



COMMISSION
DES AFFAIRES
EUROPÉENNES

LE COÛT ENVIRONNEMENTAL DE L'INTELLIGENCE ARTIFICIELLE

**Communication de Mme Danièle Obono,
députée de Paris (LFI – NFP)
et avis politique**

Mercredi 8 avril 2026

TABLE DES MATIÈRES

I. L'IA GÉNÉRATIVE : UNE RÉVOLUTION TECHNOLOGIQUE À L'EMPREINTE PHYSIQUE CONSIDÉRABLE.....	7
A. Des fondements techniques qui conditionnent l'impact environnemental.....	7
1) L'architecture des grands modèles de langage.....	7
2) Une démocratisation fulgurante depuis 2022	8
B. Les sources d'impact environnemental.....	9
1) La fabrication des équipements et des infrastructures	10
2) L'alimentation et le refroidissement des centres de données.....	10
3) La consommation en eau : un enjeu territorial.....	11
II. UNE QUANTIFICATION LACUNAIRE DANS UN CONTEXTE DE CROISSANCE EXPONENTIELLE.....	13
A. Des obstacles méthodologiques structurels.....	13
1) L'opacité des acteurs et l'absence de méthodologie commune.....	13
2) La complexité de l'imputation des usages	13
B. Les ordres de grandeur disponibles et les projections.....	15
1) Les données existantes	15
2) Des trajectoires de croissance préoccupantes.....	16
C. Les limites du narratif des bénéfiques compensatoires.....	17
1) Des cas d'usage favorables à l'environnement réels mais minoritaires	17
2) L'effet rebond et les contradictions du modèle économique.....	18
III. VERS UNE GOUVERNANCE PUBLIQUE DE L'IA SOBRE : ENJEUX DE SOUVERAINETÉ ET LEVIERS D'ACTION	19
A. La dépendance technologique comme amplificateur des impacts environnementaux.....	19
1) Une dépendance structurelle aux acteurs américains.....	19
2) L'impossible course à la puissance de calcul.....	20
B. Un cadre réglementaire en construction, encore insuffisant.....	22
1) Les avancées du cadre européen et national.....	22
2) Les lacunes persistantes.....	22
C. Recommandations	23
1) Imposer la transparence et standardiser les métriques	23
2) Planifier l'implantation des centres de données.....	24
3) Orienter l'investissement public vers des modèles frugaux	24
4) Renforcer la régulation au niveau européen	26
CONCLUSION.....	27

LISTE DES PERSONNES AUDITIONNÉES	28
EXAMEN EN COMMISSION	29
AVIS POLITIQUE LE COÛT ENVIRONNEMENTAL DE L'INTELLIGENCE ARTIFICIELLE.....	44



INTRODUCTION

L'intelligence artificielle n'est pas une technologie nouvelle. Ses fondements théoriques remontent aux travaux d'Alan Turing et ses premières applications pratiques aux années 1960. Après une période d'intense effervescence dans les années 1980, le domaine a traversé des « hivers » marqués par la déception des promesses non tenues, avant de connaître, à partir des années 2010, une renaissance portée par la conjonction de trois facteurs : la disponibilité de données en volumes massifs, la puissance de calcul des unités de traitement graphique (GPU) et les progrès algorithmiques.

Pour un état des lieux détaillés de l'intelligence artificielle, on peut utilement se référer à l'important travail de synthèse réalisé par l'Office parlementaire d'évaluation des choix scientifiques et technologiques dans un rapport publié en novembre 2024 ⁽¹⁾.

(1) Patrick CHAIZE, Corinne NARASSIGUIN et Alexandre SABATOU, [ChatGPT, et après ? Bilan et perspectives de l'intelligence artificielle](#), rapport OPECST, 28 novembre 2024

Ce qui distingue la période actuelle est l'avènement de l'intelligence artificielle générative (IAG), rendue possible par une avancée architecturale majeure : les transformeurs.

Cette avancée technologique a permis le développement de larges modèles de langage (LLM) incarnée par des outils tels que ChatGPT, dont le lancement en novembre 2022 a rencontré un large écho auprès du grand public.

Cet engouement s'accompagne d'investissements colossaux : les géants américains du numérique ont annoncé des dépenses d'infrastructure se chiffrant en centaines de milliards de dollars sur la décennie à venir. Or, derrière la dématérialisation apparente de ces services, se dissimule une réalité profondément matérielle : des centres de données énergivores, des chaînes d'approvisionnement en minerais critiques, une consommation en eau significative, et des émissions de gaz à effet de serre croissantes.

La présente communication a pour objet d'évaluer les impacts environnementaux de l'IA en s'appuyant sur les propos recueillis dans le cadre de plusieurs auditions de travail menées auprès d'administrations (ADEME, ARCEP) et de chercheur.ses. Elle s'attache en premier lieu à décrire les sources d'impact environnemental inhérentes aux architectures de l'IA générative (I), puis à mesurer l'état des connaissances et les projections disponibles dans un contexte de forte opacité des acteurs (II), avant de formuler des recommandations à l'attention des pouvoirs publics, articulées autour des enjeux de souveraineté numérique et de sobriété technologique (III).

Il convient de souligner d'emblée que les impacts sociaux, économiques et démocratiques de l'IA – au premier rang desquels figurent les questions de sécurité des données personnelles, de conditions d'emploi et d'autonomie informationnelle – ne seront pas traités dans cette communication, non qu'ils soient de moindre importance, mais parce que la présente communication fait le choix délibéré de concentrer son analyse sur la seule dimension environnementale.

Rappelons simplement ici que les investissements massifs à destination de l'IA générative sont autant d'investissements qui ne sont pas réalisés dans des domaines où les avantages sont certains (santé, éducation). Outre ce coût d'opportunité très

élevé, le développement massif des capacités de calcul allouées à l'IAG conduit à une augmentation des prix de certains composants informatiques (cartes RAM notamment).

L'IAG conduit également à une déstructuration profonde des marchés du travail dans le secteur tertiaire, faisant craindre un déclassement de nombreux travailleurs. Étant entendu que les emplois créés dans le secteur numérique ne requerront pas les mêmes compétences et seront, surtout, en nombre bien moins important – la densité d'emploi dans un centre de données étant bien plus faible que dans la plupart des métiers de services.

I. L'IA GÉNÉRATIVE : UNE RÉVOLUTION TECHNOLOGIQUE À L'EMPREINTE PHYSIQUE CONSIDÉRABLE

A. Des fondements techniques qui conditionnent l'impact environnemental

1) L'architecture des grands modèles de langage

L'intelligence artificielle générative repose sur une architecture spécifique : les grands modèles de langage (Large Language Models – LLM), également désignés « modèles de fondation » ou « modèles à usage général ». Ces modèles doivent leur existence à une avancée algorithmique déterminante : les transformeurs, introduits en 2017 par une équipe de recherche de *Google* dans l'article fondateur « *Attention is all you need* »⁽²⁾.

Ces transformeurs présentent l'avantage de s'affranchir de l'apprentissage supervisé qui repose sur l'utilisation de données étiquetées (l'étiquetage des données étant réalisé manuellement, les jeux de données étiquetées sont moins nombreux, car très coûteux)⁽³⁾. Avec les transformeurs, il est possible d'entraîner la machine avec des données non étiquetées qui sont disponibles en bien plus grand nombre. Cet entraînement dit « non supervisé » permet de faire apprendre à la machine les structures de phrases pour faire émerger des modèles linguistiques. Ensuite, ce premier apprentissage est affiné au moyen d'une phase d'entraînement supervisée, à

(2) Ashish Vaswani, Noam Shazeer, Niki Parmar, et alii, « [Attention is All you Need](#) », décembre 2017

(3) D'après les explications formulées par Conseil national des Barreaux, « [Utilisation des systèmes d'intelligence artificielle générative](#) », groupe de travail intelligence artificielle, septembre 2024

l'aide de données étiquetées et d'une phase d'apprentissage par renforcement, reposant sur un système de récompenses pour toutes les réponses correctes.

La grande avancée des transformeurs est « l'attention » selon laquelle le sens d'un mot n'est compris que dans son contexte. Les transformeurs sont capables de mettre l'attention sur le mot le plus pertinent d'une phrase en fonction des mots qui le précèdent. Ils sont alors en mesure de comprendre que le même mot peut avoir un sens différent selon la structure de la phrase que ce soit dans une phrase simple, mais également – c'est une avancée importante – dans une phrase complexe.

L'IAG apprend les mots en prenant en compte l'importance relative des mots dans leur contexte, et ce, sur des milliards d'exemples. Elle apprend la structure des phrases pour établir un modèle statistique capable de prédire le mot suivant. C'est ainsi que l'IAG est en mesure de produire un contenu nouveau (texte, image, son, vidéo) distinct de ses données d'apprentissage.

Le cycle de vie d'un modèle de langage comprend deux phases distinctes aux caractéristiques environnementales très différentes. La phase d'entraînement constitue une consommation électrique continue et très élevée sur une longue durée – des semaines, voire des mois – mais elle est pilotable et peut en principe être planifiée. La phase d'inférence, c'est-à-dire l'utilisation effective du modèle pour répondre aux requêtes des utilisateur·ices, représente une consommation marginale par requête individuelle, mais génère une charge électrique considérable en raison de la multiplication des usages à l'échelle mondiale.

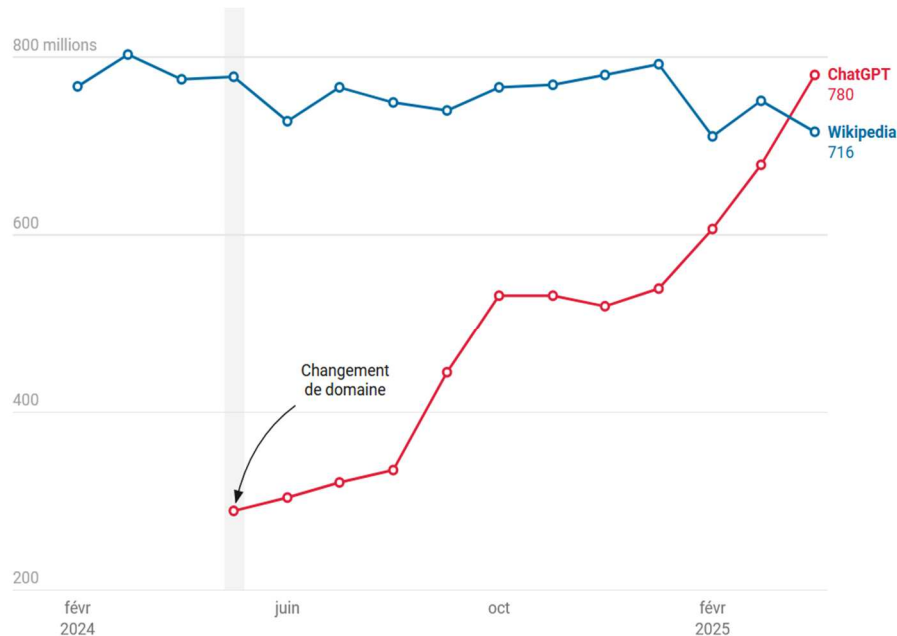
L'acronyme GPT, qui désigne les modèles phares d'OpenAI, est lui-même révélateur : « *Generative Pre-trained Transformer* ». Il traduit précisément cette architecture – transformeur – et ce mode opératoire – pré-entraînement sur données massives, puis affinage supervisé – qui caractérisent l'ensemble des grands modèles actuels.

