L'IA et l'environnement

L'IA est-elle un atout ou une catastrophe pour l'environnement ?

Guillaume Buthmann
Julie Cavé
Pierre de Ferrières de Sauveboeuf
Romain Guillaumé

Baptistin Heuberger
Juliette Jahan de Lestang
Florentin Picton

2024 - 2025

Cette publication a été réalisée par des étudiants en troisième année du cycle ingénieur de Mines Paris PSL Research University. Il présente le travail réalisé dans le cours intitulé « Descriptions de controverse », qui a pour objectif d'introduire les étudiants à l'univers incertain de la recherche scientifique et technique et de les sensibiliser aux enjeux de la participation citoyenne.

Mines Paris décline toute responsabilité pour les erreurs et les imprécisions que peut contenir cet article. Vos réactions et commentaires sont bienvenus. Pour signaler une erreur, réagir à un contenu ou demander une modification, merci d'écrire à la responsable de l'enseignement : madeleine.akrich@mines-paristech.fr

Sommaire

- Sommaire 3
- Introduction 4
- Partie 1: Quelle vocation de l'IA dans un contexte de crise environnementale ? 8
  - Le questionnement du besoin de l'IA 8
  - Des conséquences environnementales accrues 11
  - L'apport des outils numérique dans la lutte écologique 13
- Partie 2: Quels obstacles s'opposent à une mesure claire de l'empreinte environnementale de l'IA ? 16
  - Définition du spectre à prendre en compte : analyse de cycle de vie, matières premières, émissions carbone 16
  - Outils de mesure réplicables 18
  - Le problème de la transparence et de la disponibilité des données 20
  - Comptabiliser l'effet rebond 21
- Partie 3 : Compatibilité entre sobriété et compétitivité pour l'IA 23
  - Culture de l'innovation et de la croissance 23
  - Accepter une perte de rapidité et de précision pour la sobriété de l'IA 24
  - Interventionnisme et régulation des États 26
  - Approche à plusieurs échelles 28
- Conclusion 31
- Matériel et méthodes 34
- Références 35
  - Articles de presse généraliste / presse professionnelle 35
  - Articles de revue scientifique 35
  - Rapports 37
  - Sites Web 38

Introduction

Dès 2020, la presse scientifique met en exergue l'importance d'étudier conjointement les avancées de l'intelligence artificielle (IA) et son impact écologique : «*Il est devenu urgent de considérer le rôle de la technologie de l'IA dans la crise climatique.*» (Dhar, 2020)[1] L'environnement et l'IA semblent entretenir des rapports complémentaires et antinomiques dans de nombreuses sphères, qu'elles soient publiques, académiques ou industrielles. Toutefois, il est important de comprendre les définitions potentielles de l'IA afin d'étudier comment les différents acteurs les mettent en jeu dans un contexte environnemental.

En 1956, John McCarthy introduit pour la première fois le terme IA, et le décrira plus tard comme désignant *« la science et l'ingénierie de la fabrication de machines intelligentes »* (Jérémy, 2024). Néanmoins, cette définition n'est pas complète et selon un expert numérique du Shift Project, une association française créée en 2010 en tant que laboratoire d'idées qui s'est donné pour objectif l'atténuation du changement climatique, s'étonne de la variété et du nombre d'outils qui pourraient contenir de l'IA et ainsi de la définition même du terme :

> « Est-ce qu'un boîtier, un capteur connecté [...] avec un service d'IA, c'est un outil d'IA ou pas ? » (Entretien The Shift Project)

Le débat public semble manquer d'une définition claire et cette absence limite la compréhension de l'impact et du rôle de l'IA dans la lutte contre le réchauffement climatique. Le rapport de l'Association française des normes (AFNOR) 2314, propose finalement une définition d'un système IA réfléchie au sein d'un consortium de chercheurs, de législateurs (dont l'État français) et d'industriels : *« Système technique qui génère des sorties telles que du contenu, des prévisions, des recommandations ou des décisions pour un ensemble d'objectifs définis par l'humain »* (Fropier et al., 2024). Sous le prisme de cette définition, l'AFNOR propose un référentiel de l'impact environnemental de l'IA permettant de mettre en évidence ses effets ou avantages. Du fait de la complexité du domaine, la technologie avance souvent plus vite que la compréhension publique du sujet, posant alors la question de l'éthique liée aux progrès récents de l'IA. Des acteurs comme l'AFNOR choisissent donc de mettre l'accent sur la nécessité de terminologies normalisées pour tous les acteurs afin d'évaluer l'impact environnemental des systèmes.

Par ailleurs, l'émergence des questions environnementales dans le débat public a propulsé l'IA et le numérique dans son ensemble sur le devant de la scène. D'après une chargée de missions liées à l'IA au Ministère de la Transition Écologique (MTE), *« l'IA est une boîte noire pour quasi tout le monde. »* (Entretien du MTE), créant alors un paradoxe entre la technicité du sujet et ses enjeux écologiques communs. Dès 2019, la presse choisit d'utiliser des accroches chocs afin de sensibiliser ses lecteurs: *« L'ensemble du processus* [d'entraînement des modèles de langage] *émet autant de CO2 qu'un vol de 300 passagers de New York à San Francisco. »* (Niedercorn, 2019) Ce chiffre, largement dépassé

[1] Citation originale : *"It has become an urgent matter to consider the role of AI technology in the climate crisis"*

aujourd'hui, ainsi que les progrès du Deep Learning questionnent différents acteurs, dont des académiques du MIT sur la pertinence et les contreparties de l'IA du fait de *« la course à la puissance de calcul* [qui] *fait exploser le coût énergétique »* (Georges, 2020).

L'intelligence artificielle est apparue dans la lignée du développement du numérique qui se déroule déjà depuis plusieurs dizaines d'années. Le département américain de l'énergie, dans sa note *« AI For Energy »* présente l'intelligence artificielle comme un catalyseur d'innovation et la compare à une nouvelle révolution industrielle, en particulier dans le contexte de la transition vers une économie énergétique propre. On peut notamment y lire : *« Cette innovation* [confluence de changements dans le système électrique] *sera alimentée par une technologie capable de transformer de grandes quantités d'informations en capacités décisionnelles exploitables. »* (Benes et al., s. d.)[2]

Cette perception de l'IA comme une sous-catégorie du numérique nous invite à plonger brièvement au sein de la controverse autour du numérique. Les notes de l'Autorité de Régulation des Communications Électroniques, des Postes et de la distribution de la Presse (Arcep, 2019) au sujet de l'empreinte carbone du numérique décrivent les nombreux aspects discutés autour du numérique. De manière identique, un rapport du Shift Project (Ferreboeuf et al., 2020), traite de sujets liés aux numériques, tels que l'effet rebond, l'impact environnemental des usages, l'utilisation du numérique dans la transition énergétique… On remarque que les thématiques abordées sont proches des nœuds de la controverse que nous avons identifiés sur l'IA. Les discussions autour de l'IA paraissent donc n'être en réalité qu'un prolongement de la controverse autour des impacts écologiques du numérique. Pour poursuivre, en regardant les recherches Google incluant le terme IA, on remarque un intérêt en forte croissance pour le sujet dès 2016. La mise sur le marché de ChatGPT en novembre 2022 a marqué un tournant dans la perception de l'IA, notamment pour le grand public. La possibilité pour chacun de pouvoir utiliser, librement et gratuitement, des outils comprenant de l'IA a ajouté une dimension médiatique à cette controverse.

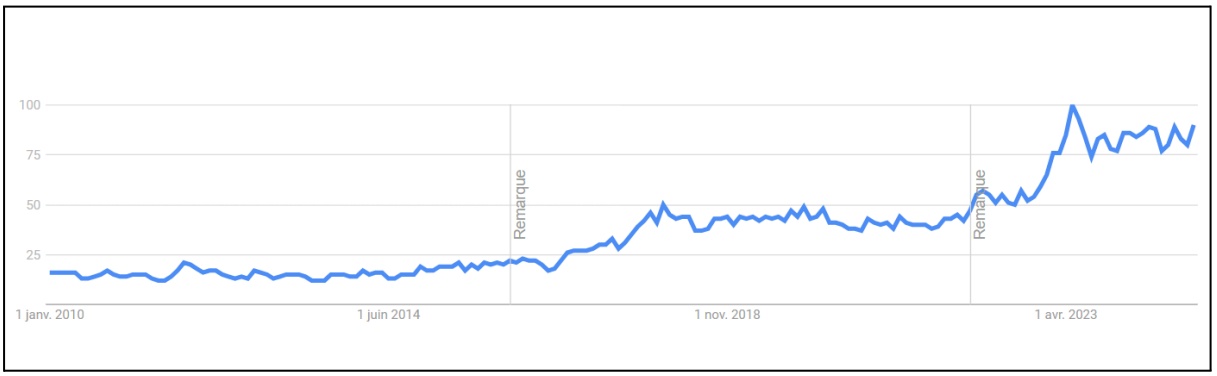


Figure 1 : Évolution des recherches sur Google contenant les mots « Artificial Intelligence » dans le monde depuis 2010.

[2] Citation originale : *"This innovation* [confluence of change across the electric power system*] will be driven by technology that can transform vast amounts of information into actionable, decision-making capabilities."*

D'un autre côté, cet engouement a poussé les grands acteurs, comme les GAFAM (Google, Amazon, Facebook, Apple, Microsoft), à investir massivement dans cette nouvelle technologie. On assiste à une nouvelle ruée vers l'or, comme décrite par le membre du Shift Project :

> « *Et pour moi, les objectifs des Big Tech, les objectifs sont directement liés aux investissements fabuleux qui sont faits, c'est juste de faire beaucoup d'argent. Très très clairement, les investissements absolument énormes qui sont générés par ces acteurs Amazon, Google, Microsoft, Oracle, Nvidia et autres… Quand on voit la capitalisation, quand on voit les attentes des investisseurs, et l'IA c'est pour faire quoi ? C'est pour avoir un retour sur investissement que l'on retrouve rarement sur le marché.* » (Entretien The Shift Project)

Les différents États se saisissent aussi à leur manière du sujet en travaillant sur une régulation adéquate et pertinente face à cette nouvelle technologie. On peut notamment citer le très récent IA act de l'Union européenne ou l'AFNOR Spec 2314 (Fropier et al., 2024). D'après un article du New York Times (Gelles, 2024), cette émergence soudaine de l'intelligence artificielle a mené à une remise en question des objectifs environnementaux de ces grandes entreprises : *« Les grands groupes technologiques misent sur une nouvelle génération de puces électroniques et de centres de données gourmands en énergie* »[3] *.*

Ainsi, la définition de l'IA a été fluctuante selon les époques. Elle l'est également aujourd'hui en fonction des acteurs rencontrés et de leur rapport à cet outil. Comment étudier le rapport de l'IA à l'environnement alors que sa définition même est problématique et non universelle ? D'autre part, après s'être penché sur l'histoire de cet outil et la construction de sa représentation dans l'imaginaire collectif, il faut noter que le développement général du numérique dont l'IA fait partie prend un rôle majeur dans nos sociétés alors même que la conscience écologique émerge. Cette histoire juxtaposée de l'utilisation accrue de l'IA et de la succession de traités environnementaux serait-elle en elle-même un paradoxe ? L'IA est à la fois présentée comme le poison et l'antidote par les nombreux acteurs. Pour certains d'entre eux, traiter cette controverse est une réelle urgence. On peut notamment retrouver cette notion d'urgence dans un article publié en 2020 dans la revue Nature Machine Intelligence (Dhar, 2020)[4] : *« Les Nations unies ont qualifié le changement climatique de « crise déterminante de notre époque » et, selon le Climate Reality Project, 97 % des climatologues s'accordent à dire que l'activité humaine en est le principal moteur.* »

De manière plus globale, on observe une augmentation des articles traitant du lien entre l'IA et l'environnement de manière significative au cours des deux dernières années, symbole d'un intérêt croissant pour cette thématique au sein de la sphère publique.

---

[3] Citation originale : "Big tech companies' bets on a new generation of power-hungry computer chips and data centers"

[4] Citation originale : *"The United Nations has called climate change a "defining crisis of our time", and, according to the Climate Reality Project, 97% of climate scientists concur that human activity is its main driver."*

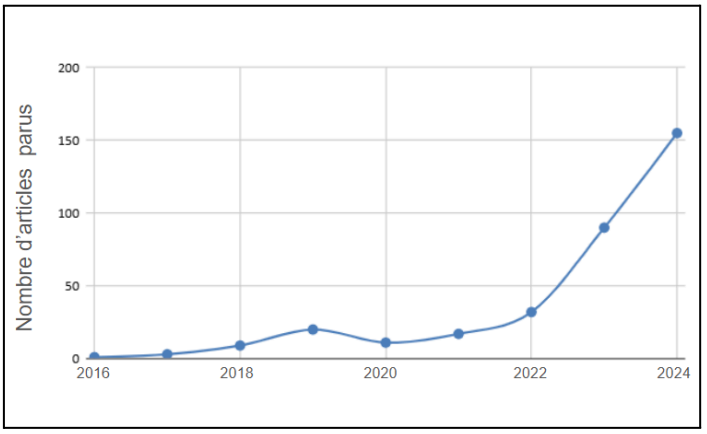


Figure 2 : Évolution du nombre d'articles parus depuis 2016 mêlant la thématique Intelligence Artificielle et l'environnement en France.

