



HAL
open science

L’empreinte environnementale complète d’un usage numérique : contribution à l’analyse de cycle de vie de services numériques

Adrien Berthelot

► To cite this version:

Adrien Berthelot. L’empreinte environnementale complète d’un usage numérique : contribution à l’analyse de cycle de vie de services numériques. Informatique [cs]. Ecole normale supérieure de lyon - ENS LYON, 2024. Français. NNT : 2024ENSL0051 . tel-04874694

HAL Id: tel-04874694

<https://theses.hal.science/tel-04874694v1>

Submitted on 8 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

en vue de l'obtention du grade de Docteur, délivré par
l'ÉCOLE NORMALE SUPERIEURE DE LYON

École Doctorale N°512
École Doctorale en Informatique et Mathématiques de Lyon

Discipline : Informatique

Soutenue publiquement le 12/11/2024, par :

Adrien BERTHELOT

L'empreinte environnementale complète d'un usage numérique : contribution à l'analyse de cycle de vie de services numériques

Devant le jury composé de :

| | | |
|---------------------|---|--------------------|
| LIGOZAT, Anne-Laure | Professeure des universités, Université Paris-Saclay | Rapporteure |
| TROUSSIER, Nadège | Professeure des universités, ENSAM | Rapporteure |
| LEGRAND, Arnaud | Directeur de recherche, UGA | Examineur |
| PIERSON, Jean-Marc | Professeur des universités, Université de Toulouse | Examineur |
| ROUVOY, Romain | Professeur des universités, Université de Lille | Examineur |
| LEFÈVRE, Laurent | Chargé de recherche HDR, ENS de Lyon | Examineur |
| CARON, Eddy | Professeur des universités, ENS de Lyon | Directeur de thèse |

Laboratoire de l'Informatique du
Parallélisme
ENS de Lyon
46, allée d'Italie
69364 LYON CEDEX 07

École doctorale InfoMaths
7 avenue Jean Capelle
Bât. Blaise Pascal
69621 VILLEURBANNE Cedex

Résumé

Dans un contexte de prise en compte des enjeux environnementaux, la réduction des impacts environnementaux du numérique s'est imposée comme une question de recherche en informatique. Cependant, une telle démarche nécessite rapidement une connaissance plus exhaustive de l'empreinte du numérique que la consommation d'électricité d'un serveur ou d'un ordinateur. Déjà présente dans d'autres industries, l'analyse de cycle de vie (ACV), semble être l'outil adéquat pour mieux connaître, contrôler et, in fine, réduire, l'empreinte environnementale du numérique. Cette thèse étudiera l'empreinte en question sous l'angle de l'ACV de services numériques. Car l'usage du numérique tel qu'expérimenté au quotidien se rapproche toujours plus d'une économie de service. Au moyen, non pas d'un unique objet, mais d'un écosystème technique, nos usages du numérique sont des consommations de services reposant autant sur des terminaux individuels que sur des infrastructures de communication, calcul et stockage. Ces co-dépendances, rendent l'empreinte de notre usage du numérique plus complexe et justifient l'application d'ACV de services numériques évaluant de bout en bout l'empreinte issue de l'usage du numérique.

Toutefois, plus de 10 ans après la parution du standard adaptant l'ACV aux biens, réseaux et services basés sur les technologies d'information et communication (TIC), même au sein des communautés intéressées, l'ACV et ses apports ne semblent toujours pas prédominants. Partant de ce constat, ce travail de thèse cherchera à expliquer la faible présence de l'ACV dans le numérique et à contribuer, conscient de ses avantages, à consolider son usage comme outil d'évaluation pour le numérique.

La contribution sera d'une triple nature : d'une part méthodologique cherchant à détailler et rendre praticable l'ACV de services numériques, d'autre part, ce travail d'évaluation servira à donner de nouvelles clés de compréhension de l'empreinte du numérique ainsi que de nouvelles grilles de lecture à même de proposer des politiques de contrôle et de réduction de cette empreinte. Enfin, ce travail de thèse contribuera à la production de résultats expérimentaux sur l'empreinte de services numériques.

Cette thèse vise à contribuer méthodologiquement à l'ACV de service numérique, en l'appliquant à des cas d'usages réels et en définissant plus précisément son application. Cela sera fait au travers de deux études de cas évaluant de bout en bout plusieurs impacts environnementaux de différents services numériques en situation réelle. Une première étude de cas se focalisera sur les services numériques basés sur l'intelligence artificielle générative (Gen-IA) en évaluant l'empreinte environnementale du Stable Diffusion, un service de génération d'images. Ensuite, une seconde étude de cas portera sur le fournisseur de services numériques libres Picasoft. Le but de cette seconde étude sera d'évaluer en

même temps plusieurs services numériques aux caractéristiques hétérogènes, hébergés par un même acteur, sans perdre le détail de la distribution de l’empreinte entre les services.

Dans l’ensemble, ce manuscrit propose d’explorer l’histoire, les potentiels, les limites et les applications de l’ACV aux services numériques. Par l’exemple de ses applications et ses méthodologies ouvertes, ce manuscrit espère contribuer à de meilleures méthodes d’évaluation des impacts environnementaux pour les services numériques.

