

● OPEN SOURCE

THE COMPANION TO THE VIDEO

The \$0 AI Stack

Free, open-source replacements for **\$179/month of AI subscriptions** — fact-checked, with the honest catches.

WHAT'S INSIDE

- ▶ 9 swaps — chat, coding, search, deep research, images, video, voice, music, avatars — each with its **direct GitHub link**.
- ▶ The **honest catch** on every one: licenses, VRAM, free tiers that rot.
- ▶ A **"how to start"** line so you can try each tonight.
- ▶ The hardware truth: **3 lanes** from a 16GB laptop to a used RTX 3090.
- ▶ And the **3 tools I'd still pay for** — because honesty cuts both ways.

9

SWAPS

\$179

/MO IN SUBSCRIPTIONS, KILLED

\$0

/MO TO RUN THEM

THE OVERVIEW

9 subscriptions. 9 open-source swaps.

ChatGPT Plus, Cursor, Perplexity, Midjourney, Runway, ElevenLabs, Suno, HeyGen — stacked together they quietly bill **\$179 every month**. Every swap below is open-source and yours to run, and every one comes with the catch named, not buried. Facts verified June 11, 2026.

#	THE SWAP	WHAT IT KILLS
01	Open WebUI + Gemma 4 CHAT	ChatGPT Plus (\$20/mo)
02	Cline CODING	Cursor Pro (\$20/mo)
03	Vane (formerly Perplexica) AI SEARCH	Perplexity Pro (\$20/mo)
04	GPT Researcher DEEP RESEARCH (BONUS)	The \$100/mo ChatGPT Pro tier (50 deep-research sessions/mo)
05	ComfyUI + FLUX.2 Klein IMAGES	Midjourney Standard (\$30/mo)
06	Wan 2.2 + LTX-2 VIDEO	Runway Pro (\$28/mo, annual rate)
07	Chatterbox VOICE	ElevenLabs Creator (\$22/mo for 100K characters)
08	ACE-Step 1.5 MUSIC	Suno Pro (\$10/mo)
09	InfiniteTalk AVATAR	HeyGen Creator (\$29/mo, ~10 min cap)

ONE HONEST NOTE UP FRONT

"\$0" means **no subscription** — not "no cost, ever." Some swaps need a GPU you own, and a couple run on API keys that cost pennies per use. Where that's true, it's flagged in the catch. What you always own is the tool, the data, and the freedom from a recurring bill.

01

CHAT

Open WebUI + Gemma 4

github.com/open-webui/open-webui github.com/ollama/ollama**KILLS**

ChatGPT Plus (\$20/mo).

- **Open WebUI** (~141K stars) is the front end; pair it with Gemma 4 12B (Apache 2.0, runs on a 16GB laptop) or Qwen3.6-27B via **Ollama**.
- Offline mode, voice calls, image generation, and web search built in.
- Chat with your documents — supports 9 different vector databases.

HOW TO START`pip install open-webui` or Docker, then `ollama run gemma4`.

△ **THE CATCH** Open WebUI's license now has a branding clause (it's not plain BSD anymore). And a local 12B model is not frontier GPT on the hardest reasoning tasks.

02

CODING

Cline

github.com/cline/cline**KILLS**

Cursor Pro (\$20/mo).

- **Cline** — ~63K stars, Apache-2.0. Plan/Act modes, approval on every edit, checkpoints.
- Runs your terminal and fixes its own compile errors. VS Code, JetBrains, and CLI.
- Free model routes: **OpenRouter**'s qwen3-coder:free (480B, 20 req/min) or fully local via **Ollama**.

HOW TO START

Install the Cline extension in VS Code, pick a model route, and approve the first edit.

△ **THE CATCH** Free endpoints rot: Gemini CLI's free tier shuts down June 18, 2026, and Qwen Code's free tier died April 15, 2026. Local models don't rot.

03

AI SEARCH

Vane (formerly Perplexica)
github.com/ItzCrazyKns/Vane
github.com/searxng/searxng
KILLS

Perplexity Pro (\$20/mo).