En outre, les effets sur l'environnement d'un objet mal défini pourraient être désastreux. Cette absence de consensus autour des métriques à utiliser, sur laquelle nous reviendrons notamment dans notre deuxième partie, invisibilise les effets de l'IA sur l'environnement. Dans l'article de Nature Machine Intelligence (Dhar, 2020)[5], on peut y lire : *« Crawford et Joler écrivent dans leur essai que les détails matériels des coûts des systèmes d'IA à grande échelle sont vagues dans l'imaginaire social, à tel point qu'un non-initié pourrait penser que la construction d'un système basé sur l'apprentissage automatique est une tâche simple* ».

Une incompréhension de ces effets pourrait, selon certains acteurs, mener à une utilisation massive de l'intelligence artificielle et une augmentation exponentielle de l'impact carbone de cette dernière, comme cela est craint dans un article publié par la revue European Radiology : « *L'IA dans les soins de santé est appelée à se développer, mais parallèlement à son développement, nous devons prendre des mesures conscientes et collectives en faveur de la durabilité à plus grande échelle.* » (Truhn et al., 2023)[6]

Nous nous attacherons à dégager une compréhension approfondie des interrogations des acteurs à propos de l'IA, de son rôle dans un contexte de crise environnementale, et de la façon dont leurs interactions successives font émerger trois problématiques sur lesquelles leurs opinions et actions convergent ou s'opposent. Dans un premier temps, nous examinerons la vocation de l'IA face à la crise environnementale : quelle place peut-elle réellement occuper ? Nous nous intéresserons dans un second temps à la mesure de son impact : comment évaluer de manière rigoureuse ses effets ? Enfin, nous analyserons la tension entre sobriété et compétitivité : est-il possible de concilier les exigences d'efficacité économique et écologique avec les efforts de réduction des impacts environnementaux ?

[5] Citation Originale : *"Crawford and Joler wrote in their essay that the material details of the costs of large-scale AI systems are vague in the social imagination, to the extent that a layperson might think that building a machine learning (ML)-based system is a simple task."*

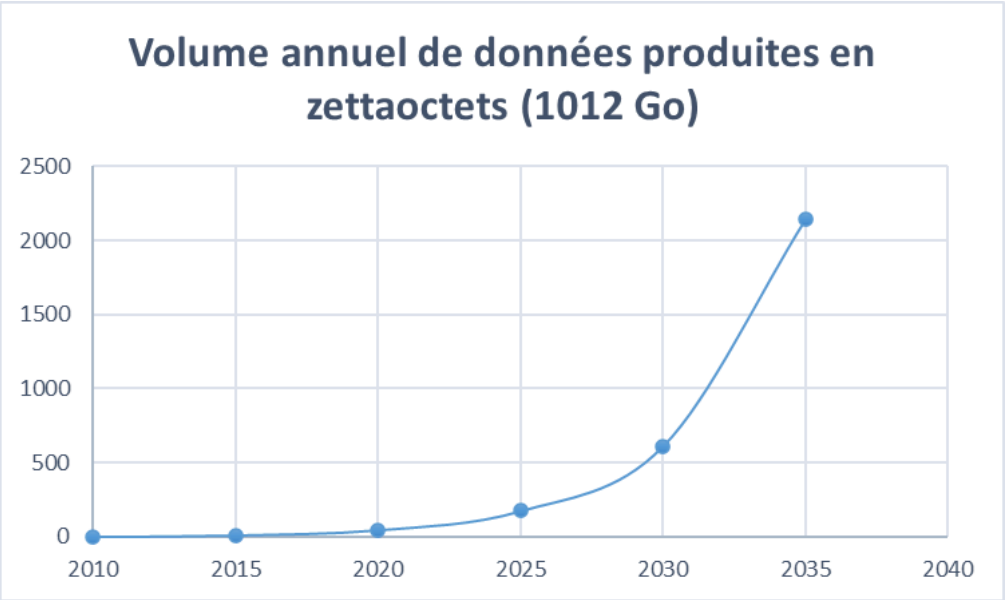
[6] Citation Originale : *"AI in healthcare is poised to grow, but alongside its development, we must take conscious, collective steps towards sustainability at a larger scale."*

Partie 1: Quelle vocation de l'IA dans un contexte de crise environnementale ?

Les enjeux liés à l'IA et la crise environnementale ont connu tous deux une forte émergence dans les dernières années. Mais quelle est la vocation de l'IA dans ce cadre ? Aucun accord sur l'utilisation qui devrait en être faite n'émerge au sein de cette controverse : des interrogations subsistent sur le besoin d'IA dans nos sociétés, alors que des débats sur les conséquences environnementales de ces outils et sur leurs contributions dans la transition écologique demeurent.

Le questionnement du besoin de l'IA

Quelque soient les usages, on assiste aujourd'hui à une explosion des données produites et échangées dans le monde. Selon Frédéric Bordage, *« Le poids d'une page web a été multiplié par 115 entre 1995 et 2005 »* (Ferreboeuf et al., 2020). C'est dans cette explosion du volume de données échangées partout autour du monde que l'IA paraît se développer, elle qui nécessite, pour son entraînement et sa pertinence un grand nombre de données à analyser.



Volume annuel de données produites depuis 2010, selon Statista, Digital Economy Compass 2019 via (Pitron, 2021).

Comme le défend la CLCV IDF (Consommation logement Cadre de vie de l'Ile de France) dans un rapport publié par l'ARCEP (Arcep, 2020), le développement des nouvelles

technologies du numérique, comme la 5G ou l'IA, est basé sur une économie du remplacement. Aucune étude d'impact n'a été faite avant le déploiement de la 5G affirme ce rapport de l'ARCEP, issu d'un groupe de travail de 2019 sur la soutenabilité économique regroupant associations, États et entreprises. Et c'est probablement dans un cadre similaire que se fait le développement à grande échelle de l'IA (Arcep, 2020; Ferreboeuf et al., 2020; Entretien The Shift Project) . *« On est sur une bulle d'investissement*[7]*, qui peut d'ailleurs être problématique »* affirme un ancien cadre de chez Oracle et membre du Shift Project (Entretien The Shift Project), faisant un parallèle avec la ruée vers l'or du XIXe aux États-Unis : *« Parce que si effectivement derrière on ne trouve pas d'or [...] on va très vite se retrouver devant un écroulement des chiffres »*. Accompagnant l'explosion des usages du numérique, le développement de l'intelligence artificielle reste concret : la consommation des data centers, indispensables pour le développement de l'IA, devrait être multipliée par 3 en vingt ans et passer de 10 à plus de 30 TWh en France selon le *Rapport d'information sur l'empreinte environnementale du numérique* publié par le Sénat en 2020 (Maurey et al., 2020).

Ce besoin est parfois remis en cause, des chercheurs universitaires Français affirment donc (Mallarino et al., 2022):

> « *Les outils numériques et la croissance continuelle des données de grande qualité et accessibles (imagerie satellite, prévisions météo, géolocalisation, sonars) facilitent l'exploitation intensive des océans, tandis que les exploitations minières et les déchets électroniques polluent pour de nombreuses années d'immenses zones aquatiques et terrestres. Les outils numériques et leur interconnexion mondiale servent également au système de production et de consommation de masse. Sans le numérique, toutes ces surexploitations seraient probablement moins faciles, moins rapides, et probablement moins destructrices.* »

On retrouve dans cette déclaration l'idée *« d'économie du remplacement »* , énoncée plus haut : avec l'émergence de technologies comme l'IA, la société ferait face à une intensification non nécessaire d'usages entraînant une exploitation toujours plus accrue des ressources naturelles. Soutenant une idée similaire, un rapport de 2023 de l'association CEMEA (CEMÉA Éducation Permanente, 2023) soutient qu'Internet s'imposerait aujourd'hui comme la réponse à nos besoins mais en créerait perpétuellement de nouveaux. Criteo, entreprise se présentant comme une agence publicitaire par IA pour des plateformes web, défend également cette position (Arcep, 2020). Dans l'industrie du numérique en général, il y a des limitations physiques du hardware et des verrous intentionnels mis sur le software, créant une obsolescence liée à l'interaction software/hardware[8]. Ainsi, l'industrie du numérique se serait construite sur la création d'un besoin par cette obsolescence, et c'est dans ce contexte que l'IA se développerait. Du côté des entreprises, on ne questionne pas cette fulgurante croissance. Comme en témoigne l'expert IA du Shift Project :

[7] L'acteur fait ici un parallèle avec une bulle spéculative où le prix d'un actif augmente de manière excessive, au-delà de sa valeur intrinsèque, sans réalité physique.
[8] Hardware désigne le matériel et les composants physiques constituant les ordinateurs. Software est un logiciel qui englobe à la fois le système d'exploitation et l'application informatique qui parcourt l'ordinateur

> « *On est au tout début de la bulle (spéculative, ndlr) avec les entreprises qui se disent : "je ne peux pas ne pas avoir cette année au moins 10 projets sur l'IA". Donc on recrute à prix d'or, où on forme en interne des ingénieurs là-dessus, et on dit : "je veux un projet IA que je montrerai sur mon site web, que je montre en interne, que je montre à mon conseil d'administration…" Et concrètement, on n'a pas encore grand chose à montrer* » (Entretien The Shift Project)

Et du côté des fournisseurs, c'est une capacité à optimiser, gagner de l'argent, de l'agilité et de la performance qui est vendue :

> « *remplacer tout un call center avec des milliers de collaborateurs par des robots qui feront la conversation téléphonique et qui parleront à d'autres robots à côté qui écouteront des conversations téléphoniques. Dans un monde en Orwell ce serait ça* » (Entretien The Shift Project)

La référence à l'auteur dystopique britannique témoigne de l'inquiétude que cette fièvre autour de l'IA procure dans les milieux associatifs. C'est en suivant les mêmes convictions qu'un membre du Ministère de la Transition écologique, chargé de mission IA affirme :

> *« il n'y a pas eu de débat démocratique dans la société pour vraiment se poser la question de : en fait qu'est-ce qu'on veut comme société ? On laisse pour l'instant complètement le marché décider en toute autonomie avec des subventions publiques »* (Entretien du MTE)

Cette crainte liée à une utilisation massive et un emballement du monde économique autour de l'IA est réfutée de manière claire par Facebook ou Google. Entreprises jouant un rôle clef dans le développement de ces techniques, elles affirment dans les groupes de travail de l'ARCEP (Arcep, 2020) qu'il n'y a *« pas de lien direct entre la limitation des volumes de données et le développement de services numériques plus durables »* ; ou que la consommation est un fait du consommateur lui-même. Ce dernier serait même incité à consommer de manière plus raisonnée. Par exemple, chaque utilisateur de la plateforme YouTube peut évaluer son temps de visionnage, choisir la qualité des vidéos qu'il regarde, désactiver la liste des vidéos additionnelles qui lui sont proposées… En outre, face à une innovation qui semble nous dépasser, se réfugier dans la tanière du connu est-il la meilleure des solutions ? Dans ce même rapport, l'entreprise suédoise de télécommunication Ericsson affirme :

> *« Il serait pour le moins étrange de refuser l'innovation technologique parce qu'il y aurait des incertitudes sur certaines conséquences indirectes possibles.* […] *L'évitement par l'immobilisme n'est pas plus une solution que la marche en avant en ignorant ces risques..»*

C'est dans cet état d'esprit que se placent les entreprises du numérique développant des projets d'IA. Les différentes interventions de ces dernières dans le colloque organisé par l'ARCEP (Arcep, 2020) laissent penser qu'elles tirent des leçons des siècles derniers, ayant vu nombre de révolutions technologiques modifier profondément le paysage économique.

Enfin, face au fort engouement suscité par les développements de l'IA, un usage raisonné des outils d'IA serait possible. Par exemple, deux techniques pour réduire les émissions

liées à l'entraînement d'un modèle d'intelligence artificielle ont été testées : retarder l'entraînement ou le suspendre en période de forte émission. Ces méthodes ouvrent la voie à des pratiques plus éco-responsables dans l'IA selon Matthew Hutson, un journaliste indépendant spécialisé en IA (Hutson, 2022). Face au triplement du trafic internet mondial depuis 2015, la consommation électrique des datacenters ne représenterait que 1% de l'électricité utilisée selon Facebook dans le rapport de l'ARCEP mentionné ci-dessus.