2) Une démocratisation fulgurante depuis 2022

Si les techniques d'IA existent depuis plusieurs décennies et si des applications sectorielles – outils d'aide à la décision, optimisation logistique, algorithmes de recommandation – sont déployées depuis longtemps, l'IAG se distingue radicalement par sa vitesse d'adoption. Comme le souligne l'ADEME, aucune technologie numérique

n'avait jusqu'alors connu une diffusion aussi rapide auprès du grand public. ChatGPT a atteint 100 millions d'utilisateurs en deux mois, là où il avait fallu neuf mois à TikTok et deux ans et demi à Instagram pour parvenir à ce seuil.

En avril 2025, le site de ChatGPT a dépassé celui de Wikipédia en nombre de visites mensuelles aux États-Unis d'Amérique.



Source – Similarweb. Graphique – Le Grand Continent

Cette adoption de masse entraîne une explosion de la demande d'inférence, c'est-à-dire du nombre de requêtes traitées quotidiennement. En outre, l'IA générative s'intègre progressivement dans des services numériques préexistants (moteurs de recherche, suites bureautiques, messageries, plateformes audiovisuelles, ...) rendant de plus en plus difficile l'isolation de son empreinte propre au sein de l'empreinte globale du numérique.

B. Les sources d'impact environnemental

Contrairement à l'imaginaire du « cloud » qui suggère une technologie immatérielle et légère, l'IA et le numérique reposent sur des infrastructures matérielles, physiques, énergivores, et qui ont donc des impacts environnementaux importants. À cet égard, nous reprenons l'image du « charbon numérique » qui avait

été avancée dans un article du Monde diplomatique (4). Ces impacts se distribuent sur trois postes principaux.

1) La fabrication des équipements et des infrastructures

La fabrication des équipements nécessaires au fonctionnement de l'écosystème IA – semi-conducteurs, cartes graphiques (GPU), serveurs, équipements réseaux, mais aussi les bâtiments abritant les centres de données – mobilise des ressources naturelles importantes. Les semi-conducteurs, fabriqués pour l'essentiel à Taïwan et en Corée du Sud, nécessitent des procédés de gravure extrêmement précis, consommateurs d'eau et de produits chimiques spécifiques. Les métaux et terres rares indispensables à leur fabrication font l'objet d'une dépendance géopolitique marquée, la Chine contrôlant une part prépondérante de leur extraction et de leur raffinage.

En termes d'émissions de gaz à effet de serre, la fabrication des équipements représente environ 25 % de l'empreinte carbone totale de la filière numérique à l'échelle mondiale. Cette proportion monte à 50 % en France, où le caractère faiblement carboné du mix électrique réduit mécaniquement le poids relatif de la consommation d'énergie à l'usage (5). La durée de vie des serveurs dédiés à l'IA demeure peu documentée, mais les praticiens soulignent l'accélération des cycles de remplacement imposée par la rapidité des évolutions technologiques, ce qui augmente l'empreinte liée à la fabrication.

2) L'alimentation et le refroidissement des centres de données

Les centres de données constituent la colonne vertébrale de l'IA. Selon le rapport de l'Agence internationale de l'énergie (AIE) de 2025, ils représentaient 1,5 % de la demande d'électricité mondiale en 2024, soit 415 TWh. L'IA contribue à elle seule pour environ 10 % de cette consommation (6). En France, on recensait, en 2025, 352 centres de données en activité pour une puissance installée totale de 2,5 gigawatts (GW), occupant quelque 2,5 millions de mètres carrés (7).

(4) Sébastien BROCA, « [Le numérique carbure au charbon, par Sébastien Broca](#) », *Le Monde Diplomatique*, mars 2020, p. 3

(5) D'après l'audition de Marlène de Bank, ingénieure, The Shift Project.

(6) Agence internationale de l'énergie (AIE), *Energy and AI*, rapport exécutif, 2025.

(7) D'après l'audition de Gauthier Roussilhe.

Les projections de consommation électrique des centres de données sont toutes alignées sur un facteur 2 ou 3, avec néanmoins de fortes incertitudes.

- Rapport de l'Agence internationale de l'énergie : x2 dans le monde entre 2023 et 2030 ⁽⁸⁾.
- Département de l'Énergie aux États-Unis d'Amérique : x2 - x3 aux US entre 2023 et 2028, pourrait dépasser les 10 % de la consommation électrique du pays ⁽⁹⁾.
- Étude ADEME pour la France : x2 entre 2024 et 2030, x3,7 entre 2024 et 2035.

L'intensité carbone de la consommation électrique des centres de données dépend directement du mix énergétique national ou régional dans lequel ils sont implantés. À l'échelle mondiale, environ 50 % de l'électricité est encore produite à partir de combustibles fossiles, ce qui explique que les émissions directes de GES de l'IA soient, pour l'essentiel, des émissions indirectes liées à la génération électrique. C'est là l'un des déterminants fondamentaux de l'empreinte carbone de la filière.

Pour l'ensemble de la filière des centres de données, les émissions mondiales étaient estimées à 400 millions de tonnes équivalent CO₂ en 2025. Les projections pour 2030 s'inscrivent dans une fourchette large, allant de 630 à 920 millions de tonnes eqCO₂ selon les hypothèses retenues quant à l'avancement des engagements climatiques des États ⁽¹⁰⁾.

3) La consommation en eau : un enjeu territorial

La consommation d'eau des centres de données est un impact environnemental encore insuffisamment pris en compte dans les débats publics. Elle se décompose en deux postes : le refroidissement direct des infrastructures informatiques et, de façon

(8) <https://www.iea.org/reports/energy-and-ai>

(9) <https://eta-publications.lbl.gov/sites/default/files/2024-12/lbni-2024-united-states-data-center-energy-usage-report.pdf>

(10) D'après l'audition de Marlène de Bank, *The Shift Project* ; cf. *The Shift Project*, « Centres de données et transition énergétique », rapport du 1^{er} octobre 2025. Chiffres 2025 : 400 MtéqCO₂ pour la filière mondiale des centres de données. Projections 2030 : entre 630 MtéqCO₂ (scénario bas, engagements climatiques respectés) et 920 MtéqCO₂ (scénario haut, transition contrariée).

souvent ignorée, l'eau consommée pour la production d'électricité qui les alimente – notamment l'hydroélectricité.

L'étude « *Making AI thirsty* » a mis en évidence que l'entraînement du seul modèle GPT-3 avait nécessité la consommation de 5,4 millions de litres d'eau ⁽¹¹⁾. Plus récemment, Google a divulgué que son ratio d'efficacité hydrique s'établissait à 1 litre d'eau par kilowattheure consommé ⁽¹²⁾. L'entraînement du modèle Mistral, documenté dans le cadre d'une analyse de cycle de vie (ACV) pilotée par l'ADEME, a mobilisé 281 000 mètres cubes d'eau ⁽¹³⁾.

Ces impacts revêtent une dimension territoriale particulièrement aiguë : les centres de données ont une durée de vie de l'ordre de trente ans, et leur implantation génère des tensions durables sur les ressources en eau locales, notamment dans les régions semi-arides ou en situation de stress hydrique. À Marseille, l'un des principaux hubs de connexion des câbles sous-marins, l'implantation de centres de données dans le Grand Port Maritime a retardé l'électrification à quai des navires et la recharge des bus électriques municipaux, des conflits d'usage reconnus par le gestionnaire du réseau de transport d'électricité (RTE) lui-même ⁽¹⁴⁾.

Par ailleurs, l'impact des centres de données sur la biodiversité est un aspect souvent ignoré. En Espagne, un projet de campus de Meta a pourtant été remis en cause en raison de son impact sur les zones de protection de l'aigle impérial ibérique et du vautour noir ⁽¹⁵⁾.

(11) Li P. et al., « Making AI Less Thirsty : Uncovering and Addressing the Secret Water Footprint of AI Models », University of California Riverside, 2023. La consommation d'eau déclarée pour l'entraînement de GPT-3 s'établit à 5,4 millions de litres.

(12) Google, « Environmental Report 2024 » ; efficacité hydrique déclarée à 1 litre d'eau par kilowattheure consommé. D'après l'audition de Marlène de Bank, *The Shift Project*.

(13) Contribution écrite de l'ADEME ; ADEME, analyse de cycle de vie (ACV) du modèle Mistral, 2024-2025. L'entraînement du modèle Mistral Large a généré 20 000 tCO₂eq et mobilisé 281 000 m³ d'eau.

(14) D'après l'audition de Marlène de Bank, *The Shift Project*.

(15) D'après l'audition de Loup Cellar (sociologue des sciences et des techniques, Observatoire sur l'impact environnemental de l'IA).

II. UNE QUANTIFICATION LACUNAIRE DANS UN CONTEXTE DE CROISSANCE EXPONENTIELLE

A. Des obstacles méthodologiques structurels

1) L'opacité des acteurs et l'absence de méthodologie commune

L'évaluation de l'empreinte environnementale de l'IA se heurte à un obstacle de premier ordre : l'absence de transparence des acteurs dominants du secteur.

Comme le souligne l'ADEME, bien qu'il soit « techniquement possible d'évaluer les impacts environnementaux des modèles d'IA, le manque de transparence des acteurs principaux rend impossible la quantification précise de ces impacts »⁽¹⁶⁾. L'INRIA confirme ce diagnostic en observant que « les acteurs du numérique publient des bilans donnant des informations parcellaires quant aux consommations énergétiques et aux impacts environnementaux des systèmes d'IA qu'ils développent »⁽¹⁷⁾.

À cette opacité s'ajoute l'absence de méthodologie commune de mesure et de *reporting*. Sans référentiel partagé, les chiffres publiés ne sont pas comparables entre eux. Une même empreinte peut varier du simple au triple selon le périmètre retenu, les hypothèses d'allocation et le facteur d'émission utilisé pour l'électricité. Une distinction fondamentale doit à cet égard être établie entre les modèles dits « propriétaires » – ceux d'OpenAI, Google ou Anthropic, pour lesquels l'opacité est quasi-totale – et les modèles « ouverts » comme Mistral, pour lesquels des données partielles sur la consommation à l'usage peuvent être obtenues. La plateforme *Hugging Face* propose ainsi un tableau de bord comparatif limité aux modèles ouverts, mais il demeure insuffisant pour rendre compte de la réalité globale du marché⁽¹⁸⁾.

2) La complexité de l'imputation des usages

La mesure « par requête » est l'unité de communication la plus fréquemment utilisée par les acteurs du secteur, mais cette métrique pose de sérieux problèmes de robustesse analytique. En premier lieu, les impacts par requête varient considérablement selon la nature de la requête – textuelle, audio, vidéo – la taille du

(16) Contribution écrite de l'ADEME.

(17) Contribution écrite de l'INRIA.

(18) [Models – Hugging Face](#) ; [AI Energy Score Leaderboard - a Hugging Face Space by AI Energy Score](#)

modèle sollicité et la qualité de la réponse attendue. En second lieu, une requête isolée n'a pas de signification environnementale pertinente : ce qui compte est l'usage annuel agrégé, qui intègre l'ensemble des phases d'entraînement, de réentraînement et d'inférence. Enfin, la métrique par requête a pour effet pervers de reporter la responsabilité environnementale sur les utilisateurs individuels, occultant la part décisive que représente l'entraînement, à la charge exclusive des développeurs de modèles.