Remerciements

À vous qui lisez ce manuscrit, le premier remerciement vous revient, l'intérêt que vous portez à ce document me touche et j'espère ne pas décevoir vos espérances à lecture de ces pages.

Cette section est naturellement beaucoup plus personnelle que le reste de ce manuscrit et bien loin de toute considération scientifique. Il est néanmoins nécessaire pour moi de profiter cet espace pour remercier un certain nombre de personnes. Cela sera peut-être long. :)

Jury

Tout d'abord, je tiens à remercier les membres de mon jury de thèse : Anne-Laure Ligozat, Nadège Troussier, Arnaud Legrand, Jean-Marc Pierson et Romain Rouvoy. Je remercie tout particulièrement Anne-Laure Ligozat et Nadège Troussier pour avoir lu et rapporté mon manuscrit.

Bien évidemment, cette thèse n'aurait pu être menée sans l'encadrement, de haute qualité scientifique et humaine, dont j'ai pu bénéficier pendant ces trois années. Merci à Alexis Nicolas pour sa patience, son empathie, sa capacité à canaliser nos réunions, à synthétiser, et pour avoir été mon guide dans l'océan OCTO. Merci à Laurent Lefèvre, pour ces années à m'accompagner en stage puis en thèse, pour avoir ouvert de nouveaux chemins dans la recherche en informatique qui ont permis l'émergence de cette thèse et pour m'avoir donné l'incalculable chance d'y participer. Merci à Eddy Caron, pour son soutien indéfectible et infatigable, pour ces longues heures de débats en salle 316, pour son ouverture d'esprit et sa sincère curiosité.

Laboratoire

J'ai eu le plaisir de réaliser cette thèse au sein du Laboratoire de l'Informatique du Parallélisme (LIP) hébergé au sein de l'ENS de Lyon. Plus particulièrement, j'ai eu le plaisir d'être membre de l'équipe Avalon, étant ainsi gratifié d'adorables collègues. Un grand merci donc à Christian Pérez pour ses conseils avisés et pour être responsable de cette belle équipe. Ces trois années n'auraient sans doute pas été aussi agréables et stimulantes sans la compagnie d'autres doctorant.es étudiant l'empreinte environnementale du numérique. Je parle de Simon, Maxime, Vladimir, Hugo et particulièrement Mathilde

grâce à qui j'ai pu écrire de très belles pages de cette thèse. Mes remerciements vont ensuite aux autres membres de l'équipe avec qui j'ai pu pendant ces trois années partager de bons moments : Pierre, Laurent, Romain, Pierre-Étienne, Thierry, Élise, Lucien, Jerry, Doreid, Étienne, Emeline, Johanna, Maxime, Thomas et Chrystelle.

Je profite de cette section pour remercier les membres de mon comité de suivi de thèse, Bora Uçar et Cédric Tedeschi.

OCTO

Mes remerciements vont premièrement à Christian Fauré pour m'avoir fait confiance pour ce projet, m'avoir invité au sein d'Octo et avoir créé au sein d'Octo cet espace de recherche. J'ai pu ainsi rencontrer d'autres doctorant.es passioné.es et passionnant.es comme Lucie, Élie et Amaranta.

Mon aventure à OCTO a été considérablement marquée par mon unique "tribu" de cœur et ses membres, l'équipe SEED. Merci à vous tous pour votre accueil chaleureux au sein d'Octo et votre curiosité. Un grand merci notamment à : Nicolas, Pierrick, Sébastien, Wojciech, Sandrine, Margaux, Sara, Adrian, Julie et Alaric.

Communauté scientifique

Cette thèse fut aussi pour moi l'opportunité de rencontres passionnantes avec des personnes travaillant sur les enjeux de l'informatique et de la soutenabilité. Je garde un particulièrement bon souvenir des écoles d'été comme SICT ou Eco-ICT, ou des conférences comme ICT4S ou Archipel. Ce fut toujours un plaisir de croiser à l'occasion de ces évènements Aina, Lou, Clément, Ludmila ou Baptiste et de pouvoir profiter de leurs discussions.

Autre élément significatif de ma thèse, j'aimerais adresser un remerciement à l'ensemble des membres de l'association Boavizta pour leur investissement dans ce projet humain et scientifique. Je suis, et continue à être, constamment enrichi par les échanges au sein de cette communauté à laquelle je dois beaucoup. Particulièrement, un grand merci à David, Gauthier, Thibault, Youen, Samuel et l'équipe d'Hubblo pour tous les échanges que j'ai eu la chance d'avoir avec eux.

Enfin, j'adresse aussi mes remerciements à Stéphane, Romain, ainsi qu'à l'association Picasoft pour m'avoir permis d'écrire un beau morceau de cette thèse.

Faculté

Les études d'informatique à l'université Lyon ont été pour moi aussi enrichissante qu'éprouvante. J'ai eu la chance de bénéficier de nombreux cours de la part de nombreux enseignants qui mériteraient une place dans ces remerciements. Toutefois, je tiens à remercier particulièrement Isabelle Guérin-Lassous et Thomas Begin pour leur rôle dans mes premières expériences dans le monde de la recherche et leurs enseignements.

