ISSUE #002 · AGENT & BOT IDENTITY

The identity acting is not the identity you think.

A Claude Code skill riding authenticated Slack sessions, a Dependabot PR auto-merged 5 minutes after malicious publish, and the AWS pattern that scopes an MCP agent per tool call.

> ATTACK

Mitiga · Idan Cohen + Yael Ben Yair · 2026-04-15 · **License to Skill, Pt. 2: Slack Compromise Through Claude Code**

An attacker publishes a Claude Code skill to npx with a benign-looking name. A developer runs it once. The skill invokes at the developer's request, but uses their authenticated Slack session, already in their browser, already trusted by the workspace, to send phishing DMs to coworkers and customers. Messages inherit the developer's identity, tone, and reach.

```
# the trust model in one line agent.run()  →  every SaaS token the user's browser holds
```

Most SaaS trust models assume a human is in the loop for each action. A skill that runs silently breaks that assumption at scale.

▸ MONDAY CHECK

Audit the Claude Code skills your team has pulled in the last 30 days. Cross-reference skill authors against an internal allowlist. If you don't have one yet, pinning skills to reviewed authors is the cheapest control in the queue.

> RULE

GitGuardian · Gaetan Ferry · 2026-04-10 · **Renovate & Dependabot: The New Malware Delivery System**

Alert on any Dependabot or Renovate PR that touches a production-path dependency AND was auto-merged within 6 hours of the upstream version publish time. Six hours sits below human-review speed. 895+ public repos auto-merged the malicious axios within 5 minutes of upstream publish. Match against your CI dependency graph, not just your GitHub org, to catch indirect pipelines.

> DEFENDER

AWS Security Blog · Riggs Goodman III · 2026-04-14 · **Secure AI Agent Access Patterns via MCP**

Scope MCP agents with per-tool STS session policies. The agent's execution role holds broad permissions; each tool invocation narrows the session via `AssumeRole` with a session policy matching only that tool's required actions. Session tags let you differentiate AI vs human actors for logging and SCPs. First named AWS pattern that treats the agent as a subject, not a service account.

> RADAR

Datadog Security Labs · Kennedy Toomey · 2026-04-16 · **The case for dependency cooldowns in a post-axios world**

A 12-hour cooldown blocks axios and s1ngularity both. A week is what most teams set. In Renovate or Dependabot, one config line. A near-zero-cost control that isn't on by default in most setups.

SOURCES

Mitiga / mitiga.io/blog/007-license-to-skill-p-2-slack-compromise-through-claude-code
GitGuardian / blog.gitguardian.com/renovate-dependabot-the-new-malware-delivery-system
AWS / aws.amazon.com/blogs/security/secure-ai-agent-access-patterns-to-aws-resources-using-model-context-protocol
Datadog / securitylabs.datadoghq.com/articles/dependency-cooldowns