Comme le note l'INRIA dans sa contribution écrite :

« La rareté des informations disponibles oblige à des estimations lacunaires. Afin de sensibiliser les utilisateurs des systèmes d'IA, de nombreuses comparaisons circulent comme certaines donnant la quantité d'eau nécessaire au refroidissement pour chaque requête effectuée sur un système d'IA générative. Si ces chiffres ont été obtenus rigoureusement, ils ne sont valables que pour un type de requête donné, sur un système donné et peuvent difficilement être généralisés ⁽¹⁹⁾. En effet, entre une simple requête textuelle et une requête nécessitant la génération d'images, l'énergie nécessaire à la réalisation des tâches peut différer de plusieurs ordres de grandeur. »

Une autre complexité tient à l'intégration croissante de l'IA dans des services numériques génériques. Les moteurs de recherche intègrent désormais nativement des résumés générés par IA – « AI Overviews » chez Google, « Copilot » chez Microsoft –, rendant de plus en plus floue la frontière entre recherche « classique » et requête IA. Une requête adressée à un moteur de recherche doté d'une couche IA serait aujourd'hui de cinq à dix fois plus consommatrice en énergie qu'une requête classique, selon plusieurs estimations ⁽²⁰⁾. L'étude de Google publiée à l'été 2024 chiffre la consommation d'une requête médiane sur son modèle Gemini à 2,4 wattheures ⁽²¹⁾. La comparabilité de ces données avec celles d'autres acteurs demeure cependant incertaine.

B. Les ordres de grandeur disponibles et les projections

1) Les données existantes

Malgré les obstacles mentionnés, quelques données fiables sont disponibles. L'analyse de cycle de vie du modèle Mistral, conduite par l'ADEME, constitue à ce jour l'effort de quantification le plus rigoureux sur un modèle européen : l'entraînement de ce modèle a généré 20 000 tonnes équivalent CO₂ et consommé 281 000 mètres cubes d'eau. Ces chiffres sont à mettre en regard du fait que Mistral est un modèle de taille

(19) Voir S. Luccioni [https://www.sshaluccioni.com/AI%20+%20Environment%20Primer%20\(Hugging%20Face\).pdf](https://www.sshaluccioni.com/AI%20+%20Environment%20Primer%20(Hugging%20Face).pdf)

(20) D'après l'audition de Gauthier Roussilhe. Une requête adressée à un moteur de recherche intégrant une couche d'IA générative consommerait de cinq à dix fois plus d'énergie qu'une requête classique, selon les estimations disponibles, avec une grande variabilité selon les périmètres méthodologiques retenus.

(21) Google, communiqué interne relayé dans la presse spécialisée, été 2024 : consommation d'une requête médiane sur le modèle Gemini établie à 2,4 wattheures (Wh).

intermédiaire : les grands modèles d'OpenAI ou de Google, dont les paramètres sont beaucoup plus nombreux, présentent une empreinte sans commune mesure.

Concernant ChatGPT, une estimation a pu être produite à partir du chiffre par requête divulgué par son directeur général, Sam Altman. Elle aboutit à une consommation quotidienne de l'ordre de 884 MWh²². Ce chiffre représente la consommation électrique de 145 000 personnes en France. Il est cependant qualifié de très conservateur par les experts auditionnés, en raison de l'absence de prise en compte des phases d'entraînement et de réentraînement.

La consommation d'électricité des centres de données en France croissait, en 2025, à un rythme de 13 % par an⁽²³⁾. Sur les 352 centres de données actifs en France, la puissance moyenne s'établissait à 7,5 mégawatts (MW) – très en deçà des mégacampus de plus de 1 GW en projet, à l'image du modèle américain⁽²⁴⁾.

2) Des trajectoires de croissance préoccupantes

Les projections disponibles convergent vers un constat de doublement à triplement de la consommation électrique des centres de données à horizon de cinq à dix ans.

Le rapport de l'AIE anticipe un doublement de la consommation mondiale des centres de données entre 2023 et 2030⁽²⁵⁾. Le Département de l'énergie des États-Unis d'Amérique projette, pour sa part, un doublement à triplement de la consommation américaine entre 2023 et 2028, avec la perspective que les centres de données dépassent 10 % de la consommation électrique totale du pays⁽²⁶⁾. L'ADEME, dans son

(22) Estimation réalisée à partir du chiffre par requête divulgué par Sam Altman, directeur général d'OpenAI (10 Wh par requête), appliqué au volume d'utilisation déclaré. Qualification de « très conservateur » : d'après l'audition de Gauthier Roussilhe. Cette estimation ne prend pas en compte les phases d'entraînement et de réentraînement.

(23) D'après l'audition de Marlène de Bank, *The Shift Project* ; The Shift Project, rapport du 1er octobre 2025. La consommation électrique des centres de données en France croissait à un rythme de 13 % par an en 2025.

(24) D'après l'audition de Gauthier Roussilhe. En France, la puissance installée pourrait passer de 2,5 GW à 9 GW selon les scénarios de développement envisagés, ce qui représente un changement d'échelle considérable par rapport à la puissance moyenne actuelle des centres de données (7,5 MW).

(25) AIE, *Energy and AI, 2025* : doublement de la consommation électrique mondiale des centres de données entre 2023 et 2030.

(26) U.S. Department of Energy, « Powering Intelligence : Analyzing Artificial Intelligence and Data Center Energy Consumption », 2024 : doublement à triplement de la consommation américaine entre 2023 et 2028, avec un risque de dépassement des 10 % de la consommation électrique nationale.

étude prospective sur la France, prévoit un doublement entre 2024 et 2030 et une multiplication par 3,7 entre 2024 et 2035 ⁽²⁷⁾.

Deux éléments méritent d'être nuancés dans l'interprétation de ces trajectoires. D'une part, le rapport de l'AIE montre que la croissance de la consommation électrique des centres de données, rapportée à la consommation mondiale totale, demeure inférieure à celle d'autres secteurs économiques en phase d'expansion rapide. D'autre part, les incertitudes liées à l'évolution des architectures de modèles, aux innovations en matière d'efficacité énergétique et aux politiques publiques rendent délicate toute projection au-delà de 2030. L'INRIA souligne à cet égard que « les innovations entourant ces modèles sont constantes » et qu'il serait imprudent d'extrapoler mécaniquement les tendances observées.

Pour autant, la trajectoire préoccupe pour deux raisons de nature systémique. En premier lieu, l'énergie allouée aux centres de données ne sera pas disponible pour décarboner d'autres secteurs : dans un système électrique où la puissance installée renouvelable et nucléaire est une ressource finie à court terme, les arbitrages entre usages sont réels. En second lieu, la concentration géographique des centres de données exerce des pressions intenses sur les réseaux électriques et les ressources en eau à l'échelle locale, sans que les planifications territoriales n'aient généralement anticipé ces besoins.

C. Les limites du narratif des bénéfices compensatoires

1) Des cas d'usage favorables à l'environnement réels mais minoritaires

Le discours dominant des grandes entreprises du secteur, incarné notamment par le narratif de Sam Altman selon lequel « l'IA aura réponse aux problèmes qu'elle pose, y compris environnementaux », repose sur l'idée que les gains d'efficacité rendus possibles par l'IA compenseront, voire excéderont, les coûts environnementaux qu'elle engendre. Une étude conjointe de PwC et Microsoft de 2019 projette des gains

(27) Contribution écrite de l'ADEME ; ADEME, étude prospective sur les centres de données en France, 2025 : doublement de la consommation entre 2024 et 2030 (x2), multiplication par 3,7 entre 2024 et 2035.

environnementaux potentiels considérables liés à l'IA dans les secteurs de l'agriculture, de l'énergie et de la mobilité ⁽²⁸⁾.

L'ADEME apporte sur ce point un éclairage plus nuancé et plus rigoureux ⁽²⁹⁾. Une distinction s'impose entre les différents types d'IA. Pour les IA dites « traditionnelles », appliquées à l'optimisation industrielle, à la maintenance prédictive, à la modélisation climatique ou à l'amélioration des prévisions météorologiques, des gains environnementaux réels peuvent être documentés dans des contextes précis et délimités. En revanche, pour l'IA générative grand public, qui concentre la quasi-totalité des investissements et génère les impacts les plus significatifs, il n'existe à ce jour aucun cas d'usage connu bénéfique pour l'environnement. L'IA est également massivement mobilisée pour localiser et extraire des énergies fossiles, ce qui constitue un contre-effet environnemental direct.

L'étude IT4Green conduite par l'ADEME illustre cette limite : les bénéfices environnementaux des solutions numériques et d'IA ne sont jamais intrinsèques à la technologie mais dépendent du contexte précis de déploiement ⁽³⁰⁾. L'exemple du télétravail est éclairant : son bilan environnemental varie radicalement selon qu'il est pratiqué en zone urbaine dense, en périurbain ou en zone rurale. Le choix du scénario de référence dans les études d'impact est donc déterminant et souvent biaisé au profit de scénarios favorables.

2) L'effet rebond et les contradictions du modèle économique

L'ADEME introduit une distinction conceptuelle essentielle entre « frugalité » (entendue comme remise en question du besoin) et « efficience » (qui désigne l'optimisation de la consommation à service rendu constant, pour continuer de satisfaire le besoin).

Des modèles plus efficaces sur le plan énergétique n'entraînent pas nécessairement une réduction de la consommation totale si la baisse du coût marginal

(28) PwC / Microsoft, « Harnessing Artificial Intelligence for the Earth », 2019. Cette étude projette des gains environnementaux potentiels de 4,4 % des émissions mondiales de GES à l'horizon 2030 grâce à l'IA dans les secteurs de l'agriculture, de l'énergie et de la mobilité.

(29) Contribution écrite de l'ADEME.

(30) [Evaluation environnementale des effets directs et indirects du numérique pour des cas d'usage](#)

de chaque requête stimule une hausse du volume global d'utilisation⁽³¹⁾. C'est précisément ce mécanisme, connu sous le nom d'effet rebond ou paradoxe de Jevons, qui explique que les progrès d'efficacité dans l'histoire du numérique ont régulièrement été absorbés et dépassés par la croissance des usages.

Cette dynamique est structurellement entretenue par le modèle économique des acteurs dominants. Nvidia, qui fournit les GPU nécessaires à l'entraînement des modèles, les hébergeurs de centres de données et les producteurs d'énergie ont un intérêt économique direct à ce que la consommation continue de croître. Parallèlement, le cycle d'investissement massif dans l'IA repose sur la promesse d'une explosion des usages qui, si elle se réalise, implique mécaniquement une aggravation de l'empreinte environnementale. La sobriété et la viabilité climatique du secteur sont donc en tension structurelle avec sa logique de valorisation boursière.

III. VERS UNE GOUVERNANCE PUBLIQUE DE L'IA SOBRE : ENJEUX DE SOUVERAINETÉ ET LEVIERS D'ACTION

A. La dépendance technologique comme amplificateur des impacts environnementaux

1) Une dépendance structurelle aux acteurs américains

La question environnementale ne peut être dissociée de la question de la souveraineté et de la dépendance technologique.