Je ne serais certainement pas sortie sur les deux pieds de ces études sans plusieurs précieux camarades. Un grand merci à Alix, Antoine, Isabelle, Pierre, Titouan, Vincent

et Nicolas (avec qui j'ai partagé et perfectionné l'art subtil d'optimiser le temps de travail). Votre présence au bâtiment Nautibus m'a été vitale.

Ami.es

Il est assez évident que je ne serais pas là où j'en suis actuellement sans de nombreux ami.es pour m'apporter soutien et sens dans cette étrange expérience qu'est la vie. Je ne les remercierai jamais assez pour cette amitié que nous partageons depuis tant d'années. J'aimerais remercier donc l'ensemble de la petite famille CCC : Pierre, Liang, Emma, Léa, Simon(P), Laure, Louise, Samir, Baptiste, Sophie, Julie, Théo, Edgar et Simon(B).

J'aimerais aussi remercier l'ensemble des ami.es avec qui j'ai pu partager ces dernières années de nombreuses aventures en GN, en vacances et dans tant d'autres moments. Un grand merci à : Georges, Xavier, Juliette, Joël, Nattes, Aaren, KT, Lambert, Florine, Morgan, Marine, Élie, Antoine, Gabriel, Alexandra, Malo, Alix, Corentin, Mista, Laureline, Bertrand, Simon(F), Janelle, Marie, Guillaume, Alban(F), Julie, Simon(B) et Alban(R).

Enfin, un dernier lot de merci aux personnes qui m'ont accompagné dans mes premières années mouvementées de GN en association, puis avec l'association GNafon. Un grand merci à : Aurore, Nicolas, Seymour et Manue.

Famille

À ma famille, pour leur amour et leur soutien inconditionnel depuis plus de 28 ans, j'aimerais adresser plusieurs remerciements sincères : À mes grands-parents, Henri, Maryvonne, Simone et Gérard, pour toutes ces vacances, ces bons moments, ces efforts pour maintenir le lien malgré la distance.

À ma sœur Aurélie et mon frère Artur, qui m'ont supporté pendant toutes ces années.

À mes parents Alain et Armelle, un immense et insondable merci.

Enfin merci à Charline pour avoir partagé avec moi ces années de thèse, avec sa patience, ses conseils et son amour. Je lui en suis infiniment reconnaissant.

Table des matières

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Empreinte environnementale | 1 |
| 1.2 | Le numérique : un <i>pharmakon</i> ? | 3 |
| 1.3 | Questions de recherches | 7 |
| 1.4 | Organisation de ce manuscrit | 8 |
| 1.5 | Note complémentaire : contexte scientifique de la thèse | 9 |
| 2 | Informatique et environnement : une prise en compte progressive | 13 |
| 2.1 | Informatique : considérer l'impact | 13 |
| 2.1.1 | Premiers travaux sur l'énergie | 13 |
| 2.1.2 | L'émergence du "Green IT" | 16 |
| 2.1.3 | Plus que de l'électricité, plus que du carbone | 20 |
| 2.2 | L'analyse de cycle de vie | 23 |
| 2.2.1 | Perspective chronologique et définition | 24 |
| 2.2.2 | Considérer les TIC puis le numérique | 27 |
| 2.2.3 | Une difficile intégration des services numériques | 30 |
| 3 | Analyse de cycle de vie de services numériques : potentiel et limites | 35 |
| 3.1 | Potentiel de l'ACV de services numériques | 35 |
| 3.1.1 | Évaluation environnementale complète | 35 |
| 3.1.2 | Éco-conception de services numériques | 39 |
| 3.1.3 | Une échelle service activable pour les organisations | 40 |
| 3.1.4 | Interopérabilité avec les autres secteurs | 41 |
| 3.2 | Les limites actuelles de l'ACV de services numériques | 42 |
| 3.2.1 | Manque de maturité globale du champ scientifique | 43 |
| 3.2.2 | Critique des rares études accessibles | 45 |
| 3.3 | Analyse de l'état de l'art | 50 |
| 3.3.1 | Comment lire une ACV de services numériques | 50 |
| 3.3.2 | Proposition d'une modélisation simplifiée de l'ACV de services numériques | 54 |
| 3.4 | Conclusion : motivation des contributions de cette thèse | 56 |
| 4 | ACV de service numérique | 59 |
| 4.1 | Une approche attributionnelle du service numérique : échelle et objectifs | 60 |
| 4.1.1 | Objectifs et périmètre | 60 |
| 4.1.2 | Données et outils | 63 |
| 4.2 | Analyse de l'inventaire du cycle de vie du service numérique | 67 |

