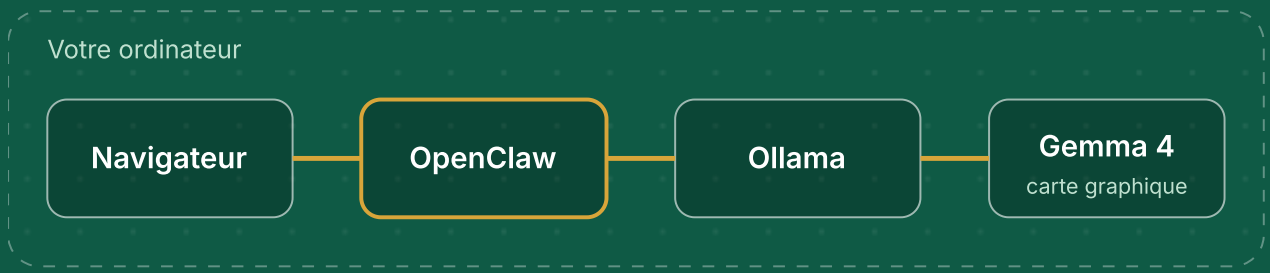
Tutoriel pas à pas · Linux

Un assistant IA sur votre ordinateur, sans cloud

Installer OpenClaw avec le modèle Gemma 4 et Ollama, pour disposer d'un assistant qui fonctionne entièrement chez vous : sans abonnement, sans clé API, et sans que vos conversations quittent votre machine.



Pierre Droulon · droulon.dev

Novembre 2026

Avant de commencer

OpenClaw est un **agent** : au-delà de répondre à vos questions, il peut lire et écrire des fichiers, lancer des commandes et naviguer sur le web avec vos droits. C'est ce qui le rend utile, et c'est aussi pourquoi ce guide insiste sur quelques règles de prudence.

Ollama est le moteur qui fait tourner le modèle d'intelligence artificielle sur votre carte graphique. Gemma 4 est le modèle lui-même, un modèle ouvert de Google, dans sa version compacte.

Ce qu'il vous faut

Élément	Recommandé	Pourquoi
Ordinateur	Linux récent (Ubuntu, Debian, Fedora…)	Les commandes de ce guide sont écrites pour Linux.
Carte graphique	NVIDIA avec 8 Go de mémoire vidéo	Le modèle prend environ 3,5 Go et tient alors entièrement sur la carte, ce qui le rend rapide.
Mémoire vive	16 Go ou plus	Confort d'utilisation pendant que le modèle tourne.
Espace disque	10 Go libres	Logiciels et modèle.
Internet	Pendant l'installation	Ensuite, tout fonctionne hors ligne, sauf si vous demandez à l'agent de naviguer.

Comptez environ **une heure**, dont une partie d'attente pendant les téléchargements.

Ce guide a été testé sous Linux avec OpenClaw 2026.9, Ollama et Gemma 4 E4B sur une carte graphique de 8 Go. Les options peuvent évoluer d'une version à l'autre : en cas de doute, la commande `--help` fait foi.

Trois gestes à connaître

- **Ouvrir un terminal** : touches `Ctrl` + `Alt` + `T` sur la plupart des distributions.
- **Coller une commande dans le terminal** : `Ctrl` + `Maj` + `V`, puis `Entrée`. Copiez une seule commande à la fois, sans les lignes qui commencent par `#`, qui ne sont que des explications.
- **Mot de passe** : quand une commande commence par `sudo`, le terminal demande votre mot de passe. Rien ne s'affiche pendant la frappe, c'est normal : tapez-le puis `Entrée`.

> **À lire avant d'utiliser l'assistant**
>
> Prenez cinq minutes pour lire les règles d'or, plus loin dans ce guide. Elles évitent l'essentiel des mauvaises surprises.

1 Vérifier la carte graphique

2 minutes

Cette commande affiche le modèle de votre carte NVIDIA et sa mémoire.

```
nvidia-smi
```

Vous devez voir un tableau avec le nom de la carte et sa mémoire, par exemple `8188MiB`. Si la commande est introuvable, installez d'abord le pilote NVIDIA recommandé par votre distribution, depuis son outil « Pilotes additionnels » ou équivalent, puis redémarrez.

> **Sans carte NVIDIA ?**
> Le guide fonctionne aussi, mais le modèle tournera sur le processeur : les réponses seront nettement plus lentes.