Le numérique est un secteur caractérisé par une forte dépendance à des acteurs étrangers et extra-européens. Pour les équipements (*hardware*), la dépendance est structurelle à l'égard de la Chine, qui contrôle une part prépondérante des capacités d'extraction et de raffinage des métaux critiques. Pour les services (*software*), les acteurs américains (États-Unis) concentrent la quasi-totalité des logiciels, des plateformes cloud et des modèles d'IA utilisés en Europe.

Selon une étude du cabinet Astérès, 80 % des dépenses logicielles et de services cloud des entreprises et administrations de l'Union européenne sont dirigées vers des

(31) Contribution écrite de l'ADEME.

acteurs américains (États-Unis), représentant 265 milliards d'euros annuels⁽³²⁾. Comme le souligne Théophile Lenoir, il est préférable de substituer au terme de « souveraineté numérique » la notion plus précise de « gestion des dépendances »⁽³³⁾. L'enjeu n'est pas de reproduire à l'identique les capacités industrielles américaines (États-Unis), ce qui est économiquement et technologiquement irréaliste pour l'Europe, mais de réduire les points de vulnérabilité les plus critiques tout en construisant des capacités autonomes dans des domaines ciblés.

L'IA renforce structurellement cet état de dépendance. Les modèles de fondation les plus puissants, ceux d'OpenAI, Google, Meta, Anthropic, sont américains (États-Unis) et bénéficient d'un accès à des ressources de calcul inaccessibles aux acteur·ices européen·nes. Le discours selon lequel l'acquisition de GPU américains et la construction de centres de données en France constitueraient un vecteur de souveraineté numérique est, à ce titre, fallacieux : disposer de centres de données sur le sol national pour faire tourner des modèles américains sur des serveurs américains ne confère aucune autonomie stratégique.

2) L'impossible course à la puissance de calcul

La tentation de répondre à la domination américaine par une stratégie de mimétisme, en cherchant à aligner les investissements publics et privés européens sur les niveaux de dépense des acteurs américains, se heurte à des limites à la fois économiques et environnementales.

Sur le plan économique, les dépenses d'investissement annuelles de TSMC, le fabricant taïwanais de semi-conducteurs, s'élèvent à 44 milliards de dollars par an, là où le « chips act » européen mobilise 43 milliards d'euros sur plusieurs années. L'écart d'échelle est structurel et ne peut être comblé à court terme (34).

(32) D'après l'audition de Théophile Lenoir (sociologue, coordinateur de l'Observatoire sur l'impact environnemental de l'IA) ; cabinet Astères, « La dépendance technologique aux softwares et cloud services américains », 2024 : 80 % des dépenses logicielles et de services cloud des entreprises et administrations européennes à destination d'acteurs américains, soit environ 265 milliards d'euros annuels.

(33) D'après l'audition de Théophile Lenoir.

(34) D'après l'audition de Marlène de Bank, *The Shift Project*. Les dépenses d'investissement annuelles de TSMC s'élèvent à environ 44 milliards de dollars par an, tandis que le « European Chips Act » mobilise 43 milliards d'euros sur plusieurs années.

Sur le plan environnemental, cette course est incompatible avec les objectifs climatiques de l'Union européenne. L'ADEME plaide en faveur d'une autre voie : celle du développement de petits modèles spécialisés, frugaux, répondant à des besoins précisément définis et compatibles avec les contraintes énergétiques du territoire. L'émergence de modèles tels que DeepSeek ou des architectures de type Mixture of Experts (MoE), qui permettent de n'activer qu'une fraction des paramètres pour chaque requête, illustre qu'il existe un espace technologique viable entre la frugalité et la puissance brute des méga-modèles.

Cependant, l'ADEME note également que la course aux investissements actuelle est très largement fondée sur l'idée d'une explosion des usages, ce qui entre en contradiction directe avec la recherche d'IAG plus sobres. Nous sommes ainsi dans une situation paradoxale où les entreprises d'IA auraient intérêt – en raisonnant à une échelle microéconomique – à aller vers une utilisation plus frugale mais où les incitations boursières – en raisonnant à l'échelle macroéconomique – les incitent à une explosion des usages, quand bien même cette explosion leur coûterait plus qu'elle ne leur rapporterait.

« Il n'est pas certain qu'une plus grande sobriété des modèles serait dans l'intérêt même des entreprises... Il existe effectivement une contradiction, en particulier avec le modèle économique actuel des fournisseurs d'IA : chaque requête leur coûte de l'argent, en fonction de la conso d'énergie associée, donc il pourrait être dans leur intérêt de réduire cette consommation.

Mais en parallèle le cycle d'investissements massifs dans l'IA repose sur l'idée que les usages vont exploser, ce qui est directement dans l'intérêt de nombreux acteurs de l'écosystème : Nvidia, hébergeurs cloud, ... Toute la puissance de calcul commandée et en installation doit donc être dépensée, sinon le cycle d'investissements va s'arrêter. Par exemple, la sortie d'une nouvelle version du modèle DeepSeek, annoncée comme beaucoup plus efficiente que les modèles états-uniens, avait fait chuter drastiquement le cours de bourse de Nvidia. »

B. Un cadre réglementaire en construction, encore insuffisant

1) Les avancées du cadre européen et national

Au niveau national, la Stratégie nationale pour l'intelligence artificielle (SNIA) adoptée fin 2021 a affiché l'objectif de développer une offre d'« IA frugale » comme l'un de ses axes prioritaires, mais sans en définir le périmètre ni en fixer les indicateurs de mesure. Des travaux de normalisation ont été engagés à l'AFNOR en lien avec le Commissariat général au développement durable (CGDD) pour élaborer un Référentiel IA frugale, dont le contenu reste à ce jour insuffisamment opérationnel.

Au niveau européen, le règlement sur l'intelligence artificielle (AI Act), adopté en 2024, constitue un cadre juridique pionnier en matière de régulation des risques liés à l'IA. Cependant, il aborde la dimension environnementale de façon marginale, privilégiant les enjeux de sécurité, de transparence algorithmique et de protection des droits fondamentaux.

La directive sur l'efficacité énergétique prévoit par ailleurs que les centres de données remontent des données sur leur consommation d'eau, ce qui constitue une avancée bienvenue mais encore insuffisante pour permettre une évaluation complète. L'ARCEP exerce, quant à elle, une fonction pionnière à l'échelle internationale en matière de collecte de données environnementales sur les opérateurs numériques.

2) Les lacunes persistantes

L'INRIA pointe que « pour une technologie se développant très vite et pour laquelle des autorisations de marché ne sont pas nécessaires, les lois, la réglementation et les normes peinent à suivre le rythme de l'innovation ». Plusieurs lacunes apparaissent structurelles.

En premier lieu, l'absence de communication des données obligatoire standardisée sur les impacts environnementaux des modèles d'IA empêche toute comparaison sectorielle et laisse le champ libre aux communications volontaires, sélectives et non vérifiables des acteurs. En second lieu, la planification de l'implantation des centres de données reste largement déficiente : l'État français a, par exemple, laissé la société américaine Digital Realty s'implanter dans le Grand Port Maritime de Marseille sans planification énergétique ou territoriale préalable.

Marlène de Bank (*The Shift Project*) souligne en outre que la directive européenne sur l'efficacité énergétique, bien que pertinente dans ses intentions, demeure « plus qu'insuffisante » au regard des trajectoires d'émissions projetées pour le secteur des centres de données. La « Décennie numérique » européenne (*Digital Decade*), dont la révision est prévue au second semestre 2026, constitue une opportunité pour intégrer des contraintes environnementales opposables aux acteurs du secteur.

C. Recommandations

1) Imposer la transparence et standardiser les métriques

La première des recommandations, sur laquelle l'ensemble des expert-es auditionné-es s'accordent, est l'obligation faite aux acteurs de l'IA de publier des données environnementales fiables, comparables et vérifiées.

Cette obligation suppose l'adoption d'une méthodologie standardisée et commune, couvrant *a minima* les émissions de gaz à effet de serre, la consommation d'électricité par phase (entraînement et inférence) et la consommation en eau. L'analyse de cycle de vie (ACV), dont l'étude Mistral a démontré la faisabilité, devrait être imposée pour tout modèle dépassant certains seuils de taille ou d'audience.

Cette exigence de transparence doit également s'appliquer aux achats publics : l'État et les collectivités territoriales devraient intégrer des critères de transparence environnementale et d'éco-conception dans leurs marchés informatiques, en s'appuyant sur le Référentiel général d'écoconception des services numériques (RGESN) de l'ARCEP et de l'Autorité de régulation de la communication audiovisuelle et numérique (ARCOM), ainsi que sur le futur Référentiel IA frugale de l'AFNOR.

Le levier de la commande public devrait, en outre, être massivement sollicité pour privilégier l'émergence et la consolidation des acteurs européens. De ce point de vue, l'Assemblée nationale pourrait donner l'exemple en dotant ses services d'un abonnement collectif à Mistral, pour inciter au développement d'un modèle d'IA ouvert et faisant l'effort d'évaluer son coût environnemental par une analyse en cycle de vie.

2) Planifier l'implantation des centres de données

L'absence de planification territoriale de l'implantation des centres de données constitue l'une des lacunes de la politique publique en faveur du numérique. L'État devrait se doter d'un cadre de planification contraignant, qui conditionne l'autorisation d'implantation des centres de données à leur compatibilité avec les trajectoires de consommation électrique définies dans la programmation pluriannuelle de l'énergie (PPE) et les schémas régionaux d'aménagement, de développement durable et d'égalité des territoires (SRADDET) ⁽³⁵⁾.

Cette planification devrait intégrer les conflits d'usage entre centres de données et besoins en énergie décarbonée des autres secteurs – transport, industrie, logement –, ainsi que les impacts sur la ressource en eau. Elle devrait également favoriser la mutualisation des rejets thermiques des centres de données dans les réseaux de chaleur urbains, qui constitue l'une des rares externalités positives exploitables de ces infrastructures.

3) Orienter l'investissement public vers des modèles frugaux

Les investissements publics dans l'IA, qu'il s'agisse des crédits du Programme et équipements prioritaires de recherche (PEPR IA) d'INRIA ou des financements mobilisés dans le cadre du plan « IA pour la France », devraient être prioritairement orientés vers le développement de petits modèles spécialisés, respectueux du règlement européen sur l'IA, répondant à des besoins précisément définis et affichant des profils d'empreinte environnementale documentés.

L'INRIA souligne à cet égard que la nécessité de proposer des systèmes d'IA performants mais aussi faciles à maintenir et ayant un coût de fonctionnement modéré est au cœur de sa stratégie de recherche. Des architectures innovantes – *Mixture of Experts (MoE)*, *Tiny Recursive Models* – ainsi que des approches bio-inspirées comme le calcul neuromorphique promettent des réductions significatives de la consommation énergétique et méritent un soutien public accru. Le ciblage des cas d'usage pertinents pour la transition écologique – amélioration des prévisions

(35) D'après l'audition de Gauthier Roussilhe (chercheur indépendant, spécialisé dans les impacts environnementaux du numérique), 2025.

météorologiques, optimisation des réseaux d'énergie, surveillance de la biodiversité – devrait guider les priorités d'allocation.

