ToolGPA

NOVEMBER 2026

# LocalAI

## The Insider's Guide

Open-source runtime for running AI models on your own hardware



**Every plan, grade, refund rule and rival among AI assistants, read from LocalAI's own pages.**

Made by Ettore Di Giacinto and the LocalAI team · checked Nov 13, 2026 · toolgpa.com

2.8
GPA



# Contents


**01 Welcome**
What it is, who it suits, its grades and the ten things to know first

**02 What it costs**
Every plan, the free plan, ways to pay less and how prices have moved

**03 What it does**
Its main features, where it runs, what it works with and who makes it

**04 Where it stands**
Its grades against its category, and whether to keep it or make your own

**05 The alternatives**
The tools it is most often weighed against, side by side

**06 Quick answers**
The questions people ask about it, each answered in a line or two

**07 About ToolGPA**
Who made this guide, what else is on the site, and how to cite it


ToolGPA put this guide together from LocalAI's own pages, its app-store listings and our grades. Every fact in it is on LocalAI's page too, where it is kept up to date. Checked Nov 13, 2026.

# Welcome

What it is, who it suits, its grades and the ten things to know first

**2.8**

GPA, above 56 of the 87 other AI assistants we grade

LocalAI is an open-source (MIT) runtime that runs language, voice, vision, image and video models on your own machines, with no GPU required. It exposes an OpenAI-compatible API, and also speaks the Anthropic, Ollama and ElevenLabs APIs, so existing clients can point at it. It also ships a web interface, a command-line chat agent and a gallery of installable models.

To get you started, here are **8 things to know about LocalAI**, taken from its own pages and our grades.

> **You install and operate it yourself, through containers, a macOS app, Linux binaries or Kubernetes.**
>
> From our grade for ease of use

**1** **Best for running models on your own machine, OpenAI-style API**

But you install and run it yourself, choose and download the models, and Windows is not listed on its install page.

**2** **It is free**

LocalAI has no paid plan at all.

### 3 It runs on 3 platforms

Web, macOS and Linux.

### 4 Drop-in API compatibility

One binary serves an OpenAI-compatible API and also speaks the Anthropic, Ollama and ElevenLabs APIs, so most clients need only a new URL.

### 5 Swappable engines

A model can run on llama.cpp, vLLM, SGLang or MLX behind the same endpoint, and each engine is pulled in as a backend only when a model asks for it.

### 6 Text, voice, vision and media

Covers language models with tool calling, transcription, speech synthesis and voice cloning, vision and detection, plus image, video and music generation.

### 7 A 2.8 GPA, above 56 of 87

That is against the other AI assistants we grade, on value, the free plan, features, ease of use and platforms.

### 8 Keep it

LocalAI is free, so making your own would save you nothing.

### LocalAI at a glance

| | |
|---|---|
| PRICING<br>Free | FREE PLAN<br>It is free |
| FREE TRIAL<br>None stated | RUNS ON<br>Web, macOS, Linux |
| MADE BY<br>Ettore Di Giacinto and the LocalAI team | CATEGORY<br>AI Assistants |
| OUR GRADE<br>2.8 GPA, good | |

## The report card

**Best for running models on your own machine, OpenAI-style API**

LocalAI is a free, MIT-licensed runtime that runs language, voice, vision and image models on your own machines and speaks the OpenAI API, so existing clients can point at it. But you install and run it yourself, choose and download the models, and Windows is not listed on its install page.

| TOOLGPA REPORT CARD | LocalAI | Graded Nov 13, 2026 |
|---|---|---|
| SUBJECT | GRADE | COMMENT |
| **Value for money** | A | Free under the MIT licence and anyone can download and run it, on a CPU or a GPU. The site invites businesses to get in touch, but no paid plan is described. |
| **Free plan and trial** | A | Free outright, with no cap from the project. Your own hardware and electricity are the cost. |
| **Features** | B | An OpenAI-compatible API, swappable engines, text, voice, vision and media, agents and MCP. No gap shows in the job, and the gallery lists 1,255 models, though bigger ones need more memory. |
| **Ease of use** | D | You install and operate it yourself, through containers, a macOS app, Linux binaries or Kubernetes. |
| **Platforms and integrations** | C | Mac, Linux and a web interface, with no phone app, and Windows is not listed. It links to llama.cpp, vLLM, MLX, whisper.cpp, Docker and Kubernetes. |
| **Grade point average** | | Good, out of 4.0 **2.8** |

An A is worth 4 points, a B 3, a C 2, a D 1 and an F nothing. The GPA is the average of the five.