- **Vane** — renamed from Perplexica in March 2026. ~35K stars, MIT, built on **SearXNG**.
- Speed / Balanced / Quality modes; searches web, discussions, and academic sources, plus image and video search.
- Deep-research mode added April 2026, and it works with **Ollama** for fully local answers.

HOW TO START Docker compose up, point it at SearXNG, and pick a model.

THE CATCH Largely a one-maintainer project, and answer quality is only as good as the model you wire in.

04

DEEP RESEARCH (BONUS)

GPT Researcher
github.com/assafelovic/gpt-researcher
KILLS

The \$100/mo ChatGPT Pro tier (50 deep-research sessions/mo).

- **GPT Researcher** — ~27K stars, Apache-2.0. Planner / executor / publisher agent architecture.
- Pulls from 20+ sources per run and writes 2,000+ word reports.
- Roughly \$0.40 and ~5 minutes per run — so 50 runs costs about \$20 of API credit instead of \$100.

HOW TO START `pip install gpt-researcher`, add your keys, ask a question.

THE CATCH Bring your own OpenAI + **Tavily** keys — it's pennies per run, not zero.

05

IMAGES

ComfyUI + FLUX.2 Klein

github.com/Comfy-Org/ComfyUI

KILLS

Midjourney Standard (\$30/mo).

- **ComfyUI** (~116K stars) plus FLUX.2 Klein 4B (Apache 2.0, ~13GB VRAM) or Qwen-Image (Apache 2.0, best text-in-image, 8–12GB quantized).
- License truth: FLUX.2-dev weights are non-commercial, BUT the outputs are explicitly free for commercial use.
- VRAM ladder: 8GB = SDXL / Qwen quantized · 12–16GB = Klein · 24GB = full FLUX.2-dev.

HOW TO START Install ComfyUI, download Klein 4B, load the default FLUX workflow.

△ **THE CATCH** "Free" here means you own the GPU and pay the electricity.

06

VIDEO

Wan 2.2 + LTX-2

github.com/Wan-Video/Wan2.2
github.com/Lightricks/LTX-Video

KILLS

Runway Pro (\$28/mo, annual rate).

- **Wan 2.2** — ~16K stars, Apache 2.0; the 5B variant does 720p on one consumer card.
- **LTX-2** — ~10K stars, Apache 2.0; the only open model generating video + synced audio in one pass, up to 4K, 20s clips.

HOW TO START Both run inside ComfyUI — install the Wan or LTX nodes and start with the 5B variant.

△ **THE CATCH** A 12–24GB card buys you 5–10s clips that take minutes each. TRUST WARNING: "Wan 2.7 open source" articles are SEO fakes — Alibaba never open-sourced past 2.2.

07

VOICE

Chatterboxgithub.com/resemble-ai/chatterbox**KILLS**

ElevenLabs Creator (\$22/mo for 100K characters).

- **Chatterbox** — ~25K stars, MIT. In blind tests, preferred 65.3% vs ElevenLabs' 24.5%.
- Clones a voice from 5–10 seconds of audio; emotion dial; [laugh] / [cough] tags.
- Turbo variant: 350M params, 75ms latency, runs on a 6GB GPU. Watermarks its output.

HOW TO START`pip install chatterbox-tts`, feed it a 10-second sample.

△ **THE CATCH** Turbo is English-only, the multilingual variant is not the model that won the blind test, and long audio needs chunking.

08

MUSIC

ACE-Step 1.5github.com/ace-step/ACE-Step-1.5 github.com/fspecii/ace-step-ui**KILLS**

Suno Pro (\$10/mo).

- **ACE-Step 1.5** — ~11K stars, MIT. A full song in under 10 seconds on an RTX 3090; turbo mode runs in ≤6GB VRAM.
- Songs from 10 seconds to 10 minutes, 50+ languages, runs on Mac / AMD / Intel / NVIDIA.
- Community **Spotify-style UI**, and LoRA fine-tuning on your own tracks takes ~1 hour.

HOW TO START

Clone the repo, run the gradio app, type a style prompt.

△ **THE CATCH** The devs' own words: quality sits "between Suno v4.5 and v5" — one generation behind. Fine for content music, not release-grade vocals.