Vers des modèles d'IA plus sobres ? ⁽³⁶⁾

Il y a une nécessité pour les acteurs du numérique de proposer des systèmes d'IA performants mais aussi faciles à maintenir et ayant un coût de fonctionnement modéré. Ceci implique donc des innovations et de la recherche quant à la conception et à l'entraînement des systèmes d'IA (Retrieval Augmented Generation - RAG, Model Context Protocol - MCP, pruning, quantisation, ...).

Parmi ces innovations, on peut mentionner des techniques d'architectures de modèles récentes, utilisées dans des modèles tels que Mistral ou *DeepSeek*, telles que les Mixture of Experts (MoE), qui permettent de rendre des modèles plus sobres tout en maintenant des performances élevées. Dans cette architecture, seul 1 à 3 « experts » (sur 8 à 64) sont activés par requête, contre l'intégralité des paramètres dans un modèle dense. Le MoE permet d'augmenter la taille du modèle sans augmenter proportionnellement sa consommation énergétique.

Les recherches sur l'IA frugale sont au cœur de la stratégie d'INRIA et des projets de recherche sélectionnés et financés. Ainsi, au sein du PEPR IA (Programme et équipements prioritaires de recherche pour l'intelligence artificielle) ⁽³⁷⁾, plusieurs projets de recherche sont spécifiquement dédiés à l'IA frugale ⁽³⁸⁾.

Des technologies d'IA basées sur des principes bio-inspirés font l'objet d'intenses recherches, certaines commencent à être opérationnelles. Elles promettent des réductions significatives de l'impact environnemental.

Enfin, de nouvelles approches frugales, comme les *Tiny Recursive Models*, qui sont des modèles compacts, avec peu de paramètres, et récursifs (c'est-à-dire que le modèle s'auto-appelle ou réutilise ses propres résultats intermédiaires pour traiter des tâches complexes) semblent fournir des alternatives à des modèles intégrant de très grandes quantités de paramètres. Ces approches récentes laissent entrevoir que la recherche scientifique sur le développement de modèles frugaux (et davantage adaptés à l'usage final) reste dynamique.

(36) D'après la contribution écrite de l'INRIA.

(37) Voir <https://www.pepr-ia.fr/>

(38) Voir notamment <https://www.pepr-ia.fr/projet/holigrail/> et <https://www.pepr-ia.fr/projet/sharp/>

4) Renforcer la régulation au niveau européen

L'essentiel des leviers d'action pertinents excède le cadre national et appelle une réponse européenne coordonnée. La révision de la Décennie numérique prévue en 2026 constitue une occasion à saisir pour intégrer des objectifs contraignants de réduction de l'empreinte environnementale du numérique et de l'IA, en cohérence avec la Stratégie nationale bas-carbone (SNBC) et le pacte vert européen. Le Parlement européen devrait plaider pour que l'AI Act soit complété par un volet environnemental opposable, incluant des exigences de reporting et des seuils d'empreinte maximums pour les modèles mis sur le marché.

La création d'une agence internationale de régulation de l'IA, sur le modèle de l'Agence internationale de l'énergie atomique (AIEA), dotée d'un mandat couvrant explicitement les impacts environnementaux serait, *in fine*, la solution la plus ambitieuse à l'échelle internationale.

L'ARCEP fait figure de modèle international en matière de collecte de données sur l'empreinte des acteurs numériques. Au niveau européen, l'AI Act est le cadre réglementaire le plus complet à l'échelle internationale en matière d'IA. La France et l'Union européenne seraient donc légitimes à prendre la tête des négociations internationales pour la création de cette agence.

Le discours selon lequel « moins de régulation renforce notre compétitivité » est à cet égard un contre-sens : l'absence de régulation environnementale avantage structurellement les acteurs américains et fragilise les acteurs européens dont la différenciation ne peut reposer que sur la qualité, la fiabilité et la soutenabilité de leurs services.

CONCLUSION

L'intelligence artificielle générative représente un défi environnemental de premier plan. Son empreinte ne se limite pas à la consommation électrique des centres de données mais s'étend à l'extraction de ressources naturelles rares, à la consommation d'eau, aux impacts sur la biodiversité et, de façon indirecte, à l'ensemble des effets de second rang qu'induisent ses usages sur les comportements et la consommation énergétique des ménages et des entreprises.

Le premier impératif est celui de la transparence. Tant que les acteurs dominants ne seront pas tenus de publier des données environnementales fiables et comparables, les pouvoirs publics seront condamnés à légiférer à l'aveugle.

Le second impératif est celui du choix de société. Le développement de l'IA générative grand public n'est pas une fatalité technologique : c'est le produit de choix d'investissement, de modèles économiques et de politiques publiques. La question de savoir quelle IA nous voulons – à quels usages elle répond, pour qui elle crée de la valeur et à quels coûts environnementaux et sociaux – est une question éminemment politique, qui ne saurait être laissée à la seule appréciation des acteurs de marché.

Les auditions conduites dans le cadre de cette communication ont mis en évidence l'existence d'une voie alternative crédible : celle du développement de modèles frugaux, spécialisés, transparents et souverains, construits sur la base d'une coopération scientifique européenne renforcée et d'une régulation ambitieuse. Cette voie est plus étroite que celle de la démesure des méga-modèles, mais elle est la seule compatible avec les engagements climatiques de la France et de l'Union européenne.

LISTE DES PERSONNES AUDITIONNÉES

Auditions réalisées

- Agence de la transition écologique (ADEME)
- Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (ARCEP)
- M. Gauthier Roussilhe, chercheur indépendant spécialisé dans les impacts environnementaux du numérique
- Mme Marlène de Bank, ingénieure, The Shift Project
- M. Loup Cellar, sociologue des sciences et des techniques, Observatoire sur l'impact environnemental de l'IA
- M. Théophile Lenoir, sociologue, coordinateur de l'Observatoire sur l'impact environnemental de l'IA

Contributions écrites

- Institut national de recherche en sciences et technologies du numérique (INRIA)
- Agence de la transition écologique (ADEME)

EXAMEN EN COMMISSION

M. le vice-président Laurent Mazaury. L'ordre du jour appelle la présentation de Mme Danièle Obono d'une communication sur le coût environnemental de l'intelligence artificielle (IA) et d'un projet d'avis politique sur le sujet.

Mme Danièle Obono, rapporteure. Le développement massif des systèmes d'intelligence artificielle générative et leur appropriation par le grand public sont l'un des faits structurants de la décennie 2020 et s'inscrivent, plus largement, dans la numérisation croissante de nos sociétés, à l'œuvre depuis le début du siècle.

La révolution numérique peut être une chance immense pour l'humanité. Elle peut permettre un progrès humain, scientifique et démocratique inédit, pour autant que cette technologie soit utilisée à ces fins. Or l'IA générative accentue, à bien des égards, les problèmes déjà soulevés par le numérique : accaparement par les multinationales, manipulation algorithmique, fracturation sociale, addictions aux écrans, violation de la vie privée, dépendances à des acteurs étrangers, et impact environnemental non négligeable. C'est sur ce dernier aspect, trop souvent invisibilisé, que porte ma communication, car la crise climatique n'en finit pas de s'accélérer.

Selon l'Organisation météorologique mondiale, le climat de la Terre n'a jamais été aussi déséquilibré depuis le début de ses observations. Les concentrations de gaz à effet de serre entraînent un réchauffement continu de l'atmosphère et de l'océan ainsi que la fonte des glaces. L'Europe est le premier continent qui se réchauffe le plus vite. Un tiers de la population vit déjà dans des zones touchées par la sécheresse. Aujourd'hui 60 à 70 % des sols européens sont dégradés, et notre biodiversité s'effondre. Plus de 80 % des habitats naturels du continent sont dans un mauvais état de conservation.

Ces changements rapides et à grande échelle se sont produits en quelques décennies mais auront des répercussions néfastes pendant des centaines, voire des milliers d'années. Il y a urgence à engager la bifurcation écologique à tous les niveaux, et notamment en matière de numérique.

Dans le cadre de cette communication, j'ai auditionné différentes chercheuses et chercheurs ainsi que les agences de l'État compétentes : l'agence de la transition

écologique (ADEME) et l'autorité de régulation des communications électroniques, des postes et de la distribution de la presse (ARCEP).

Avant d'exposer leurs enseignements, j'aimerais insister sur un point : les entreprises d'IA états-uniennes ont construit leur narratif marketing autour de l'idée que l'intelligence artificielle sera le grand remède miracle permettant l'éradication du cancer et le développement des technologies contre le changement climatique. Ce récit fallacieux est destiné précisément à faire parler, à attirer consommateurs et investisseurs afin de leur faire croire qu'ils participent, en achetant un abonnement, à une grande révolution.

Il existe une dimension potentiellement révolutionnaire de certaines applications d'intelligence artificielle à certains domaines précis. C'est ce qu'illustre par exemple le prix Nobel de chimie de l'année 2024 attribué à une équipe de chercheurs en intelligence artificielle pour leurs travaux sur la synthèse des protéines. Mais ces progrès ne sont pas le fait des grands modèles de langage promus à grand renfort de communication. Ils sont le fait de modèles spécialisés, précisément entraînés à des fins de recherche, sur des données ciblées.

Il ne faut donc pas confondre :

- les usages frivoles et énergivores des grands modèles de langage servant à synthétiser au choix une multitude d'images allant du souverain pontife en vêtement de bain aux parodies de personnalités politiques ou des vidéos de chiens et de chats se comportant comme des humains ;

- et les usages utiles, bien plus frugaux et silencieux de modèles d'intelligence artificielle appliqués à certaines recherches scientifiques de pointe.

Cette précision faite, je voudrais vous présenter les principales observations de ma communication sur les différents aspects du coût environnemental de l'IA et les difficultés de leur prise en compte dans le cadre réglementaire actuel, avant d'esquisser quelques recommandations.

Le coût environnemental de l'IA est lié à trois facteurs principaux :

- la consommation d'électricité des centres de données et le type d'énergie utilisée pour produire cette électricité (charbon, hydroélectricité, nucléaire...) ;

- la fabrication des équipements : semi-conducteurs, cartes graphiques (GPU), serveurs, équipements réseaux, mais aussi les bâtiments abritant les centres de données ;
- la consommation d'eau des centres de données pour leur refroidissement.

S'agissant de l'électricité, les projections de consommation électriques des centres de données sont toutes alignées sur un facteur 2 ou 3. L'IA contribue à environ 10 % de la consommation des centres de données et cette proportion pourrait bien sûr augmenter compte tenu de la hausse rapide des usages de ces systèmes par le grand public.

S'agissant des équipements, la fabrication des équipements représente environ 25 % de l'empreinte carbone totale de la filière numérique à l'échelle mondiale. Cette proportion monte à 50 % en France, où le caractère faiblement carboné du mix électrique réduit mécaniquement le poids relatif de la consommation d'énergie à l'usage.

S'agissant de l'eau, l'étude *Making AI thirsty* a mis en évidence que l'entraînement du seul modèle GPT-3 avait nécessité la consommation de 5,4 millions de litres d'eau. Plus récemment, Google a divulgué que son ratio d'efficacité hydrique s'établissait à 1 litre d'eau par kilowattheure consommé. L'entraînement du modèle Mistral, documenté dans le cadre d'une analyse de cycle de vie (ACV) pilotée par l'ADEME, a, quant à lui, mobilisé 281 000 mètres cubes d'eau.

