L'IA générative peut-elle être écologiquement responsable ?

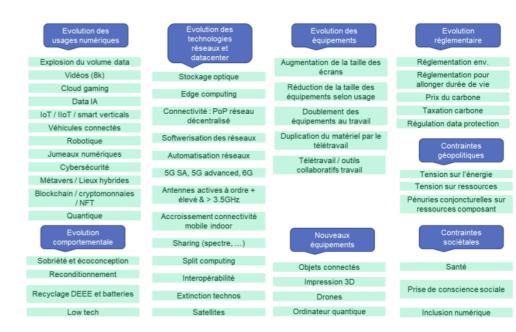
Midisciences 2025 Pierre Jourlin

Les gaz à effet de serre (GES)

- Vapeur d'eau (H2O)
- *《* Dioxyde de carbone (CO2)
- d Méthane (CH4)
- **†** Tétra-fluorométhane (CF4)
- Mexafluorure de soufre (SF6)

Quelle part de l'IA dans les émissions de GES ?

Numérique = 3–4 % des GES mondiaux 2.5 % en France Source : ARCEP-ADEME



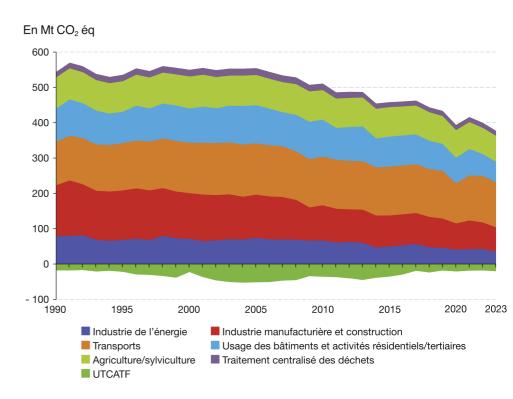
Répartition par secteurs

en France

source: [Datalab2024]

émissions de GES en France ⊂ émissions de GES induites par la France

objectif EU: -55 % net en 2030 par rapport à 1990 objectif de réduction de 5 % par an entre 2022 et 2030.



Situation mondiale

- En moyenne : 4 tonnes de CO2e / an / personne
- Mais très grande inégalités :
- 【 Ces émissions ≈ 1,3 million de décès supplémentaires liés à la chaleur
- Resultation x7 dans pays inégalitaires

Source: [Oxfam2023]

Compte carbone?

France

- Départ : 9 tonnes CO2e/an/personne (~160 MWh/an/personne)
- Objectif : Réduction de 6% par an
- Problème : comment suivre la consommation par personne ? (ex. datacenters)

→ Rappel: 1 kWh = 1 heure d'un radiateur 1kW



Intensité carbone de l'électricité [Ember2023]

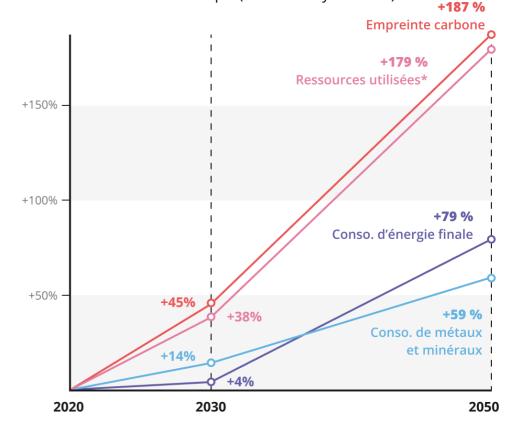
Pays	Kg CO2e / kWh
 France	0,056
USA	0,369
Allemagne	0,381
Russie	0,441
C Ouzbékistan	1,168

Projection 2030, 2040, 2050

- Numérique = 3–4 % des GES mondiaux
- 2,5 % en France
- +60 % prévu d'ici 2040
- Source : ARCEP-ADEME

Sans actions pour la réduire, l'empreinte carbone pourrait presque tripler en 2050, la consommation d'énergie doubler

Evolution du scénario tendanciel de 4 indicateurs de l'impact environnemental du numérique (sur tout le cycle de vie).

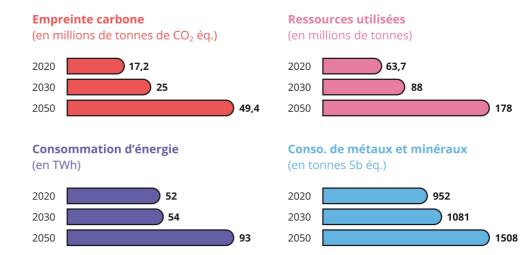


^{*} Définition MIPS prenant en compte les matériaux utilisés, la biomasse, les déplacements de terre mécaniques ou par érosion, l'eau, et l'air.

Projection d'indicateurs

Des indicateurs issus d'une méthode normalisée d'analyse de cycle de vie qui comporte des définitions précises

Evolution de 4 indicateurs de l'impact environnemental du numérique dans le scénario tendanciel, en valeurs absolues.



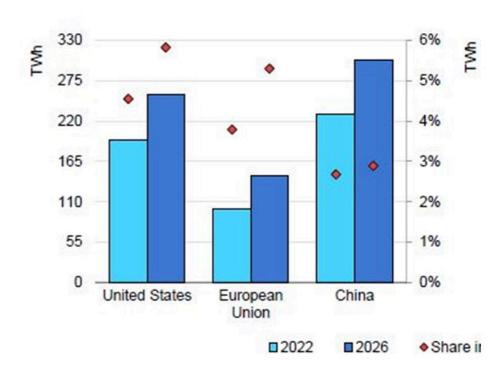
Empreinte carbone: émissions de gaz à effets de serre exprimées en équivalent CO₂.

