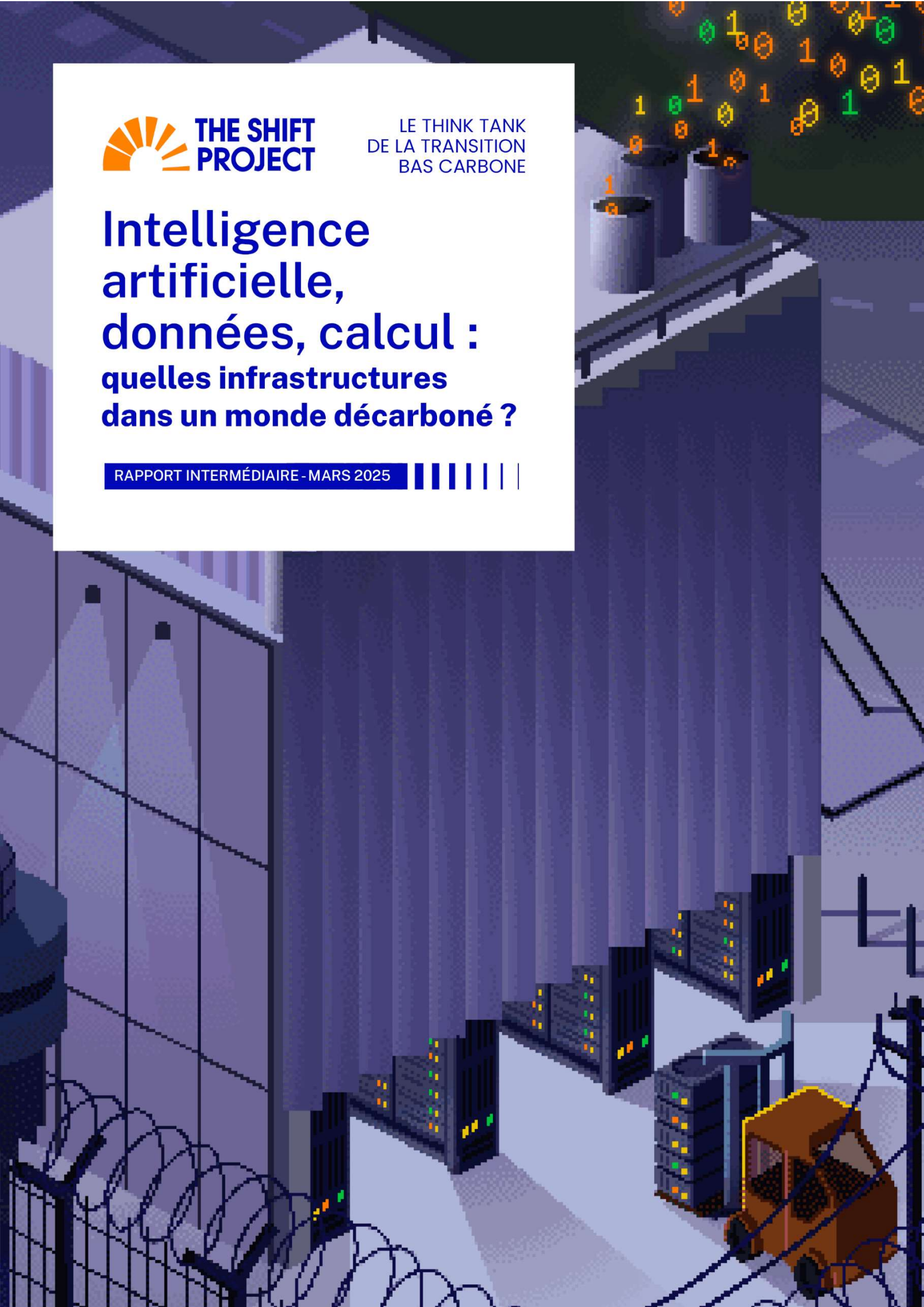




LE THINK TANK
DE LA TRANSITION
BAS CARBONE

Intelligence artificielle, données, calcul : quelles infrastructures dans un monde décarboné ?

RAPPORT INTERMÉDIAIRE - MARS 2025



A propos : Le think tank *The Shift Project* (TSP)

Le *Shift Project* est un *think tank* qui œuvre en faveur d'une économie libérée de la contrainte carbone. Association loi 1901 reconnue d'intérêt général et **guidée par l'exigence de la rigueur scientifique**, sa mission est d'éclairer et d'influencer le débat sur la **transition énergétique et climatique** en Europe. Il est soutenu depuis 2010 par des grandes **entreprises** françaises et européennes ainsi que par des **organismes publics**, des associations d'entreprises et des PME. Il est épaulé par un réseau de dizaines de milliers de bénévoles regroupés au sein d'une association loi 1901 : ***The Shifters***, créée en 2014 pour apporter un soutien bénévole au Shift Project.

Il ambitionne de mobiliser les entreprises, les pouvoirs publics et les corps intermédiaires sur les risques, mais aussi et surtout sur les opportunités engendrées par la « **double contrainte carbone** » que représentent ensemble les tensions sur **l'approvisionnement énergétique et le changement climatique**. Sa démarche est ainsi marquée par un **prisme d'analyse particulier**, fondé sur la conviction que l'énergie est un facteur de développement de premier ordre : dès lors, les risques induits par le changement climatique, intimement liés à l'usage de l'énergie, relèvent d'une **complexité systémique** et transdisciplinaire particulière. Les enjeux climat-énergie conditionnent l'avenir de l'humanité ; il est donc nécessaire d'intégrer cette dimension le plus rapidement possible à notre modèle de société.

Votre participation aux travaux : relectures et contributions

Le travail qui vous est présenté ici est exploratoire : il vise à initier de nouvelles discussions et pose sur de nombreux sujets davantage de questions qu'il n'en résout. Bien qu'il soit déjà le fruit d'un travail collectif, **ce rapport intermédiaire est encore un document de travail imparfait, incomplet et évolutif.**

Comme vous le constaterez, **les résultats de nos analyses quantitatives n'ont pas encore été pleinement réalisés.** Pour les phases de travail à venir, **toute source de données que vous trouveriez essentielles à intégrer à nos travaux seront les bienvenues.**

Vos retours seront également précieux sur l'ensemble du document et des éléments qui vous y sont présentés : méthodologie, approche générale et choix des angles choisis sont ici présentés pour être soumis aux avis des parties prenantes du sujet des technologies numériques, de leur décarbonation et de leurs enjeux de résilience.

Dans cette logique, nous vous prions d'envoyer vos remarques, critiques et propositions aux contacts indiqués en fin de rapport.

Vous n'avez bien entendu pas besoin de lire l'ensemble de ce long document pour nous aider à l'améliorer : toutes les contributions sont les bienvenues.

Nous avons besoin de vous !

Comité de rédaction

Les auteurs

Sylvain BAUDOIN (Membre du Groupe expert, The Shift Project)

Aurélie BUGEAU (Professeur des universités, Université de Bordeaux, LaBRI)

Marlène DE BANK (Ingénieure de recherche « Numérique », The Shift Project)

Pauline DENIS (Chargée de projet « Numérique », The Shift Project)

Maxime EFOUI-HESS (Coordinateur du programme « Numérique », The Shift Project)

Thomas FELTIN (Ingénieur de recherche, Cisco Systems France)

Hugues FERREBOEUF (Chef de projet « Numérique », The Shift Project)

Arnaud GUEGUEN (Consultant Climat, indépendant)

Siba HAIDAR (Enseignante-chercheuse, ESIEA)

Alexis LEAUTIER (Ingénieur et membre de l'association The Shifters)

Anne-Laure LIGOZAT (Professeur des universités, ensIIE, LISN)

Ludovic MOULARD (Directeur Développement Durable, fifty-five / The Brandtech Group)

Denis PERILHOU (Expert associé, The Shift Project)

Alexandre THEVE (Directeur R&D, Davidson Consulting)

Aurore VAUDATIN (Product Environmental Manager, OVHcloud)

Tristan VUILLIER (Responsable Sustainability France, Devoteam)

Martin WILLAME (Doctorant, Université Catholique de Louvain)

Les interprétations, positions et recommandations figurant dans ce rapport ne peuvent être attribuées ni aux contributeurs, ni aux relecteurs, ni aux membres du groupe de travail cités ci-dessus. Le contenu de ce rapport n'engage que The Shift Project.

Partenaires mécènes

Un grand merci aux mécènes engagés à nos côtés en faveur de la sobriété numérique, qui rendent possible la publication de ce travail :



Table des matières

A PROPOS : LE THINK TANK <i>THE SHIFT PROJECT</i> (TSP)	1
Votre participation aux travaux : relectures et contributions	2
Comité de rédaction	3
Table des matières	5
Table des figures	7
INTRODUCTION	9
I. Le numérique, à la fois outil et défi pour la décarbonation de l'économie	9
II. Une trajectoire insoutenable qu'il s'agit d'infléchir	10
III. Pourquoi travailler sur l'empreinte énergie-climat de l'intelligence artificielle et du développement des capacités informatiques ?	12
IV. Les centres de données : aujourd'hui une observable difficilement observée, demain un indicateur de pilotage	14
QUELLES EVOLUTIONS DES CAPACITES INFORMATIQUES MONDIALES EN RAISON DU « PHENOMENE IA GENERATIVE » ?	16
I. Les leçons du passé	17
II. L'état des lieux des tendances actuelles du secteur	21
A. Etat des lieux	21
B. Evaluation de la consommation électrique et mise en perspective	24
III. Quelle empreinte carbone pour les centres de données mondiaux ?	25
IV. Les dynamiques mondiales informatiques et énergétiques : l'exemple des Etats-Unis	27
V. Quels impacts sur les terminaux et sur les réseaux ?	31
CAPACITES DE CALCUL EN FRANCE ET EN EUROPE : DIMENSIONNEMENTS POSSIBLES, TRAJECTOIRES D'EMISSIONS DE GAZ A EFFETS DE SERRE, QUELLES GOUVERNANCES POUR ETRE A LA HAUTEUR DES ENJEUX	37
I. Le cas de l'Irlande	37
II. L'état des lieux en France	40
A. Numérique & climat	40
B. Numérique & électricité	41
C. Numérique & électrification des autres secteurs	44
III. Une première réflexion sur les modélisations possibles et leurs objectifs	52

A. Objectifs d'éclairage énergétique et climatique	52
B. Critères, contraintes, périmètre, découpage.....	52
C. Modélisation	54
D. Scénarisation	56
IV. Intégration territoriale et régionale des centres de données : de premières questions pour une stratégie résiliente ?	57
V. L'état des lieux en Europe	58
UNE IA AU SERVICE DE L'INTERET GENERAL DOIT ETRE A LA HAUTEUR DES ENJEUX CLIMATIQUES ET DE TRANSITION ENERGETIQUE	63
I. Approche par cas d'usage : quels services et promesses technologiques pour l'IA ?.....	63
A. IA et IA générative : une rupture technologique ?.....	63
B. Les cas d'usage : une illustration du panorama d'applications et services possibles d'intelligence artificielle	65
II. Impact énergie-climat de l'IA : analyse fonctionnelle des cas d'usage	67
A. Approche par cas d'usage : critères retenus et description technologique	67
B. La "Boussole de l'IA" : de l'analyse fonctionnelle du cas d'usage à la qualification de ses effets sur les impacts énergie-climat.....	69
C. Quels sont les facteurs inflationnistes	70
D. Vision systémique : sortir de la vision sectorielle.....	71
III. Quantification sur un cas d'usage.....	71
IV. Pour des approches "bilan net" et systémique.....	71
RECOMMANDATIONS	73
CONCLUSIONS	74
REFERENCES	76
ANNEXE 1 : TYPOLOGIE DE CENTRES DE DONNEES	88
ANNEXE 2 : UN PREMIER CALCUL : "QUELLE EMPREINTE CARBONE POUR LES PROJETS DE CENTRES DE DONNEES MONDIAUX ?"	95
ANNEXE 3 : REALISER ET/OU INTERPRETER UNE ESTIMATION DE LA CONSOMMATION ELECTRIQUE DES CENTRES DE DONNEES : QUELLES QUESTIONS SE POSER ?	96
ANNEXE 4 : LES CRYPTO-DEVICES, UN USAGE DES INFRASTRUCTURES NUMERIQUES TOUJOURS PRESENT ?	99

Table des figures

Figure 1 - Distribution de l’empreinte carbone du numérique mondial en 2019, par poste pour les phases de production (40 %) et d’utilisation (60 %). Source : (The Shift Project, 2021)	10
Figure 2 - Nos usages et nos réseaux sont les deux faces d’une même dynamique. Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport	13
Figure 3 - Historique des estimations en consommation électrique des centres de données dans le monde entre 2020 et 2030 (échelle verticale démarrant à 100 TWh pour le graphique de gauche). Source : (LBNL et al., 2024).....	19
Figure 4 - Historique des estimations en consommation électrique des centres de données entre 2020 et 2030 aux Etats-Unis Source : (LBNL et al., 2024)	19
Figure 5 - Historique de la consommation électrique des centres de données entre 2014 et 2023 aux Etats-Unis. Source : LBNL, 2024	20
Figure 6 - Puissance électrique : historique 2018 - 2023. Sources : (DCByte, 2024; Jefferies, 2024).....	20
Figure 7 - État des lieux des prévisions de demande de puissance électrique pour l’IT ou pour les centres de données dans les centres de données mondiaux. Sources : (DCByte, 2024; Jefferies, 2024; McKinsey & Company, 2024; McKinsey, 2024; Schroders, 2024; SemiAnalysis et al., 2024).....	22
Figure 8 - État des lieux des prévisions de consommation électrique annuelle des centres de données mondiaux. Sources : (Deloitte, 2024; Goldman Sachs, 2024; Hintemann R. & Hinterholzer S., 2020; IEA, 2024a, 2024b; Jefferies, 2024; LBNL et al., 2024; McKinsey & Company, 2024; SemiAnalysis et al., 2024)	22
Figure 9 - Traduction en émissions de gaz à effets de serre (GtCO ₂ e) de plusieurs estimations. Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport.....	25
Figure 10 - Investissements dans les centres de données aux Etats-Unis (à gauche), coût des processeurs graphiques (au centre), taille et complexité des modèles d’intelligence artificielle (à droite). Source : (IEA & IEA Paris, 2024c, 2024b, 2024a)	27
Figure 11 - Augmentation des densités de puissance dans les centres de données. Source : (Uptime Intelligence, 2023, 2024).....	28
Figure 13 – Sondage Uptime Intelligence sur les applications « denses ». Source : (Uptime Intelligence, 2024).....	28
Figure 14 - Evolution du PUE. Sources : (Arcep, 2023; Uptime Intelligence, 2024)	29
Figure 15 - Projections de trafic internet et d’usages à 2033. Source : (Nokia, 2024)	34
Figure 16 - Projections de trafic internet et d’usages à 2033. Source : (Nokia, 2024)	34
Figure 17 - Projections de trafic internet et d’usages à 2033. Source : (Nokia, 2024)	34
Figure 18 - Projection de trafic internet à 2033, décomposition entre consommateur et entreprise, et croissance générée par l’IA. Source : (Nokia, 2024).....	35
Figure 19 - Demande d’électricité par secteur entre 1990 et 2023 en Irlande. Source : (Prof. Hannah Daly, University College Cork, 2024)	38
Figure 20 - Consommation électrique des centres de données en Irlande, dépassant la consommation urbaine résidentielle. Source : (Central Statistics Office (CSO), 2024; The Journal, 2024)	39
Figure 21 - Estimation de la consommation d’électricité des centres de données et de sa part dans la demande totale d’électricité en Irlande. Source : (IEA, 2024a)	39

Figure 22 - La carte régionale des 35 sites de centres de données dédiés à l'IA identifiés par le gouvernement. Source : (DCMag, 2025a)	43
Figure 23 - Consommation d'électricité en 2050 selon les principaux scénarios de transition français Source : Site comprendre2050.fr, décryptage "Quelle évolution de la consommation électrique pour une France bas carbone ?" (ADEME, négaWatt, The Shift Project, 2025)	44
Figure 24 - Consommation d'électricité dans le secteur des transports en 2050. Source : Site comprendre2050.fr, décryptage « Quelle consommation d'électricité dans les transports de demain ? » (ADEME, négaWatt, The Shift Project, 2025)	45
Figure 25 - Consommation d'électricité dans le secteur des transports en 2015, 2030 et 2050. Source : (ADEME, 2021; RTE, 2022).....	45
Figure 26 - Consommation d'électricité dans le secteur des transports en 2050. Source : Site comprendre2050.fr, décryptage « Quelle consommation d'électricité des industries en 2050 ? » (ADEME, négaWatt, The Shift Project, 2025).....	46
Figure 27 - Consommation d'électricité dans le secteur de l'industrie en 2015, 2030 et 2050. Source : RTE Futurs énergétiques, ADEME Transition(s) 2050.....	46
Figure 28 - Consommation d'électricité dans le secteur résidentiel entre 2005 et 2050 dans le scénario de référence de RTE. Source : (RTE, 2022).....	47
Figure 29 - Projections de consommation d'électricité en France (écarts à 2020, projections pour les centres de données et pour l'électrification de la mobilité). Sources : (ADEME & Arcep, 2023; RTE, 2022, 2023, 2024).....	Erreur ! Signet non défini.
Figure 30 - Scénarios possibles pour l'évolution de la demande énergétique des centres de données dans l'UE28 jusqu'en 2030. Source : (European Commission, 2020).....	58
Figure 31 - Estimations pour la consommation énergétique des centres de données européens. Source : (European Commission, Kamiya G., Bertoldi P., 2024)	58
Figure 32 - Projection de consommation énergétique des centres de données en Europe en TWh. Source : (McKinsey & Company, 2024)	59
Figure 33 - Projection de demande d'installation de capacités informatiques en Europe GW. Source : (McKinsey & Company, 2024)	60
Figure 34 - Estimation de la consommation d'énergie des centres de données par pays en EU27 en 2022 en TWh , Source : (European Commission, Kamiya G., Bertoldi P., 2024).	61
Figure 35 - Distribution des serveurs par type de centres de données. Source : (LBNL et al., 2024).....	89
Figure 36 - Les composants fonctionnels d'un data center. Source : (L'institut Paris Région, 2023).....	90
Figure 37 - Structure d'un centre de données. Source : (Ahmed M. U. et al., 2021).....	91
Figure 38 - Répartition de la consommation électrique d'un centre de données. Source : (Ahmed M. U. et al., 2021).....	92
Figure 39 - Répartition des impacts du "Scope 3 Empreinte embarquée" d'un centre de données. Source : (Schneider Electric, 2023).....	92
Figure 40 - Indicateurs de la directive U4E. Source : (U4E et al., 2025).....	93
Figure 41 - Valeurs des indicateurs de la directive U4E. Source : (U4E et al., 2025)	93
Figure 42 - Système de notation proposé dans la directive U4E. Source : (U4E et al., 2025)	94
Figure 43 - Estimation de la consommation électrique pour les centres de données dans le monde. Source : (IEA, 2024a), p. 35.....	99

Introduction

I. Le numérique, à la fois outil et défi pour la décarbonation de l'économie

Les technologies de l'information, aujourd'hui centrales dans des activités essentielles de nos sociétés, pourraient jouer un rôle crucial dans la transformation de notre économie. Si les équipements numériques et les usages qu'ils permettent – et promettent – semblent être conçus pour relever des défis toujours plus grands, cela ne doit pas les affranchir d'une réflexion sur leur pertinence environnementale. En effet, dans un monde où les ressources sont finies, il est important de se souvenir que chaque action réclame de l'énergie et de la matière, y compris transformer, créer, stocker ou échanger des informations. Les technologies numériques ne font donc pas exception : ce ne sont pas des outils virtuels, mais bien des systèmes et supports physiques, bien que nous n'en percevions pas toujours directement la matérialité.

Les technologies numériques forment un système d'envergure mondiale : les terminaux (smartphones, ordinateurs, tablettes etc.) se connectent entre eux via des infrastructures réseaux (câbles terrestres et sous-marins, antennes de réseaux mobiles, fibres optiques etc.) afin d'échanger des informations stockées et traitées dans les centres de données, cœurs battants de ce système. Or, chacun de ces éléments nécessite de l'énergie non seulement pour fonctionner (phase d'utilisation) mais également, avant cela, pour être construit (phase de production) : extraction minière et raffinage des matières premières, processus industriels de transformation et de fabrication puis livraison aux consommateurs et consommatrices : tout le cycle de vie de ces éléments nécessite des ressources biotiques et abiotiques.

Chaque service numérique s'appuie sur des infrastructures physiques dont la résilience et la pertinence vis-à-vis de la **double contrainte carbone (réduire les émissions carbonées de nos activités ; s'affranchir de notre dépendance aux énergies fossiles)** doivent être interrogées. Le numérique est un catalyseur : là où il est déployé, il permet d'optimiser, accélérer, fluidifier, paralléliser... L'intelligence artificielle est l'un de ces outils, et l'IA générative vient encore y ajouter de nouvelles couches d'usages.

Déployer tous ces systèmes sans stratégie de long terme intégrant la double contrainte carbone mène à l'accélération de toutes les dynamiques actuelles, y compris les plus éloignées de nos objectifs de résilience. Faire du numérique un véritable outil de réinvention de nos activités pour les rendre compatibles avec les limites planétaires réclame une prise en compte systémique des impacts du numérique.

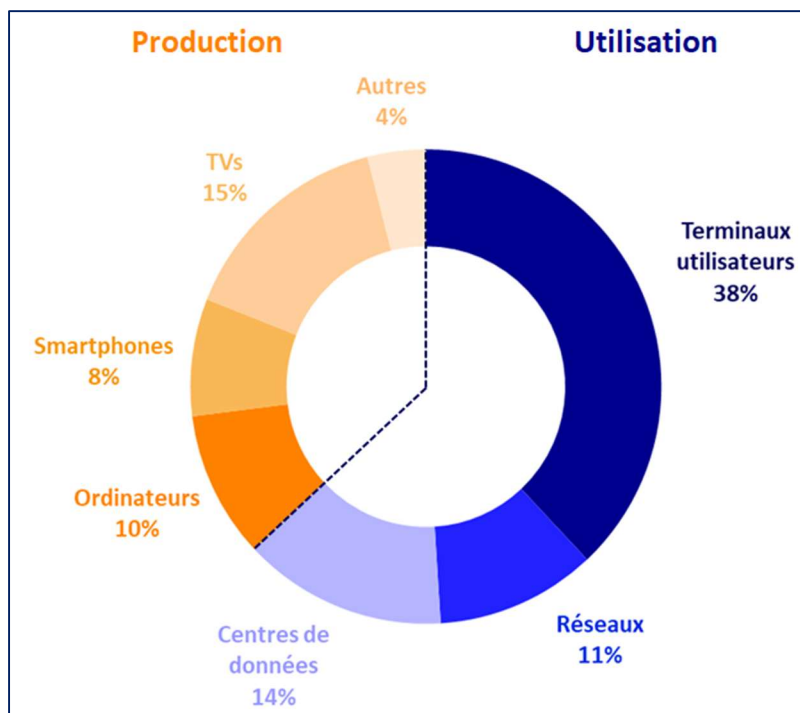


Figure 1 - Distribution de l'empreinte carbone du numérique mondial en 2019, par poste pour les phases de production (40 %) et d'utilisation (60 %). Source : (The Shift Project, 2021)

II. Une trajectoire insoutenable qu'il s'agit d'infléchir

Le numérique représente déjà près de 4 % des émissions mondiales (The Shift Project, 2021), soit du même ordre que l'intégralité des véhicules utilitaires lourds dans le monde (IEA, 2021a). A l'échelle française, il représente 4,4 % de l'empreinte carbone du pays¹ (ADEME, 2025).

La particularité du secteur numérique tient à la rapidité d'augmentation de ses émissions, qui croissent selon une tendance particulièrement incompatible avec sa décarbonation : + 6 %/an en moyenne au niveau mondial (The Shift Project, 2021) et + 2 à 4 %/an en France (ADEME & Arcep, 2023; HCC, 2020; Sénat, 2020). Les optimisations techniques et opérationnelles ne parviennent pas à compenser le déploiement soutenu de ses infrastructures, parcs et flux (ADEME & Arcep, 2023; Bol et al., 2020; European Commission, 2020; GreenIT.fr, 2019; IEA, 2022, 2024a; LBNL et al., 2024; The Shift Project, 2023). Ce constat continue de se vérifier et s'est illustré au cours des cinq dernières années, qui devaient pourtant selon certaines études marquer un plafonnement de ces impacts grâce au progrès technologique (IEA, 2019; ITU-T, 2020; Masanet E. et al., 2020). Les déploiements à large échelle de l'IA et de sa composante

¹ Données pour l'année 2022

généraliseraient et pourraient aggraver ces dynamiques déjà insoutenables. Ce rapport aspire donc à les éclairer.

A l'échelle de la France comme à l'échelle mondiale, le numérique représente en 2022 de l'ordre de 10 % de la consommation électrique totale (ADEME, 2025; The Shift Project, 2021, 2023), une part significative qui devrait lui valoir le même niveau d'attention que d'autres secteurs. Dans un contexte d'électrification intense des usages (mobilité, bâtiment, industrie etc.), on comprend qu'il se trouve lui aussi au cœur des enjeux de planification de la transformation de nos systèmes et de priorisation d'accès aux ressources désormais en tension, dont l'électricité fait partie (The Shift Project, 2023b).

Rendre le numérique compatible avec la double contrainte carbone consiste donc non pas seulement à renforcer des leviers d'optimisation déjà déployés, mais à le placer sur une trajectoire fondamentalement différente de celle qu'il suit actuellement. Au même titre que les autres secteurs de l'économie, il doit atteindre son objectif de décarbonation, que les acteurs industriels (GSMA, GeSI²) se sont eux-mêmes fixés par le biais de l'initiative SBTi et sur la base d'une recommandation de l'ITU (SBTi et al., 2020) (p. 9) à - 45 % en 2030 par rapport à 2020 au niveau mondial³.

The Shift Project propose de prendre cet objectif comme base de la construction de la trajectoire nationale, en l'adaptant aux spécificités de la décarbonation déjà importante du mix électrique du pays. The Shift Project préconise de construire la trajectoire française autour de cet objectif SBTi recalculé pour l'adapter au cas français, de - 30 % des émissions du secteur à 2030 par rapport à 2020 (The Shift Project, 2023).

La construction de notre système numérique se fait au travers d'interactions multiples entre le système technique et les usages qu'il sous-tend. L'analyse des enjeux énergétiques et climatiques selon une approche systémique met en évidence que la maîtrise des impacts des technologies numériques nécessite une réflexion approfondie sur le déploiement des offres, l'adoption des usages que nous favorisons ou non, ainsi que la place de la sobriété. La mobilisation des leviers de sobriété (inflexion des volumes de terminaux, de données et de calculs) est même l'une des conditions *sine qua non* à la maîtrise de la consommation énergétique grâce aux gains d'efficacité énergétique (The Shift Project, 2023).

² GSMA : GSM Association, association rassemblant les acteurs internationaux de la connectivité mobile (constructeurs, opérateurs etc.). | GeSI : Global enabling Sustainability Initiative, groupement d'acteurs internationaux du numérique et des télécommunications, dont la mission est d'œuvrer sur le numérique durable.

³ Un cadre d'engagements non quantitatif existe au niveau national, pris par les industriels français dans le cadre des feuilles de routes de décarbonation du secteur du numérique.

III. Pourquoi travailler sur l’empreinte énergie-climat de l’intelligence artificielle et du développement des capacités informatiques ?

Le 30 novembre 2022, l’intelligence artificielle générative a fait une entrée fracassante dans le débat public avec la sortie de l’agent conversationnel ChatGPT. L’adoption fut fulgurante : 1 million d’utilisateurs en 5 jours, 100 millions en 2 mois (Sagot B., 2023), avec 400 millions de visites hebdomadaires (L’usine digitale & Seramour C., 2025).

Depuis, les nouveaux usages permis par l’intelligence artificielle générative⁴ viennent développer un système numérique et plus largement une économie déjà soumise dans leur ensemble à la double contrainte carbone. A l’instar des autres plateformes et moteurs de recherche en ligne ciblant le grand public, ChatGPT repose sur une stratégie volumique visant une adoption par le plus grand nombre, mais se distingue par une intensité de calcul bien plus élevée à la fois en amont (à la construction du service) et en aval (à l’usage du service). Cela entraîne mécaniquement un coût marginal environnemental plus important : **l’adoption massive d’un service par construction intense en calcul amplifie significativement la pression s’exerçant sur les ressources électriques et l’environnement.**

Cette intelligence artificielle générative grand public est le résultat :

- De recherches académiques et de R&D dans le domaine du traitement automatique des langues (et son intégration dans une interface d’interaction homme-machine « directe » et « naturelle »),
- De stratégies technologiques d’accroissement de puissance des processeurs et des processeurs graphiques,
- De la délégation des capacités de stockage et de calcul vers des infrastructures à forts effets d’économie d’échelle (cloud),
- Du développement et de l’accessibilité de grands corpus de textes, souvent enrichis d’annotations linguistiques permis par la collecte et le stockage de données.

Ces dynamiques, propres à l’intelligence artificielle générative, s’intègrent dans une logique systémique : le numérique fonctionne en système et ses trois tiers (terminaux, infrastructures réseaux, centres de données) évoluent de concert, rendus interdépendants par les échanges de données. Les choix de déploiements faits au niveau des usages et infrastructures de la donnée impactent l’ensemble du système numérique tout en étant le résultat de la trajectoire générale donnée au système (Figure 2) :

- La volonté de développer les nouveaux usages promis par le moment « IA générative » justifie le déploiement de nouvelles capacités techniques (amélioration des capacités des processeurs, augmentation des capacités des baies de serveurs, déploiement des centres de données pour atteindre les capacités de calcul et stockage nécessaires etc.) : c’est l’effet

⁴ Voir paragraphe « IA et IA générative : une rupture technologique ? »

d'usage ;

- Le déploiement de nouvelles capacités techniques entraîne le développement de nouveaux usages (possibilité d'avoir accès rapidement et simplement à une réponse ou une image produite grâce des capacités de calculs très importantes etc.) : c'est l'effet d'offre.

Usages et infrastructures restent les deux faces d'une même pièce. Ce qui est peut-être particulier dans ce moment « IA générative », c'est la rapidité avec laquelle la boucle est activée : les investissements financiers, humains et techniques sont mobilisés de manière considérable, à la fois pour alimenter le déploiement massif de centres de données et pour catalyser l'adoption des usages alors rendus possibles, justifiant ainsi la poursuite des dynamiques.

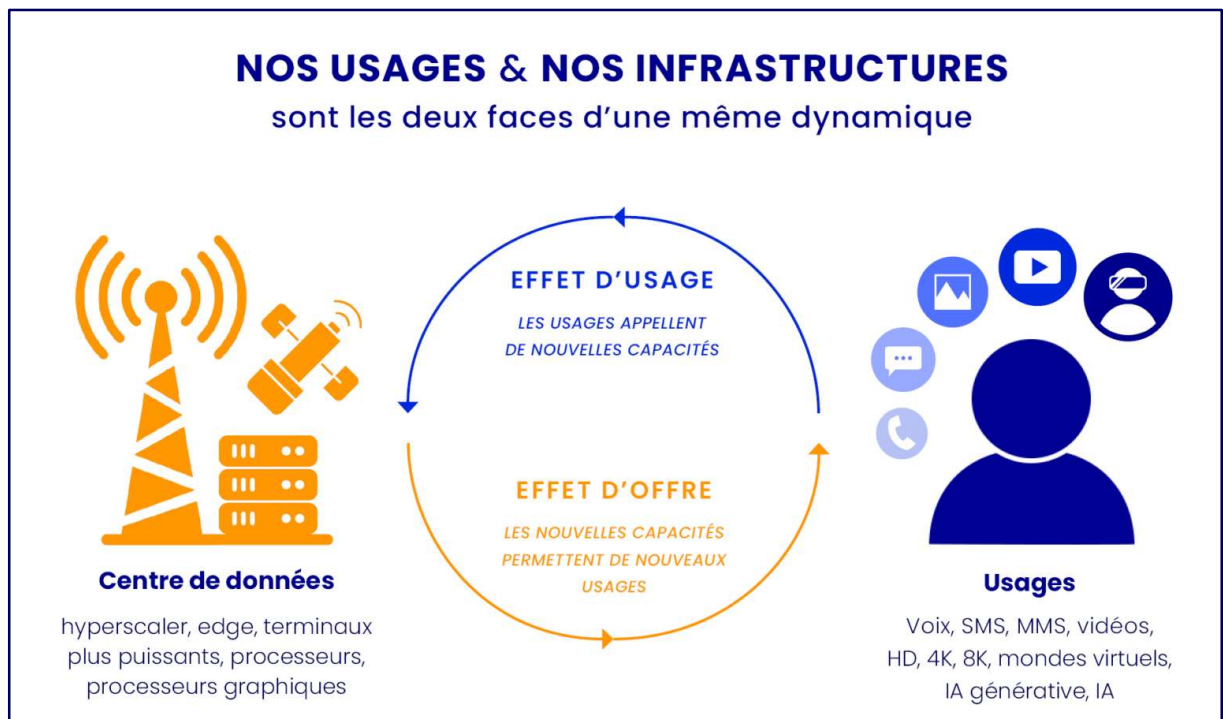


Figure 2 - Nos usages et nos réseaux sont les deux faces d'une même dynamique.
Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport

Ce rapport vise à étudier les dynamiques structurant les usages de la données (infrastructures de calcul, de stockage, dynamiques de déploiement des applications intenses en ressources informatiques etc.) avec une attention particulière sur les modifications du système numérique induites par les nouveaux usages de l'intelligence artificielle générative et ses briques technologiques et infrastructurelles constituantes (nouveaux processeurs, remplacement de serveurs, nouveaux centres de données et de nouveaux types, remplacement accéléré d'équipements obsolètes pour l'IA, terminaux avec plus de puissance informatique etc.).

Les axes de travail sont les suivants :

- Selon quelles tendances évolue l'empreinte énergie-carbone des centres de données à l'échelle mondiale ? L'amplification récente des dynamiques numériques avec le « phénomène IA » (sur les centres de données, les terminaux et les réseaux) retarde-t-elle la possibilité d'inversion des tendances d'évolution des émissions de gaz à effets de serre du système numérique, indispensable à l'atteinte des objectifs du secteurs (SBTi et al., 2020) (p9) ?

