Hive Pro

Threat Level

Amber

HiveForce Labs

THREAT ADVISORY

ATTACK REPORT

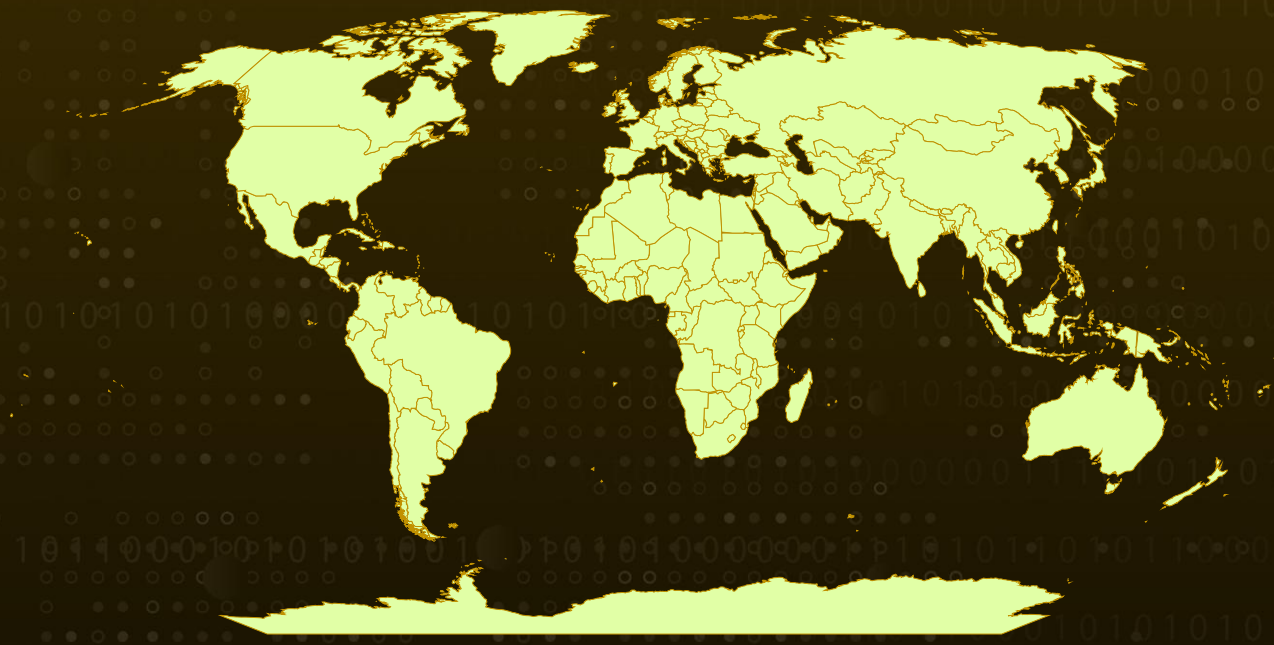
When Your AI Agent Hands You the Malware: Inside FakeGit's AgentBaiting Campaign

Date of Publication	Admiralty Code	TA Number
July 23, 2026	A1	TA2026209

Summary

First Seen: March 2026
Targeted Regions: Worldwide
Targeted Platforms: Windows
Targeted Industries: Cross-industry and opportunistic; developers and enterprise users of AI Skills and MCP tooling
Malware: SmartLoader, StealC
Campaign: FakeGit
Attack: FakeGit is a large-scale operation of roughly 7,600 malicious GitHub repositories, created by about 6,600 profiles, that disguise SmartLoader malware as ordinary software downloads. More than 800 of these repositories pose as AI Skills or MCP servers and have surfaced over 600 times across public AI registries, giving them a false sense of legitimacy. Victims who download and run the ZIP "installer" trigger a LuaJIT loader chain that deploys SmartLoader, which then establishes persistence and delivers the StealC information stealer. The campaign's most notable evolution, dubbed AgentBaiting, sees AI agents independently discover these repositories during routine tasks and relay the attacker's installation instructions to their users.

