



MINISTÈRE
DE LA TRANSITION
ÉCOLOGIQUE,
DE LA BIODIVERSITÉ,
DE LA FORÊT, DE LA MER
ET DE LA PÊCHE

Liberté
Égalité
Fraternité

Inria

DOCUMENT DE SYNTHÈSE

Les principaux défis à relever pour **favoriser** **la performance** **environnementale** **de l'IA**



SOMMET
POUR **L'ACTION**
SUR **L'IA**

Février 2025

Préambule

Le développement de l'intelligence artificielle (IA) est rapide et hétérogène, tant en termes de méthodologie que d'applications. Il est fort probable que le nombre de domaines qui ne sont pas touchés par l'IA soit largement inférieur à ceux que cette nouvelle technologie est en train de transformer.

Dans le présent document, l'IA est considérée comme un domaine théorique et pratique dont l'objectif est de comprendre les mécanismes de la cognition et de la réflexion et de les imiter à l'aide d'ordinateurs et de logiciels. Selon la loi européenne sur l'IA (article 3), un système d'intelligence artificielle est un système basé sur une machine qui est conçu pour fonctionner avec différents niveaux d'autonomie. Il peut faire preuve d'adaptabilité après son déploiement et, pour des objectifs explicites ou implicites, il déduit à partir des données qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer des environnements physiques ou virtuels. Ainsi, l'IA est principalement considérée du point de vue de l'infrastructure matérielle et des algorithmes, qui, après une phase d'apprentissage pouvant nécessiter une très grande quantité de données, permettent d'effectuer diverses tâches, en particulier :

- **Percevoir l'environnement et prédire son évolution ;**
- **Traiter l'information**, y compris par l'analyse, l'indexation et l'extraction de connaissances ;
- **Prendre des décisions, choisir des actions** et exécuter des tâches pour atteindre des objectifs spécifiques.

Dans le cadre du sommet pour l'action sur l'intelligence artificielle, un grand nombre de questions liées au développement de l'IA sera abordé. L'objectif est de permettre à chacun d'identifier les risques associés à cette technologie, ainsi que ses opportunités et ses avantages. Cela nous permettra de tracer une voie qui soit conforme à l'intérêt public et à la diversité des points de vue, pour l'évolution et le développement de l'IA. Ce document s'intéresse aux impacts environnementaux¹ de l'IA et plus spécifiquement aux composants matériels et logiciels sur lesquels les outils d'IA sont basés, ainsi qu'à la conception de ces outils. Plus précisément, l'objectif de ce document est de présenter cinq défis clés à relever pour favoriser la performance environnementale de l'IA. Ce document ne fournit pas une liste exhaustive des défis. Il est conçu pour évoluer au fil du temps afin d'intégrer les progrès et les défis émergents dans le cadre de processus de consultation similaires.

S'appuyant sur les récents développements de la technologie de l'IA, plusieurs défis ont été mis en avant par un large panel de scientifiques, de chefs d'entreprise, d'organisations internationales, d'autorités administratives et d'autres parties prenantes. Nous visons à créer une coalition de parties prenantes prêtes à relever ces défis afin de maximiser les impacts positifs de l'IA, notamment en termes environnementaux, tout en minimisant l'empreinte environnementale du déploiement de solutions basées sur l'IA.

Pour relever chaque défi, les contributeurs ont cherché à établir des objectifs collectifs ambitieux basés sur des découvertes et des développements récents. Cet effort a cependant été compliqué par le manque de données précises sur l'empreinte environnementale de l'IA. Pour palier ce problème, certains des défis présentés dans ce document visent à combler le manque de données disponibles.

¹ Selon le règlement de l'UE établissant un cadre pour la définition d'exigences en matière d'écoconception applicables aux produits durables, on entend par impact environnemental toute modification de l'environnement, qu'elle soit négative ou bénéfique, résultant totalement ou partiellement d'un produit au cours de son cycle de vie.

Introduction

et conclusions

Bien que reposant parfois sur des méthodes algorithmiques utilisées depuis des années, **le développement de l'IA progresse aujourd'hui à un rythme très rapide**, suscitant de nombreuses controverses qui soulèvent à la fois des opportunités et des inquiétudes. La contribution de l'IA à la protection de l'environnement et au changement climatique fait partie de ces controverses. D'une part, **l'IA est une technologie puissante qui peut nous aider à mieux comprendre les phénomènes géophysiques**, à affiner les prévisions en matière de changement climatique et à soutenir la décarbonation de divers secteurs tels que l'agriculture, l'industrie, la mobilité et l'énergie, pour ne citer que quelques exemples. D'autre part, le développement et le fonctionnement des systèmes d'IA, en particulier à grande échelle, ont une incidence négative importante sur l'environnement, en raison notamment **d'une consommation élevée d'eau, d'électricité et de ressources, ainsi que du faible taux de recyclage des composants**.

Bien que les technologies numériques ne représentent actuellement que 3 % environ des émissions mondiales de gaz à effet de serre², ce chiffre est en forte augmentation. Parallèlement, **les technologies numériques représentent déjà jusqu'à 12 % de la consommation mondiale d'électricité**, et une augmentation importante de leur empreinte énergétique dans les années à venir pourrait compromettre les objectifs de production d'électricité liés à la transition écologique. À l'heure actuelle, le rythme du développement de l'IA dépasse largement celui de la capacité de production d'électricité à partir de sources d'énergie renouvelables telles que les panneaux photovoltaïques ou les éoliennes³. Dans les grandes économies telles que les États-Unis, la Chine et l'Union européenne, les centres de données ne représentent aujourd'hui que 2 à 4 % environ de la consommation totale d'électricité. Toutefois, en raison de leur plus grande concentration spatiale par rapport à d'autres infrastructures à forte intensité énergétique, leur impact local peut être significatif. Par exemple, le secteur a déjà dépassé 10 % de la consommation d'électricité dans au moins cinq États américains. En Irlande, il représente désormais plus de 20 % de la consommation totale d'électricité⁴.

