

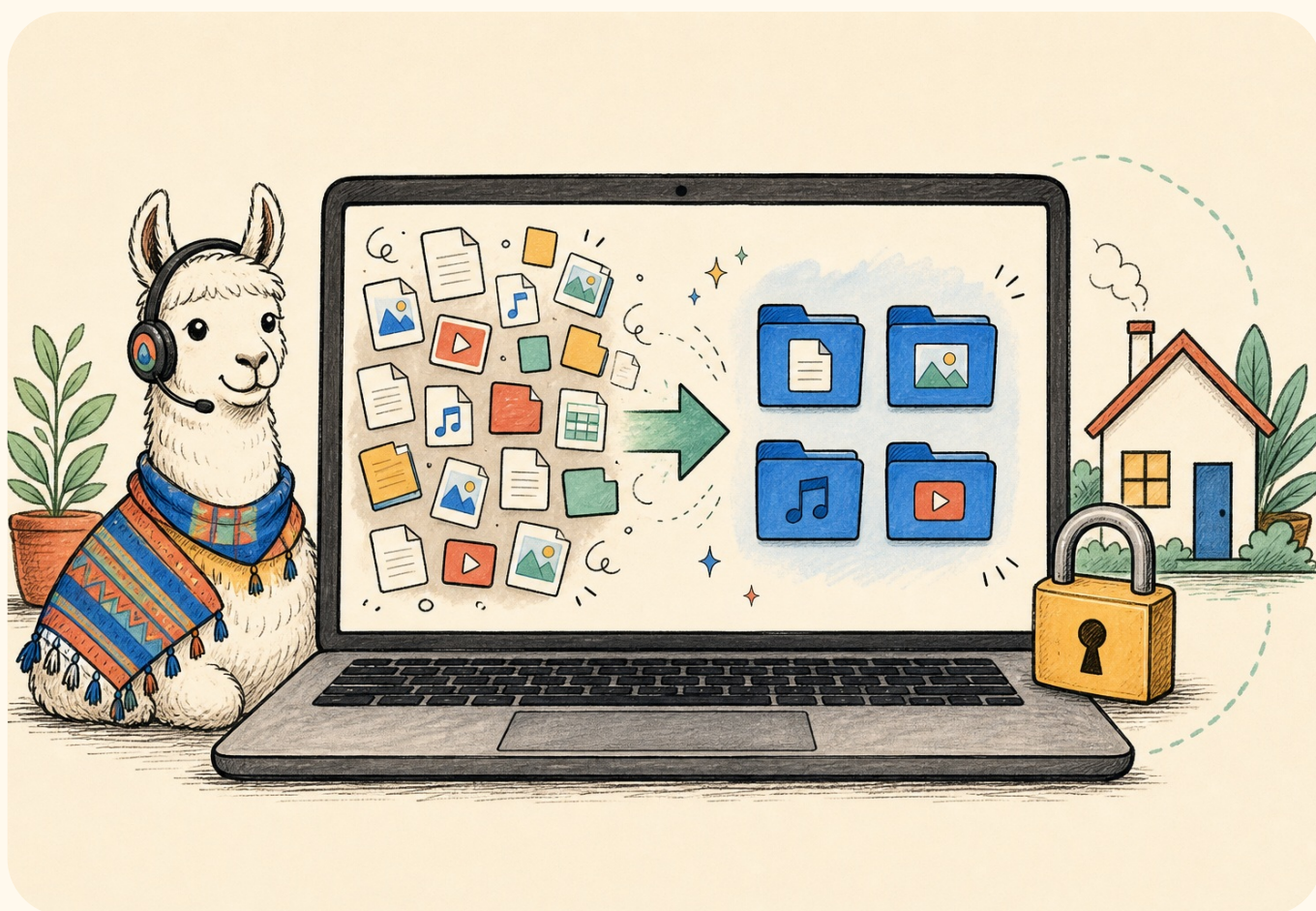


راهنمای خیلی ساده

اولین ایجنت هوش مصنوعی تو

رایگان، آفلاین و روی کامپیوتر خودت

با Llama، Ollama و OpenCode دسکتاپ شلوغ را فقط با تأیید خودت مرتب کن.



مناسب مبتدی

داده روی دستگاه

بدون هزینه API

نکته: سخت‌افزار، برق و دانلود مدل رایگان نیستند؛ اما برای پیام‌ها هزینه API نمی‌دهی.