- Quelles sont les implications pour la demande et le secteur électrique ? Quels sont les enjeux à intégrer dans les réflexions stratégiques à l'échelle française et européenne ?
- Comment placer l'IA, et plus largement l'accroissement des capacités informatiques, sur une trajectoire compatible avec les contraintes physiques ? Quels choix sont à faire pour les rendre pertinentes dans leur contribution à la décarbonation du secteur numérique et à celles des autres secteurs ?

IV. Les centres de données : aujourd'hui une observable difficilement observée, demain un indicateur de pilotage

Si les effets du déploiement large spectre de l'intelligence artificielle (générative) et de ses dynamiques dépassent largement le seul tiers "centres de données", entraînant une déformation et une inflation du système numérique dans son ensemble, **les centres de données sont aujourd'hui un élément central de cette nouvelle dynamique numérique.**

Leur consommation énergétique a été estimée à plusieurs reprises par le passé mais n'est pas suivie en temps réel. A l'échelle mondiale, le saut de 200 TWh à 460 TWh entre les deux publications IEA (IEA, 2021b, 2024a) ou entre celles de Masanet et LBNL (LBNL et al., 2024; Masanet E. et al., 2020) est révélateur de cette difficulté à mesurer et prévoir les consommations énergétiques à l'échelle de 3 à 5 ans. De même, à l'échelle nationale, les travaux "Futurs énergétiques 2050", "Bilan Prévisionnel 2035", "Schéma Décennal de Développement du Réseau 2040" (RTE, 2022, 2023, 2024) font des estimations ponctuelles mais l'absence d'inventaire public des projets de centres de données en cours est préoccupante, d'autant plus compte-tenu du changement d'échelle annoncé lors du Sommet pour l'Action sur l'Intelligence Artificielle tenu en 2025 à Paris⁵. **L'empreinte carbone embarquée des centres de données reste quant à elle un angle mort des analyses et projections**, tout comme les enjeux locaux tels que l'utilisation de la ressource en eau et l'artificialisation des sols. **La mise en place d'inventaires de projets de centres de données pour suivre leurs implantations, leurs consommations énergétiques et leurs émissions de gaz à effets de serre est de plus en plus essentielle** mais est aujourd'hui manquante pour les scénarios de transition.

Faire des émissions de gaz à effets de serre des centres de données sur l'ensemble de leur cycle de vie un indicateur de pilotage des politiques publiques dédiées permettrait de s'assurer que les rythmes d'efficacité énergétique, d'intensité carbone et de demande sont compatibles avec les trajectoires et objectifs de référence du secteur et du territoire.

Aux risques de résilience et de souveraineté posés par le numérique, et l'IA en particulier, ces dynamiques portent ceux de la mise en danger de la transition énergie-climat et de la fragilisation

⁵ Lors du Sommet pour l'action sur l'Intelligence Artificielle ayant pris place en France les 10-11 février 2025 (Elysée, 2025), des annonces de centres de données de grande ampleur ont été faites : 1 GW sur un campus axé IA (Le Monde, 2025a) et 1 GW à Cambrai (Le Monde, 2025b). Ceci place certains centres de données au même niveau que certaines industries d'ampleur comme l'illustre l'IEA : les centres de données sont plus concentrés spatialement que les usines sidérurgiques, mines de charbon et de minerais, centrales électriques et entrepôts (fig 4.12 dans (IEA, 2024b)). Ce changement d'échelle de la puissance de calcul du tertiaire vers l'industriel est un phénomène à acter afin de proposer des politiques adaptées (Carnino G. & Marquet C., 2022), notamment climatiques.

des réseaux de transport d'électricité, tant au niveau mondial qu'en France. Le cas de l'Irlande et les décisions d'utilisation d'énergies fossiles aux Etats-Unis pour alimenter les centres de données, sur lesquels nous reviendrons dans la suite de ce rapport, sont particulièrement illustratifs de ces enjeux. De la même manière, d'autres critères environnementaux devraient être suivis à l'échelle locale voire nationale, et ce d'autant plus qu'il existe certaines réglementations les concernant. Pour en citer trois d'entre eux : l'artificialisation des sols due à la construction de nouveaux centres de données, à suivre dans le cadre de la loi ZAN⁶ du 20 juillet 2023 ; la préservation des milieux naturels et de la biodiversité⁷ suivie par exemple dans le cadre de la loi pour la reconquête de la biodiversité, de la nature et des paysages ; les prélèvements et la consommation d'eau encadrés par la directive-cadre sur l'eau.

⁶ Zéro artificialisation nette

⁷ Les constructions de centre de données peuvent nécessiter le défrichage de bois accueillant différentes espèces protégées. Voir par exemple https://www.mrae.developpement-durable.gouv.fr/IMG/pdf/2023-08-09_marcoussis_91__projet_extension_data_center4_avis_delibere_.pdf.

Quelles évolutions des capacités informatiques mondiales en raison du « phénomène IA générative » ?

Cette question intervient alors que l'intelligence artificielle (IA) et en particulier l'intelligence artificielle générative (IA générative, ou GenAI) sont invoquées par les grands acteurs du numérique pour justifier la mauvaise orientation de leurs trajectoires climatiques (Google, 2024; Microsoft, 2024) qui s'avèrent s'écarter de façon vertigineuse des objectifs qu'ils avaient annoncés il y a quelques années. De la même façon, l'intelligence artificielle est le mot d'ordre choisi pour justifier le déploiement à marche forcée des centres de données aujourd'hui et pour la dizaine d'années à venir⁸ avec :

- A l'échelle française, la proposition d'1 GW pour l'IA générative d'ici 10 ans (France Datacenter & EY Parthenon, 2024), rendue obsolète après le Sommet pour l'action sur l'Intelligence Artificielle (Elysée, 2025) lors duquel les annonces suivantes ont été faites : 1 GW sur un campus axé IA (Le Monde, 2025a), 1 GW à Cambrai (Le Monde, 2025b), 1 troisième GW (L'usine digitale, 2025) et 35 centres de données totalisant 1200 hectares (DCMag, 2025a) (voir figure 18),
- A l'échelle européenne, le déploiement de 35 GW d'ici 2030 (McKinsey & Company, 2024),
- A l'échelle mondiale, le déploiement de 96 GW d'ici 2026, dont au moins entre 1.5 à 3 GW pour chaque GAFAM d'augmentation de puissance IT installée (Schroders, 2024).

Un énoncé un peu simpliste consisterait à dire que le développement des centres de données ne fait que répondre à la demande d'IA et de GenAI, nous complétons cet énoncé afin d'extraire les dynamiques du développement des centres de données :

- D'une part, l'offre se développe : les capacités de calcul (que ce soit dans les centres de données ou dans les terminaux) se développent (Epoch AI, s. d.-b). La configuration matérielle des centres de données répondait auparavant prioritairement (et en moyenne) au besoin de stockage de données alors que le recours à l'IA générative et autres applications intenses en traitement implique de faire croître très fortement les capacités de calcul.
- D'autre part, il faut être conscient qu'implanter un centre de données prend 3 à 5 ans (DCByte, 2024), que les acteurs se basent sur des anticipations et ne réagissent pas simplement à une demande réelle et instantanée pour réaliser des investissements qu'ils auront à cœur de rentabiliser à l'avenir.

⁸ Les chiffres qui suivent sont de premiers ordres de grandeur de ce qui se dit en termes d'annonces et de projections. Les parties qui suivent auront à cœur d'éclairer quant à la lisibilité (ou non) de ces chiffres.

- Certains acteurs de référence n'expliquent même pas le développement des centres de données par l'IA générative, mais choisissent plutôt de mettre en évidence l'engagement au numérique qui a été créé pendant la pandémie de COVID-19 ainsi que le développement du cloud (DCByte, 2024).
- Enfin, il n'est pas impossible que nous soyons face à une période de surinvestissement, les acteurs du secteur annonçant eux-mêmes qu'il s'agit avant tout d'anticiper une croissance future, imaginée gigantesque mais par définition incertaine⁹.

Pour éclairer la question de "quel est l'impact de l'IA générative seule ?" et de son attributionnalité (quelle est la part courante de l'empreinte climatique attribuable à l'IA générative si on répartissait l'infrastructure sur les usages), les calculs bottom-up partant d'un nombre de GPU livrés (de Vries A., 2023) sont intéressants : ils mettent en évidence l'empreinte liée à la fabrication des équipements et ils permettent de sensibiliser à l'impact environnemental de l'IA générative. Cependant, ils peuvent ne pas nous renseigner suffisamment sur les dynamiques structurelles globales afin de les comprendre et les anticiper.

Enfin, que l'intelligence artificielle générative soit le facteur inflationniste ou non, elle est avant tout le symbole de la continuité de l'extension du système numérique vers plus de puissance de calcul et elle possède une aura médiatique (éventuellement en représentant des enjeux économiques et géopolitiques jugés suffisamment impérieux) qui évite d'avoir à justifier cette expansion, quitte à mettre en péril les engagements climatiques.

Cette partie aura donc vocation à traiter de l'empreinte climatique de cette extension autour des centres de données mais plus largement de ce qui relève du stockage et du calcul (et donc aussi les terminaux et les réseaux).

Au stade de la publication de ce rapport intermédiaire, nous souhaitons favoriser les échanges et les retours, si possible constructifs. Nous souhaitons aussi encourager les échanges sur les hypothèses et les données considérées dans les prévisions de consommations énergétiques, souvent passées sous silence alors que pouvant jouer au 1er ordre sur les résultats. Toutes vos remarques, critiques, propositions sur les hypothèses et les données considérées dans les prévisions de consommations énergétiques (souvent passées sous silence alors que pouvant jouer au 1er ordre sur les résultats), ainsi que toute piste de recherche de données ou entretiens à mener seront donc un ingrédient clé pour les suites de ces travaux !

I. Les leçons du passé

A l'échelle mondiale, jusqu'en 2021, l'IEA évaluait une consommation d'électricité des centres de données d'environ 200 TWh (IEA, 2021b) et quasiment stable malgré l'augmentation forte des usages.

⁹ Pour exemple, les déclarations de Sundar Pichai (PDG d'Alphabet, société mère de Google), « Le risque de sous-investir dans l'IA est terriblement supérieur à celui de trop investir », ou de Mark Zuckerberg (PDG de Meta), « [l'IA] vise à construire de la capacité avant que nous n'en ayons besoin, plutôt que trop tard » (Challenges, 2024).

Face au manque crucial de suivi et de remontée d'informations (pas de système de suivi de l'évolution de la consommation électrique réelle des centres de données par exemple), les seuls points de référence ont longtemps été les estimations et modélisations réalisées dans le cadre de travaux académiques, dont les conclusions ont quelquefois dûes être utilisées à la périphérie voire en-dehors de leur domaine de validité.

En 2020, les travaux de Masanet et al. avaient mis l'accent sur les gains d'efficacité énergétique à ne pas sous-estimer, base de l'énoncé ayant accompagné les publications de l'IEA jusqu'en 2021 : *"si les tendances actuelles en matière d'efficacité du matériel et de l'infrastructure des centres de données peuvent être maintenues, la demande énergétique mondiale des centres de données peut rester presque stable jusqu'en 2022, malgré une augmentation de 60 % de la demande de services"* (IEA, 2021b; Masanet E. et al., 2020).

En 2022, l'IEA affiche un intervalle de 240 à 340 TWh (IEA, 2022). En 2023, l'EDNA commence à se demander si la demande n'est pas trop forte comparée aux gains d'efficacité (EDNA, Technology collaboration programme by IEA, 2023). En 2024, l'IEA produit le chiffre de 460 TWh pour l'année 2022 (IEA, 2024a).

L'absence de suivi de l'évolution des centres de données ces dernières années et la confiance excessive placée dans le potentiel des gains d'efficacité semblent avoir freiné toute préparation concernant la production d'électricité pour ces acteurs, la gestion des réseaux de transport d'électricité ainsi que toute vision d'encadrement des émissions de gaz à effets de serre (secteur non inclus dans le marché de permis d'émissions EU-ETS, par exemple). **Ce manque de préparation se fait sentir, dans un secteur numérique aux dynamiques particulièrement intenses.**

Il serait précieux pour l'action publique et internationale de retracer l'historique réel du déploiement des centres de données ainsi que d'évaluer les décisions prises en conséquence. C'est ce qu'amorce la publication du LBNL en revenant sur l'historique des *estimations* passées (Figure 3, (LBNL et al., 2024)). A cet historique des *estimations*, pourrait être ajoutée l'estimation Borderstep qui dès 2020 anticipait cette évolution : 350 TWh en 2017, 500 TWh en 2022 jusqu'à 880 TWh en 2030 (Hintemann R. & Hinterholzer S., 2020). The Shift Project avait quant à lui publié, en 2021, 4 scénarios descriptifs de différentes tendances en usages et en efficacité : partants de 420 TWh de consommation d'énergie finale en 2019, évoluant vers 370 TWh à 549 TWh en 2022, et vers 331 TWh à 734 TWh en 2025 (The Shift Project, 2021).

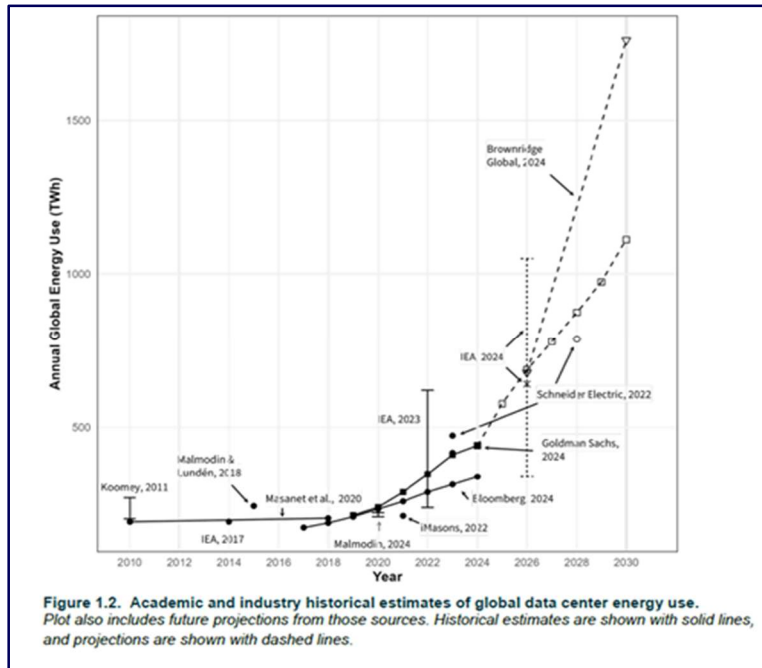


Figure 3 - Historique des estimations en consommation électrique des centres de données dans le monde entre 2020 et 2030 (échelle verticale démarrant à 100 TWh pour le graphique de gauche). Source : (LBNL et al., 2024)

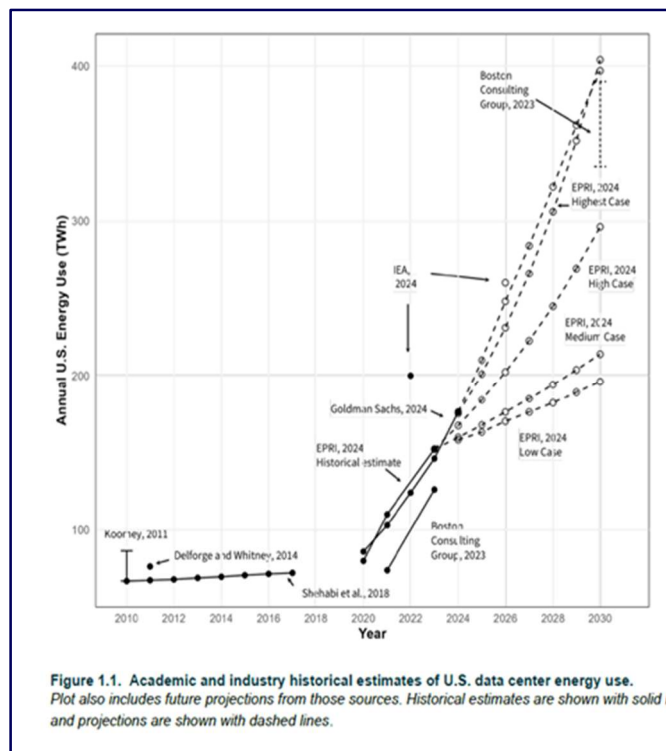


Figure 4 - Historique des estimations en consommation électrique des centres de données entre 2020 et 2030 aux Etats-Unis Source : (LBNL et al., 2024)

LBNL retrace aussi la croissance historique aux Etats-Unis : une multiplication par deux entre 2014 et fin 2021, date de l'introduction de l'intelligence artificielle générative grand public (figure 4).

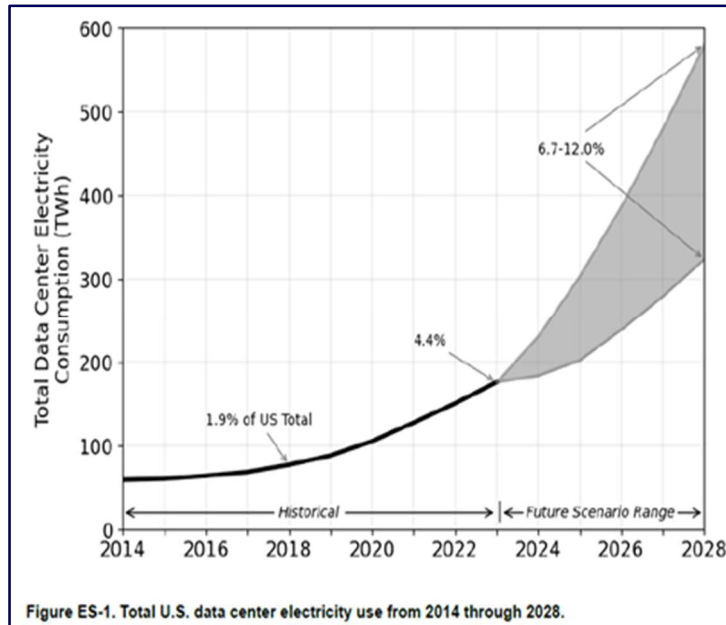


Figure 5 - Historique de la consommation électrique des centres de données entre 2014 et 2023 aux Etats-Unis. Source : LBNL, 2024

A l'échelle mondiale, cette amorce de réflexivité sur les estimations antérieures pourra également être enrichie par des relevés sur les capacités des centres de données : les données DCByte étant en demande de puissance électrique pour l'IT et les données Jefferies en demande de puissance électrique totale pour les centres de données (Figure 6). Ces données mettent en évidence le déploiement à croissance constante des capacités entre 2018 et 2023, et l'accélération depuis 2023 (d'autant plus visible sur la Figure 7).

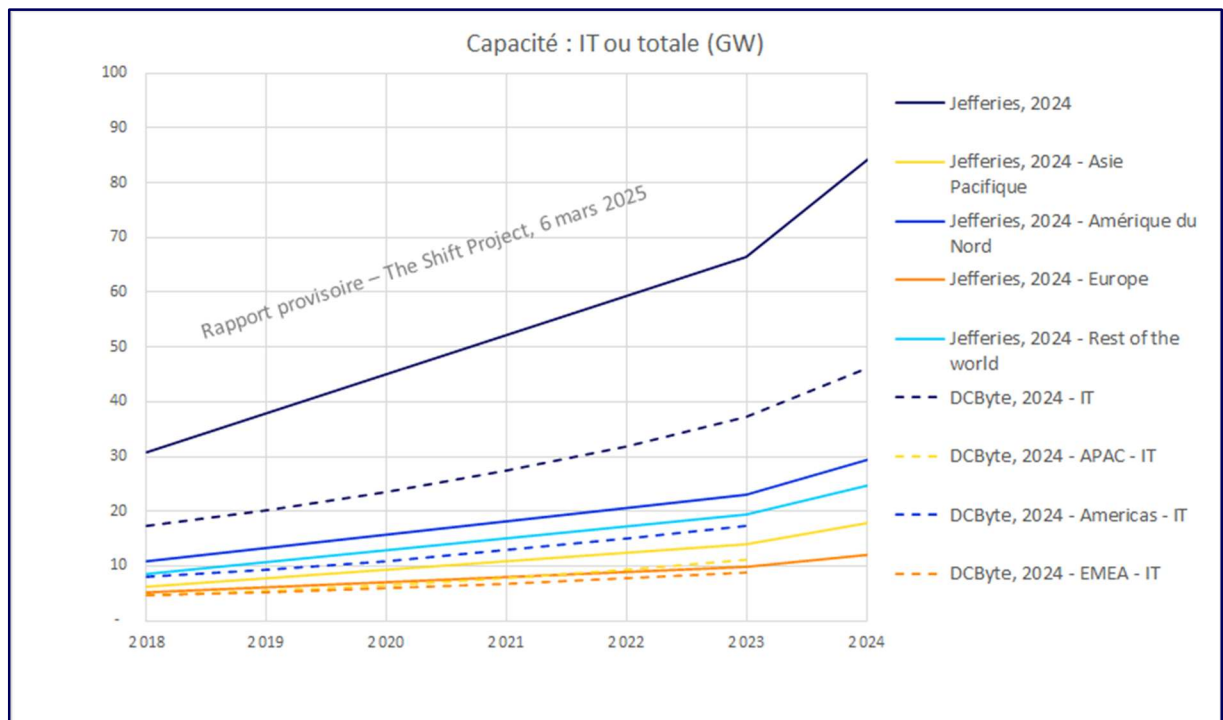


Figure 6 - Puissance électrique : historique 2018 - 2023. Sources : (DCByte, 2024; Jefferies, 2024)

Au vu des dynamiques numériques, ceci doit maintenant précipiter la mise en place de dispositifs de suivi des centres de données, de leurs consommations électriques et de leurs émissions de gaz à effets de serre sur l'ensemble de leur cycle de vie. L'EED, en imposant un reporting pour les centres de données de plus de 500 kW dans une base de données européenne, est en ce sens un point de départ (Journal officiel de l'Union européenne, 2024).

II. L'état des lieux des tendances actuelles du secteur

A. Etat des lieux

Les deux graphiques ci-dessous rassemblent diverses projections de la consommation électrique (TWh) des centres de données et des capacités informatiques (GW) réalisées à la fois par : des acteurs de l'énergie, des analystes de cabinet de conseil en stratégie (généralistes ou spécialisés sur la filière data center), de laboratoires de R&D, d'académiques et de banques d'investissements.

Deux précautions de lecture sur ces graphiques : la première concerne la description des hypothèses dans les rapports consultés bien trop minime au vu de l'importance du sujet ; la seconde concerne les périmètres qui peuvent être différents selon les études. Ce sont deux angles morts qu'il faudra combler à l'avenir dans les bilans mondiaux :

- **Jusqu'ici, de nombreuses projections n'incluent pas tous les types de centres de données et se concentrent uniquement sur les plus grands** : hyperscalers, cloud, colocation (voir « Annexe 1 : Typologie de centres de données »). Les centres de données des entreprises en propre sont encore trop souvent oubliés alors qu'ils peuvent représenter de 5 % à 50 %¹⁰ de la consommation électrique selon la zone géographique et l'année considérées. De plus, les centres de données type « edge » pourraient être plus nombreux à l'avenir, voire la puissance informatique pourrait être beaucoup plus décentralisée (voir « Quels impacts sur les terminaux et sur les réseaux ? ») si notre système numérique évolue vers un modèle en couches où la puissance de traitement est plus distribuée¹¹.
- **Jusqu'ici les rapports se concentraient sur seulement 3 grandes zones** : Etats-Unis, Chine, et Europe. Mais en 2023, on comptait déjà 2,1 GW de puissance IT installée pour la zone South East Asia¹² (BCG, 2024) et 11 GW pour la zone Asia Pacific¹³ (DCByte, 2024) avec 1,2 GW de projets opérationnels ou en cours de réalisation à Mumbai Au Brésil aussi, São Paulo compte déjà 0,7 GW de puissance IT installée et des annonces se font pour un projet à hauteur de 5 GW (Data Center Dynamics, 2024c).

¹⁰ Selon les zones géographiques et les périodes temporelles. Par exemple : 5% projeté au niveau monde en 2028 (LBNL et al., 2024), 23% projeté en Europe en 2028 (McKinsey & Company, 2024), 35% au niveau monde en 2018 (Masanet E. et al., 2020), 50% pour le comptage France territorial en 2020 (ADEME & Arcep, 2023)

¹¹ Ce sujet est d'autant plus important à prendre en compte que ces centres de données pourraient être moins puissants et moins efficaces que les centres de données centraux du fait de mutualisations plus limitées.

¹² Singapour, Malaisie (Johor), Indonésie (Batam), Thaïlande, Philippines, Vietnam

¹³ South East Asia + Inde, Australie, Nouvelle-Zélande, Japon, Corée du Sud + Hong Kong, Taïwan

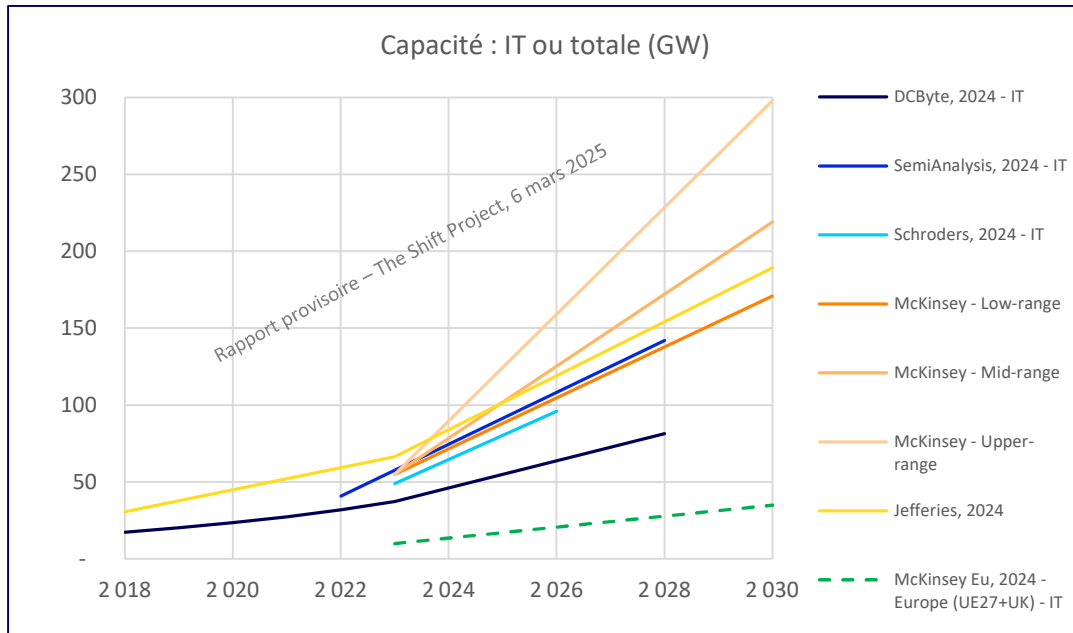


Figure 7 - État des lieux des prévisions de demande de puissance électrique pour l'IT ou pour les centres de données dans les centres de données mondiaux. Sources : (DCByte, 2024; Jefferies, 2024; McKinsey & Company, 2024; McKinsey, 2024; Schrodgers, 2024; SemiAnalysis et al., 2024)

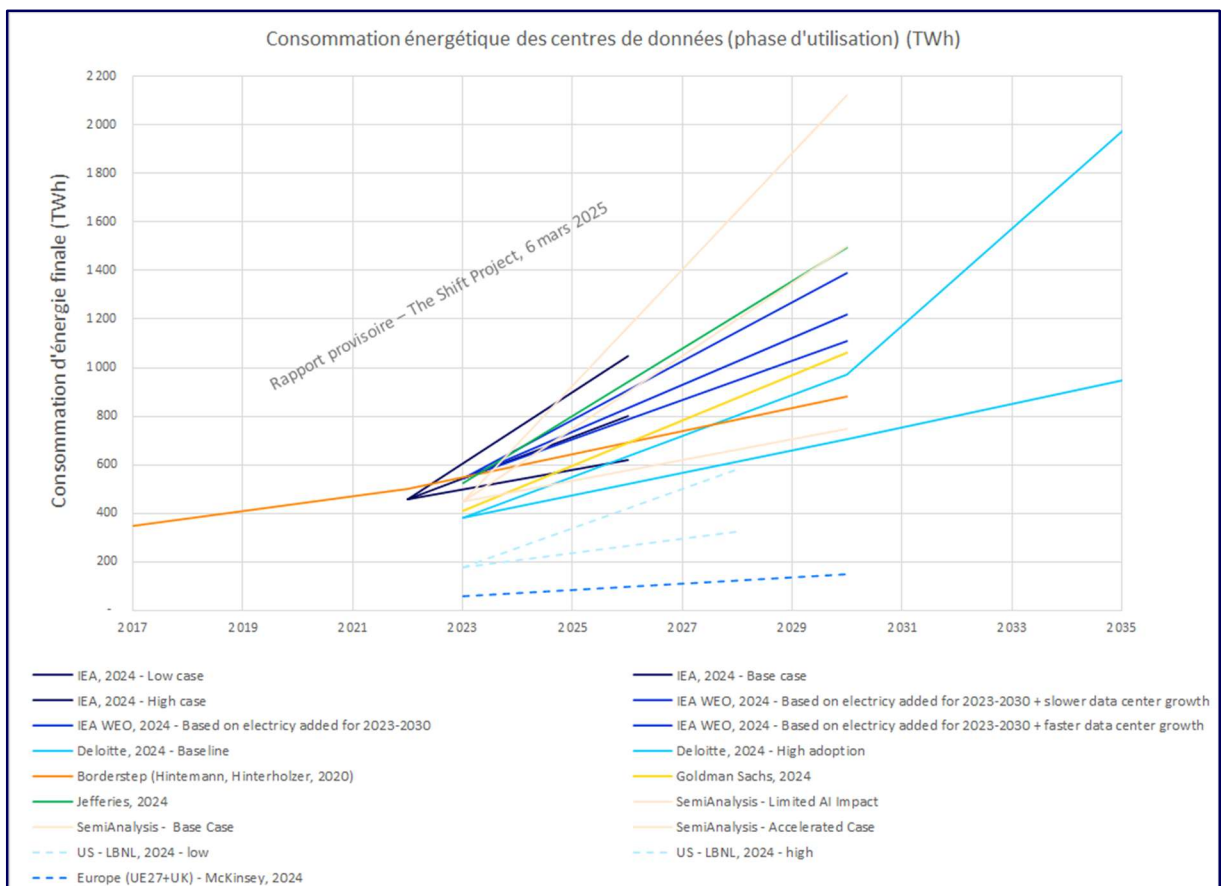


Figure 8 - État des lieux des prévisions de consommation électrique annuelle des centres de données mondiaux. Sources : (Deloitte, 2024; Goldman Sachs, 2024; Hintemann R. & Hinterholzer S., 2020; IEA, 2024a, 2024b; Jefferies, 2024; LBNL et al., 2024; McKinsey & Company, 2024; SemiAnalysis et al., 2024)

Quelles conclusions ?

- **Le statu quo qui prévalait jusqu'en 2021 d'une consommation électrique mondiale annuelle des centres de données quasi-constante autour de 200 TWh est définitivement obsolète.**
- DCByte prévoit 89 GW de puissance IT opérationnelle en 2028, là où Jefferies en estime 189 GW pour les centres de données en 2030.
- En consommation d'énergie, les projections actuelles pour 2030 varient de 700 TWh à 2 100 TWh en 2030, ce qui montre des incertitudes importantes, que tous les acteurs relèvent. Les scénarios de l'AIE dans son "WEO report 2024" vont de 1100 TWh à 1400 TWh environ alors que ceux indiqués dans le "Electricity report 2024 to 2026" amènent (en prolongeant à 2030 les tendances indiquées jusqu'à 2026) à 840 TWh - 1400 TWh - 2400 TWh selon le cas.
- **La majorité des projections sont en consommation d'électricité (TWh) ou en puissance installée (GW) ce qui ne reflète ni le type d'électricité employée ni l'empreinte carbone embarquée**, deux éléments critiques puisque :
 - Certains centres de données (beaucoup aux Etats-Unis) ont annoncé recourir au gaz ou au charbon, impliquant des non fermetures de centrales normalement prévues, ce qui entraîne au moins un facteur 10 sur le facteur d'émission de l'électricité (voir fin du paragraphe « Les dynamiques mondiales informatiques et énergétiques : l'exemple des Etats-Unis »).
 - Jusqu'ici l'empreinte carbone embarquée pouvait représenter entre 1/3¹⁴ de l'empreinte carbone liée à la consommation d'électricité et autant que celle-ci¹⁵(selon le type d'électricité). De plus, les taux de remplacement pourraient évoluer (éventuellement pour aller vers des serveurs plus performants) (peu de données disponibles à date sur ce sujet).
- Jusqu'ici, la consommation électrique annuelle des centres de données augmentait au rythme déjà plus que considérable d'un doublement en dix ans, ceci en raison du développement des usages numériques et en dépit des gains d'efficacité énergétique issus du passage au cloud.

Le déploiement actuel de l'IA générative exacerbe cette tendance d'ici 2030 en induisant une séquence d'investissements massifs sur tous les continents dont la nécessaire rentabilisation **pourrait entraîner un triplement de la consommation en huit ans** (en se basant sur la vision haute de l'IEA dans IEA WEO, aboutissant à une consommation de 1391 TWh).

¹⁴ Dans l'étude (Schneider Electric, 2023) : pour le cas Etats-Unis, le "Scope 3 empreinte embarquée" représente 35 % du Scope 2 ; pour le cas France, le "Scope 3 empreinte embarquée" représente 104 % du Scope 2. Les valeurs sont prises à 15 ans de durée de vie pour un centre de données défini en introduction du rapport.

¹⁵ Idem en prenant un facteur d'émission de l'électricité représentatif de la France.

B. Evaluation de la consommation électrique et mise en perspective

Cette partie est l'un des grands axes de travail restant à poursuivre suite à ce rapport intermédiaire.

Ces travaux viseront à :

- Produire une modélisation prospective transparente de l'évolution de la consommation électrique des centres de données ;
- Mettre en perspective les résultats produits dans la littérature (cf. partie précédente) ;
- Mettre en lumière les facteurs et variables desquelles dépendent les résultats finaux au premier ordre, afin d'éclairer la comparaison des différentes estimations.

Les premiers éléments d'analyses issus de l'amorce de ce travail sont les suivant :

- Il existe aujourd'hui une variabilité très importante dans les résultats (pouvant aller jusqu'à un facteur 2 à horizon 5 ans), au vu du manque de consolidation de certaines données clés et de la sensibilité de certaines hypothèses ;
- La structure de l'outil de modélisation est disponible en « Annexe 3 : Réaliser et/ou interpréter une estimation de la consommation électrique des centres de données : quelles questions se poser ? » et ressource complémentaire (tableur) de ce rapport.