Notons que la consommation d'eau revêt une dimension territoriale particulièrement aiguë : les centres de données ont une durée de vie de l'ordre de trente ans, et leur implantation génère des tensions durables sur les ressources en eau locales, notamment dans les régions semi-arides ou en situation de stress hydrique. À Marseille par exemple, l'implantation de centres de données dans le Grand Port Maritime a retardé l'électrification à quai des navires et la recharge des bus électriques municipaux.

Ce coût environnemental est difficilement maîtrisable en l'état et n'est, de fait, pas abordé par le règlement sur l'IA adopté en 2024. Cela tient à trois principales difficultés.

La première de ces difficultés, évoquée par l'ensemble des acteurs lors des auditions menées, est l'absence de données disponibles.

En effet, les données relatives à la pollution des grands modèles de langage massivement utilisés de nos jours (comme ChatGPT, Claude, Gemini, Grok...) ne sont pas disponibles et les pollutions imputables à l'intelligence artificielle sont souvent elles-mêmes incluses sans plus de précision dans les pollutions numériques...

Pour prendre un exemple : il est difficile de savoir précisément la puissance de calcul allouée à des modèles d'IA au sein d'un centre de donnée pouvant servir à d'autres usages numériques. Cette difficulté tient au fait que la plupart des data centers commerciaux fonctionnent sur un modèle multi-tenants où plusieurs entreprises avec des usages différents peuvent cohabiter et utiliser la puissance de calcul pour leurs propres usages.

Ce problème est bien connu et nous a été rappelé par les acteurs auditionnés. Cependant, à ce jour, le cadre réglementaire européen, via l'AI Act adopté en 2024, reste centré sur la protection de l'utilisateur et omet la dimension environnementale.

La deuxième difficulté est davantage méthodologique et recouvre la question de la métrique. Autrement dit : comment mesurer le coût environnemental de l'IA ? La mesure « par requête » est l'unité de communication la plus fréquemment utilisée par les professionnels du secteur, mais cette métrique pose de sérieux problèmes de robustesse. En premier lieu, les impacts par requête varient considérablement selon la nature de la requête – textuelle, audio, vidéo – la taille du modèle sollicité et la qualité de la réponse attendue. En second lieu, une requête isolée n'a pas de signification environnementale pertinente : ce qui compte est l'usage annuel agrégé, qui intègre l'ensemble des phases d'entraînement, de réentraînement et d'inférence. Enfin, la métrique par requête a pour effet pervers de reporter la responsabilité environnementale sur les utilisateurs individuels, occultant la part décisive que représente l'entraînement, à la charge exclusive des développeurs de modèles.

Une autre complexité tient à l'intégration croissante de l'IA dans des services numériques génériques. Les moteurs de recherche intègrent désormais nativement des résumés générés par IA – « AI Overviews » chez Google, « Copilot » chez Microsoft –, rendant de plus en plus fine la frontière entre recherche « classique » et requête IA.

Une requête adressée à un moteur de recherche doté d'une couche IA serait aujourd'hui de cinq à dix fois plus consommatrice en énergie qu'une requête classique, selon plusieurs estimations.

Enfin, la troisième difficulté concerne la problématique du coût environnemental est elle-même intriquée aux dépendances que nous subissons en tant qu'Européens : dépendances à la Chine pour le hardware (équipements) et dépendances aux États-Unis pour les softwares (logiciels).

Le problème du coût environnemental de l'IA n'est pas isolé. Un chiffre établi par le cabinet Astérès en 2024 mérite d'être rappelé : 80 % des dépenses logicielles et de services cloud des entreprises et administrations de l'Union européenne sont dirigées vers des acteurs états-uniens, représentant 265 milliards d'euros annuels.

Ce chiffre met en lumière notre dépendance massive qui fait de l'Europe est une colonie numérique des États-Unis. Mais ce n'est pas tout. Vu sous l'angle de la sobriété et de l'empreinte environnementale, ce chiffre raisonne amèrement avec le célèbre adage populaire : « on n'est jamais mieux servi que par soi-même ».

De fait, la meilleure manière de construire des modèles d'IA vertueux, respectueux de l'environnement, est encore de disposer d'entreprises d'IA souveraines, françaises ou européennes.

Soyons clairs. Le discours selon lequel « moins de régulations renforcent notre compétitivité » est à cet égard un contresens : l'absence de régulation environnementale avantage structurellement les acteurs américains et fragilise les acteurs européens dont la différenciation ne peut reposer que sur la qualité, la fiabilité et la soutenabilité de leurs services.

Chercher à protéger le consommateur européen est une bonne chose. Mais cette ambition ne peut pas être la seule fin de l'action politique européenne. Il nous faut également obliger les grands modèles de langage à la transparence environnementale. Cette exigence sera également de nature à mettre en avant nos entreprises d'IA nationales et européennes, dont les efforts ne sont, pour lors, pas récompensés à leur juste valeur.

Dans le prolongement de ce propos, plusieurs recommandations peuvent être formulées.

D'abord et avant tout : il nous faut obliger les entreprises d'IA à une transparence absolue sur leur coût environnemental, tant dans la phase d'entraînement des modèles que dans leur phase d'inférence, c'est-à-dire d'utilisation par les consommateurs finaux.

Aucune action de politique publique ne pourra être entreprise en la matière tant que nous n'aurons pas d'indicateur fiable et harmonisé : MWh consommés, litres d'eau consommés, tonnes d'équivalent CO2 émises.

Ces données sont nécessaires à l'action publique autant qu'à la recherche pour améliorer la sobriété des modèles.

J'ajoute que la publication de ces données permettrait, sans doute, d'objectiver et de valoriser le caractère plus vertueux des solutions françaises et européennes.

Ainsi, l'entreprise d'IA française Mistral s'est prêtée en 2024 à une analyse en cycle de vie en lien avec l'ADEME qui a permis de déterminer précisément le coût environnemental de son modèle. Cette démarche, singulière dans l'écosystème, est à saluer et devrait être un passage obligé pour l'ensemble des modèles souhaitant accéder au marché français ?

Ensuite et cela rejoint le point précédent : l'État doit être exemplaire. Nous disposons pour cela d'un outil, je dirais même d'une « arme d'orientation massive » du marché en faveur de solutions nationales et européennes : je veux parler de la commande publique. Il n'appartient qu'à nous de doter les agents publics français et européens et les administrations critiques de solutions nationales et européennes pour donner l'exemple.

Enfin, il me semble essentiel de poursuivre les recherches dans le domaine de l'IA pour assurer des solutions plus sobres et mieux adaptées aux usages pertinents. Mais ces recherches doivent se faire tout en encadrant les usages compulsifs qui sont énergivores et abrutissants. Je pense à la génération d'images et de vidéos toutes plus inutiles les unes que les autres et parfois dangereuses pour l'information de qualité. Ces images et ces vidéos mobilisent des capacités de calcul et se déversent allègrement

sur des réseaux qui n'ont plus de « sociaux » que le nom. Il nous faut en avoir conscience, pour les encadrer en conséquence.

C'est l'objet de l'avis politique que je soumets à votre sagacité.

Mme Liliana Tanguy (EPR). L'intelligence artificielle prend une place croissante dans notre économie comme dans nos administrations. Elle constitue un levier important d'innovation, de compétitivité et, dans certains cas, d'amélioration de l'action publique.

Mais son développement a également un coût environnemental qu'il ne faut ni ignorer, ni surestimer. Les systèmes d'intelligence artificielle reposent sur des capacités de calcul très importantes qui impliquent des centres de données fortement consommateurs d'électricité et d'eau. À mesure que ces usages se développent, leur impact environnemental progresse.

Ce constat doit être posé avec lucidité. Pour autant, la réponse ne peut pas être de freiner par principe le développement de l'intelligence artificielle. Elle doit être de mieux l'encadrer.

C'est d'ailleurs le sens des travaux déjà engagés, en France comme au niveau européen. Sous l'impulsion du ministère de la Transition écologique, un référentiel pour une intelligence artificielle dite 'frugale' a été élaboré afin de définir des indicateurs de bonne pratique pour réduire concrètement son empreinte environnementale. Au niveau européen, le règlement sur l'intelligence artificielle, entré en vigueur en août 2024, est en cours de mise en œuvre.

Il nous faut donc poursuivre les efforts sans céder aux simplifications. Si cet avis comporte certaines intentions que nous pouvons entendre, nous ne pourrions pas voter en sa faveur.

D'abord parce qu'il se fragilise lui-même par des considérants qu'on qualifierait d'outranciers lorsqu'ils affirment, par exemple, que la numérisation de nos sociétés conduirait à une déconnexion du vivant, à une désagrégation du lien social. Un tel raccourci dogmatique n'apporte rien à l'analyse du coût environnemental de l'intelligence artificielle. Au contraire, il brouille le débat.

Ensuite, parce que cet avis révèle sa véritable arrière-pensée politique. Lorsqu'il propose de faire passer sous contrôle public les infrastructures du numérique et des télécommunications, nous ne sommes plus dans l'encadrement écologique de l'intelligence artificielle, ni dans la recherche d'une plus grande sobriété numérique. Nous sommes dans une logique de quasi-nationalisation, profondément anti-libérale, qui dépasse très largement l'objet du texte et qui, à nos yeux, n'apporte pas de réponse crédible aux défis environnementaux soulevés.

Pour toutes ces raisons, le groupe EPR votera contre cet avis politique.

Mme Danièle Obono, rapporteure. Je relève des points de constat convergents, mais également des caricatures un peu simplistes. Il me semble avoir indiqué dans la présentation de cette communication et dans l'avis, qu'il ne s'agit pas de stopper ou de freiner le développement de ces technologies numériques, mais de les encadrer et de les réguler pour qu'elles servent précisément des buts et des objectifs politiques.

C'est bien notre rôle et notre responsabilité de nous assurer que des secteurs de développement économique et industriel peuvent servir l'intérêt général. Or, le modèle de développement inspiré notamment des États-Unis d'Amérique ne correspond pas à ce qu'on pourrait attendre en matière environnementale. Les propositions qui sont portées sont le fruit d'une discussion avec les acteurs du secteur, qui parfois vont beaucoup plus loin que ce que j'ai pu présenter. Vous avez dû avoir connaissance d'une tribune, signée par 800 personnalités, qui exigent l'arrêt du développement de l'IA. Parmi ces personnalités, il y a des spécialistes du domaine, de grands chercheurs qui demandent un moratoire sur ce développement.

Nous ne portons pas cette proposition. Mais elle sonne peut-être l'alerte sur la manière dont se développe ce secteur. Nous ne sommes pas les seuls, et pas plus déraisonnables que l'ensemble de ces personnalités pour alerter sur l'impasse dans laquelle nous pourrions nous retrouver si nous ne prenons pas nos responsabilités de régulateurs. C'est le sens de cet avis politique.

Mme Sophia Chikirou (LFI-NFP). Je partage l'analyse dont vous nous rendez compte et les recommandations que vous faites. Elles sont courageuses, audacieuses, et, vous l'avez rappelé, ce ne sont pas les vôtres, en quelque sorte. Ce sont des

recommandations qui émanent de vos auditions, de l'ADEME, de l'ARCEP, de Shift Project, de l'Institut de recherche en sciences et technologies du numérique.