| | | |
|----------|--|------------|
| 4.2.1 | Clés d'allocation | 69 |
| 4.2.2 | Terminaux utilisateurs | 69 |
| 4.2.3 | Réseaux | 71 |
| 4.2.4 | Datacenters | 73 |
| 4.3 | Évaluation des impacts | 75 |
| 4.4 | Analyse des résultats | 77 |
| 4.5 | Conclusion et limites | 77 |
| 5 | Étude de cas d'un service unique : Stable Diffusion | 81 |
| 5.1 | Contexte IA et environnement | 82 |
| 5.1.1 | IA et Gen-AI au sein du numérique : définitions et enjeux | 82 |
| 5.1.2 | L'empreinte environnementale de l'IA : un travail encore incomplet | 83 |
| 5.2 | Méthodologie : L'IA comme service | 85 |
| 5.2.1 | Équipement d'accès - Terminaux utilisateurs | 87 |
| 5.2.2 | Équipement d'accès - Réseaux | 87 |
| 5.2.3 | Serveur de datacenters - Hébergement web | 87 |
| 5.2.4 | Serveur de datacenters - Inférence du modèle | 88 |
| 5.2.5 | Serveur de datacenters - Entraînement du modèle | 88 |
| 5.3 | Étude de cas : Stable Diffusion | 89 |
| 5.3.1 | Cadrage de l'étude | 89 |
| 5.3.2 | Résultats | 91 |
| 5.3.3 | Analyses de sensibilités | 92 |
| 5.4 | Discussions, limites et perspectives | 94 |
| 5.4.1 | L'enjeu de l'approche service dans un contexte comme celui de l'IA général | 96 |
| 5.4.2 | Obsolescence hardware et software | 97 |
| 5.4.3 | Limites actuelles | 98 |
| 5.4.4 | Perspectives | 99 |
| 5.5 | Conclusion | 101 |
| 6 | Étude de cas d'un ensemble de services : le CHATONS Picasoft | 103 |
| 6.1 | Contexte : Impact environnemental des services numériques | 104 |
| 6.2 | Méthodologie générale | 105 |
| 6.2.1 | Outils mis en œuvre | 105 |
| 6.2.2 | Distribuer l'empreinte de l'hébergement entre les services | 106 |
| 6.2.3 | Inclure les parties réseaux et terminaux utilisateurs | 109 |
| 6.3 | Étude de cas : Les services Picasoft | 111 |
| 6.3.1 | Contexte | 111 |
| 6.3.2 | Objectif et périmètre | 111 |
| 6.3.3 | Résultats | 113 |
| 6.4 | De nouvelles stratégies de réduction de l'empreinte environnementale des services | 114 |
| 6.4.1 | Empreinte par visiteur comme indicateur d'efficacité | 117 |
| 6.4.2 | Le coût de l'imprévisibilité | 117 |
| 6.4.3 | Le coût de l'utilisateur : la grande inconnue restante | 118 |
| 6.5 | Limites | 122 |

TABLE DES MATIÈRES

| | | |
|----------|---|------------|
| 6.6 | Discussions | 122 |
| 6.7 | Perspectives | 123 |
| 6.8 | Conclusion | 124 |
| 7 | Conclusion | 125 |
| 7.1 | Conclusion générale | 125 |
| 7.2 | Discussion : de la mesure des impacts environnementaux | 127 |
| 7.2.1 | Que faire de l'ACV | 129 |
| 7.2.2 | Que faire de l'évaluation environnementale | 130 |
| 7.3 | Perspectives | 131 |
| | Annexe 1 : ACV accéléré d'un ensemble de services numériques | 133 |
| | Publications et présentations | 135 |
| | Postface : Bilan carbone de cette thèse | 137 |
| | Bibliographie | 143 |

Introduction

*Ce n'est qu'au cours de la période
représentée par le siècle actuel
qu'une espèce - l'homme - a acquis
le pouvoir significatif de modifier la
nature de son monde.*

Rachel Carson - 1962

1.1 Empreinte environnementale

Les activités humaines ont, dans leurs interactions avec la biosphère, toujours participé à modifier celle-ci. L'époque contemporaine peut être caractérisée par une croissance du nombre de problèmes environnementaux, i.e. de dégradation des conditions de vie des populations, et ceci de manière globalisée. Que cet état de fait, soit analysé sous le prisme du concept de *Grande accélération* [1] ou celui d'*Anthropocène* [2], le poids de la technosphère sur la biosphère n'a jamais été aussi grand et, si la dynamique actuelle continue, les modifications pourraient être significativement durables à l'échelle de l'Humanité. Actuellement, une partie de la gestion de ces problèmes environnementaux consiste à modifier ces dynamiques, avec une certaine urgence nécessaire [3], notamment en allégeant le plus possible le poids environnemental des activités humaines, de cette technosphère, sur la biosphère.

La Figure 1.1, issu du rapport du GIEC [4], illustre probablement le problème environnemental le plus connu. Différentes activités de la technosphère humaine émettent depuis l'ère industrielle un volume anormal, au regard de la dynamique propre de la biosphère, de gaz à effet de serre (GES). Ces émissions de GES entraînent un phénomène de changement climatique perturbant de manière hétérogène, mais globale, les sociétés humaines. De la conférence de Rio en 1992 à l'accord de Paris en 2015, le problème du changement climatique a fait l'objet d'une attention particulière. Les concepts d'émissions et d'empreinte carbone [5] se sont popularisés. Derrière ces concepts, on trouve un modèle caractérisant la contribution au changement climatique des nombreux GES par rapport

au plus connu, le dioxyde de carbone CO_2 [6]. On observe alors pour essayer de résoudre ce problème la recherche de solutions techniques et politiques, premièrement dans le but de limiter les émissions de GES, de prévenir ; le processus complémentaire, "réparer", étant bien plus périlleux à mettre en œuvre.