### Who LocalAI is for

- ✓ Fits if you want models running on your own hardware, with no data leaving it
- ✓ Works for developers who want a drop-in replacement for the OpenAI API

### Who it is not for

- ✗ Skip it if you do not want to run software yourself, as it is a runtime to install and operate
- ✗ Too light for non-technical use, as you pick and download the models
- ✗ Not for you if you use Windows, as it is not listed on the install page
- ✗ Watch out: bigger models need more memory or a GPU for good speed
- ✗ Watch out: documentation pages have moved between addresses

## Good to know

### Good for

- ✓ Developers who want a local, OpenAI-compatible endpoint for existing tools
- ✓ People who need data to stay on their own hardware
- ✓ Teams running models on a mix of CPU-only machines and GPUs

### Worth knowing

- ⓘ It is a runtime to install and operate yourself, through containers, a macOS app, Linux binaries or Kubernetes. Windows is not listed on the install page.
- ⓘ You choose and download the models. The gallery lists 1,255 models, and larger ones need more memory or a GPU for good speed.
- ⓘ No paid plan or hosted service is described on the pages read. The homepage offers a business contact link.
- ⓘ Documentation pages have moved between addresses, for example from /installation/ to /docs/installation/.

# What it costs

Every plan, the free plan, ways to pay less and how prices have moved

**Free**

No paid plan at all

LocalAI is free to use.

> **Free under the MIT licence and anyone can download and run it, on a CPU or a GPU.**
>
> From our grade for value for money

OPEN SOURCE

## $0

MIT-licensed software you run yourself, on a CPU or on a GPU. The site also invites businesses to get in touch.

Read from localai.io on Nov 13, 2026.

PART 03 · FEATURES

# What it does

Its main features, where it runs, what it works with and who makes it

3

Places it runs: Web, macOS and Linux

Open-source runtime for running AI models on your own hardware. Here is what it does, as its own site describes it, and what it covers: open source, self-hostable, local-first, public API, chat assistant, AI agents, text to speech.

> **An OpenAI-compatible API, swappable engines, text, voice, vision and media, agents and MCP.**
>
> From our grade for features

### Drop-in API compatibility

One binary serves an OpenAI-compatible API and also speaks the Anthropic, Ollama and ElevenLabs APIs, so most clients need only a new URL.

### Swappable engines

A model can run on llama.cpp, vLLM, SGLang or MLX behind the same endpoint, and each engine is pulled in as a backend only when a model asks for it.

### Text, voice, vision and media

Covers language models with tool calling, transcription, speech synthesis and voice cloning, vision and detection, plus image, video and music generation.

### Agents and MCP

The local-ai chat command starts an agent that can run shell commands behind an approval gate, delegate to sub-agents and load MCP servers, plugins and skills.

## CPU-first and distributed

The site says every feature ships a CPU path, with GPU support for CUDA, ROCm, SYCL, Metal and Vulkan, and routing across several machines.

### Where it runs

Web · macOS · Linux

### What it works with

llama.cpp · vLLM · MLX · whisper.cpp · Docker · Kubernetes

LocalAI is made by Ettore Di Giacinto and the LocalAI team.

PART 04 · STANDING

# Where it stands

Its grades against its category, and whether to keep it or make your own

**56**/87

Other AI assistants it is graded above

LocalAI's GPA of 2.8 is higher than 56 of the 87 other AI assistants we grade, and level with 11.

