

Trystram, D., & Ménessier, T. (2024, mai 13). *L'IA peut-elle vraiment être frugale ?* The Conversation. <http://theconversation.com/lia-peut-elle-vraiment-etre-frugale-226274>

Academic rigour, journalistic flair

# L'IA peut-elle vraiment être frugale ?

Published: May 13, 2024 4.51pm CEST

## Authors



1. [Denis Trystram](#)

Professeur des universités en informatique, Université Grenoble Alpes (UGA)



2. [Thierry Ménessier](#)

Professeur de philosophie politique, Université Grenoble Alpes (UGA)

## Disclosure statement

Denis Trystram participe à l'écriture d'une Spec AFNOR pour encadrer l'IA frugale. Le présent texte a bénéficié de la dynamique de ce groupe mais n'engage que son auteur.

Thierry Ménessier a reçu des financements de l'Agence Nationale de la Recherche dans le cadre du plan de financement France 2030 pour l'institut pluridisciplinaire d'intelligence artificielle de Grenoble, MIAI.

## Partners

[Université Grenoble Alpes \(UGA\)](#) provides funding as a founding partner of The Conversation FR.

[View all partners](#)

Si l'IA peut avoir des applications vertueuses pour l'environnement, cela vaut-il tout le carbone émis lors de son développement et de son usage ? Kak (France)/L'Opinion-Cartooning for Peace, Author provided (no reuse)

Autour de nous, nous observons le numérique envahir tous les autres secteurs. L'[intelligence artificielle](#) (IA) est l'un des derniers maillons de ce bouleversement technologique : elle accompagne désormais tout traitement automatique qui exploite le déluge de données numériques. Mais au vu des enjeux écologiques auxquels nous faisons face aujourd'hui, sera-t-il possible de concevoir une IA respectueuse des contraintes environnementales ?

Avant de rentrer dans le sujet de l'IA frugale, il convient de poser le décor. La crise climatique sans précédent à laquelle nous faisons face a commencé avec la révolution industrielle, au milieu du XIX<sup>e</sup> siècle, qui a planté les germes de notre société de consommation actuelle. Le [changement climatique](#) n'est pas la seule menace environnementale : stress hydrique, épuisement des ressources, perte de la biodiversité... Mais c'est sans doute la plus visible et la plus documentée, et donc celle qui peut nous aider à mieux appréhender les autres.

## Un secteur qui croît toujours plus vite

Le secteur du numérique n'est pas facile à cerner, car [il est dilué partout](#). Selon l'ADEME, il représente [2,5 % des émissions carbone de la France en 2022](#). Ces dernières années, le [domaine a connu une forte croissance](#) et les [études prospectives](#) envisagent principalement des scénarios de poursuite de cette croissance, au moins à moyen terme.

### A weekly email for Europeans by European scholars

Un petit calcul réalisé à partir de données publiques sur le scénario SSP1-19 du GIEC, un des plus optimistes, souligne l'aberration de cette croissance. Si le secteur croît selon la prévision la plus basse de croissance, le numérique émettrait 6 fois plus que l'objectif du scénario de décroissance des [émissions mondiales de CO<sub>2</sub>](#) d'ici à 2050 ! Même si la croissance du secteur stagnait au niveau d'aujourd'hui, il représenterait trois quarts des émissions totales... Dans un tel monde, que nous resterait-il pour le reste ?

Si on se focalise sur l'IA, on observe une rupture claire à partir de 2012. La croissance du secteur s'emballe alors avec un doublement des besoins en puissance de calcul tous les 5-6 mois au lieu de 24 mois, chiffre jusqu'alors stable de la classique [loi empirique de Moore](#). Cette date correspond au développement des modèles d'IA reposant sur l'apprentissage profond, ou deep learning, rendus possibles par l'utilisation de processeurs graphiques (GPU) pour effectuer les calculs à la base de l'apprentissage profond et par le développement des données ouvertes sur Internet. Rappelons que l'IA n'est pas réduite à l'apprentissage par réseaux de neurones profonds, mais ce sont incontestablement ces derniers qui sont les plus gourmands.

Un nouveau palier a été atteint en 2023, avec l'explosion des modèles génératifs comme l'agent conversationnel ChatGPT. Même s'il est difficile d'avancer des chiffres précis, étant donné que les « géants de la tech » comme OpenAI, Meta ou Microsoft qui sont à l'origine des plus gros modèles ne communiquent plus sur ces données, cette diffusion à large échelle est très inquiétante.

## **Le poids de l'IA générative sur le climat**

ChatGPT est basé sur le modèle GPT-3, remplacé aujourd'hui par une version améliorée GPT-4. Ce n'est pas le seul, mais c'est le plus populaire et un de ceux pour lequel il existe des données. Le modèle sur lequel il s'appuie possède 176 milliards de paramètres et a nécessité 552 tonnes d'équivalent CO<sub>2</sub> pour son entraînement en Californie. En termes de consommation électrique (indicateur plus objectif au sens où il ne dépend pas du mix énergétique), le modèle a tourné des jours sur [près de 4 000 gros GPU de Nvidia dont la consommation a été estimée à 1 283 MWh](#) (megawatt-heure, soit 1 000 kWh).