L'un des objectifs cruciaux des suites de ce travail sera donc de récolter, critiquer et consolider les valeurs des paramètres clés de la modélisation, afin de construire une vision d'ensemble transparente. Toutes vos remarques, critiques, propositions sur le contenu ou la structure actuelle de la modélisation, ainsi que toute piste de recherche/obtention de données seront donc un ingrédient clé pour les suites de ces travaux !

III. Quelle empreinte carbone pour les centres de données mondiaux ?

Ce calcul a pour objectif de traduire les trajectoires suivies par les centres de données en émissions de gaz à effets de serre, en incluant : les émissions de gaz à effets de serre à la consommation d'énergie et l'empreinte embarquée liée à la phase de production.

Trois points ont été regardés :

- L'année 2022 se basant sur la consommation de 460 TWh (IEA, 2024a),
- L'année 2030 avec deux consommations de 1220 TWh et 1390 TWh pour prendre en compte le cas de référence et le cas de croissance rapide (IEA, 2024b).

Avec deux hypothèses de facteur d'intensité carbone de l'électricité :

- Le facteur évolue de 460 gCO₂e/kWh en 2022 à 312 gCO₂e/kWh en 2030 (IEA, 2024b) (Stated Policies, p301),
- Le facteur reste à 460 gCO₂e/kWh pour modéliser un ajout de centre de données dont toute la nouvelle demande ne serait satisfaite que par du gaz¹⁶.

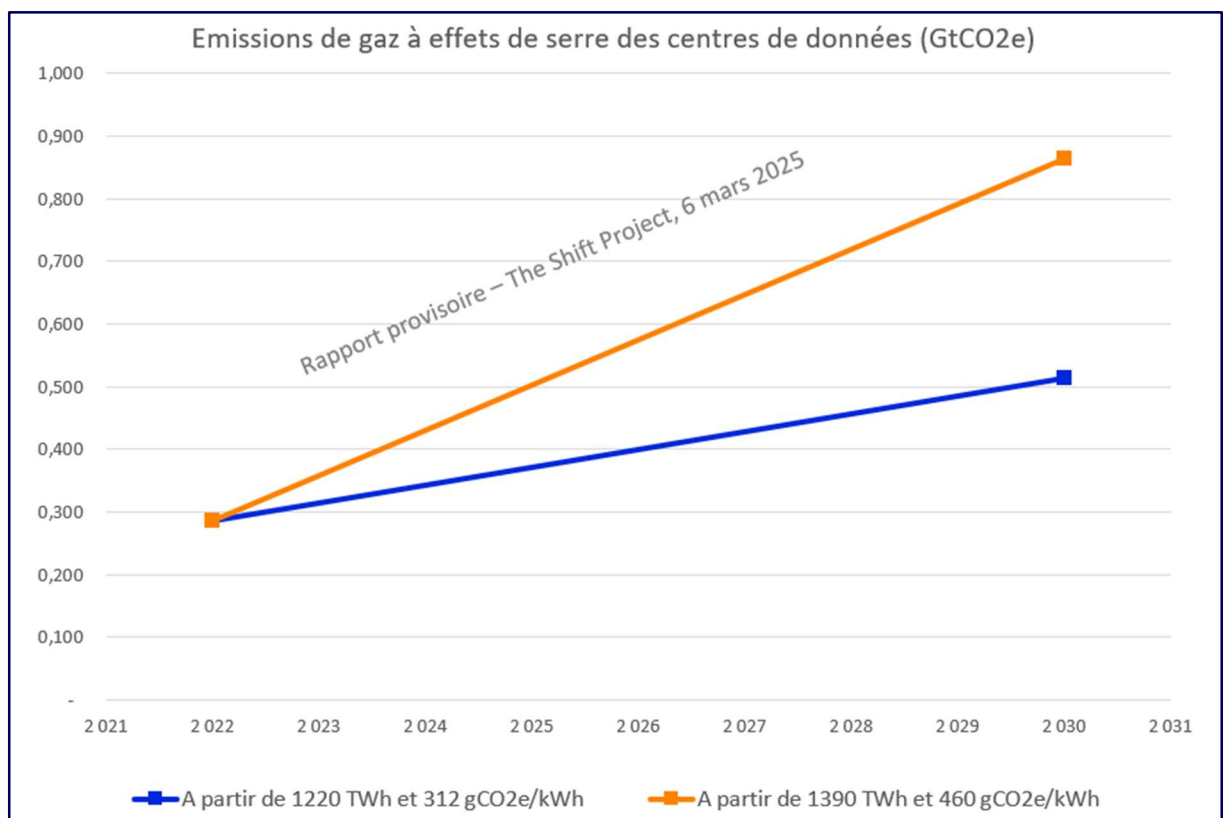


Figure 9 - Traduction en émissions de gaz à effets de serre (GtCO₂e) de plusieurs estimations. Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport

¹⁶ Dont le facteur d'émission est d'environ 450 gCO₂e/kWh.

Entre 2022 et 2030, les émissions de gaz à effets de serre évolueraient vers 514 à 864 MtCO₂e, soit une augmentation de 80 % à 200 %.

Dans 5 ans, à l'horizon 2030, les choix actuels de construction de centres de données à l'échelle mondiale pourraient impliquer des émissions de gaz à effets de serre de l'ordre de 0,86 GtCO₂e, ce qui représente :

- Plus de 3/4 des émissions CO₂ seules du secteur de l'aviation commerciale civile¹⁷,
- 1,6 % à 2,0 % des émissions en 2030, pour les centres de données seuls, selon les scénarios de comparaison¹⁸.

Ce calcul se base sur des projections (décrites dans le paragraphe précédent), dont on rappelle qu'elles peuvent avoir leurs limites : à la fois du côté des gains d'efficacité, mais aussi des disponibilités (énergétique, dont bas-carbone, matérielle, financière, offre rencontrant effectivement sa demande).

Suite à ce rapport intermédiaire, pour améliorer ces projections en empreinte carbone, un travail de scénarisation sur la consommation électrique pourrait être réalisé avec la prise en compte de différentes trajectoires possibles sur l'intensité carbone, sur l'efficacité énergétique, sur l'offre et la demande.

Le calcul est détaillé en « Annexe 2 : Un premier calcul : “Quelle empreinte carbone pour les projets de centres de données mondiaux ?” », veuillez nous contacter pour tout échange technique ou concernant des données supplémentaires que vous pourriez nous fournir. Toutes vos remarques, critiques, propositions sur les orientations explorées à ce stade, ainsi que toute piste de recherche/obtention de données seront donc un ingrédient clé pour les suites de ces travaux !

¹⁷ Synthèse Pouvoir Voler en 2050, p2, 1.1 GtCO₂e en 2018 (The Shift Project & Supaero Decarbo, 2021)

¹⁸ Par comparaison à 54 GtCO₂eq en 2030 en limitant le réchauffement à 3°C (>50%), Table SPM.2 (Intergovernmental Panel On Climate Change (ipcc), 2023), on obtient 1,6 %.

Si on compare à 44 GtCO₂eq en 2030 (en limitant le réchauffement à 2°C (>67%), Table SPM. 2 (Intergovernmental Panel On Climate Change (ipcc), 2023), on obtient 2,0%. La compatibilité entre ces scénarios sur les centres de données et une trajectoire 2°C (>67%) à l'échelle mondiale étant plus que discutable, puisque cette “hausse surprise” de 0,86 GtCO₂e n'est pas compensée par une prise en compte dans les scénarios des autres secteurs.

IV. Les dynamiques mondiales informatiques et énergétiques : l'exemple des Etats-Unis

Au cœur du moteur de ces dynamiques, on trouve l'investissement croissant dans les centres de données aux États-Unis, la réduction significative du coût des processeurs graphiques, ainsi que l'augmentation récente et drastique de la taille des modèles d'intelligence artificielle (Figure 10).

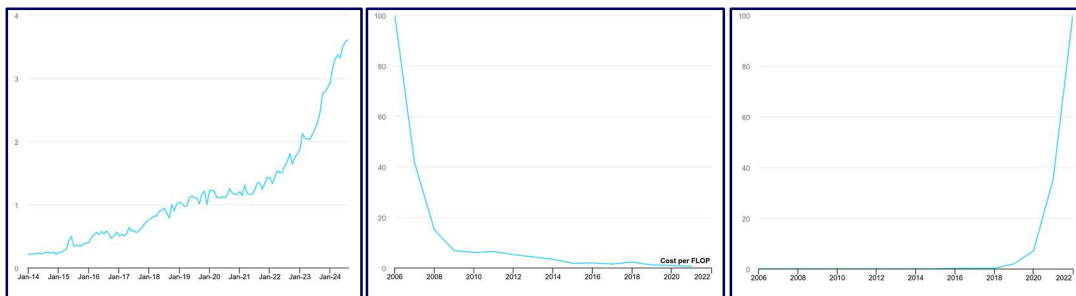


Figure 10 - Investissements dans les centres de données aux Etats-Unis (à gauche)¹⁹, coût des processeurs graphiques (au centre)²⁰, taille et complexité des modèles d'intelligence artificielle (à droite)²¹. Source : (IEA & IEA Paris, 2024c, 2024b, 2024a)

Côté machine learning, selon le site Epoch AI, on constate une augmentation des performances matérielles et algorithmiques (à iso-usage), une augmentation des performances des modèles et une augmentation de la puissance de calcul nécessaire pour entraîner des modèles de machine learning « remarquables » (Epoch AI, s. d.-b).

Côté processeurs graphiques, la puissance des processeurs graphiques, évaluée selon leur TDP²², était de l'ordre de 250 W il y a 2-3 ans, est maintenant de l'ordre du kW, et pourrait encore augmenter²³ d'ici 2 ans.

Pour un parc fixe, cela implique donc une augmentation de la consommation électrique ainsi qu'un remplacement accéléré d'équipements moins puissants (ces équipements pouvant être revendus et utilisés dans un autre parc ou non).

Les densités de puissance augmentent donc, au niveau du serveur comme du centre de données, ce qui peut rendre les structures non adaptées, en refroidissement comme en capacité

¹⁹ Index en ordonnée valant 1 en décembre 2019.

²⁰ Index du coût par flop en ordonnée valant 100 en 2006. Notes: FLOP= Floating Point Operations.

²¹ Index en flop valant 100 en décembre 2019. Notes: The underlying data refers to the Training Compute (FLOP).

²² Thermal design power, à savoir le transfert thermique vers l'extérieur dont doit pouvoir bénéficier ce composant pour fonctionner correctement.

²³ A partir des Thermal Design Power (TDP), en faisant l'approximation qu'ils sont proches de la consommation énergétique (hypothèse qui peut se tenir si le matériel fonctionne à sa capacité maximale), 300W pour le A100 (A100, s. d.), 700W de TDP pour le H100 (H100, s. d.), 2700W pour la GB200 Superchip (GB200 Superchip, s. d.) (si on considère qu'il s'agit d'une seule puce).

Si le site Epoch AI n'est pas certain de la tendance sur 20 ans, il estime qu'il est possible qu'il y ait une augmentation du TDP pour les derniers GPU ML (Epoch AI, s. d.-c).

d'alimentation électrique, si bien que l'Uptime Institute rapporte déjà que 29 % des opérateurs de centres de données mettent à niveau leurs salles²⁴ (Uptime Intelligence, 2024)²⁵.

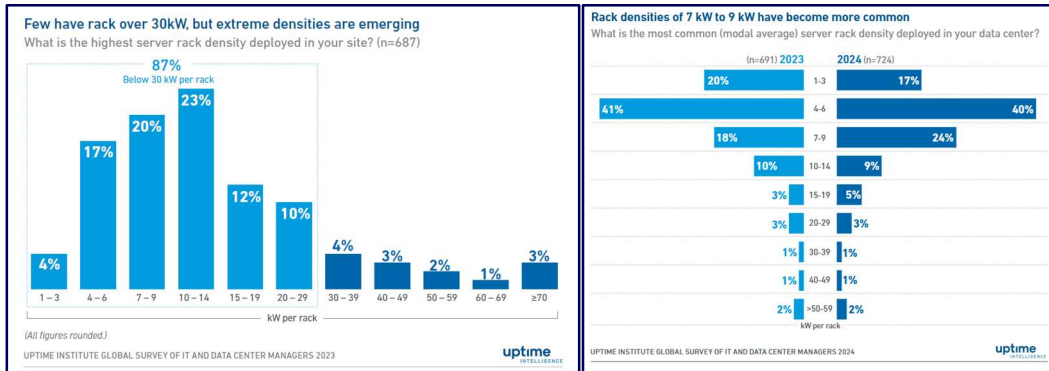


Figure 11 - Augmentation des densités de puissance dans les centres de données. Source : (Uptime Intelligence, 2023, 2024)

Enfin, si on couple cela à une augmentation des usages (permis par une disponibilité de l'offre) : entraînements de modèles, usages « individuels » ou « de bureau » (dont les inférences en augmentation constante (L'usine digitale & Seramour C., 2025)), autres services très intenses plébiscités dans le sondage Uptime Intelligence (Figure 12), ceci conduit à une augmentation de la charge sur les centres de données ; ceux-ci étant construits à l'avance pour que les services soient toujours disponibles.

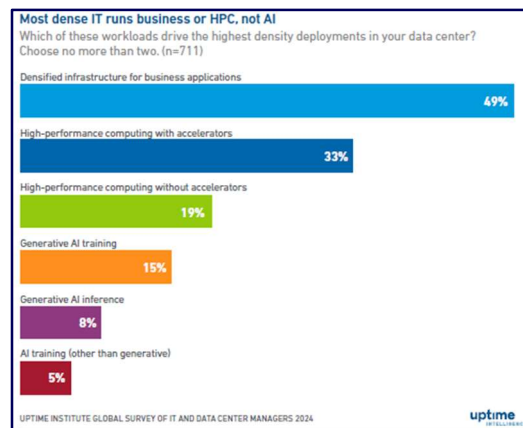


Figure 12 – Sondage Uptime Intelligence sur les applications « denses ». Source : (Uptime Intelligence, 2024)

²⁴ Dans le cas ultime, le centre de données entier devra être remplacé, ce qui se traduira vraisemblablement par la construction d'un nouveau centre de données, l'existant étant conservé pour fournir des services de plus basse valeur ajoutée.

²⁵ Les méthodes utilisées par Uptime Intelligence sont des sondages par e-mail à destination de 850 propriétaires de centres de données et 750 fournisseurs et consultants, voir dernière figure de leur rapport pour plus d'informations.

Côté gains d'efficacité énergétique au niveau des centres de données, on constate une amélioration du PUE²⁶ avec une moyenne autour de 1.55²⁷ (Figure 13).

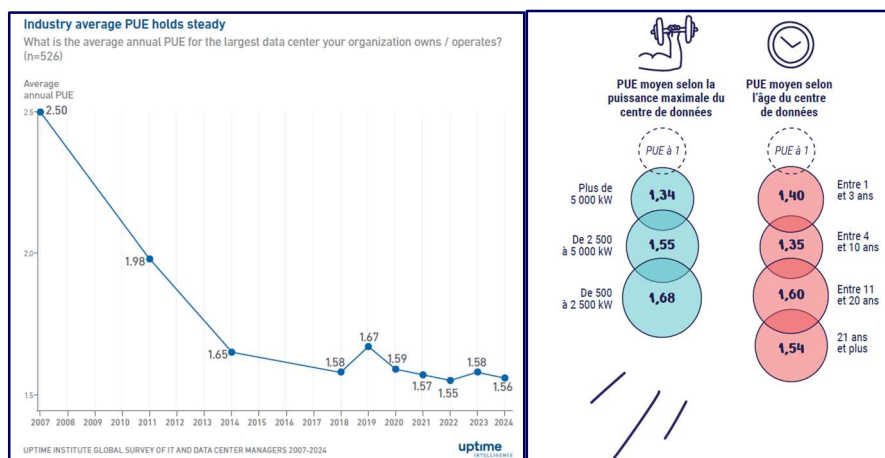


Figure 13 - Evolution du PUE. Sources : (Arcep, 2023; Uptime Intelligence, 2024)

Côté énergie, l'implantation de ces nouveaux centres de données impose de nouvelles consommations d'électricité. Mais aux Etats-Unis, c'est un mur énergétique aux conséquences carbonées qui se profile depuis début 2024 :

- **L'IA, le cloud computing et les crypto-monnaies, via les centres de données, exercent une contrainte forte et imprévue sur les réseaux de transport d'électricité** vieillissants, réseaux de transport qui sont déjà en croissance pour absorber électrification des usages et réindustrialisation (Washington Post, 2024).
- **Face à ces difficultés énergétiques, les grands acteurs cherchent de la production d'électricité bas-carbone** : Amazon a signé un accord avec Dominion pour explorer le déploiement de 0,3 GW de SMR (Data Center Dynamics, 2024d), de même pour Oracle qui cherche à équiper de 3 réacteurs nucléaires modulaires un futur centre de données, et Microsoft relance la centrale nucléaire précédemment mise à l'arrêt de Three Mile Island. Sam Altman, PDG d'Open AI, finance quant à lui une start-up de fusion nucléaire : Helion Energy (Le Monde & Alexandre Piquard, 2024).
- **Mais les besoins énergétiques pour l'IA sont immédiats. En attendant les SMR ou la fusion nucléaire, les acteurs se tournent vers des énergies fossiles et notamment le gaz.** C'est le cas du centre de données de xAI, Colossus, qui s'est équipé de 35 MW de générateurs mobiles au gaz naturel (Data Center Dynamics, 2024b) et de Meta qui annonce 3 centrales à turbines à gaz naturel pour un total de 2,3 GW (The Register & Brandon Vigliarolo, 2024).
- **Les réseaux d'électricité du sud-est des Etats-Unis prévoient l'ajout de 20 GW de centrales électriques au gaz naturel** (ce qui impliquerait 80 MtCO_{2e} annuelles) (Data

²⁶ Power Usage Effectiveness : indicateur d'efficacité énergétique utilisé pour qualifier l'efficacité énergétique d'un centre de données. C'est le ratio entre l'énergie totale consommée par l'ensemble du centre d'exploitation (avec entre autres, le refroidissement, le traitement d'air, les onduleurs...) et la partie qui est effectivement consommée par les systèmes informatiques que ce centre exploite (serveurs, stockage, réseau).

²⁷ Attention à ne pas extrapoler de futurs gains d'une telle courbe : (i) le PUE ne descendra pas en-dessous de 1, ce qui limite le gisement d'efficacité, (ii) concernant les PUE, il faut noter que les PUE en usage nominal sont très différents des PUE pendant la phase de remplissage du centre de données (nous en reparlerons plus bas, comme tout indicateur d'efficacité, il est dégradé quand le volume est faible).

Center Dynamics, 2025) et l'entreprise First Energy a **annulé la fermeture de deux centrales à charbon**, les maintenant jusqu'en 2035-2040 (Financial Times, 2024).

Pour éviter cette poursuite des énergies fossiles, les clefs sont dans les mains du secteur des technologies de l'information et de la communication.

Au final, l'étendue des effets de ces différentes dynamiques sera fonction de :

- **L'offre et la demande** : de l'ampleur du développement de l'offre IA et de l'adoption de celle-ci (inférences mais aussi volumes et intensités des entraînements), du taux de croissance et de l'intensité des usages numériques en général, de la disponibilité financière et des investissements, de la production de processeurs et des chaînes d'approvisionnement.
- **L'efficacité énergétique du système numérique** : du développement et de l'adoption de modèles plus efficaces voire frugaux²⁸, du développement et de l'adoption de processeurs plus efficaces et à l'empreinte carbone réduite et de la gestion des centres de données (refroidissement, réutilisation de chaleur fatale, parcs optimisés).
 - Mais aussi de la **vitesse de renouvellement pour les équipements numériques** dans les centres de données²⁹.
 - Mais aussi du développement de l'**edge**, à savoir de centres de données au plus proche des consommateurs³⁰.
- **Lien entre l'efficacité énergétique et l'offre** : l'efficacité est-elle au service d'une demande plus large ? En résultera-t-il un effet rebond ? (Luccioni A. S. et al., 2025) Dans la mesure où la puissance de calcul est rendue disponible, l'efficacité peut aboutir à une diffusion à très large échelle même sur des cas d'usages initialement jugés comme marginaux ou inutiles, ceux-ci ayant désormais la possibilité d'être déployés (effet d'offre).
- **L'intensité carbone de l'électricité** : la disponibilité de puissance électrique, notamment décarbonée.

Ces incertitudes sont importantes, sans compter celles imputables au fait que les données publiques sur ces sujets ne sont jamais complètes (peu de sources, hypothèses rarement explicitées, pas de détail sur les modélisations), et que les fabricants et fournisseurs ne sont pas transparents sur les données énergétiques et environnementales.

Néanmoins, même dans les scénarios les plus bas, **l'évolution du numérique et en particulier de l'IA est trop rapide en regard des délais structurels de transformation des infrastructures électriques et d'efficacité énergétique**. Cet écart de vitesse met en évidence le manque de responsabilité sur ces rythmes d'évolution et l'échec d'une politique de non-encadrement marquant un désalignement d'objectifs de décarbonation entre acteurs publics et acteurs privés.

²⁸ (AFNOR, 2024) : A noter que ce référentiel ne contient pas seulement des bonnes pratiques d'efficacité

²⁹ Comment le remplacement prématuré de matériel joue sur les empreintes énergétiques et climatiques du système numérique ? (Empreinte embarquée à la fabrication additionnelle, mais quelle seconde vie pour l'équipement en question ?)

³⁰ Comment l'edge joue sur les empreintes énergétiques et climatiques du système numérique ? (Allègement des réseaux, mais l'edge utilise généralement des infrastructures existantes, à l'efficacité énergétique moindre et de plus grande taille. Avec des latences plus faibles, car plus proches des lieux de production des données ou de consommation des services, il peut être utilisé pour l'inférence. Avec des capacités plus faibles, il peut être utilisé pour les réentraînements.

Dans les travaux qui feront suite à ce rapport intermédiaire, il s'agira d'échanger avec les parties prenantes du secteur pour soulever les questions qui seront à traiter par les acteurs, une fois un constat consolidé posé par nos travaux : en plus d'une nécessaire transparence, quelles options sont déjà explorées ou dorénavant à considérer pour permettre l'alignement des objectifs : mécanismes de notations ? Mécanismes et règles de marché, signal-prix, permis d'émissions ? Incitations ? Politiques de pénalités ? Comment et à partir de quels standards seraient conçus ces mécanismes afin d'agir au bon niveau et à la bonne échelle ? Toutes vos remarques, critiques, propositions ou pistes de recherche/entretiens sur les questions que se posent et devront se poser les parties prenantes seront donc un ingrédient clé pour les suites de ces travaux !

V. Quels impacts sur les terminaux et sur les réseaux ?

De manière générale, bien que l'IA vienne dans certains cas se substituer à d'autres solutions, son déploiement vient s'ajouter au système numérique déjà existant.

Décrire ses impacts sur les terminaux et réseaux peut se faire en adoptant une approche "conséquentielle". Plutôt que de tenter d'évaluer des impacts directs des usages IA au travers de la manière dont ils sollicitent les infrastructures et supports, ce type d'approche consiste à poser la question des conséquences de l'introduction des services : quels nouveaux dimensionnements et nouvelles caractéristiques techniques des systèmes (terminaux, réseaux et centres de données) vont être engendrés et motivés dans le cas d'un niveau de déploiement donné des services d'IA dans la société et l'économie ? Cette approche présente l'intérêt de questionner le sujet sous l'angle des déploiements d'application, services et usages (l'effet d'usage), tout en liant aux projections infrastructurelles qui sont censées les permettre (l'effet d'offre) (Figure 2 - Nos usages et nos réseaux sont les deux faces d'une même dynamique. Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport Figure 2).

De la même manière que sont étudiés, dans ces travaux, les effets du déploiement de l'IA sur la répartition de la capacité de calcul (edge, centralisée) ou sur le dimensionnement macroscopique des infrastructures centres de données, l'approche "conséquentielle" permet de faire émerger et d'éclairer les effets possibles sur les capacités des terminaux ou les spécifications des réseaux appelées par l'IA à moyen et long terme.

L'ACV conséquentielle

Une ACV conséquentielle (ACV-C) est une méthodologie codifiée dont l'objectif est de modéliser l'ensemble des impacts environnementaux consécutifs à un changement survenant dans le cycle de vie d'un produit (Dandres, 2012)³¹. L'objectif est donc sensiblement différent de celui d'une ACV attributionnelle (ACV-A) qui consiste à étudier le cycle de vie d'un produit. Alors que l'ACV-A est réalisée dans un état statique où le cycle de vie du produit n'évolue pas dans le temps, l'ACV-C repose sur l'évaluation des conséquences causées par le passage d'un état A à un état B du cycle de vie d'un produit. Il y a donc une notion temporelle à prendre en compte en ACV-C

³¹ https://publications.polymtl.ca/881/1/2012_ThomasDandres.pdf

correspondant à la période de temps nécessaire pour que le cycle de vie du produit passe de l'état A à l'état B.

La démarche de ce rapport, qui vise à éclairer des dynamiques macroscopiques des composantes du système numérique, **ne vise pas à produire une analyse de cycle de vie**, qui vise à documenter les impacts d'une unité fonctionnelle définie selon un cahier des charges précis. **Ce que les travaux ici présentés intègrent, c'est la dimension dynamique dans l'étude des conséquences des trajectoires choisies pour un système donné. Le terme « conséquentiel » ne sera donc pas utilisé au sens de l'ACV-C dans la suite de ce document.**

Quels sont les effets induits par l'IA sur les terminaux ?

Les centres de données ne sont pas les seules instances de stockage et de calcul que nous considérons dans nos analyses. Le déploiement massif de l'IA peut en effet avoir des conséquences sur les terminaux, qui sont liés au reste de l'écosystème numérique et à ses dynamiques par des phénomènes comme le déploiement de nouvelles architectures, l'appel de nouvelles capacités (traitement et stockage de la donnée) par des services émergents ou les effets d'obsolescence (logicielle, fonctionnelle, marketing etc.) associés.

Comprendre ces interactions permet d'identifier les effets d'offre et d'usage (Figure 2 - Nos usages et nos réseaux sont les deux faces d'une même dynamique. Source : Rapport intermédiaire, The Shift Project, dans le cadre de ce rapport Figure 2) affectant les terminaux, ainsi que les dynamiques qui en découlent :

- De manière générale, la prise en compte des nouveaux besoins en ressources numériques (mémoire et stockage, capacités de traitement) pour pouvoir rendre des services d'IA locaux (au niveau des terminaux) et déconnectés pourrait aboutir à l'augmentation de l'empreinte environnementale des terminaux, tant à la production qu'à l'usage. Le passage du concept d'IoT à celui de Smart IoT et de thin edge à thick edge par exemple, va engendrer une intensité énergie et carbone beaucoup plus importante dans la phase de fabrication des équipements, de par le recours à des modules de traitement et de connectivité plus puissants (Pirson T., Bol D., 2021).
-
- Les gains d'efficacité énergétique (qui ne concernent donc que la phase d'usage) au niveau des terminaux pourraient être contrebalancés par le renouvellement du parc dû à l'obsolescence (provoquée par divers facteurs : incompatibilité avec les nouvelles couches logicielles, déclenchement de l'acte d'achat d'un nouveau terminal simplement par attractivité des services offerts avec etc.) et les impacts engendrés par la production de ces nouveaux équipements.
- Par effet d'usage, l'ancrage de nouvelles applications IA dans des systèmes et habitudes d'usages au fil de leur adoption peut pousser au déploiement de nouveaux terminaux spécifiquement conçus pour l'IA. Cela contribuera non seulement directement à l'augmentation du parc (et aux effets environnementaux associés), mais également à la pénétration des services d'IA dans des usages (quotidiens, industriels, institutionnels, multiplication d'assistants et d'agents à fort degrés de personnalisation et de spécialisation, process d'automatisation de tâches complexes etc.) qui deviennent alors dépendants de ces nouvelles technologies des matériels et des services de haut niveau qui les

composent³².

- Les potentiels effets rebonds économiques pourraient alors stimuler l'offre de services d'IA, l'abaissement potentiel des coûts des nouveaux terminaux IA dans le cas d'un déploiement massif rendant ceux-ci plus abordables.

Afin d'évaluer et projeter les effets induits de l'IA sur les terminaux, la mesure (en FLOP, par exemple) du besoin en capacité de traitement requis par les différentes tâches appelées par les services pourrait permettre de :

- Quantifier et suivre les risques et phénomènes d'obsolescence qu'engendrerait le déploiement d'un service donné ;
- Quantifier et suivre l'évolution de l'intensité des services d'IA déployés ou projetés, en amont de leur adoption (une brique technologique requérant davantage de FLOP nécessitera par exemple soit un terminal plus puissant soit des temps de calcul plus longs) ;
- Mesurer les conséquences de l'effet de parc, c'est-à-dire de la stagnation de la puissance de calcul disponible dans un parc donné. Stagnation non alignée avec les besoins théoriques croissants en capacités dans le cas de déploiement de services toujours plus intenses, et due en particulier à l'allongement, souhaité ou subi, de la durée de vie de certains équipements³³. En effet, il n'est pas garanti qu'il y aura substitution totale des terminaux d'ancienne génération par les terminaux IA : des contradictions existent entre d'une part les objectifs de durabilité des terminaux et donc d'allongement de leur durée de vie et d'autre part les besoins croissants de puissance de calcul notamment sur les terminaux pour exécuter tout ou partie des services d'IA (déploiement d'assistants personnels dans des suites bureautiques par exemple).

Le déploiement à large échelle des terminaux conçus pour l'IA pourrait donc entraîner sur le long terme une intensification en ressources informatiques d'activités-clés ou quotidiennes de nos sociétés. L'augmentation des volumes de données, produits et traités par les terminaux, qui découlerait de cette dynamique affecterait à son tour les infrastructures réseaux, selon la même logique "effet d'offre - effet d'usage".

³² Notons qu'il n'est pas nécessairement de l'intérêt des fournisseurs de services et de matériels que les services d'IA locaux soient sobres et frugaux, i.e. qu'ils puissent s'exécuter localement avec des terminaux simples et courants, non équipés pour l'IA. En effet, prévus pour l'IA, les terminaux deviennent des inducteurs de la dynamique d'augmentation des besoins en services d'IA et de la possibilité de proposer davantage de ces services, une fois déployés les terminaux permettant de les exécuter).

³³ En 2021, l'ARCEP notait que la durée de vie des smartphones s'était allongée notablement entre 2013 et 2019. Plus récemment, le développement du reconditionné et les politiques RSE volontaristes des entreprises laissent penser que cette tendance se confirme (ARCEP, 2021 : <https://www.arcep.fr/actualites/actualites-et-communications/detail/n/environnement-120721.html>).

Quels sont les effets induits par l'IA sur les réseaux ?

Au niveau mondial, le trafic sur les réseaux croît, à la fois chez les consommateurs et les entreprises, essentiellement porté par le streaming vidéo. Par exemple, sur les réseaux mobiles, le trafic a augmenté de 30% par an de 2020 à 2024³⁴ passant de 55 Eo à 157 Eo par mois (Ericsson, 2025).

A l'avenir, certaines projections estiment par exemple, pour 2033, 3 scénarios (moderate, aggressive, disruptive) multipliant le trafic mondial entre 5 et 9³⁵ (Nokia, 2024). Cette évolution se justifierait avec la décomposition d'usages suivante : la vidéo et les réseaux sociaux restent prédominants pour les usages mobiles, tandis qu'applications d'entreprises, XR et IOT se répartissent le trafic d'entreprises qui serait multiplié par un facteur 8 à 21.

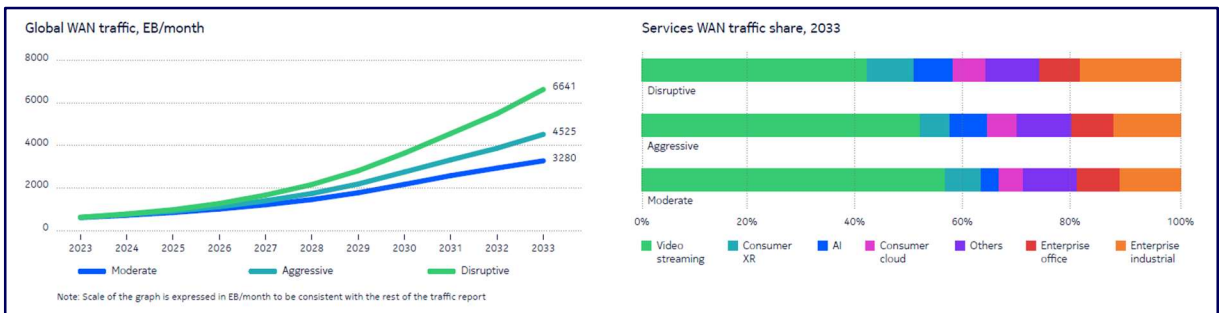


Figure 14 - Projections de trafic internet et d'usages à 2033. Source : (Nokia, 2024)

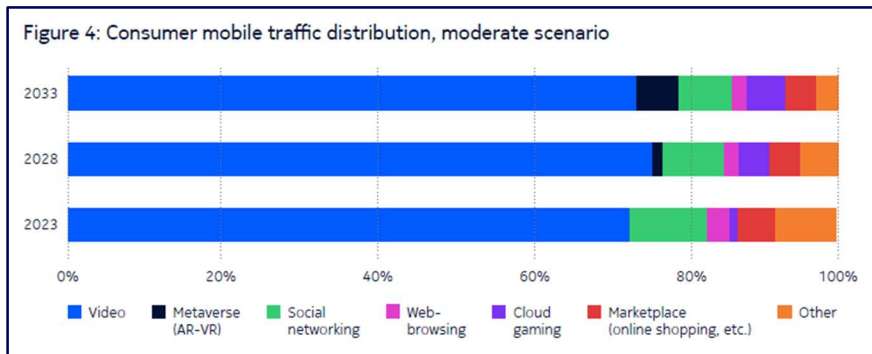


Figure 15 - Projections de trafic internet et d'usages à 2033. Source : (Nokia, 2024)

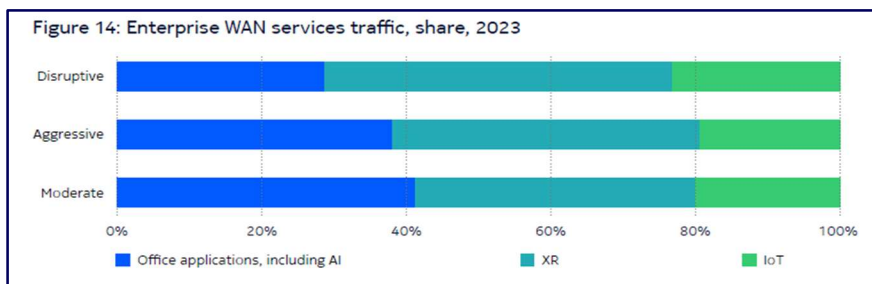


Figure 16 - Projections de trafic internet et d'usages à 2033. Source : (Nokia, 2024)

³⁴ T3 à T3

³⁵ En France aussi, la FFT prévoit une croissance du trafic IP d'un facteur 6 de 88 Eo en 2020 vers 526 Eo en 2030, portée par l'augmentation des usages du numérique, l'apparition de nouvelles technologies et la croissance des équipements utilisateurs (FFT, 2024).