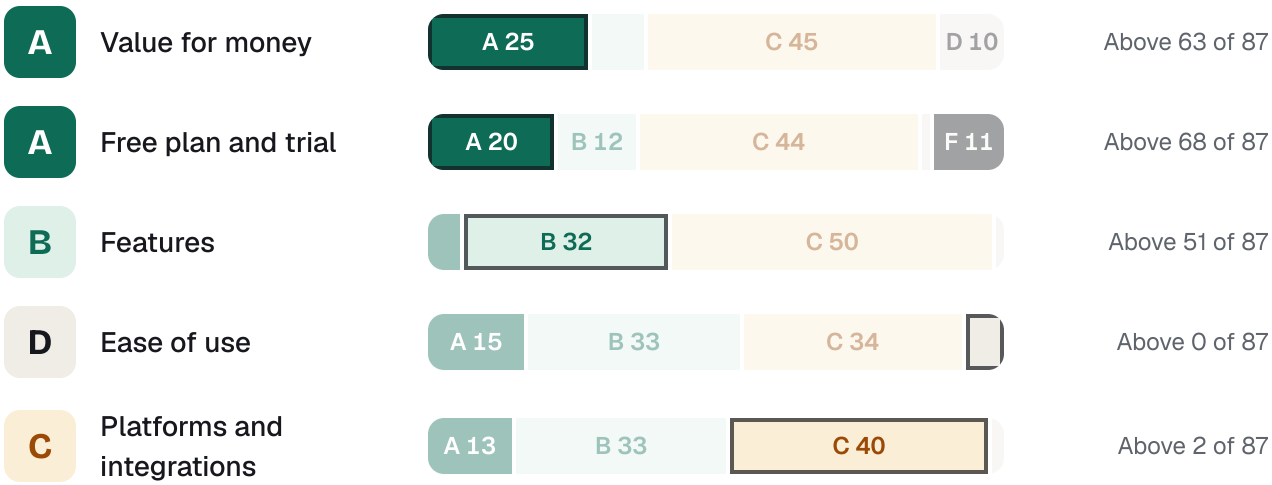
It stands best on free plan and trial (A, above 68 of 87) and worst on ease of use (D, above 0).

LocalAI has no paid plan.

On features, ease of use and platforms it scores 2.0 out of 4. Nothing scores higher for less, though at least one tool matches that score at a lower price.

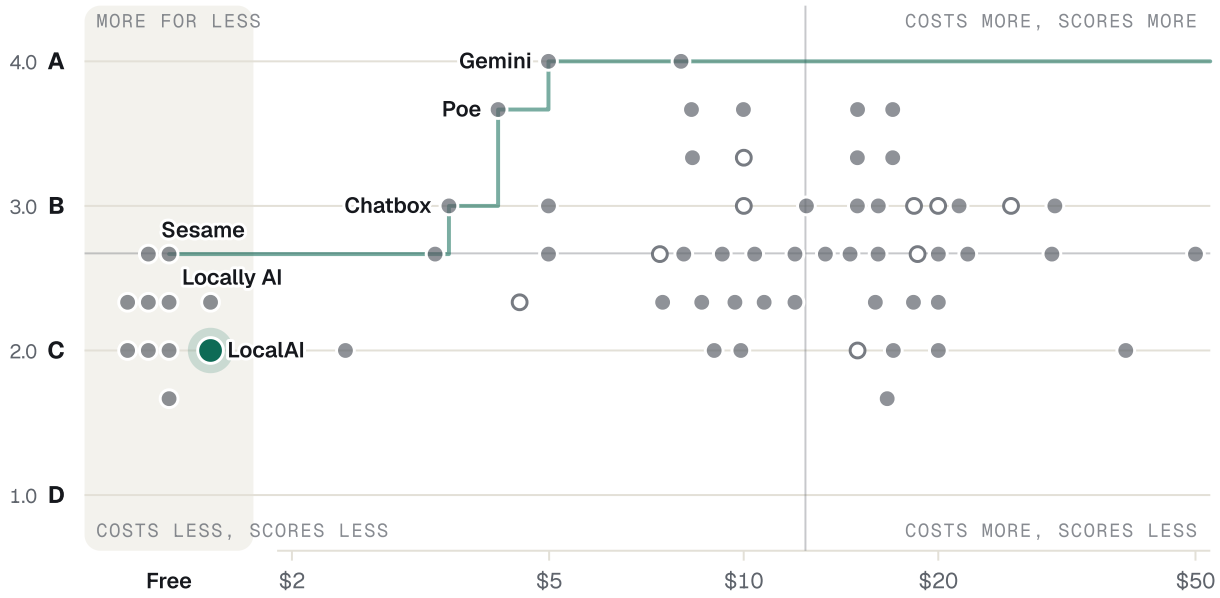
## Each grade against the category

Every bar is all 88 AI assistants we grade, cut by the letter each got. LocalAI's own letter is outlined.



