

AI Spy — DEV-12

Forensic view · generated 2026-09-11 19:42 UTC · AIRES v2

DEV-12

76.2 Critical

IP DEV-12 · standard_workstation · p_risk 0.89 · confidence medium ±9.2

What was detected

- An autonomous AI agent with action capability (shell / filesystem) is configured.
- Unsanctioned / concealed AI usage was detected (governance risk).

How it contributes to the score

Dimension	Score	Meaning
Capability	84.3	local AI compute power
Egress	40.8	ways data can leave
Data Exposure	0.0	sensitive data near AI
Credential	0.0	exposed AI/API secrets
Governance	75.9	unsanctioned / evasive use

Threat scenarios (coupling ×1.396)

- **Unrestrained Agent (permissions bypassed + shell execution)** (×1.396) — An AI host with permissions bypassed AND a hook/skill that runs shell: every tool call can execute arbitrary commands with no human gate.

Top remediation: Removing the filesystem (MCP server) drops this endpoint from 76 to 57.

Forensic detail

Signals 5 → 5 exposures after entity resolution · environment ×1.0

Resolved exposures (confidence-weighted)

Exposure	Kinds	Conf.	Contribution
filesystem (MCP server)	mcp_server	0.88	0.669
claude code:PostToolUse:* (ai hook)	ai_hook	0.88	0.572
claude code (permission bypass)	permission_bypass	0.88	0.572
Claude Code (LLM client)	llm_client	0.85	0.383
deploy (ai skill)	ai_skill	0.5	0.25

Tools

Name	Category	Evidence	Path
			C:\Users\amart\claude

MCP servers

Name	Capabilities	Risk
filesystem	filesystem	high

AI Exposure Register

Raw AIRES **76.2** (Critical) → Adjusted AIRES **76.2** (Critical)

Findings (lifecycle)

Finding	Category	Status	Owner	First seen	Last seen	Seen
mcp server — filesystem	MCP	Open		2026-09-11	2026-09-11	x1
ai hook — claude code:PostToolUse:*	hook	Open		2026-09-11	2026-09-11	x1
permission bypass — claude code	perm	Open		2026-09-11	2026-09-11	x1
ai skill — deploy	skill	Open		2026-09-11	2026-09-11	x1
llm client — Claude Code	Tool	Open		2026-09-11	2026-09-11	x1