Des conséquences environnementales accrues

Alors que certains critiquent le besoin même et les usages de l'IA, les organes publics et internationaux mettent en parallèle des objectifs de développement durable et une utilisation grandissante de l'IA. Cette juxtaposition peut apparaître contradictoire pour certains acteurs, notamment les associations favorables à un numérique plus raisonné comme GreenIT ou le CEMÉA. Cette intensification de l'utilisation de l'IA mène en effet à une intensification de ses conséquences environnementales. Ainsi, l'ambiguïté de cette technologie peut être mise en évidence par la mise en perspective des impacts de l'IA sur l'environnement, face à ses usages et sa finalité. Jusqu'à quel point est-il possible de justifier une technologie malgré ses effets ? Quelles conséquences environnementales peuvent être considérées comme acceptables, et selon quels besoins et quelle utilisation ?

Les effets de l'IA sont de plus en plus décriés dans la presse comme le montre l'article du New York Times en octobre dernier qui rend compte de l'augmentation des émissions CO2 de Microsoft de 30% liées directement à l'installation de ses data centers (Gelles, 2024). Un article du Figaro de mars 2024 met en évidence ce que représente cette utilisation du numérique avec des exemples plus concrets :

> *« Des études ont montré que la création d'un large modèle de langage va émettre en moyenne 284 tonnes équivalent CO2. Concrètement, cela représente 315 vols aller-retour New York-San Francisco ou encore les émissions au cours de l'ensemble de leur existence de cinq voitures, depuis l'extraction des minerais nécessaires à leur fabrication jusqu'à leur mise à la casse.»* (Mollaret, 2024)

Ces éléments sont corroborés par toute une partie de la communauté scientifique qui utilise des chiffres similaires pour montrer l'impact CO2 de ces technologies (Dhar, 2020). Il semble aujourd'hui apparaître un consensus sur la part du numérique dans les émissions de gaz à effet de serre qui était *« en 2018 de 4% des émissions globales mondiales »*, supérieure aux émissions du secteur aérien (CEMÉA Éducation Permanente, 2023; Loup, 2022).

Au-delà des émissions de CO2, l'impact de l'IA touche aussi aux ressources planétaires par sa matérialisation avec une multiplication des datacenters. En effet, comme le souligne *L'enfer numérique* (Pitron, 2021), 12,5% de la production mondiale de cuivre et 7% de l'aluminium seraient destinés à la construction de datacenter. Le numérique engloutit aussi une large part de la production mondiale de petits métaux (70% du gallium, 80% du germanium, 23% de l'Ag...). Il est essentiel de prendre en compte tout le cycle de vie lié à

l'IA pour en comprendre l'étendue des effets sur l'environnement qui peuvent être de nature très différente. C'est ce que rappelle la *Commission de l'IA : Notre ambition pour la France* :

> *« L'empreinte carbone des modèles d'IA ne vient pas que de leur entraînement mais de l'ensemble de leur cycle de vie : la fabrication et le transport des équipements (20 à 30 % de l'empreinte) ; le développement et l'utilisation du modèle (50 à 60% : entraînement et inférence) et les déchets (10 à 20 %) »* (Commission de l'intelligence artificielle, 2024)

De plus, la consommation énergétique des installations numériques qui supportent l'IA croît de façon exponentielle au fil des années comme le montre le rapport de The Shift Project sur la *Sobriété Numérique* avec une augmentation *« d'environ 9% par an (période 2015 à 2020), une tendance largement supérieure au scénario moyen « expected case » prévu par Andrae et Edler (4%) et à peine inférieure à leur scénario catastrophe « worst case » (10%).»* (Entretien The Shift Project). Il est estimé que *« l'IA pourrait consommer 85 à 134 TWh d'électricité en 2027, soit une consommation équivalente à celle de l'Argentine ou de la Suède. »* (Commission de l'intelligence artificielle, 2024). Les datacenters qui sont en partie l'incarnation matérielle du traitement de données utilisée par l'IA *« représentent 1% de la demande mondiale d'énergie. L'énergie utilisée par Google est à 15% utilisée pour les datacenters »* (Trihinas et al., 2022)[9]. Pour réduire leur empreinte carbone, les entreprises utilisent des schéma de fléchage vers des productions d'électricité bas carbone. Un exemple donné par Fabrice Flipo dans *The Conversation* est le cas d'Amazon qui a installé un site avec des datacenters à côté d'une centrale nucléaire aux États-Unis pour atteindre ses objectifs d'émissions (Leparmentier, 2024). Il montre dans le même article que l'augmentation substantielle de la consommation d'électricité à cause des datacenters localement mène à des pénuries qui sont palliées par l'utilisation d'énergie fossile pour les habitants environnants. C'est pourquoi les élus s'opposent aux incitations du gouvernement pour promouvoir le développement des centres de données et de la big data utilisée en masse par l'IA (Flipo, 2022). En s'assurant de la consommation d'électricité bas carbone, les entreprises maintiennent au mieux leur impact environnemental mais ne prennent pas en compte les autres besoins en électricité qui s'y superposent naturellement et n'ont pas les moyens d'être fléchés. Que se passe-t-il lorsqu'une industrie doit s'électrifier pour les besoins de la transition écologique et que la puissance disponible est réservée pour le développement de futurs centres de données ? La communauté scientifique alerte sur les effets de ces derniers sur des ressources essentielles :

> *« les data centers mettant notamment en péril le bon fonctionnement des réseaux de distribution électrique ou produisant des pressions sur les ressources en eau potable, qu'ils utilisent pour refroidir leurs installations. »* (Marquet, 2024)

Le débat se cristallise majoritairement entre les grandes entreprises de la tech, les chercheurs et les associations mais touche peu l'individu en tant que citoyen. Les organisations internationales semblent aussi plus difficiles à mobiliser *« Si l'ONU y voit un risque en termes de respect de la vie privée et pointe le danger que représente le monopole*

[9] Citation Originale : "equivalent to 1% of the global energy demand [6]. Moreover, Google reports that approximately 15% of its energy use is attributed to AI/ML"

privé de ces données par les GAFAM, aucun mot n'est dit en termes de danger pour l'environnement. » (CEMÉA Éducation Permanente, 2023). Ceci semble majoritairement causé par *« le mensonge de l'immatériel »* (Pitron, 2021) qu'impose l'IA et le numérique plus généralement. Même en tenant compte des différents spectres associés à la définition même de l'IA, que ce soit par les experts ou le grand public, elle repose sur des systèmes numériques existants décuplés afin de stocker et traiter des quantités de données colossales et d'ainsi en délier des tendances, prédire. Alors que les associations et les États prennent conscience peu à peu de l'enjeu derrière la matérialisation du numérique et en particulier de l'IA, il est difficile pour le grand public de concevoir qu'une requête sur ChatGPT, pour ne citer que lui, consomme *« trente fois plus d'électricité que sur un moteur de recherche normal »* d'après un agent du Ministère de la Transition écologique (Entretien du MTE). En effet, c'est ce phénomène de dématérialisation de l'IA qui éloigne le grand public de ses conséquences environnementales présentes comme dans tout autre secteur. Le CEMÉA fait paraître également un article en décembre 2023 sur le numérique soutenable mettant l'accent sur l'importance de la sensibilisation du public par les associations au coût matériel réel du numérique auquel l'IA appartient en rappelant que : *« dématérialiser, c'est matérialiser autrement ». Nous pouvons même parler de multi matérialisation ou multiplication de la matière ! »* (CEMÉA Éducation Permanente, 2023). Si les conséquences environnementales de l'IA sont discutées dans l'arène publique c'est d'abord par cette difficulté de lier matériel et IA car *« c*[e n'est] *pas tangible »* (Entretien du MTE). Il y a une question même de la réalité des conséquences environnementales de cet outil. La finalité de l'IA trouve aussi sa place dans l'utilisation qu'en fait le grand public. Si le déni de l'immatériel venait à être percé, le grand public pourrait donner un nouveau cadre d'utilisation à l'objet, plus respectueux de l'environnement en s'affranchissant de la finalité présupposée par les développeurs de l'outil et son script (Pitron, 2021; Entretien avec GreenIT).

L'apport des outils numérique dans la lutte écologique

Ces externalités négatives souvent oubliées peuvent-elles être justifiées par un nombre croissant d'utilisateurs montrant un besoin irrésistible pour ce nouvel outil ? C'est le point que défend Éric Schmidt, PDG de Google, lorsqu'il suggère que l'essor de l'IA est trop important pour être freiné par des préoccupations climatiques (Gelles, 2024). Mais au-delà de la démonstration de la nécessité de l'IA, c'est l'argument de son utilisation au service de la lutte contre le changement climatique qui est souvent avancé par les entreprises développant ces outils. Les chercheurs et associations critiquent d'ailleurs cette position par laquelle le cadre de la controverse se trouve redéfini. Il ne s'agit plus d'opposer environnement et IA mais de les rassembler sous un même enjeu où l'IA pourrait agir comme solution technologique pour contribuer aux problématiques de la crise environnementale. Une nouvelle justification naît alors sur la capacité des usages de l'IA pour l'environnement à compenser ses externalités négatives comme le suggère Françoise Berthoud en parlant des prévisions des entreprises travaillant sur l'IA *« C'est David contre Goliath. L'industrie a confisqué le discours sur les bénéfices environnementaux d'Internet.*

Elle a essayé par tous les moyens de vendre cette idée et elle continue » (CEMÉA Éducation Permanente, 2023). L'IA devient une solution aux problèmes auxquels elle contribue et cela permet de légitimer sa place. Le New York Times relaye lui-même des déclarations de Nvidia, acteur californien dominant dans le matériel utilisé dans des projets d'IA, qui affirme contribuer à rendre les entreprises plus efficaces sur le plan énergétique grâce à l'IA, citant l'exemple de Foxconn[10] qui a réduit ses besoins énergétiques de 30 % dans une nouvelle usine conçue avec l'aide de l'IA (Gelles, 2024). Cette position peut aussi être identifiée notamment dans le rapport SMARTer 2030 : ICT solutions for 21st Century Challenges qui affiche en 2015 des conclusions optimistes et très positives sur les effets de l'IA comme le fait que *« les émissions évitées par les TIC sont 10x plus importantes que celles générées par le déploiement de ces technologies »* (Pitron, 2021). Il est rédigé par les entreprises membres du GeSI (Global e-Sustainability Initiative), i.e. AT&T, Dell, Huawei, IBM, Swisscom etc., et Accenture Strategy. Une opposition se crée donc entre les scientifiques et chercheurs qui reprochent le manque de consistance scientifique de ces propos et les entreprises telles que les GAFAM qui mettent en valeur leur contribution sur l'atténuation et l'adaptation au changement climatique. L'administration publique se fixe alors comme objectif de faire en sorte que ce bilan positif soit communément construit et non pas une vitrine pour les GAFAM : *« pour l'instant on n'a pas retiré grand-chose grâce à l'IA donc je pense que ce calcul il est pas possible de le faire pour l'instant. Mais nous ce qu'on essaie de prôner c'est qu'effectivement il y ait un impact enfin un bilan à impact positif »* ((Entretien du MTE). Pour autant, certains collectifs comme le Shift Project n'hésitent pas à affirmer selon Guillaume Pitron : *« La transition numérique telle qu'elle est actuellement mise en oeuvre participe au dérèglement climatique plus qu'elle n'aide à la prévenir" Lean ICT : pour une sobriété numérique »* (Pitron, 2021). Pourtant, certaines applications contribuent à mieux appréhender et quantifier les effets du changement climatique afin de l'atténuer et de s'y adapter.

Au-delà du domaine du numérique, l'IA pourrait ainsi jouer un rôle important dans la transition écologique et la gestion des risques climatiques. La DOE (Department Of Energy aux États-Unis) avance les capacités de l'IA à atténuer les émissions de méthane, un puissant gaz à effet de serre, par la détection et la prévision de fuites (Benes et al., s. d.). Dans le même sens, l'université de Parma montre l'utilité de l'IA dans l'optimisation de la consommation d'eau dans l'agriculture. L'agriculture utilisant 70% de l'eau mondiale, dans des zones parfois en stress hydrique, un tel modèle pourrait être disruptif (Preite & Vignali, 2024). De nombreux avantages du même type sont avancés par l'ICEF (Innovation for cool Earth Forum) dans son rapport de 2023 sur les apports de l'IA face à la crise climatique (Sandalow et al., 2023).