# Price against what you get

Each dot is one of 65 AI assistants. Further left costs less; higher up scores better on features, ease of use and platforms. LocalAI is the marked dot.




What paid plans start at, a month (each step along is roughly double) ⍰

# Keep it, or make your own?

Whether this is something you could build for yourself with an AI app builder, and what you would give up.

## Keep it

LocalAI is free, so making your own would save you nothing.

And it is not the kind of software a weekend replaces.

### What you could not

- ✕ An AI model, which you can rent by the request but not make.
- ✕ An AI agent, with the models and the safeguards behind it.
- ✕ A voice model, which you can rent by the minute but not make.
- ✕ A real app for macOS and Linux. Something home-made would be a web page.

- It plugs into other products, llama.cpp, vLLM and MLX among them, and has an API.
- The depth: we grade its features B, which means it covers the whole job.

**What this rests on.** LocalAI's listing: what it does, where it runs, what it plugs into and what it costs. It is a judgement about the kind of software this is. We have not built a replacement for it, and making your own means looking after it too. How we decide, and every tool's verdict.

# The alternatives

The tools it is most often weighed against, side by side



Ollama, LM Studio, Jan, Open WebUI and AnythingLLM

The 5 tools LocalAI is most often weighed against, with where each one differs from it.

2.6 GPA

## Ollama

Free plan · from $20/mo

Run open AI models on your computer or in Ollama's cloud.

- Only LocalAI lists a web app.
- Only Ollama lists a Windows app.
- LocalAI is open source and self-hostable, according to its own site.

LocalAI and Ollama compared in full

---

3.0 GPA

## LM Studio

Free plan · from $20/mo

Desktop app for running open AI models on your own computer.

- Only LocalAI lists a web app.
- Only LM Studio lists a Windows app.
- LocalAI is open source and self-hostable, according to its own site.

LocalAI and LM Studio compared in full

👋

3.0
GPA

## Jan

Free

Open-source desktop app for running local and cloud AI models.

- Only LocalAI lists a web app.
- Only Jan lists a Windows app.
- LocalAI is self-hostable, according to its own site.

LocalAI and Jan compared in full

OI

3.2
GPA

## Open WebUI

Free plan

Self-hosted chat interface for local and cloud AI models.

- Only Open WebUI lists a Windows app.
- LocalAI is open source, according to its own site.
- LocalAI also covers text to speech.

LocalAI and Open WebUI compared in full

3.2
GPA

## AnythingLLM

Free plan · from $50/mo

Open-source AI app for documents, agents and local models.

- Only AnythingLLM lists Windows and Android apps.
- LocalAI also covers text to speech.
- AnythingLLM also covers meeting notes.

LocalAI and AnythingLLM compared in full

PART 06 · QUESTIONS

# Quick answers

The questions people ask about it, each answered in a line or two

13

Questions, each answered in a line or two

The short version of everything above, for when you only need the answer. Each one is told in full in its part of the guide.

### Is LocalAI free?

Yes. LocalAI is free to use.

### How much does LocalAI cost?

LocalAI lists one plan. Open source: $0.

### What platforms does LocalAI work on?

LocalAI is available on Web, macOS and Linux.

### Does LocalAI have an Android app?

No Android app was listed when its pages were read on Nov 13, 2026. LocalAI is listed for Web, macOS and Linux.

### Does LocalAI have an iPhone app?

No iPhone app was listed when its pages were read on Nov 13, 2026. LocalAI is listed for Web, macOS and Linux.

### Is LocalAI open source?

Yes. LocalAI is open source, according to its own site.

### Does LocalAI have an API?

Yes. LocalAI offers a public API, according to its own site.

### What does LocalAI integrate with?

LocalAI's listing names llama.cpp, vLLM, MLX, whisper.cpp, Docker and Kubernetes.

### Who makes LocalAI?

LocalAI is made by Ettore Di Giacinto and the LocalAI team.

