hex pinned SHA. The sequence the agents run —

git clone <repo> git checkout aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

— resolves to the attacker's branch rather than the pinned object.

"every one of the four agents checks out the pinned commit without verifying the checkout landed there, letting an attacker swap in malicious code while the pin still looks intact."

Why it is zero-click

Marketplaces bump their pinned SHA over time. Agents re-run the checkout in the background when that happens, so the malicious tree arrives without any user action on Claude Code and Codex.

Two attack vectors

Publish a functional plugin and turn it malicious after review; or hijack an existing trusted repository and reach every installed instance.

Which hosts allow the branch name

GitHub rejects a 40-hex branch name outright. Some hosts forbid it, "others, Bitbucket among them, and any self-hosted git server, allow it." Marketplaces on those hosts are a supported configuration — Anthropic's own documentation lists Bitbucket and self-hosted git as valid marketplace backends.

The fix

Resolve the commit actually in the working tree and abort unless it equals the pinned SHA:

test "$(git rev-parse HEAD)" = "<pinned-sha>"

Status

Discovered May 2026 with a working proof of concept against all four agents. Coordinated disclosure to all vendors June 2026. Anthropic confirmed the Claude Code fix on 17 June 2026 (2.1.179). Codex 0.146.0 verified fixed 12 August 2026. Google confirmed on 4 August 2026 that Gemini CLI would not be fixed (deprecated). GitHub Copilot: no fix shipped.

Updating is the only complete mitigation where one exists. Gemini CLI users should migrate to Antigravity.

Zero-click RCE vulnerability hit four major AI coding agents, two remain unpatched

SECONDARY

Publisher	Help Net Security
Published	2026-09-18
Retrieved	2026-09-21
Cite key	helpnetsecu2026zeroclick
URL	https://www.helpnetsecurity.com/2026/09/18/plugin4shell-ai-coding-agents-vulnerability/

Assessment. Trade coverage one day after the research post. Used to corroborate patch status and the framing of which two agents remain exposed. Coverage date is 18 September; the event date is the researchers' 17 September publication.

Help Net Security, 18 September 2026.

GitHub Copilot and Gemini CLI remain without patches. Microsoft has not shipped a fix for Copilot; Google deprecated Gemini CLI rather than patching it.

Patched: Claude Code 2.1.179 (Anthropic); Codex 0.146.0 (OpenAI).

No CVE identifiers are given.

AIR researchers found the bug in May 2026 and disclosed it to vendors the following month, with working proof-of-concept exploits against all four agents.

The flaw bypasses SHA pinning — the mechanism that locks plugins to specific reviewed code versions. As the researchers put it: "every one of the four agents checks out the pinned commit without verifying the checkout landed there, letting an attacker swap in malicious code while the pin still looks intact."

It is zero-click because background auto-updates re-run the vulnerable checkout without user action on Claude Code and Codex.

Attackers can either publish functional plugins that become malicious after review, or hijack existing trusted repositories to distribute compromised code to all installed instances.

AI coding agents' 0-click RCE flaw could hand attackers keys to the kingdom

SECONDARY

Publisher	The Register
Published	2026-09-17
Retrieved	2026-09-21
Cite key	theregister2026ai
URL	https://www.theregister.com/security/2026/09/17/ai_coding_agents_0-click_rce_flaw_could_hand_attackers_keys_to_the_kingdom/

Assessment. Same-day trade coverage. Adds that GitHub implemented SHA-name restrictions at the forge and that Google points Gemini CLI users at Antigravity CLI. Used only to corroborate details the primary source also carries.

The Register, 17 September 2026.

Disclosed by Air, a security startup specialising in enterprise AI agent protection. Reported to vendors June 2026.

"The agent checks out the exact commit the marketplace pinned but never verifies it landed there."

Affected: Anthropic Claude Code, OpenAI Codex, Google Gemini CLI, GitHub Copilot.