قبل از هر کاری: اسم ابزارها و لینک‌ها

فعلا چیزی اجرا نکن. فقط این فهرست را ببین و ابزار مناسب سیستم خودت را دانلود کن.

Ollama

۱

مدل Llama را روی کامپیوتر اجرا می‌کند.

ollama.com/download

OpenCode

۲

ایجنتی است که ابزار خواندن و جابه‌جایی فایل دارد.

opencode.ai/docs

Llama 3.1

۳

مدل پیشنهادی؛ با اولین اجرا خودکار دانلود می‌شود.

ollama.com/library/llama3.1

Windows فقط - WSL

۴

محیط لازم برای اجرای ساده OpenCode در ویندوز.

learn.microsoft.com/windows/wsl/install

چیست؟ Terminal

یک برنامه آماده داخل Linux، macOS، و WSL است. لازم نیست جداگانه دانلودش کنی؛ فقط دستورها را داخل آن می‌چسبانی.



ایجنت پولی یا ایجنت رایگان؟

اینجا «رایگان» یعنی بدون هزینه API. این جدول ساده کمک می‌کند انتخاب کنی.

ایجنت رایگان آفلاین

Llama + Ollama + OpenCode

- | | |
|------|---------------------------------------|
| مزیت | برای هر پیام هزینه API نداری |
| مزیت | بعد از دانلود بدون اینترنت کار می‌کند |
| مزیت | فایل‌ها روی کامپیوترت می‌مانند |
| ضعف | نصب و رفع خطا با خودت است |
| ضعف | به RAM، دیسک و برق نیاز دارد |
| ضعف | مدل کوچک دقت کمتری دارد |

ایجنت ابری پولی

مثل سرویس‌های آماده آنلاین

- | | |
|------|---|
| مزیت | نصب ساده‌تر و شروع سریع‌تر |
| مزیت | معمولا مدل قوی‌تر و اجرای ابزار دقیق‌تر |
| مزیت | روی کامپیوتر ضعیف هم قابل استفاده |
| ضعف | اینترنت لازم است |
| ضعف | اشتراک، اعتبار یا محدودیت مصرف دارد |
| ضعف | داده‌ها از دستگاه خارج می‌شوند |

پیشنهاد این آموزش

اگر حریم خصوصی و نداشتن هزینه API برایت مهم است و نصب چند ابزار برایت اشکالی ندارد، مسیر آفلاین را ادامه بده. اگر فقط می‌خواهی سریع شروع کنی، ایجنت ابری ساده‌تر است.



کدام مدل روی کامپیوتر من اجرا می‌شود؟

جواب به مدل، مقدار Context و روش اجرای CPU یا GPU بستگی دارد. این راهنما تقریبی است، نه تضمین قطعی.

سه چیزی که باید بشناسی

حافظه اصلی سیستم: RAM:

بخشی از حافظه یکپارچه است Apple Silicon حافظه کارت گرافیک؛ در: VRAM:

فضای لازم برای دانلود و نگهداری فایل مدل: Disk:

۸ گیگابایت رم - بدون کارت گرافیک قوی

حداقل برای آزمایش

llama3.2:1b | file 1.3 GB | context 4K

برای چت ساده؛ برای ایجنت فایل ضعیف‌تر است.

۱۶ گیگابایت رم یا حدود ۸ گیگابایت حافظه گرافیکی

شروع قابل قبول

llama3.2:3b or llama3.1:8b | context 4K-8K

ممکن است بخشی روی CPU اجرا و کند شود.

۳۲ گیگابایت رم یا ۱۲ تا ۱۶ گیگابایت حافظه گرافیکی

انتخاب بهتر

llama3.1:8b | context 16K-32K

برای ابزارها بهتر؛ Context بزرگ‌تر باز هم حافظه بیشتری می‌خواهد.

هم اجرا می‌شود، اما معمولاً کندتر است GPU الزامی نیست؛ بدون GPU



یعنی چه؟ Context Window

VRAM و RAM مثل اندازه میز کار مدل است: هرچه بزرگتر باشد، متن و ابزارهای بیشتری را همزمان می‌بیند؛ اما Context بیشتری مصرف می‌کند.

پیش‌فرض فعلی Ollama بر اساس VRAM

Default context: 4K	حافظه گرافیکی کمتر از ۲۴ گیگابایت
Default context: 32K	حافظه گرافیکی ۲۴ تا ۴۸ گیگابایت
Default context: 256K	حافظه گرافیکی ۴۸ گیگابایت یا بیشتر

نکته مهم برای ایجنت‌ها

برای کارهای ایجنتی کانتکست بزرگ‌تری پیشنهاد می‌کند؛ اما این مقدار می‌تواند حافظه زیادی بخواهد. عدد حداکثر روی OpenCode صفحه مدل به معنی قابل اجرا بودن آن روی هر لپ‌تاپ نیست.