Changement de la température à la surface du globe par rapport à la période 1850–1900

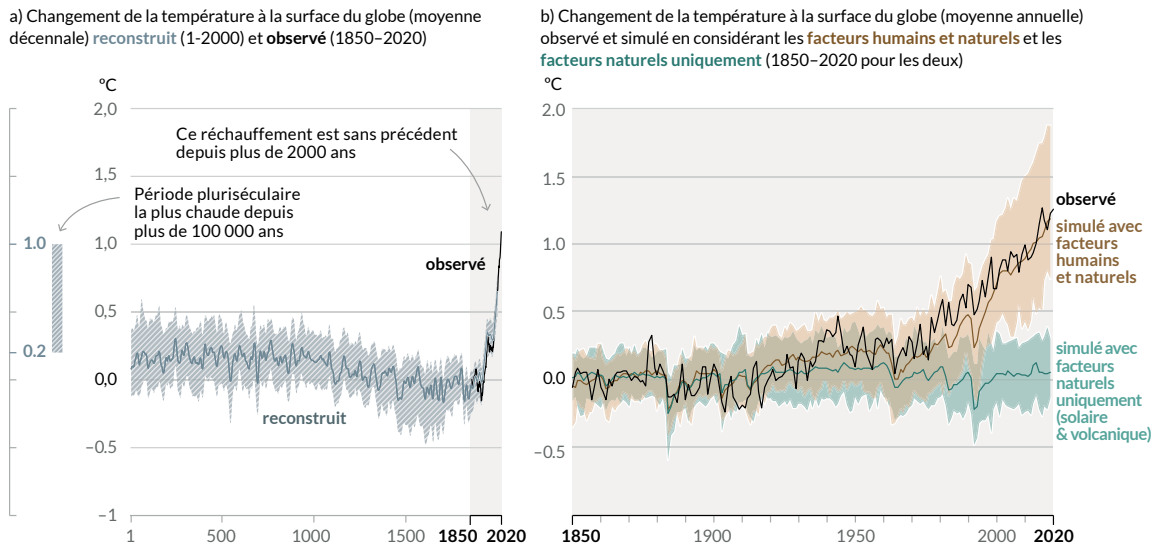


FIGURE 1.1 – Évolution du changement climatique depuis l'ère industrielle, de son origine anthropique, illustration issu du rapport du GIEC [4]

La pré-dominance actuelle, parmi les enjeux environnementaux, du réchauffement climatique ne doit pas cacher que cela n'a pas toujours été le cas. Le problème du trou de la couche d'ozone, par exemple, représente une forme de succès collectif face à un problème au schéma similaire [7]. Des activités industrielles émettent des chlorofluorocarbures (CFC) qui endommagent la couche d'ozone dégradant les conditions de vie humaine et, en réponse, le protocole de Montréal est signé en 1987. Les émissions de substances appauvrissant la couche d'ozone (SAO) sont sévèrement contrôlées et réduites. Cependant, en rapport avec le précédent problème, si la diminution des SAOs a évité des émissions de GES, certains SAOs étant aussi des GES, le protocole a dû être amendé en 2016. L'amendement de Kigali a été adopté pour restreindre l'usage de produits remplaçant les CFC mais étant de puissants GES. Cet historique autour des SAOs et de l'amendement de Kigali en 2016 illustre bien la multiplicité des problèmes environnementaux et leurs potentielles interdépendances.

Cette interdépendance entre les différents problèmes environnementaux est à garder à l'esprit. En observant la Figure 1.2, on constate que les enjeux environnementaux sont multiples. La modélisation de chacun d'entre eux est un travail toujours en cours. Ces limites planétaires sont une représentation de la dégradation des conditions d'habitabilité de la Terre au-delà des dynamiques durables, i.e. le "safe operating space" en vert. Le levier d'action le plus efficace contre le franchissement de ces limites est la réduction, voire l'arrêt des activités à la source de ces problèmes. Cependant, le plus souvent, ce levier va se traduire par le choix d'autres technologies pour soutenir l'activité humaine.

Tout comme dans le cadre des émissions de SAO, les solutions technologiques adaptées aux enjeux environnementaux nécessitent une approche globale. Une solution à un problème peut aussi bien avoir un effet de bord positif comme négatif sur une autre situation.