09

AVATAR

InfiniteTalkgithub.com/MeiGen-AI/InfiniteTalk**KILLS**

HeyGen Creator (\$29/mo, ~10 min cap).

- **InfiniteTalk** — ~6.8K stars, Apache-2.0, v1.5 released May 21, 2026.
- One photo + an audio file → unlimited-length talking video; syncs lips, head, and body.
- Pair it with **Chatterbox** for your own cloned voice on the avatar.

HOW TO START Clone the repo, download the v1.5 weights, feed it one photo + one WAV.

⚠ **THE CATCH** The weakest swap of the nine — a research-grade pipeline: 16–24GB VRAM, minutes per clip, 720p ceiling.

THE RUNNING TALLY

ChatGPT Plus \$20 + Cursor Pro \$20 + Perplexity Pro \$20 + Midjourney \$30 + Runway \$28 + ElevenLabs \$22 + Suno \$10 + HeyGen \$29 = **\$179/month killed** — plus the \$100/mo ChatGPT Pro deep-research tier replaced for ~\$20 of API credit.

BEFORE YOU INSTALL ANYTHING

Hardware: the 3 lanes

Everything in this guide runs somewhere on this ladder. Find your lane first — it decides which swaps you can run tonight and which need an upgrade.

LANE 1 · WHAT YOU ALREADY OWN

A 16GB laptop

Enough for the everyday swaps **today**: chat (Open WebUI + Gemma 4 12B), AI search (Vane), and music (ACE-Step turbo runs in ≤6GB VRAM). No purchase required — start here.

LANE 2 · THE ONE-TIME BUY

A 24GB GPU — used RTX 3090 ≈ \$700–800, once

Unlocks **everything in this guide**: full FLUX.2-dev images, Wan 2.2 / LTX-2 video, InfiniteTalk avatars, instant ACE-Step songs. One payment roughly equal to four months of the full subscription stack — then it's yours.

LANE 3 · THE FREE-CLOUD SHORTCUT

Free cloud tiers

They work — and **they rot**. See the June 18, 2026 Gemini CLI free-tier shutdown, and Qwen Code's free tier before it. Fine for trying things; don't build your workflow on someone else's promo budget.

HONESTY CUTS BOTH WAYS

The 3 I'd still pay for

A \$0 stack is a tool, not a religion. These three are where the paid version still earns its bill.

1

Claude

Frontier agentic coding has **no open equal** right now. If code is how you make money, this is the one subscription that pays for itself.

2

A \$10 flat-rate coding plan — if you don't own a GPU

Flat rates are **subsidized**; API tokens aren't. If you're in Lane 1 with no GPU, a cheap flat-rate plan beats burning metered API credit on every edit.

3

Suno — if music is the product you sell

ACE-Step's own devs put it "between Suno v4.5 and v5." For background and content music, free wins. If the **song itself is what you sell**, the paid generation gap matters.

⚠ THE LICENSE WATCH-LIST

FLUX.2-dev — outputs are commercially OK, but serving the weights as a service is not. **HunyuanVideo** — license is void in the EU, UK & South Korea. **Zero-asterisk picks** (truly permissive): Wan 2.2, LTX-2, Chatterbox, ACE-Step, InfiniteTalk, Qwen-Image, and Klein 4B.

BUILT BY HYPERAUTOMATION LABS

Stop renting. Start owning.

You now hold 9 open-source swaps for \$179/month of AI subscriptions — with every catch named, the hardware lanes mapped, and the 3 bills still worth paying. Pick your lane, install one swap tonight, and cancel the bill it makes pointless.

WANT THIS PDF + FUTURE STACKS?

ZEROSTACK

Comment the keyword **ZEROSTACK** and we'll send the link straight to you.

Where to go next

- ▶ Website: hyperautomationlabs.co
- ▶ YouTube — deep dives like this one: [@hyperautomationlabs](https://www.youtube.com/@hyperautomationlabs)
- ▶ Instagram — bite-size AI drops: [@hyperautomationlabs](https://www.instagram.com/hyperautomationlabs)