OpenCode recommendation: context >= 64K

از ChatGPT بخواه برای انتخاب کند

این متن را همراه مشخصات کامپیوترت بفرست

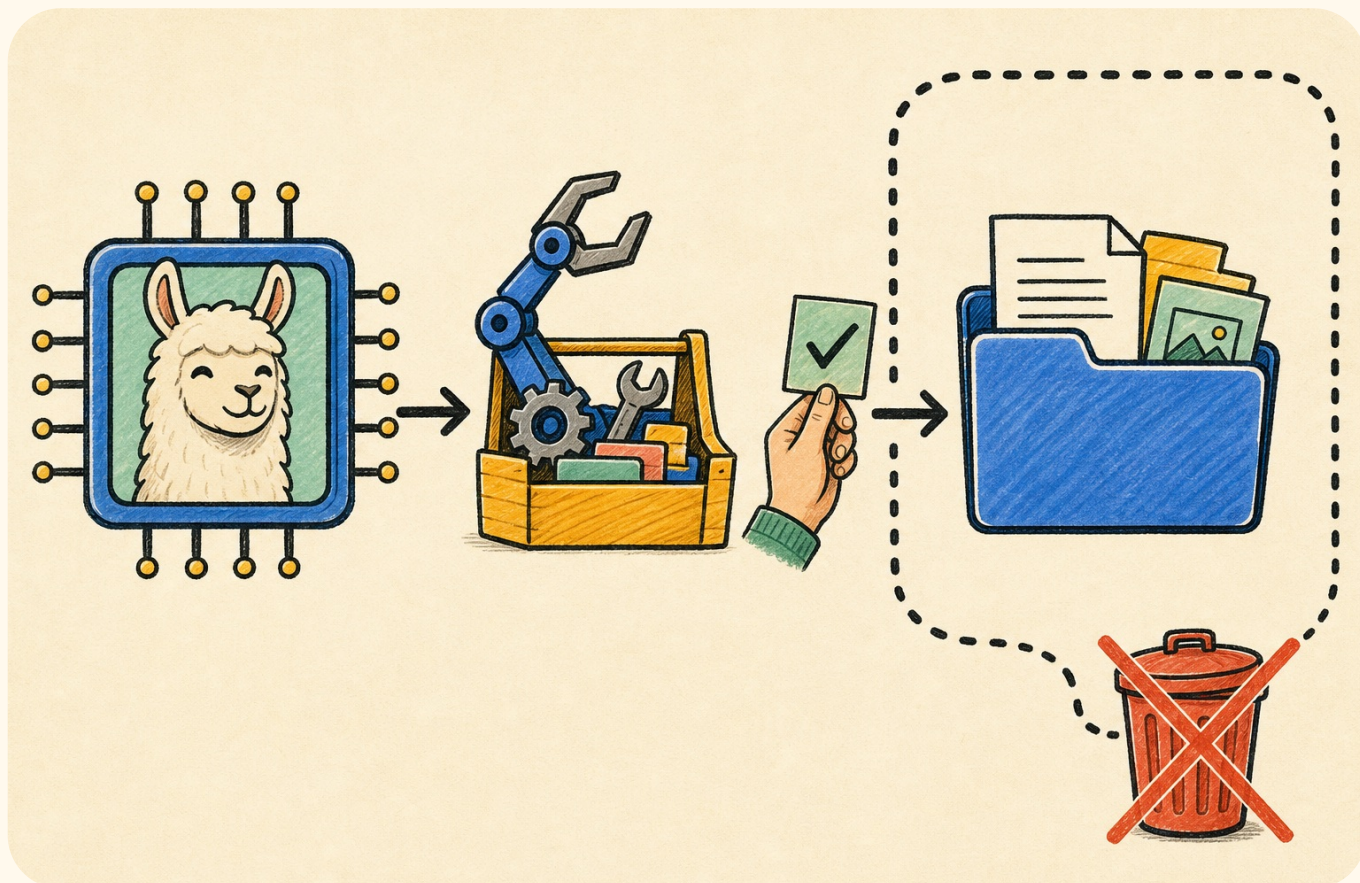
I want to run Ollama with OpenCode for a Desktop file-organizing agent. My computer: OS = [macOS/Windows/Linux], CPU = [...], RAM = [...], GPU = [...], VRAM = [...], free disk = [...]. Compare llama3.2:1b, llama3.2:3b, and llama3.1:8b. Recommend the safest model and a realistic context size from 4K, 8K, 16K, 32K, or 64K. Explain expected speed, whether it fits in memory, and give the exact Ollama model name. Do not guess missing specs; ask me for them.