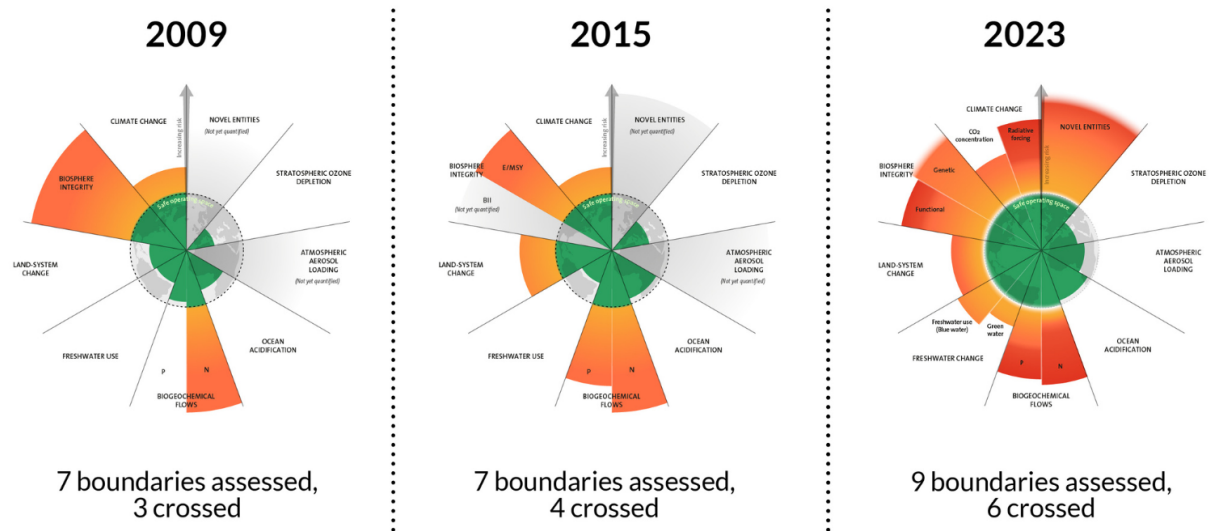


FIGURE 1.2 – Évolution du nombre de limites planétaires observées et de leur dépassement au cours des différents travaux du *Stockholm Resilience Center*, jusqu'au plus récent [8]

Pour répondre à cette nécessaire approche globale, le présent manuscrit se concentrera sur les méthodes d'évaluations des impacts environnementaux directs de solutions technologiques comme l'analyse de cycle de vie (ACV). Il ne faut toutefois pas ignorer les impacts sociaux, et encore moins les croire déconnectés des enjeux environnementaux. Ces catégories "environnemental" et "social" sont interconnectées et se recoupent bien souvent. On pourrait d'ailleurs critiquer cette séparation qui est faite entre ces catégories, rappelant la commune division "nature" et "culture". Par exemple, lorsqu'en 1950, Karl William Kapp publie ses recherches, dans la continuité d'économistes comme Pigou, il liste un long ensemble de coûts ignorés par les entrepreneurs. Sa liste de coûts inclut entre autres : la pollution de l'air, du sol, de l'eau, l'épuisement des ressources renouvelables et fossiles, ou encore l'obsolescence programmée, et pourtant ces éléments ne sont jamais présentés comme des "coûts environnementaux" mais bien des coûts sociaux [9]. La notion d'impacts environnementaux relève dans ce manuscrit d'un choix méthodologique, d'une approche ayant guidé les travaux d'évaluations des technologies étudiés, mais qui bien sûr n'est qu'une réponse partielle aux questions que posent les technologies d'information et de communication (TIC) à l'heure de l'Anthropocène.

1.2 Le numérique : un *pharmakon* ?

Le concept du numérique comme *pharmakon* est dû à Stiegler [10] qui emprunte le terme à Derrida, qui l'emprunte à Platon qui caractérisait ainsi une des premières technologies d'information : l'écriture. Le numérique comme *pharmakon* symbolise une vision ambivalente de la technologie comme simultanément "remède" et "poison", faisant écho aux questions sur les impacts positifs et négatifs des TIC.

Si cette question est si prégnante, c'est que cette dynamique de grande accélération évoquée à la section précédente concerne grandement le numérique. Les dernières décennies sont caractérisées par une explosion à la fois des capacités, mais aussi de l'usage des TIC [11]. D'un mouvement commun, le coût, l'utilisation et les capacités de stockage, calcul et communications ont radicalement augmenté. Un phénomène qui est parfois qualifié de "transition" numérique, même s'il est difficile de savoir où est le point d'arrivée de cette transition, tant la dynamique semble parfois exponentielle.

Le poids du numérique sur l'environnement est non négligeable et surtout c'est un poids grandissant, suivant la croissance de ce secteur. L'empreinte carbone du numérique est ainsi devenue un sujet de préoccupation, dans la communauté académique et ceci jusqu'aux gouvernements. La question des émissions de GES du numérique est d'autant plus critique, technique et politique, que les orientations décidées arrivent à des moments décisifs au regard des enjeux environnementaux comme vu précédemment [3]. Ces choix sont d'autant plus non triviaux qu'il est considéré que les TIC pourraient aider à résoudre d'autres problèmes environnementaux [12, 13], et que ces mêmes affirmations sont débattues scientifiquement [14, 15]. Illustration de ces choix d'orientation, la figure 1.3, issue du rapport ADEME-Arcep [16], représente quatre scénarios possibles d'empreinte carbone du numérique.