Pour décrire la place que l'IA occuperait dans ce trafic, ces projections le décompose en :

- Trafic IA direct : le trafic avec des applications IA connues comme telles ou intégrées dans des applications de productivité d'entreprises,
- Trafic IA indirect : la croissance du trafic créée par **l'engagement des consommateurs via les algorithmes de recommandations et la personnalisation des contenus**.

L'approche est semblable à d'autres analyses (Ericsson, 2024), qui identifient comme principal facteur d'augmentation **la création de contenu hyper-personnalisé à grande échelle**, qui pourrait favoriser une croissance du trafic mobile. L'utilisation accrue d'assistants vidéo et d'interactions immersives basés sur la GenAI pourrait augmenter à la fois le **trafic montant et descendant** (Ericsson, 2024).

Si l'étude (Ericsson, 2024) ne communique pas pour l'instant de prévisions chiffrées relatives au trafic supplémentaire issu de l'IA, l'analyse des trajectoires produites dans l'étude prospective (Nokia, 2024) montre que le trafic (total, fixe et mobile) pourrait être en 2033 environ 50 % plus élevé qu'il ne le serait en l'absence d'IA.

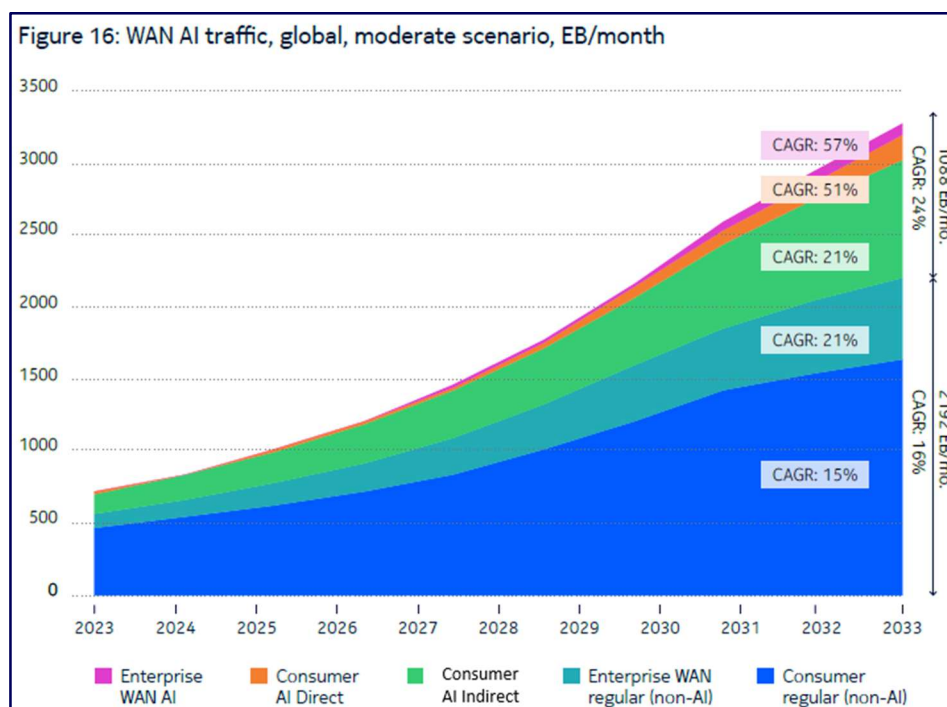


Figure 17 - Projection de trafic internet à 2033, décomposition entre consommateur et entreprise, et croissance générée par l'IA. Source : (Nokia, 2024)

Ces premières estimations doivent malgré tout être nuancées :

- Elles sont produites par des organisations étant parties prenantes de la chaîne de valeur, pour qui l'anticipation de forte croissance de trafic est liée à des enjeux commerciaux ;
- Il faut aussi tenir compte des adaptations réseaux possibles pour répondre à cette croissance du trafic. Cela inclut des technologies de compression plus efficaces, la gestion des congestions, des architectures dédiées à l'IA.

Au-delà de ces considérations centrées sur la croissance en volume du trafic, deux aspects semblent potentiellement impactant :

- La multiplication de cas d'usage ayant des contraintes de performance de type "temps réel" obligerait à réduire la latence effective de la chaîne technique utilisée, conduisant éventuellement à une densification accrue des points de présence mais aussi à une recherche d'optimisation au sein du continuum cloud-réseau, tout ceci se traduisant notamment par le développement de l'edge computing ;
- La capacité donnée aux utilisateurs d'IA de générer de la vidéo ou de faire interpréter des images fixes ou animées augmenterait fortement le trafic montant (upstream) alors que l'architecture actuelle des réseaux est conçue pour écouler le plus efficacement possible du trafic descendant (downstream).

Nous n'avons pas pu à ce stade quantifier ces conséquences qui pourraient cependant s'avérer significatives.

L'analyse que nous avons menée à ce stade nous pousse à identifier les questions suivantes pour des travaux futurs, suite à ce rapport intermédiaire :

- Aujourd'hui, le trafic lié à l'IA paraît mineur et n'est pas mesuré. Comment pourrait-on publiquement observer cette évolution ?
- L'intelligence artificielle peut-elle autant multiplier la création de contenus personnalisés et être à l'origine d'une hausse de trafic possible ?
- Quels sont les effets indirects des algorithmes de recommandation (et donc de l'IA) sur les réseaux ?
- L'IA et le numérique dans des cas d'application très décentralisées (robotique agricole, mobilité autonome) pourrait (ou pas) induire un déploiement des réseaux amplifié (bande-passante additionnelle, équipements plus performants pour réduire les latences) ?

Toutes vos remarques, critiques, propositions sur les données et directions explorées à ce stade, ainsi que toute piste de recherche de données ou entretien à mener seront donc un ingrédient clé pour les suites de ces travaux !

Capacités de calcul en France et en Europe : dimensionnements possibles, trajectoires d'émissions de gaz à effets de serre, quelles gouvernances pour être à la hauteur des enjeux

Si le chapitre précédent décrit des dynamiques mondiales et vise à éclairer leurs conséquences climatiques, le moteur des dynamiques mondiales actuelles n'est pas (en grande partie) entre nos mains. Cette partie-ci ambitionne d'instruire les questions que ces dynamiques impliquent en France et en Europe et les gammes de choix dont nous disposons, sur les volets climatiques, électriques comme sur les implications territoriales.

En plus du mur énergétique américain (voir partie précédente), le cas de l'Irlande par lequel nous commençons cette partie sera univoque. Se rendre compte de la situation actuelle en France sera un prérequis : quels sont les besoins énergétiques du numérique mais aussi des autres secteurs ?

L'ambition de ce rapport intermédiaire sera aussi d'échanger sur les besoins en modélisation au niveau français.

Toutes vos remarques, critiques, propositions sur les données et directions explorées à ce stade, ainsi que toute piste de recherche de données ou entretien à mener seront donc un ingrédient clé pour les suites de ces travaux !

I. Le cas de l'Irlande

L'Irlande est un pays qui s'est positionné comme l'un des leaders européens du secteur numérique avec des politiques très volontaristes (notamment fiscales) pour attirer les

investissements. Des problématiques majeures ont émergé ces dernières années concernant le dimensionnement des infrastructures électriques, les conflits d'usage entre les différents secteurs d'activités ainsi que des pénuries d'énergie renouvelable pourtant nécessaire pour soutenir la croissance du déploiement des centres de données tout en restant en maîtrise de la trajectoire carbone du pays.

Si l'on exclut les centres de données, la demande en électricité en Irlande est restée relativement stable ces dernières années. Cependant, en intégrant les centres de données, la demande a augmenté de 24,7 % entre 2012 et 2022 (en rouge sur la figure ci-dessous) (Prof. Hannah Daly, University College Cork, 2024).

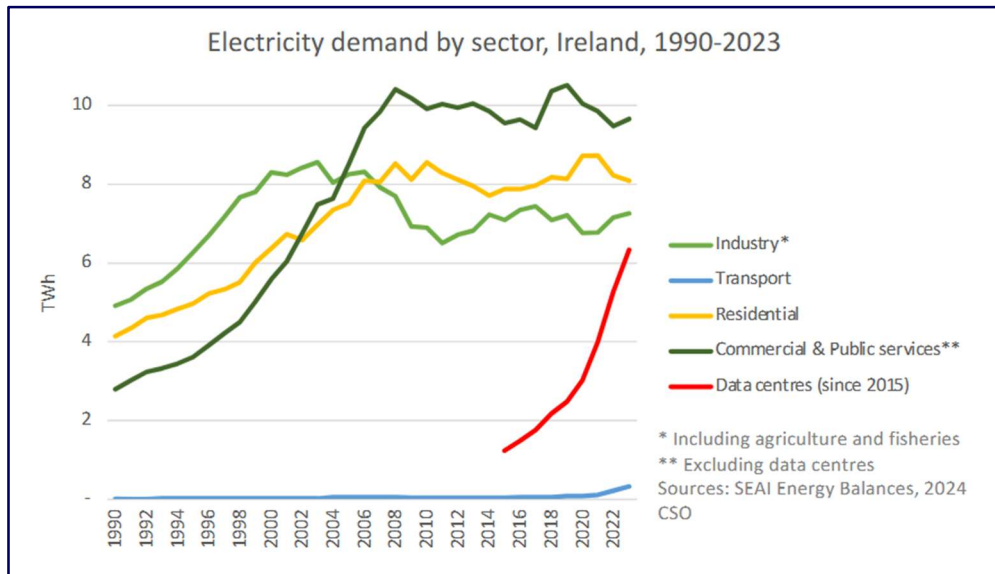


Figure 18 - Demande d'électricité par secteur entre 1990 et 2023 en Irlande. Source : (Prof. Hannah Daly, University College Cork, 2024)

Selon les statistiques officielles irlandaises, les centres de données consomment déjà plus de 20 % de l'électricité disponible, dépassant la consommation électrique des zones résidentielles urbaines (Central Statistics Office (CSO), 2024; The Journal, 2024).

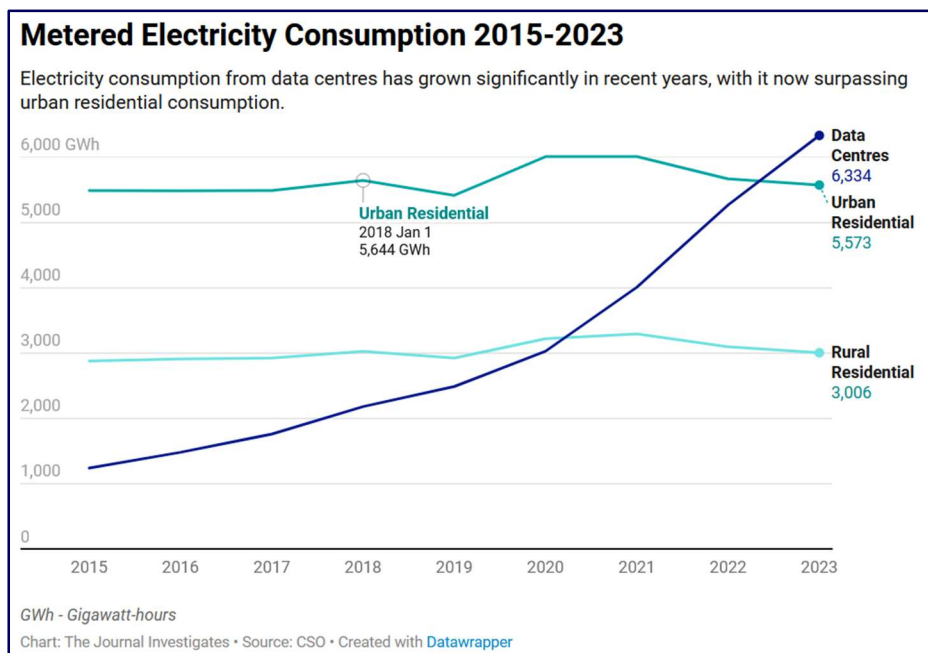


Figure 19 - Consommation électrique des centres de données en Irlande, dépassant la consommation urbaine résidentielle. Source : (Central Statistics Office (CSO), 2024; The Journal, 2024)

Certaines projections prévoient que le secteur atteindra autour de 30 % de la consommation globale du pays en 2028, et le secteur pourrait même être, sous certains scénarios, le plus consommateur en électricité, dépassant les secteurs industriels et des services d'ici quelques années (IEA, 2024a; Prof. Hannah Daly, University College Cork, 2024).

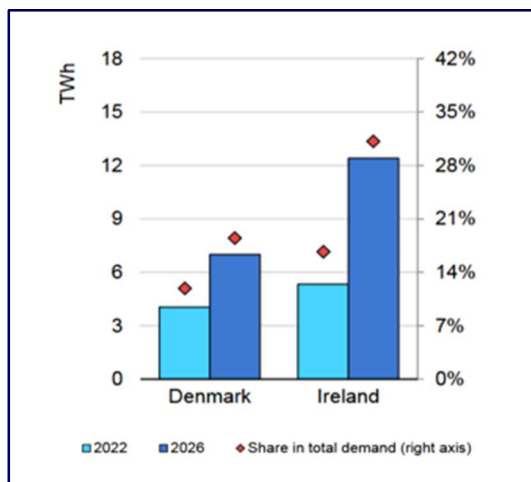


Figure 20 - Estimation de la consommation d'électricité des centres de données et de sa part dans la demande totale d'électricité en Irlande. Source : (IEA, 2024a)

Le secteur des énergies renouvelables du pays, dont le développement a été planifié pour assurer la migration du mix énergétique global dans un contexte de croissance modérée de la demande, se montre désormais incapable de pouvoir satisfaire les nouvelles demandes. Pour assurer leur développement, les centres de données se tournent vers des alimentations au gaz naturel notamment, et multiplient l'utilisation en local de générateurs de secours et d'urgence alimentés aux combustibles fossiles. En 2023, l'Environmental Protection Agency (EPA) irlandaise, responsable de la délivrance des permis de licences industrielles pour les installations dont la capacité dépasse 50 MW, a délivré 13 licences à des centres de données pour l'exploitation de

générateurs électriques. Ce chiffre, qui ne prend même pas en compte les petites installations ne nécessitant pas de licence industrielle, soulève de sérieuses préoccupations et témoigne d'une dépendance accrue aux combustibles fossiles.

Ces tensions d'accès à l'énergie constituent un goulet d'étranglement pour le développement de nouveaux projets et incitent les autorités à renforcer les critères d'octroi des autorisations pour de nouveaux centres de données. En 2021, EirGrid, l'opérateur national du réseau électrique, a ainsi imposé un moratoire pour la région de Dublin, où se concentre la majorité des centres de données, sur les nouvelles demandes d'implantation jusqu'en 2028. En 2024, le conseil du district de Dublin Sud met en avant « la capacité insuffisante du réseau électrique et l'absence d'énergie renouvelable significative sur le site pour alimenter les datacenter » refusant ainsi un projet sur son territoire par un acteur majeur du domaine.

En complément de l'exemple américain détaillé dans la partie précédente, le cas de l'Irlande illustre que des politiques favorisant à tout va l'implantation de centres de données peuvent s'avérer risquées et contre productives au développement du secteur, en particulier en l'absence de stratégies long-terme suffisantes sur les systèmes énergétiques, mais aussi nuisibles quand le recours aux énergies fossiles se fait. Le cabinet d'intelligence économique Gartner et le site Epoch AI identifiant même l'énergie comme facteur limitant des centres de données, à 2027 pour l'un, avec plus de 40% des centres de données d'IA existants d'ici 2027 qui pourraient rencontrer des contraintes opérationnelles en raison de pénuries d'énergie, à 2030 pour l'autre (Epoch AI, s. d.-a; Gartner, 2024).

II. L'état des lieux en France

A. Numérique & climat

En France, l'empreinte carbone du numérique pour l'année 2020 a été estimée à 17 MtCO₂e sans inclure les émissions importées des centres de données (ADEME & Arcep, 2023). Une double mise à jour publiée en janvier 2025 a révisé cette évaluation en portant sur l'année 2022 et en intégrant les émissions importées des centres de données, ce qui amène ainsi l'empreinte carbone du secteur numérique à 29.5 MtCO₂e (ADEME, 2025).

Pour 2022, l'empreinte carbone du secteur numérique représente 4,4 % de l'empreinte carbone totale de la France, dont 46 % attribuables aux centres de données, en tenant compte des émissions importées (ADEME, 2025). Toutefois, comme l'indique d'ailleurs l'ADEME, ce chiffre de 4,4 % est probablement sous-estimé, ne prenant pas en compte l'arrivée récente de l'intelligence artificielle générative ni toutes les évolutions du parc de terminaux.

Cette mise à jour soulève des questions sur la fréquence de mesure de l'empreinte énergétique et climatique pour suivre la rapidité des évolutions dans le secteur numérique. Cette rapidité des évolutions amplifie également l'importance de la transparence sur les données carbone, car perdre des mois ou années de recherche pour faire la rétro-ingénierie de données non fournies ne permet pas d'avoir en temps réel les bonnes informations pour prendre des décisions éclairées.

B. Numérique & électricité

En France, la consommation électrique du secteur numérique est de 51,5 TWh en 2022 soit 11 % de la consommation électrique française (ADEME, 2025).

Pour les centres de données, depuis 5 ans, les projections ne cessent de s'accroître pour la consommation d'électricité territoriale :

- En 2022, RTE, dans son étude « Futurs énergétiques 2050 », déclare 3 TWh de consommation électrique pour 2019 et prévoit 5 TWh pour 2030 dans la catégorie "secteur tertiaire" (RTE, 2022).
- En 2023, l'ADEME-Arcep, dans son étude « Evaluation de l'impact environnemental du numérique en France et analyse prospective », modélise 11,6 TWh de consommation électrique pour 2020 prenant en compte plusieurs types de centres de données : colocation, HPC³⁶, traditionnels pour acteurs publics et entreprises (ADEME & Arcep, 2023). Au passage, en 2023 le cabinet Hubblo modélise 11,3 TWh pour 2020 de consommation cloud à l'étranger (Hubblo & Fourboul E., 2023).
- En mars 2023, l'ADEME-Arcep projette dans le scénario tendanciel 16 TWh pour 2030 et 39 TWh pour 2050 (ADEME & Arcep, 2023).
- En décembre 2023, RTE, dans son bilan prévisionnel pour 2035, prévoit entre 23 et 28 TWh pour 2035, toujours dans la catégorie "secteur tertiaire" (RTE, 2023) (chapitre 2, p. 99, dont l'annexe 2A sur les hypothèses est attendue), et réévalue à 10 TWh sa vision pour le début de la décennie.
- En 2024, RTE, dans son schéma de développement du réseau, étudie 8 GW de demande pour les centres de données³⁷. Ces 8 GW sont traduits en 25,3 TWh en Ile-de-France et 3,8 TWh à Marseille à 2040 (RTE, 2024) (document B). Sous l'hypothèse d'un taux d'utilisation de 60 %, cette demande de 8 GW correspondrait à 42 TWh.
- En 2025, RTE agrège cette fois l'industrie et les centres de données (ne fournissant plus le détail pour les centres de données donc, ni pour l'industrie d'ailleurs) et projette 21 GW de contrats de raccordement signés (RTE, 2025) (Synthèse 2025, en attente de la fiche n°5).

L'incertitude sur l'état actuel de la consommation énergétique des centres de données est déjà plus qu'importante, d'un facteur de 1 à 4 : entre 3 TWh (RTE, 2022) et 12 TWh (ADEME & Arcep, 2023). Sur la base de ces deux mêmes études, à 2030, l'écart en 2030 est d'un facteur 1 à 3 : entre 5 TWh (RTE, 2022) et 16 TWh (ADEME & Arcep, 2023). Une part de cette incertitude sur l'état des lieux peut résider dans les périmètres considérés pour ces données³⁸, ce qui représente un axe d'amélioration pour les futures publications sur ce sujet, propice à une meilleure organisation de la transition énergétique.

³⁶ Calcul haute performance

³⁷ En septembre 2024, RTE disait avoir signé pour 4,5 GW d'offres de raccordement de centres de données et « un volume équivalent » était en cours d'instruction (LeMagIT & Raoul G., 2025).

En février 2025, dans ses 140 projets signés pour 21 GW, RTE inventorie 40 projets de datacenters, pour une moyenne de 130 MW chacun, soit 5.2 GW (LeMagIT & Raoul G., 2025).

³⁸ Secteur tertiaire côté RTE, ou puissance/voltage suffisamment important pour être considéré par RTE, ou catégorie de centres de données.

Il est impératif de progresser rapidement sur le partage et la qualité des données ainsi que sur la modélisation de la demande : une incertitude aussi forte sur une demande d'ores et déjà significative et de toute façon en forte croissance crée un risque certain à l'horizon 2033-2035.

D'autant plus que, l'installation de centres de données sur le territoire français est encouragée via :

- Un allègement significatif de taxes sur la consommation d'électricité (TICFE) pour les opérateurs de centres de données (Legifrance, 2022)³⁹,
- Une politique incitative avec le projet de loi simplification qui vise à faire des centres de données des Projet d'Intérêt National Majeur (PINM) ce qui permettra d'accélérer certaines procédures (compatibilité des documents d'urbanisme, raccordement au réseau électrique, reconnaissance de raisons impératives d'intérêt public majeur - RIIPM) (Vie publique, 2024)⁴⁰,
- Une politique stratégique et commerciale nationale à la fois menée par :
 - EDF proposant 4 sites sur son foncier pour 2 GW et 2 nouveaux sites à venir (DCMag, 2025b; Le Figaro, 2025),
 - L'Elysée et rendue visible lors du Sommet pour l'action sur l'IA avec les annonces de 3 sites de 1 GW (Le Monde, 2025a, 2025b; L'usine digitale, 2025),
 - Le gouvernement annonçant 35 sites prêts à l'emploi pour une surface totale de l'ordre de 1 200 hectares (DCMag, 2025a; Le Monde, 2025a) (carte à l'échelle régionale administrative sur la Figure 21), avec 15 des 35 sites en mesure d'être raccordés au réseau à haute tension et pouvant atteindre 750 mégawatts de puissance (LeMagIT & Raoul G., 2025)⁴¹.

³⁹ Relève d'un tarif réduit de l'accise, pour la fraction qui excède un gigawattheure sur une année civile, l'électricité consommée pour les besoins de l'infrastructure immobilière qui répond aux conditions cumulatives suivantes :

1° Elle est consacrée au stockage physique, au traitement, au transport et à la diffusion de données numériques ;

2° Son accès est sécurisé ;

3° Elle comprend des dispositifs spécifiques et dédiés de contrôle de son environnement thermique, de la qualité de son air, de son alimentation en énergie et de prévention des incendies ;

4° Elle intègre un système de management de l'énergie conforme aux critères prévus au second alinéa de l'article [L. 233-2](#) du code de l'énergie ;

5° L'exploitant adhère à un programme, reconnu par une autorité publique, nationale ou internationale, de mutualisation des bonnes pratiques de gestion énergétique des centres de données incluant : a) L'écoconception des centres de stockage de données ; b) L'optimisation de l'efficacité énergétique ; c) Le suivi de la consommation énergétique et la réalisation de comptes rendus périodiques y afférents ; d) La mise en œuvre de technologies de refroidissement qui répondent à des critères de performance.

6° La chaleur fatale qu'elle génère est valorisée au sein d'un réseau de chaleur ou de froid ou l'installation respecte un indicateur chiffré sur un horizon pluriannuel en matière d'efficacité dans l'utilisation de la puissance, déterminé par décret ;

7° L'eau qui y est utilisée à des fins de refroidissement est limitée selon un indicateur chiffré sur un horizon pluriannuel, déterminé par décret ;

8° Le niveau d'électro-intensité, apprécié à l'échelle de cette installation, est au moins égal à 2,25 %.

⁴⁰ Pour les centres de données de dimension industrielle, le projet de loi prévoit de les qualifier de projets d'intérêt national majeur (PINM), ce qui permettra d'accélérer certaines procédures (compatibilité des documents d'urbanisme, raccordement au réseau électrique, reconnaissance de raisons impératives d'intérêt public majeur - RIIPM).

⁴¹ RTE liste environ 50 postes sources capables d'accueillir 250 MW et une vingtaine capable de supporter 750 MW, pas seulement dédiés aux centres de données (LeMagIT & Raoul G., 2025).

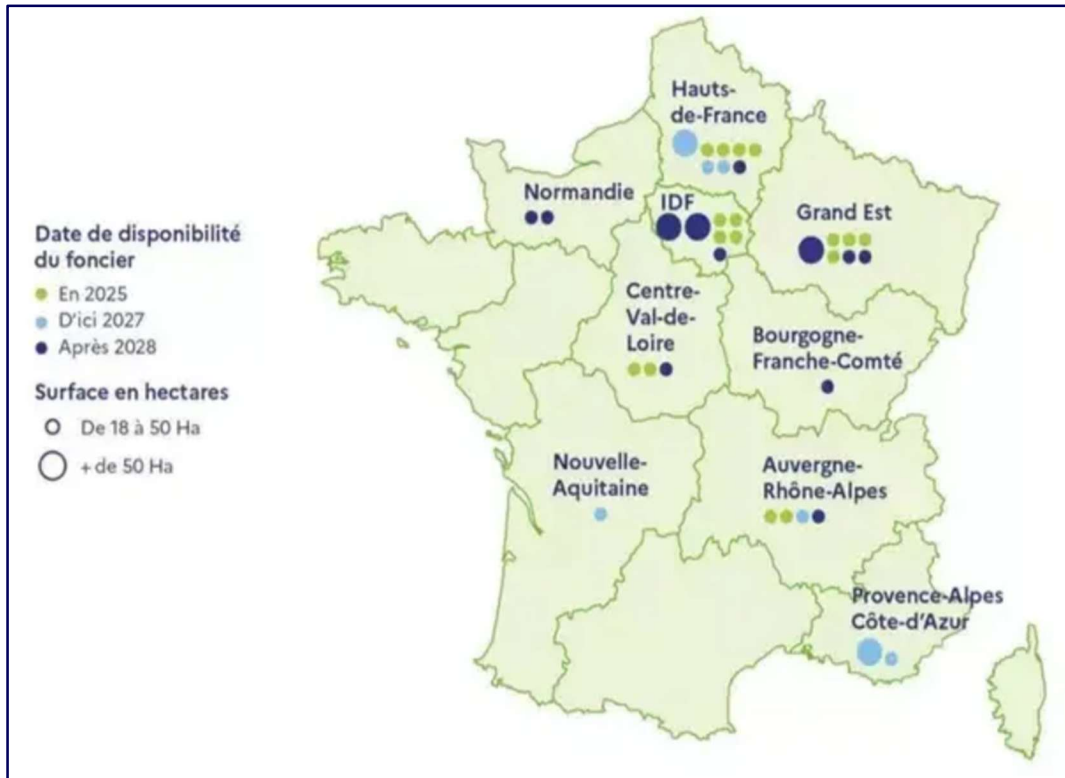


Figure 21 - La carte régionale des 35 sites de centres de données dédiés à l'IA identifiés par le gouvernement.
Source : (DCMag, 2025a)

La consommation électrique du secteur numérique est un sujet bien identifié depuis au moins 5 ans (Sénat, 2020). Toutefois, sur la connaissance et le suivi du parc de centres de données, il existe encore beaucoup (trop) de questionnements. Par exemple :

- Quels sont les inventaires publics disponibles ? Sur les projets en cours, à venir ? Quelles sont leurs fréquences d'actualisation ?
- Quels observatoires existent ? Quels sont les besoins d'anticipation ?
- Quels sont les canaux de débat public ?
- Comment s'effectue le suivi de l'évolution des centres de données plus diffus : centres de données d'entreprise et centres de données de petite taille ?
- Certains centres de données concentrent sur de petites surfaces de fortes capacités : cela peut-il constituer un risque pour la gestion de l'équilibrage du réseau ?

Ces questions semblent essentielles face à cette prévision d'intégration territoriale de centres de données, et dans un contexte de transition énergétique et écologique, et d'autant plus que les projections sont sans cesse revues à la hausse.

C. Numérique & électrification des autres secteurs

1. La transition énergétique nécessite une électrification des activités et une maîtrise de la demande

La transition énergétique envisagée par les scénarios prospectifs français (RTE, ADEME notamment) repose sur :

- **L'électrification des secteurs économiques majeurs** : tous les scénarios prévoient une électrification importante des équipements et procédés, avec une intensité variable selon les trajectoires, notamment dans le bâtiment, la mobilité ou les activités industrielles. L'énergie électrique y devient majoritaire. Ainsi, le besoin de production d'électricité est en augmentation dans la quasi-totalité des scénarios, mais à des degrés divers. Par exemple, dans le scénario de référence RTE (RTE, 2022), la consommation d'électricité totale à 454 TWh en 2020, devrait atteindre 508 TWh en 2030, 567 TWh en 2040 et 645 TWh en 2050.
- **La maîtrise de la demande** : la réduction de la consommation énergétique est indispensable à l'atteinte de la neutralité carbone, qui ne peut être réalisée uniquement par la décarbonation de la production énergétique. Cette diminution repose sur deux piliers : la sobriété et l'efficacité énergétique.

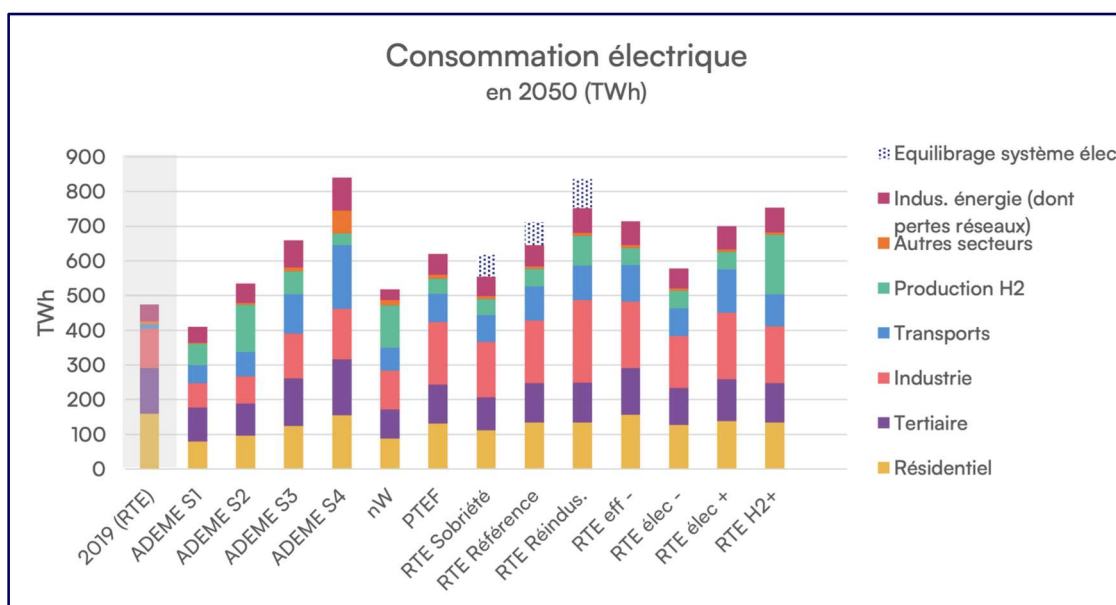


Figure 22 - Consommation d'électricité en 2050 selon les principaux scénarios de transition français

Source : Site [comprendre2050.fr](https://www.comprendre2050.fr), décryptage "Quelle évolution de la consommation électrique pour une France bas carbone ?" (ADEME, négaWatt, The Shift Project, 2025)

2. Analyse sectorielle des besoins en électricité à horizon 2030-2035

Au niveau sectoriel, l'exemple des besoins en électricité à horizon 2030 et 2050 de plusieurs secteurs (transport, industrie, résidentiel, hydrogène bas-carbone) permettra de donner une représentation de la transition, dans son ensemble, afin d'illustrer les enjeux de concurrence d'usage de l'électricité entre les secteurs, et d'illustrer les efforts anticipés sur la demande ; illustrations qui pourront être mises en perspective de la transition numérique.

Pour le **secteur des transports**, la demande en électricité croît selon tous les scénarios, à 2030 et à 2050.

D'ici à 2030, la consommation électrique des transports devrait augmenter significativement, passant de 12 TWh en 2020 à des niveaux atteignant jusqu'à 57 TWh dans certains scénarios. Pour les scénarios ayant des hypothèses intermédiaires (voir Figure 24) :

- **Scénario tendanciel de l'ADEME** (ADEME, 2021) : 35 TWh en 2030 (+23 TWh, soit une multiplication par 3 environ).
- **Scénario de référence de RTE** (RTE, 2022) : 33 TWh en 2030 (+21 TWh, soit une multiplication par un peu moins de 3).