La phase d'usage est bien plus consommatrice encore ! Chaque jour, les [quelque dix millions d'utilisateurs mobilisent 564 MWh d'électricité](#). Les annonces récentes des patrons d'OpenAI et Microsoft sur des commandes de centaines de milliers de GPU pour alimenter les futures versions sont vertigineuses en termes de consommation et d'impact environnemental. Avec sa capacité de production actuelle, le constructeur Nvidia est loin de pouvoir en produire autant.

ChatGPT n'est que l'élément visible de cette galaxie. Aujourd'hui, l'IA est un moteur de la croissance exponentielle du secteur du numérique, avec une explosion du nombre d'applications et services qui utilisent l'IA générative. Le développement de l'IA à ce rythme n'est bien entendu pas soutenable tel quel.

## **Comment penser une IA plus frugale ?**

On ne pourra soutenir cette croissance que si l'IA permet des économies d'émissions considérables dans tous les autres secteurs. C'est la voix majoritaire qui porte le message d'une IA qui va nous aider à sortir de la crise. Malgré de trop nombreuses applications inutiles ou questionnables, il existe des apports bénéfiques pour la société notamment pour simuler et analyser des phénomènes physiques complexes comme [l'étude de scénarios pour contrer la crise climatique](#). Encore faut-il que ces solutions ne soient pas *in fine* pires que le mal ! Par exemple, l'IA va permettre aux entreprises exploitant les énergies fossiles d'optimiser leur activité et donc [d'émettre encore plus de CO<sub>2</sub>](#).

Partout, on entend parler d'IA frugale sans que ce terme soit clairement défini. Dans le langage usuel, la sobriété est souvent entendue comme la réaction adéquate face à une consommation abusive d'alcool. Dans le contexte de l'IA, cela renvoie plutôt à la simplicité (ce qui est clairement insuffisant ici), à la modération, voire l'abstinence. Frugalité et sobriété sont souvent considérées comme synonymes ; il est également possible de considérer que la frugalité concerne le fonctionnement des systèmes techniques tandis que la sobriété renvoie à leur usage dans le cadre des pratiques sociales.

Les deux dimensions se complètent dans le sens où tout système technique s'adresse à des usages qui se trouvent de la sorte facilités et encouragés. Ainsi, plus le système apparaît propice à

l'usage, plus son impact s'accroît : c'est ce que l'on appelle l'effet rebond. Cependant, le plus pertinent est la définition en creux : le contraire de la frugalité est ainsi qualifié de glotonnerie selon *Le Robert*. Il est donc possible de considérer que la frugalité-sobriété comme une vertu qui s'apprécie en négative, en fonction de la quantité de ressources que l'on ne consomme pas.

Or, caractériser une IA frugale s'avère difficile pour plusieurs raisons. D'une part, les analyses existantes ciblent souvent l'entraînement des modèles et/ou la phase d'usage, mais ignorent le cycle de vie complet du service ou du produit. Cela inclut la production, l'utilisation et le stockage des données, et l'infrastructure matérielle mise en œuvre, depuis la fabrication jusqu'à la fin de vie de tous les équipements impliqués. D'autre part, pour un service reconnu comme utile pour la société, il conviendrait d'estimer les volumes de données impliquées dans le processus et les effets positifs indirects induits par son déploiement. Par exemple, un système d'optimisation énergétique pour un appartement peut permettre une augmentation de confort ou le déploiement de nouveaux services grâce aux économies réalisées.

## **Mettre l'IA au régime, une démarche nécessaire, mais insuffisante**

Aujourd'hui, les termes de frugalité ou de sobriété sont souvent synonymes d'efficacité énergétique : on imagine et développe une solution sans prendre en compte son coût environnemental, puis on l'améliore de ce point de vue dans un second temps. Il faudrait au contraire s'interroger en amont sur les effets avant le déploiement du service, quitte à y renoncer.

L'IA frugale est donc caractérisée par une contradiction intrinsèque, au vu de la boulimie d'énergie et de données aujourd'hui nécessaire à l'entraînement des gros modèles et à leurs usages, au mépris des risques considérables pour l'environnement. En matière d'IA, la frugalité doit aller bien plus loin que la simple efficacité : elle doit d'abord être compatible avec les limites planétaires. Elle doit aussi interroger les usages en amont, jusqu'au renoncement de certains services et pratiques, en se basant sur des analyses de cycle de vie complètes et rigoureuses.

Les finalités que recouvrent ces développements technologiques devraient au moins être collectivement débattues. Derrière l'argument d'une efficacité accrue se cache la compétition entre souverainetés nationales ou la concurrence entre des firmes intéressées par des profits colossaux. Il n'y a rien dans ces finalités qui ne soit considéré à l'aune d'une approche éthique.

Une évaluation des systèmes d'algorithmes à l'aide des [éthiques environnementales contemporaines](#) permet même de fonder la notion de sobriété sur d'autres bases. En effet, et en dépit de [leur variété](#), ces éthiques ne considèrent pas la Nature (l'eau, l'air, les matériaux et les vivants) comme des ressources à disposition de la seule espèce humaine, engagée dans la compétition technologique et l'hédonisme industriel. En conclusion, on pourrait affirmer que s'ouvre aujourd'hui pour la recherche responsable en IA une perspective aussi formidable que difficile à réaliser : proposer des modèles et des systèmes les plus compatibles possibles avec une telle définition « forte » de la sobriété.

We have a weekly email, which I help curate. It brings you a selection of the latest news, views and analysis, straight from the world's leading experts. Sign up and we'll deliver it to your inbox every Thursday.

Claudia Lorenzo Rubiera  
Culture Editor