C'est dans un élan comparable que le Ministère de la Transition écologique a fait un appel à projet qui s'appelait *« démonstrateur d'IA frugale au service de la transition écologique dans les territoires »*. Cet appel à projet aurait fait émerger douze lauréats de projets très divers selon le ministère : utilisation de l'IA pour prédire et identifier des fuites d'eau dans le réseau d'eau potable, modèles pour analyser le cadastre solaire et faciliter le développement de

[10] Foxconn est un groupe industriel Taiwanais spécialisé dans la fabrication de produits électroniques.

projets photovoltaïques, un modèle d'IA de recherche et de tri de solution écologique sur la base des rapports du GIEC et de l'IPBES… Selon le Ministère de la Transition écologique, l'IGN (Institut géographique national) travaille sur des modèles d'IA, en particulier un jumeau numérique de la France pour des simulations et prévisions de catastrophes naturelles locales (inondation, incendies…). (Entretien du MTE)

> *« je pense qu'il y a des cas très spécifiques où ça peut permettre de réduire l'impact* [...] *tout ce qui est optimisation vraiment de haute intensité* [...] *ça commence à devenir intéressant si on le sérialise et qu'on le systématise. » (*Entretien du MTE)

C'est cette optique que défend un membre du ministère de la transition écologique, qui se trouve alignée avec les préconisations du DOE aux États-Unis. On assiste donc à une révolution des outils du numérique avec l'émergence de l'intelligence artificielle. Ces outils paraissent en capacité de se mettre au service de la crise environnementale. Dans certains cas, l'IA peut même contribuer à la réduction des émissions pour des établissements utilisant des supercalculateurs comme chez Météo France (Entretien du MTE). Néanmoins, la priorité devant cette révolution serait d'informer au mieux l'ensemble des acteurs sur l'impact environnemental de leurs achats d'après un rapport du Sénat (Maurey et al., 2020).

On remarque donc une forte ambivalence de l'IA au service de l'environnement. D'un côté, des entreprises du numérique affirment que ce type de technologie permet de rendre le monde plus vert : l'IA permet l'optimisation de la consommation d'énergie, le télétravail limite les déplacements… D'un autre côté, les associations et certains consommateurs pointent un manque de réflexion et de réalisme quant aux usages réels des technologies d'intelligence artificielle. L'ONU met elle-même en tension les objectifs de développement durable et l'usage massif de données avec l'explosion des technologies du numériques, dont l'IA. Bien que dans certains cas d'urgence climatique l'IA pourrait être utile, *« Il faut donc être critique vis-à-vis de la politique du big data comme solution possible à la crise écologique, tant elle semble contradictoire »* affirme certaines associations qui luttent pour un numérique durable (CEMÉA Éducation *Permanente, 2023).* Le discours des entreprises du numérique semblerait oublier une partie des effets réels : *« Il ne faut pas que ce soit l'arbre qui cache la forêt »* scande un Shifter (Entretien The Shift Project).

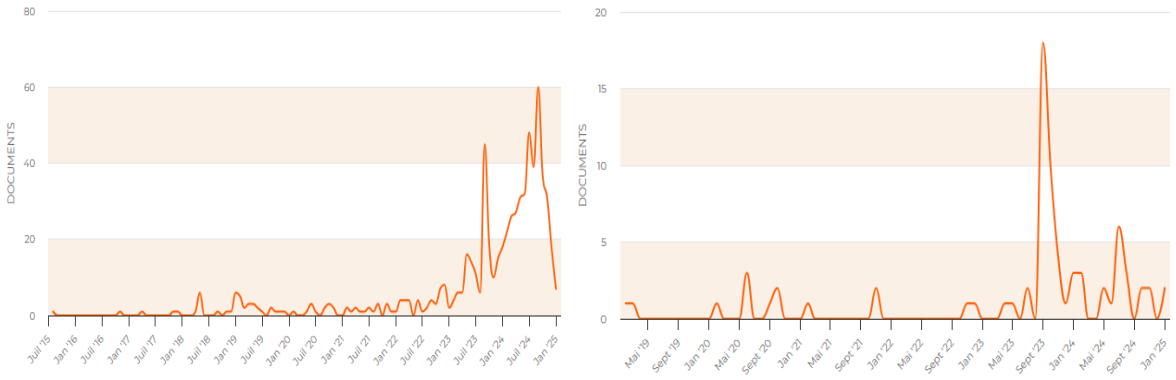
Entreprises développant des modèles d'IA	- Besoin croissant - Atout pour la lutte contre la crise écologique - Innovation nécessaire pour la compétitivité
Ministère de la Transition écologique	- Usage de l'IA pour des usages préservant l'environnement - Interrogation sur ses usages et ses besoin - Adaptation à cette innovation et garantir une compétitivité nationale
Associations pour une IA plus raisonnée	- Questionnement les usages et la finalité de l'IA - Dénonciation un mythe de l'immatériel - Remise en cause de la course à l'innovation face aux exigences environnementales
Individus, société civile	- Utilisation croissante de l'outil - Besoins émergeants - Méconnaissance des enjeux environnementaux de l'IA

Tableau explicatif et simplifié des enjeux portés par les différents acteurs

L'essentiel paraît donc résider dans l'évaluation des conséquences environnementales de l'intelligence artificielle. Entre consommation accrue d'énergie et de matériaux, explosion des usages et optimisation des besoins d'énergie, d'émission et prévision de phénomènes naturels, de nombreuses facettes de l'IA sont défendues par les acteurs. Mais est-il possible de mesurer objectivement les impacts environnementaux de l'intelligence artificielle ?

Partie 2: Quels obstacles s'opposent à une mesure claire de l'empreinte environnementale de l'IA ?

Afin d'étudier l'empreinte environnementale de l'intelligence artificielle, il est nécessaire de définir et construire des outils de mesure concrets, précis et robustes qui pourront servir de base à l'établissement de normes et législations. La conception d'outils pertinents et complets est une tâche complexe, exigeant une prise en compte de multiples dimensions. Cette difficulté émerge de la diversité des impacts environnementaux associés à l'IA et de la nécessité d'adopter un cadre d'analyse global, s'étendant au-delà de la simple consommation énergétique. Nous montrerons d'abord que l'étendue des paramètres à prendre en compte est difficile à définir, et que la généralisation d'une méthode de calcul est un point délicat également, avant de dire qu'un des obstacles actuels à l'obtention de mesures précises et complètes est le manque de données fiables en provenance des différents acteurs de la chaîne de production d'une intelligence artificielle : (industrie minière / fournisseurs de matières premières), constructeurs de matériel informatique, géants du numérique, pour enfin aborder l'effet rebond.



Répartition chronologique des articles de presse concernant l'IA et l'environnement (figure de gauche) et les articles mentionnant la mesure de l'impact de l'IA parmi ces articles (figure de droite) (Europresse)

Définition du spectre à prendre en compte : analyse de cycle de vie, matières premières, émissions carbone

Le déploiement d'un modèle d'IA nécessite l'intervention de nombreux acteurs de secteurs différents et l'on peut vouloir considérer l'empreinte environnementale, dans toute sa diversité, de chacune de ces étapes de production. Certaines méthodes déjà existantes prennent en compte l'entièreté de cette chaîne. Par exemple, le bilan carbone permet de suivre les émissions de gaz à effet de serre causées directement et indirectement par les activités d'une entreprise. L'analyse de cycle de vie (ACV) quant à elle permet d'évaluer les

impacts environnementaux d'un système sur l'ensemble de son cycle de vie. Ces indicateurs sont très informatifs et complets mais complexes à établir. Ils reposent sur d'autres outils de mesures plus spécialisés. Dans le cas de l'intelligence artificielle, les impacts environnementaux se concentrent principalement sur les phases de production matérielle, d'entraînement du modèle et d'utilisation du modèle.

La fabrication des équipements nécessaires au déploiement de l'IA (processeurs spécialisés, cartes graphiques, serveurs, centres de données) concentre une part majeure de la consommation en énergie et ressources à travers l'extraction et l'utilisation des matières premières et la construction des infrastructures. Par exemple, la fabrication des puces électroniques, malgré leur faible masse, mobilise une grande quantité de ressources : jusqu'à 32 kg de matières premières (eau, métaux rares) pour une puce de seulement 2 grammes (Ferreboeuf et al., 2020). Les taux de recyclage extrêmement faibles de certains matériaux comme le tantale ou le germanium (< 1 %) aggravent cette empreinte (Ferreboeuf et al., 2018). La construction des centres de données est responsable quant à elle d'une grande part des émissions de gaz à effet de serre attribuées au cycle de vie d'un modèle d'IA : plus de la moitié des émissions pour certains modèles selon une étude (Luccioni et al., 2024). La présence et l'activité de ces infrastructures peut également provoquer des perturbations locales sur la biodiversité, notamment par la création d'un îlot de chaleur.

Dans un second temps, les phases d'entraînement et d'inférence nécessitent aujourd'hui une importante consommation énergétique. L'entraînement de modèles de grande taille, comme les grands modèles de langage (LLM), est particulièrement gourmand en énergie électrique. Les moyens de production de cette électricité ont donc une influence majeure sur les émissions de gaz à effet de serre attribuables aux modèles d'IA. La phase d'inférence, c'est-à-dire la phase d'utilisation quotidienne des modèles une fois mis en service, peut également générer un coût significatif qui croît avec la durée de vie des modèles. D'autre part, les centres de données, qui hébergent la puissance de calcul nécessaire, nécessitent des systèmes de refroidissement intensifs et donc une quantité importante d'eau. L'utilisation de gaz fluorés pour refroidir ces infrastructures est également de plus en plus courante. Constituant une nouvelle technologie, leur utilisation est souvent mal régulée et leur impact n'est pas encore bien mesuré (Ferreboeuf et al., 2020). Ils constituent donc un exemple d'une source additionnelle d'empreinte environnementale à prendre en compte.
On voit ainsi qu'il est difficile de considérer l'ensemble des impacts environnementaux d'un système en plein évolution comme l'IA car ceux-ci se diversifient et changent rapidement et car les estimations sont longues complexes à réaliser.

Enfin, il convient d'intégrer également dans l'analyse l'impact lié au domaine d'application des modèles. Certains modèles sont ainsi censés optimiser un processus en permettant des gains de temps ou des analyses plus complètes et précises. Par exemple, le modèle GenCast de Google DeepMind permet de générer des prévisions météorologiques nécessitant une puissance de calcul et des temps de calculs bien inférieurs aux méthodes classiques qui requièrent des supercalculateurs. Pour de nombreux modèles, cet apport bénéfique peut être difficile à évaluer et doit être mis au regard d'un potentiel effet rebond. Il

existe aussi des utilisations très spécifiques au domaine de l'intelligence artificielle : les modèles dits *« de fondation »* servent de point de départ pour entraîner d'autres modèles. D'une part, ils contribuent à mutualiser l'énergie requise pour entraîner plusieurs modèles, optimisant ainsi les ressources nécessaires. D'autre part, ils ont également facilité la création de nombreux modèles gourmands en énergie, qui n'auraient pas vu le jour sans leur existence. Par conséquent, dans certains cas, intégrer le domaine d'utilisation d'un modèle dans l'évaluation de son empreinte environnementale peut présenter une certaine ambiguïté.

Outils de mesure réplicables

L'intérêt principal est de pouvoir mesurer le coût de l'usage d'un modèle d'IA pour l'environnement, et de pouvoir comparer cette mesure à celle du gain éventuel engendré par cet usage, mais aussi à d'autres modèles d'IA. Cela nécessite une méthode de calcul réplicable, c'est à dire qu'elle soit applicable à n'importe quel usage de l'IA, peu importe le modèle ou l'ordinateur utilisé. Ces indicateurs peuvent être un outil à la décision.

Il faut d'abord pouvoir quantifier la consommation directe d'un modèle l'IA, selon les différentes variables retenues comme étant d'intérêt. Le choix de l'outil de mesure est un premier défi : ceux existant varient dans leur facilité d'utilisation et dans les résultats obtenus.

L'entreprise Hugging Face, par exemple, a fait une analyse de cycle de vie d'un modèle d'IA appelé BLOOM, en prenant en compte l'énergie nécessaire à la fabrication du hardware, à l'entraînement, et la phase d'inférence. Pour le calcul de l'énergie pendant la phase d'inférence, ils ont utilisé le logiciel CodeCarbon (Kirkpatrick, 2023). Il s'agit d'un code prêt à l'emploi, qui est facile d'utilisation mais pour lequel les paramètres d'intensité carbone de l'énergie sont fixés. Du côté du gouvernement français, un outil appelé Green Algorithms, développé par un chercheur de l'Université de Cambridge, au Royaume-Uni, est utilisé pour estimer cet impact. Par exemple, lors d'un appel à projets d'IA pour la transition lancé par l'Ecolab, les porteurs de projets ont dû

> *« Calculer l'impact énergétique de leurs algorithmes grâce à un outil qui s'appelle Green Algorithms. Et c'est un outil qui est très cool mais qui calcule que l'énergie consommée, et donc pas l'eau, pas les matériaux. »* (Entretien du MTE)

Cet outil estime l'impact de l'entraînement d'un modèle d'IA grâce à la durée d'entraînement et aux caractéristiques du matériel (hardware) utilisé.

