

## Forum:[Hack Forums](#)

Thread: [CVE-2026-82533: DeepSeek Harness Vulnerability Lets AI Agents Escape Its Own Sandbox](#)

**MadScientist** - September 10, 2026, 6:13 am

OX Research found and disclosed a critical vulnerability in DeepSeek Harness, DeepSeek's open-source AI coding-agent harness, that allowed a sandboxed AI agent to disable its own confinement with a single shell command – on shipped defaults, with no network exposure and no credentials.

Vulnerability Details

CVE: CVE-2026-82533

CWE: CWE-807 Reliance on Untrusted Inputs in a Security Decision

CVSS: 9.4

CVSS Vector: CVSS:4.0/AV:N/AC:L/AT:N/PR:N/UI:P/VC:H/VI:H/VA:H/SC:H/SI:H/SA:H

Description: DeepSeek Harness exposed its agent-control API on a local HTTP port without authentication, relying solely on the client-supplied 'Host' request header to determine whether a request was trusted rather than verifying the connection's actual peer address. Because the product's OS sandbox restricted file writes but left loopback networking open, a sandboxed agent could use a single shell command to call that API and elevate its own session to 'danger-full-access' with approval prompts disabled — effectively disabling its own sandbox on the shipped default configuration. If the port was exposed to a network, the same interface also allowed an unauthenticated remote attacker to take full control of the agent and export every stored conversation without a key.

Read More -

<https://www.ox.security/blog/cve-2026-82533-deepseek-harness-ai-agent-sandbox-escape/>

**XtremeDiscount** - September 10, 2026, 4:02 pm

Oh wow, thanks for the read. That's cool you can make it escape the sandbox