

FREE TUTORIAL · SECURITY

Cara **Aman Install Skill** GitHub ke Claude Code Pakai NVIDIA SkillSpector

Install repo atau skill GitHub sembarangan ke Claude Code bisa bahaya. Pelajari cara pasang security gate otomatis dari NVIDIA yang nge-scan tiap repo sebelum masuk — cukup 3 langkah utama.

 Waktu setup: ~20 menit

FREE

Gratis 100%



Claude Code + SkillSpector

KENAPA INI PENTING

Install Skill GitHub Sembarangan Itu Nyata Bahayanya

Claude Code dan Codex punya ekosistem skill yang luas — ribuan repo GitHub bisa langsung lo install dan jalankan sebagai agent. Masalahnya, nggak semua repo itu aman. Ada yang sengaja dibuat untuk menyisipkan perintah tersembunyi (prompt injection), ada yang diam-diam mencuri kredensial, dan ada yang membajak supply chain dependency lo.

Bedanya dengan malware biasa: ancaman di sini nyasar langsung ke AI agent lo. Sebuah SKILL.md yang terlihat normal bisa berisi instruksi tersembunyi di HTML comment yang nyuruh Claude kirim API key lo ke server luar — dan Claude nurutin aja karena dia "baca" file itu sebagai instruksi.

Tiga ancaman paling umum di skill/repo GitHub untuk AI agent:

- **Prompt Injection** — instruksi tersembunyi di file teks yang memanipulasi perilaku AI
- **Supply Chain Attack** — dependency dengan CVE atau yang fetch kode dari URL berbahaya
- **Malware** — script yang baca ~/.ssh, env vars, atau buka reverse shell

Dengan SkillSpector terpasang, kamu bisa:

- Install skill GitHub dengan tenang — ada layer scan otomatis sebelum eksekusi
- Lihat hasil scan detail: severity, kategori ancaman, dan file mana yang mencurigakan
- Atur sendiri mana yang di-block dan mana yang cukup warning, sesuai kebutuhan lo

TOOLS

Apa Itu NVIDIA SkillSpector?

SkillSpector adalah tool open-source buatan NVIDIA — bukan pihak ketiga sembarangan — yang dirancang khusus untuk men-scan skill, plugin, dan repo yang akan digunakan oleh AI agent. Dibuat mengikuti panduan keamanan internal NVIDIA untuk semua penggunaan AI dan agent.

Cara kerjanya: lo pasang SkillSpector sebagai **PreToolUse hook** di Claude Code. Artinya setiap kali Claude mau menjalankan perintah yang melibatkan install atau clone repo dari luar, hook ini berjalan dulu, scan targetnya, baru memutuskan boleh lanjut atau diblok.

Repo resmi: <https://github.com/NVIDIA/SkillSpector>

Gratis, open-source, dikembangkan langsung oleh tim AI security NVIDIA.

3 Kategori Fokus Scanning

SkillSpector defaultnya sensitif banget dan sering false positive. Supaya workflow lo nggak lambat, fokuskan scan ke 3 kategori ini saja:



Prompt Injection

Hidden instruction, zero-width char, HTML comment berisi perintah AI



Supply Chain Attack

Typosquatting, dependency dengan CVE, remote code fetch dari URL tak dikenal



Malware

YARA match, exfiltration kredensial/env var, reverse shell, persistence cron

TUTORIAL

3 Langkah Setup SkillSpector sebagai Security Gate

1

Temukan dan Install SkillSpector sebagai CLI Global

Buka terminal, pastikan kamu punya `Python >= 3.12` dan `uv` terinstall. Kalau belum ada, install `uv` dulu dengan: `curl -LsSf`

`https://astral.sh/uv/install.sh | sh`

Setelah itu install SkillSpector langsung dari GitHub:

```
uv tool install git+https://github.com/NVIDIA/SkillSpector
```

Verifikasi instalasi berhasil dengan menjalankan:

```
skillspector --help
```

 dan

```
skillspector scan --help
```

Baca output-nya dan catat flag yang tersedia — jangan nebak flag yang tidak ada.

💡 Test dulu ke satu folder skill yang sudah ada di `~/.claude/skills` buat mastiin CLI jalan sebelum pasang hook.

2

Gunakan Prompt Setup untuk Pasang Hook Otomatis di Claude Code

Ini langkah utama yang paling sering dilewat orang. Lo nggak perlu setup manual — cukup kasih prompt setup di bawah ini ke Claude Code, dan dia akan otomatis:

- Bikin script `~/.skillspector-gate/gate.sh` yang nge-scan repo sebelum dieksekusi
- Daftarin hook ke `~/.claude/settings.json` sebagai PreToolUse Bash
- Bikin file `~/.skillspector-gate/policy.json` yang fokus ke 3 kategori ancaman
- Setup logging di `~/.skillspector-gate/logs/`
- Test dengan 2 folder dummy: satu bersih, satu "jahat" palsu

💡 Backup `settings.json` kamu sebelum menjalankan prompt ini — Claude akan otomatis melakukan backup tapi lebih aman kalau kamu backup manual juga.

