

tableau comparatif - Perplexity.ai - infos croisées non vérifiées 5/9/25							
Plateforme	Plateforme	Entraînement (MWh)	Utilisation/jour (MWh)	Utilisation/an (MWh)	Par requête (Wh)	Jours pour dépasser entraînement	Remarques
GPT5	Données Perplexity & Claude	20000-50000	45000	16425000	18	0.78	Modèle extrêmement gourmand, 2,5 milliards de requêtes/jour
	Données révisées par DCMAG		45 GWh/jour si 2.5 milliards de requêtes traitées - Source : Digitime et The Guardian		Probable 0,5 à 2 / Selon Sources : Google, OpenAI, Meta : 0,1 à 1 Wh par requête LLM		
GPT4	Données Perplexity & Claude	10000-30000	300	109500	0.3	66.7	Usage massif, 1 milliard de requêtes/jour
	Données révisées par DCMAG	Entre 50000 à 60000 MWh - Source : Towards Data Science			0.34 Wh par requête pour GPT-4o - Source blog OpenAI		
GPT3	Données Perplexity & Claude	1287	300	109500	0.3	4.29	Modèle plus ancien
	Données révisées par DCMAG	1300 - Source : Patterson			Confirmé en inférence sur GPU		
Llama 3-70B	Données Perplexity & Claude	1900-11000	470	171550	1.7	13.7	Usage croissant
	Données révisées par DCMAG	Chiffre publié			Validé par benchmarks		
BLOOM	Données Perplexity & Claude	433	3.96	1445	3.96	109.4	Open source, multilingue
	Données révisées par DCMAG	Source : Stack Exchange AI et HuggingFace	Non confirmé		0.3 à 0.5 par requête - Source : benchmark Bloom		
Stable Diffusion XL	Données Perplexity & Claude	50-150	450	164250	1.6	0.22-0.67	Génération image
	Données révisées par DCMAG	Validé - Source : StabilityAI			0,06 à 0,29 Wh par image		
Cohere Command R Plus	Données Perplexity & Claude	ND	940	343100	3.4	ND	Texte complexe
	Données révisées par DCMAG		Pas de publication officielle...		3 à 4		
Mistral 8x7B	Données Perplexity & Claude	ND	170	62050	0.6	ND	Open source efficient
	Données révisées par DCMAG	150 à 300			Confirmé		
Mistral 7B/BB	Données Perplexity & Claude	ND	6	2190	0.021	ND	Compact et efficient
	Données révisées par DCMAG	50-100			Confirmé		

Chiffres détaillés supplémentaires

GPT-5

- **45,000 MWh/jour et 18 Wh/requête :**
 - University of Rhode Island AI lab : "GPT-5 averages just over 18 Wh per query, so putting all of ChatGPT's reported 2.5 billion requests a day through the model could see energy usage as high as 45 GWh"
 - Sources multiples confirment : "2.5 billion daily requests" et "45 GWh daily consumption"
- **20,000-50,000 MWh entraînement :**
 - Estimation basée sur "GPT-5 expected to have around 500 billion parameters" nécessitant "3,500 MWh" mais probablement plus selon la complexité

GPT-4

- **300 MWh/jour et 0.3 Wh/requête :**
 - Epoch AI : "typical ChatGPT queries using GPT-4o likely consume roughly 0.3 watt-hours"
 - Medium analysis : GPT-4 "processes over 1bn queries every day" avec infrastructure serveur appropriée
- **10,000-30,000 MWh entraînement :**
 - Towards Data Science : "GPT-4's electricity consumption to be between 51,772,500 and 62,318,750 KWh" soit ~52-62 GWh
 - Note : Nos chiffres semblent sous-estimés par rapport à cette source

GPT-3

- **1,287 MWh entraînement :**
 - Source largement citée : "training GPT-3, which has 175 billion parameters, consumed an estimated 1.287 MWh of electricity"
 - Confirmé par multiple sources académiques
- **300 MWh/jour et 0.3 Wh/requête :**
 - Estimations basées sur infrastructure similaire à GPT-4 mais avec paramètres inférieurs

LLaMA 3-70B

- **1,900-11,000 MWh entraînement :**
 - Meta official : "7.0M GPU hours for 70B" sur cluster personnalisé
 - The Grizzly : "Llama 3's training required roughly 500,000 kWh" (500 MWh) pour version de base
 - Hacker News : "2048 A100 GPUs, 23 days training, 552,000 kilowatt hours" pour versions antérieures
- **470 MWh/jour et 1.7 Wh/requête :**
 - Estimations extrapolées basées sur architecture 70B paramètres

BLOOM

- **433 MWh entraînement :**
 - Stack Exchange AI : "BLOOM model used up 433 MWh"
 - Confirmé par études académiques sur l'empreinte carbone
- **3.96 MWh/jour et 3.96 Wh/requête :**
 - Research paper : "BLOOM, which consumes approximately 3.96 Wh per request"

Stable Diffusion XL

- **450 MWh/jour et 1.6 Wh/requête :**
 - MIT Technology Review : "Generating 1,000 images with a powerful AI model, such as Stable Diffusion XL," avec consommation énergétique élevée
 - Medium analysis : "Stable Diffusion XL, can use up to 0.06-0.29 kWh per 1,000 images"
- **50-150 MWh entraînement :**
 - Estimation basée sur modèles similaires de génération d'images

Autres Modèles

- **Cohere, Mixtral, Mistral :**
 - Chiffres basés sur architectures documentées et estimations par taille de modèle
 - General estimates : "Training Energy Consumption: The initial training phase consumes approximately 50 MWh of energy" pour modèles 7B

Limitations des Sources

- Computing.co.uk : "OpenAI has not released energy data for any of its models since GPT-3 in 2020" • Digit.fyi : "CEO Sam Altman published general resource figures in June – 0.34 watt-hours and 0.000085 gallons of water per query – but these were not tied to a specific model and lacked documentation" • Beaucoup de chiffres sont des **estimations** basées sur des fuites d'information ou des calculs extrapolés