L'augmentation de l'utilisation de l'IA entraîne une augmentation des besoins informatiques correspondants, qui est perceptible depuis plus de cinq ans, les demandes d'entraînement de modèles étant multipliées par 15 tous les deux ans. Cette tendance s'est encore accélérée au cours des deux dernières années, sous l'effet du développement rapide de l'IA générative, les besoins en énergie pour l'entraînement ayant été multipliés par 750 au cours de cette période, et les coûts environnementaux associés à l'entraînement de grands modèles étant particulièrement élevés (voir les figures 1-a et 1-b). Il convient toutefois de noter qu'**un découplage récent est apparu entre la croissance des FLOP⁵ et la taille des modèles ou des ensembles de données d'apprentissage⁶**.

² Rapport sur l'économie numérique 2024 de la CNUCED

³ Goldman Sachs AI, data centers and the coming US power demand surge 2024

⁴ AIE Ce que le boom des centres de données et de l'IA pourrait signifier pour le secteur de l'énergie, 2024

⁵ Opérations en virgule flottante (FLOP)

⁶ J. Hoffmann et al. [Training Compute-Optimal Large Language Models 2023](#)

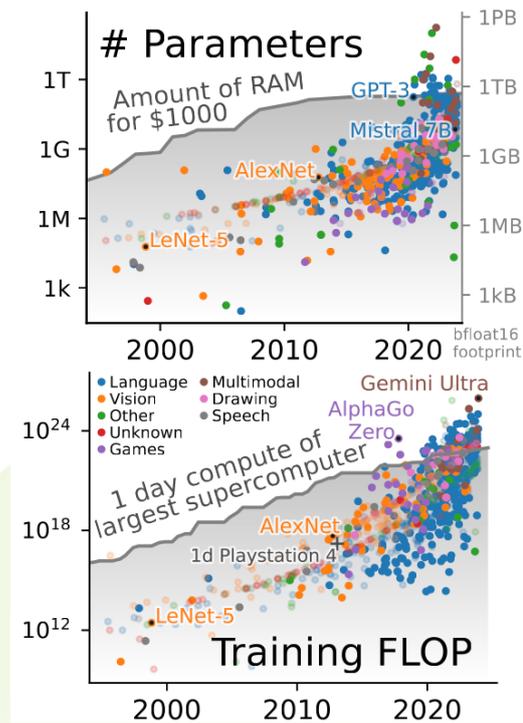


Figure 1-a : Une explosion de la taille des modèles.
 En haut : L'augmentation⁷ de la taille des modèles rend leur fonctionnement de plus en plus coûteux en termes de mémoire vive⁸.
 En bas : les besoins en ressources augmentent plus rapidement que la disponibilité de celles-ci⁹.

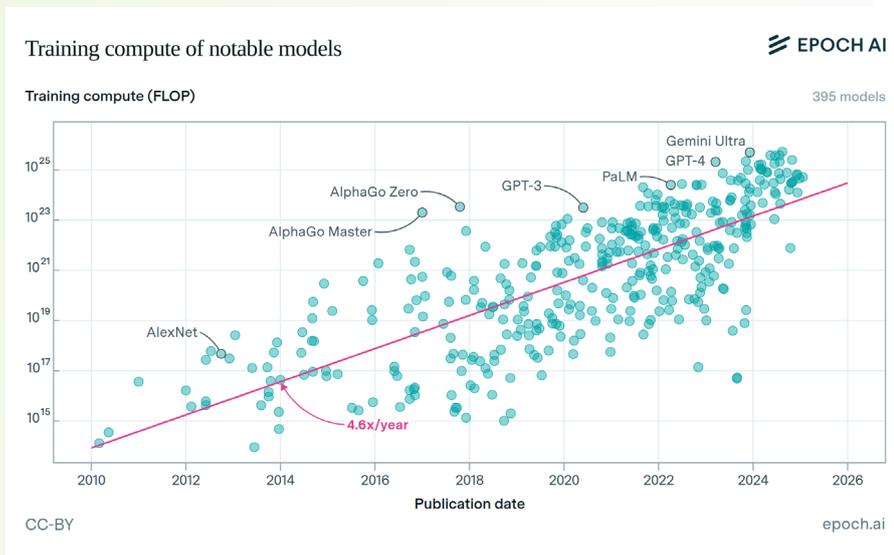


Figure 1-b : Les modèles d'IA sont de plus en plus exigeants en termes de capacité de calcul - vous pouvez voir ici le nombre total d'opérations nécessaires pour entraîner chaque modèle d'IA, en fonction du temps¹⁰.

7 G. Varoquaux, A. S. Luccioni, et M. Whittaker. Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI 2024

8 Mémoire vive (RAM)

9 Epoch, Parameter, compute and data trends in machine learning, 2023 (Paramètres, tendances en matière de calcul et de données dans l'apprentissage automatique).

10 J. Sevilla et E. Roldán, Training Compute of Frontier AI Models Grows by 4-5x per Year, Epoch AI – 2024

Alors que les plus grandes machines de calcul actuellement disponibles¹¹ supportent des charges de calcul de type exaflopique (10^{18} opérations par seconde), les exigences des derniers modèles d'IA requièrent une puissance cumulée de l'ordre de 10^{25} opérations.

Même si les chiffres présentés dans ce document se concentrent explicitement sur le développement de modèles, l'évaluation des performances des systèmes d'IA - en particulier en termes d'impact environnemental - reste souvent très difficile. En effet, les informations nécessaires telles que la taille du système, le corpus d'apprentissage, la fréquence des mises à jour ou le volume d'inférence ne sont pas toujours accessibles au public et doivent être estimées à l'aide de variables de substitution.

