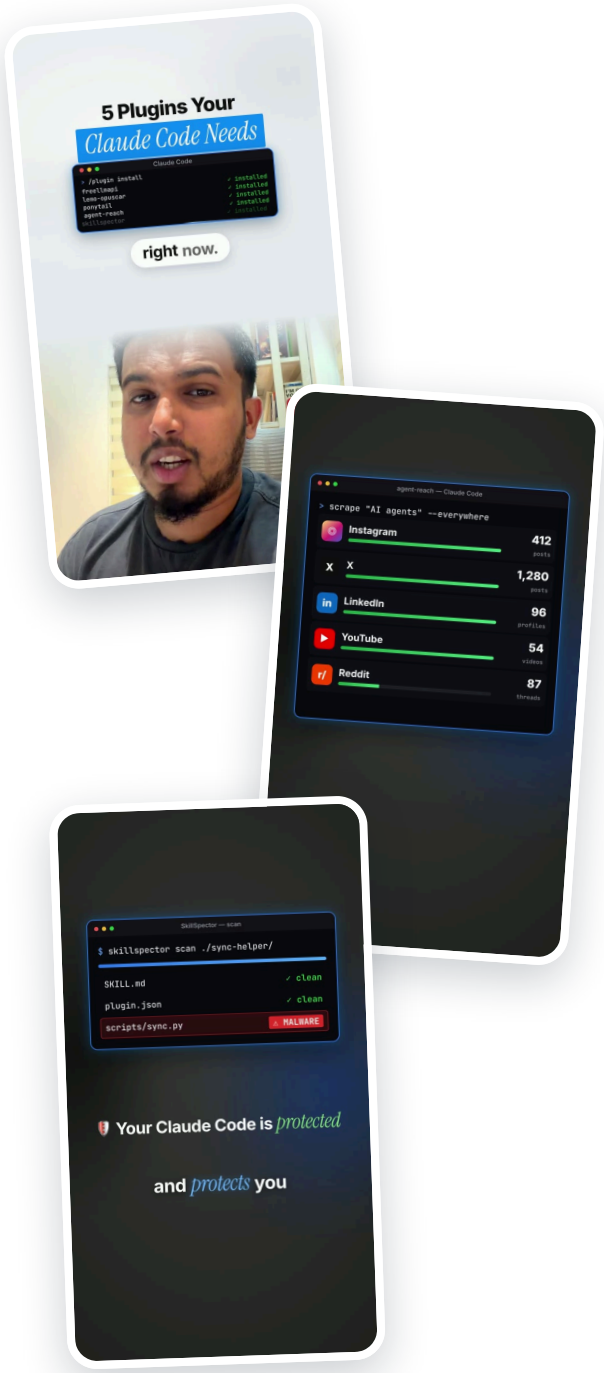
FREE GUIDE · CLAUDE CODE PLUGINS

5 Claude Code Plugins You Need

The five plugins from my reel. For each one: what it does, the real numbers, how to install it step by step, the prompt to give Claude, and what to watch out for.

Atheef Abdul Rahman
Ailutions



Each plugin fixes a different problem: running out of usage, making animated videos, wasting tokens, getting data off the internet, and keeping your setup safe. Install them in the order on page 8 and you will feel the difference the same day.

How plugins work in Claude Code

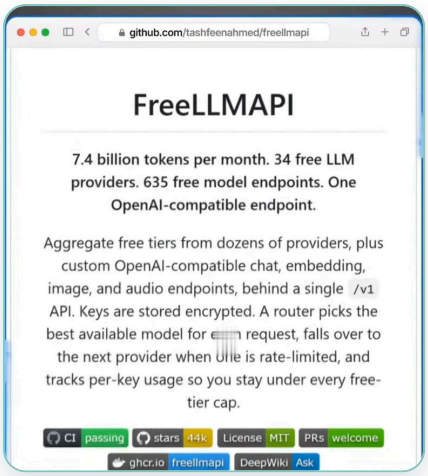
A plugin adds new skills, commands or tools to Claude Code. Most install with two commands. You do this once per plugin.

1. **Open Claude Code**
   Use the Claude desktop app (Code tab) or the `claude` command in Terminal. Plugins work in both.
2. **Add the marketplace**
   A marketplace is the GitHub repo that hosts the plugin: `/plugin marketplace add owner/repo`.
3. **Install the plugin**
   `/plugin install name@marketplace`. Type `/plugin` on its own to browse, enable, disable or remove plugins.
4. **Start a new session**
   New skills and commands load when a session starts. Type `/help` to see the new commands.
5. **Scan before you trust**
   Plugins can run code on your computer. Install from the official repo only, and scan unknown ones with SkillSpector (plugin 5).

