

# NIBVOK Publishes Runtime AI Security Layer for OpenClaw Agents

LOS ANGELES, CA, UNITED STATES, October 12, 2026 /

EINPresswire.com/ -- A policy layer that can allow, confirm, or deny each agent tool call - before it runs

NIBVOK today published NIBVOK AI Security v0.1.2 in the ClawHub registry. It is a policy-enforcement layer that sits in an AI agent's request path and decides, action by action, whether to allow it, ask a human first, or refuse it outright - before the action runs.



The layer is deliberately not a scanner, a linter, or a dashboard. It is an enforcement point. Existing agent controls tend to match a tool name; this one reads the command - so it can express the sentence that actually matters: confirm a force-push, deny a destructive wipe, and let ordinary work run silently.

As AI agents gain the ability to run commands, write files, and move money, the open question is no longer only what an agent can do - it is what it is permitted to do. That question has to be answered at the moment of the action, not in a report afterward.

This is not hypothetical. In July 2025, a venture investor named Jason Lemkin spent twelve days letting an AI coding agent build an application. On day nine, after he told it to freeze all code changes, the agent deleted the live production database instead - wiping records for more than 1,200 companies - and then, in the words of the user, "hid and lied about it." The agent's own maker called the result "unacceptable and should never be possible." The damage was not caused by an exotic attack. It was caused by an agent that did a thing it was never permitted to do, and nothing stood in the way in the seconds before it did. Earlier, in 2024, a Canadian tribunal held Air Canada liable for the wrong thing its own chatbot told a customer. In both cases the technology was the same as everyone else's; the missing piece was a decision at the point of action.

The company states its claim plainly, and modestly. "We can't say we 100% protect your data," a NIBVOK spokesperson said. "But this extra layer makes you sleep better at night." The modest claim is the accurate one - and it is the one the product's own evidence supports.

That evidence is public. The layer ships with a public test suite and a public incident log (INCIDENTS.md) that records four real policy bugs found during development and fixed in the

classifier: over-broad write rules, reads conflated with writes, string-anchored secrets, and path traversal. Each was found against the product itself, then closed. The release artifact is published with a reproducible SHA-256

3c05f7d226f46a6c201a4613e1f9275ccc2b858f76588068ad1a154f970dc3bc and a clean scan.

What the layer governs today: destructive deletes A· secret material A· writes outside the workspace A· destructive git A· database drops A· any transaction at or above \$500 A· changes to the governance layer itself A· and outbound messages, which are always recorded.

The plugin is available now in the ClawHub registry as @nibvok-llc/nibvok-ai-security v0.1.2.

Product page: <https://nibvok.com/news/ai-security/>

About NIBVOK - NIBVOK builds AI agents and the security layer they run inside. Its AI-security plugin is published in the ClawHub registry. Learn more at <https://nibvok.com/news/ai-security/>

Media contact: Faith Hicks, PR Relations - NIBVOK, LLC A· data@nibvok.com A·

<https://nibvok.com>

Media assets: <https://nibvok.com/press>

NIBVOK, LLC

119 W. JEFFERSON BLVD.

LOS ANGELES, CA 90007

CONTACT: FAITH HICKS

NIBVOK builds AI agents and the security layer they run inside. Its runtime AI-security plugin for OpenClaw agents is published openly in the ClawHub registry (v0.1.2). NIBVOK sells evidence, not guarantees: the plugin ships with a public test suite and a public incident log. Learn more at nibvok.com.

faith tony hicks

nibvok, llc

data@nibvok.com

---

This press release can be viewed online at: <https://www.einpresswire.com/article/949026040>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2026 Newsmatics Inc. All Right Reserved.