Attack Regions



Targeted Non-Targeted

Attack Details

#1

FakeGit works by making malware look like a routine software install, and it is not new. It is assessed as a continuation of an older operation that pushed Lumma Stealer, tracked by Trend Micro under the threat actor name Water Kurita. Operators copy or fabricate GitHub projects and host them under lookalike developer identities, often just one character off a real handle, then dress them up with fake star and fork counts, copied descriptions, and polished READMEs. Many impersonate familiar tools like Gmail, WhatsApp, Databricks, Jenkins, and Docker, and more than 800 pose as AI Skills or MCP servers, amplified by over 600 listings across public registries such as LobeHub, Glama, MCP.so, and MCP Market. The key evolution is AgentBaiting: an AI agent hunting for a capability can find one of these repositories on its own, treat the attacker's README as legitimate documentation, and pass the install steps to its user. In Island's testing, Claude Code, Gemini, and ChatGPT all surfaced malicious repositories without being shown a link.

#2

The attack starts the moment the victim opens the downloaded ZIP. A representative package holds no real software, just three files: a small launcher script (for example, application.cmd), a renamed LuaJIT-style runtime (luau.exe), and an obfuscated Lua payload disguised as a text or icon file (ico64.txt). The launcher passes the payload to the runtime, executing a roughly 300 KB single-line Lua program. Names vary across the campaign, but the structure is constant: a .cmd or .bat launcher, a LuaJIT-style runtime, and a payload masquerading as a text, icon, license, or data file. This loader chain drops SmartLoader, which hides its console window and persists through scheduled tasks under %LOCALAPPDATA%.

#3

Instead of moving laterally, the campaign focuses on the endpoint and quietly stages its next payload. SmartLoader pulls its current command-and-control address from a value stored in a Polygon smart contract. This blockchain dead drop makes takedown harder and then retrieves encrypted stages from GitHub. Those stages lead to a PE crypter that injects the StealC information stealer into another process. Because the host is already compromised at execution, the fake "integration" never needs the access it advertised.

#4

StealC handles the theft once it is running in memory. It targets browser passwords, cookies, active sessions, extension data, email and remote-access credentials, screenshots, and host details, then exfiltrates them over its C2 channel. Crucially for responders, StealC steals live sessions, not just stored passwords, so resetting credentials is not enough. Active browser sessions, OAuth grants, API tokens, and cloud and developer credentials must all be revoked to fully close out an intrusion.

Recommendations

- **Build a Curated Capability Catalog:** Maintain an approved, internally reviewed catalog of Skills, MCP servers, and agent plugins so users and agents pull capabilities from a trusted source instead of open, unvetted discovery, which is exactly the behavior this campaign depends on.
- **Sandbox Every New Capability Before Rollout:** Evaluate any new Skill or MCP server in an isolated environment that has no browser sessions, cloud credentials, SSH keys, or production data before allowing it into wider use.
- **Reject Executable-Delivered Capabilities on Sight:** Treat any Skill or MCP server that installs through a downloaded executable or ZIP launcher as malicious. A legitimate capability is source code with a manifest, not a Windows ZIP holding a launcher script, a renamed runtime, and a payload disguised as a text file.
- **Verify the Publisher, Not Just the Project:** Independently confirm developer identities before trusting a repository. This campaign relies on one-character-off usernames, copied profiles, and modest star counts, so treat star and fork counts as unreliable signals.
- **Treat Registry Listings as No Signal of Trust:** Do not assume a Skill or MCP server is safe because it appears on LobeHub, Glama, MCP.so, MCP Market, or similar catalogs. These registries reproduced attacker READMEs and download links, adding a false layer of credibility.
- **Monitor Agent-Initiated Activity:** Apply the same scrutiny to downloads, git clones, and shell commands originating from AI agents as you would to user-initiated browser downloads. AgentBaiting means an agent completing a routine task can retrieve and stage malware on its own.
- **Maintain an AI Capability Inventory with Hashes:** Keep an inventory of every AI capability in use, including source repository, commit, version, and package hash, so newly disclosed campaign repositories can be matched against your environment quickly.

