Tech Latte

FREE GUIDE

Run Claude Code on free AI models

How to connect Claude Code to FreeLLMAPI, a free, open-source project that pools the free tiers of dozens of AI providers behind one setup. The link, the setup steps, and the limits to know first.

INSIDE	What it is, what to know before you start, and eight setup steps.
COST	The software is free (MIT license). It runs on free accounts you open with each AI provider. An optional $19-a-year plan gets new models faster.
FROM	Tech Latte · techlatte.ai

Checked on the project's GitHub page on 23 November 2026

1 What it is

Claude Code is the app you type into. It reads your files, writes and edits them, and runs tasks for you. Normally the thinking behind it comes from Anthropic's Claude models, under your Claude plan's limits.

FreeLLMAPI is a free app that runs on your own computer. You connect free accounts from other AI providers, such as Google, Groq, Cerebras, Mistral and OpenRouter. Claude Code then sends its work through FreeLLMAPI to those free models instead.

When one model hits its free limit, FreeLLMAPI moves to the next one on your list, called the fallback chain. It can also pass along a short handoff note, so the next model knows what you were working on. That note is optional.

github.com/tashfeenahmed/freellmapi

Providers	87 free AI providers
Models	736 model families (600+ in our video)
Free capacity	About 10.1 billion tokens a month, added up across all of those free tiers
Popularity	About 57,000 stars on GitHub · MIT license

Figures from the project's own GitHub page, 23 November 2026.

2 Know this first

> - **The 10.1 billion is pooled.** It is the free capacity of all 87 providers added together, not one person's allowance. What you get depends on which accounts you connect.
> - **You open each account yourself.** Each provider's own limits and terms still apply, even when your work goes through FreeLLMAPI.
> - **These are not Anthropic's Claude models.** Claude Code looks the same, but the answers come from the free models you connect. The maker says the top models aren't included and quality can dip late in the day as the best free models hit their daily caps. They reset at midnight UTC.
> - **Your Claude plan doesn't change.** This does not raise Anthropic's limits on Claude itself. Your normal Claude Code still works the way it did.
> - **It's for practice.** The maker says it is for personal experimentation and learning, not for real business work. Start with a practice project, and keep client files and private information out of it. Free tiers have their own terms, so read them before you send anything you'd want kept private.

3 Set it up

These steps use the FreeLLMAPI desktop app for Mac or Windows, the simplest way in. Plan on about 30 minutes.

1

Install Claude Code

Follow Anthropic's setup page. Then open Terminal (Mac) or PowerShell (Windows) and type `claude --version`. If you see a version number, it's ready.

docs.claude.com/en/docs/claude-code/setup

2 Install Node.js

The connect command in step 7 needs Node.js version 20 or newer. Download the LTS version, then check with `node --version`.

nodejs.org

3 Install the FreeLLMAPI desktop app

On a Mac, download the .dmg from the latest release: arm64 for Apple Silicon, x64 for Intel. On Windows, use the Microsoft Store version, which is signed and updates itself. The app sits in your menu bar or system tray. There is no password to set up.

github.com/tashfeenahmed/freellmapi/releases/latest

4 Add your free provider keys

Click the tray icon, then **Open Dashboard**, then the **Keys** page. Open a free account with a provider, create an API key there, and paste it in. Start with two or three providers; you can add more later. The full list, with each one's limits, is at freellmapi.co/models.html.

5 Set your fallback chain

In the dashboard, put the models in the order you want them tried. When the first one runs out, FreeLLMAPI moves to the next. The **Playground** page lets you send a test message and see which model answered.

6 Copy your one key

The tray popover (or the top of the Keys page) shows a single key that starts with `freellmapi-`. It is the only key Claude Code needs. Keep it private, like a password.

7 Open Claude Code through it, in a practice folder

In Terminal, make a practice folder and start Claude Code there through FreeLLMAPI. Replace YOUR-KEY with the key from step 6.

```
mkdir practice-project
cd practice-project
npx freellmapi launch --url http://127.0.0.1:31415 --api-key
YOUR-KEY
```

This runs Claude Code on your free models for this session only and saves nothing to your settings. Typing plain `claude` later still uses your normal Claude account.

8 Optional: make it your default

If you want every Claude Code session to use FreeLLMAPI, the project has a setup command. Run it with `--dry-run` first to see what it would change. It backs up your settings before it writes anything.

```
npx freellmapi setup-claude --url http://127.0.0.1:31415 --api-
key YOUR-KEY --dry-run
```

The handoff note from section 1 is a separate, optional setting. The project's guide explains it under "Context Handoff". You can skip it.

---

4 If something goes wrong

Nothing answers	Check the FreeLLMAPI app is running in your menu bar or tray, and that the key you pasted is the one from step 6.
A provider fails	Open the Keys page. Each key has a status dot. Re-create the key on the provider's site if it shows a problem.
Answers get worse	The best free models may have hit their daily caps. They reset at midnight UTC, or you can add another provider.
You want to stop	Close the Claude Code session and use plain `claude` again. To remove FreeLLMAPI, quit and delete the app, then its data folder (the project's FAQ lists where).

The masterclass: Claude, ChatGPT and AI employees

The masterclass from our video is in the Agentic AI Academy, an online community where business owners learn to put Claude, ChatGPT and AI employees to work, with courses and weekly live coaching. No coding needed.

Join the Academy

We're members and this is our referral link, so we may earn a commission if you join. It costs you nothing extra.

Tech Latte helps owner-led businesses put AI to work safely.
techlatte.ai · PO Box 113, Boyds, MD 20841
FreeLLMAPI was checked on its GitHub page on 23 November 2026. Providers, limits and prices change, so check the project's own page before you rely on it. Independent Tech Latte guide. We are not affiliated with FreeLLMAPI, Anthropic or any provider named.