

IA et environnement

L'IA est une technologie à double tranchant. Elle offre des **opportunités pour accompagner la transition écologique** : optimisation de systèmes existants (ex : agriculture de précision), gestion intelligente de ressources et réseaux (ex : transport, eau, électricité, tri des déchets), prédiction et simulation de scénarios climatiques. Cependant le développement actuel des systèmes d'IA menace d'annuler ces bénéfices car le secteur a un **coût environnemental élevé**. Le numérique représente environ 4% des émissions de gaz à effet de serre (GES) dans le monde et l'IA une fraction croissante de ces émissions.

L'empreinte environnementale des systèmes d'IA

L'empreinte environnementale des systèmes d'IA doit se calculer en tenant compte de l'ensemble de leur **cycle de vie** : fabrication des infrastructures (datacenters, serveurs, processeurs graphiques), fonctionnement (réseau électrique, refroidissement), entraînement, utilisation, etc.

◆ Fabrication du matériel (hardware)

- > La production des serveurs et processeurs nécessaires au numérique et à l'IA s'appuie sur des **métaux** et **terres rares** (lithium, cobalt, tantale). Leur extraction provoque déforestation, pollution des sols et des eaux via l'usage de produits chimiques, une forte consommation d'eau et d'énergie, et s'accompagne d'enjeux géopolitiques et socio-économiques (violations des droits humains, destruction d'habitats, conflits...).

◆ Entraînement des modèles

- > L'entraînement des types d'IA est particulièrement coûteux en ressources. Concernant l'IAg, les grands modèles de langage (LLM), de plus en plus gros (nombre de paramètres croissants), demandent pour leur entraînement des **besoins en calcul et en énergie** toujours plus élevés (cf. fiche **Fonctionnement de ChatGPT**). On estime que l'entraînement de GPT-3 (ChatGPT) aurait consommé 1,287 GWh (équivalent à 552 tonnes CO₂).

552000 kg de CO₂ en France

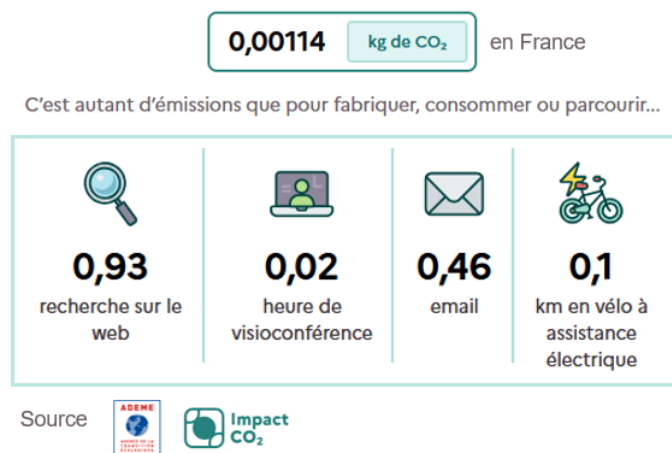
C'est autant d'émissions que pour fabriquer, consommer ou parcourir...



Source  

◆ Utilisation : focus particulier sur les grands modèles génératifs

- > L'utilisation des IA représente un coût énergétique non négligeable à grande échelle. L'impact de l'utilisation dépend de **plusieurs facteurs** : la taille du modèle (nombre de paramètres), la longueur du prompt, le contexte cumulatif ([RAG](#), historique de la conversation), la taille de la réponse produite (tokens, pixels, etc.), la nature de la tâche demandée (impact d'une vidéo > image > texte), etc.
- > Une page de texte (400 tokens) générée par le modèle Large 2 de Mistral AI consommerait en France 5 cL d'eau, et 5 Wh (équivalent à 1,14g de CO₂) *, mais **l'analyse d'un document de 200 pages ou la génération d'une image photoréaliste pourrait consommer dix fois plus**. * Le calcul d'équivalence Wattheures/CO₂ diffère selon le mix énergétique de chaque pays.



- > Des efforts visent à développer des modèles plus petits et spécialisés, moins gourmands en données et en paramètres. Cependant, ces optimisations peuvent accroître les usages et annuler ces bénéfices (**effets rebonds**).

◆ Infrastructures numériques et IA

- > Les **data centers** hébergent les données et les serveurs de calcul nécessaires au fonctionnement des IA. La consommation électrique de ces centres (actuellement 1,5% de la consommation mondiale) devrait "plus que doubler" d'ici 2030, selon un rapport de l'Agence internationale de l'énergie (AIE). Or, dans certains pays, l'électricité provient encore largement de sources fossiles, facteur de hausse des émissions de CO₂.
- > Des initiatives pour réduire l'empreinte écologique existent : construction de data centers en France (projet de Mistral AI), récupération de la chaleur des serveurs pour chauffer des habitations (comme le fait l'IAg Suisse [Eurìa](#), et le prévoit le projet bordelais [BXIA](#)).