The 5 plugins at a glance

Plugin	What it fixes	Cost	Setup	Page
FreeLLMAPI	Running out of usage	Free (needs Docker)	Medium	3
Lemo-Opuscar	Making animated videos	Free (uses your tokens)	Easy	4
Ponytail	Wasted tokens and bloated code	Free	Easy	5
Agent Reach	Getting data off the internet	Free (STT key optional)	Medium	6
SkillSpector	Malicious skills and plugins	Free	Medium	7

> **A note on the numbers**
> All figures in this guide come from each project's own README or benchmark in November 2026. Free tiers and features change often, so check the GitHub page before you rely on a number.



PLUGIN 01 · FREE MODELS

FreeLLMAPI

Pools the free tiers of 34 AI providers behind one address, so you keep working when a limit runs out.

7.4B	635	34
free tokens a month	free model endpoints	AI providers

github.com/tashfeenahmed/freellmapi — MIT · free

A small router that runs on your own computer. You add free API keys from providers like Groq, Cerebras, Gemini and Mistral. It gives you one OpenAI-compatible address and one key. Every request goes to the best free model available, and when one hits its rate limit it switches to the next, so your work carries on.

How it works

- Six routing strategies, automatic failover and per-key usage tracking, so you stay under every free cap.
- Speaks the OpenAI `/v1`, Anthropic `/v1/messages` and Gemini formats, which is how Claude Code can use it.
- Keys are stored encrypted (AES-256-GCM) on your machine. Dashboard at `localhost:3001`.
- Each answer carries an `X-Routed-Via` header showing which provider replied.

Install

```
# needs Docker installed
curl -fsSL https://freellmapi.co/install.sh | bash

# open http://localhost:3001, add your free provider keys,
# then copy the unified freellmapi-... key from the Keys page
npx freellmapi setup-claude --url http://localhost:3001 --api-key <unified-key>
```

How to use it Start Claude Code through the router with `npx freellmapi launch`. Set the model to `auto`, `auto:fast` or `auto:smart`, or pick a specific model ID.

PROMPT TO CLAUDE

Use the FreeLLMAPI router for this task. Draft 20 caption ideas for my reel about [topic], then group them by hook style.

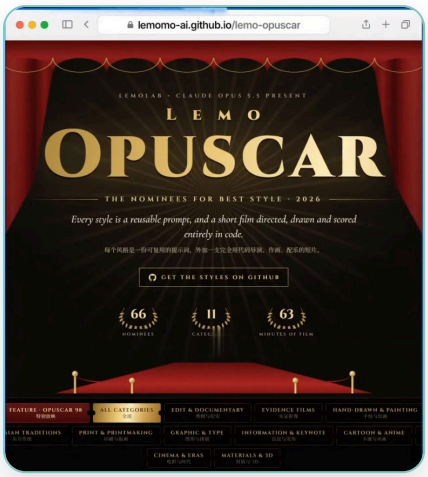
BEST FOR

Side projects, drafts, bulk tasks, background agents and learning without paying per token.

WATCH OUT

No frontier-class models, so do not expect Claude Opus quality. Speed varies and there is no uptime promise. Quality can drop late in the day as top models hit their caps (they reset at midnight UTC). The project describes itself as for personal experimentation only.

Pro tip. Keep your normal Claude login for important work and client data. Use FreeLLMAPI for the high-volume, low-risk jobs.



PLUGIN 02 · ANIMATED VIDEO

Lemo-Opuscar

66 film styles that teach Claude Code to make animated videos entirely in code.

66	11	0
film styles	style categories	stock footage or AI video

github.com/lemomo-ai/lemo-opuscar — MIT code · free

A style library from LemoLab. Each style is a reusable prompt that tells Claude how to direct, draw and score a short film in code. There is no AI video generation and no stock footage. You bring the story and Claude builds every frame, note and cut.

How it works

- Categories include Hand-drawn & Painting, East Asian Traditions, Print & Printmaking, Cartoon & Anime, Games, Cinema & Eras, Materials & 3D and Information & Keynote.
- Example styles: Watercolor Brush, Ukiyo-e, Swiss Motion Graphics, Risograph Print, Chinese Ink Wash, 16-bit Pixel RPG.
- Make a style your own: custom styles are saved in `~/lemo-my-styles/` and survive updates.
- Needs Node 20+, ffmpeg 7+ and Python 3.11+ (or uv). Claude installs the rest.