L'augmentation considérable de la taille des modèles peut être attribuée à trois facteurs :

- ▶ La volonté de développer des modèles d'IA généralistes.
- ▶ L'amélioration significative des performances des grands modèles de langage (LLM) au-delà d'une certaine taille critique, comme le montre **la figure 2**.
- ▶ L'approche commune consistant à améliorer les performances des LLM en augmentant le volume des données utilisées pour l'apprentissage.

Les conséquences de cette course au gigantisme vont au-delà de l'impact environnemental. Le coût du développement et de la maintenance des grands systèmes d'IA augmente si rapidement que les propriétaires sont encouragés à favoriser les applications qui génèrent des revenus substantiels au détriment de celles qui soutiennent le bien commun.

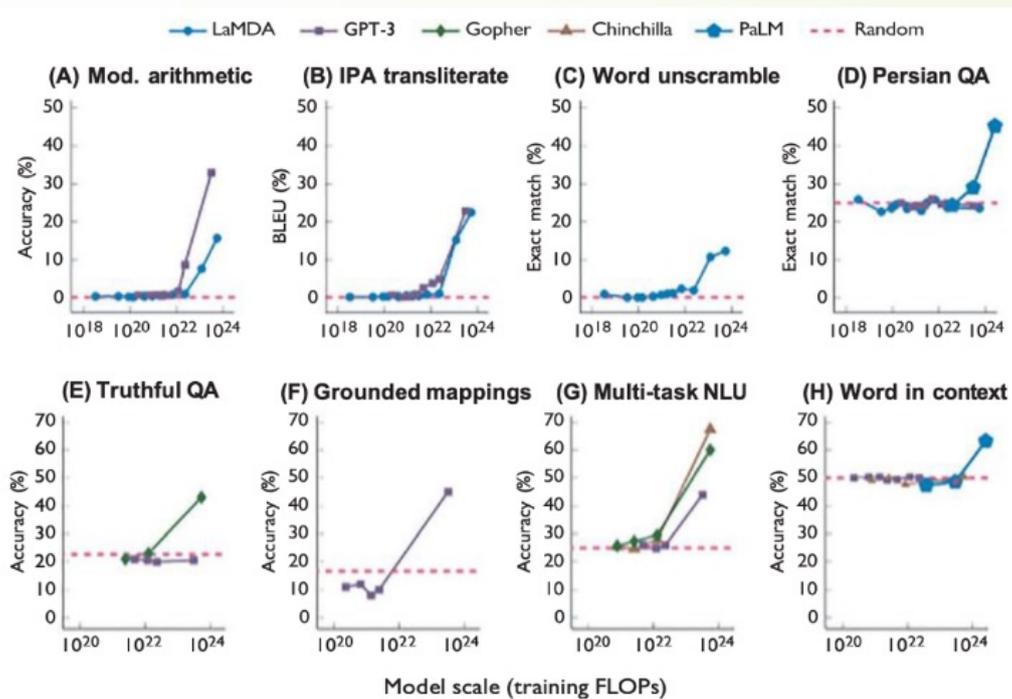


Figure 2 : Sauts dans les capacités sur huit points de référence de raisonnement pour cinq modèles de langage génératif différents en fonction du nombre d'opérations en virgule flottante (FLOP) investies dans l'optimisation des paramètres du modèle pendant l'apprentissage. Les sauts ont été désignés comme l'émergence de capacités spécifiques à des seuils particuliers de la taille du modèle¹².

¹¹ Voir <https://top500.org/>

¹² J. Wei et al, *Emergent Abilities of Large Language Models*, 2022.

Toutefois, pour de nombreuses tâches, les performances des systèmes d'IA tendent à diminuer à mesure que la taille du système (en particulier l'empreinte mémoire) augmente, comme le montrent la figure 3 et la référence¹³.

Ces tendances plaident en faveur du développement de systèmes d'IA qui se concentrent sur la qualité des données utilisées pour l'apprentissage plutôt que sur la quantité et qui exploitent des informations a priori - telles que la symétrie, l'invariance et les connaissances d'experts - pour améliorer les performances tout en réduisant le coût de la phase d'apprentissage.

Des SLM (petits modèles de langage) comportant seulement quelques milliards de paramètres ont récemment été proposés, offrant des performances élevées tout en restant économes en ressources et suggérant un certain découplage entre les performances et la taille.

