



Retourner au contenu associé (lien : [Clubic] Six semaines sans cloud : j'ai confié tout mon travail à une IA locale à 4 500 €)

[^] # Re: GTT

Posté par **bubar** 🐧 le 19 juillet 2026 à 17:00. En réponse au lien [Clubic] Six semaines sans cloud : j'ai confié tout mon travail à une IA locale à 4 500 €. Évalué à 7 (+4/-0). Dernière modification le 19 juillet 2026 à 17:04.

Dans les grandes lignes :

Le sucre autour, c'est Newell qui s'en charge. Et bien qu'il apporte un *overhead* si on compare à LMstudio, ce dernier ~~pue~~ ~~fort~~ n'est pas libre et d'autre part je suis un fervent supporter de Newell du début et tjs à ce jour :-)



Pour la conf :

- llama-server v9748 (livré par l'équipe Newell)
 - reasoning-format désactivé
 - reasoning-budget désactivé
 - semantic-memory activée
- gemma-4-e4b.gguf
 - température à 0
 - Pénalité de fréquence à 2
 - Native tools calling
 - désactivation du TextStreamer pour éviter le "think leak" sans désactiver le "think"
 - Mmproj activé pour la reconnaissance d'image
 - accélération matérielle activée
- RAG actif / accès à certains dossiers de mes documents

Reconnaissance vocale :

- Whisper C++ (ggml-small-bin)
 - model small
 - VAD activé (silence à 1 / speech-duration à 5)
 - accélération matérielle activée

Synthèse vocale :

- Kokoro (82m / voice v1)

Utilitaire d'affichage gpu :

- radeontool

En espérant avoir répondu à tes questions :)

Répondre

Se connecter

Proposer un contenu

Identifiant

Identifiant

Mot de passe

Mot de passe

☐ Connexion automatique

Se connecter

Pas de compte ? S'inscrire...

Revenir en haut de page

Derniers commentaires

- Re: résumé
- Re: web
- web
- Re: Re
- Re: Fr
- No future...
- Re: refuser les track...
- Re: Youtube
- scrapping
- plausible
- PhotoRec
- Moi aussi je fais un ...

Étiquettes (tags) populaires

- intelligence_artificielle
- grands_modèles_de...
- merdification
- états-unis
- administration_fran...
- capitalisme
- réchauffement_clim...
- datacenter
- france
- canicule
- bigtech
- jihad_butlérien

Sites amis

- Agenda du Libre
- April
- Éditions D-BookeR
- Éditions Diamond
- Éditions Eyrolles
- Éditions ENI
- En Vente Libre
- Framasoft
- La Quadrature du Net
- Lea-Linux
- Open Source Initiative
- Imprimerie Grafik Plus

À propos de LinuxFr.org

- Mentions légales
- Faire un don
- L'équipe de LinuxFr....
- Informations sur le s...
- Aide / Foire aux que...
- Suivi des suggestion...
- Wiki du site
- Règles de modération
- Statistiques
- API pour le développ...
- Code source du site
- Plan du site