Anthropic: Claude Code 2.1.179 patched. OpenAI: Codex 0.146.0 patched. Google: deprecated Gemini CLI, directing users to Antigravity CLI. Microsoft: no patch for Copilot; GitHub implemented SHA-name restrictions.

Two exploitation routes: submit a benign plugin that later receives a malicious update, or hijack a legitimate repository and push a compromised version to all installed instances.

Germany Transposes NIS 2 Directive — Increased Cybersecurity Requirements for Businesses

SECONDARY

Publisher	Covington & Burling
Published	2026-01
Retrieved	2026-09-21
Cite key	covingtonbu2026germany
URL	https://www.globalpolicywatch.com/2026/01/germany-transposes-nis-2-directive-increased-cybersecurity-requirements-for-businesses/

Assessment. Law-firm analysis of the German transposition. Used for the BSIG section numbering and the substance of the supply-chain and management-body duties. Its date for the act ('5 December 2025') reads as adoption/promulgation; two other sources give 6 December 2025 as entry into force, which is the date used.

Covington and Burling, Global Policy Watch, January 2026.

Act: Gesetz zur Umsetzung der NIS-2-Richtlinie und zur Regelung wesentlicher Grundzuege des Informationssicherheitsmanagements in der Bundesverwaltung (NIS2UmsG / NIS2UmsuCG). Adopted and binding 5 December 2025.

Amends the BSIG (Gesetz ueber das Bundesamt fuer Sicherheit in der Informationstechnik und ueber die Sicherheit in der Informationstechnik von Einrichtungen).

Scope: approximately 29,500 companies.

Registration: no later than three months after first or again qualifying (Section 33 new BSIG).

Section 30 new BSIG requires establishments to implement risk-management measures covering policies, incident handling, business continuity, and "supply chain security, including security-related aspects concerning relationships between each establishment and its direct suppliers".

Section 38 new BSIG imposes implementation and monitoring obligations on management bodies, requires them to "regularly participate in training courses", and makes them liable for damages caused through fault in accordance with corporate law.

Germany — NIS 2 Directive transposition status

SECONDARY

Publisher	NIS 2 Directive transposition tracker
Published	2026-08-06
Retrieved	2026-09-21
Cite key	nisdirectiv2026germany
URL	https://www.nis-2-directive.com/Transposition/Germany.html

Assessment. Transposition tracker. Used for the entry-into-force date (6 December 2025) and the supervisory authority. Corroborated on the date by secjur.

NIS 2 Directive transposition tracker, Germany.

The NIS 2 implementation in Germany became effective on 6 December 2025.

As of 6 August 2026 Germany has "completed the principal transposition of NIS 2"; focus has shifted to "registration, operational compliance, management accountability and coordination."

The Federal Office for Information Security (BSI) is the primary supervisory authority and remains responsible for cybersecurity governance at federal level.

NIS2 Haftung Geschaeftsfuehrer: Paragraph 38 Pflichten und Enthaftung 2026

SECONDARY

Publisher	secjur
Published	2026
Retrieved	2026-09-21
Cite key	secjur2026nis
URL	https://www.secjur.com/blog/nis2-haftung-geschaeftsfuehrer

Assessment. German compliance vendor's legal explainer. Used only as the second source for the 6 December 2025 entry-into-force date and for the three Section 38 duties. Its reading that liability cannot be waived in advance was not used in the edition.

secjur, NIS2 Haftung Geschaeftsfuehrer.

"seit dem 6. Dezember 2025 in Kraft" — the NIS2UmsuCG entered into force on 6 December 2025.

Section 38 BSIG imposes three non-delegable duties on management:

Billigung — management must actively approve the risk-management measures; informal authorisation is insufficient.

Ueberwachung — management must actively control implementation and receive regular status reports.

Schulungen — Section 38(3) requires management to participate regularly in cybersecurity training sufficient to recognise risks and assess their operational impact.