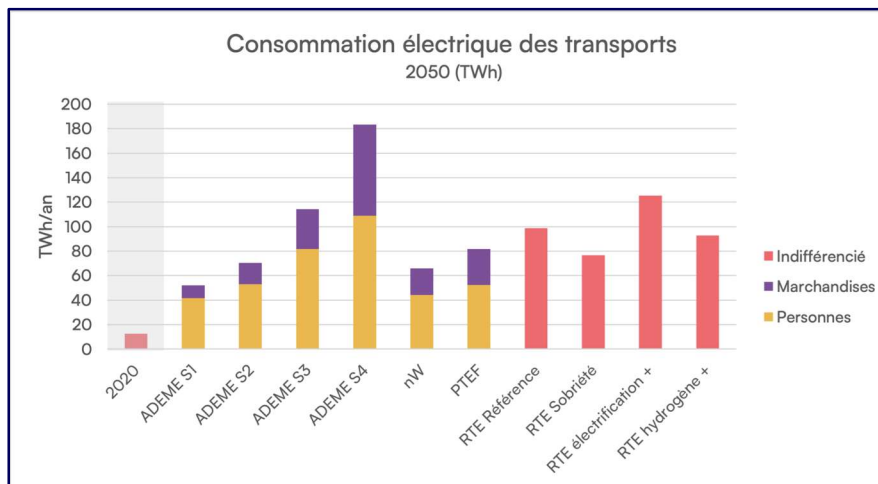


Figure 23 - Consommation d'électricité dans le secteur des transports en 2050. Source : Site comprendre2050.fr, décryptage « Quelle consommation d'électricité dans les transports de demain ? » (ADEME, négaWatt, The Shift Project, 2025)

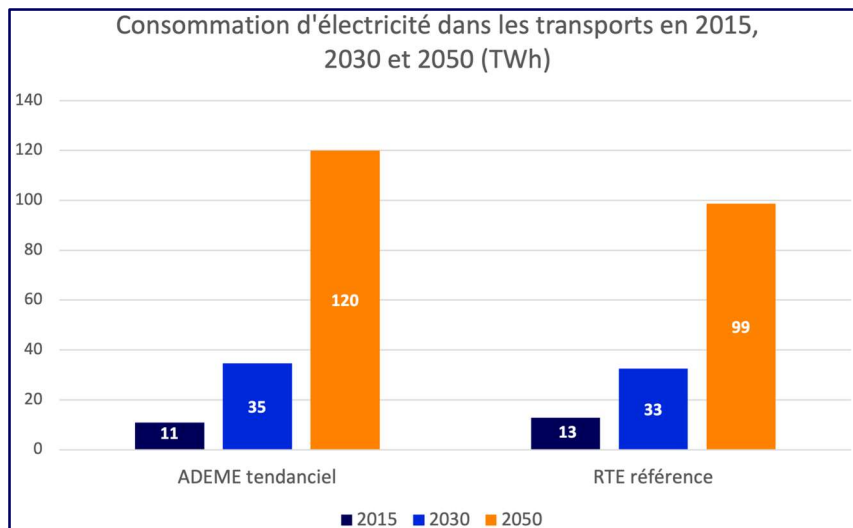


Figure 24 - Consommation d'électricité dans le secteur des transports en 2015, 2030 et 2050. Source : (ADEME, 2021; RTE, 2022)

Pour le **secteur de l'industrie**, la demande en électricité croît selon la plupart des scénarios.

La consommation électrique de l'industrie passe de 115 TWh en 2020 à des niveaux atteignant jusqu'à 123 TWh dans certains scénarios. Pour les scénarios ayant des hypothèses intermédiaires (voir Figure 26) :

- **Scénario tendanciel de l'ADEME** (ADEME, 2021) : 127 TWh en 2030 (+12 TWh).
- **Scénario de référence de RTE** (RTE, 2022) : 113 TWh en 2030 (-2 TWh).

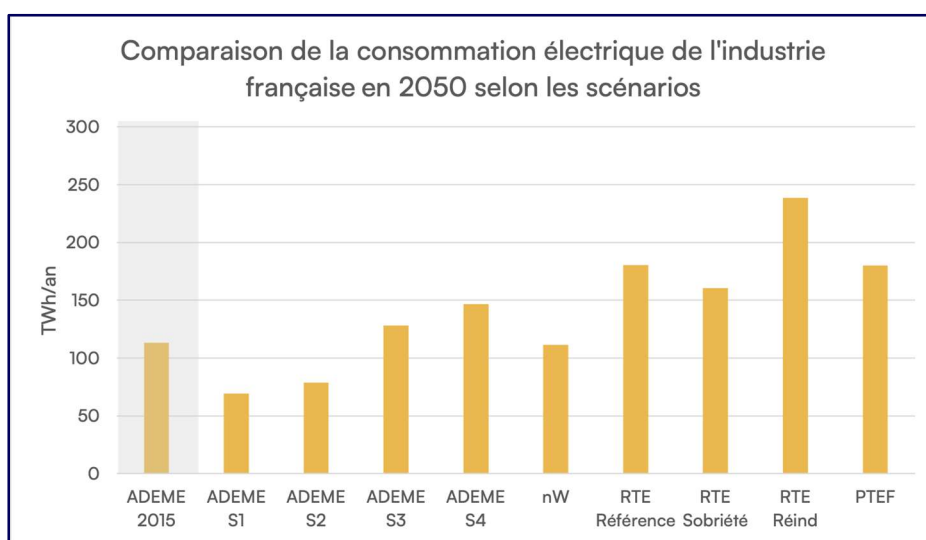


Figure 25 - Consommation d'électricité dans le secteur des transports en 2050. Source : Site comprendre2050.fr, décryptage « Quelle consommation d'électricité des industries en 2050 ? » (ADEME, négaWatt, The Shift Project, 2025)

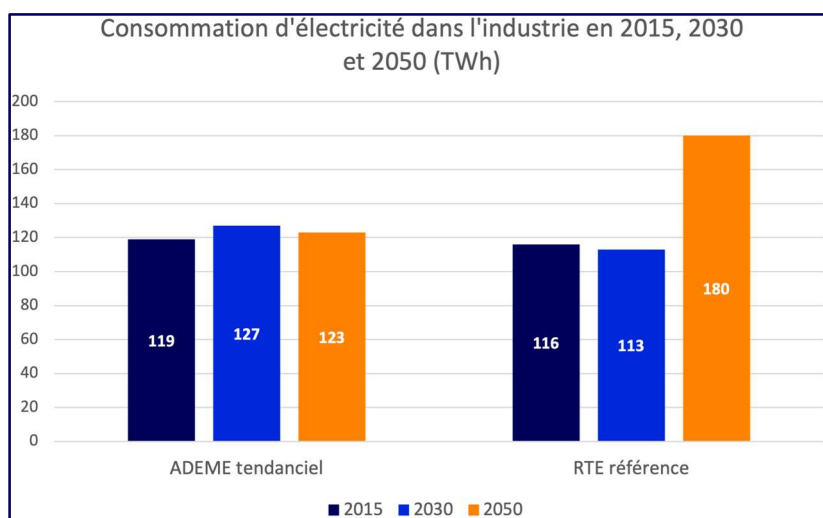


Figure 26 - Consommation d'électricité dans le secteur de l'industrie en 2015, 2030 et 2050. Source : RTE Futurs énergétiques, ADEME Transition(s) 2050

Pour le **secteur résidentiel**, la consommation électrique dans la trajectoire de référence de RTE est orientée sensiblement à la baisse sous l'effet prégnant de l'amélioration de l'efficacité énergétique :

- A l'horizon 2030 : de 159 TWh en 2020 à 151 TWh en 2030, soit –6 TWh par rapport à 2020
- A l'horizon 2050 : de 159 TWh en 2020 à 134,1 TWh en 2030, soit –25 TWh par rapport à 2020

Ainsi, la forte électrification du chauffage (70% de logements chauffés électriquement en 2050 contre 40% environ aujourd'hui) doit être plus que compensée par les effets baissiers de la performance des équipements, d'une rénovation importante des logements existants et de la nouvelle réglementation environnementale pour la construction neuve. La consommation d'électricité pour l'usage du chauffage baisserait ainsi de 11,5 TWh soit 25% en 2050 par rapport à 2019.

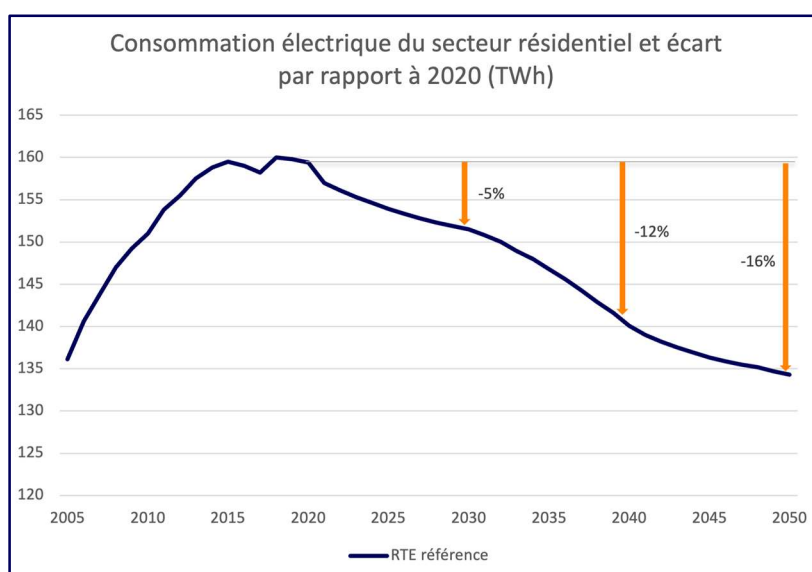


Figure 27 - Consommation d'électricité dans le secteur résidentiel entre 2005 et 2050 dans le scénario de référence de RTE. Source : (RTE, 2022)

Concernant le **développement de l'hydrogène bas-carbone**, c'est un axe majeur de la stratégie énergétique française. D'ici à 2050, son usage pourrait être multiplié par un facteur allant de 2 à 7 selon les scénarios. L'électrolyse à partir d'électricité décarbonée deviendra le mode de production central.

D'ici à 2030, la consommation électrique de ce secteur passe de 0 TWh en 2020 à des niveaux atteignant jusqu'à 86 TWh dans certains scénarios. Pour les scénarios ayant des hypothèses intermédiaires :

- **Scénario tendanciel de l'ADEME** (ADEME, 2021) : 22,3 TWh en 2030.
- **Scénario de référence de RTE** (RTE, 2022) : 25,1 TWh en 2030.

Aujourd'hui, RTE assure un accès transparent et non-discriminatoire aux utilisateurs du réseau. L'arbitrage quant à l'utilisation de l'électricité par les différents secteurs ne lui revient pas : une planification nationale via la feuille de route de la transition énergétique française, la SNBC, est donc essentielle. Cette planification nécessite de prendre en compte les besoins et caractéristiques de chaque secteur en termes d'électrification :

- En termes de consommation électrique (TWh), de capacités demandées (GW),
- Mais aussi en échelles de temps : par exemple, la constitution d'une filière d'hydrogène bas-carbone visée par les politiques publiques est sur un horizon de 10-15 ans,
- Et considérant les enjeux pour réaliser les transformations : déploiement des capacités technologiques, moyens financiers nécessaires etc.

3. La place des centres de données et du secteur numérique dans la transition énergétique est d'autant plus incertaine

Les estimations de la consommation actuelle du secteur sont incertaines (voir paragraphe « Numérique & électricité » Numérique & électricité) ce qui rend nécessairement la prospective à 2030 et à 2050 d'autant plus délicate.

En 2022, RTE retenait l'hypothèse d'une hausse de la demande électrique des centres de données mais basé sur un faible volume. Ainsi, dans la trajectoire de référence (en bleu foncé sur la **Erreur ! Source du renvoi introuvable.**), RTE anticipe un triplement de la consommation des centres de données passant de 3 TWh en 2019 à 5 TWh en 2030 et 9,5 TWh en 2050 (en bleu outremer sur la **Erreur ! Source du renvoi introuvable.**).

En 2023, l'étude ADEME-Arcep propose un scénario tendanciel pour la consommation annuelle d'électricité des centres de données : pour 2030 et 2050, elles sont respectivement de 16,4 TWh et 39,0 TWh (en orange sur la **Erreur ! Source du renvoi introuvable.**). En orange sont indiquées d'autres projections de la consommation des centres de données. Et enfin sont aussi indiquées les projections d'un autre secteur concurrentiel : la mobilité (en jaune sur la **Erreur ! Source du renvoi introuvable.**).

Ces valeurs sont selon des scénarios tendanciels : elles ne sont donc pas les valeurs maximales possibles.

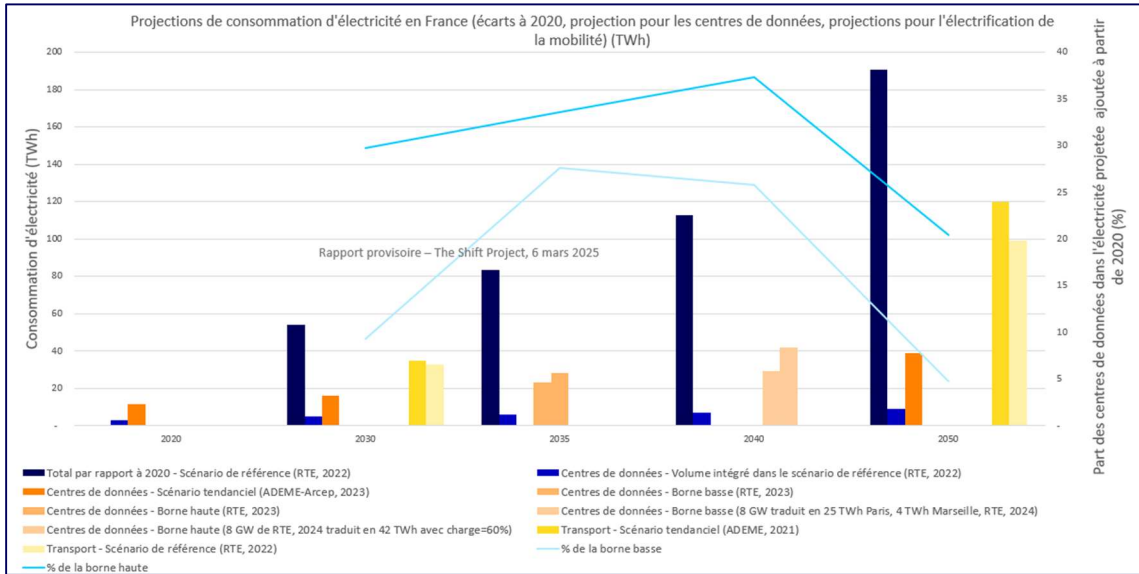


Figure 28 - Projections de consommation d'électricité en France (écarts à 2020, projections pour les centres de données et pour l'électrification de la mobilité). Sources : (ADEME & Arcep, 2023; RTE, 2022, 2023, 2024)⁴²

En bleu turquoise (lecture sur l'axe à droite), sont représentées les parts de la consommation des centres de données dans les projections d'ajout d'électricité à partir de 2020 (référence à 454 TWh) à l'horizon 2030 (508 TWh, soit + 54 TWh par rapport à 2020) et 2050 (645 TWh soit + 191 TWh par rapport à 2020).

- **L'addition de 11 TWh (16 TWh - 5 TWh, pris en compte) entre 2020 et 2030 selon l'ADEME-Arcep représenterait 30 % de l'ajout d'électricité du scénario de référence de RTE entre ces mêmes dates.**
- **L'addition de 20 TWh entre 2020 et 2035 selon le bilan prévisionnel de RTE représenterait 31 % de l'ajout d'électricité du scénario de référence de RTE entre ces mêmes dates.**
- L'addition de 30 TWh entre 2020 et 2050 selon l'ADEME-Arcep représenterait 20 % de l'ajout d'électricité du scénario de référence de RTE entre ces mêmes dates.

La dynamique actuelle du déploiement des centres de données soulève donc des questions quant à la répartition de l'électricité ajouté au réseau qui serait consommée en France⁴³ selon les scénarios et donc, en pratique, quant à la quantité d'électricité disponible (production + imports).

Au-delà des incertitudes conséquentes autour du numérique, une question ne s'est pas encore posée et doit l'être rapidement : quelles parts occuperont la sobriété et l'efficacité énergétique dans ce secteur ? Alors que les scénarios sont parfois déjà très optimistes quant aux efforts de sobriété et d'efficacité énergétique, peut-on parier sur cela dans un secteur avec une telle dynamique actuelle que le numérique ? Dans quels secteurs, la sobriété est-elle préférable : la mobilité, l'industrie ou le numérique ?

⁴² Pour 2040, le Schéma décennal de développement du réseau (SDDR) publié par RTE en 2024 annonce jusqu'à +8 GW en capacité pour les centres de données en 2040. En supposant un taux d'utilisation de 60% cela représenterait une consommation de 42 TWh en 2040.

⁴³ Ainsi que du rôle d'équilibrage des réseaux voisins (au Royaume-Uni et en Allemagne) durant l'hiver notamment.

4. Les centres de données peuvent aussi être des producteurs d'électricité bien que consommateurs nets d'électricité

L'un des éléments clé de la transformation des systèmes électriques vers leur décarbonation consiste à déployer des sources d'énergies renouvelables variables, comme l'énergie solaire et éolienne, dont la production dépend des conditions météorologiques et n'est donc pas garantie en permanence. Cela pose des enjeux d'équilibre offre-demande instantané pour le réseau électrique. Dans les différents scénarios énergétiques, plus le niveau d'intégration des énergies renouvelables intermittentes est important, plus l'enjeu de déploiement de capacités flexibles est crucial (pour le cas français par exemple : (RTE, 2022)). La flexibilité d'un moyen de production, de consommation ou de stockage d'électricité est sa capacité à modifier le moment d'apport ou de consommation d'électricité du réseau.

Cette caractéristique est souvent évoquée en ce qui concerne les centres de données. Afin d'assurer la continuité d'activité requise pour les services les plus critiques, les centres de données doivent s'équiper de systèmes de production d'électricité déconnectés du réseau qui prendront le relais en cas d'incident sur ce dernier. Pour ce faire, les centres de données sont classiquement équipés de groupes électrogènes. Les objectifs climatiques et d'intégration d'énergies renouvelables des acteurs, combinée à l'amélioration des performances des batteries électriques, mènent cependant certains d'entre eux à s'équiper de systèmes de stockage d'énergie par batterie (BESS en anglais⁴⁴) en lieu et place de ces groupes électrogènes⁴⁵.

Outre la gestion de l'intermittence, la mobilisation très rapide de l'énergie stockée via les BESS peut également permettre aux opérateurs de centre de données de vendre, en plus des services IT, des services d'effacement⁴⁶ ou d'équilibrage des réseaux électriques en utilisant la puissance et l'énergie disponibles au sein de leurs systèmes de batterie. Cette posture permet aux centres de données de doublement justifier leur statut d'industrie stratégique et prioritaire : à la fois en tant qu'acteur de premier plan de la digitalisation de l'économie, mais aussi en tant que potentiel partie prenante de la décarbonation de l'économie⁴⁷. Cette posture peut et doit interroger :

- En positionnant les centres de données comme des fournisseurs d'énergie ou de services aux gestionnaires de réseaux électriques, le système électrique pourrait en partie devenir dépendant du secteur numérique, a minima localement.
- Les capacités de batteries déployables étant limitées (contraintes de ressources physiques, de structuration et mobilisation des chaînes de valeur etc.), se pose la question de la concurrence d'accès et d'usage des batteries électriques, dans un contexte de besoins intenses d'électrification dans toute l'économie pour sa décarbonation (mobilité, industrie, utilisation directe par les infrastructures électriques pour leur équilibrage etc.).
- Le développement intense des centres de données posant des enjeux cruciaux quant à la disponibilité énergétique et les conditions de compatibilité de ces dynamiques avec la décarbonation du numérique et de l'économie dans son ensemble, la possibilité d'un rôle

44 Battery Energy Storage System

45 Voir <https://blog.se.com/datacenter/2024/05/01/the-rise-of-bess-powering-the-future-of-data-centers/>

46 Méthode de gestion de la consommation d'électricité, consistant pour un industriel à réduire volontairement sa consommation (contre rémunération), sur demande du gestionnaire de réseau. L'effacement peut permettre de soulager le réseau lorsque la demande en consommation est trop forte par rapport à la production.

47 Voir à ce titre les commentaires sur le projet de décret et d'arrêté relatifs aux mesures d'urgence définies en application des articles L. 321-17-1 et L. 321-17-2 du code de l'énergie fixant la liste des catégories de sites de consommation exemptées de l'obligation de mise à disposition des groupes électrogènes de secours

(<https://www.consultations-publiques.developpement-durable.gouv.fr/projets-de-decret-et-d-arrete-relatifs-aux-mesures-a2754.html>)

structurant et essentiel des centres de données pour la décarbonation de l'économie via ce positionnement reste à démontrer, notamment au regard des phénomènes documentés dans ce rapport.

III. Une première réflexion sur les modélisations possibles et leurs objectifs

Nous proposons dans cette section du rapport intermédiaire d'explorer un certain nombre de questions en lien avec la modélisation des impacts des centres de données, et les objectifs climatiques et énergétiques associés. Nous serions ravis de recevoir vos retours sur l'ensemble de ce chapitre.

Quels intérêts à modéliser les différentes évolutions possibles pour les centres de données : évaluer les consommations d'électricité et les émissions de gaz à effets de serre des nouveaux centres de données pour en décrire les effets sur les transitions énergétiques et écologiques ? Quelles sont les modélisations existantes ? Éclairent-elles déjà des décisions publiques ?

Quels sont les différents types de centres de données ? Les questions qui se posent sont-elles généralistes ou dépendent-elles des différents types de centres de données ? Quels sont les périmètres et les granularités géographiques pertinentes ? Quels sont les différents types d'offres, d'usages, de besoins ? Y a-t-il différents besoins en stockage de données et en calculs ? Quelles sont les trajectoires d'atténuation et de changement climatique à considérer ? Quels scénarios serait-il pertinent de décrire ?

Quelles sont les informations nécessaires pour parvenir à modéliser les consommations énergétiques et prévoir les impacts climatiques des centres de données ? Quelles sont les données qu'il faut dorénavant recueillir pour mettre en place des politiques adaptées sur les centres de données ? Quelles sont les sources ou les référentiels existants ?

A. Objectifs d'éclairage énergétique et climatique

Prévoir les différentes options pour les centres de données, et plus largement pour la transition énergétique sur laquelle le secteur numérique grignote, semble devenir incontournable. L'intérêt de ces prévisions paraît d'autant plus important que les politiques commerciales et fiscales à la manœuvre du déploiement massif de centres de données ne semblent pas alignées avec des politiques énergétiques et territoriales pérennes et résilientes. Derrière les investissements en centres de données et en IA, se trouvent bien des décisions structurelles, qui peuvent aisément fragiliser la transition électrique, d'autant plus si celle-ci est désordonnée.

Du point de vue climatique, la part du secteur numérique dans l'empreinte carbone augmente, sans que les autres secteurs ne prévoient d'émettre moins pour compenser les émissions supplémentaires du numérique. Cette désorganisation doit être maîtrisée et rectifiée rapidement au risque de mettre en péril nos objectifs climatiques. Quelles sont les émissions de gaz à effets de serre des nouveaux centres de données ?

B. Critères, contraintes, périmètre, découpage

- Quels sont les critères à considérer et les questions à se poser pour les identifier ?

- Consommation électrique, émissions de gaz à effets de serre territoriales, empreinte carbone française ?
- Comment mettre en évidence les transferts d'impacts possibles entre énergie et carbone (renouvellement pour des équipements plus efficaces par exemple) ? Quel optimum global rechercher ?
- Comment modéliser les effets d'autres critères d'impacts tels que les consommations d'eau, les déchets générés, l'artificialisation des sols ou encore les pressions sur les ressources minérales, en complément de l'approche classique (énergie, carbone) du Shift Project ?
- Quelles sont les contraintes à prendre en compte ?
 - Compte-tenu des trajectoires de réchauffement sur lesquelles nous sommes, les caractéristiques des lieux d'implantation des centres de données définissent-elles de nouvelles contraintes : zones interdites car ne disposant pas du volume d'eau requis pour les centres de données en période de sécheresse et de forte chaleur ? zones interdites car susceptibles de créer des îlots de chaleur en ville ?
 - Compte-tenu des trajectoires d'atténuation sur lesquelles nous sommes engagés, comment modéliser l'impact de l'artificialisation des sols liée à l'implantation de nouveaux centres de données, en particulier sur des zones naturelles (prairies, forêts, friches urbaines) ?
 - Y a-t-il des limites énergétiques, climatiques, matérielles et financières à modéliser comme contraintes ?
- Quel périmètre géographique serait à considérer ?
 - France dans un premier temps.
 - Quels seraient les intérêts à modéliser les centres de données en Europe ?
 - Un périmètre permettant de décrire la part "importée" de nos consommations numériques, et les dépendances qui en résultent ?
- Quel découpage géographique choisir pour ce modèle ?
 - Par régions pour pouvoir prendre en compte des contraintes locales : que ce soit pour les sols, pour les postes RTE, pour des enjeux de latence (proximité avec des interconnexions (IXP), proximité avec utilisateurs ou données) ?
 - Par pays pour pouvoir prendre en compte des facteurs d'émissions de l'électricité différents (si le périmètre géographique considéré est européen) ?
- Le mot "centre de données" est aujourd'hui un mot valise qui recoupe des tailles (en m² ou GW d'IT), des modèles d'affaires (on-prem, fournisseurs de services, fournisseurs d'infrastructures de colocation), des usages (polyvalents, spécialisés dans le cloud, le compute, le storage, l'entraînement, l'inférence...) et des caractéristiques techniques très différentes. Quelle typologie de centres de données définir puisque :
 - Les matériels (types de serveurs, fréquences de renouvellement) sont différents selon les types de centres de données.
 - Les installations de refroidissement et énergétiques sont différentes selon les types de centres de données.
 - Quelles sont les caractéristiques qui seront disponibles dans le relevé EED ?

- La typologie que nous proposons est disponible en Annexe 1 : edge, entreprise, fournisseurs de services de télécommunication, colocation ou co-hébergement, hyperscaler.
- Y-a-t-il un intérêt à séparer les centres de données par type d’approvisionnement électrique (Enedis, RTE) ? Voire même à modéliser des réseaux d’électricité privatisés ?

C. Modélisation

D’un point de vue énergie / climat, une façon de démarrer ce travail consiste à se poser les trois questions successives :

- Quelles trajectoires d’usages ?
- Quelles trajectoires d’efficacité ?
- Quelles trajectoires énergétiques ?

Cela revient à démarrer à partir d’une modélisation proche de l’équation de Kaya (comme dans la publication (Ni W. et al., 2024)) :

$$C = D \times \frac{E_{data}}{D} \times \frac{E}{E_{data}} \times \frac{C}{E} = D \times e \times PUE \times f$$

Où :

- C sont les émissions de gaz à effets de serre totales (en MtCO₂e),
- D représente la mesure de la demande en calcul des centres de données considérés (en EFLOPs, opérations en virgule flottante par seconde)⁴⁸,
- E_{data} représente la consommation d’énergie IT dans les centres de données (en kWh),⁴⁹
- E représente la consommation totale d’énergie des centres de données mesurée (en kWh),
- e = E_{data} / D représente la consommation d’énergie par unité de calcul, ou intensité énergétique des centres de données (en kWh/ EFLOPs),
- PUE = E/E_{data} représente la mesure couramment utilisée pour mesurer l’efficacité énergétique des centres de données,⁵⁰
- f = C/E représente l’intensité globale des émissions des centres de données (représentatif du facteur d’émission de l’électricité, mais aussi de l’empreinte embarquée des équipements, en divisant par la durée d’utilisation dans le cas d’un modèle “stock”).

⁴⁸ Cette demande en calcul pouvant être exprimée par rapport à la capacité disponible via la charge.

⁴⁹ La consommation d’énergie en veille, dite “idle”, doit être attentivement prise en compte dans E_{data} ou dans E.

⁵⁰ Attention différents types de PUE coexistent : PUE 1 = tout ce qui sort des onduleurs ; PUE 2 = tout ce qui entre dans la multiprise alimentant un rack ; PUE 3 = tout ce qui sort de la multiprise alimentant un rack.

A partir de cette base, et par des jeux de sommation, il est possible de préciser ce modèle avec par exemple :

- Différentes formulations de l'offre et de la demande (des opérations mais aussi le trafic par exemple),
- Différentes régions (en France, ou en Europe),
- Différents types de centres de données, qui ont des efficacités et des PUE différents.

Toujours à partir de cette base, il pourrait être judicieux d'ajouter :

- Une variable représentative de la durée d'installation d'un centre de données, qui pourrait agréger des durées comme celles de mise en compatibilité des plans locaux d'urbanisme, d'agrément, d'autorisation environnementale, de permis de construire, d'acquisition du foncier, de raccordement électrique, de raccordement numérique et de raccordement aux productions ENR (voir graphique p22 (L'institut Paris Région, 2023)),
- Un profil de remplissage du centre de données, et l'identification de variables qui sont dépendantes du remplissage (ex : PUE), éventuellement tenant compte de la disponibilité matérielle des serveurs,
- Un profil de PUE, selon les zones géographiques (température et selon les technologies de refroidissement) éventuellement,
- Un profil de taux de charge, éventuellement modélisant différentes politiques énergétiques,
- La variation de consommation au cours des journées ou de l'année afin de modéliser les pics de consommation, ou non. Ceci permettrait de réfléchir aux surdimensionnements possibles, aux stratégies d'allocation voire d'hyper-allocation courantes des acteurs du numérique (en réponse à de la rentabilité matérielle voire même des stratégies commerciales de services vendus en "heures creuses numériques"), aux systèmes utilisés aujourd'hui en cas de pannes (batteries "tampon" et volants d'inertie avant l'allumage des générateurs de secours au diesel gaz) et aux demandes de flexibilités possibles.
- La prise en compte des durées d'amortissement matérielles : quelles sont les durées de vie des serveurs et que deviennent les éléments une fois remplacés ? Et comment prendre en compte une obsolescence différenciée des types d'équipements IT dans les serveurs puisque les impacts à la fabrication et à la consommation sont différents aussi : quelles durées de vie respective pour des CPU, GPU, RAM, alimentation, cartes, disques ? et quelles évolutions technologiques différenciées ?
- De manière générale, considérer l'évolution de ces variables selon différentes phases du centre de données : montée en charge (~3 ans ?), utilisation nominale (~15 ans ?), fin de vie.

Si on se concentre sur la modélisation de l'offre et de la demande :

- La modélisation d'un nombre de racks et d'une puissance kW/rack serait une variable représentative de l'offre installée (c'est par exemple le cas dans l'étude (Schneider Electric, 2023)).

- La modélisation de la surface utilisée et d'un profil d'évolution de la densité de puissance (kW/m²) par surface serait une variable similaire (c'est par exemple le cas dans (ADEME & Arcep, 2023)).
- Plus macroscopiquement, la répartition des centres de données avec leur puissance (GW) serait une variable intéressante en tant qu'elle permettrait de s'appuyer sur le système commun de notation des centres de données à l'échelle de l'Union européenne à venir (Journal officiel de l'Union européenne, 2023). Mais comment modéliser les évolutions possibles : à partir d'une liste de projets en cours et des probabilités de réalisation ?
- Il serait également possible de séparer les capacités informatiques entre "data" (stockage de données) et "compute" (traitement de données), une façon de prendre en compte les évolutions vers plus d'activités de calcul , ainsi que modéliser les évolutions par type de matériel, les tendances associées, afin d'identifier les leviers d'optimisation sur les différents types de matériels (stockage, compute traditionnel, IA, ...).
 - Dans le "Référentiel méthodologique d'évaluation environnementale des services d'hébergement informatique en centre de données et de services cloud" (ADEME, 2023), il y a deux unités fonctionnelles distinctes pour : "mettre à disposition un serveur physique hébergé en centres de données avec une capacité de calcul donnée" ou de "stockage donnée".
 - Pour les centres de données spécialisés en IA, la demande pourrait potentiellement être modélisée en « token ».
- Il pourrait être intéressant de séparer par types d'utilisations : cloud, stockage, IA, IA générative, entraînement, inférence, afin par exemple de modéliser divers taux d'évolution pour chacune de ces utilisations.
- Il pourrait être intéressant de séparer par types d'utilisateurs : entreprises, particuliers et même "robot"⁵¹.
- La modélisation de l'offre / demande doit être prise en compte de façon différenciée selon les types de centres de données pour inviter notamment à la réflexion sur l'évolution des capacités informatiques dans les entreprises, à l'edge, ainsi que sur les centres de données spécifiques IA voire spécifiques entraînements (à ce sujet voir l'étude de De Vries sur le *bottom-up* en partant de nombre de GPU vendus ou prévus (de Vries A., 2023).

D. Scénarisation

Différentes scénarisations peuvent être envisagées :

- Le scénario de référence pourrait être un scénario "Tendanciel + Sustainable AI", une façon de refléter la position du Sommet pour l'action sur l'IA (Elysée, 2025), à savoir un scénario tendanciel sur la demande, incluant des leviers d'écoconception. Quels leviers avec quelles amplitudes seraient représentatifs de ce "Sustainable AI" ?
- Un scénario uniquement « Tendanciel » ;

⁵¹ Au sens où les bots sont de plus en plus nombreux et génèrent "seuls" ou plutôt de façon "automatisée" des usages (Thales Group, 2024).

- Deux scénarios inflationnistes, avec ou sans dépassement des capacités de production d'électricité bas-carbone ;
- Un scénario "Ecoconception et sobriété" ;
- Des scénarios touchant aux limites matérielles, foncières, énergétiques, ou climatiques.

IV. Intégration territoriale et régionale des centres de données : de premières questions pour une stratégie résiliente ?

Nous proposons dans cette section du rapport intermédiaire de lister certains effets, d'explorer les questionnements que cela doit soulever, sans toutefois avoir la prétention d'épuiser le sujet.

Vos retours, critiques ou propositions sur les questions posées et les orientations explorées à ce stade seront des éléments clés des prochaines étapes de ces travaux !

Afin de permettre un développement pérenne, à la fois pour les régions, territoires et filières économiques, les régions et territoires d'accueil et d'intégration des centres de données vont dorénavant avoir à intégrer plusieurs incontournables dans leurs stratégies dont :

- **Les trajectoires de réchauffement** sur lesquelles nous sommes, qui peuvent définir de nouvelles contraintes (ex : zones interdites car ne disposant pas du volume d'eau requis pour les centres de données en période de sécheresse et de forte chaleur, zones interdites car susceptibles de créer des îlots de chaleur en ville, etc),
- **Les trajectoires d'atténuation** sur lesquelles nous sommes engagés, qui peuvent aussi définir de nouvelles contraintes (ex : objectifs de Zéro Artificialisation Nette, préservation de zones naturelles (prairies, forêts, friches urbaines, etc)),
- **Le changement d'échelle des centres de données**, qui pour certains peuvent dorénavant être au même niveau de concentration que certaines industries d'ampleur⁵² actant un changement d'échelle du tertiaire vers l'industriel (Carnino, Marquet, 2022).

À la lumière de ces trois grandes trajectoires, si les externalités positives et négatives des centres de données (emploi, attractivité, pollution de l'air, ressources limitées en eau, sols, énergie, climat, environnement, etc.) restent relativement stables en nature, leurs impacts quantitatifs peuvent varier, voire franchir des seuils critiques. Ainsi, une externalité jugée mineure hier pourrait devenir significative demain.