Il est déjà difficile d'estimer l'impact environnemental d'une IA lors de sa phase d'entraînement, mais une autre difficulté est l'estimation de son impact liée à la phase d'inférence. Une étude a été menée sur différents modèles d'IA, et conclut que bien que la phase d'entraînement soit plus coûteuse en énergie que celle d'inférence, de l'ordre de 100 000 000 fois plus, cet ordre de grandeur d'inférences peut être atteint rapidement pour certains modèles populaires (pensons par exemple à GPT-3, un des agents

conversationnels de l'entreprise OpenAI). Les modèles dits *« de fondation »*, versatiles et pouvant être utilisés pour de multiples applications sans avoir besoin de les réentraîner en entier, sont parfois considérés comme plus écologiques car, quoique très lourds à entraîner, évitent de recréer un modèle pour chaque usage. Cependant, cette même étude souligne le fait qu'une inférence avec un modèle se basant sur un modèle de fondation est plus énergivore qu'une inférence avec un modèle plus léger construit uniquement en vue d'une application, ce qui nuance cette perception, au vu du nombre d'inférences parfois exponentiel (Luccioni et al., 2024). Certains acteurs insistent aussi sur la nécessité d'une uniformisation de ces méthodes de mesure. Pour légiférer, notamment, une chargée de mission au Ministère de la Transition Écologique le dit :

> *« L'UNESCO a fait aussi un genre de référentiel pour mesurer l'empreinte environnementale de l'IA. Il y a genre l'institut international des télécommunications, des choses comme ça, qui, chacun a fait son propre référentiel. Et on a aussi engagé des travaux pour essayer d'arriver à une feuille de route commune quoi, enfin un référentiel commun. Parce que ça sert un peu à rien que chacun fasse le sien. »* (Entretien du MTE)

A cet effet, le GIEC, depuis 2014, s'efforce de fixer des valeurs pour les intensités carbone des différentes sources d'énergie par pays, ce qui est nécessaire au calcul de l'empreinte carbone de tâches numériques. Cependant, cette intensité varie même au sein d'une journée, en fonction des changements du mix énergétique en temps réel, ce qui rend difficile les estimations précises.

A Chypre, par exemple, un même modèle entraîné vers midi, heure de pic de la production d'énergie solaire et éolienne, avait une empreinte carbone inférieure de 1.64kg au même modèle entraîné à 18h et de 2.3kg de moins que le même modèle entraîné à 21h. L'objet utilisé pour les inférences, aussi, dans le cas de l'Internet des Objets, peut faire varier du simple au double l'énergie nécessaire à ces inférences. Il faudrait aussi évaluer la performance énergétique d'un modèle, plutôt que simplement sa précision (parfois une architecture peut être considérablement moins énergivore tout en ne baissant que peu en précision par rapport au *« meilleur »* modèle, conservant des performances techniques acceptables et devenant plus frugales) (Trihinas et al., 2022).

Ainsi, il est compliqué de construire des outils de mesure de cette consommation en énergie et en émission carbone. Les modélisateurs s'opposent selon le degré de précision qu'ils décident d'adopter. Ces outils devraient en effet effectuer une comparaison approfondie de ces paramètres en temps réel en fonction de la composition physique du mix énergétique, évaluer la puissance consommée pour les calculs dédiés à l'IA spécifiquement… cela demande de naviguer entre plusieurs interfaces et de nombreuses couches de complexité à l'outil de mesure.

Le problème de la transparence et de la disponibilité des données

Les méthodes et outils développés dans le but d'évaluer quantitativement les impacts environnementaux des modèles d'IA présupposent la disponibilité de certaines données concernant ces modèles et les infrastructures qui les supportent. Aujourd'hui, les modèles les plus gros et les plus utilisés tels que les modèles génératifs (texte, image) sont possédés par entreprises privées. Pour la plupart de ces modèles, leur architecture, c'est-à-dire leur nombre de paramètres et le détail de leur fonctionnement, est gardée secrète. De plus, les calculs nécessaires à l'inférence sont exécutés au sein des data-centers, dont il n'existe bien souvent que peu de données portant sur leur consommation énergétique et leurs émissions de gaz à effet de serre. En outre, les constructeurs de matériel informatique communiquent également peu sur les émissions liées à la fabrication de leurs puces (quelques chiffres évoqués précédemment existent quand même). Ce manque de transparence des géants du numérique rend difficiles les estimations d'impacts et constitue un frein à l'élaboration de nouveaux outils de mesures. Ainsi, de nombreux chercheurs le déplorent et appellent à une plus grande collaboration de la part des constructeurs de puces et des concepteurs de modèles. Par exemple, des auteurs (Luccioni et al., 2024)[11] espèrent :

> *« [...] que les chercheurs et les praticiens de l'apprentissage automatique pratiqueront la transparence en ce qui concerne la nature et les impacts de leurs modèles, afin de permettre une meilleure compréhension de leurs impacts sur l'environnement. »*

Pour Jesse Dodge, chercheur à l'Institut Allen pour l'IA, le plus grand obstacle à la compréhension de l'empreinte carbone de l'IA est le manque de transparence des entreprises à but lucratif de ce domaine (Kirkpatrick, 2023).

De manière plus générale, ce manque de transparence et de données concerne le secteur du numérique dans sa globalité, comme cela est souligné dans le rapport (Ferreboeuf et al., 2018):

> *« Il n'existe pas de chiffres globaux issus de mesures (même à l'échelle d'un pays) de la consommation d'énergie induite par les usages du Numérique. Les estimations disponibles sont obtenues soit par projection de mesures réalisées sur des échantillons (par exemple, un ensemble de data centers), soit par utilisation de modèles plus ou moins sophistiqués, mais détaillant rarement l'ensemble des hypothèses adoptées. ».*

Afin de renverser cette tendance à l'opacité, il est nécessaire de normaliser au sein des entreprises du numérique le partage systématique de certaines données permettant l'estimation des impacts environnementaux. Une mise en place de nouvelles régulations pourrait contraindre les différents acteurs à communiquer. Néanmoins, pour certaines informations comme l'architecture des modèles, le secret industriel prévaut. Enfin, des initiatives, comme le Référentiel Environnemental du Numérique du Shift Project évoqué

[11] Citation originale : *"that both ML researchers and practitioners will practice transparency regarding the nature and impacts of their models, to enable better understanding of their environmental impacts."*

dans (Ferreboeuf et al., 2018), ont également été prises afin de rendre facilement accessibles les données concernant l'empreinte environnementale du matériel informatique.

Comptabiliser l'effet rebond

Nous l'avons vu, l'impact direct de l'IA, lié à la construction du matériel de calcul et à l'entraînement puis l'inférence des modèles, est déjà difficile à estimer. Cependant, lorsqu'on parle de l'impact d'une technologie, un autre aspect à prendre en compte est son impact indirect, ou effet rebond.

L'effet rebond, ou paradoxe de Jevons, est l'accroissement de la consommation provoquée par l'augmentation de l'efficacité avec laquelle une ressource est employée (d'où le paradoxe, croyant améliorer l'efficacité on accroît la consommation de la ressource…) qui étaient jusque-là posées à l'usage d'un bien, d'un service ou d'une technologie. En termes d'IA, cet effet rebond peut être considérable, parce que c'est une technologie qui permet d'optimiser de nombreux procédés, et le risque est que les économies réalisées soient réinvesties pour faire, avec les mêmes ressources, plus d'opérations, plutôt que de profiter de cette optimisation pour diminuer la consommation de ressources.

Ainsi, comme le dit (Fropier et al., 2024),

> *« l'évaluation des effets de 1er et 2nd ordre est nécessaire mais pas suffisante. Des hypothèses et des mesures de contrôle doivent être identifiées et parfaitement documentées pour évaluer et garder les effets d'ordres supérieurs, notamment les effets rebond, sous contrôle ».*

(Ferreboeuf et al., 2018) ajoute que

> *« [les] gains positifs liés aux progrès sur l'efficacité (en temps, ressources énergétiques, etc.), [...] sont en très grande majorité surestimés du fait de la non-prise en compte des "effets rebond", qui conduisent à augmenter la consommation des ressources dont on a augmenté l'efficacité ou la facilité d'usage ».*

Un exemple de cet effet rebond est donné en exemple avec le cloud gaming (Arcep, 2019).Cette, pratique consistant à mutualiser les « ressources informatiques dédiées aux jeux vidéo », allonge la durée de vie des terminaux et réduit le taux d'équipement nécessaire. Mais bien que la quantité de ressource nécessaire à usage constant diminue grâce au cloud computing, elle rend également accessible à de nombreuses personnes des jeux vidéos auxquels ils n'auraient pas joué sans le cloud, ce qui peut conduire à *« une augmentation d'usage susceptible de contrebalancer les éventuels gains énergétiques qu'il aurait permis d'exploiter »*.

En effet, historiquement, les nouvelles technologies destinées à diminuer l'empreinte d'une activité humaine sur l'environnement ont plutôt eu l'effet inverse, comme le soulignent certaines associations (Ferreboeuf et al., 2018) :

> *« Autant le Numérique pourrait servir à repenser les usages et modifier les modes de fonctionnement et de production pour aboutir à une réduction des émissions de GES, voire à une réduction de la consommation énergétique (réseaux électriques intelligents, mobilité et transports intelligents, monitoring environnemental et urbain, dématérialisation, télétravail et visioconférence, bâtiments intelligents et écoconception logicielle), autant ceci n'aura une chance de se produire qu'à la condition de maîtriser les « effets rebond » dans ces secteurs. Or, jusqu'ici, les effets rebond se sont montrés plus importants que les gains apportés par l'innovation technologique* [...]. *Par exemple, les bénéfices environnementaux du télétravail sont largement inférieurs à ceux intuitivement escomptés, en tout cas lorsqu'il n'est pas combiné à d'autres changements de l'écosystème social. Il est donc illusoire de considérer a priori que le bilan énergétique global d'une Transition Numérique* [...] *soit voué à être largement vertueux, voire même simplement équilibré. »*

La mention de l'effet rebond revient également dans les entretiens avec les acteurs, notamment associatifs (Entretien The Shift Project) : *« Est-ce que vous pouvez trouver des innovations qui n'ont pas fait l'objet d'un effet rebond ? »*.

Dans plusieurs rapports, les acteurs mettent en avant le risque d'une numérisation croissante et non régulée. Par exemple, l'effet rebond est parfois mentionné sans être nommé (Demailly et al., 2017) :

> *« Le numérique est aujourd'hui le support technique privilégié de nos modes de vie, que nous savons trop dispendieux pour l'équilibre de notre planète. Et, dans notre système actuel, le risque demeure que les gains d'efficience apportés par le numérique soient réinvestis dans toujours plus de consommation et de croissance matérielle, un phénomène que nous constatons depuis des décennies. »*

Plus pernicieux encore, les outils de mesure de l'impact de l'IA ont également un coût environnemental, et un coût qui peut être assez important (Marquet, 2024) :

> *« En étudiant comment des logiciels de calcul parfois très lourds sont employés pour rendre compte de l'empreinte environnementale du big data, nous observons une récursivité qui n'est peut-être pas si anodine ».*

Ainsi, comme nous avons pu le voir, mesurer l'empreinte environnementale de l'intelligence artificielle est une tâche complexe, entravée par des obstacles multidimensionnels. La diversité des impacts à considérer, la difficulté à établir des outils de mesure universels, le manque de transparence des acteurs industriels et l'effet rebond limitent la fiabilité des estimations actuelles. Ces défis soulignent l'urgence de normaliser la mesure et de promouvoir une plus grande collaboration pour la collecte et la communication des données.

Une approche rigoureuse et systématique est essentielle évaluer précisément l'empreinte environnementale de l'IA et donc pouvoir orienter son développement vers une compatibilité avec les enjeux actuels.

Partie 3 : Compatibilité entre sobriété et compétitivité pour l'IA

La compatibilité entre sobriété et compétitivité dans le domaine de l'IA soulève des questions complexes et multidimensionnelles. Cette réflexion s'articulera autour de plusieurs axes principaux : la contribution de l'IA à une culture d'innovation et de croissance économique, les défis environnementaux associés à son développement rapide, et enfin, les solutions potentielles pour concilier ces enjeux, allant des initiatives de sobriété volontaire à l'intervention des pouvoirs publics.