2 Installer Ollama

5 minutes

Ollama s'installe avec son script officiel, puis tourne en arrière-plan.

```
curl -fsSL https://ollama.com/install.sh | sh
```

Par défaut, Ollama donne au modèle une mémoire de conversation trop courte pour un agent. On l'agrandit avec un petit fichier de réglages. Copiez ce bloc en entier, d'un seul coup :

```
sudo mkdir -p /etc/systemd/system/ollama.service.d
sudo tee /etc/systemd/system/ollama.service.d/override.conf <<'EOF'
[Service]
Environment="OLLAMA_CONTEXT_LENGTH=65536"
Environment="OLLAMA_FLASH_ATTENTION=1"
Environment="OLLAMA_KV_CACHE_TYPE=q8_0"
EOF
```

Puis appliquez les réglages :

```
sudo systemctl daemon-reload && sudo systemctl restart ollama
```

Réglage	Effet
`CONTEXT_LENGTH`	Le modèle se souvient d'environ 64 000 « jetons » de conversation, le minimum conseillé pour OpenClaw.
`FLASH_ATTENTION`	Calcul plus rapide et moins gourmand en mémoire.
`KV_CACHE_TYPE`	Compresse la mémoire de conversation pour qu'elle tienne sur la carte.

Piège fréquent

La ligne `[Service]` est indispensable. Sans elle, les réglages sont ignorés sans aucun message d'erreur.

3 Télécharger et tester le modèle

5 à 15 minutes selon la connexion

```
ollama pull gemma4:e4b
```

Une fois le téléchargement terminé, posez-lui une première question, directement :

```
ollama run gemma4:e4b "Bonjour, présente-toi en une phrase."
```

Le modèle affiche parfois une phase « Thinking… » avant de répondre : il réfléchit, c'est normal. Vérifiez ensuite qu'il tourne bien sur la carte graphique :

```
ollama ps
```

Ce que vous devez voir

Une ligne `gemma4:e4b` avec `100% GPU` et `65536` dans la colonne du contexte. Si la liste est vide, c'est que le modèle s'est mis en veille : relancez la commande `ollama run` ci-dessus, puis `ollama ps`.

4 Installer OpenClaw

10 minutes

Le script officiel installe OpenClaw et, si besoin, une version récente de Node.js. Lancez-le **avec votre compte habituel, sans** `sudo` : il demandera lui-même votre mot de passe s'il en a besoin.

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

Fermez ensuite le terminal et ouvrez-en un nouveau. Sans cela, la commande `openclaw` reste introuvable.

Lancez enfin l'assistant de configuration, qui installe aussi le service en arrière-plan :

```
openclaw onboard --install-daemon
```

Quand il vous demande le fournisseur de modèle, choisissez **Ollama** (local). Allez jusqu'au bout des questions : un assistant interrompu laisse une configuration incomplète.

5 Relier OpenClaw au modèle

5 minutes

Ces trois réglages assurent qu'OpenClaw trouve le modèle et démarre correctement, même si l'étape précédente a été interrompue.

Choisir Gemma 4 comme modèle par défaut, puis accepter les réparations que propose le diagnostic :

```
openclaw models set ollama/gemma4:e4b
openclaw doctor
```

Indiquer que le service fonctionne en local :

```
openclaw config set gateway.mode local
```

Signaler au service qu'Ollama est disponible. OpenClaw ne voit Ollama que si une variable précise existe dans son environnement. Copiez ce bloc en entier :

```
mkdir -p ~/.config/systemd/user/openclaw-gateway.service.d
tee ~/.config/systemd/user/openclaw-gateway.service.d/ollama.conf <<'EOF'
[Service]
Environment="OLLAMA_API_KEY=ollama-local"
EOF
```

```
systemctl --user daemon-reload
openclaw gateway restart
```

> **Ce n'est pas un vrai mot de passe**
>
> La valeur `ollama-local` est simplement le signal qu'attend OpenClaw pour un Ollama installé sur la même machine. Elle ne protège rien et n'a pas besoin d'être gardée secrète.

6 Vérifier et ouvrir l'assistant

2 minutes

```
openclaw models list --provider ollama --refresh
```

Ce que vous devez voir

La ligne `ollama/gemma4:e4b` avec `yes` dans les colonnes `Local` et `Auth`.

Ouvrez enfin l'interface dans votre navigateur :

```
openclaw dashboard
```

Utilisez toujours cette commande pour l'ouvrir : elle ajoute automatiquement la clé d'accès à l'adresse. Envoyez un premier message dans une nouvelle conversation : votre assistant fonctionne.

7 Sécuriser, une fois pour toutes

5 minutes

Dans sa configuration de départ, l'agent exécute des commandes sur votre ordinateur sans rien demander. Deux réglages changent cela.

Lancer l'audit de sécurité et appliquer les corrections sans risque qu'il propose :

```
openclaw security audit
openclaw security audit --fix
```

Exiger votre accord avant toute commande que vous n'avez pas déjà autorisée. Vérifiez d'abord les options exactes de votre version avec `openclaw exec-policy --help`, puis :

```
openclaw exec-policy set --host gateway --security allowlist --ask on-miss
openclaw gateway restart
```

Pour tester, demandez à l'assistant : « Affiche la date avec la commande date. » L'interface doit vous demander votre accord avant d'exécuter quoi que ce soit.

> **Un garde-fou, pas un coffre-fort**
>
> Les approbations évitent les erreurs, mais n'isolent pas complètement l'agent de votre système. Pour un usage personnel, elles constituent, avec les règles d'or, un niveau de sécurité raisonnable.