Windows: Settings > System > About و Task Manager > Performance
macOS: Apple menu > About This Mac و System Settings > General > Storage



سه ابزار کنار هم یک ایجنت می‌سازند

محدوده کار است Desktop ابزار می‌دهد و OpenCode. به‌تنهایی فایل جابه‌جا نمی‌کند Llama



تو تأیید می‌کنی

۳

قبل از تغییر فایل‌ها، برنامه را بررسی می‌کنی.

اجرا می‌کند OpenCode

۲

ابزار ساخت پوشه و جابه‌جایی فایل را دارد.

فکر می‌کند Llama

۱

برنامه و پاسخ را روی کامپیوتر می‌سازد.

مرزهای این آموزش

فقط Desktop؛ حذف ممنوع؛ تغییر نام ممنوع؛ بازنویسی ممنوع؛ فایل مشکوک داخل Review-Later؛ و اجرای کار فقط بعد از تأیید تو.



نصب قدم به قدم

۱. Ollama را نصب کن

۱

لینک را باز کن و نسخه سیستم خودت را دانلود کن.

<https://ollama.com/download>

۲. اگر Windows داری، WSL را نصب کن

۲

باز کن Run as administrator را با PowerShell.

`wsl --install`

بعد از اجرا، کامپیوتر را Restart کن

۳. OpenCode را نصب کن

۳

در macOS/Linux از Terminal و در Windows از Ubuntu/WSL استفاده کن.

`curl -fsSL https://opencode.ai/install | bash`

مدل Llama جداگانه نصب نمی‌شود؛ هنگام اولین اجرا Ollama آن را دانلود می‌کند.



اول وارد Desktop شو، بعد ایجنت را اجرا کن

این کار باعث می‌شود محدوده اصلی ایجنت همان پوشه Desktop باشد.

macOS و Linux

```
cd ~/Desktop  
ollama launch opencode --model llama3.1
```

Windows / Ubuntu داخل WSL

```
cd /mnt/c/Users/YOUR-NAME/Desktop  
ollama launch opencode --model llama3.1
```

در اولین اجرا چه اتفاقی می‌افتد؟

باز می‌شود. این مرحله OpenCode را دانلود می‌کند؛ حجم مدل حدود ۴/۹ گیگابایت است. سپس llama3.1 مدل Ollama ممکن است طول بکشد و به اینترنت نیاز دارد.

اگر کامپیوتر ضعیف است

به جای llama3.1 از llama3.2 استفاده کن؛ حدود ۲ گیگابایت است، اما ممکن است ابزارها را کم‌دقت‌تر اجرا کند. برای توقف فوری Ctrl + C را بزن.



این پرامپت را داخل OpenCode بچسبان

نسخه انگلیسی برای اجرای ابزارها قابل اعتمادتر است؛ فقط متن را کپی و ارسال کن.

پرامپت آماده

Review the files in this Desktop folder and propose a simple folder structure by file type and purpose. First show only the plan and exact folder names, then wait for my approval. Do not delete, rename, overwrite, or move anything before I approve. After approval, create only the approved folders and move files into them. Put duplicates and uncertain files in Review-Later. Stay inside this Desktop folder. Stop if a filename collision or unexpected error occurs, and report it. Finish with a short list of every moved file.

اگر برنامه درست بود

I approve this exact plan. Proceed.

اگر مطمئن نیستی

Do not change anything. Stop now.



اگر این پنج مورد درست است، آماده‌ای

نصب شده‌اند Ollama و OpenCode

1

اجرا کرده‌ام Desktop را از داخل OpenCode

2

پرامپت حذف، تغییر نام و بازنویسی را ممنوع کرده است.

3

ایجنت اول فقط برنامه را نشان داده است.

4

در صورت خطا Ctrl + C را می‌زنم و توقف می‌کنم.

5

فرمول ایجنت آفلاین امن

مدل روی کامپیوتر + ابزار محدود + Desktop + تأیید تو

منابع رسمی

ollama.com/download | opencode.ai/docs

docs.ollama.com/integrations/opencode | learn.microsoft.com/windows/wsl/install