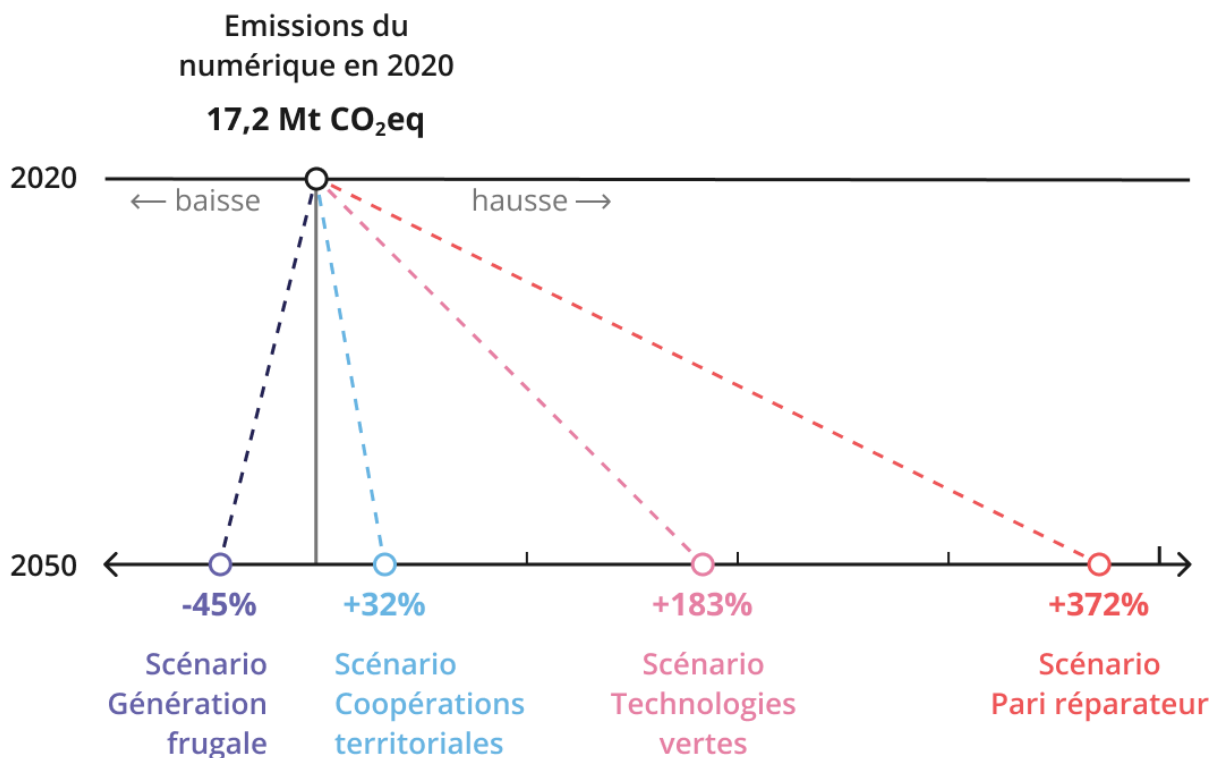


FIGURE 1.3 – Prévisions, selon les 4 scénarios prospectifs, des taux d'évolution des émissions de CO₂eq du numérique en 2050 (sur tout le cycle de vie), en France, par rapport à 2020 [16]

Les 4 scénarios mentionnés dans le rapport de concernant les TIC [16] sont issus d'une étude prospective plus large [17] décrivant différents futurs technologiques, sociétaux

et industriels à l'échelle de la France. Il est important de noter que dans le cadre de l'étude, bien que l'empreinte carbone du numérique varie, tous les scénarios pour 2050 sont considérés comme "neutres en carbone". Ils considèrent à différents degrés le recours à de futures technologies de captage et de stockage du carbone à l'heure actuelle non-existante. Différentes hypothèses sont ainsi faites sur la capacité des TIC à pouvoir à l'avenir contribuer à la création et utilisation de ces nouvelles technologies à même de décarboner les activités humaines. Cependant, le rapport [16] souligne le fait suivant :

« Si ce développement du numérique permet en partie de réduire d'autres impacts environnementaux dans d'autres secteurs (mobilité par exemple), les consommations qu'il engendrerait en électricité et en ressources posent de toute façon la question de leur faisabilité. »

Cette conclusion illustre l'importance d'une évaluation plus complète du coût environnemental. Afin de parvenir à des choix éclairés sur la proportion d'utilisation du numérique, sur le dosage de ce *pharmakon*, il faut des outils d'estimation plus globaux. Les enjeux environnementaux ne se limitent malheureusement pas au réchauffement climatique. Par conséquent, l'empreinte environnementale du numérique doit prendre en compte l'ensemble de ses impacts environnementaux, au-delà de l'empreinte carbone et de la consommation d'électricité. Cette évaluation plus complète est nécessaire pour établir des politiques, faire des choix plus informés sur la place des TIC face aux enjeux environnementaux. La Figure 1.4 est issue d'une étude [18] estimant dans une approche ACV les impacts environnementaux du secteur numérique dans l'Union européenne (UE) en 2019. Cette approche permet d'estimer les impacts en prenant en compte l'ensemble du cycle de vie des produits, de la fabrication à la fin de vie, en prenant en compte un ensemble de catégories d'impacts environnementaux. Ces catégories répondent chacune à un "problème" environnemental et sont quantifiées par un modèle propre. La Figure 1.4 présente les résultats de manière pondérée, selon leur importance relative, dans une méthodologie établie par le *Joint Research Center* (JRC) [19]. Cette représentation permet de classer les impacts du numérique par importance. On observe alors que, dans ce cadre, le changement climatique n'est pas l'impact le plus important du numérique, bien qu'il soit conséquent.

