ROUNDUP

Jan Alternatives: 6 Local AI Apps Compared for 2026

Six alternatives to Jan for running LLMs on your own hardware: LM Studio, Ollama, Msty, AnythingLLM, Open WebUI and GPT4All, with pricing and trade-offs.

Why look past Jan?

Jan is a good default for running language models on your own machine. Per its docs, it is open source under Apache 2.0, runs offline by default, pulls open models such as Llama, Gemma and Qwen from Hugging Face, connects to cloud providers like OpenAI, Anthropic and Groq when you want them, supports MCP, and exposes an OpenAI-compatible server on localhost:1337. It costs nothing.

People still go looking for alternatives, usually for one of a few reasons. Some want a headless model server they can script against. Others need chat over their own documents built in, or a web front end a whole team can log into. A few just want a different interface. The six tools below cover those cases.

How scores work here

scored.tools rates tools on five criteria: quality of output, ease of use, pricing value, API and integration quality, and how well the tool solves the problem it claims to solve. None of the tools in this roundup has a published scored.tools rating yet, so no entry carries a score, and the list follows no ranking. Each entry describes what the product does and what its own pages and documentation say about pricing.

LM Studio

Score: not yet rated

Lm Studio is the closest like-for-like swap for Jan. It is a desktop app for finding, downloading and chatting with local models, plus a local server that speaks the OpenAI API format and an SDK for developers. The biggest change in its history concerns licensing. Until July 2025 the app terms told companies to contact LM Studio for a separate commercial licence. That requirement was dropped, and the company's blog now says you no longer need a commercial licence to use it at work, with no

form to fill in. The catch for open-source purists is that the app itself is closed source, where Jan's code is public. If that matters to your security review, Jan keeps the edge. If you want a polished model browser and a free licence that now covers office use, LM Studio is the obvious first stop.

Best for: individuals and teams who want a desktop model runner with a strong model browser and don't need the source code.

Pricing: Free for personal and work use, covering the app, SDK and Hub. An Enterprise plan adds SSO, model gating and private sharing; pricing is by quote.

Ollama

Score: not yet rated

Ollama is a model runtime first and a chat app second. You pull a model with one command, and it serves it over a local REST API that also offers OpenAI-compatible endpoints. That makes it the default backend for a long list of other tools, including Open WebUI further down this page. Ollama now also sells cloud access to larger models. Third-party write-ups list Free, Pro at $20 a month and Max at $100 a month, with Pro described as 50 times the free allowance. Several of those sources report that Ollama moved its paid cloud tiers to usage-based billing on 31 August 2026, while others still describe session limits, so check ollama.com/pricing before you budget for it. Local use is unmetered either way. If you liked Jan's local API server more than its chat window, Ollama is the cleaner fit.

Best for: developers who want a scriptable local model server to sit behind their own apps or another front end.

Pricing: Local use is free. Cloud tiers are reported at $20/month (Pro) and $100/month (Max), with billing terms recently in flux.

Msty

Score: not yet rated

Msty (now sold as Msty Studio) is a desktop client that puts local and cloud models side by side in one interface, so you can compare answers across providers in the same chat. Its paid tier, Aurum, is where most of the extras live. Msty lists access to Msty Studio Web, Forge Mode, real-time data options through the API, Turnstiles automation and advanced performance metrics. In a November 2025 post Msty announced price rises: Aurum Yearly went from $129 to $149 and Aurum Lifetime from $249 to $349. That makes Msty the only paid-first option here for solo users, and the lifetime price is high next to Jan's zero. The case for paying is the workflow polish on top of model access, which Jan and LM Studio don't try to match.

Best for: solo users who switch between local and cloud models all day and will pay for a more refined workspace.

Pricing: Free tier available. Aurum is $149/year or $349 lifetime per Msty's November 2025 announcement; confirm current prices on its pricing page.

AnythingLLM

Score: not yet rated

Anythingllm is built around chatting with your own documents. You create workspaces, drop in files, and it handles embedding and retrieval so the model answers from your material. It can run models itself or connect to Ollama, LM Studio and the main cloud APIs, and it has agent features and a developer API. The docs promise that AnythingLLM Desktop "will and forever will be free to use for 99% of our features." A self-hosted Docker version covers multi-user setups, and managed cloud hosting is listed by third-party sites at around $50 a month for the basic tier, with higher tiers and enterprise plans above that. Jan has added assistants and MCP, but if retrieval over a folder of PDFs is the main job, AnythingLLM is purpose-built for it.

Best for: people whose main use is asking questions of their own files, privately.

Pricing: Desktop app is free. Managed cloud reportedly starts around $50/month; check the official pricing page for current tiers.

Open WebUI

Score: not yet rated

Open Webui is a self-hosted, browser-based chat interface that usually sits on top of Ollama or any OpenAI-compatible endpoint. It's the pick when you want to give a whole team a ChatGPT-style front end running on your own server, with user accounts and permissions. Watch the licence. Since v0.6.6 in April 2025, Open WebUI has used a BSD-3–style licence with an added branding clause. Per its docs, you can use it for personal, business or teaching purposes as long as the "Open WebUI" branding stays intact. White-labelling needs an Enterprise licence, and third-party guides note a small-deployment exception of up to 50 users in a rolling 30 days. For an internal tool that keeps its branding, it costs nothing beyond your hosting.

Best for: teams that want a shared, self-hosted chat UI on top of their own model server.

Pricing: Free to self-host with original branding. Enterprise licence (white-labelling, support) is quoted by sales.

GPT4All

Score: not yet rated

Gpt4all from Nomic was one of the first easy local chat apps, and its LocalDocs feature lets you chat with files on your machine. It is free and open source. The concern is maintenance. Real Python notes

that GPT4All hasn't shipped a desktop release since version 3.10.0 in February 2025, and a 2026 guide says you shouldn't expect new features or an up-to-date inference engine. Newer model architectures land in Jan, LM Studio and Ollama first. GPT4All still runs, but picking it today means accepting a stalled project, so it ranks as a fallback for older setups rather than a recommendation.

Best for: existing users with a working LocalDocs setup who don't need newer models.

Pricing: Free and open source.

Comparison table

Tool	Score	Type	Open source	Local API	Price
Jan	Not rated	Desktop app	Yes (Apache 2.0)	OpenAI-compatible	Free
Lm Studio	Not rated	Desktop app	No	OpenAI-compatible	Free; Enterprise by quote
Ollama	Not rated	Model runtime	Yes	Native REST plus OpenAI-compatible	Free locally; cloud from $20/mo (reported)
Msty	Not rated	Desktop client	No	Client, connects to others	Free tier; Aurum $149/yr or $349 lifetime
Anythingllm	Not rated	Document chat app	Yes	Developer API	Desktop free; cloud from about $50/mo
Open Webui	Not rated	Self-hosted web UI	Yes, with branding clause	Front end for Ollama or OpenAI-compatible APIs	Free with branding; Enterprise by quote
Gpt4all	Not rated	Desktop app	Yes	Local API server	Free

Final picks

For a straight replacement for Jan's desktop experience, go with Lm Studio. It's free for work since July 2025, and the only real trade is closed source code. If you build things, start with Ollama as your local model server, and add Open Webui on top when other people need a login. Choose Anythingllm when document Q&A is the main job. Pick Msty only if you'll pay $149 a year for a smoother multi-model workspace. Skip Gpt4all for new setups until Nomic ships again. If none of

those needs applies to you, Jan's mix of an open licence, offline-first defaults and a free price is hard to beat, and staying put is a fine answer.

Generated 2026-10-30 · Read the live article at scored.tools/blog/jan-alternatives/