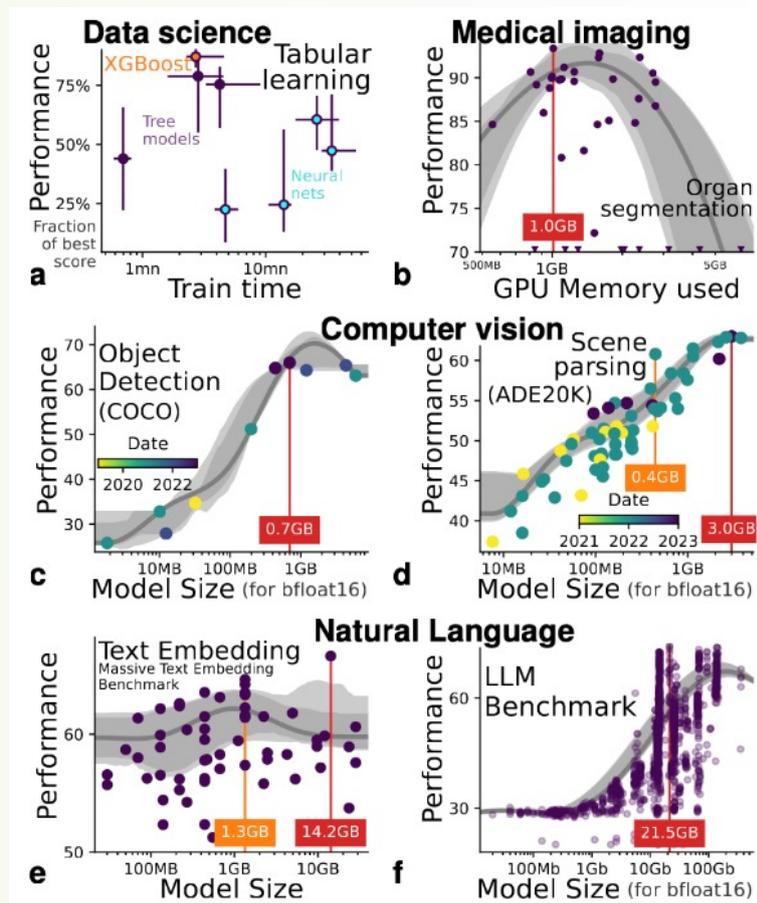


Figure 3 : Graphiques de performances en fonction de l'échelle (empreinte temporelle ou empreinte mémoire) sur des données de référence provenant de a) l'apprentissage tabulaire, b) un défi de segmentation d'images médicales, c) la détection d'objets par vision informatisée, d) l'analyse de scènes, e) l'intégration de texte et f) la compréhension de texte.

¹³ E. Horvitz et T. M. Mitchell, *Scientific Progress in Artificial Intelligence: History, Status, and Futures 2024*. In *Realizing the Promise and Minimizing the Perils of AI for Science and the Scientific Community*. Presses de l'Université de Pennsylvanie

Les technologies de l'information et de la communication (TIC) ont toujours fonctionné à l'interface entre des systèmes centralisés (serveurs, centres de données) et des systèmes décentralisés (terminaux, interfaces). On assiste aujourd'hui à l'émergence de systèmes d'IA fonctionnant sur des téléphones mobiles, voire des appareils connectés. Cette tendance va s'intensifier, avec plusieurs conséquences :

- ▶ Moins de trafic sur les réseaux et moins de calcul dans les centres de données ;
- ▶ Une utilisation accrue de terminaux éventuellement moins optimisés que les centres de données ;
- ▶ Une accélération possible de l'obsolescence des appareils qui ne supportent pas ces nouveaux systèmes, conduisant à un possible remplacement anticipé des équipements numériques dans les années à venir.

Bien qu'il soit difficile de tracer une voie claire pour les développements futurs de l'IA, **nous pouvons tout de même tirer quelques conclusions :**

- ▶ **Le potentiel de l'IA est vaste et réel**, et il serait irréaliste d'essayer de se passer entièrement de cette technologie ;
- ▶ Il faut trouver une voie pour le développement de l'IA qui **concilie la préservation de notre planète et l'innovation** ;
- ▶ **Le développement de l'IA soulève de nombreuses questions, dont beaucoup vont au-delà de l'impact environnemental.** Il est donc essentiel de pouvoir évaluer les systèmes d'IA, afin que ces nouvelles technologies puissent démontrer de manière factuelle leur potentiel au service de l'intérêt général.

Les défis présentés ci-dessous ont deux objectifs. Le but est bien évidemment de **limiter l'empreinte environnementale de l'IA**. Dans le même temps, le développement d'outils d'IA durables doit permettre des avancées scientifiques et des innovations. Ces avancées, qui vont bien au-delà de l'optimisation des systèmes existants, peuvent transformer profondément les pratiques actuelles et faciliter la transition écologique.

Propositions

Défi n° 1

→ Des technologies performantes sur le plan environnemental

Les impacts environnementaux d'un système d'IA sont multiples, depuis la fabrication, la durée de vie, l'utilisation, le recyclage, jusqu'aux changements d'usage. Ils doivent être explorés dans une perspective holistique, qui prenne en compte le matériel, les données et les algorithmes du système¹⁴. La consommation d'énergie¹⁵ - qui est calculée tout au long du cycle de vie et pondérée en fonction de l'utilisation - est un bon indicateur de ces impacts, mais d'autres types d'impacts, tels que la consommation d'eau, doivent être pris en compte.

Un grand nombre de projets de recherche sont actuellement en cours et de nombreuses technologies verront le jour pour améliorer les performances, en particulier l'efficacité énergétique, de l'architecture de l'IA, telles que :

- Accélérateurs numériques à haute efficacité énergétique,
- Microarchitectures spécialisées, chiplets spécialisés, intégration 3D,
- Intégration d'accélérateurs analogiques et photoniques compatibles avec le calcul en mémoire,
- Traitement vectoriel, calcul tensoriel¹⁶, calcul hyperdimensionnel, calcul quantique dans le flux de travail de l'IA,
- IA périphérique ou intégrée, systèmes d'IA à distance sur des terminaux,
- Réduction des coûts d'inférence grâce à des modèles plus petits et plus faciles à exécuter,
- Nouvelles techniques de refroidissement (liquide, immersif) et réutilisation de l'énergie extraite.

Ces améliorations matérielles nécessitent des algorithmes adaptés à cette nouvelle architecture :

- Quantification, précision multiple, réduite et adaptée, ainsi que des stratégies d'optimisation de la topologie du réseau ;
- Nouveaux algorithmes d'optimisation et de rétropropagation ;
- Algorithmes et modèles pour l'IA basée sur les agents, l'apprentissage fédéré et l'apprentissage distribué qui contribueront au développement de systèmes d'IA plus flexibles, plus respectueux de la vie privée et plus évolutifs.

¹⁴ C. Wu et al., *Sustainable AI: Environmental implications, challenges and opportunities 2022*.

¹⁵ AIE Électricité 2024, Paris. Ce rapport et la section consacrée à l'intelligence artificielle apportent un éclairage sur la manière de construire un système énergétique plus optimisé et montrent le poids relatif de l'IA par rapport aux autres secteurs consommateurs d'électricité.

¹⁶ N. P. Jouppi et al. *In-Datacenter Performance Analysis of a Tensor Processing Unit*, 2017.

Il convient également de noter que des pratiques responsables de gestion des données peuvent jouer un rôle clé dans la réduction du stockage de données inutiles, y compris le gaspillage de données, les données redondantes, inutilisées ou gérées de manière inefficace.