Culture de l'innovation et de la croissance auquel le développement de l'IA contribue

L'IA a connu, au cours des dix dernières années, un développement fulgurant. D'abord, dans la recherche: les publications scientifiques se sont multipliées pour proposer des modèles de plus en plus performants et aux capacités de plus en plus impressionnantes. Progressivement, les avancées de l'IA ont permis une démocratisation de cet outil. L'utilisation de l'IA est désormais banalisée dans la plupart des entreprises, avec une part importante de marché pour les entreprises qui développent les modèles utilisés, telles que Meta, OpenAI ou encore Google. Enfin, le grand public a maintenant accès à la majorité des technologies développées dans ce domaine.

Pour les entreprises, les raisons de cette démocratisation de l'IA sont principalement économiques. Pour celles qui développent de nouveaux modèles, le revenu dépend directement du nombre de clients qu'elles réussissent à séduire avec leurs produits. Pour des raisons de compétitivité, il est donc naturel que ces entreprises entretiennent une forte culture de l'innovation et de la croissance. Pour la majorité des entreprises restantes, qui mobilisent l'IA dans leurs processus de production, l'enjeu économique est également important. En effet, les mêmes raisons de compétitivité s'appliquent, puisqu'une entreprise ayant amélioré ses processus par l'IA sera plus performante que la concurrence. De plus, l'IA est de plus en plus perçue comme un moyen de remplacer le travail de certains employés, et donc maintenir, voire améliorer, la production de l'entreprise à moindre coût économique (Entretien avec GreenIT). Ces raisons économiques justifient les investissements massifs qui sont faits et qui alimentent la dynamique de croissance du domaine (Entretien du MTE).

Les investissements du secteur peuvent conduire les développeurs et les utilisateurs de l'IA à ne pas se soucier des conséquences environnementales associées à cette croissance. Tandis que *« l'ONU met en tension les objectifs de développement durable et l'utilisation des Big Data »* (CEMÉA Éducation Permanente, 2023), les médias font de plus en plus le constat d'une croissance incompatible avec les objectifs de neutralité carbone (Séramour,

2024). Les conséquences environnementales proviennent d'abord du développement de modèles plus volumineux et coûteux, particulièrement en ressources nécessaires au fonctionnement des datacenters. On peut aussi citer les entreprises développant l'IA, qui semblent privilégier la performance des modèles plutôt que leur adéquation avec le développement durable (Kirkpatrick, 2023), ce qui conduit les chercheurs à privilégier la précision de leurs algorithmes au cours du développement (Dhar, 2020). La difficulté à quantifier les émissions induites par le développement de modèle contribue à accentuer cette dynamique.

Récemment, cette croissance à tout prix de l'IA est cependant remise en cause. Perçue par certains organismes tels que le Shift Project comme *« une bulle d'investissement, qui peut [...] être problématique »* (Entretien The Shift Project), la soudaine croissance économique de l'IA questionne. Par conséquent, la question de l'utilité de l'IA s'est récemment installée dans le débat sur sa croissance fulgurante. On retrouve au premier plan l'usage qui est fait de l'IA, qui peut aujourd'hui sembler excessif : dans les entreprises, certaines tâches sont désormais déléguées à l'IA, qui est parfois plus polluante que les anciennes techniques utilisées. De même, l'usage de services comme ChatGPT plutôt qu'un moteur de recherche classique pour une requête sur Internet interroge quant à sa pertinence comme évoqué en première partie (Mollaret, 2024). Enfin, l'utilité de l'IA dans la lutte contre le réchauffement climatique devrait être justifiée, de sorte que son utilisation ne soit pas plus néfaste que les effets bénéfiques qu'elle apporte pour contribuer à la résolution de cette crise (Entretien du MTE). Plus généralement, un rapport des Ceméa soutient que *« la plupart du temps, le progrès technique est automatiquement vu comme bénéfique »* et que le *« solutionnisme technique privilégie donc l'usage des technologies, plutôt que leur critique »* (CEMÉA Éducation Permanente, 2023). Le questionnement sur l'intelligence artificielle semble donc évident pour certains chercheurs et organismes publics ou privés qui s'intéressent à son impact environnemental. La notion d'une utilisation plus frugale de l'IA, guidée par une volonté de sobriété, apparaît tout de même comme un point central du débat autour de sa croissance économique. Les conséquences environnementales de l'IA, tout comme les outils qu'elle apporte pour lutter contre le réchauffement climatique et la croissance économique qu'elle permet à de nombreuses entreprises, semblent donc devoir être équilibrées.

Accepter une perte de rapidité et de précision pour la sobriété de l'IA

Les récents débats autour de l'impact environnemental de l'IA semblent donc souligner la nécessité de décroissance de cette technologie. Un rapport des Ceméa affirme que :

> *« le lien entre l'injonction individuelle à la décroissance et le progrès technologique est incohérent. Au sein de nos sociétés de consommation, le désir de changement se heurte bien souvent à la volonté de faire partie du groupe, de ne pas être à la marge, au désir créé pour l'acte de consommation »* (CEMÉA Éducation Permanente, 2023)

Il apparaît donc que la décroissance ne sera pas immédiate, notamment pour les citoyens qui pourront avoir des difficultés à se priver de privilèges dont les autres continuent à jouir. Pour les entreprises également, la décroissance pourrait-être difficile à accepter : comment accepter de réduire l'utilisation d'un outil qui leur permet d'augmenter leur productivité ? Enfin, on peut également lire dans la presse que *« stimuler la croissance et, « en même temps », appeler à la sobriété a quelque chose de kafkaïen, qui se résout par une sobriété de second rang, que l'on appelle d'ordinaire l'efficacité »* (Flipo, 2022). L'implémentation de cette efficacité comme forme de décroissance posant donc des problèmes, on observe dernièrement l'étude de plusieurs possibilités visant à inciter à la décroissance.

On observe dans ce nœud de discussion que la décroissance est au centre du débat pour permettre une réduction des impacts environnementaux. L'accent est mis sur la nécessité de sobriété en questionnant régulièrement la pertinence des usages de l'IA, ou en permettant aux développeurs de l'IA de réaliser l'impact de leurs modèles et de les inciter à proposer des modèles plus sobres. Le principe premier de la frugalité, c'est de *« questionner le besoin en IA. [...] il faut que l'IA soit réellement utile, réponde à un réel besoin »* (Entretien du MTE). Dans le monde de la recherche, il a également été suggéré d'introduire une nouvelle métrique (floating-point operations) et de demander aux chercheurs d'accompagner leurs présentations de modèles de cette métrique (Tamburrini, 2022). Une telle mesure permettrait de faire prendre en compte l'impact environnemental des modèles lors de leur évaluation par les pairs, et de potentiellement inciter à proposer des modèles plus sobres. Plus généralement, l'intervention d'un gouvernement peut être une solution pour inciter à la sobriété. On peut citer le *« plan de sobriété énergétique »* mis en place récemment par Elisabeth Borne (Flipo, 2022) qui présente une feuille de route avec de nombreuses mesures et informations permettant aux citoyens et entreprises d'évaluer les moyens à leur disposition pour prendre en compte la sobriété dans leurs opérations.

Au-delà des incitations et contraintes gouvernementales, on observe les entreprises se fixer des objectifs de neutralité carbone. Parmi les acteurs majeurs du numérique, Microsoft prend en 2020 un engagement pour atteindre la neutralité carbone d'ici 2030 (Leparmentier, 2024). De tels engagements avaient déjà été mis en place avant l'ère de l'IA, mais le développement fulgurant de celle-ci semble perturber les objectifs et risque de contraindre les entreprises à revoir leurs objectifs, ou agir sur leur usage de l'IA. On peut noter que le

président de Microsoft reconnaît que l'objectif de 2030 de neutralité carbone de l'entreprise est désormais cinq fois plus éloigné qu'en 2020, principalement à cause des besoins en électricité de l'IA (Séramour, 2024). Google constate également un bond de 48% de ses émissions de gaz à effet de serre dus uniquement au développement de l'IA, ce qui remet en cause les objectifs de neutralité carbone de la multinationale. Dans le même sens, l'Agence Internationale de l'Énergie prévoit que la consommation d'électricité pourrait doubler entre 2023 et 2026 du fait de l'utilisation de l'IA, des data centers et de la cryptomonnaie (Entretien du MTE). De telles prévisions, accompagnées des constats récents faits par les grandes entreprises développant l'IA, semble pousser ces entreprises à revoir leurs objectifs pour les rendre plus réalistes, ou à changer leur utilisation de l'IA pour corriger leur trajectoire, par exemple en mobilisant l'IA pour réduire les émissions, ou plus radicalement en changeant la manière d'utiliser l'IA.

Au-delà des impacts positifs que l'IA peut avoir pour la lutte contre le changement climatique, le choix des modèles adaptés et le questionnement de la pertinence de l'usage de l'IA apparaissent comme une solution adaptée à la situation actuelle, où les entreprises voient leurs objectifs de neutralité carbone remis en question par la croissance de l'IA et la compétitivité renforcée qu'elle a introduit. Aujourd'hui, les incitations et encouragements de la communauté à développer des modèles plus respectueux de l'environnement se multiplient, notamment avec certains chercheurs qui suggèrent de ne pas développer de nouveaux modèles à la performance légèrement supérieure quand des modèles existants plus simples peuvent convenir (Six, 2024). Le projet *« Ecology by design »* vise également à *« mieux intégrer l'enjeu écologique à la fois dans la conception de ses services et dans sa contribution à la société »* (Demailly et al., 2017). Enfin, on observe notamment au Danemark un partenariat entre le Copenhagen Centre on Energy et les universités et laboratoires de recherche du pays pour implémenter de bonnes pratiques et une législation, qui favorisent notamment l'utilisation de modèles simples par rapport à leurs alternatives en apprentissage profond (Dhar, 2020). Commencer par réformer les pratiques d'implémentation et d'utilisation auprès des chercheurs et étudiants qui seront amenés à définir l'usage de l'IA à l'avenir apparaît comme un moyen efficace de démocratiser la sobriété dans ce domaine encore nouveau.

■ Interventionnisme et régulation des États

La puissance publique commence à s'imprégner de la question : la chargée de mission au ministère de la transition écologique parle de deux objectifs principaux *« l'IA au service de la transition écologique, et aussi l'IA plus frugale »* , c'est à dire *« une IA qui se veut le moins, la moins consommatrice de ressources possible »* . Des travaux sont lancés sur ces thèmes en collaboration entre plusieurs agences publiques telles que l'ADEME (agence de l'environnement et de la maîtrise de l'énergie), l'ARCEP (autorité de régulation des communications électronique, des postes et de la distribution de la presse) ou encore le CEREMA (centre d'étude et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement) (Entretien du MTE). De manière analogue, aux États-Unis, le DOE (Department of Energy) joue un rôle dans l'élaboration de politiques et de lignes directrices

pour un déploiement responsable de l'IA, en collaboration avec les organismes de réglementation et l'industrie, comme ils l'indiquent dans leur rapport (Benes et al., s. d.) AI for Energy - Opportunities for a Modern Grid and Clean Energy Economy :

> *« Le DOE* [...] *a organisé une série d'ateliers avec des parties prenantes externes représentant le monde universitaire, l'industrie et les organisations non gouvernementales afin d'approfondir les cas d'utilisation de l'IA dans le secteur de l'énergie et d'identifier les tendances et les préoccupations actuelles »*[12]

En France, l'AFNUM (Alliance française des industries du numérique) demande la création d'un observatoire de l'impact environnemental du numérique afin d'établir un état des lieux précis de la situation. Dans un second temps, le but est de responsabiliser tous les acteurs de la chaîne de valeur et d'ajouter un cadre législatif (plutôt au niveau européen) réaliste par rapport à la mondialisation (Arcep, 2020).

Un article du CEMEA établit l'importance du lien entre éducation et politique et écrit dans (CEMÉA Éducation Permanente, 2023) :

> *« Historiquement, le Développement Durable comprenait quatre piliers : l'écologie, l'économie, le social et la gouvernance. Avec le temps, le dernier pilier a eu tendance à disparaître. Le lien entre le fait de pouvoir décider pour soi, à être auteur-autrice plutôt que simple consommateur-consommatrice nous semble pourtant essentiel. Le lien entre l'éducation et la politique est à rétablir. Pour un numérique choisi demain, il faut des citoyen-ne-s capables d'en comprendre les enjeux aujourd'hui, capables de résister et de questionner »*

Cette citation fait écho aux mots de la porte-parole de GreenIT qui évoque les questionnements démocratiques qui devraient être considérés avant que l'innovation et le marché dictent nos manières de vivre et de consommer (Entretien avec GreenIT). Dans un article du Figaro publié en 2024, on peut lire «On émet soixante fois plus de carbone en opérant une requête sur ChatGPT que sur Google» écrit *« Il convient déjà de se poser la question de savoir si nous avons réellement besoin d'une IA générative pour réaliser le problème qu'on lui soumet. »* Il remet à son tour en question le réel besoin de l'IA générative qui fait partie des plus consommatrices en ressource (Mollaret, 2024).