Les effets rebonds liés aux usages de l'IA

- ◆ Effets directs : A mesure que les modèles d'IA se font légers et moins chers, leur intégration se généralise (dans les logiciels, les emails, les véhicules, etc.), ce qui annule les gains écologiques par une **explosion de leur utilisation**. La quête de la performance pousse à un remplacement prématuré des équipements (processeurs, serveurs, terminaux utilisateurs), produisant une augmentation de **déchets électroniques**, peu recyclés (1,2 à 5 millions de tonnes supplémentaires d'ici 2030).
- ◆ Effets indirects : L'IA est susceptible d'augmenter l'intensité ou la **profitabilité d'activités déjà largement émettrices**. Quelques exemples :
 - > Industrie textile (fast fashion) : certaines entreprises utilisent l'IA pour analyser les tendances, accélérer le processus de conception et adapter la production en temps réel. Cette optimisation peut soutenir un renouvellement plus rapide des collections dans une industrie déjà à l'origine d'environ 10 % des émissions mondiales de GES.
 - > Energies fossiles : les sociétés pétrolières exploitent l'IA pour mieux localiser les gisements, optimiser le forage et augmenter les rendements d'extraction, ce qui pourrait prolonger la rentabilité des ressources fossiles.
 - > Publicité ciblée : les plateformes numériques (dont [ChatGPT Ads](#)) affinent le ciblage et la personnalisation des annonces publicitaires, ce qui peut encourager la surconsommation.

L'IA au service de la transition écologique

L'IA peut devenir un levier pour la crise écologique, à condition qu'elle soit dirigée vers des usages pertinents et sobres. Selon un [rapport du Sénat](#) de 2025 recense de nombreuses applications et possibilités de l'IA au service de la transition écologique :

- > **Mieux appréhender l'environnement, améliorer la prise de décision** (analyse de données satellites et de capteurs, modélisation des écosystèmes, suivi de la biodiversité)
- > **Prévoir et s'adapter aux crises** (prévisions météorologiques, scénarios climatiques)
- > **Diminuer les émissions** en optimisant les réseaux et les systèmes énergétiques (chauffage, climatisation, éclairage), tout comme les transports, l'irrigation et les processus industriels.

Le Gouvernement s'est doté d'une [feuille de route consacrée à l'IA et à la transition écologique](#), intégrant à la fois le développement d'une **IA frugale** (réduction de son empreinte environnementale) et l'encouragement d'**usages orientés vers la préservation** des ressources, du climat et de la biodiversité.

Il convient d'aborder l'évaluation de l'impact environnemental de l'IA avec nuance, sans perdre de vue d'autres pratiques de consommation courantes plus polluantes et déjà non soutenables. Cependant la massification des usages s'ajoute à nos habitudes et pressurise toujours plus des réseaux locaux d'eau et d'électricité en tension.

Bonnes pratiques d'usage des systèmes d'IAg

◆ Questionner la nécessité :

- > En ai-je vraiment besoin ? Y a-t-il des alternatives ? (moteur de recherche, Wikipedia, banque d'images ou de vidéos). Est-ce que cela me fait gagner beaucoup de temps (ex : une heure de gagnée sur un logiciel gourmand en énergie comme Photoshop). Privilégier l'utilisation de systèmes d'IAg lorsque la **valeur ajoutée** est avérée.

◆ Optimiser ses prompts et conversations

- > Être **précis** dans son prompt pour obtenir une réponse satisfaisante du premier coup.
- > Être **concis** pour réduire la puissance de calcul (éviter les mots inutiles (ex : "~~P~~eux-tu donner", "~~s'il te plaît~~", "~~merci~~"). Voir cet [outil d'optimisation de prompts](#).
- > **Grouper** ses questions dans un même prompt.
- > Certains verbes (ex : analyser, expliquer, justifier, créer, recommander) incitent le modèle à être plus bavard : plutôt que de demander "Analyse ce texte", préférer "Classe les idées principales de ce texte" ou "Identifie les points clés".
- > Demander des réponses concises ou sous forme de liste.
- > Arrêter la conversation (bouton stop) si l'IAg hallucine ou ne part pas dans la direction souhaitée.
- > Mesurer s'il est plus intéressant de créer un assistant pré-configurés (par ex. GPT, Gem) ou un flux d'agents pour des tâches récurrentes.
- > Faire attention aux usages « récréatifs » ou aux [compagnons IA](#).
- > Éviter les modes "Raisonnement" si non nécessaires. Ces modèles génèrent des milliers de "tokens de pensée" (matérialisés en gris dans l'exemple ci-dessous) qui augmentent drastiquement la consommation. Exemple avec l'IAg Deepseek :

Qu'est ce qui commence la nuit et termine le matin ?



☒ Réfléchi pendant 8 secondes ▾

- Hmm, l'utilisateur pose une devinette en français. C'est une question simple mais qui demande une réflexion sur le sens des mots. La structure "commence la nuit et termine le matin" suggère qu'il s'agit d'un jeu de mots ou d'une expression courante.

Je pense immédiatement au mot "Nuit" lui-même - il commence par la lettre "N" et se termine par "T", mais cela ne correspond pas vraiment à la formulation. Ah, peut-être qu'il s'agit du mot "matinée"? Non, ça ne colle pas non plus.