Le développement de technologies efficaces nécessite de relever des défis majeurs, tels que la création de plateformes logicielles co-intégrées et de piles de logiciels d'application conçues pour le développement de réseaux neuronaux (apprentissage et inférence) dans des environnements aux ressources limitées et dépendants du contexte. **Nous pouvons également envisager de nouvelles approches basées, par exemple, sur des technologies inspirées de la nature¹⁷ telles que :**

- **Architectures neuromorphiques** pour le calcul clairsemé et de faible précision ;
- **Réseaux neuronaux à impulsions** pour le traitement des événements à faible consommation d'énergie ;
- **Réseaux neuronaux oscillants** pour la résolution de problèmes d'optimisation combinatoire et la recherche de motifs ;
- **Mémoires associatives** pour la recherche de motifs et l'estimation de la distance ;
- **D'autres approches bio-inspirées utilisant des dispositifs memristifs¹⁸**, des réseaux neuronaux photoniques pour des calculs ultra-rapides et l'informatique basée sur l'ADN pour le parallélisme massif.

Défi n°2

→ Vers des modèles spécialisés et agiles, formés sur des ensembles de données fiables

Lorsqu'une nouvelle technologie émerge, il y a une tendance naturelle à développer des outils généralistes utilisant cette technologie. Toutefois, des outils spécialisés apparaissent souvent peu de temps après.

C'est ce que l'on observe dans le développement de l'IA générative, avec l'émergence initiale de grands modèles qui fournissent des fonctions à usage général. Nous devons renforcer le développement d'outils d'IA plus spécifiques à une tâche donc plus petits, car ces outils, contrairement aux grands modèles, sont plus susceptibles de concilier la protection de l'environnement et l'innovation.

Comme le montre la **figure 3**, l'augmentation du volume de données utilisées pour l'apprentissage de l'IA n'améliore pas nécessairement les performances d'un système d'IA. En fait, les performances peuvent diminuer lorsque les données utilisées sont redondantes. L'objectif est donc de définir un cadre propice au développement de solutions à la fois petites, suffisantes et résilientes. Des solutions d'IA sophistiquées conçues pour résoudre des problèmes spécifiques peuvent parfois être préférables au perfectionnement de modèles à usage général.

¹⁷ D. S. Modha et al., *Neural inference at the frontier of energy, space, and time - Science* 382, pp. 329–335, 2023.

¹⁸ Les dispositifs memristifs sont des commutateurs électriques qui conservent un état de résistance basé sur l'historique de la tension et du courant passés, leur permettant de stocker et de traiter des informations. Un type clé, basé sur le mouvement ionique, consiste en un empilement de couches minces conducteur/isolant/conducteur. Proposés pour la première fois dans les années 1960, les progrès récents ont permis de mettre au point des dispositifs rapides, à faible consommation d'énergie et à haute endurance, dont l'échelle peut être inférieure à 10 nm et qui peuvent être empilés en 3D. Toutefois, les mécanismes fondamentaux de ces dispositifs restent obscurs, ce qui entrave leur application à plus grande échelle [Yang, J., Strukov, D. & Stewart, D. *Memristive devices for computing. Nature Nanotech* 8, 13–24 (2013)].

Les principaux problèmes auxquels nous sommes confrontés pour relever ce défi sont les suivants :

→ Favoriser la disponibilité et accroître la visibilité d'ensembles de données précises, de données spécifiques à certains domaines ou à certaines applications. Ces jeux de données, qui devraient être conformes à la législation en vigueur, peuvent être utilisés par un grand nombre d'acteurs, en particulier les acteurs émergents, pour l'entraînement, ainsi que pour l'évaluation des systèmes mis au point.

→ Favoriser le développement de communs numériques, tels que des outils à code source ouvert, qui rassembleront non seulement des ensembles de données, mais aussi des composants et architectures logiciels.

Défi n° 3

→ De nouvelles méthodes et de meilleures données pour évaluer l'empreinte environnementale de l'IA

L'un des moyens de répondre aux préoccupations soulevées par l'IA est de faire preuve de transparence quant à son impact, en particulier sur l'environnement. L'impact environnemental de l'IA fait partie intégrante de son évaluation. Pour limiter ces impacts, il faut les comprendre en détail car leur évaluation va au-delà de l'analyse du cycle de vie et des exigences en matière d'écoconception¹⁹.

Les impacts environnementaux directs de l'IA, notamment la fabrication d'équipements, l'utilisation d'eau douce, de matériaux abiotiques, d'éléments terrestres rares et la consommation d'énergie, ou les émissions de gaz à effet de serre, commencent à être étudiés. Grâce à ces connaissances, les scientifiques seront en mesure d'élaborer des méthodologies et des indicateurs de performance clés pour évaluer l'empreinte environnementale d'un système d'IA, à condition de disposer des informations et des quantités physiques caractérisant son fonctionnement.

Les modèles open source peuvent encourager le partage des ressources et éviter la répétition de l'entraînement des modèles pour des utilisations similaires, ce qui garantirait une consommation plus efficace de l'énergie et des ressources.

Les principaux problèmes auxquels nous sommes confrontés pour relever ce défi sont les suivants :

→ Définir des paramètres significatifs²⁰ qui peuvent ensuite être utilisés pour construire des métriques et des indicateurs afin de quantifier l'impact environnemental de l'IA. Il est essentiel de prendre en compte les impacts de bout en bout - des terminaux clients qui utilisent les réseaux aux serveurs nuage - à chaque étape du processus d'IA (apprentissage, inférence) et avec une vision complète du cycle de vie des ressources matérielles et logicielles²¹.

¹⁹ T. Duke, P. Giudici, P., *Responsible AI in Practice: A Practical Guide to Safe and Human AI*, Apress L.P., 2025

²⁰ On peut penser à des données telles que : la taille du système, le processus d'apprentissage, le volume d'inférence, la fréquence de mise à jour, la performance énergétique des processeurs et leur durée de vie.