Ressources utilisées: indicateur MIPS qui considère cinq types de ressources, comprenant les ressources abiotiques (matériaux, énergie fossile...), la biomasse, les déplacements de terre mécaniques ou par érosion, l'eau, et l'air. Il donne une idée de l'effort effectué pour produire nos biens et services.

Consommation d'énergie finale : désigne l'énergie directement utilisée par l'utilisateur final, sous forme d'électricité ou de carburant.

Consommation de métaux et minéraux : cet indicateur évalue la quantité de ressources minérales et métalliques extraites de la nature en équivalent antimoine (un élément chimique dont on retrouve le symbole Sb dans le tableau périodique des éléments). C'est un standard des analyses de cycle de vie qui permet de mesurer l'épuisement des ressources naturelles.

Consommation centres de données Source [IEA2024]



Al Générative & environnement

[Zewe2025]

- Impact écologique difficile à évaluer
- Impact écologique difficile à atténuer
- Impact Carbone :
 - Consommation électrique
 - émissions de GES
 - f pressions sur le réseau électrique.

[Chen2025] [Strubell2019]

- Fabrication matériel informatique
- Extraction = pollution + produits chimiques
- 2023 : 3,85M GPU | 2024 : 2,67M GPU
- Prévision : +320% sur 5 ans [Muthoni2024]
- → épuisement ressources
- → pertubation écosystème
- Transport matériaux et matériel

- Pré-entraînement IA = 300 T CO2e / modèle
 - → ≈ 125 vols A/R NY ↔ Pékin
 - \rightarrow ≈ empreinte carbone d'une vie (≈75 ans)
- Réglage fin
- Renforcement (RLHF)
- Inférence : $coût \times nombre$ d'utilisateurs
- Impact ressources en eau :
 - CPU/GPU/TPU → +
 - **8** Refroidissement = consommation d'eau
 - 2 L / kWh [Bashir2024]

- Data centers (Centres de Données, CD)
- 1940 : ENIAC, premier ordinateur numérique à usage général. Université de Pennsylvanie.
- \approx 10.000 CD dans le monde :
- 1 CD "hyperscale" (AWS, Google, Azure) entre
 100k et 1M serveurs.
- 1 serveur = plusieurs gros CPU, GPU, NPU

- IA → Accélération du rythme de construction
- CD échelle mondiale :
 - 460 térawatts en 2022.
 - 11e plus grand consommateur d'électricité au monde
 - Arabie saoudite (371 térawatts)
 - France (463 térawatts) source : OCDE
 - CD mondial, 2026 : 1 050 térawatts (5e place entre le Japon et la Russie)
- 1 milliard de radiateurs (~8 milliards d'individus)

Noman Bashir, chercheur en informatique et impact climatique au MIT Climate and Sustainability Consortium (MCSC) et postdoctorant au Computer Science and Artificial Intelligence Laboratory (CSAIL)

La demande de nouveaux centres de données ne peut être satisfaite de manière durable. Le rythme auquel les entreprises construisent de nouveaux centres de données signifie que la majeure partie de l'électricité nécessaire à leur alimentation doit provenir de centrales à combustibles fossiles. [Bashir2024]

Exemple : entraı̂nement GPT-3 → 1 287 mégawattheures

- = 120 foyers américains moyens pendant un an
- = 552 tonnes de dioxyde de carbone = 123 véhicules diesels pour 1 an [Patterson2021]

Entrainement IA => Fluctuation rapide de la consommation d'énergie => générateurs diesels

1 requête ChatGPT consomme environ cinq fois plus d'électricité qu'une simple recherche sur le Web. [Bashir2024]

Faible durée de vie des LLM (+d'1 modèle par semaine)

Consommation globale CD

- Amérique du Nord
 - 2022:2688 MW → 2023:5341 MW
- Monde
 - 2022 : 460 TWh (11e consommateur mondial)
 - 2026 : 1 050 TWh (≈ 5e place)

Anatomy of an Al System [Crawford2018] (Amazon echo)

X Extraction

- Cobalt, lithium
- Travail précaire (ex : RDC)

Travail invisible

- Assemblage, étiquetage, modération
- Humains essentiels mais invisibilisés

📡 Données

- Collecte continue des assistants vocaux
- Questions éthiques : consentement, vie privée

Infrastructure

- Centres de données = énorme consommation
- Réseaux + Cloud = empreinte massive



- Obsolescence rapide
- Déchets exportés, impact environnemental + sanitaire

Pouvoir & asymétrie

- GAFAM = contrôle des données + infrastructures
- Manque de transparence, déséquilibre

Que ne pas faire

- X IA générative pour des futilités
- Étre poli avec une IA
- X LLM vs algo classique (plus lent, plus coûteux)
- Attention au piège "productivité = IA"

Que faire?

- Recherche indépendante
- Recherche d'intérêt public
- Recherche bas carbone
- Régulation (ARCEP, UE, etc.)
- Mobilisation citoyenne [Techworkers2020]
- Intelligence collective

Recherche

- Distillation (réduire le nb de paramètres)
- Quantisation (réduire la taille des paramètres)
- Modèles "ouverts" (permet installation locale)

Optimiser l'usage

- Finesse de gravure -> Effet rebond
- Modèles locaux, légers :
 - souveraineté individuelle!
 - confidentialité
 - Martin Mart
 - <u>M</u> Benchmarks Modèles légers [HuggingFace]
 - Article [Dini2025]
 - Financements?

Ressources et liens

- Zewe2025
- def2
- Crawford2018
- Dhar2020
- def5
- Bashir2024
- Patterson2021
- Muthoni2024

- Ember2023
- IEA2024
- Datalab2024
- Strubell2019
- Techworkers2020
- Chen2025
- Oxfam2023
- Siblini2025
- HuggingFace
- Dini2025