Les règles d'or

1. L'interface reste accessible **uniquement depuis votre ordinateur**. N'ouvrez jamais son accès au réseau ou à Internet.
2. Ne lancez jamais OpenClaw avec `sudo` ni en administrateur.
3. Ne demandez jamais à l'agent de « faire ce que dit » une page web ou un fichier que vous n'avez pas écrit : ils peuvent contenir des instructions cachées.
4. N'installez des extensions (skills) que depuis des sources de confiance, après avoir lu leur audit.
5. Relisez toute commande qui supprime, déplace ou modifie des fichiers avant de l'autoriser.
6. Faites une sauvegarde avant chaque mise à jour.
7. Relancez `openclaw security audit` après chaque changement de réglage ou nouvelle extension.
8. Ne confiez jamais de mot de passe, de clé ou de donnée bancaire à l'assistant.

Au quotidien

Bien formuler ses demandes

Gemma 4 en version compacte est rapide, mais il s'égare sur les tâches longues. Quatre habitudes suffisent à obtenir des résultats fiables :

1. **Une tâche par message.** « Liste les fichiers .log de mon dossier Projets » plutôt que « range tout et fais-moi un résumé ».
2. **Des chemins précis.** Indiquez le chemin exact du dossier plutôt que « mon dossier météo ».
3. **Un plan avant l'action.** « Propose les commandes que tu comptes lancer, n'exécute rien. » Puis validez.
4. **Une vérification après.** L'agent peut annoncer une réussite alors qu'une commande a échoué : demandez-lui d'afficher le résultat.

Bon à savoir

- **Une nouvelle conversation par sujet.** Une conversation trop longue est résumée automatiquement, et des détails se perdent.
- **Une première réponse plus lente après une pause.** Le modèle se met en veille après 5 minutes sans message et doit se recharger.
- **Arrêter une tâche en cours** : envoyez simplement le mot `stop`. En dernier recours, dans un terminal : `openclaw gateway stop`.

- **Masquer les messages techniques** : envoyez `/verbose off`, `/trace off` ou `/reasoning off`.

Sauvegarder et mettre à jour

Créez une sauvegarde vérifiée maintenant, puis avant chaque mise à jour :

```
mkdir -p ~/Sauvegardes/openclaw
openclaw backup create --output ~/Sauvegardes/openclaw --verify
```

Pour mettre à jour, prévisualisez d'abord, puis appliquez et vérifiez :

```
openclaw update --dry-run
openclaw update
openclaw doctor
```

> **Deux interdits**
>
> Ne copiez jamais les fichiers internes d'OpenClaw à la main pendant qu'il tourne : la copie serait corrompue. Et ne mettez jamais à jour avec `sudo npm install -g openclaw` : utilisez `openclaw update`.

En cas de problème

Commencez toujours par ces deux commandes : elles disent ce qui ne va pas, et proposent souvent la solution.

```
openclaw gateway status
openclaw doctor
```

Ce que vous voyez	Cause probable	Solution
`openclaw: command not found`	Terminal ouvert avant l'installation.	Fermez-le et ouvrez-en un nouveau.
Le service ne démarre pas, le journal parle de `gateway.mode`	Configuration initiale interrompue.	`openclaw config set gateway.mode local` puis `openclaw gateway restart`.
« The configured model is unavailable »	Le service ne voit pas Ollama.	Refaites la troisième partie de l'étape 5, puis vérifiez `Auth: yes` à l'étape 6.
L'interface demande une clé	Adresse tapée à la main.	Ouvrez-la avec `openclaw dashboard`.
Réponses très lentes	Le modèle ne tient pas sur la carte.	`ollama ps` : visez 100 % GPU. Fermez les jeux ou logiciels qui utilisent la carte.
Plus aucune réponse	Ollama arrêté.	`sudo systemctl restart ollama` puis `openclaw gateway restart`.
Le modèle écrit du code au lieu d'agir	Demande trop complexe pour un petit modèle.	Découpez la demande en étapes plus simples.
Réglages d'Ollama sans effet	Ligne `[Service]` manquante.	Recopiez le bloc de l'étape 2 en entier.

Si rien ne fonctionne, relancez tout dans l'ordre :

```
openclaw gateway stop
sudo systemctl restart ollama
openclaw doctor
openclaw gateway start
openclaw dashboard
```

Aide-mémoire

Commande	Usage
`openclaw dashboard`	Ouvrir l'assistant dans le navigateur.
`openclaw gateway status` / `restart`	État du service, redémarrage.
`openclaw doctor`	Diagnostic et réparations guidées.
`openclaw security audit`	Contrôle de sécurité.
`ollama ps`	Le modèle est-il chargé, sur la carte graphique ?
`openclaw <commande> --help`	Aide exacte de votre version.
docs.openclaw.ai	Documentation officielle : installation, sécurité, sauvegardes, dépannage.