De nombreuses régulations sont tout de même mises en place afin d'endiguer un développement trop important de l'IA et de l'empreinte associée. En Chine, les data-centers possédés par des organismes publics doivent améliorer leur efficacité énergétique et devront, d'ici 2032, être entièrement alimentés par des énergies renouvelables (5% au moins en 2023). Aux États-Unis, depuis 2020, une loi impose de mener des études sur la consommation des data-centers en énergie et en eau (pour pouvoir par la suite faire des lois). Certains États comptant de nombreux data-centers ont déjà imposé des pratiques plus durables afin de réduire l'impact des datacenters. Du côté de l'UE, le Climate Neutral Data Center Pact (2020), fixe des objectifs pour la neutralité carbone en 2030. Il a été signé par

[12] *Citation originale : "DOE [...] held a series of workshops with external stakeholders representing academia, industry, and non-governmental organizations to further develop AI use cases for the energy sector and identify current trends and concerns"*

plus de soixante opérateurs, principalement des grandes entreprises et des acteurs majeurs de l'industrie des data-centers (Séramour, 2024b). Plus particulièrement, la France met en place un ensemble d'acteurs et de régulations pour tenter d'encadrer le développement de l'IA.

> *« La loi REEN répond donc en favorisant l'allongement de la durée de vie des terminaux, les usages écologiquement vertueux et notamment l'écoconception des services numériques ; en réduisant la consommation d'énergie des centres de données et exigeant des collectivités territoriales de mettre en place une "stratégie numérique responsable". »*

Cette loi est accompagnée d'une feuille de route « Numérique et environnement » ainsi que de la loi AGEC. En parallèle, *« le gouvernement [...] a inauguré un Haut Comité pour un numérique éco responsable »* composé de 5 groupes de travail et ayant pour but de suivre et proposer des axes d'amélioration de l'empreinte environnementale du numérique (Flipo, 2022). Néanmoins, la chargée de mission du Ministère de la Transition écologique rappelle que la plupart des mesures se veulent non contraignantes dans un premier temps (Entretien du MTE). Pourtant, malgré les efforts des organismes de régulation, le développement de l'IA se fait à un rythme très rapide et les réglementations ont du mal à suivre. C'est ce qu'on peut lire dans (Marquet, 2024).

> *« Les conflits liés aux data centers et au manque de réglementation et d'anticipation politique à leur sujet se traduisent notamment par l'existence de deux moratoires, à Dublin et à Amsterdam, suspendant tout permis de construire pour ces installations. »*

Quelques incohérences restent néanmoins régulièrement pointées du doigt. *« L'ONG Réseau Action Climat notait le caractère contradictoire des mesures gouvernementales, en termes de résultat à atteindre »* comme par exemple *« Stimuler la croissance »* et, *« en même temps*, appeler à la sobriété *»* (Flipo, 2022). Certains opérateurs comme Bouygues font le constat suivant : *« En dépit de leurs efforts (...) les opérateurs ne pourront à la fois réduire drastiquement leurs émissions de gaz à effet de serre et, dans le même temps, augmenter la capacité de leurs réseaux pour répondre à la demande toujours plus forte de leurs clients, provoquée en partie par l'adoption par la société de comportements plus sobres (télétravail par exemple). »* Selon eux, la solution réside en un effort collectif et équitablement supporté par l'ensemble des acteurs de la filière. Ils parlent aussi de la nécessité d'une analyse précise des impacts du secteur en comptant aussi les externalités positives, thème déjà abordé précédemment (Arcep, 2020).

■ Approche à plusieurs échelles

Certains acteurs soulignent l'importance de traiter la question à plusieurs échelles afin de dégager des solutions pérennes. C'est notamment le point de vue proposé par le rapport Référentiel général pour l'IA frugale de l'AFNOR qui évoque la nécessité d'une approche multidimensionnelle impliquant tous les acteurs du cycle de vie de l'IA, des développeurs

aux utilisateurs finaux, en passant par les décideurs politiques (Fropier et al., 2024). C'est aussi ce que la chargée de mission IA au Ministère de la Transition écologique soutient :

> *« On veut que ce soit contraignant, mais on veut pas que ce soit juste à l'échelle de la France donc on se dit que le cadre propice à ça c'est l'échelle européenne, donc on est en train déjà de travailler à un document, pareil que pour nous en France, mais au niveau européen, avec le comité européen de normalisation. Et à partir de là, plus tard, on aimerait pousser pour que ça devienne contraignant. »*

À cela, un article sur l'empreinte écologique de l'IA au service du médical ajoute l'importance de l'action collective et des décisions à grande échelle pour un impact significatif sur la durabilité. *« Ces actions représentent des changements à un petit niveau, et souvent à un niveau individuel. L'ampleur des émissions de gaz à effet de serre, l'urgence d'atténuer le changement climatique et la menace existentielle qui pèse sur notre planète exigent une action collective »* (Truhn et al., 2023)[13].

Un article étudiant la conciliation entre transition numérique et écologique évoque deux solutions complémentaires en matière de réglementation. La première serait d'augmenter le niveau d'exigence des régulations *« en abordant la question de leur empreinte écologique et de leurs impacts potentiels, positifs comme négatifs »*. La deuxième serait de les rendre *« plus créatives, en s'intéressant moins exclusivement à la performance technologique et économique et davantage à l'exploration de modèles alternatifs : l'open source comme vecteur de diffusion de solutions vertueuses ; le partage, la collaboration et les « communs » comme alternatives positives à la surconsommation ; le « pair à pair » et les systèmes d'échange alternatifs…»* (Demailly et al., 2017). C'est aussi ce que propose l'article de l'association CEMÉA pour un numérique soutenable (CEMÉA Éducation Permanente, 2023).

> *« Pour permettre cette transition, nous devons articuler usage individuel et politique, éducation et législation. Le numérique responsable, la sobriété numérique doivent, non seulement imposer des restrictions aux entreprises et être force de propositions et de régulation, mais également "faire sens" collectivement. »*

Ainsi, on comprend la nécessité d'une approche multi-échelle afin de trouver des solutions pérennes et efficaces.

Finalement, la question de la compatibilité entre la croissance de l'intelligence artificielle et les impératifs de sobriété écologique s'inscrit dans un contexte où innovation technologique et contraintes environnementales se heurtent de plus en plus fréquemment. D'après certains organismes publics et certains acteurs privés, les différentes approches explorées – de la frugalité dans le développement des modèles d'IA à l'interventionnisme gouvernemental – montrent que des solutions existent pour réduire l'impact environnemental de l'IA tout en maintenant son utilité et sa pertinence.

[13] Citation originale : "these actions represent changes at a small level, and often at an individual level. The magnitude of greenhouse gas emissions, the urgency of mitigating climate change, and the existential threat to our planet demand collective action."

La culture de l'innovation qui sous-tend le développement de l'IA contribue indéniablement à la compétitivité des entreprises et au progrès technologique. D'après les organismes privés et publics s'intéressant à l'impact environnemental de l'IA, cette dynamique devrait être réorientée pour mieux intégrer les enjeux de durabilité, notamment en promouvant des usages réellement justifiés et en limitant la croissance des modèles énergivores. Les incitations à adopter des pratiques plus sobres, combinées à des initiatives collectives et à une régulation proactive, pourraient constituer un levier efficace pour équilibrer ces impératifs.

Enfin, une approche multidimensionnelle, impliquant acteurs publics, privés et citoyens, paraît s'imposer, selon de nombreux acteurs, comme une nécessité pour répondre aux défis posés par le développement de l'IA. En adoptant une vision globale et collaborative, il serait possible d'envisager un avenir où la transition numérique s'inscrit dans une perspective durable et responsable, conciliant innovation et respect des limites planétaires.

Conclusion

Au terme de notre exploration, la controverse autour de l'intelligence artificielle et de son impact sur l'environnement met en lumière des tensions profondes entre les ambitions technologiques et les impératifs environnementaux. Ces tensions reflètent les visions contrastées des différents acteurs impliqués, chacun contribuant à dessiner les contours de cette controverse.

L'explosion récente des usages de l'intelligence artificielle et l'essor de l'économie qui l'accompagne ont suscité de nombreuses interrogations sur le besoin réel qu'elle prétend satisfaire. Tandis que les grandes entreprises internationales (ex. : GAFAM, Nvidia) défendent son expansion en mettant en avant son utilité et ses bénéfices économiques tout en promouvant les solutions durables qui leur seraient associées, les associations environnementales dénoncent la création artificielle de besoins et les impacts cachés d'une telle croissance. Les organisations publiques, comme l'ARCEP, tentent d'encadrer ce développement dans un cadre durable, mais peinent à contrebalancer l'influence des grandes entreprises.

Ces débats sur la nécessité de l'IA s'intensifient face aux conséquences environnementales qu'elle génère. Les émissions de CO_2 croissantes, la forte consommation de ressources liée à la *« dématérialisation »*, ou encore le *« mensonge de l'immatériel »* dénoncé par la communauté scientifique, illustrent les limites d'une technologie souvent présentée comme immatérielle. La complexité d'évaluer l'impact réel des objets numériques alimente également les critiques portées par les associations et les think tanks.

Paradoxalement, l'IA est aussi perçue comme une solution pour atténuer l'impact environnemental du numérique. Des initiatives soutenues par le Ministère de la Transition écologique ou le Département américain de l'énergie montrent son potentiel à optimiser les ressources et réduire les émissions dans des secteurs variés. Certains industriels et chercheurs soulignent même sa capacité à transformer profondément des domaines clés grâce à des optimisations de haute intensité.

Ainsi, tous les acteurs s'accordent sur une interrogation centrale : quelle est la véritable finalité de l'IA ? Si son potentiel pour relever des défis environnementaux est indéniable, ses impacts restent souvent sous-évalués ou occultés, ce qui nourrit un débat essentiel sur l'équilibre à trouver entre innovation technologique et durabilité.

La finalité de l'IA dans un contexte environnemental pousse les acteurs à s'interroger sur les meilleures manières de mesurer son impact. Chercheurs, législateurs et industriels tentent de s'accorder sur l'évaluation de cet impact, depuis la création et l'entraînement des modèles jusqu'à leur fin de cycle de vie. Ils proposent des bilans carbone et des analyses de cycle de vie à l'aide d'outils spécialisés. Ces évaluations incluent les infrastructures utilisées, leur construction, ainsi que leur consommation en électricité, ressources minérales ou

hydriques, en particulier pour la création et le refroidissement des data centers et processeurs.

Parmi les bonnes pratiques envisagées pour limiter les émissions inutiles figurent l'utilisation de modèles plus spécialisés et énergétiquement efficients, ou encore le réentraînement partiel de modèles existants afin de réduire les coûts fixes d'entraînement.

Pour permettre des comparaisons fiables entre les modèles, les outils de mesure doivent être réplicables. Les acteurs insistent sur l'importance de prendre en compte l'entraînement, l'inférence, le mix énergétique de la région où se situent les data centers et de distinguer de manière systématique la consommation liée à chaque ordinateur des utilisations spécifiques de l'IA, comme une requête ChatGPT.

Cependant, chercheurs, associations et presse s'accordent sur le fait que ces mesures ne gagneront en fiabilité que si les entreprises technologiques deviennent plus transparentes concernant leurs émissions de carbone et leur consommation énergétique. Cette exigence croissante de transparence est cruciale pour les chercheurs, qui estiment qu'elle favoriserait la prévention, l'optimisation et la régulation des impacts.

Enfin, académiques et législateurs soulignent la nécessité de prendre en compte les effets rebonds : si l'IA optimise certains processus, comme la production, cela peut entraîner une augmentation de la production, générant ainsi indirectement davantage d'émissions.

Toutes ces interrogations autour de l'IA soulèvent la question de la compatibilité entre sobriété et compétitivité. L'IA contribue fortement à la culture de l'innovation et de la croissance promue par les grandes entreprises, mais les importants intérêts économiques et les investissements massifs creusent le fossé entre ces entreprises technologiques et les associations ou régulateurs prônant la sobriété.

D'après les ONG, il est indispensable d'accepter une perte de rapidité et de précision des modèles d'IA pour aller vers plus de sobriété. Cependant, les entreprises ne semblent pas prêtes à renoncer à leurs avantages compétitifs si leurs concurrents ne font pas de même. Les législateurs tentent donc de favoriser une approche plus sobre énergétiquement, avec des plans environnementaux proposant des feuilles de route. Toutefois, ces recommandations restent souvent de simples incitations. On observe une divergence entre les objectifs annoncés, comme la neutralité carbone à l'horizon 2030, et la croissance rapide des parts de marché des entreprises technologiques dans l'IA.