⁵² Comme l'illustre l'IEA, aux Etats-Unis, les centres de données sont plus concentrés spatialement que les usines sidérurgiques, mines de charbon et de minerais, centrales électriques et entrepôts (IEA, 2024b) (fig 4.12).

V. L'état des lieux en Europe

L'échelle européenne est intéressante pour observer à la fois ce qui se passe dans d'autres pays notamment l'adéquation entre les différentes politiques énergétiques (fiscales, industrielles, numériques, climatiques, énergétiques...) mais aussi pour encadrer les consommations énergétiques et les émissions de gaz à effets de serre à son échelle.

L'institut Borderstep a développé un modèle pour les centres de données au niveau européen, alimentant une scénarisation proposée en 2020 dans le cadre des travaux de la Commission européenne sur le cloud et les centres de données (Figure 29) (European Commission, 2020).

En 2024, la Commission européenne a conduit une étude sur la consommation énergétique des centres de données en Europe (EU27) (Figure 30) (European Commission, Kamiya G., Bertoldi P., 2024).

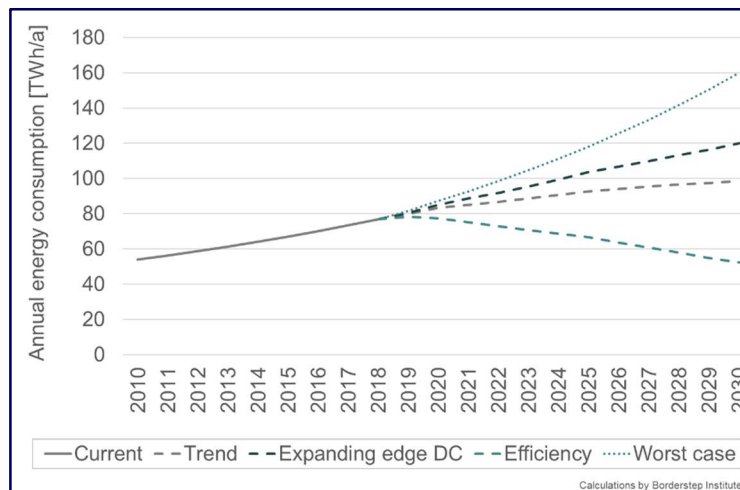


Figure 29 - Scénarios possibles pour l'évolution de la demande énergétique des centres de données dans l'UE28 jusqu'en 2030. Source : (European Commission, 2020).

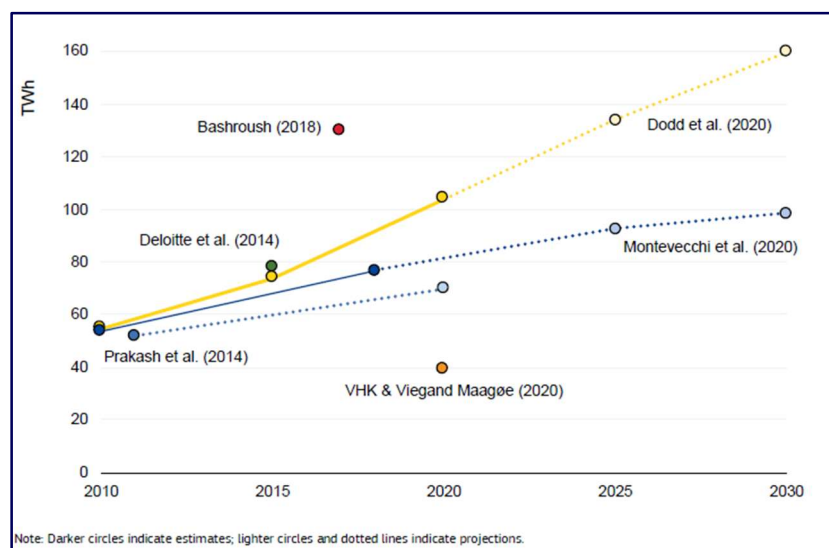


Figure 30 - Estimations pour la consommation énergétique des centres de données européens. Source : (European Commission, Kamiya G., Bertoldi P., 2024)

En 2018, les centres de données de l'UE (UE28) consommaient 77 TWh d'énergie.

D'après la Commission européenne (European Commission, 2024), cette consommation augmenterait de manière significative d'ici 2030 : en 2018, une hausse de 28 % était déjà prévue, mais désormais, avec l'IA – générative-, une augmentation de deux à trois fois pour certains pays est anticipée. Au sein de l'Union, les centres de données représentaient 2,7 % de la demande d'électricité en 2018 et pourraient atteindre au moins 3,2 % d'ici 2030 si le développement se poursuit sur la trajectoire actuelle (European Commission, 2024).

D'après un rapport de la Commission européenne (European Commission, Kamiya G., Bertoldi P., 2024), les centres de données de l'UE27 consomment entre 45 et 65 TWh d'électricité en 2022, soit entre 1,8 et 2,6 % de la consommation totale d'électricité. Cette estimation est légèrement inférieure à celle de (Montevecchi, F. et al., 2020) qui estiment que les centres de données représentaient 2,7 % de la consommation totale d'électricité de l'UE28 en 2018.

En 2024, l'Agence Internationale de l'Énergie a estimé une consommation électrique de 100 TWh pour l'Europe⁵³ en 2022, avec une projection à 150 TWh pour 2026 (IEA, 2024a).

De son côté, le cabinet de conseil en stratégie McKinsey propose un scénario de référence visant 150 TWh en 2030 pour l'Europe (UE27 + Royaume-Uni), soit une augmentation de 85 TWh entre 2023 et 2030. Cela porterait la part des centres de données dans la demande d'électricité européenne à 4.5 %. Cette croissance étant notamment attribuée aux hyperscalers (70% de la demande prévue d'ici à 2028) et centres de données de colocation. Selon McKinsey toujours, 25 GW de capacités IT pourraient être à installer d'ici à 2030, ce qui représenterait 15 à 25 % de l'ensemble de la nouvelle demande européenne nette jusqu'en 2030 (McKinsey & Company, 2024).

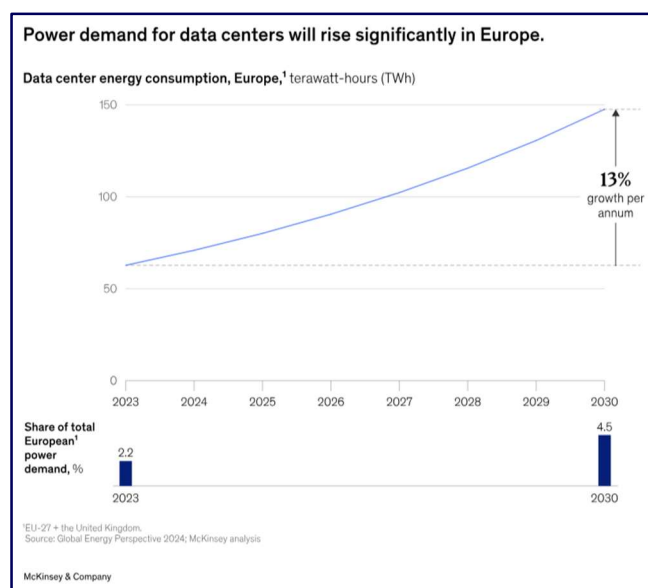


Figure 31 - Projection de consommation énergétique des centres de données en Europe en TWh. Source : (McKinsey & Company, 2024)

⁵³ Périmètre géographique à bien identifier selon les études.

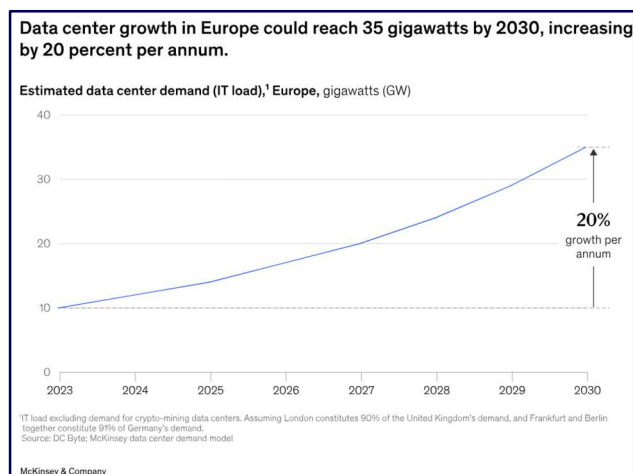


Figure 32 - Projection de demande d'installation de capacités informatiques en Europe GW. Source : (McKinsey & Company, 2024)

Dans ce contexte, il est crucial de noter que ces dynamiques sont très concentrées (Figure 33) sur quelques zones, avec :

- Deux tiers de la consommation énergétique des centres de données pour quatre pays : Allemagne, France, Pays-Bas et Irlande, alors qu'ils ne rassemblent que 40 % de la population européenne,
- La part dans la consommation nationale d'électricité étant très hétérogène, atteignant 18% en Irlande, 5,6% aux Pays Bas ou encore 4,9% au Luxembourg et 4,6% au Danemark,
- Les douze principaux marchés, que cette étude a estimés sur la base de données nationales, représentent environ 95 % de la consommation d'énergie des centres de données de la région (European Commission, Kamiya G., Bertoldi P., 2024).

L'écosystème européen se caractérise donc par de fortes disparités régionales dans les dynamiques d'implantation des centres de données. Dans ce contexte, la Commission européenne et chaque pays doivent pouvoir garantir le développement maîtrisé des projets de centres de données, dans le respect des contraintes régionales, des feuilles de route de développement des capacités électriques nationales, et des engagements de réduction de l'empreinte carbone.

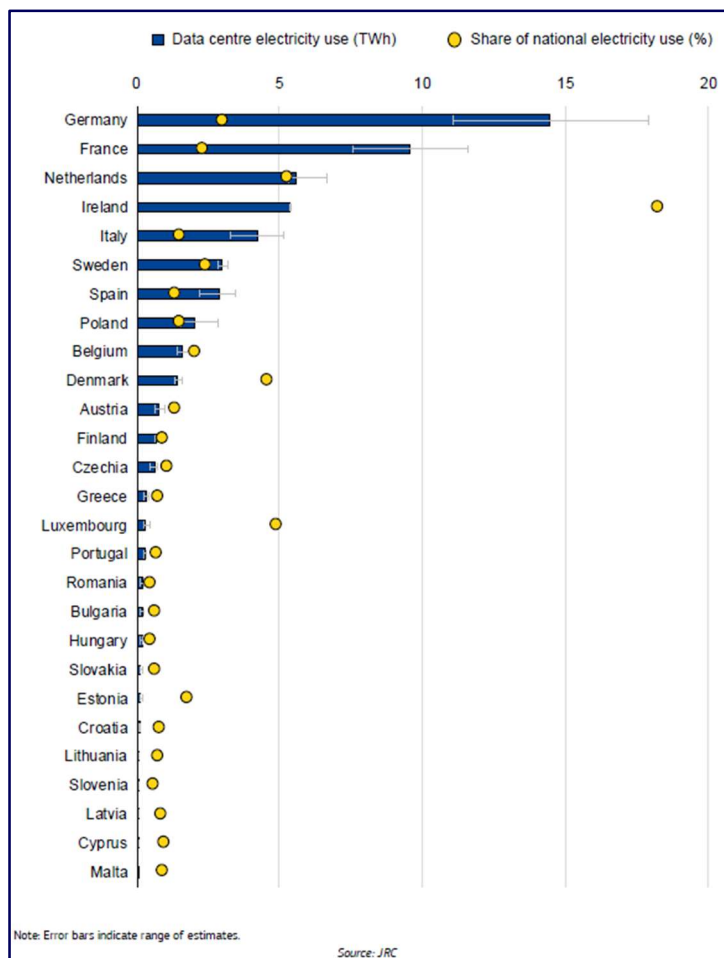


Figure 33 - Estimation de la consommation d'énergie des centres de données par pays en EU27 en 2022 en TWh , Source : (European Commission, Kamiya G., Bertoldi P., 2024).

Pour la Commission européenne, les grands centres de données (notamment centres de données de colocation et hyperscalers) représentent environ 65 % du total, contre 35 % pour les centres de données d'entreprise⁵⁴ (European commission, 2024).

Il nous paraît important dans la suite des travaux, suite à ce rapport intermédiaire, de mettre en regard face à ces projections haussières pour les centres de données, les capacités énergétiques européennes.

Au niveau européen, la directive relative à l'efficacité énergétique (Energy Efficiency Directive), qui structure la réglementation de l'Union européenne de promotion de l'efficacité énergétique et de réduction de la consommation d'énergie, a été révisée et impose désormais aux exploitants de centres de données de communiquer leurs indicateurs de performance clés tels que la consommation électrique, les performances informatiques ou les bandes passantes nécessaires pour le réseau (Journal officiel de l'Union européenne, 2024).

⁵⁴ Le rapport compare cette répartition aux estimations de (Dodd N. et al., 2020) qui estimaient une répartition 56/44 en 2020 et prévoient une répartition 66/34 en 2025.

Cette mise à jour de la directive vise à accroître la transparence dans le domaine, en établissant une base de données européenne de données pertinentes pour la performance énergétique (et l'empreinte hydrique) des centres de données et également, à promouvoir de nouvelles conceptions et de nouveaux développements en matière d'efficacité.

Avec le règlement délégué sur le système de notation des centres de données, l'Union Européenne va pouvoir collecter des informations sur les centres de données et attribuer des indicateurs de soutenabilité, ce qui est un premier pas mais probablement insuffisant pour infléchir les courbes de consommation énergétique et d'émissions de gaz à effets de serre.

Beaucoup de questions peuvent se poser à l'échelon européen :

- *Quel sera l'avenir des émissions carbone des centres de données en Europe ? Les émissions pourront-elles être insérées dans le marché des permis d'émissions EU-ETS ?*
- *La vision de transition jumelle ("twin transition", combiné des transitions écologiques et numériques) incorporera-t-elle les réflexions sur la soutenabilité climatique des infrastructures numériques ?*
- *Quelle sera la place des politiques européennes en complément des politiques nationales et locales sur l'objectif climatique du secteur numérique ?*

Une IA au service de l'intérêt général doit être à la hauteur des enjeux climatiques et de transition énergétique

I. Approche par cas d'usage : quels services et promesses technologiques pour l'IA ?

Usages et infrastructures sont les deux faces de la même pièce : la dimension que nous donnons aux services et applications IA est directement liée à celle des supports physiques de calcul, de traitement, de stockage et de transmission de la donnée que nous déployons pour les rendre possibles. Les dynamiques décrites dans la partie précédente permettent de comprendre les trajectoires physiques macroscopiques, dans lesquelles pourront s'inscrire nos usages. Cependant, de nombreuses possibilités existent quant au mix de services et applications qui se développeront effectivement sur ces bases.

A. IA et IA générative : une rupture technologique ?

La définition de "l'intelligence artificielle" est moins normative que dynamique : l'horizon technologique se déplace. En simplifiant, on pourrait dire que l'IA désigne les applications d'automatisation les plus avancées, à un moment donné, en termes de traitement de l'information, de complexité des tâches, de précision et de fiabilité.

Le principe du numérique est de produire des données de sortie en réponses à des informations en entrée sur lesquelles il applique des règles déterminées. Les paradigmes technologiques de nos machines de traitement de l'information se sont cependant diversifiés, notamment avec le développement de l'apprentissage machine. Ce dernier automatise, au moins en partie, la construction des règles à appliquer : plutôt que de s'appuyer exclusivement sur l'humain, le système numérique est d'abord chargé d'élaborer ses propres règles. Il les génère sur la base

d'une analyse statistique des liens existants entre des données d'entrée et les sorties attendues⁵⁵, avant de les appliquer à son tour.

Les travaux présentés dans le cadre de ce rapport rejoignent en cela sensiblement les conclusions du rapport de la Commission de l'intelligence artificielle (Commission de l'intelligence artificielle & Gouvernement français, 2024) quant à la définition à donner à notre objet d'étude : les nouvelles dynamiques applicatives de l'IA découlent d'une disponibilité accrue de données stockées et valorisables par des infrastructures de calcul aux dimensions et à la complexité croissantes.

L'intelligence artificielle générative en est l'une des avancées les plus visibles. La spécificité de cette nouvelle famille d'algorithmes est de pouvoir générer de l'information, forte de quatre caractéristiques généralement retrouvées dans les services qui la composent (Commission de l'intelligence artificielle & Gouvernement français, 2024) :

- **L'information produite est souvent perçue comme réaliste**, c'est-à-dire qu'il est difficilement identifiable qu'il s'agit d'un contenu produit par un algorithme (textes cohérents, images réalistes, contenu vocal aux intonations convaincantes etc.) ;
- **L'information est produite rapidement** : les interactions avec ces outils sont fluides et permettent une réponse rapide à l'échelle humaine, de l'ordre de quelques secondes ;
- **La polyvalence : les informations produites peuvent être de natures très différentes, et parfois complexes, tout en gardant un haut degré de fiabilité**, c'est-à-dire que les informations restent réalistes y compris pour des tâches avec des instructions multiples et précises ;
- **L'interaction humain-machine nécessaire à la production d'informations est simple** (chatbot, interactions vocales etc.), c'est-à-dire que l'interaction a l'air naturelle pour un être humain.

Ces quatre caractéristiques, piliers de l'originalité et du taux de pénétration rapide des services s'appuyant sur l'IA générative, sont rendues possible par les grandes quantités de données et de capacités de calcul disponibles pour l'entraînement des modèles. C'est la mise en place d'une phase d'entraînement intense en amont de la phase d'inférence (phase d'utilisation), qui permet de produire un système qui puisse être particulièrement réactif et adaptable tout en restant suffisamment fiable. Les "Large Language Models" (LLM) constituent une brique centrale de ces nouveaux systèmes.

Si l'offre en promesses et projections de services aujourd'hui classés dans la catégorie "intelligence artificielle" est dopée par le déploiement large spectre des applications génératives, il se cache en réalité une multitude de fonctionnalités possibles, s'appuyant sur la technologie générative ou non. Les briques technologiques algorithmiques et d'acquisition/traitement de données étant partie intégrante de nombreuses applications déjà déployées, il convient de considérer l'IA générative comme une brique technologique supplémentaire mobilisable pour rendre un service donné. Au vu de son intensité infrastructurelle particulièrement importante en comparaison des autres types de solutions technologiques, poser la question de sa nécessité pour rendre le service souhaité sera inévitable dans le cadre de la conception de biens et services numériques compatibles avec la double contrainte carbone.

⁵⁵ Par exemple, en lui demandant de calculer la "probabilité" d'occurrence d'un mot suivant les mots qui le précèdent, sur la base d'un corpus de textes.

Si l'axe précédent du rapport permet de dresser le tableau macroscopique des enjeux carbone-énergie des infrastructures de la donnée et du calcul, il s'agira ici de donner une grille de lecture repartant des services souhaités. Le travail présenté dans les pages qui suivent vise à fournir des points de repère aux acteurs pour façonner et orienter la conception et le déploiement des applications estampillées « IA » de manière à les rendre compatibles avec des dynamiques alimentant les trajectoires soutenables du numérique, plutôt que de les rendre délétères. La méthode propose d'identifier et caractériser :

- Les briques technologiques sous-jacentes à une proposition de service (segmentation technologique des fonctionnalités composant le service),
- Les infrastructures sollicitées par chacune des briques,
- Les impacts macroscopiques alimentés par les choix de mobilisation d'une fonctionnalité ou brique technologique.

B. Les cas d'usage : une illustration du panorama d'applications et services possibles d'intelligence artificielle

L'approche par cas d'usage développée dans ce rapport vise à dresser un panorama illustratif, plutôt qu'exhaustif, des applications possibles de l'IA. Une liste de cas d'usage a été construite de manière à visibiliser les questions et développer des démarches méthodologiques à répliquer par les parties prenantes des services IA lors des choix de conception (R&D, développement de produit, orientation d'investissements etc.) et de déploiement (achat de services, stratégie publique de déploiement etc.).

Cette démarche ne couvre donc pas la diversité des applications technologiques possibles pouvant tomber sous la définition "d'intelligence artificielle", présentes dans tous les secteurs de l'économie et parmi lesquelles on peut par exemple trouver⁵⁶ :

- Des analyses embarquées dans les caméras de vidéosurveillance (analyse de la circulation, analyses comportementales, déploiement en espaces publics ou privés, détection d'objets ou déchets sur l'espace public etc.) ;
- Des assistants conversationnels pédagogiques (apprentissage des langues, maths, français etc.) par interactions écrites ou verbales ;
- Des assistants conversationnels commerciaux (service client, simple conseil à identification de produits, accompagnement jusqu'à l'acte d'achat à partir d'une instruction simple etc.) dans de nombreux secteurs (assurance, tourisme, retail etc.) ;
- Des outils de détection de contenus pour assistance à la modération de plateformes et réseaux sociaux (contenus violents, contenus générés par IA etc.) ;

⁵⁶ Les exemples intégrés dans cette liste sont issus des échanges avec les parties prenantes du secteur. Le choix a été fait, à ce stade, de ne pas les accompagner d'exemple concrets, pour éviter la mise en avant de certaines solutions commercialisées par rapport à d'autres, la liste ne pouvant être absolument exhaustive."

- Des assistants de génération de code informatique (génération d'un code en réponse à un cahier des charges, traduction d'un langage informatique à un autre etc.) ;
- Des outils de modélisation et d'amélioration des capacités de calcul haute performance dans la recherche et les sciences appliquées (chimie et pharmacologie, matériaux, modélisation climatique etc.) ;
- Des assistants de recherches d'informations au sein de corpus documentaires données (démarches administratives, corpus légaux, synthèses de rapports techniques etc.) ;
- Des applications dans l'énergie et l'assistance à pilotage de systèmes électriques ;
- Des applications industrielles pour la maintenance prédictive ou l'optimisation logistique ;
- etc.

Dans les applications possibles de l'IA, on retrouve les différentes catégories des services numériques au vu de leur positionnement dans le contexte des contraintes environnementales :

- Les applications visant une réduction des impacts environnementaux (du numérique lui-même ou bien d'autres secteurs), communément désignés par l'appellation "AI for Green", l'intelligence artificielle au service des objectifs environnementaux (ces objectifs n'étant atteints que si le bilan net entre bénéfiques impacts environnementaux de la solution elle-même est positif) ;
- Les applications visant d'autres objectifs que des objectifs environnementaux : qualité des services de santé, amélioration de la productivité, efficacité ou la rentabilité de certaines activités, amélioration de la fiabilité de certains processus etc.

Au sein de cette seconde catégorie, on trouve certaines applications dont l'objectif est d'améliorer les rendements et l'efficacité de processus directement liés aux industries et infrastructures énergétiques fossiles. Même dans le cas où ils permettent de réduire les impacts unitaires de procédés fossiles, ces cas d'usage, parfois désignés par l'appellation "AI for Brown", ne peuvent être considérées comme des applications "AI for Green" (Carbone 4, 2024).

L'objectif du travail ici produit étant de questionner les conditions d'un développement et d'un déploiement des technologies de type "IA" compatible avec la double contrainte carbone, toute application s'inscrivant dans un scénario de maintien des énergies fossiles n'est donc, par construction, pas compatible avec notre périmètre.

Le numérique, et l'IA tout particulièrement, est un catalyseur du système dans lequel on l'intègre (amélioration de l'efficacité, de la rapidité d'exécution, de la productivité etc.). Le déployer sans mettre en place de trajectoire de référence et de point de repère pour en orienter la conception et les infrastructures mènera donc mécaniquement à placer l'IA dans un paradigme systémique qui augmente les impacts environnementaux.

Déployer des cas d'usage de l'IA sans l'intégrer dans une stratégie globale de décarbonation volontaire, pilotée et cohérente, c'est construire une IA qui restera tout autant fossile que l'économie dans laquelle on la place. Tant dans les ressources qu'elle mobilisera pour son propre fonctionnement que dans les dynamiques qu'elle participera à alimenter dans le reste de l'économie.

II. Impact énergie-climat de l'IA : analyse fonctionnelle des cas d'usage

A. Approche par cas d'usage : critères retenus et description technologique

Critères de sélection des cas d'usage

Les cas d'usage explorés dans ce rapport ont été sélectionnés de manière à éclairer les questions clés des phénomènes liés à l'IA et à ses infrastructures :

- **Le phénomène "IA générative"** : quelles implications a le déploiement de la technologie générative sur la structuration des services "IA" et de leurs impacts énergie-climat ?
- **L'orientation de l'innovation** : quelles différentes combinaisons de fonctionnalités et briques technologiques (taille, infrastructures sollicitées, générative ou non etc.) sont possibles pour concevoir un service IA compatibles avec les contraintes énergie-climat ?
- **La vision systémique du secteur numérique** : quels sont les impacts des services IA sur le système numérique dans son ensemble (terminaux, réseaux, centres de données) ?
- **La vision systémique de l'économie** : comment les services IA interagissent-ils avec les transformations en cours ou à mener pour décarboner les autres secteurs (mobilité, agriculture etc.)⁵⁷ ?
- **La vision quantitative** : comment quantifier les impacts énergie-carbone d'un cas d'usage défini, en intégrant les impacts indirects et la vision systémique dans l'analyse ?

Liste des cas d'usage retenus

Sur la base de ces critères, une liste de cas d'usage a été retenue afin d'en effectuer une analyse qualitative (puis, dans un deuxième temps, quantitative pour l'un d'entre eux) :

- **Mobilité autonome**
Technologies IA pour la mobilité dans différents environnements (complexité de l'espace, niveau de connectivité de la zone etc.) : routier urbain (voiture ou transports en commun/navette), routier non urbain (camion, machine agricole), maritime, rail, aérien.
- **Système d'optimisation des épandages**
Technologies d'analyse de données relevées en zone de culture (détection de zones à plus faible croissance, détection de ravageurs etc.) afin de suivre et optimiser les épandages d'engrais et pesticides.
- **Assistant personnel de compte-rendu**
Assistant (intégré à une application ou terminal d'échange avec contenu vocal) produisant des comptes-rendus sur la base des échanges vocaux des personnes utilisatrices.

⁵⁷ Le prisme d'analyse est avant tout énergie-climat dans le cadre de ces travaux, afin d'éclairer le cadre lié à cette double contrainte, à reconnecter avec les autres limites planétaires.

- **Outil de recherche en ligne**
Utilisation de l'IA générative comme outil de recherche d'informations en ligne, à l'image d'un moteur de recherche.
- **Outil d'aide au diagnostic**
Technologie IA d'analyse d'images, pour aide au diagnostic en anatomopathologie (détection de cellules cancéreuses par analyse de lames d'échantillons).
- **Production de création pour spot publicitaire**
Technologies d'IA générative au sein des services de création publicitaire pour produire des contenus images et vidéos.

Cas d'usage	Briques et niveaux technologiques possibles ?	Secteur concerné
Assistant personnel de compte-rendu	<ul style="list-style-type: none"> ● Analyse de contenu audio ● Technologie générative de contenu texte ● Possibilité de différents niveaux de fonctionnalité : simple compte-rendu texte, « enrichissement » du compte-rendu par proposition de contenus complémentaires, intervention de l'assistant directement dans les échanges par texte ou verbal, traduction en direct sur le terminal etc. 	Grand public, "end-users"
Outil de recherche en ligne	<ul style="list-style-type: none"> ● Détection et classification de contenu pour leur indexation ● Technologie générative de contenu texte ● Intégration dans des systèmes algorithmiques de moteurs de recherche 	Grand public, "end-users"
Production de création pour spot publicitaire	<ul style="list-style-type: none"> ● Technologie générative de contenu image et vidéo 	Commerce & publicité

Mobilité autonome	<ul style="list-style-type: none"> • Analyse et reconnaissance d'images et vidéo • Apprentissage par renforcement • Localisation de la capacité de calcul (edge/embarquée ou centralisée/déportée) 	Transport
Système d'optimisation des épandages d'engrais et pesticides	<ul style="list-style-type: none"> • Acquisition des données : capteurs in situ, acquisition satellite, relevés aériens (par drones etc.) • Analyse des données : synthèse de données, analyse et corrélation, analyse et reconnaissance d'images etc. 	Agriculture
Outil d'aide au diagnostic médical	<ul style="list-style-type: none"> • Analyse et reconnaissance d'images • Niveau de déploiement/massification (spécifiques ou généralisée, nombre de sites ou services de soins, intégration dans les parcours médicaux etc.) 	Santé

Table 1 - Cas d'usage

B. La "Boussole de l'IA" : de l'analyse fonctionnelle du cas d'usage à la qualification de ses effets sur les impacts énergie-climat

La méthodologie d'analyse qualitative construite dans le cadre de ce rapport vise à répondre à la question suivante : quelles sont les conditions de compatibilité d'un service IA avec la contrainte énergie-carbone ?

L'analyse économique n'est en effet pas suffisante pour savoir si une solution est sobre ou frugale, même lorsqu'elle permet de traduire une réduction des consommations énergétiques directes du service : les postes d'impact ne sont pas nécessairement directs mais sont au contraire mutualisés dans les infrastructures (réseaux, centres de données, chaînes de valeur des terminaux, capteurs et électronique). Les cas unitaires sont ainsi rarement de gros postes directs d'émissions : c'est l'échelle de déploiement (attendue ou réalisée) qui structure les

infrastructures nationales et mondiales dans leur dimensionnement et leurs impacts et les effets systémiques sont rarement évalués.

Dès lors, comment éclairer ses choix pour contribuer à alimenter des trajectoires énergétiquement et climatiquement soutenables, en tant que partie prenante de la conception et du déploiement de services s'appuyant sur l'IA ?

La "Boussole de l'IA" présentée dans ce rapport et la méthode qui l'accompagne ont pour objet de permettre une discussion éclairée entre trois types de parties prenantes du service d'IA considéré :

- Les sphères stratégiques et décisionnelles, qui pilotent les grandes orientations technologiques (fonctionnalités souhaitées, modèle et conditions de rentabilité ou pertinence du service, scénarios de déploiement/mise sur le marché/achat etc.) ;
- Les sphères de conception technique, qui effectuent les choix technologiques visant à répondre aux grandes orientations ;
- Les sphères techniques de l'impact environnemental, capables de mener l'analyse de cycle de vie, de mener l'étude d'impact quantitative et/ou l'analyse de compatibilité avec la trajectoire sectorielle ou de l'organisation.

La Boussole permet en effet de faire apparaître qualitativement les liens entre les choix fonctionnels d'un service et les dynamiques que son déploiement va induire sur les infrastructures numériques qu'il sollicite. Elle propose une grille de lecture émergeant de l'application au cas spécifique des applications "IA" des quatre grands déterminants de l'impact d'une technologie numérique :

La rentabilité ou pertinence de la solution dépend-elle de...

- ... la généralisation d'un nouvel équipement (capteur, terminal utilisateur etc.) ?
- ... la multiplication des capacités de calcul disponibles ?
- ... l'augmentation des volumes de données valorisables et/ou stockées ?
- ... l'augmentation des capacités réseaux (bande passante, couverture, latence) ?

C. Quels sont les facteurs inflationnistes

L'objectif des travaux qui suivront ce rapport intermédiaire sera de construire, consolider et déployer une méthodologie, la « Boussole de l'IA », visant à permettre de relier les choix de fonctionnalités faits lors de la conception et/ou du déploiement, avec les trajectoires d'impacts carbone-énergie dans lesquelles ces choix peuvent s'inscrire. Toutes vos remarques, critiques, propositions sur les orientations prises à ce stade, les questions et critères identifiés et toute piste d'entretiens pertinents à mener seront un ingrédient clé pour les suites de ces travaux !

D. Vision systémique : sortir de la vision sectorielle

Dans le cadre de ce rapport intermédiaire, les analyses présentées sont principalement centrées sur la méthodologie permettant de replacer les services d'IA et leurs fonctionnalités dans une vision systémique.

Ces analyses devront cependant être complétées, dans le cadre des travaux qui mèneront au rapport final, par une recontextualisation de ces conclusions au sein d'une vision systémique :

- Au niveau du système numérique en tant que tel, les dynamiques alimentées par les services étudiés créant des effets indirects sur les différents tiers du numérique (intensification des besoins en ressources pour maintenir et sécuriser une infrastructure numérique complexifiée, impacts des paradigmes et performances des algorithmes de recommandation sur les volumes de contenus consommés et donc les dimensionnements infrastructurels en découlant) ;
- Au niveau de l'économie dans son ensemble, les impacts des solutions numériques influant sur les trajectoires suivies par les autres secteurs (mobilité, bâtiment, industrie etc.), directement (surcoût carbone découlant d'une nouvelle couche de connectivité dans un bâtiment ou une infrastructure par exemple) ou indirectement (effet rebond engendré par une solution d'IA sur les usages de mobilité longue par exemple).

III. Quantification sur un cas d'usage

Dans le cadre de ce rapport intermédiaire, seule l'analyse fonctionnelle et qualitative a été réalisée.

Prolonger cette analyse avec une étude quantitative de l'un des cas d'usage sélectionné fait partie des objectifs du travail à réaliser dans le cadre de la réalisation du rapport final.

Toutes vos remarques, critiques, propositions sur les orientations prises à ce stade, les questions et critères identifiés et toute piste ou proposition de partage de retours d'expérience ou d'analyses menées sur des cas réels seront un ingrédient clé pour les suites de ces travaux !

IV. Pour des approches “bilan net” et systémique

Les analyses n'ayant pas encore été intégralement réalisées dans le cadre de ce rapport intermédiaire, les conclusions consolidées sur les points de vigilance quant aux conditions de pertinence des applications "IA" restent à construire.

Le cahier des charges d'une approche rigoureuse et au bon niveau de l'analyse de la pertinence d'une solution numérique reste cependant bien valable pour l'intelligence artificielle et ses applications. La pertinence énergie-climat d'une technologie numérique doit être :

- **Systematique** : les conditions de pertinence d'une solution dépendent en grande partie des conditions et choix de déploiement. Il ne peut donc y avoir de réponse générique à la question de la pertinence d'une solution numérique du point de vue énergie-carbone : celle-ci doit être effectuée dans chaque contexte spécifique avant de valider son déploiement.
- **Exhaustive, en "bilan net"** : la pertinence énergie-climat d'une solution numérique ne peut s'évaluer qu'en prenant en compte tous les impacts liés à son cycle de vie, y compris la phase de production des supports physiques qu'elles nécessitent (terminaux, réseaux, serveurs et centres de données). Un bilan qui ne prendrait en compte que les impacts directs n'apporte pas d'indication d'intérêt du point de vue des enjeux physiques.
- **Systemique** : les effets et impacts pris en compte dans l'évaluation de la pertinence d'une solution doivent intégrer les effets indirects (effets rebonds, transformation des usages, des besoins de maintenance etc.). Se passer de cette analyse, c'est priver la prise de décision d'un outil permettant de piloter les choix de déploiement en anticipant les effets délétères qui diluent sinon les apports énergie-climat initialement prévus.