Attendez, si on prend la phrase littéralement : quelque chose qui commence pendant la nuit et finit le matin... Ça pourrait être le sommeil, un rêve, ou une activité nocturne. Mais comme c'est présenté comme une devinette, il doit y avoir un double sens linguistique.

Oh, je vois! La solution classique à cette devinette est effectivement le mot "N" - première lettre de "nuit" et dernière lettre de "matin". C'est un jeu de lettres simple mais intelligent. L'utilisateur cherche probablement cette réponse traditionnelle.

Cette devinette classique a pour réponse : **la lettre N**.

En effet :

- Le mot **nuit** commence par la lettre **N**.
- Le mot **matin** se termine par la lettre **N**.

◆ Choisir le plus petit modèle adapté

- > Si l'outil d'IAg le permet, utiliser le plus petit modèle capable de réussir la tâche. Plus un LLM est grand (en milliards de paramètres = variables apprises), plus il lui faut d'énergie pour produire un token. **Un gros modèle peut avoir une consommation énergétique des milliers de fois plus importante** qu'un petit modèle spécialisé. Par exemple, l'outil [DuckAI](#) de DuckDuckGo permet de sélectionner de plus petits modèles et de personnaliser la longueur des réponses. Les plateformes [Compar:IA](#) ou [Vittascience](#) permettent de comparer l'empreinte environnementale de différents LLM et choisir des modèles plus frugaux.

◆ Limiter l'usage du RAG

- > La technologie [RAG](#) (Retrieval-Augmented Generation) combine la récupération d'informations (provenant de documents chargés par l'utilisateur ou de ressources internet récupérée par l'IA), le traitement de ces contenus et la génération de texte, ce qui le rend plus gourmand qu'une requête simple sans RAG à modèle de langage égal. Son efficacité en termes de gain de temps et de pertinence peut compenser cet impact lorsqu'elle est utilisée de façon ciblée.
- > Réserver le mode RAG aux questions nécessitant réellement une **synthèse sourcée** et non pour une simple requête de navigation.
- > **Limiter la quantité** de documents ou de passages importés aux informations strictement indispensables afin de réduire le nombre de tokens d'entrée et donc la consommation.

◆ Prendre en compte les limites de l'IA embarquée selon le matériel

- > L'IAg "embarquée" ou "en local" consiste à faire tourner des modèles plus petits (Small Language Models - SLM) directement sur l'appareil de l'utilisateur (ordinateur, smartphone), sans envoyer de données vers un data center. Or une étude montre que **le matériel de l'utilisateur compte** : s'il est ancien ou non optimisé, faire tourner un modèle de langage sur le processeur de l'utilisateur peut consommer jusqu'à 29 fois plus d'énergie qu'un appel vers un modèle distant de type ChatGPT.

En conclusion, le développement et l'usage de l'IA exigent une **approche responsable**, soutenue par des **efforts de transition énergétique** pour ne pas aggraver des modèles de consommation déjà insoutenables. Son **usage doit être ciblé, proportionné**, et privilégié aux cas où le **bénéfice net est avéré**. Enfin, l'accès à des données transparentes est un indispensable pour évaluer et réguler efficacement l'empreinte réelle de ces technologies.

Sources :

- Adamska, M., et al. (2025). *Green prompting*. arXiv. <https://doi.org/10.48550/arXiv.2503.10666>
- Café IA. (2025). *IA & environnement*. <https://cafeia.org/wp-content/uploads/2025/06/IA-Environnement-Module-Cafe-IA-V1.0-28-mai-2025.pdf>
- Cows, J. et al. (2023). The AI gambit: Leveraging AI to combat climate change. *AI & Soc*, 38(1), 283-307. <https://doi.org/10.1007/s00146-021-01294-x>
- LeMagIT. (2026). *L'AI Energy Score affiche un écart énergétique colossal entre les modèles d'IA*. <https://www.lemagit.fr/actualites/366619735/LAI-Energy-Score-affiche-un-ecart-energetique-colossal-entre-les-modeles-dIA>
- Greenspector. (2025). *What is the environmental impact of local AI on our smartphones?* <https://greenspector.com/en/artificial-intelligence-smartphone-autonomy/>
- Le Monde. (2025). *Dopée par l'IA, la demande d'électricité des centres de données doublera d'ici 2030*. [lemonde.fr/article/6593594](https://www.lemonde.fr/article/6593594)
- Luccioni, A. S., et al. (2025). From efficiency gains to rebound effects. *Proc. FAccT '25*, 76–88. <https://doi.org/10.1145/3715275.3732007>
- Mister IA. (2025). *Consommation énergétique de l'intelligence artificielle en 2025*. <https://www.mister-ia.com/article/ia-consommation-energetique>
- Mistral AI. (2025). *Notre contribution pour la création d'un standard environnemental mondial pour l'IA*. <https://mistral.ai/fr/news/our-contribution-to-a-global-environmental-standard-for-ai>
- You, J. (2025). *How much energy does ChatGPT use?* Epoch AI. <https://epoch.ai/gradient-updates/how-much-energy-does-chatgpt-use>