²¹ A. Berthelot, E. Caron, M. Jay, L. Lefevre, *Estimating the environmental impact of Generative-AI services using an LCA-based methodology*. CIRP LCE 2024-31e Conférence sur l'ingénierie du cycle de vie, juin 2024.

→ **Encourager les entreprises à partager davantage de données**, car la plupart des données sur les impacts environnementaux des systèmes d'IA proviennent actuellement des fournisseurs d'IA, tels que les exploitants de centres de données à grande échelle, ou de leurs clients. Cela peut se faire par le biais de différentes méthodes, telles que la divulgation volontaire ou la déclaration de données (par exemple, le système européen de déclaration des centres de données). Une normalisation est également nécessaire pour garantir que les rapports de données soient facilement comparables et exploitables²².

→ **Proposer des mesures logicielles de la consommation d'énergie** avec des niveaux d'abstraction capables de s'adapter à l'hétérogénéité des technologies et des architectures utilisées par les fabricants. Il s'agit notamment de définir des approches permettant de comprendre la combinaison matérielle et logicielle la plus optimale en termes de performance environnementale.

→ **Quantifier les impacts positifs probables ou possibles de l'IA dans des secteurs spécifiques**. Cela pourrait s'avérer extrêmement utile pour inciter les personnes et les organisations soucieuses du développement durable à explorer de nouvelles applications de cette technologie pour le bien public. Estimer les effets des systèmes d'IA sur la réduction de l'empreinte environnementale d'autres secteurs économiques est une vaste entreprise. Pour y parvenir, nous devons d'abord nous limiter à des cas d'utilisation et à des applications spécifiques ou à quelques secteurs bien définis, par exemple l'agriculture, les transports, les télécommunications ou le logement.

Défi n° 4

→ Appliquer les principes de l'économie circulaire au matériel utilisé pour l'IA

Au cours des dernières décennies, des progrès ont été réalisés dans la mise en œuvre des principes de l'économie circulaire et dans l'amélioration du recyclage au sein de l'industrie du numérique²³. En ce qui concerne la législation, un exemple notable est l'adoption du règlement sur l'écoconception des produits durables (ESPR) à l'échelle de l'Union européenne.

Ce règlement établit un cadre réglementaire général pour l'adoption de nouvelles normes d'écoconception plus ambitieuses pour les produits liés à l'énergie, les technologies de l'information et de la communication (TIC) et d'autres biens électroniques.

Toutefois, **cela ne signifie pas que tous les défis techniques et sociétaux aient été résolus** lorsqu'il s'agit de généraliser les approches de l'économie circulaire dans les chaînes d'approvisionnement de l'IA. Cette section du document suggère que, même si nous ne partons pas de zéro, ce défi reste à relever.

Le principal obstacle réside dans le fait que **le déploiement de l'IA est susceptible d'entraîner une augmentation du flux de déchets électroniques**, qui pourrait atteindre une accumulation totale de 1,2 à 5 millions de tonnes au cours de la période 2020-2030, en fonction des différents scénarios de développement de l'IA générative²⁴. Cela s'explique notamment par le fait que **le matériel et les logiciels doivent être adaptés aux nouvelles technologies de l'IA**. L'identification et la mise en œuvre de stratégies d'économie circulaire

²² La Commission adopte un système européen de notation de la durabilité des centres de données.

²³ Chersan, Ionela-Corina & Paunescu, Mirela & Nichita, Mirela & Dumitru, Valentin & Manea, Lidia. (2023). *Circular Economy Practices in the Electrical and Electronic Equipment Sector in the European Union*. Amfiteatru Economic. 25. 80-100. 10.24818/EA/2023/62/80.

²⁴ P. Wang et al, *The challenges of generative artificial intelligence in electronic waste*, Nature Computational Science, 2024.

à un stade précoce - comme c'est déjà le cas pour certains appareils et technologies électroniques et électriques classiques - et tout au long de la chaîne de valeur de l'IA sont donc essentielles et pourraient permettre de réduire la production de déchets électroniques de 16 à 86 %.

En outre, les stratégies qui augmentent la modularité et la réparabilité des équipements - qu'il s'agisse de terminaux ou d'équipements connectés - en permettant **le remplacement de composants individuels plutôt que d'unités entières**, stimuleront la dynamique de l'économie circulaire. Elles simplifieront également l'intégration ou l'ajout de solutions d'IA intégrées (Edge AI, Tiny Edge) dans les équipements industriels, ce qui leur permettra d'exploiter les avantages de l'IA, comme l'amélioration de la maintenance prédictive, par exemple. Cela pourrait augmenter la durée de vie des composants matériels utilisés pour alimenter la technologie de l'IA.

Si l'IA peut contribuer à prolonger la durée de vie du matériel, il est essentiel de prendre en compte l'impact environnemental du matériel lui-même. Des recherches ont déjà été menées sur ce sujet, et certaines industries ont élaboré des stratégies visant à réduire l'empreinte environnementale des technologies numériques courantes, et pas seulement de celles qui sont spécifiques à l'IA. Bien que ces questions ne soient pas propres à l'IA, elles deviennent de plus en plus importantes avec l'essor de l'IA, en particulier de l'IA générative.

L'essor des systèmes d'IA pourrait également accroître la demande pour certains métaux. Le renforcement des stratégies de collecte et de recyclage des équipements peut réduire l'impact environnemental du matériel utilisé par les systèmes d'IA, en fournissant une source secondaire de métaux critiques et en contribuant à garantir un approvisionnement plus fiable et durable de ces minéraux essentiels²⁵. Le principal défi du recyclage réside dans la complexité de la récupération des métaux rares présents dans les composants électroniques. Par conséquent, en plus de créer un marché pour les pièces détachées récupérées, nous devons soutenir les efforts en matière de recherche et développement pour établir une industrie axée sur le traitement et la récupération des composants qui ne peuvent pas être directement réutilisés dans de nouveaux équipements.