Install

```
claude plugin marketplace add lemomo-ai/lemo-opuscar
claude plugin install lemo-opuscar@lemolab

# then, inside Claude Code
/lemo-opuscar:lemo-opuscar
```

How to use it Run the command above, or just describe the film. Default output is 1920 x 1080 at 24 fps, so ask for vertical if it is for Reels.

PROMPT TO CLAUDE

Make a 45-second film in the watercolor style about [your topic]. Make it vertical, 1080 x 1920, for Instagram Reels. Use this script: [paste].

BEST FOR

Animated explainers, faceless reels, brand stories, product intros and lesson videos.

WATCH OUT

One film takes about 30 to 60 minutes and a fair amount of tokens. The styles are tuned for Claude Opus 5.5, so other models may not match them. Three narration tracks in the demo films are not licensed for reuse.

Pro tip. Start with a 20 to 30 second film to learn a style before you commit to a long one.



PLUGIN 03 · FEWER TOKENS

Ponytail

Makes Claude Code write less code and use fewer tokens, without cutting the parts that matter.

–45%	–53%	98%
tokens	lines of code	risky logic tested

github.com/DietrichGebert/ponytail — MIT · free

A skill that makes your agent "think like the laziest senior dev in the room". Before writing anything it climbs a ladder: is this code needed at all? Does it already exist in the project? Does the standard library or platform already do it? Only then does it write the minimum.

How it works

- Reads the relevant code first, then picks the lowest step on the ladder that works.
- Never cuts validation, error handling, security or accessibility.
- Adds one small test for risky logic (branches, loops, parsers, money, security).
- Ends each reply with what it skipped or did not check.

Install

```
/plugin marketplace add DietrichGebert/ponytail
/plugin install ponytail@ponytail
```

How to use it Commands: `/ponytail lite | full | ultra | off` sets the intensity. `/ponytail-review` checks your current changes, `/ponytail-audit` checks the whole repo, `/ponytail-gain` shows the scoreboard, `/ponytail-help` lists everything.

PROMPT TO CLAUDE

/ponytail full. Build a simple contact form that saves entries to a Google Sheet. Keep it as small as possible and tell me what you skipped.

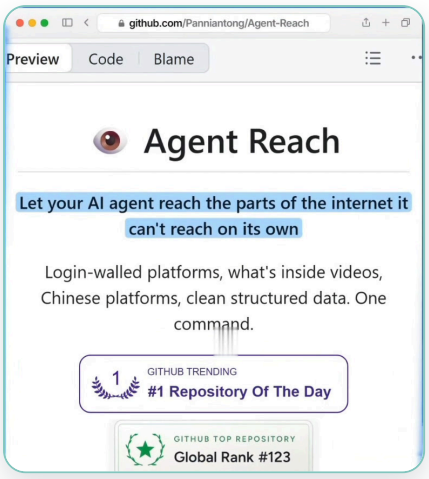
BEST FOR

Building apps, scripts and automations, and keeping projects small and easy to maintain.

WATCH OUT

The numbers come from the author's own benchmark (Claude Code with Opus 5.5, 39 tasks, 5 runs each: –53% code, –41% time, –26% cost, –45% tokens), so your results will vary. Install only from the official GitHub repo or npm package. Copies with `.exe` or `.dll` files are not from the author.

Pro tip. Run `/ponytail-review` before every commit. It catches missing tests and shortcuts early.



PLUGIN 04 · INTERNET DATA

Agent Reach

Lets Claude search and read the internet, including login-walled platforms, with zero API fees.

14	1	$0
platform channels	command line tool	API fees

github.com/Panniantong/Agent-Reach — MIT · free

A read-only tool your agent uses to search and read content it normally cannot reach: YouTube subtitles, Reddit threads, GitHub, LinkedIn, X, Instagram and more. It only reads. It never posts, comments, likes or sends messages.

How it works

- `agent-reach read <url>` picks the right channel for any link.
- Works with no login: YouTube, GitHub, web search, Bilibili, V2EX. Optional: Reddit, LinkedIn and more. Experimental: X, Instagram, Facebook.
- Login platforms reuse your Chrome login inside a separate browser. It does not read your saved passwords.
- YouTube speech-to-text needs a Groq or OpenAI key: run `agent-reach configure groq-key` in your own terminal.

Install

```
# paste this message into Claude Code:
Install Agent Reach: https://raw.githubusercontent.com/Panniantong/Agent-Reach/main/docs/install.md

# do NOT pip install agent-reach (that is a different package)
```

How to use it Claude installs it, may ask to install `uv`, and asks which optional channels you want. Later just say "check Agent Reach status" or "upgrade Agent Reach".