Potential MITRE ATT&CK TTPs

Tactic	Technique	Sub-technique
Resource Development	T1585: Establish Accounts	
	T1587: Develop Capabilities	T1587.001: Malware
	T1608: Stage Capabilities	T1608.001: Upload Malware
Initial Access	T1195: Supply Chain Compromise	
Execution	T1204: User Execution	T1204.002: Malicious File
	T1059: Command and Scripting Interpreter	T1059.003: Windows Command Shell
Defense Evasion	T1036: Masquerading	T1036.005: Match Legitimate Name or Location
	T1027: Obfuscated Files or Information	
	T1140: Deobfuscate/Decode Files or Information	
	T1564: Hide Artifacts	T1564.003: Hidden Window
	T1055: Process Injection	
Persistence	T1053: Scheduled Task/Job	T1053.005: Scheduled Task

Tactic	Technique	Sub-technique
Command and Control	T1102: Web Service	T1102.001: Dead Drop Resolver
	T1105: Ingress Tool Transfer	
Credential Access	T1555: Credentials from Password Stores	T1555.003: Credentials from Web Browsers
	T1539: Steal Web Session Cookie	
Collection	T1113: Screen Capture	
	T1005: Data from Local System	
Exfiltration	T1041: Exfiltration Over C2 Channel	

Indicators of Compromise (IOCs)

TYPE	VALUE
SHA256	216a2c99fd42c00f9323d8b16dd19f622f7f4778b2b1d7cf07a3de5621fd1546, 91e5dbfaf45edf25fbc2168f92083e05dfa427afa7633e991392e33cc7427dad, 498fe8fb806cd0e6685f97fc7d74de769dae5a28cdc821557b7585ad5ad83147, 62744baa8077bb8be237647fd78e3bea2ca0932bf4be3d5618600f97118095f8, 1da8df487d30b988f3c350c065206726aaa13f079a07151cd42ab5579994b9de, c15693106682f2ddb26649cab6e1962a64537627cde4c5d3c79d5a0be8c1b5a8, 66afc7d87d10dbe392898c4e5c613e0442fabb396415c2bef3a5ef2ac752c5ad, a33f40cab1ab7f971d3464af3e7595918107332b9e83342007571842b9e22826, 3c858facbad66f5479e2c4add171421dc1b6488b36f33e7cff073aba585954a7, fc1278f419e611bf40ca414099bfd9ad98a31ffb054371e8cb65a84849b00eaf

TYPE	VALUE
Filenames	yu-ai-agent-1.0-beta.3.zip, awesome-skills-claude-3.3.zip, agent_ai_awesome_skills_2.0.zip, for-security-mcp-3.3.zip, plugins_claude_awesome_code_2.4.zip, mcp-walmart-2.2.zip, server_databricks_mcp_1.6.zip, mcp-server-jenkins-3.2.zip, gateway-docker-mcp-v1.6-alpha.5.zip, alibabacloud-skills-bigdata-v1.7.zip
GitHub Repository	hfgwygey/yu-ai-agent, Mann1988/awesome-claude-skills, h4vzz/awesome-ai-agent-skills, StanLeyJ03/mcp-for-security, xbim08/awesome-claude-code-plugins, DomingosNgongo/walmart-mcp, 45d5r/databricks-mcp-server, MauManto/jenkins-mcp-server, waynestimulative605/docker-mcp-gateway, lucaducapuca/alibabacloud-bigdata-skills, Naveenkm007/spaceship-mcp, hahaha-saygex/gmail-mcp, adlaiponderous700/claude-skill-cinematic-prompt

References

https://www.island.io/blog/agentbaiting-how-800-fake-ai-skills-and-mcp-servers-delivered-malware

What Next?

At **Hive Pro**, it is our mission to detect the most likely threats to your organization and to help you prevent them from happening.

Book a Demo of **HivePro**.

REPORT GENERATED ON

July 23, 2026 • 08:00 AM

More at www.hivepro.com