Les principaux problèmes auxquels nous sommes confrontés pour relever ce défi sont les suivants :

- **Identifier et mettre en œuvre des stratégies et leviers qui renforceront les efforts visant à maintenir la dynamique d'allongement de la durée de vie des équipements** (serveurs et terminaux) et à réduire l'obsolescence prématurée²⁶, notamment par l'optimisation logicielle, l'écoconception du matériel (pour améliorer la modularité, l'évolutivité et le démantèlement en fin de vie), ou la réutilisation des équipements de génération précédente pour des applications moins énergivores telles que l'exécution de modèles de petite et moyenne taille, de projets spécialisés, etc. ;
- **Identifier des stratégies permettant d'éviter que les composants (tels que les cartes, les puces, etc.) ne deviennent des déchets**, notamment en améliorant les équipements de démantèlement utilisés pour récupérer des composants qui peuvent être réutilisés pour de nouveaux usages lorsqu'ils ne sont plus compatibles avec le fonctionnement des systèmes d'IA ;
- **Renforcer les activités de récupération des métaux** afin que lorsque les composants ne peuvent plus être réutilisés et deviennent des déchets électroniques, nous puissions augmenter la circularité des matières premières et réduire notre dépendance aux métaux critiques provenant de l'exploitation minière. Nous devons également continuer à nous attaquer collectivement aux obstacles structurels et techniques,

²⁵ <https://www.iea.org/reports/recycling-of-critical-minerals>

²⁶ Selon le règlement de l'UE établissant un cadre pour la définition d'exigences en matière d'écoconception pour les produits durables, « l'obsolescence prématurée » désigne une caractéristique de la conception d'un produit ou une action ou omission ultérieure qui fait que le produit devient non fonctionnel ou moins performant sans que ces changements de fonctionnalité ou de performance soient le résultat d'une usure normale. Les exigences en matière d'écoconception doivent également tenir compte des pratiques associées à l'obsolescence prématurée. De telles pratiques ont un impact globalement négatif sur l'environnement, sous la forme d'une augmentation des déchets et de la consommation d'énergie et de matériaux, que l'on peut réduire grâce aux exigences d'écoconception tout en contribuant à une consommation durable.

notamment en veillant à ce que les métaux recyclés conservent les mêmes performances que leurs homologues d'origine.

Défi n° 5

→ Changer l'image de l'IA pour promouvoir le développement d'outils d'IA frugaux et leur utilisation rationnelle

De nombreux scientifiques travaillent à la mise au point de systèmes d'IA frugaux, économiques, voire low-tech. Il est essentiel de stimuler ces travaux, de les récompenser et de les rendre visibles. Pour ce faire, nous devons changer la manière dont nous mesurons et évaluons les performances.

Le développement de systèmes d'IA à très grande échelle est en partie dû aux critères utilisés pour évaluer la performance des algorithmes et des technologies dans la communauté scientifique. Ces critères de performance privilégient actuellement la performance quantitative, par exemple au travers de benchmarks généraux, au détriment de benchmarks "spécialisés" ou "centrés" sur quelques problèmes qui permettraient de mettre en évidence les avantages d'une IA "spécifique".

Pour que les utilisateurs aient un comportement responsable, des modèles d'IA frugale doivent exister et faire l'objet d'une publicité appropriée. Par le biais de partenariats public-privé ou de labels, par exemple, les fournisseurs d'IA devraient être encouragés à développer de tels modèles d'IA ciblés et spécifiques à une tâche, et à en faire la publicité auprès des clients B2B et du public.

Les principaux problèmes auxquels nous sommes confrontés pour relever ce défi sont les suivants :

→ **Donner de la visibilité aux travaux scientifiques sur l'IA frugale et durable** à l'échelle internationale, notamment par le développement de nouvelles revues ou conférences facilitant la publication et la présentation des travaux des chercheurs et praticiens travaillant sur l'IA et l'environnement ;

→ **Développer la formation initiale et continue à l'IA.** Les contenus pédagogiques doivent intégrer des compétences vertes et des outils et solutions conçus pour un monde sous contrainte. La rapidité avec laquelle de nouveaux concepts d'IA émergent souligne l'importance de l'apprentissage tout au long de la vie ;

→ **Développer des contenus éducatifs** (par exemple sous l'égide de l'UNESCO ou de l'OCDE) tels que des MOOC, pour aider les personnes à comprendre le fonctionnement de l'IA, en soulignant ses forces et ses faiblesses, et en encourageant l'utilisation raisonnée des outils d'IA. Des chiffres et des exemples concrets aident le grand public à comprendre les divers impacts de l'IA, notamment sur l'environnement.

Annexe

Il existe des liens profonds entre l'IA et l'environnement. D'une manière générale, ils peuvent être répartis entre trois catégories :

- ▶ **L'IA qui modélise l'environnement, le climat**, ainsi que les risques naturels et environnementaux ;
- ▶ **L'IA comme outil de décarbonation de certaines industries** (agriculture, mobilité, industrie, logement et construction) ;
- ▶ **L'empreinte environnementale de l'IA**, et plus particulièrement l'impact des composants matériels et logiciels sur lesquels reposent les outils d'IA.

Le présent document traite du dernier point. Dans ce qui suit, nous décrivons brièvement les contributions de l'IA dans les deux premiers domaines.

Modélisation environnementale : comprendre, simuler et prévoir

La simulation numérique de modèles mécanistes a permis de comprendre et d'anticiper les phénomènes géophysiques (météorologie, océanographie, climat, risques naturels et environnementaux, etc.) Néanmoins, la taille des modèles, leur paramétrisation et leur coût (temps de calcul et coût environnemental associé) font obstacle aux avancées scientifiques. Dans le même temps, il existe une multitude de données (capteurs, données satellitaires, enquêtes réalisées sur de nombreuses années) permettant l'utilisation de la science des données et de l'IA en particulier. Ces nouvelles technologies permettent des avancées majeures²⁷, parfois avec une hybridation entre les modèles traditionnels et les techniques d'IA.