### What are the alternatives to LocalAI?

The closest alternatives to LocalAI in our directory are Ollama, LM Studio, Jan, Open WebUI and AnythingLLM.

### How does LocalAI compare with other AI assistants?

LocalAI's GPA of 2.8 is higher than 56 of the 87 other AI assistants we grade, and level with 11.

### Is LocalAI good value for money?

ToolGPA grades LocalAI A for value for money. LocalAI has no paid plan.

### Could I make my own version of LocalAI?

LocalAI is free, so making your own would save you nothing. And it is not the kind of software a weekend replaces.

## Where this guide comes from

The prices, plans, platforms and features were read from LocalAI's own pages on Nov 13, 2026, and the grades were given on Nov 13, 2026. Graded from this listing, whose facts were read from LocalAI's own pages on 13 November 2026. Not based on using the product, and shorter than our full reviews.

Every tool we list gets five grades, A to F, each with its reason, and a GPA that averages them. How each is decided is at toolgpa.com/grades. This guide is ToolGPA's own work; it is not made, paid for or approved by Ettore Di Giacinto and the LocalAI team. Prices change, so check localai.io before you buy.

localai.io

localai.io/docs/installation/

github.com/mudler/LocalAI

PART 07 · ABOUT

# About ToolGPA

Who made this guide, what else is on the site, and how to cite it

**7,721**

Tools graded, in 95 categories

ToolGPA grades software the way a school report grades a pupil: five subjects, a letter for each with the reason written beside it, and a GPA that averages them. Prices, plans and platforms are read from each maker's own pages and dated, so you always know how fresh a figure is.

Every grade, review and ranking on ToolGPA is our own judgement. No company can pay for a grade or change one. Every tool on the site has a guide like this one, free to download and pass on. It all lives at [toolgpa.com](toolgpa.com).

| 7,721 | 95 | 739 | 579 |
|---|---|---|---|
| tools graded | categories | price changes on record | apps priced in 12 countries |

## What else is on toolgpa.com

### LocalAI, kept up to date

Everything in this guide, with prices and grades as they are today

toolgpa.com/tools/localai

### How grades are given

The rules behind every letter and every GPA

toolgpa.com/grades

### Value charts

What each tool costs against what it scores, for every category

toolgpa.com/charts

### Compare

Any two tools side by side, plan by plan

toolgpa.com/compare

### Price changes

Every price rise and cut on record, each with its source

toolgpa.com/price-changes

### Prices by country

What apps cost on the App Store in 12 countries

toolgpa.com/prices-by-country

### Pay less

Every public discount: yearly billing, students, nonprofits

toolgpa.com/save

### Keep or make

Which tools are worth paying for, and which you could build yourself

toolgpa.com/keep-or-make

**Scan for the live page**

Prices and grades as they are today

### Quoting this guide?

Please name ToolGPA and link to the live page, so readers can check the figures for themselves:

```
ToolGPA, "LocalAI: The Insider's Guide",
checked Nov 13, 2026.
https://toolgpa.com/tools/localai
```

### Know someone choosing between AI assistants?

Send them this guide: toolgpa.com/tools/localai/guide

ToolGPA elsewhere: youtube.com/@toolgpa

# See our other guides

The best-graded AI assistants after LocalAI, each with an Insider's Guide of its own.

### AI Assistants Insider's Guide

Every one of the 88 AI assistants we grade, compared

### Gemini Insider's Guide

Google's AI assistant, tied into Gmail, Docs and other Google apps

3.8 GPA

### Chatbox Insider's Guide

AI chat client for your own API keys or a hosted model plan

3.4 GPA

### ChatGPT Insider's Guide

OpenAI's general-purpose AI assistant for chat, images and voice

3.4 GPA

ToolGPA

## Choosing software? Every tool graded, every price dated.

Compare tools at toolgpa.com ›

Everything in this guide is ToolGPA's own summary of what LocalAI's pages, its app-store listings and public reviews said on the dates shown. Grades are our judgement under the rules at toolgpa.com/grades. Every grade, review and ranking on ToolGPA is our own judgement. No company can pay for a grade or change one. LocalAI is a trademark of its owner, and this guide is not made, paid for or approved by Ettore Di Giacinto and the LocalAI team. Prices and plans change: check localai.io before you buy.