3

Konfigurasi False Positive: Fokus ke 3 Kategori, Bukan Semua

SkillSpector defaultnya blok hampir segalanya. Kalau dibiarkan, setiap perintah `npm install` atau `pip install` bakal kena scan dan memperlambat kerja lo secara signifikan.

Yang benar: set `policy.json` buat blok HANYA finding dengan severity HIGH/CRITICAL di 3 kategori di atas. Finding LOW/MEDIUM atau kategori lain → cukup print warning, jangan blok.

Tambahkan **allowlist** untuk repo yang lo percaya (contoh: repo resmi Anthropic, NVIDIA, tools yang lo sendiri bikin) supaya nggak ter-scan ulang tiap kali.

💡 Hook hanya trigger untuk perintah yang benar-benar ambil kode eksternal: git clone, pip install dari git+https, download ke folder skills. Perintah lain seperti cd, ls, echo → langsung lolos tanpa scan.

PROMPT SETUP

Copy Prompt Ini dan Jalankan di Claude Code

Copy seluruh prompt di bawah dan paste langsung ke Claude Code. Claude akan menjalankan semua langkah setup otomatis — install, konfigurasi hook, setup policy, sampai test.

PROMPT SETUP SKILLSPECTOR

Lo adalah security engineer. Tugas lo: pasang NVIDIA SkillSpector sebagai "security gate" otomatis di environment gue, supaya setiap kali gue (atau agent) install skill / clone repo GitHub baru, repo itu di-scan DULU sebelum dipakai.

KONTEKS

- Repo: <https://github.com/NVIDIA/SkillSpector>
- Target: Claude Code (wajib) + Codex CLI (kalau terinstall)
- OS gue: deteksi sendiri (macOS/Linux/WSL)

LANGKAH 1 – INSTALL

1. Cek Python ≥ 3.12 dan uv. Kalau belum ada, kasih tau gue dulu.
2. Install SkillSpector sebagai CLI global (pakai uv tool install dari repo

VERIFIKASI

Cara Test: Apakah Security Gate Berhasil Terpasang?

Setelah Claude selesai setup, ada dua cara verifikasi cepat. Pertama, cek apakah hook sudah terdaftar di settings Claude Code:

JALANKAN DI TERMINAL

```
cat ~/.claude/settings.json | grep -A5 "PreToolUse"
```

Output yang benar akan menampilkan entry hook yang mengarah ke `~/.skillspector-gate/gate.sh`. Kedua, coba jalankan clone repo kecil yang aman — perhatikan apakah muncul output scan SkillSpector sebelum clone dieksekusi.

TEST CLONE REPO (CLAUDE CODE)

```
"Clone repo ini untuk coba:  
https://github.com/NVIDIA/SkillSpector"
```

Kalau setup benar, Claude akan menampilkan hasil scan SkillSpector dulu sebelum memulai clone. Lo bisa lihat log lengkapnya di folder `~/.skillspector-gate/logs/`.

SETELAH SETUP

Yang Bisa Kamu Lakukan Sekarang

- ✓ Install skill atau repo GitHub baru ke Claude Code tanpa khawatir ada backdoor atau prompt injection tersembunyi
- ✓ Lihat history semua scan di folder logs — tahu persis repo mana yang pernah dicurigai dan kenapa
- ✓ Tambah repo terpercaya ke allowlist sendiri supaya nggak di-scan ulang setiap saat (hemat waktu)
- ✓ Atur policy.json sesuai kebutuhan — mau lebih ketat atau lebih longgar tergantung workflow lo
- ✓ Pasang hal yang sama ke Codex CLI kalau kamu pakai dua AI coding agent sekaligus
- ✓ Tetap cek sumber repo secara manual untuk repo yang belum dikenal — tools ini ngurangin risiko, bukan ngilangin risiko sepenuhnya

Ingat: SkillsSpector adalah lapisan tambahan, bukan pengganti judgement lo. Repo dari sumber yang lo kenal dan percaya tetap lebih aman daripada repo random walaupun lolos scan. Security berlapis selalu lebih baik.



Ikuti untuk Tutorial AI Tools Terbaru

Tutorial ini cuma satu dari banyak konten AI tools gratis yang saya share setiap minggu di Instagram.

Follow @rifqyyusuf di Instagram

Dibuat dengan Claude · Tutorial AI Tools untuk Kreator Indonesia