Les gouvernements et institutions collaborent à différents niveaux pour réguler l'impact environnemental de l'IA et du numérique, notamment à travers des lois et initiatives comme la loi REEN en France ou le Climate Neutral Data Center Pact en Europe. Cependant, beaucoup de ces mesures restent non contraignantes. Malgré ces efforts, des contradictions persistent entre les objectifs de croissance économique et la sobriété énergétique, comme le soulignent les ONG et certains acteurs du secteur, qui appellent à une approche plus cohérente et partagée.

Une transition réussie nécessite une coordination multidimensionnelle et internationale impliquant tous les acteurs. Elle pourrait s'appuyer sur des modèles alternatifs, la collaboration et les systèmes de partage, pour concilier durabilité et innovation.

Ainsi, il apparaît que l'IA n'est ni intrinsèquement une menace ni un remède pour l'environnement. Tout dépend de la manière dont elle est mise en œuvre et encadrée. La controverse révèle un besoin urgent de coordination entre acteurs, non seulement pour établir des normes partagées, mais aussi pour repenser les priorités du développement technologique dans un contexte de crise climatique.

En fin de compte, cette controverse n'est pas qu'une opposition entre des visions antagonistes : elle est aussi une opportunité. Elle pose un débat pour redéfinir les conditions d'un progrès technologique en harmonie avec les exigences écologiques. Mais cette opportunité ne pourra être saisie que si entreprises, gouvernements et société civile parviennent à converger vers des objectifs communs, où l'innovation cesse d'être une fin en soi et devient un moyen au service d'un avenir durable.

Matériel et méthodes

Le présent travail résulte tout d'abord d'une analyse de la presse nationale et internationale, depuis les années 2015, ainsi qu'une analyse quantitative réalisée sur des corpus d'environ 500 articles de presse à l'aide d'Europresse et d'environ 900 articles scientifiques regroupés sur Scopus. L'utilisation de Cortext nous a permis de mettre en évidence les connexions entre les différents nœuds de notre sujet, les acteurs clés ainsi que l'évolution temporelle du débat.

Ce travail initial a permis l'élaboration de grilles de questions adressées à trois acteurs de la controverse analysée, au cours d'entretiens semi-directifs. Les témoignages de ces acteurs, qui occupent tous des fonctions différentes, ont été retranscrits puis analysés, et des extraits pertinents ont été exploités dans ce document. Nous avons eu l'occasion d'échanger avec :

- un expert Data Centers, IA et Énergie au sein du Shift Project, instructeur et animateur de la fresque du numérique.
- une porte-parole de l'association GreenIT
- une chargée de mission IA au sein du Ministère de la Transition Écologique

Ces entretiens ont été menés à partir de questions préparées en amont mais ont également laissé une part importante à une expression libre de nos interlocuteurs. Les transcriptions ont fait l'objet d'une analyse spécifique, pour en faire ressortir les nœuds de controverse, notamment ceux pouvant y émerger sans être ressortis de notre lecture d'articles de presse.

Le contenu de notre étude a par la suite été étoffé grâce aux éléments trouvés dans des rapports d'experts ou institutionnels. Les données techniques qu'ils recèlent sont venues enrichir notre travail et l'éclairer sous un angle nouveau, différent mais complémentaire des synthèses de la presse et des prises de position affichées par les divers acteurs dans leurs interviews. Nous nous sommes en particulier penchés sur les rapports effectués par l'ARCEP, l'AFNOR, le département américain de l'énergie ou encore du Shift Project.

Tous les documents, rapports ou autres utilisés lors de ce travail, cités ou non ci-dessus, sont listés dans les références présentes ci-dessous.

Il convient toutefois de souligner que notre enquête a été effectuée dans un laps de temps assez restreint - trois mois - et que le corpus de témoignages aurait gagné à être enrichi. Nous regrettons par exemple l'absence de témoignages d'acteurs américains directement impliqués dans les affaires les plus médiatisées de la controverse, tels que les GAFAM.

Références

Articles de presse généraliste / presse professionnelle

- Charruau, M. (2023, juin 27). Intelligence artificielle : « Certes complexe, l'IA peut aider à la décarbonation ». Sud Ouest.
- Flipo, F. (2022, novembre 29). Sobriété énergétique, écoresponsabilité numérique... De quoi parle-t-on exactement ? The conversation.
- Gelles, D. (2024, octobre 22). The A.I. Power Grab. The New York Times.
- Georges, B. (2020, novembre 23). Le défi énergétique de l'intelligence artificielle. LesEchos.
- Hutson, M. (2022, juin 26). Measuring AI's Carbon Footprint New tools track and reduce emissions from machine learning. IEEE Spectrum.
- Jérémy, R. (2024, septembre 3). Intelligence Artificielle : Définition, histoire, enjeux. Datascientest.
- Leparmentier, A. (2024, juin 15). Aux Etats-Unis, l'insatiable appétit d'électricité de la tech inquiète. Le monde.
- L'intelligence artificielle au service du bilan carbone. (2024, juin 24). climatico.
- Mollaret, G. (2024, mars 22). «On émet soixante fois plus de carbone en opérant une requête sur ChatGPT que sur Google». Le Figaro.
- Niedercorn, F. (2019, juin 11). Quand une IA émet autant de CO2 que cinq voitures. LesEchos.fr.
- Pitron, G. (2021). L'enfer numérique : Voyage au bout d'un Like. Éditions les Liens qui libèrent.
- Renouard, G. (2023, mai 24). BTP : Comment l'intelligence artificielle peut accélérer la réduction des émissions de CO2. La tribune.
- Séramour, C. (2024a, mai 17). Les plans des géants technologiques pour devenir carbone neutre mis à mal par l'IA. L'usine digitale.
- Séramour, C. (2024b, mai 17). Les plans des géants technologiques pour devenir carbone neutre mis à mal par l'IA. L'usine digitale.
- Six, N. (2024, mars 25). Intelligence artificielle : Le bilan carbone de la génération d'images, de textes ou de sous-titres. Le Minde.

Articles de revue scientifique

- Awesso, D.-N. E. (2021). À la recherche d'une (in)compatibilité entre IA, Blockchain et Droit de l'environnement. HAL Open Sciences.

- CEMÉA Éducation Permanente. (2023). Un numérique soutenable est-il possible ? CEMÉA Éducation Permanente.

- Demailly, D., Francou, R., Kaplan, D., & Saujot, M. (2017). Faire converger les transitions numérique et écologique: Annales des Mines - Responsabilité et environnement, N° 87(3). https://doi.org/10.3917/re1.087.0013

- Devaux, C., & Nicolaï, J.-P. (2023). Vers un devoir de vigilance numérique (ou comment intégrer le problème de la pollution numérique dans la gouvernance des entreprises). Revue Interdisciplinaire Droit et Organisations-RIDO, 4, 21.

- Dhar, P. (2020). The carbon impact of artificial intelligence. Nature Machine Intelligence, 2(8). https://doi.org/10.1038/s42256-020-0219-9

- Duberry, J. (2021). Intelligence artificielle, données volumineuses et conservation de la biodiversité. VertigO, Volume 21 Numéro 2. https://doi.org/10.4000/vertigo.32470

- Gloukhovtsev, M. (2024). Making IT Sustainable : Techniques and Applications (1st ed). Elsevier Science & Technology.

- Guillaumont, M. (2023). Méthodologie pour mesurer et réduire l'impact environnementale des algorithmes de Deep Learning afin de concevoir des projets en IA éco-responsables. Capgemini, 2.

- Han, B. A., Varshney, K. R., LaDeau, S., Subramaniam, A., Weathers, K. C., & Zwart, J. (2023). A synergistic future for AI and ecology. Proceedings of the National Academy of Sciences, e2220283120. https://doi.org/10.1073/pnas.2220283120

- Istrate, R., Tulus, V., Grass, R. N., Vanbever, L., Stark, W. J., & Guillén-Gosálbez, G. (2024). The environmental sustainability of digital content consumption. Nature Communications, 3724. https://doi.org/10.1038/s41467-024-47621-w

- Le Goff, T. (2024). Recommandations pour une action publique en faveur d'une IA générative respectueuse de l'environnement. HAL Open Sciences.

- Loup, C. (2022, juin 28). Les implications écologiques des data centers à Marseille. Séminaire MédiaLab, SciencePo.

- Luxery-Bitri, A., Mallarino, D., & Ninassi, B. (2023). La sobriété numérique : Un impératif envrionnemental ? La collection numérique.

- Mallarino, D., Le Bras, S., & Bonamy, C. (2022). Les impacts environnementaux et sociétaux des données : Un défi pour l'avenir. JRES.

- Marquet, C. (2020). Binaire Béton : Quand les infrastructures numériques aménagent la ville. Université Paris Saclay.

- Marquet, C. (2024). Environnement—Le big data au secours de la crise climatique ? HAL Open Sciences.

- Munister, V. D., Zolkin, A. L., Malikov, V. N., Kosnikova, O. V., & Poskryakov, I. A. (2021). Computational analysis of the digital footprint using machine learning and artificial intelligence. Journal of Physics: Conference Series, 2094(3), 032003. https://doi.org/10.1088/1742-6596/2094/3/032003

- Pachot, A., & Patissier, C. (2022). Chapitre III. L'IA réduit la facture énergétique. In Intelligence artificielle et environnement : Alliance ou nuisance ?. Dunod.

- Preite, L., & Vignali, G. (2024). Artificial intelligence to optimize water consumption in agriculture : A predictive algorithm-based irrigation management system. Computers and Electronics in Agriculture, 109126. https://doi.org/10.1016/j.compag.2024.109126

- Tamburrini, G. (2022). The AI Carbon Footprint and Responsibilities of AI Scientists. Philosophies, 4. https://doi.org/10.3390/philosophies7010004

- Trihinas, D., Thamsen, L., Beilharz, J., & Symeonides, M. (2022). Towards Energy Consumption and Carbon Footprint Testing for AI-driven IoT Services. 2022 IEEE International Conference on Cloud Engineering (IC2E). https://doi.org/10.1109/IC2E55432.2022.00011

- Truhn, D., Müller-Franzes, G., & Kather, J. N. (2023). The ecological footprint of medical AI. European Radiology, 34(2). https://doi.org/10.1007/s00330-023-10123-2

- Zavyalova, E. B. (2022). Artificial Intelligence—Reducing the Carbon Footprint? In Industry 4. 0 : Fighting Climate Change in the Economy of the Future. Springer International Publishing AG.

Rapports

- Arcep. (2019). L'empreinte carbone du numérique (No. 5; Réseaux du futur). Arcep.

- Arcep. (2020). Rapport pour un numérique responsable. Arcep.

- Benes, K. J., Porterfield, J. E., & Yang, C. (2024). AI for Energy : Opportunities for a Modern Grid and Clean Energy Economy. U.S. Department of Energy.

- Commission de l'intelligence artificielle. (2024). AI: Notre ambition pour la France. Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique de France.

- Ferreboeuf, H., Efoui-Hess, M., & Kahraman, Z. (2018). LEAN ICT-Pour une sobriété numérique. The Shift Project.

- Ferreboeuf, H., Efoui-hess, M., Marrauld, L., & Lescop, C. (2020). Déployer la sobriété numérique. The Shift Project.

- Fropier, J., Mongiet, P., Ribeaucourt, G., Riou, P., Semede, A., & Trystram, D. (2024). Afnor Spec 2314 [Norme]. Afnor.

- Kirkpatrick, K. (2023). The Carbon Footprint of Artificial Intelligence. Communications of the ACM, 66(8). https://doi.org/10.1145/3603746

- L'Impact Écologique de l'Intelligence Artificielle : Un Défi à l'Ère du Numérique. (2024). Marketing Digital HEC Montreal.
- Linhart, D., Carles, O., & Le doré, A. (2020). Impact carbone de la connexion internet. Objectif Carbone.
- Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing : Watts Driving the Cost of AI Deployment? The 2024 ACM Conference on Fairness, Accountability, and Transparency. https://doi.org/10.1145/3630106.3658542
- Maurey, H., Chaize, P., CHEVROLLIER, G., & HOULLEGATTE, J. (2020). Rapport d'information sur l'empreinte environnementale du numérique. Sénat.
- NVIDIA. (2024). NVIDIA Sustainability Report Fiscal Year 2024.
- Sandalow, D., McCormick, C., Kucukelbir, A., Friedmann, J., Nagrani, T., Fan, Z., Halff, A., d'Aspremont, A., Glatt, R., Méndez Leal, E., Karl, K., & Ruane, A. (2023). Artificial Intelligence for Climate Change Mitigation Roadmap. ICEF Innovation Roadmap Project.

Sites Web

- Aruba. (s. d.). Energy production. Aruba.it. Consulté le 19/12/2024.