Cette approche permet de poser les questions de la pertinence énergie-carbone d'une solution, jusqu'à la pertinence même du recours au numérique pour répondre au besoin.

Appliquée à l'IA, cette démarche rejoint certains éléments de la définition de "l'IA frugale" formulée par l'AFNOR (AFNOR, 2024) :

« Un service frugal d'IA est donc un service pour lequel :

- La nécessité de recourir à un système d'IA plutôt qu'à une autre solution moins consommatrice pour répondre au même objectif a été démontrée ;
- De bonnes pratiques (...) sont adoptées par le producteur, le fournisseur et le client d'IA pour diminuer les impacts environnementaux du service utilisant un algorithme d'IA ;
- Les usages et les besoins visent à rester dans les limites planétaires et ont été préalablement questionnés. »

Recommandations

Au stade de ce rapport intermédiaire, nous n'avons pas établi de liste de recommandations.

Nous pouvons en revanche déjà dire quelques mots sur les besoins en recommandations pour les sujets suivants :

- *La dérive climatique induite par les émissions de GES des centres de données nécessiterait un encadrement puisque les émissions totales ne cessent d'augmenter, s'éloignant de l'engagement SBTi du secteur et amplifiant les conséquences du changement climatique ;*
- *La consommation électrique des centres de données nécessitera une planification, celle-ci étant dimensionnante au premier ordre, en ce qu'elle peut fragiliser les transitions énergétiques et carbone, sans pour autant être organisée ;*
- *Éclairer les avancées à venir sur l'IA générative semble primordial. En particulier, la transparence de la part des fournisseurs de services et le suivi des données d'impact et de consommation des solutions et centres de données est un ingrédient préliminaire et indispensable à établir une direction systémique cohérente pour les infrastructures numériques et les systèmes dans lesquels elles s'intègrent.*

Enfin, nos recommandations seront classées en 4 catégories, qui restent valables pour les différentes composantes et les différents usages numérique (The Shift Project, 2023) :

- *Mesure & transparence,*
- *Optimisation & écoconception systémique,*
- *Réorganisation collective vers la sobriété,*
- *Compétences et formation.*

Conclusions

Ce rapport étant un rapport intermédiaire, sa production intervient à mi-chemin des travaux de recherche et de réflexion engagés depuis le mois de septembre 2024. Les conclusions que nous pouvons en tirer sont donc, par nature, provisoires et partielles.

En outre, le manque de transparence des acteurs dominants du domaine de l'IA, le manque de disponibilité des données ainsi que la documentation plus que perfectible du parc existant de centres de données et des nouveaux projets, notamment en France et en Europe, rend difficile toute évaluation quantitative des impacts énergie-carbone du « phénomène IA générative », même en ce qui concerne les impacts directs.

Il semble malgré tout possible – et nécessaire compte tenu de la vitesse de propagation du « phénomène IA générative » – de livrer d'ores et déjà un certain nombre de constats, dont certains sont aussi des alertes :

- Alors que l'empreinte énergie-carbone du numérique augmentait d'ores et déjà à un rythme soutenu avant l'apparition il y a deux ans du « phénomène IA générative », celui-ci aggrave considérablement la tendance. L'utilisation à grande échelle et de façon indifférenciée de l'IA générative et de grands modèles de langage pour en faire une technologie d'application générale⁵⁸ y joue un rôle central. Il est donc urgent de **déconstruire un premier mythe** selon lequel l'IA serait forcément de l'IA générative s'appuyant sur des modèles de fondation gigantesques et absolument polyvalents.
- Au-delà de l'effet néfaste sur l'empreinte propre du numérique, le « phénomène IA générative » est constitutif d'un risque important – clairement avéré aux Etats-Unis – de déséquilibre du système électrique. Risque d'ailleurs identifié par l'Agence internationale de l'Energie : **le rythme de croissance de la consommation électrique des centres de données devient beaucoup trop rapide pour y répondre de façon optimale, que ce soit en termes de transport ou de production d'électricité**. Les conséquences attendues vont des conflits d'usage avec d'autres secteurs (transports, logement, voire industrie) à l'impossibilité de tenir les trajectoires de décarbonation de l'électricité. Il faut donc renverser la logique et identifier au moins en France et en Europe un **budget énergie-carbone (et une trajectoire pluriannuelle)** disponible pour le numérique (dont l'IA) au sein duquel peuvent s'élaborer les stratégies des acteurs sans craindre de neutraliser nos objectifs communs.
- Il existe des solutions numériques décarbonantes (« *green digital solutions* ») au sens défini par exemple au sein de la European Green Digital Coalition (EGDC, 2024) et il existe donc aussi des cas d'usage où le recours à l'IA peut permettre de réduire les impacts environnementaux d'un système – sous certaines conditions contextuelles et une fois explicitée la trajectoire de référence à laquelle se comparer. En première analyse, il semble toutefois que l'IA utilisée dans ces cas d'usage soit rarement de l'IA générative. Dans ces

⁵⁸ « General Purpose Technology »

conditions il est à priori injustifié de prétendre que les coûts environnementaux directs du « phénomène IA générative », bien réels et liés au recours à l'IA générative, seraient surcompensés par des gains indirects qui, même lorsqu'ils se matérialisent, ne sont pas dus à l'IA générative : **rien ne le démontre à ce stade.**

- L'IA - et le numérique dans son ensemble - agit comme un catalyseur. Y avoir recours sans discernement au sein de modes de production et de consommation linéaires et reposant encore substantiellement sur l'utilisation d'énergies fossiles risque d'aboutir à une augmentation des émissions plutôt qu'à une baisse. Le « référentiel général pour l'**IA frugale** » développé par l'AFNOR constitue une première réponse en ce qu'il conduit à questionner dans un premier temps le recours à une solution IA puis à répondre au besoin en mobilisant le moins de ressources techniques possibles. La « Boussole de l'IA » que nous cherchons à développer ambitionne d'éclairer le plus en amont possible les enjeux relatifs à des cas d'usage.

Il y a urgence à éclairer avec des données objectives la problématique énergie-climat du « phénomène IA générative ». Ce rapport montre que les risques environnementaux qu'il recèle sont avérés, réels et substantiels. Il doit donc faire l'objet d'un débat sociétal informé et ceci d'autant plus en France et en Europe où la réduction de la dépendance aux énergies fossiles se double d'enjeux de souveraineté criants. Nous espérons que le travail amorcé dans le cadre de ce rapport intermédiaire y contribuera, tout en mobilisant les expertises et le temps des parties prenantes du secteur afin d'aboutir à une analyse et des recommandations consolidées dans la suite des travaux à venir.

Références

- A100. (s. d.). <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/a100/pdf/nvidia-a100-datasheet-nvidia-us-2188504-web.pdf>
- ADEME. (2021). *Transition(s) 2050. Choisir maintenant. Agir pour le climat.* <https://librairie.ademe.fr/societe-et-politiques-publiques/5072-prospective-transitions-2050-rapport.html>
- ADEME. (2023). *Evaluation environnementale des Services d'hébergement informatique en centre de données et de Services Cloud.* <https://librairie.ademe.fr/produire-autrement/6031-evaluation-environnementale-des-services-d-hebergement-informatique-en-centre-de-donnees-et-de-services-cloud.html>
- ADEME. (2025). *Numérique & environnement: Entre opportunités et nécessaire sobriété.* <https://librairie.ademe.fr/consommer-autrement/7883-avis-de-l-ademe-numerique-environnement-entre-opportunités-et-nécessaire-sobriété.html>
- ADEME, & Arcep. (2023). *Evaluation de l'impact environnemental du numérique en France et analyse prospective—Analyse prospective à 2030 et 2050 (3/3).* https://www.arcep.fr/uploads/tx_gspublication/etude-prospective-2030-2050_mars2023.pdf
- ADEME, négaWatt, The Shift Project. (2025). *Site comprendre2050.fr, décryptage "Quelle évolution de la consommation électrique pour une France bas carbone? "* <https://comprendre2050.fr/>
- AFNOR. (2024). *Référentiel général pour l'IA frugale, Un référentiel pour mesurer et réduire l'impact environnemental de l'IA, Afnor Spec 2314.* <https://www.afnor.org/actualites/referentiel-pour-mesurer-et-reduire-impact-environnemental-de-ia/>

- Ahmed M. U., Bollen M. H. J., & Alvarez M. (2021). *A Review of Data Centers Energy Consumption and Reliability Modeling*. <https://ieeexplore.ieee.org/document/9599719>
- Arcep. (2023). *Enquête annuelle « Pour un numérique soutenable »* [2ème édition - année 2021]. <https://www.arcep.fr/cartes-et-donnees/nos-publications-chiffrees/impact-environnemental/derniers-chiffres.html>
- BCG. (2024). *Accelerating compute needs underpin Southeast Asia's rapid data center growth*. <https://www.bcg.com/publications/2024/southeast-asia-accelerating-compute-needs-underpin-southeast-asias-rapid-data-center-growth>
- Bol, D., Pirson, T., & Dekimpe, R. (2020). *Moore's Law and ICT Innovation in the Anthropocene* (Electronic Circuits and Systems group, ICTEAM Institute, Université catholique de Louvain.).
- Carbone 4. (2024). *Net Zero Initiative for IT. Understanding the role of digital solutions in the global net zero effort*. <https://www.carbone4.com/publication-nzi-it>, p. 21
- Carnino G., & Marquet C. (2022). *Cooling, quick fix et spaghetti cloud dans l'univers du datacenter*. <https://journals.openedition.org/artefact/13419>
- Central Statistics Office (CSO). (2024). *Data Centres Metered Electricity Consumption 2023*. <https://www.cso.ie/en/releasesandpublications/ep/p-dcmec/datacentresmeteredelectricityconsumption2023/keyfindings/>
- Challenges. (2024, novembre 21). *Le grand mirage financier de l'IA : les investisseurs s'impatientent*. https://www.challenges.fr/entreprise/tech-numerique/le-grand-mirage-financier-de-l-ia-les-investisseurs-s-impatientent_912087
- Commission de l'intelligence artificielle, & Gouvernement français. (2024). *IA : Notre ambition pour la France*. <https://www.gouvernement.fr/upload/media/content/0001/09/4d3cc456dd2f5b9d79ee75fee63b47f10d75158.pdf>
- Data Center Dynamics. (2024a, juillet 22). *Revealing full data center environmental faces thanks to life cycle analysis*. <https://www.datacenterdynamics.com/en/opinions/revealing-full-data-center-environmental-faces-thanks-to-life-cycle-analysis/>

- Data Center Dynamics. (2024b, août 27). *Elon Musk's xAI data center 'adding to Memphis air quality problems'—Campaign group*. <https://www.datacenterdynamics.com/en/news/elon-musks-xai-data-center-adding-to-memphis-air-quality-problems-campaign-group/>
- Data Center Dynamics. (2024c, septembre 11). *Scala AI City : Scala pitches \$50bn Brazil data center campus of up to 4.7GW*. <https://www.datacenterdynamics.com/en/news/scala-ai-city-scala-pitches-50bn-brazil-data-center-campus-of-up-to-47gw/>
- Data Center Dynamics. (2024d, octobre 18). *Dominion Energy outlines long-term strategy for Virginia's power infrastructure*. <https://www.datacenterdynamics.com/en/news/dominion-energy-outlines-long-term-strategy-for-virginia-power-infrastructure/>
- Data Center Dynamics. (2025, janvier 30). *Southeastern US utilities to add 20GW of natural gas by 2040 to meet data center demand—Report*. <https://www.datacenterdynamics.com/en/news/southeastern-us-utilities-plan-more-than-20gw-of-new-natural-gas-capacity-by-2040-to-meet-unprecedented-data-center-demand-report/>
- DCByte. (2024). *Global Data Centre Index*. <https://www.dcbyte.com/global-data-centre-index/2024-global-data-centre-index/>
- DCMag. (2025a, février 9). *La carte régionale des 35 sites de data centers dédiés à l'IA identifiés par le gouvernement*. <https://dcmag.fr/la-carte-regionale-des-35-sites-de-data-centers-dedies-a-lia-identifies-par-le-gouvernement/>
- DCMag. (2025b, février 13). *EDF confirme : 4 + 2 sites de data centers pré-identifiés sur son foncier et 2 GW*. <https://dcmag.fr/edf-confirme-4-2-sites-de-data-centers-pre-identifies-sur-son-foncier-et-2-gw/>
- de Vries A. (2023). *The growing energy footprint of artificial intelligence*. [https://www.cell.com/joule/pdf/S2542-4351\(23\)00365-3.pdf](https://www.cell.com/joule/pdf/S2542-4351(23)00365-3.pdf)
- Deloitte. (2024). *Powering artificial intelligence A study of AI's environmental footprint—Today and tomorrow*. <https://www.deloitte.com/global/en/issues/climate/powering-ai.html>
- Dodd N., Alfieri, F., Gama Caldas M., & Maya-Drysdale, L. (2020). *Development of the EU Green Public Procurement (GPP) criteria for data centres, server rooms and cloud services—*

- Final technical report.* <https://op.europa.eu/en/publication-detail/-/publication/89971797-a9fa-11ea-bb7a-01aa75ed71a1/language-en>
- EDNA, Technology collaboration programme by IEA. (2023). *Policies for Data Centre Energy Efficiency: Scope, Trends and Availability of Data.* <https://www.iea-4e.org/wp-content/uploads/publications/2023/02/DC-Workstream-activity-1-publication-version.pdf>
- EGDC. (2024). *Net Carbon Impact Assessment Methodology for ICT Solutions.* <https://www.greenigitalcoalition.eu/net-carbon-impact-assessment-methodology-for-ict-solutions/>
- Elysée. (2025). *Sommet pour l'action sur l'Intelligence Artificielle.* <https://www.elysee.fr/sommet-pour-l-action-sur-l-ia>
- Epoch AI. (s. d.-a). *Can AI Scaling Continue Through 2030?* Consulté 12 février 2025, à l'adresse <https://epoch.ai/blog/can-ai-scaling-continue-through-2030>
- Epoch AI. (s. d.-b). *Key Trends and Figures in Machine Learning. Published online at epoch.ai.* Consulté 26 février 2025, à l'adresse <https://epoch.ai/trends>
- Epoch AI. (s. d.-c). *The computational performance of machine learning hardware has doubled every 2.7 years.* Consulté 17 janvier 2025, à l'adresse <https://epoch.ai/data/machine-learning-hardware#data-insights>
- Ericsson. (2024). *Ericsson Mobility Report—Novembrer 2024.* <https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/november-2024>
- Ericsson. (2025). *Ericsson Mobility Visualizer.* <https://www.ericsson.com/en/reports-and-papers/mobility-report/mobility-visualizer?f=1&ft=2&r=2,3,4,5,6,7,8,9&t=1,2,3,4,5,6,7&s=4&u=1&y=2024,2030&c=3>
- European Commission. (2020). *Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market—Final study report.* <https://digital-strategy.ec.europa.eu/en/library/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market>
- European Commission. (2024, mars 15). *Commission adopts EU-wide scheme for rating sustainability of data centres.* https://energy.ec.europa.eu/news/commission-adopts-eu-wide-scheme-rating-sustainability-data-centres-2024-03-15_en

- European Commission, Kamiya G., Bertoldi P. (2024). *Energy Consumption in Data Centres and Broadband Communication Networks in the EU*.
<https://publications.jrc.ec.europa.eu/repository/handle/JRC135926>
- Financial Times. (2024, juillet 4). *We're just honest': FirstEnergy CEO on coal plant retirement delays*. <https://www.ft.com/content/052fc091-3720-4d8d-9493-2a16d100220c>
- France Datacenter, & EY Parthenon. (2024). *Étude d'impact économique, social et environnemental de la filière des datacenters en France*.
<https://www.francedatacenter.com/ressources/>
- FS blog. (2024, août 10). *Guide Complet sur les Classifications et Catégories des Centres de Données*. <https://community.fs.com/fr/article/comprehensive-guide-to-data-center-classifications-and-categories.html>
- Gartner. (2024, novembre 12). *Gartner Predicts Power Shortages Will Restrict 40% of AI Data Centers By 2027*. <https://www.gartner.com/en/newsroom/press-releases/2024-11-12-gartner-predicts-power-shortages-will-restrict-40-percent-of-ai-data-centers-by-20270>
- GB200 Superchip. (s. d.). https://cdn.prod.website-files.com/61dda201f29b7efc52c5fbaf/6602ea9d0ce8cb73fb6de87f_nvidia-blackwell-architecture-technical-brief.pdf
- Gnibga W., Blavette A., & Orgerie A. (2023). *Renewable Energy in Data Centers : The Dilemma of Electrical Grid Dependency and Autonomy Costs*. <https://hal.science/hal-04189173v1/document>
- Goldman Sachs. (2024, mai 14). *AI is poised to drive 160% increase in data center power demand*. <https://www.goldmansachs.com/insights/articles/AI-poised-to-drive-160-increase-in-power-demand>
- Google. (2024). *Environmental report 2024*.
<https://www.gstatic.com/gumdrop/sustainability/google-2024-environmental-report.pdf>
- GreenIT.fr. (2019). *Empreinte environnementale du numérique mondial*. GreenIT.fr.
<https://www.greenit.fr/empreinte-environnementale-du-numerique-mondial/>
- H100. (s. d.). <https://resources.nvidia.com/en-us-tensor-core/nvidia-tensor-core-gpu-datasheet?ncid=no-ncid>

- HCC. (2020). *Maîtriser l'impact carbone de la 5G* [Haut conseil pour le climat].
<https://www.hautconseilclimat.fr/publications/maitriser-limpact-carbone-de-la-5g/>
- Hintemann R., & Hinterholzer S. (2020). *Borderstep 2020—Energy consumption of data centers worldwide*.
https://www.researchgate.net/publication/341427004_Energy_consumption_of_data_centers_worldwide_How_will_the_Internet_become_green
- Hubblo, & Fourboul E. (2023). *Impacts importés des datacenters : L'angle mort des analyses territoriales des impacts du numérique*. <https://hubblo.org/fr/blog/datacenters-imported-impacts/>
- IEA. (2019). *Data centres and energy – from global headlines to local headaches?*
<https://www.iea.org/commentaries/data-centres-and-energy-from-global-headlines-to-local-headaches>
- IEA. (2021a). *Global CO2 emissions from transport by subsector, 2000-2030*.
<https://www.iea.org/data-and-statistics/charts/global-co2-emissions-from-transport-by-subsector-2000-2030>
- IEA. (2021b, février 15). *Data Centres and Data Transmission Networks—Tracking report*.
<https://www.iea.org/reports/data-centres-and-data-transmission-networks>
- IEA. (2022). *Data Centres and Data Transmission Networks—Tracking report*. International Energy Agency. <https://www.iea.org/reports/data-centres-and-data-transmission-networks>
- IEA. (2024a). *Electricity 2024—Analysis and forecast to 2026*.
<https://www.iea.org/reports/electricity-2024>
- IEA. (2024b). *World Energy Outlook 2024*. <https://www.iea.org/reports/world-energy-outlook-2024>
- IEA, & IEA Paris. (2024a). *Artificial intelligence model size and complexity, 2006-2022*.
<https://www.iea.org/data-and-statistics/charts/artificial-intelligence-model-size-and-complexity-2006-2022>

- IEA, & IEA Paris. (2024b). *Cost of graphics processing units, 2006-2022*. <https://www.iea.org/data-and-statistics/charts/cost-of-graphics-processing-units-2006-2022>
- IEA, & IEA Paris. (2024c). *Investment in data centres in the United States, January 2014 to August 2024*. <https://www.iea.org/data-and-statistics/charts/investment-in-data-centres-in-the-united-states-january-2014-to-august-2024>
- Intergovernmental Panel On Climate Change (Ippc) (Éd.). (2023). Summary for Policymakers. In *Climate Change 2022—Mitigation of Climate Change* (1^{re} éd., p. 3-48). Cambridge University Press. <https://doi.org/10.1017/9781009157926.001>
- ISO. (2021). *ISO/IEC 22237, Data Center Facilities and Infrastructures*. <https://www.iso.org/standard/78550.html>
- ITU-T. (2020). *Greenhouse gas emissions trajectories for the information and communication technology sector compatible with the UNFCCC Paris Agreement* [L.1470. Series L: environment and icts, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant]. <https://www.itu.int/rec/T-REC-L.1470-202001-I/en>
- Jefferies. (2024). *Data Centers & AI: Powering the Future*. <https://www.jefferies.com/insights/the-big-picture/how-data-centers-are-shaping-the-future-of-energy-consumption/>
- Journal officiel de l'Union européenne. (2023). *Directive (UE) 2023/1791 du Parlement européen et du Conseil du 13 septembre 2023 relative à l'efficacité énergétique et modifiant le règlement (UE) 2023/955 (refonte) (Texte présentant de l'intérêt pour l'EEE)*. <https://eur-lex.europa.eu/eli/dir/2023/1791/oj?locale=fr>
- Journal officiel de l'Union européenne. (2024). *Règlement délégué (UE) 2024/1364 de la Commission du 14 mars 2024 sur la première phase de la mise en place d'un système commun de notation des centres de données à l'échelle de l'Union*. https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:L_202401364
- LBNL, Lawrence Berkeley National Laboratory, Berkeley, California, Shehabi A., Smith, S.J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., & Koomey J., Masanet, E., Sartor, D. (2024). *2024 United States Data Center Energy Usage Report*. <https://eta->

publications.lbl.gov/sites/default/files/2024-12/lbnl-2024-united-states-data-center-energy-usage-report.pdf

Le Figaro. (2025, janvier 30). *Le patron d'EDF en opération séduction pour attirer les data centers.*

o <https://www.lefigaro.fr/economie/le-patron-d-edf-en-operation-seduction-pour-attirer-les-data-centers-20250130>

Le Monde. (2025a, février 6). *IA : les Emirats arabes unis annoncent la construction en France*

d'un data center géant. https://www.lemonde.fr/economie/article/2025/02/06/intelligence-artificielle-les-emirats-arabes-unis-construiront-en-france-un-data-center-geant_6535048_3234.html

Le Monde. (2025b, février 9). *Sommet de l'IA : le fonds canadien Brookfield va investir 20 milliards*

d'euros en France, notamment pour développer ses data centers. https://www.lemonde.fr/economie/article/2025/02/09/sommet-de-l-ia-le-fonds-canadien-brookfield-va-investir-20-milliards-d-euros-en-france-notamment-pour-developper-ses-data-centers_6538362_3234.html

Le Monde, & Alexandre Piquard. (2024, septembre 23). *Les géants du numérique se*

convertissent au nucléaire pour éteindre les besoins énergétiques toujours plus importants de l'IA. https://www.lemonde.fr/economie/article/2024/09/23/les-geants-du-numerique-se-convertissent-au-nucleaire_6329378_3234.html

Legifrance. (2022, août 18). *Code des impositions sur les biens et services.*

https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000046196792

LeMagIT, & Raoul G. (2025, février 17). *Datacenters IA en France : Promesses contre*

promesses. <https://www.lemagit.fr/actualites/366619359/Datacenters-IA-en-France-promesses-contre-promesses>

L'institut Paris Région. (2023). *Le développement des data centers en Ile-de-France. Éléments*

pour une stratégie régionale et territoriale. <https://www.institutparisregion.fr/nos-travaux/publications/le-developpement-des-data-centers-en-ile-de-france/>

Luccioni A. S., Strubell E., & Crawford K. (2025). *From Efficiency Gains to Rebound Effects : The*

Problem of Jevons' Paradox in AI's Polarized Environmental Debate. <https://arxiv.org/pdf/2501.16548>

- L'usine digitale. (2025, février 12). *Sommet IA : Fluidstack veut construire un supercalculateur décarboné de 1 GW d'ici 2028*. <https://www.usine-digitale.fr/article/sommet-ia-fluidstack-veut-construire-un-supercalculateur-decarbone-de-1-gw-d-ici-2028.N2227340>
- L'usine digitale, & Seramour C. (2025, février 21). *ChatGPT passe la barre des 400 millions d'utilisateurs hebdomadaires*. <https://www.usine-digitale.fr/article/chatgpt-passe-la-barre-des-400-millions-d-utilisateurs-hebdomadaires.N2227779>
- Malmodin J. et al. (2023). *ICT Sector Electricity Consumption and Greenhouse Gas Emissions—2020 Outcome*. <https://dx.doi.org/10.2139/ssrn.4424264>
- Masanet E., Shehabi A., Lei N., Smith S., & Koomey J. (2020). *Recalibrating global data center energy-use estimates*. https://datacenters.lbl.gov/sites/default/files/Masanet_et_al_Science_2020.full_.pdf
- McKinsey. (2024). *AI power: Expanding data center capacity to meet growing demand*. <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand>
- McKinsey & Company. (2024). *The role of power in unlocking the European AI revolution*. <https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/the-role-of-power-in-unlocking-the-european-ai-revolution>
- Microsoft. (2024). *How can we advance sustainability? 2024 Environmental Sustainability Report*. <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1IMjE>
- Montevecchi, F., Stickler, T., Hintemann, R., & Hinterholzer, S. (2020). *Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market : Final study report*. <https://data.europa.eu/doi/10.2759/3320>
- Ni W., Hu X., Du H., Kang Y., Ju Y., & Wang Q. (2024). *CO2 emission-mitigation pathways for China's data centers*. <https://www.sciencedirect.com/science/article/abs/pii/S0921344923005177?via%3Dihub>
- Nokia. (2024). *Global network traffic report*. <https://onestore.nokia.com/asset/213660>
- Pirson T., Bol D. (2021). *Assessing the embodied carbon footprint of IoT edge devices with a bottom-up life-cycle approach*. <https://arxiv.org/abs/2105.02082>

- Prof. Hannah Daly, University College Cork. (2024). *Data centres in the context of Ireland's carbon budgets*.
https://www.friendsoftheearth.ie/assets/files/pdf/data_centres_and_the_carbon_budgets_-_prof_hannah_daly_dec_2024.pdf
- RTE. (2022). *Futurs énergétiques 2050—Rapport complet* [Février 2022]. <https://www.rte-france.com/analyses-tendances-et-prospectives/bilan-previsionnel-2050-futurs-energetiques>
- RTE. (2023). *Bilan prévisionnel—Édition 2023—Futurs énergétiques 2050—2023-2035 : Première étape vers la neutralité carbon*. <https://assets.rte-france.com/analyse-et-donnees/2024-07/Bilan%20previsionnel%202023%20Chapitre%20%20Consommation.pdf>
- RTE. (2024). *Schéma décennal de développement du réseau—Edition 2024—Consultation publique—Hypothèses et analyses technico-économiques—Document B*. <https://assets.rte-france.com/prod/public/2024-03/SDDR2024-volet-etudes-technico-%C3%A9conomiques-doc-B.pdf>
- RTE. (2025). *Schéma décennal de développement du réseau. Synthèse*. https://assets.rte-france.com/prod/public/2025-02/RTE_SDDR2025_Synthese_2.pdf
- Sagot B. (2023). *Apprendre les langues aux machines*. <https://youtu.be/uPiD8SFv41A?feature=shared>
- SBTi, ITU, GeSI, & GSMA. (2020). *Guidance for ICT companies setting science based targets—Mobile networks operators, fixed networks operators and data centres operators*. <https://sciencebasedtargets.org/sectors/ict>
- Schneider Electric. (2023). *Quantifying Data Center Scope 3 GHG Emissions to Prioritize Reduction Efforts—White Paper* 99. https://www.se.com/ww/en/download/document/SPD_WP99_EN/
- Schroders. (2024). *Power-hungry AI applications are demanding a significant expansion in global energy capacity*. <https://www.schroders.com/en/global/individual/insights/power-hungry-ai-applications-are-demanding-a-significant-expansion-in-global-energy-capacity/>

- SemiAnalysis, Patel D., Nishball D., & Eliahou Ontiveros J. (2024, mars 13). *AI Datacenter Energy Dilemma – Race for AI Datacenter Space Gigawatt Dreams and Matroshka Brains Limited By Datacenters Not Chips*. <https://semianalysis.com/2024/03/13/ai-datacenter-energy-dilemma-race/#datacenter-math>
- Sénat. (2020). *Pour une transition numérique écologique. Rapport d'information n°555 (2019-2020)*. https://www.senat.fr/rap/r19-555/r19-555_mono.html
- Thales Group. (2024, avril 16). *Bots Now Make Up Nearly Half of All Internet Traffic Globally*. Bots Now Make Up Nearly Half of All Internet Traffic Globally
- The Journal. (2024, novembre 28). *Ireland's data centres turning to fossil fuels after maxing out country's electricity grid*. <https://www.thejournal.ie/investigates-data-centres-6554698-Nov2024/>
- The Register, & Brandon Vigliarolo. (2024, décembre 5). *Day after nuclear power vow, Meta announces largest-ever datacenter powered by fossil fuels*. https://www.theregister.com/2024/12/05/meta_largestever_datacenter/https://www.theregister.com/2024/12/05/meta_largestever_datacenter/
- The Shift Project. (2021). *Impact environnemental du numérique : Tendances à 5 ans et gouvernance de la 5G*. The Shift Project. <https://theshiftproject.org/article/impact-environnemental-du-numerique-5g-nouvelle-etude-du-shift/>
- The Shift Project. (2023). *Planifier la décarbonation du système numérique en France : Cahier des charges*. <https://theshiftproject.org/article/planifier-la-decarbonation-du-systeme-numerique-en-france-cahier-des-charges/>
- The Shift Project, & Supaero Decarbo. (2021). *Pouvoir voler en 2050, quelle aviation dans un monde contraint ?* <https://theshiftproject.org/article/quelle-aviation-dans-un-monde-contraint-nouveau-rapport-du-shift/>
- U4E, United for efficiency, & UN Environment Programme. (2025). *Centres de données et serveurs informatiques performants pour une transformation durable du marché*. https://united4efficiency.org/wp-content/uploads/2025/02/U4E-for-AI-Summit_FR_7Feb25-1.pdf

University of Cambridge. (2025). *Cambridge Bitcoin Electricity Consumption Index*.
<https://ccaf.io/cbnsi/cbeci>

Uptime Intelligence. (2023). *Uptime Institute Global Data Center Survey 2023*.
<https://uptimeinstitute.com/resources/research-and-reports/uptime-institute-global-data-center-survey-results-2023>

Uptime Intelligence. (2024). *Uptime Institute Global Data Center Survey 2024*.
<https://datacenter.uptimeinstitute.com/rs/711-RIA-145/images/2024.GlobalDataCenterSurvey.Report.pdf?version=0>

Vie publique. (2024, octobre 22). *Projet de loi de simplification de la vie économique*.
<https://www.vie-publique.fr/loi/293913-entreprises-projet-de-loi-de-simplification-de-la-vie-economique>

Washington Post. (2024, mars 7). *Amid explosive demand, America is running out of power*.
<https://www.washingtonpost.com/business/2024/03/07/ai-data-centers-power/>

Annexe 1 : Typologie de centres de données

Un centre de données est une structure dédiée à l'hébergement, l'interconnexion et l'exploitation centralisées des équipements de stockage, calcul et transport de données (ISO, 2021). En plus de ces services, ces infrastructures contiennent généralement des systèmes robustes de distribution d'énergie, de contrôle environnemental et de sécurité, qui garantissent la disponibilité des services qu'elle héberge.

Cette définition générale recouvre toutefois plusieurs types de centres de données qui varient par leur taille et leur fonction, allant de petit centre de données d'entreprise sur site, qui servent une seule organisation, aux larges centres de co-hébergement (ou colocation) et *hyperscale*, qui offrent des infrastructures et des services partagés pour plusieurs clients à très large échelle. Ces distinctions sont liées à des différences fonctionnelles, dépendant de (i) leur objectif (installations d'entreprise, de colocation, de co-hébergement ou d'opérateur de réseau), (ii) leur niveau de sécurité, (iii) leur taille physique, et (iv) leur mode d'hébergement (constructions mobiles, temporaires ou permanentes). Il est toutefois possible de distinguer quelques grandes catégories d'espaces :

- **Edge** : déploiements de micro-centres de données, en périphérie de réseau, géographiquement proche des utilisateurs finaux, afin de répondre à des demandes spécifiques, souvent de latence et/ou de sécurité.
- **Centres de données entreprise** : centres de données internes, gérés par des entreprises (ou des acteurs publics) et pour leurs usages, qui peuvent varier par leur taille (PME, branche ou grande entreprise).
- **Fournisseurs de services de télécommunication** : centres de données gérés par des entreprises de télécommunications pour soutenir les services internes nécessaires à la fourniture de leurs services réseau.
- **Colocation ou co-hébergement** : centres de données qui offrent aux entreprises la possibilité d'héberger leur matériel hors site, fournissant les services essentiels de gestion de l'alimentation électrique, de refroidissement, de la sécurité et de la mise en réseau⁵⁹.
- **Hyperscale** : Centres de données construits par des entreprises qui déploient des services et plateformes de services à très grande échelle.

Plus simplement, il est possible de distinguer entre centres de données **on-premise** (ou **on-prem**) lorsqu'ils sont situés sur le site de l'entreprise qui les utilise, en **colocation** lorsqu'un tiers gère le site d'hébergement, et **cloud** lorsque le site d'hébergement et les serveurs sont gérés par un tiers.

Bien que basées sur des distinctions d'usages, ces nuances vont en général de pair avec des considérations de superficie. Les petits centres de données ont en moyenne une superficie de 15 mètres carrés et correspondent aux centres de données edge, de PME ou de branche

⁵⁹ A noter que 61% des fournisseurs de services de colocation hébergent notamment des locataires hyperscale (figure 14 (Uptime Intelligence, 2024)).

entreprise. Les tailles intermédiaires incluent les centres de données internes (250 mètres carrés) et les centres de données de fournisseurs de services (650 mètres carrés). Enfin, sont considérés de grande échelle les espaces de colocation (pour une superficie moyenne de plus de 1000 mètres carrés) ou hyperscale (pour presque 2800 mètres carrés).⁶⁰

Depuis 2010, la répartition entre ces catégories a largement évolué en faveur des centres de colocation et hyperscale, passant de 10% à près de 80% de la part des serveurs installés en centre de données aux États-Unis en 15 ans. Cette évolution est due à l'utilisation étendue de technologies de virtualisation, qui offrent une grande flexibilité de déploiement, et qui répondent rapidement aux évolutions en capacité de calcul du marché, souvent à plus bas coût. Cette migration est d'autant plus avantageuse qu'elle permet un accès à du matériel spécialisé, tel que des GPUs modernes, nécessaires à l'entraînement et l'inférence d'algorithmes d'IA.