PROMPT TO CLAUDE

Use Agent Reach to search Reddit and YouTube for what people are saying about [topic] this month. Give me the top 10 posts with links and a 3-line summary of each.

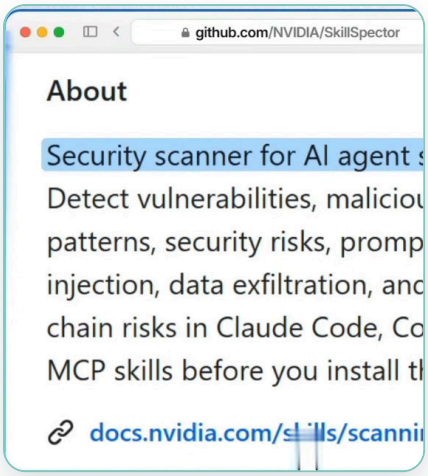
BEST FOR

Content research, competitor and trend tracking, reading YouTube videos as text, and market research.

WATCH OUT

X, Instagram and Facebook are experimental and can stop working at any time. Platforms may restrict automated access, so use a secondary account, not your main one. Respect each site's terms and people's privacy. The project has no token or coin; any crypto using its name is a scam.

Pro tip. Turn transcripts of the top YouTube videos in your niche into a list of hooks for your next reels.



PLUGIN 05 · SECURITY

NVIDIA SkillSpector

A security scanner that checks skills, plugins and MCP servers before you install them.

71	17	0 to 100
risk patterns	categories	risk score

github.com/NVIDIA/SkillSpector — Apache 2.0 · free

Part of NVIDIA's Verified Skills pipeline. It reads a skill (a folder, a `SKILL.md`, a zip or a GitHub link) without running it, and flags prompt injection, data exfiltration, credential theft, privilege escalation, malicious code, typosquatted dependencies and MCP tool poisoning. You get a risk score and a clear verdict.

How it works

- Score 0 to 20: **SAFE**. 21 to 50: **CAUTION**. 51 and up: **DO NOT INSTALL**.
- Reports as terminal output, JSON, Markdown or SARIF (`--format`).
- `--no-llm` runs static checks only, so file contents never leave your machine.
- Needs Python 3.12+. Install with uv or pip, or run it in Docker.

Install

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

# scan before you install anything
skillspector scan https://github.com/user/some-skill

# optional: let Claude Code call it directly
uv tool install 'skillspector[mcp] @ git+https://github.com/NVIDIA/skillspector.git'
claude mcp add skillspector -- skillspector mcp
```

How to use it Run a scan from the terminal, or with the MCP server added, ask Claude to scan a link for you before you install it.

PROMPT TO CLAUDE

Before I install this plugin, scan it with SkillSpector and tell me the verdict, the risk score and the top findings: [GitHub link].

BEST FOR

Checking any third-party skill, plugin or MCP server, and checking skills before you share them with your team.

WATCH OUT

It is a scanner, not a sandbox or a real-time antivirus. It warns you before you install, but it cannot contain a skill you install anyway. With LLM analysis on (the default), file contents are sent to your configured AI provider; use `--no-llm` to keep everything local.

Pro tip. Scan every plugin in this guide too. Good habits start with the tools you trust most.

Install them safely

Plugins run on your computer with access to your files. Five habits keep you safe.

- ✓ **Install from the source.** Use the GitHub links in this guide. Fake copies of popular plugins appear on npm, PyPI and file-sharing sites.
- ✓ **Read what it can touch.** Check the README for what the plugin runs, which files it reads and what it sends over the internet.
- ✓ **Scan it.** Run `skillspector scan <link>` and do not install anything scored 51 or higher.
- ✓ **Keep keys out of chat.** Enter API keys and logins in your own terminal or the plugin's dashboard, never in a chat message.
- ✓ **Use secondary accounts.** For scraping tools, log in with a spare account so your main one is never at risk.

The order I recommend

#	Plugin	Why this order
1	**SkillSpector**	So you can scan everything else first.
2	**Ponytail**	Cuts token use on everything you do after it.
3	**Agent Reach**	Research and data for your content and projects.
4	**Lemo-Opuscar**	Turn that research into animated videos.
5	**FreeLLMAPI**	Add free models for bulk and background jobs.

Your turn

Install one plugin today, starting with SkillSpector so you can scan the rest. Then add the others one at a time and test each one on a small task. Send me a message and tell me which one saved you the most time.

Atheef Abdul Rahman · Ailutions