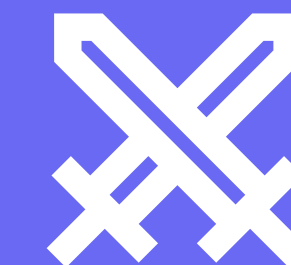




# Les duels de l'IA

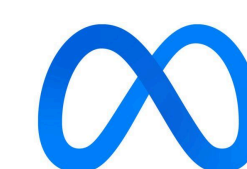
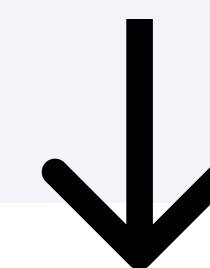
par compar:IA



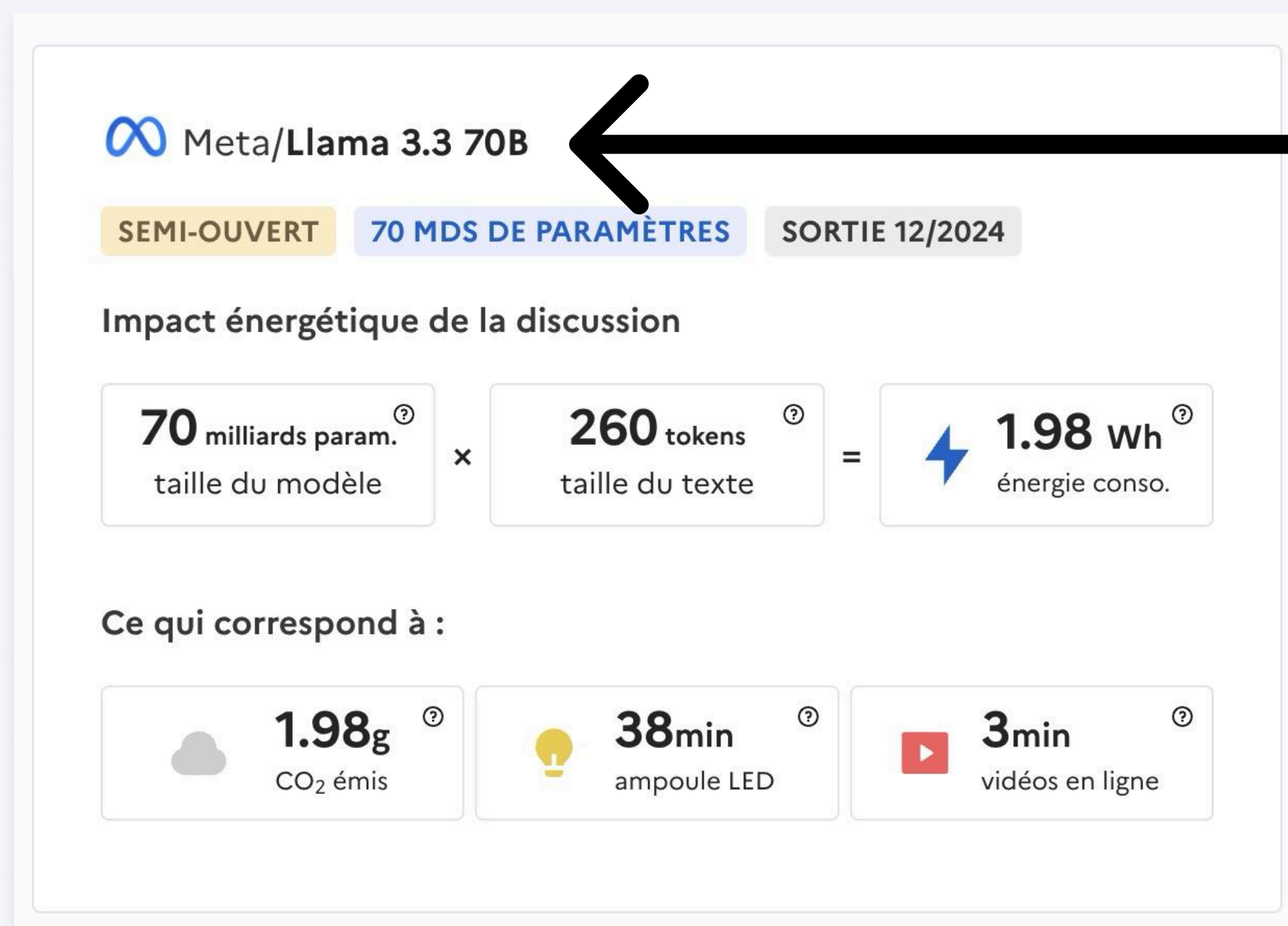
Plus un modèle est **grand**,  
plus il est "**intelligent**",  
mais plus il a besoin de **puissance de calcul**  
et donc plus il consomme de **l'énergie**.

## Quantité de **paramètres**

70B = 70 milliards de paramètres



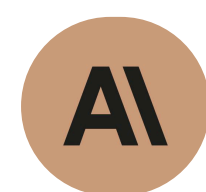
Meta/Llama 3.3 70B



# Qu'est-ce que la "taille" d'un modèle ?



Les duels de l'IA  
par compar:IA



Anthropic/Claude 3.5 Sonnet

→ **300** milliards de **paramètres** (env) → **XL**



Meta/Llama 3.3 70B

→ **70** milliards de **paramètres** → **L**



Mistral/Mistral Small

→ **24** milliards de **paramètres** → **M**



Google/Gemma 3 4b

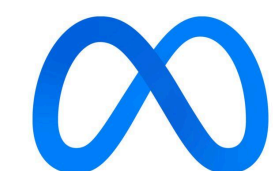
→ **4** milliards de **paramètres** → **XS**

# La distillation

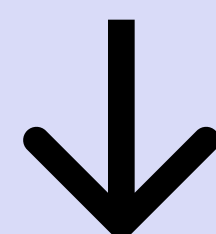


**Les duels de l'IA**  
par compar:IA

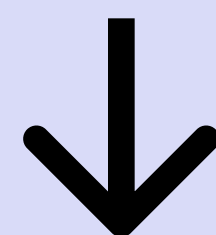
Famille Llama 3



Meta/Llama 3.1 405B



Meta/Llama 3.3 70B



Meta/Llama 3.3 8B

La distillation d'IA consiste à **entraîner un petit modèle en imitant un grand pour le rendre plus rapide et efficace.**