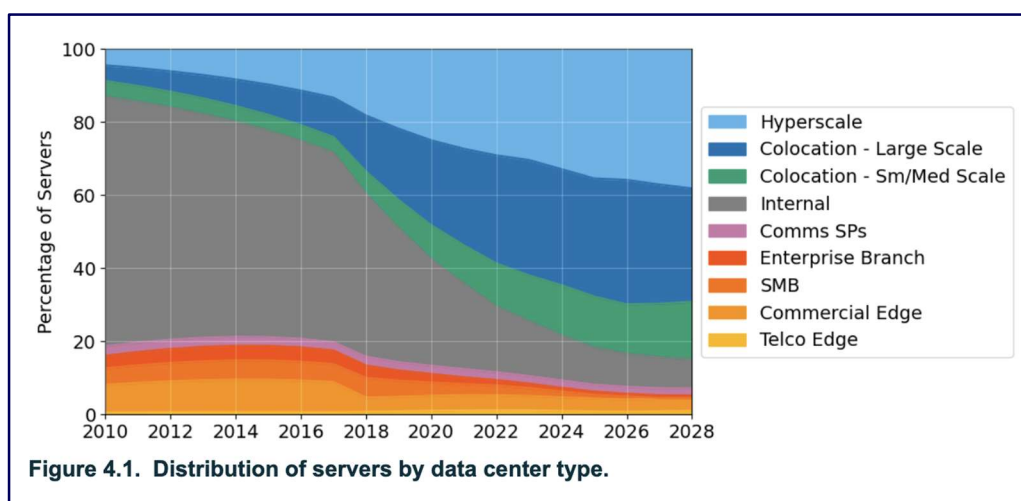


Figure 34 - Distribution des serveurs par type de centres de données. Source : (LBNL et al., 2024)

Ces proportions peuvent varier selon les zones géographiques, les périodes temporelles, ou l'indicateur regardé (nombre de serveurs ou consommation électrique). Par exemple, si on regarde la part de la consommation électrique des centres de données n'étant ni colocation ni hyperscale sur l'ensemble de la consommation électrique des centres de données : Masanet l'évaluait au niveau monde à 35% en 2018 (Masanet E. et al., 2020), l'ADEME l'évaluait au niveau France (territorial) à 52% en 2020 (ADEME & Arcep, 2023), elle est projetée au niveau monde évoluant de 21% en 2023 vers 5% en 2028 (LBNL et al., 2024) et au niveau Europe évoluant de 40% en 2023 vers 22% en 2028 (McKinsey & Company, 2024).

Dans son étude prospective, l'ADEME-Arcep a constitué des catégories plutôt par types d'acteurs : public local, public national, entreprises hors acteur du digital, cloud, HPC, edge. Vis-à-vis des catégories ci-dessus, cloud et HPC sont plutôt du type co-hébergement ou hyperscale, et les centres de données recensés dans public de même type que ceux des entreprises.

A l'échelle d'un pays ou d'une région, il est aussi possible de séparer :

- **Les centres de données installés sur son territoire** : vision inventaire national, permettant notamment d'identifier les besoins en ressources territoriales (énergie, foncier, eau) et les émissions de gaz à effets de serre territoriales,

⁶⁰ Shehabi, Arman, et al. "2024 United States Data Center Energy Usage Report." (2024)

- **Les capacités de stockage et de calcul associées à l'utilisation de services à l'étranger** : vision empreinte carbone, permettant d'identifier les dépendances et les impacts climatiques de nos usages.

Enfin, il est aussi possible de les classer par objectifs (polyvalents ou spécialisés), par modèles d'affaires (on-prem, colocation, cloud), et sur d'autres critères encore (performance, tolérance aux pannes) (FS blog, 2024).

Les centres de données ont pour principaux postes de consommation énergétique (i) les charges de ses équipements informatiques, (ii) celles des équipements de refroidissement et de contrôle environnemental, et (iii) le système de conditionnement de puissance interne, ainsi que les équipements de sécurité, d'éclairage et de bureaux. L'alimentation électrique d'un centre de données est soumise à des contraintes de continuité, avec des systèmes d'alimentation de secours, en cas de panne du système d'alimentation principal, afin de prévenir les interruptions de service.

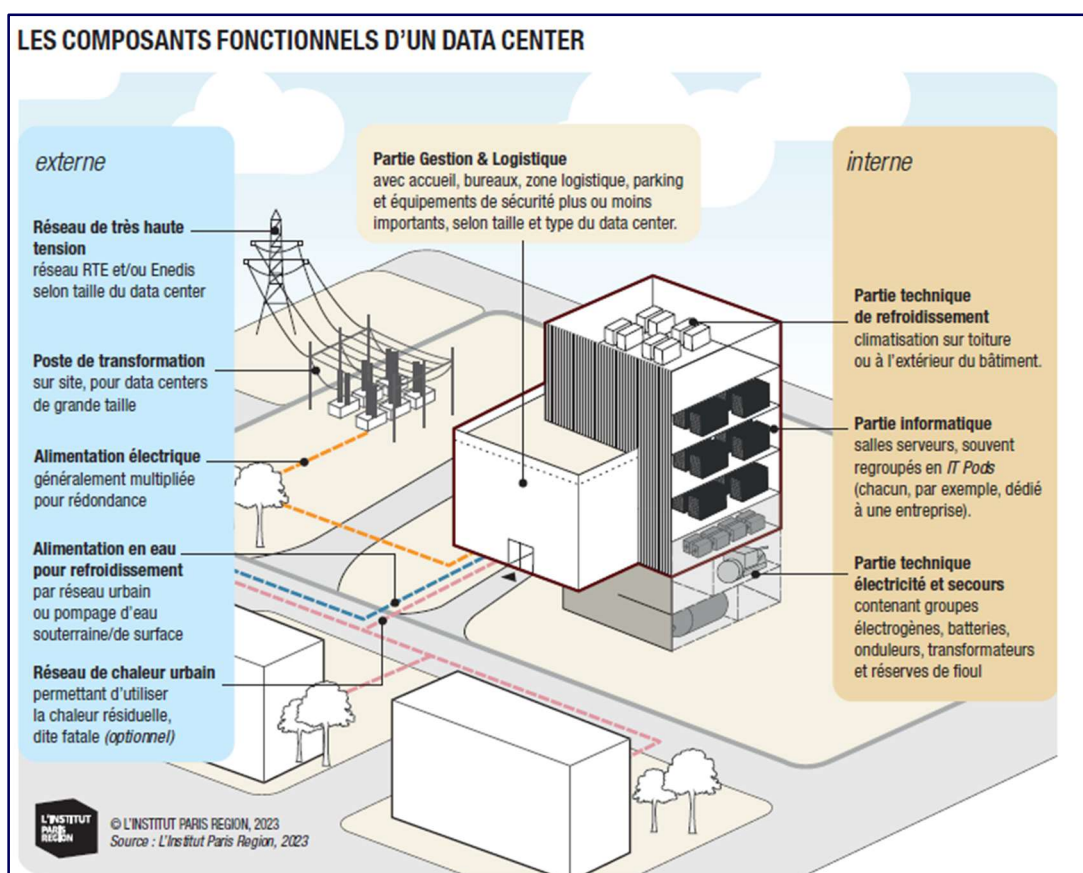


Figure 35 - Les composants fonctionnels d'un data center. Source : (L'institut Paris Région, 2023)

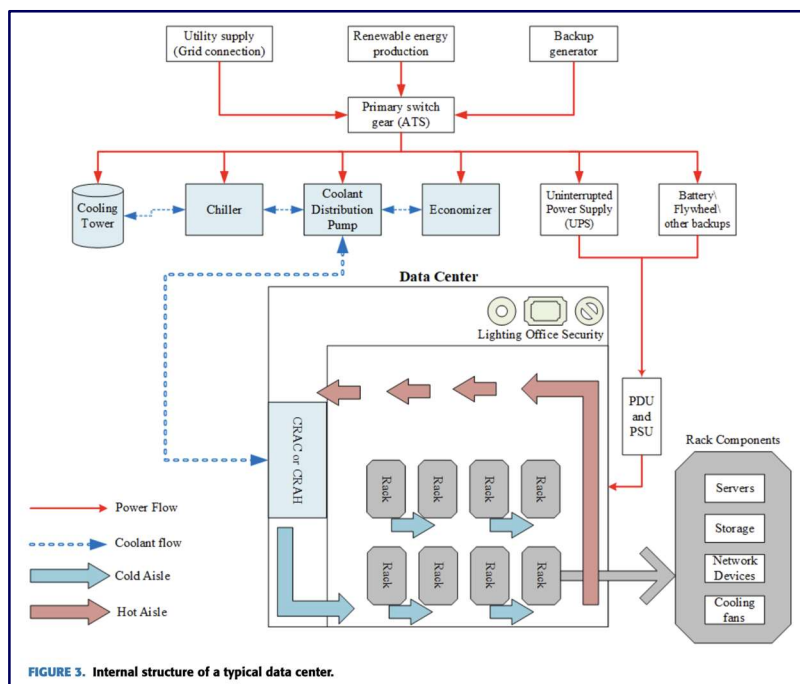


Figure 36 - Structure d'un centre de données. Source : (Ahmed M. U. et al., 2021)

La consommation d'un centre de données dépend de la consommation de ses équipements informatiques et de son système de refroidissement. Si bien que l'efficacité énergétique d'une installation est mesurée par son indicateur d'efficacité énergétique (souvent en anglais PUE pour *Power Usage Effectiveness*), défini comme la consommation énergétique totale du centre de données divisée par la consommation de ses équipements informatiques. Plus cet indice est proche de 1, plus le centre de données est efficace, car il utilise moins d'énergie pour ses fonctions *auxiliaires* (refroidissement, conditionnement de puissance, éclairage, etc.). En 2024, le PUE moyen mondial se situe autour de 1.5, avec toutefois des disparités régionales et temporelles. Ce facteur dépend principalement des systèmes de refroidissement utilisés (en lien avec la modernité des infrastructures), et du climat dans lequel se situe le centre de données (Ahmed, 2021).

Concernant le PUE, différents sujets sont à améliorer : plusieurs aspects nécessitent des améliorations : bien qu'une norme de calcul existe (ISO/IEC 30134-2:2016), la standardisation reste insuffisante, avec un manque de transparence concernant les hypothèses et les définitions utilisées. La prise en compte de la consommation "idle" notamment fait partie des sujets à améliorer. Enfin, comme pour tout indicateur d'efficacité, il est important de considérer que l'augmentation du volume vient artificiellement améliorer cet indicateur.

Les centres de données peuvent être équipés de sources d'énergie renouvelable pour compléter la puissance perçue du réseau local. Il y a des enjeux d'interconnexion avec la grille (Gnibga W. et al., 2023). Ils peuvent aussi souscrire à des contrats énergétiques plus verts pour améliorer leurs rapports d'émissions. Rappelons toutefois la différence entre les émissions basées sur la localisation du centre de données (*location-based*), qui suivent les moyennes du réseau électrique dans un lieu spécifié, et les émissions basées sur le marché (*market-based*), qui concernent l'énergie achetée sous contrat, comme l'énergie renouvelable provenant d'un fournisseur extérieur au principal fournisseur de services publics d'une région. Ces normes d'attribution peuvent être trompeuses dans les rapports d'émission, car les émissions market-based sont en général abaissées par des valeurs plus basses d'intensité carbone.

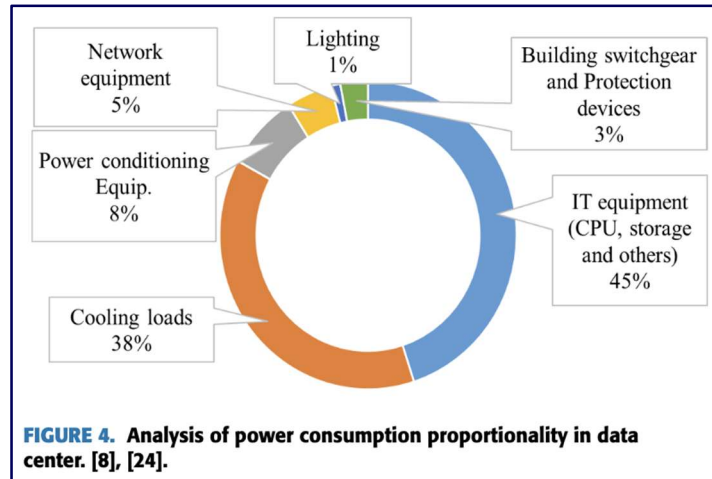


Figure 37 - Répartition de la consommation électrique d'un centre de données. Source : (Ahmed M. U. et al., 2021)

Enfin, en empreinte carbone, les répartitions peuvent être différentes, Schneider Electric en propose une décomposition pour un centre de données type⁶¹ (Schneider Electric, 2023)⁶², mettant en évidence la part majeure du scope 3, dont l'empreinte embarquée.

Step	Carbon focus	Findings	Proposed actions
1	Total carbon	<ul style="list-style-type: none"> Scope 3 represents 38-69% Electricity has a Scope 3 component 	<ul style="list-style-type: none"> Use more renewable/clean energy
2	Scope 3 emissions	<ul style="list-style-type: none"> Capital goods represents 46-71% Fuel- and energy-related activities represents 13-47% Core & shell represents a small percentage 	<ul style="list-style-type: none"> Purchase low carbon capital goods Use more renewable/clean energy
3	Embodied carbon	<ul style="list-style-type: none"> Manufacturing represents ~ 90% IT represents 57-83% Facility infrastructure represents 17-43% 	<ul style="list-style-type: none"> Extend server lifespans Design and operate for high utilization from IT to the facility Optimize IT demand
4	Facility infrastructure embodied carbon	<ul style="list-style-type: none"> Power system represents ~ 30% Cooling system represents ~ 30% Core & shell represents 8-15% 	<ul style="list-style-type: none"> Purchase efficient and low carbon products Reuse existing building for data centers instead of new construction
5	Sub-system embodied carbon	<ul style="list-style-type: none"> Concrete represents 85% LV switchgear represents ~ 30% VRLA battery represents 6-21% Air-cooled chiller represents ~50% 	<ul style="list-style-type: none"> Evaluate modular and prefabricated construction methods Purchase efficient and low carbon products

Figure 38 - Répartition des impacts du "Scope 3 Empreinte embarquée" d'un centre de données. Source : (Schneider Electric, 2023)

⁶¹ Dont les caractéristiques sont les suivantes : 1 MW, charge de 50%, 6 kW / rack, PUE de 1.34, facteur d'émission de l'électricité des Etats-Unis, 5 tCO₂e / kW d'IT en empreinte embarquée, 94 % de serveurs de stockage et 6 % d'équipements réseaux, peuplés à 50 %. Mais avec la possibilité de faire varier ces caractéristiques dans un simulateur en ligne <https://www.se.com/ww/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-lifecycle-co2e-calculator/>

⁶² Schneider Electric. (2023). Quantifying Data Center Scope 3 GHG Emissions to Prioritize Reduction Efforts—White Paper 99. https://www.se.com/ww/en/download/document/SPD_WP99_EN/

Pour flécher les achats durables dans les centres de données et les serveurs informatiques, U4E en concertation avec des acteurs comme Uptime Institute, France Datacenter, L'Alliance Française des Industries du Numérique, EQUINIX, GIMELEC, Google, Microsoft, LBNL, IBM Corporation établit des lignes directrices recommandant de suivre les indicateurs suivants (Figure 39), fournit des objectifs (Figure 40) et établit un score sur base de ces indicateurs (Figure 41) (U4E et al., 2025)⁶³.

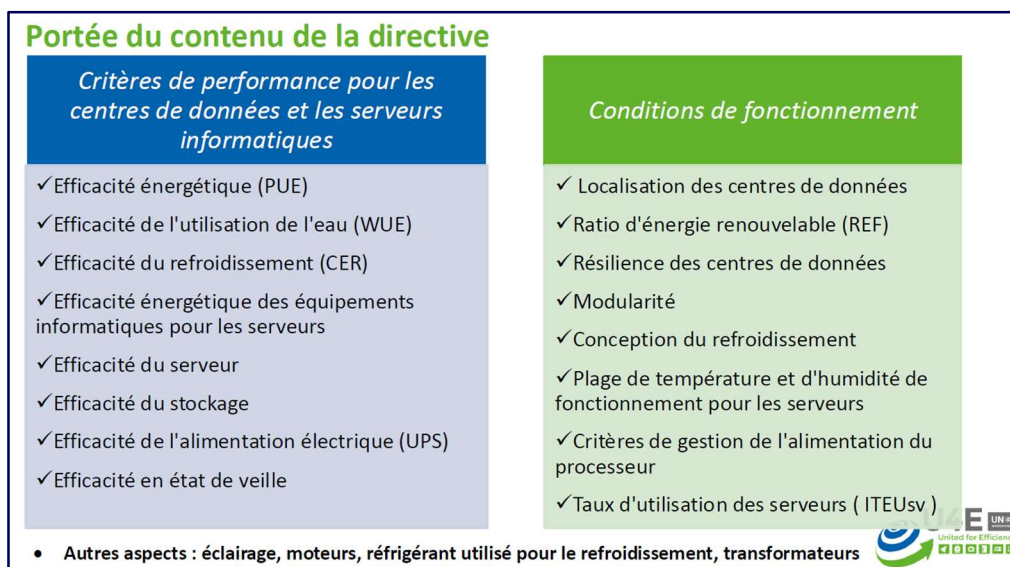


Figure 39 - Indicateurs de la directive U4E. Source : (U4E et al., 2025)

Valeurs des indicateurs clés (KPI)

		2025	2027	2029	2031
Centre de données de colocation existant pour héberger les données	PUE	≤ 1,5 HH : ≤ 1,7	≤ 1,4 HH : ≤ 1,6	≤ 1,3 HH : ≤ 1,5	≤ 1,2 HH : ≤ 1,4
	WUE	≤ 1,5 L/kWh	≤ 1 L/kWh	≤ 0,5 L/kWh	≤ 0,2 L/kWh
	RÉF	≥ 50 %	≥ 60 %	≥ 70 %	≥ 80 %
	CER	≥ 2,5	≥ 2,9	≥ 3,8	≥ 5,7
	ITEUsv	≥ 50 %	≥ 60 %	≥ 70 %	≥ 80 %
Nouveau bâtiment du centre de données - By design / après 3 ans de fonctionnement	PUE	≤ 1,4 / ≤ 1,5 HH : ≤ 1,6 / ≤ 1,7	≤ 1,3 / ≤ 1,4 HH : ≤ 1,5 / ≤ 1,6	≤ 1,2 / ≤ 1,3 HH : ≤ 1,4 / ≤ 1,5	≤ 1,1 / ≤ 1,2 HH : ≤ 1,3 / ≤ 1,4
	WUE	≤ 1,5 L/kWh	≤ 1 L/kWh	≤ 0,5 L/kWh	≤ 0,2 L/kWh
	RÉF	≥ 50 %	≥ 60 %	≥ 70 %	≥ 80 %
	CER	≥ 2,9 / ≥ 2,5	≥ 3,8 / ≥ 2,9	≥ 5,7 / ≥ 3,8	≥ 10 / ≥ 5,7
	ITEUsv après 3 ans	≥ 50 %	≥ 60 %	≥ 70 %	≥ 80 %

HH : Climat chaud et humide (zones climatiques ASHRAE 0A, 1A, 2A, 3A)




Figure 40 - Valeurs des indicateurs de la directive U4E. Source : (U4E et al., 2025)

⁶³ U4E, United for efficiency, & UN Environment Programme. (2025). *Centres de données et serveurs informatiques performants pour une transformation durable du marché*. https://united4efficiency.org/wp-content/uploads/2025/02/U4E-for-AI-Summit_FR_7Feb25-1.pdf

Indicateurs de performance clés	Points	Pondération
Gestion de l'énergie (PUE)	De 2 à 1,2 noté sur 5 points	30 %
Efficacité du refroidissement (CER)	De 2,5 à 10 noté sur 5 points	20 %
Consommation d'eau (WUE)	De 2 à 1,2 L/kWh noté sur 5 points	20 %
Ratio d'énergie renouvelable (REF)	De 50% à 90% noté sur 5 points	20 %
Taux d'utilisation des serveurs : taux d'utilisation des serveurs (ITEUsv)	De 30% à 70% noté sur 5 points	10 %




Figure 41 - Système de notation proposé dans la directive U4E. Source : (U4E et al., 2025)

Annexe 2 : Un premier calcul : “Quelle empreinte carbone pour les projets de centres de données mondiaux ?”

Cette annexe fait référence à l'onglet « Calculette GES à partir de TWh » dans le fichier « VERSION PROVISOIRE - Etat des lieux Monde.xlsx ».

Calcul :

Dans celui-ci, à partir d'une estimation en TWh, nous proposons de calculer les émissions de gaz à effets de serre. Ce premier calcul, très simple, se base sur :

- La prise en compte de l'intensité carbone de l'électricité,
- Un ratio pour estimer l'empreinte carbone embarquée des centres de données à partir de leur consommation d'électricité.

Hypothèses :

Le facteur d'émission de l'électricité considéré est le facteur d'émission monde : 460 gCO₂e/kWh en 2022 (Stated Policies) (IEA, 2024b) et 312 gCO₂e/kWh en 2030. Cette hypothèse pour 2030 est contestable puisque la croissance des centres de données vient justement demander une consommation électrique, souvent non anticipée, qui pourra être carbonée (figure 4.1 dans (IEA, 2024b)). Par ailleurs, sur chaque réseau de transport d'électricité, les facteurs d'émissions de l'électricité sont différents et chaque centre de données peut avoir un facteur d'émission affiché (grâce à l'achat de PPA) ou un facteur d'émission réel différent de celui du réseau. Ce choix d'une valeur moyenne monde permet de refléter l'impact des centres de données compte-tenu des capacités de production électriques mondiales actuelles.

Afin de prendre en compte l'empreinte carbone embarquée, nous avons pris comme hypothèse que l'empreinte carbone embarquée représentait 35% de l'empreinte carbone de l'utilisation à partir de quelques chiffres macro : ~26%⁶⁴ (Data Center Dynamics, 2024a), ~32% (Malmodin J. et al., 2023)⁶⁵, ~35%⁶⁶ (Schneider Electric, 2023).

⁶⁴ Chiffre de Lescuyer pour des centres de données en Europe en considérant une durée de vie de 20 ans.

⁶⁵ A partir de l'empreinte embarquée de 30 MtCO₂e et de l'empreinte totale de 95 MtCO₂e (p9 de la publication, besoin du Supplément à la publication pour donner plus de détails sur les hypothèses).

⁶⁶ Chiffres obtenus pour un centre de données de : 1 MW, charge de 50%, 6 kW / rack, PUE de 1.34, facteur d'émission de l'électricité des Etats-Unis, 5 tCO₂e / kW d'IT en empreinte embarquée, 94 % de serveurs de stockage et 6 % d'équipements réseaux, peuplés à 50 %. Mais avec la possibilité de faire varier ces caractéristiques dans un simulateur en ligne <https://www.se.com/ww/en/work/solutions/system/s1/data-center-and-network-systems/trade-off-tools/data-center-lifecycle-co2e-calculator/>

Annexe 3 : Réaliser et/ou interpréter une estimation de la consommation électrique des centres de données : quelles questions se poser ?

Cette annexe a vocation à renforcer et détailler les questions du paragraphe « Evaluation de la consommation électrique et mise en perspective ».

Elle se veut essentiellement méthodologique au sens où elle nous permet de partager des questions que nous estimons aujourd'hui avoir seulement partiellement éclairées et qui mettent en lumière des incertitudes pouvant expliquer de forts écarts d'une étude à une autre.

Elle est accompagnée de l'onglet « Trajectoire GES à partir de GW » dans le fichier « VERSION PROVISOIRE - Etat des lieux Monde.xlsx » afin de permettre la familiarisation avec ces questions. Cet onglet n'est pas un calculateur mais bien un outil de réflexion et de simulation, en tout cas à date.

Projeter la consommation électrique des centres de données à un horizon de quelques années peut être réalisé à partir de la connaissance de la demande de puissance électrique⁶⁷ à installer (GW) provenant de sites existants et de projets à différents stades de maturité. En effet, un projet de centre de données nécessite de 2 à 5 ans pour devenir opérationnel.

Pour cet exercice, prenons par exemple les données fournies dans le rapport DCByte (DCByte, 2024). Celles-ci agrègent :

1. Des demandes d'installation de puissance IT
2. Pour près de 7 000 centres de données
3. A travers le globe, avec des données parcellaires pour la Russie et la Chine
4. A partir d'analyses de marché, d'images d'observation par satellite, de communiqués, d'entretiens, d'inspections physiques
5. Des historiques d'évolution pour chacune de ces 3 catégories :
 - **Les projets en phase opérationnelle** : ce sont les projets pour lesquels la puissance IT est déjà déployée, qu'elle soit louée ou non.

⁶⁷ D'autres méthodes sont possibles, s'appuyant par exemple sur des estimations historiques et prévisionnelles de parcs et de livraisons d'équipements informatiques (serveurs, baies de stockage, etc.), dont nous ne traiterons pas ici, à ce stade.

- **Les projets dans les carnets de commande** : ce sont à la fois **les projets en cours de déploiement et les projets sécurisés** c'est-à-dire ceux sur lesquels la confiance est élevée avec des éléments acceptés tels que terrain, puissance.
- **Les projets « en phase précoce »** : ce sont ceux qui ont été annoncés ou ceux sur lesquels existent une spéculation mais qui ne sécurisent pas tous les éléments requis (terrain, puissance, etc.).

Ces 5 points sont autant de points d'attention puisque :

1. Les historiques, liste de projets, projections peuvent être en *puissance IT* (la puissance électrique des capacités informatiques installées) ou en puissance électrique totale pour l'ensemble du centre de données y compris *l'environnement* nécessaire au bon fonctionnement de l'IT (les équipements permettant le refroidissement des serveurs notamment). Auquel cas, on ajouterait ou non la consommation énergétique de *l'environnement*, ce qui peut constituer jusqu'à 50 % de la consommation IT en fonctionnement nominal et beaucoup plus en phase de montée en charge. Dans ce rapport, les estimations sont en puissance IT.
2. Le nombre de centres de données et/ou le type de centres de données considérés donnent une indication du périmètre considéré : avec 7 000 centres de données sur 12 000 environ dans le monde dans ce rapport, une partie seulement est couverte. Mais est-ce bien grave? Cela dépend de la question à laquelle ces estimations cherchent à répondre : informer sur l'état du marché lié aux deux plus grandes catégories de centres de données (hyperscalers et colocation) ? Ou estimer les capacités de productions électriques requises pour alimenter l'ensemble des capacités informatiques mondiales ?
3. Le périmètre géographique est aussi important pour les mêmes raisons. Dans la majorité des rapports, le périmètre Chine n'est jamais détaillé et souvent dit « partiel ». Avec ou sans, l'estimation se décale d'environ 25 – 30 %. Idem pour les pays dans lesquels la filière centres de données émerge, et qui ne sont pas toujours représentés alors que la croissance dans les années à venir peut y être très forte
4. La question de la construction et de la fiabilité des données primaires. Dans ce rapport, la méthodologie pour construire les données informe (un peu) sur la fiabilité à accorder aux projets « en phase précoce » par exemple.
5. La compréhension des différentes maturités de projet peut être pertinente pour comprendre à quelles dates auront vraiment lieu les demandes en consommation électrique éventuellement modéliser des durées de mise en service et de montée en charge.

Ensuite à partir d'un scénario de demande de puissance IT, il s'agit de considérer différents profils :

- **Un profil de charge.** Sachant que plusieurs phénomènes se recouvrent : pour les centres de données déjà installés, il est possible que *le taux de charge*⁶⁸ s'améliore. Pour les centres de données en cours d'installation, ils sont peu chargés au démarrage.
- **Un profil de PUE.** Encore une fois plusieurs phénomènes se recouvrent : le PUE étant un indicateur d'efficacité, il diminue avec le volume (quand les centres de données se remplissent). A l'ouverture, les PUE sont mauvais pendant plusieurs années.

⁶⁸ Le taux de charge qui nous intéresse ici va être le pourcentage moyen d'utilisation de la puissance électrique de l'IT installée ; celui-ci découle (sans que ce soit proportionnel) du taux de charge IT, c'est-à-dire le pourcentage moyen de la puissance de calcul utilisée.

Pour aller jusqu'au calcul d'émissions de gaz à effets de serre, cela se complète par :

- Un profil d'intensité carbone de l'électricité,
- Une façon de prendre en compte les émissions embarquées.

A date, les valeurs dans l'onglet "Trajectoire GES à partir de GW" dans le fichier « VERSION PROVISOIRE - Etat des lieux Monde.xlsx » sont simplement indicatives, pour ouvrir la discussion sur justement :

- *La modélisation,*
- *La sélection des données,*
- *L'intérêt de réaliser ce raisonnement par type de zone géographique ou par type de centres de données pour ajuster certaines valeurs.*

Vos contributions lors de l'atelier n°1, ou par retour d'e-mail, nous seront précieuses.

Annexe 4 : Les crypto-devises, un usage des infrastructures numériques toujours présent ?

Les transactions des crypto-devises reposent sur des blockchains publiques, plateformes décentralisées qui permettent à tout un chacun de tracer les transactions (depuis/vers des comptes encodés) mais aussi de participer à la sécurisation de ce mécanisme. Pour avoir le droit de valider les transactions - et être récompensé en monnaie nouvellement créée à l'occasion - des compétitions entre validateurs ont lieu à chaque validation de blocs de transactions. Majoritairement, deux types de mécanismes dominant :

- Le "Proof of Work" (PoW) consiste à mobiliser de la capacité de calcul afin d'être le premier à résoudre un problème mathématique donné. Les infrastructures mobilisées sont par exemple des fermes de serveurs équipés de cartes spécialisées (ASIC/GPU). C'est le mécanisme sur lequel s'appuie le Bitcoin, premier actif en capitalisation.
- Le "Proof of Stake" (PoS) repose sur la "mise en jeu" de ses actifs en crypto-devises pour avoir une chance de gagner le droit de valider. C'est le mécanisme sur lequel s'appuie l'Ethereum, second actif en capitalisation.

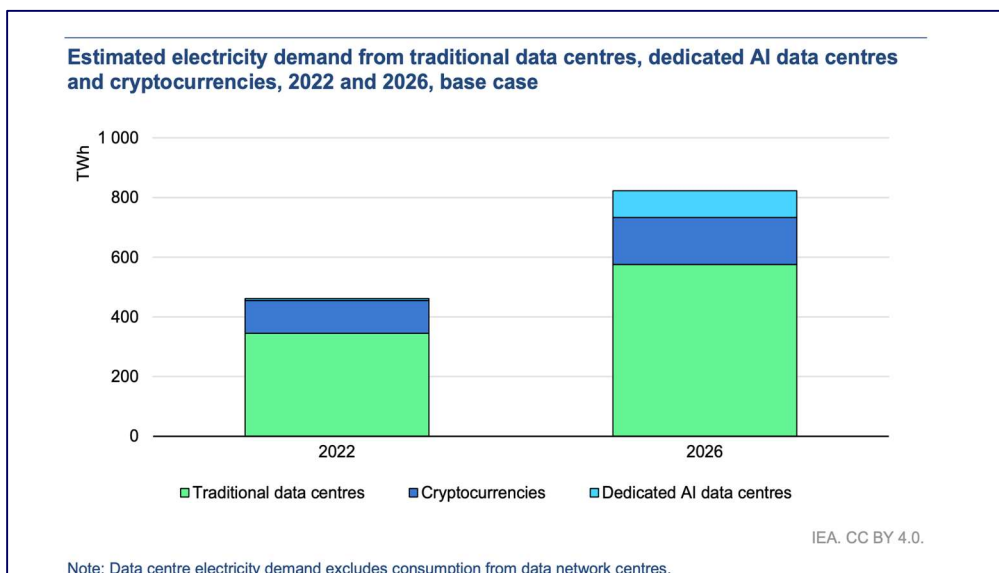


Figure 42 - Estimation de la consommation électrique pour les centres de données dans le monde. Source : (IEA, 2024a), p. 35

C'est le mécanisme "Proof of Work" qui représente la très grande majorité de la consommation électrique des crypto-devises : 120 TWh sur les 130 TWh de consommation totale des crypto-devises en 2023 (IEA, 2024a) (University of Cambridge, 2025).

Historiquement, la croissance de cette consommation est ininterrompue (IEA, 2024a) (University of Cambridge, 2025) , alimentée par deux facteurs : l'amélioration des performances et de l'efficacité des cartes de calcul, ainsi que le nombre de mineurs en compétition⁶⁹. Sa répartition territoriale dans le monde dépendra cependant des dynamiques insufflées aux infrastructures de calcul qui y sont dédiées : disponibilité de la ressource informatique, disponibilité en électricité, réglementations et politiques publiques accompagnant, encadrant et autorisant/interdisant leur implantation. Aujourd'hui, un pays comme la France n'est pas directement exposé à ces enjeux d'infrastructures, contrairement aux Etats-Unis⁷⁰, mais les effets de concurrence et de disponibilité de ressources pour les centres de données et leur implantation pourraient influencer sur leur répartition future.

⁶⁹ Le nombre de mineurs est au moins en partie corrélé au cours du bitcoin et du coût de l'électricité, facteur clé dans les modèles économiques de cette activité (Berkeley, 2024).

⁷⁰ La France représente 0,21 % du hashrate total mondial, soit environ 360 MWh/an ; les Etats-unis, 37,84 %. (University of Cambridge, 2025b).

THE SHIFT PROJECT

The Shift Project est un think tank qui œuvre en faveur d'une économie post-carbone. Association loi 1901 reconnue d'intérêt général et guidée par l'exigence de la rigueur scientifique, notre mission est d'éclairer et influencer le débat sur la transition énergétique en Europe. Le développement exponentiel du numérique, et la façon dont ce développement peut interagir avec les objectifs de décarbonation de nos sociétés, constitue l'un des angles essentiels des enjeux de la transition carbone.

Nos membres sont de grandes entreprises qui veulent faire de la transition énergétique leur priorité.

Contacts :

Hugues Ferrebœuf

Chef de projet

hugues.ferreboeuf@theshiftproject.org

Marlène de Bank

Ingénieure de recherche

marlene.debank@theshiftproject.org

Maxime Efoui Hess

Coordinateur de projet

maxime.efoui@theshiftproject.org

Pauline Denis

Chargée de projet

pauline.denis@theshiftproject.org

Ilana Toledano

Responsable Communication

ilana.toledano@theshiftproject.org