Les services climatiques (la fourniture et l'utilisation de données, d'informations et de connaissances sur le climat pour soutenir la prise de décision) jouent un rôle clé dans de nombreux secteurs de notre économie.

Contribution de l'IA à la décarbonation et à la transition écologique – Atténuation

Les technologies numériques et l'IA transforment des secteurs très divers et constituent souvent un levier pour aider à réduire l'empreinte environnementale de ces secteurs. Les avancées sont particulièrement notables dans les secteurs suivants : agroécologie (polyculture, outils d'aide à la décision pour les agriculteurs, diversité génétique, robotique flexible, etc.), mobilité (multimodalité, outils de planification, etc.), industrie (écoconception, optimisation des centres de données, jumeaux numériques), énergie (réseaux électriques intelligents, énergies renouvelables). Cette liste n'est pas exhaustive et, à ce stade, peu de résultats quantitatifs sont disponibles.

Il convient de souligner que l'utilisation de l'IA dans ce domaine ne se limite pas à l'optimisation des processus existants. L'IA peut avoir un effet transformateur, en proposant une trajectoire totalement différente (innovations radicales ou percées créatives).

²⁷ V. Eyring, W.D. Collins, P. Gentine et al, [Pushing the frontiers in climate modeling and analysis with machine learning](#). Nat. Clim. Chang. 14, 916–928 2024. GenCast.
G. Couairon, C. Lessig, A. Charantonis et C. Monteleoni, [ArchesWeather: An efficient AI weather forecasting model at 1.5 deg resolution 2024](#).

Ce document de synthèse a été coordonné par l'Institut national de recherche en sciences et technologies du numérique (INRIA), la Direction du numérique (DNUM) et le Laboratoire d'innovation (ECOLAB) du ministère de la Transition écologique, de la Biodiversité, de la Forêt, de la Mer et de la Pêche.

Contributeurs

La création de ce document de synthèse est le fruit d'un effort de collaboration qui n'aurait pu être mené à bien sans les éclairages apportés par les chercheurs, les chefs d'entreprise et les experts en IA qui y ont participé. Veuillez noter que les contributeurs ne soutiennent pas nécessairement le document de synthèse dans son intégralité.

Fabien Abrikh, Institut du numérique responsable (INR)

Ludovic Argà, Orange

Cedric Auliac, Commissariat à l'énergie atomique et aux énergies alternatives (CEA) / Agence de programme Appareils, systèmes numériques et infrastructures

Elise Behm, ministère de la Transition écologique

Louis Billerot, Ministère de l'économie et des Finances

Jean-Philippe Bourgoïn, Commissariat à l'énergie atomique et aux énergies alternatives (CEA) / Agence de programme Appareils, systèmes numériques et infrastructures

Chloé Broguet, Lenovo

Maxime Chantillon, Centre interuniversitaire de microélectronique (imec)

Nathalie Charbonniaud, Orange

Bertrand Charpentier, Prune AI

Franck Coisnon, Eviden/Atos

Philippe de Cuetos, Alliance française des industries du numérique (Afnium)

Fabien Demange, Eviden/Atos

Sabine Demey, Programme flamand de recherche sur l'IA / Centre interuniversitaire de microélectronique (imec)

Claire Dorville, ministère de la Transition écologique

Diane Dufoix, IBM France

Tamar Eilam, IBM

Cyril Fakiri, Lenovo

Julie Galland, Commissariat français à l'énergie atomique et aux énergies alternatives (CEA)

Emilie Gaudu, Lenovo

Clara Grojean, Alliance française des industries du numérique (Afnium)

Yves Haggiag, Lenovo

Prof. Ian R. Hodgkinson, Université de Loughborough

Prof. Thomas W. Jackson, Université de Loughborough

Pierre Jaeger, IBM

Dr. Somya Joshi, Institut de l'environnement de Stockholm

Léo Lafarge, Alliance française des industries du numérique (Afnium)

Alexandra Laffitte, Lenovo

Patricia Lago, Université libre d'Amsterdam

Régis Le Drézen, Think Smartgrids

Grégory Lebourg, OVHcloud

Laurent Lefevre, École normale supérieure Lyon / Institut national de recherche en sciences et technologies du numérique (Inria)

Thierry Lochon, ministère de la Transition écologique

Sasha Luccioni, École normale supérieure / Université Laval Programme des Nations Unies pour l'environnement

Jimena Luna, Lenovo

Caroline Marcouyoux, Amazon Web Services

Sébastien Massart, Dassault Systèmes

Stella Morabito, Alliance française des industries du numérique (Afnm)

Sheila O' Hara, IBM France

Josh Parker, NVIDIA

Aaron Pietzonka, ministère de la Transition écologique

Anne-Cécile Rémond, Lenovo

Andres Rodriguez, IBM

David Rolnik, Université McGill / Climate Change AI

Jacques Sainte-Marie, Institut national de recherche en sciences et technologies du numérique (Inria)

Chiara Sandionigi, Commissariat français à l'énergie atomique et aux énergies alternatives (CEA)

Pierre Seroul, Eviden/Atos

Sabrina Stanisl, Qualcomm

Durabilité de l'entreprise, Business AI, BTP et équipes chargées des affaires gouvernementales, SAP

Stuart Sweeney Smith, Google

Denis Thystram, Université Grenoble Alpes / MIAI cluster IA

Jérôme Totel, Data4

Laurent Vanel, IBM France

Matthieu Wong Hang, Lenovo.

Sponsors



Contact

sustainability-ai-summit@developpement-durable.gouv.fr