Tous les experts du domaine disent qu'il est temps de prendre conscience que nous devons agir sur la régulation, poser des limites et trouver les responsables de la pollution. Le principe que vous essayez de poser, c'est celui du pollueur-payeur. L'impact que vous décrivez, que ce soit sur les territoires, sur l'eau, sur les usages, sur l'aménagement du territoire, c'est notre responsabilité politique.

Je tiens aussi à rappeler que la France insoumise n'est pas dans le « technosolutionnisme ». M. Sam Altman dit, à propos de la pollution et de l'impact environnemental de l'IA, que « l'IA aura réponse aux problèmes qu'elle pose, y compris environnementaux ». Nous n'y croyons pas. Nous pensons que le monde étant fini, nous risquons d'atteindre ces limites avant même d'avoir trouvé la solution à son propre renouvellement, ce que nous avons appelé la règle verte.

La France insoumise n'est pas non plus dans le « technoscepticisme ». C'est une erreur que de nous interpréter de la sorte. En 2016-2017, le premier homme politique français qui aborde dans la campagne présidentielle le sujet du numérique et de l'intelligence artificielle, à une époque où personne ici ne savait même ce que cela voulait dire, c'est Jean-Luc Mélenchon. Nous le faisons avec un meeting, le 5 février 2017, où nous parlons des frontières de l'humanité et dans lequel nous posons la question du numérique. Il est le tout premier à parler d'IA générative, qui s'est développée à partir de 2022 et s'est démocratisée.

Vous avez choisi, je parle au groupe EPR, de vous opposer à cet avis pour une seule et unique raison : il pose des limites au capitalisme financier et vorace qui est en train de détruire notre souveraineté nationale, la souveraineté populaire, puisque nous n'avons pas la maîtrise de nos centres de données en France. Il nous impose également un modèle qui n'a pas d'intérêt public. Mais il profite de nos ressources, il profite de nos lois, il profite de notre argent. Puisque de l'argent public dans l'installation de centres de données, dans le développement, la recherche, il y en a beaucoup, et c'est une des niches fiscales les plus importantes dans notre pays.

Concernant la question du pollueur-payeur, centrale au sein de l'Union européenne mais aussi de l'ONU : comment rendons-nous aujourd'hui responsables les

générateurs de pollution qui sont complètement déresponsabilisés par rapport à la pollution qu'ils produisent ?

Mme Danièle Obono, rapporteure. Je vous remercie d'avoir rappelé le caractère résolument pragmatique de notre approche. Nous estimons en effet que ces technologies peuvent être mises au service de l'humanité et du progrès humain, à condition d'en assurer un contrôle politique effectif et de définir collectivement les finalités auxquelles elles doivent répondre, ce qui, à ce jour, fait défaut.

S'agissant de la question soulevée par Sophia Chikirou, elle touche précisément à un enjeu central. À l'échelle européenne, les grandes plateformes du secteur tirent aujourd'hui parti de la difficulté à établir avec précision le niveau de pollution généré, et, par conséquent, à en déterminer les responsabilités. Il en résulte un brouillage permanent entre, d'une part, la phase amont précédant l'usage individuel, qui relève des plateformes et des entreprises développant des systèmes d'intelligence artificielle, et, d'autre part, les usages individuels, largement médiatisés.

Cette confusion conduit à faire peser l'essentiel de la responsabilité sur les utilisateurs, ce qui constitue, selon nous, une difficulté majeure, dans la mesure où une part significative de cette responsabilité demeure invisibilisée. Par ailleurs, même en se limitant à la phase d'usage, il convient de rappeler que le modèle économique de ces industries repose sur l'incitation à la consommation, en particulier de contenus de divertissement, qui fonde largement leur valorisation économique et financière. À ce titre, ces acteurs portent une responsabilité dans les stratégies de marketing et de publicité qui encouragent ces dynamiques de consommation, voire de surconsommation.

C'est pourquoi il apparaît indispensable de préciser, dans le cadre de la législation européenne, des mesures et des indicateurs permettant d'encadrer ces pratiques, ainsi que d'imposer des obligations accrues de transparence. Il s'agit là du premier niveau de responsabilité sur lequel nous concentrons notre attention et pour lequel des avancées sont nécessaires.

Mme Marietta Karamanli (SOC). Je souhaite tout d'abord saluer l'intérêt du travail présenté. Ce rapport a, à tout le moins, le mérite de poser un débat que nous ne pouvons éluder, à savoir celui du caractère matériel, énergivore et consommateur

de ressources de l'intelligence artificielle. À cet égard, le diagnostic proposé est utile et, dans l'ensemble, les constats formulés apparaissent convaincants. Toutefois, certaines zones d'ombre subsistent.

La première tient à l'absence quasi totale d'évaluation économique des mesures proposées. Si le rapport formule des recommandations, parfois très ambitieuses, il ne comporte ni chiffrage des coûts de mise en œuvre, ni estimation des investissements publics ou privés nécessaires, ni analyse budgétaire des conséquences pour les collectivités, les opérateurs et l'État. Il aurait été souhaitable d'approfondir cet aspect.

La seconde zone d'ombre est d'ordre méthodologique. Le rapport décrit l'empreinte environnementale de l'intelligence artificielle, sans proposer de véritable comparaison entre différentes trajectoires possibles, qu'il s'agisse d'une politique de sobriété forte, d'une option intermédiaire, du statu quo ou d'un scénario d'accélération technologique. En l'absence de tels éléments comparatifs, il demeure difficile d'identifier un point d'équilibre pertinent.

Le troisième point concerne la question de la souveraineté. Le rapport évoque cette problématique sans toutefois organiser réellement l'arbitrage entre dépendance et coûts. Il souligne à juste titre la dépendance européenne à l'égard des acteurs américains et des chaînes de valeur extra-européennes. Néanmoins, la question demeure de savoir s'il convient de restreindre certaines technologies au risque de voir nos concitoyens recourir à des solutions extérieures, porteuses de normes et de valeurs qui ne sont pas les nôtres, ou s'il faut accepter un coût environnemental encadré afin de réduire ces dépendances. Le rapport gagnerait à mieux expliciter ces termes du débat et les choix politiques qui en découlent.

Enfin, si la communication met l'accent sur la transparence, la planification et la frugalité, elle n'indique pas suffisamment ce qu'il conviendrait d'accepter ou, au contraire, de proscrire. À titre d'exemple, l'intelligence artificielle peut, dans des domaines tels que la médecine ou la santé, sans se substituer à l'humain, apporter une valeur ajoutée significative. La médecine personnalisée, notamment fondée sur la génomique, suppose l'exploitation de volumes importants de données et peut être à

l'origine de bénéfices considérables. Une telle perspective aurait pu être davantage développée dans le rapport.

Pour conclure, si l'avis peut recueillir le soutien de notre groupe, des réserves demeurent. J'aurais souhaité qu'elles puissent être prises en compte dans la version définitive du rapport qui sera publiée ultérieurement.

Mme Danièle Obono, rapporteure. J'entends votre intervention comme une invitation à poursuivre les travaux. Le format retenu, celui d'une communication, présente en effet un caractère particulier, en ce qu'il ne constitue pas un rapport à proprement parler mais vise à ouvrir des pistes de réflexion. C'est dans cet esprit que j'ai souhaité formuler un avis, afin d'en poser les bases.

Un des enjeux principaux, en lien avec les travaux menés à l'échelle européenne dans lesquels nous nous inscrivons, réside dans l'accès aux données et dans l'exigence de transparence. À cet égard, l'obtention de ces informations ne représente pas un coût significatif. Elle suppose en revanche une volonté politique, la mise en place d'outils adaptés et l'instauration d'un cadre contraignant minimal permettant d'assurer la transmission effective de ces données par les acteurs concernés. Tel est, à mon sens, l'objectif principal vers lequel doivent converger nos efforts.

Par ailleurs, un élément marquant des auditions, que je n'avais pas encore évoqué, tient à la difficulté d'esquisser une véritable troisième voie française et européenne. Cette difficulté s'explique notamment par l'insuffisance de réflexion stratégique sur les objectifs poursuivis. À cet égard, la protection des consommateurs, qui constitue aujourd'hui la principale orientation affirmée par l'Union européenne, ne peut pas, à elle seule, servir de ligne directrice pour le développement de l'intelligence artificielle, en particulier si la dimension environnementale et les coûts associés ne sont pas pleinement pris en compte. L'avis que je présente se veut ainsi une invitation à approfondir cette réflexion.

Les auditions ont également mis en évidence le nombre limité d'exemples et de retours d'expérience relatifs aux stratégies mises en œuvre dans d'autres pays. L'exemple de l'Espagne a été évoqué, avec la tentative d'élaborer une stratégie distincte de la dynamique dominante à l'échelle internationale, souvent caractérisée

par un développement massif des infrastructures reposant largement sur les mécanismes de marché et une dérégulation poussée, à l'image du modèle américain. C'est précisément sur ce point que j'insiste dans l'avis. Le développement d'une stratégie européenne suppose de réguler et de protéger les initiatives existantes ou émergentes au sein des modèles français et européens.

À cette fin, il apparaît nécessaire de placer l'ensemble des acteurs sur un pied d'égalité. La transparence constitue, à cet égard, l'outil central dont nous disposons au niveau européen. Je partage pleinement l'idée qu'il convient de poursuivre ce travail et me tiens tout à fait disposée à le faire dans un esprit transpartisan au sein de notre commission.

M. Nicolas Bonnet (EcoS). Tout d'abord, je voudrais commencer par remercier Danièle Obono pour le travail accompli et les questions soulevées dans ce rapport, qui sont éminemment importantes.

À l'heure actuelle, tout le monde parle d'intelligence artificielle, mais sans forcément se poser les bonnes questions sur ce qu'il y a derrière. Les pouvoirs publics, notamment Mme la ministre, s'appuient sur les enjeux de souveraineté pour justifier le fait qu'il faudrait recourir à l'intelligence artificielle de manière systématique. Je l'ai même entendu dire en audition que si nos entreprises ne passaient pas toutes à l'IA, une bonne partie d'entre elles finirait par s'effondrer.

Je pense qu'à ce niveau de manque de vision critique, il est nécessaire de se poser davantage de questions. L'intelligence artificielle repose d'abord et avant tout sur le numérique. Or, le numérique est souvent qualifié de dématérialisé. En réalité, il repose profondément sur du matériel informatique et électronique. Aujourd'hui, lorsque l'on utilise un téléphone ou un ordinateur, on ne perçoit pas nécessairement tout ce qu'il y a derrière, qu'il s'agisse des serveurs, des data centers ou de l'ensemble des infrastructures nécessaires à leur construction et à leur production. Néanmoins, cela génère de nombreuses externalités négatives. Il est donc essentiel de les prendre en compte, de les examiner et d'avoir la lucidité ainsi que le courage de les regarder en face.

En effet, la production de matériel implique d'abord l'extraction de nombreuses matières premières, un processus très énergivore et souvent très polluant. Il a

également lieu, dans de nombreux cas, dans des pays où les populations sont exploitées. Ensuite, le raffinage de ces matériaux requiert également une quantité importante d'énergie, tout comme la fabrication du matériel électronique. Enfin, une fois ce matériel produit, son utilisation continue de nécessiter beaucoup d'énergie. S'agissant des data centers, leur multiplication s'explique par les besoins croissants de l'intelligence artificielle en capacité de calcul et de stockage. Leur fonctionnement nécessite également des ressources en eau pour le refroidissement.

L'ensemble de ces éléments souligne l'importance de rendre visibles ces externalités négatives. C'est précisément ce à quoi s'attache ce rapport, qui propose également de nombreuses pistes d'action. J'ai entendu qu'il ne résolvait pas l'ensemble des problèmes, mais cela ne constitue pas une difficulté : ce n'est pas son objectif. Son but est, selon moi, de nous éclairer sur les problématiques liées à ces enjeux.

C'est également ce que j'ai eu l'occasion de constater lors d'un colloque récent, organisé avec différentes associations de la société civile travaillant sur ces sujets. Nous parvenons toutes et tous aux mêmes conclusions : un besoin de régulation. Il s'agit d'ailleurs de notre rôle, tant au sein de cette assemblée qu'au niveau de l'Union européenne. J'appelle donc chacun et chacune d'entre nous à se saisir de ce rapport afin d'en tirer les régulations nécessaires pour prévenir les dérives qui y sont clairement exposées.

Mme Danièle Obono, rapporteure. Je souhaiterais appuyer le propos sur la question de la souveraineté, notion largement mobilisée. S'il y a un domaine dans lequel nous ne disposons d'aucune souveraineté, ni au niveau français ni au niveau européen, c'est bien celui de l'intelligence artificielle. Comme je l'ai indiqué, nous dépendons entièrement des États-Unis d'Amérique pour les logiciels et les services, et de la Chine pour les matières premières. En la matière, nous n'avons donc pas progressé. C'est la raison pour laquelle je rappelais l'intérêt des exemples français, qui sont notables, y compris en matière de sobriété, et que nous pouvons valoriser.

Cela s'inscrit également dans une volonté de considérer que la question de la bifurcation écologique doit être planifiée, et que l'État, non seulement au niveau national mais aussi au niveau des collectivités, doit jouer un rôle. Lorsqu'il est question,

par exemple, de conflits d'usage de l'eau pour décider de l'implantation d'un data center dans un territoire confronté à des problématiques d'accès et de distribution de l'eau, il s'agit aussi d'une question locale relevant des collectivités. Il est donc nécessaire de pouvoir en débattre de manière démocratique.

Il convient ainsi d'organiser l'ensemble de ces éléments. C'est pourquoi la dimension de régulation et l'élaboration d'une stratégie de l'IA en lien avec la soutenabilité écologique constituent des enjeux majeurs, sur lesquels nous avons pris du retard et qu'il nous faut désormais rattraper. Si nous voulons réduire nos dépendances - car nous ne sommes pas encore au stade de la souveraineté - il est nécessaire de progresser vers cet objectif, afin d'atteindre progressivement une souveraineté numérique et écologique.

M. le vice-président Laurent Mazaury. Aucun amendement n'ayant été déposé, je vous propose de mettre aux voix le projet d'avis politique.

Le projet d'avis politique est ***adopté***.

AVIS POLITIQUE
LE COÛT ENVIRONNEMENTAL DE L'INTELLIGENCE ARTIFICIELLE

La commission des Affaires européennes,

Vu l'article 88-4 de la Constitution,

Vu le traité sur le fonctionnement de l'Union européenne, en particulier ses articles 4, 11 et 191,

Vu la Charte des droits fondamentaux de l'Union européenne, en particulier son article 37 imposant un niveau élevé de protection de l'environnement,

Vu la communication de la Commission européenne du 11 décembre 2019 détaillant les grandes orientations du Pacte vert pour l'Europe,

Vu le règlement (UE) 2021/1119 du 30 juin 2021 établissant le cadre requis pour parvenir à la neutralité climatique, aussi connu sous le nom de « loi européenne pour le climat »,

Vu le règlement (UE) 2024/1689 du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle,

Vu le règlement (UE) 2026/667 du 11 mars 2026 modifiant le règlement (UE) 2021/1119 en ce qui concerne l'établissement d'un objectif intermédiaire de l'Union en matière de climat pour 2040,

Vu la seizième édition du rapport sur l'écart des émissions, publié en novembre 2025 qui révèle que les projections du réchauffement climatique au cours de ce siècle, basées sur la mise en œuvre complète des contributions déterminées au niveau national (CDN), sont maintenant de 2,3 à 2,5 °C, tandis que celles basées sur les politiques actuelles sont de 2,8 °C, soit au-delà des 1,5 °C visés par l'Accord de Paris et désormais hors d'atteinte ;

Considérant les alertes de l'observatoire Copernicus sur les conséquences climatiques liées à la hausse des émissions de gaz à effet de serre ;

Considérant que les projections de réchauffement climatique ne dispensent pas de l'obligation d'agir et supposent, au contraire, une exigence accrue à l'égard de l'ensemble des secteurs les plus émissifs ;

Considérant que la plupart des projections internationales et nationales font état d'une hausse d'un facteur deux ou trois de la consommation électrique des centres de données d'ici à 2030, principalement du fait de la nouvelle demande en capacités de calcul des IAG ;

Considérant les conséquences de l'installation des centres de données sur la biodiversité ;

Considérant que les centres de données représentaient 1,5 % de la demande d'électricité mondiale en 2024 d'après l'Agence internationale de l'énergie, soit 415 TWh et que l'IA contribuait à elle seule pour environ 10 % de cette demande ;

Considérant le fait que l'intensité carbone de la consommation électrique des centres de données dépend directement du mix énergétique national ou régional dans lequel ils sont implantés et qu'à l'échelle mondiale, environ 50 % de l'électricité est encore produite à partir de combustibles fossiles ;

Considérant, plus largement, l'ensemble des sources de pollution le long de la chaîne de valeur de l'intelligence artificielle (fabrication des équipements, consommation d'eau pour refroidir les centres de données...) ;

Considérant la hausse rapide des capacités de calcul requises par les modèles d'intelligence artificielle générative et leur adoption massive par le grand public ;

Considérant les risques à la fois sociaux, économiques et environnementaux que font peser ces technologies sur les sociétés contemporaines ;

Considérant la Coalition mondiale pour une intelligence artificielle durable lancée lors du sommet de Paris sur l'IA du 11 février 2025 ;

Considérant que la numérisation croissante des sociétés participe d'une déconnexion du vivant et d'une désagrégation du lien social caractéristiques du vivre-ensemble, insensibilisant une partie croissante de la population à l'impératif de protection de l'environnement et augmentant l'isolement des individus ;

Considérant que les arguments classiquement avancés par les entreprises d'IAG dominantes sur le marché, selon lesquels l'IAG permettrait de résoudre les grands problèmes humains comme le cancer et le changement climatique, relèvent d'un récit marketing avant tout destiné à rendre leurs produits attractifs, sans aucune garantie effective que les coûts environnementaux de ces modèles soient compensés par des gains futurs largement hypothétiques ;

Considérant l'opacité des entreprises d'IA quant aux coûts environnementaux de leurs modèles, tant dans la phase d'entraînement que dans la phase d'inférence, avec l'exception notable de l'entreprise française Mistral ;

Considérant que l'article 1er du règlement du 13 juin 2024 relatif à l'IA dispose que son objectif est « d'améliorer le fonctionnement du marché intérieur et de promouvoir l'adoption d'une intelligence artificielle (IA) axée sur l'humain et digne de confiance, tout en garantissant un niveau élevé de protection de la santé, de la sécurité et des droits fondamentaux consacrés dans la Charte, notamment la démocratie, l'état de droit et la protection de l'environnement, contre les effets néfastes des systèmes d'IA dans l'Union, et en soutenant l'innovation » et, qu'en tout état de cause, le choix devrait être fait de limiter le déploiement de l'IA dans le cas où les atteintes à l'environnement seraient inévitables ;

Estime qu'aucune action politique crédible de limitation de l'intelligence artificielle ne pourra être entreprise tant que des données transparentes et harmonisées sur le coût environnemental des modèles ne seront pas publiées ;

Estime qu'en l'absence de coopération internationale, l'option d'une troisième voie portée par la France, entre les stratégies chinoises et américaines, doit passer par le soutien des entreprises d'IA françaises et européennes qui se soumettent aux exigences de transparence et de réduction de leur impact environnemental ;

Invite le Gouvernement français et la Commission européenne à la création d'une mission nationale de maîtrise de l'intelligence artificielle, associant des chercheur·ses en intelligence artificielle et en sciences humaines et sociales, des associations de défense des libertés, des expert·es, avec pour rôle d'instruire et de conseiller la décision publique en la matière ;

Invite le Gouvernement français et la Commission européenne à la création d'une agence publique des logiciels libres chargée de planifier et financer leur développement stratégique domaine par domaine, notamment pour les besoins liés aux services publics, à la souveraineté numérique de la France et de l'Union européenne et à la bifurcation écologique ;

Invite le Gouvernement français et la Commission européenne à soutenir la création d'une agence de gouvernance mondiale d'Internet dédiée dans l'ONU, pour faire respecter la souveraineté numérique des États membres de l'Union européenne, protéger les libertés individuelles et mettre un terme à la mainmise des organisations de droit américain sur l'Internet mondial ;

Invite le Gouvernement français et la Commission européenne à soutenir l'ouverture et la publication des données d'intérêt général détenues par des acteurs privés pour en faire des biens communs ;

Invite le Gouvernement français et la Commission européenne à soutenir via la commande publique des services et logiciels français et européens grâce à une loi d'achat de biens produits spécifique ;

Invite le Gouvernement français à soutenir l'hébergement sur des serveurs de droit français situés en France des données des services publics et des secteurs clés de l'économie ;

Invite le Gouvernement français et la Commission européenne à soutenir le développement d'un cloud souverain, garantissant une maîtrise totale de la chaîne de production (matériel et logiciel) sans recours à de la sous-traitance ;

Invite le Gouvernement français et la Commission européenne à procéder au traitement et au stockage des données publiques, et particulièrement des données personnelles liées à la santé et à l'éducation, en les confiant à des structures publiques françaises ;

Invite le Gouvernement français et la Commission européenne à passer sous contrôle public les infrastructures du numérique et des télécommunications ;

Invite le Gouvernement français et la Commission européenne à inciter à un usage sobre et raisonné du numérique afin de lutter contre la surconsommation et la pollution numériques ;

Invite le Gouvernement français et la Commission européenne à interdire les usages inutiles et réguler les usages énergivores des serveurs ;

Invite le Gouvernement français et la Commission européenne à agir sur les modèles économiques des GAFAM reposant sur la captation de l'attention et la consommation sans limite de médias ;

Invite le Gouvernement français et la Commission européenne à planifier l'aménagement des centres de stockage de données pour garantir notamment la réutilisation systématique de la chaleur générée par leur fonctionnement et garantir un objectif zéro artificialisation nette du territoire ;

Invite le Gouvernement français et la Commission européenne au développement de la programmation éco-responsable (*green coding*) et l'éco-conception des outils

numériques matériels et immatériels, en l'introduisant dans les formations initiales et dans les critères de la commande publique ;

Invite le Gouvernement français et la Commission européenne à systématiser l'usage du référentiel général d'écoconception de services numériques dans les développements numériques des administrations publiques.