Au contraire, on constate une diversité des catégories d'impact significative pour le secteur numérique. Ces catégories sont pour certaines spécifiquement liées à certaines phases du cycle de vie. L'utilisation de ressources minières est fortement liée à la phase de fabrication, les radiations ionisantes sont liées à la phase d'usage, du fait de la présence de nucléaire dans le mix électrique européen, enfin, les émissions de particules sont liées au transport des biens. Au sein de cette diversité des sources d'impacts, la grande tendance qui semble mise en valeur par le graphique est la prédominance des terminaux utilisateurs dans l'impact environnemental du numérique. Il pourrait alors être logique de se focaliser ce point. Toutefois, afin de mieux comprendre les conclusions possibles de cette information, il est nécessaire de décrire les trois niveaux utilisés dans la figure :

- **Terminaux utilisateurs** : Ensemble des équipements des utilisateurs finaux, tels que smartphones, ordinateurs, écrans, tablettes, télévisions, objets connectés, ...
- **Réseaux** : Infrastructures réseaux filaires et mobiles, cœur de réseau, équipements réseaux des utilisateurs finaux type *box* internet, ...

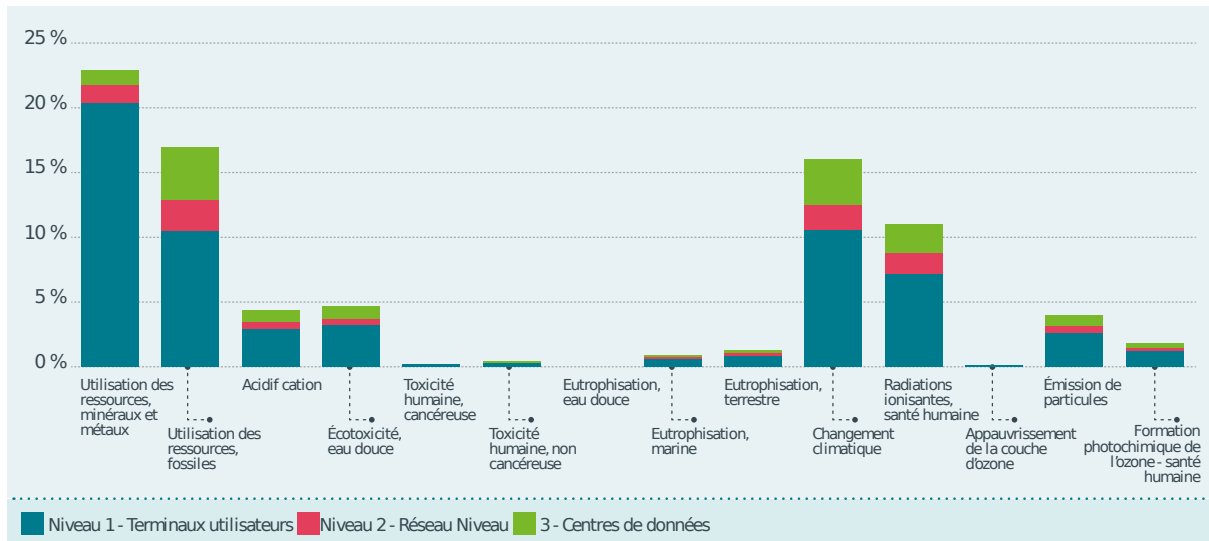


FIGURE 1.4 – Résultats normés et pondérés (méthode EF 3.0) des impacts environnementaux du numérique dans l’Union européenne en 2019 [18]

- **Centres de données** : Ensemble des équipements utilisés dans les *datacenters*, qu’il s’agisse de moyen de stockage, calcul ou communication, ou bien d’infrastructure électrique ou de refroidissement.

En premier lieu, il faut rappeler que ces catégories sont un choix arbitraire destiné à représenter la part de secteurs économiques. Elle se concentre sur le matériel, ignorant la partie logicielle du numérique qui peut être un élément transversal entre ces parties. Si cette séparation remonte à, au moins 2008, dans un rapport justement écrit par et pour des industriels [20], elle n’est pas immuable. Certains choix comme l’inclusion ou non de télévisions influencent les résultats et ne comportent pas de consensus [21, 22]. Le calcul et la répartition de l’empreinte du numérique à grande échelle est un objet de controverse aux enjeux certains [23, 24]. Quel sous-secteur porte la plus grande responsabilité ? La focalisation sur l’empreinte des terminaux est fréquente dans de nombreux travaux, et elle est justifiée. Cependant, il faut garder à l’esprit que ces 3 secteurs forment un écosystème d’interdépendances. Aucun consommateur n’achète un smartphone uniquement pour l’objet en lui-même. L’existence d’infrastructures réseaux et datacenters sont la condition *sine qua none* de l’achat de la majorité des terminaux.

Considérant ce fait, cette thèse prend le parti d’étudier l’empreinte du numérique à l’échelle service. Le service numérique est considéré comme l’interaction concertée entre ces 3 niveaux, i.e. Terminaux-Réseaux-Datacenters, afin de répondre au besoin d’un utilisateur. Cette échelle est issue de la structure, en 3 tiers elle aussi, proposée par la norme ITU L.1410 [25] sous la forme Équipements-Réseaux-Services, mais où chaque partie englobe les précédentes. Ainsi, le service comprend l’ensemble des équipements et infrastructures pour les coordonner afin de répondre à un usage dans ce qui tend à être une économie de fonctionnalité [26].

L’économie de fonctionnalité est le cadre dans lequel pour répondre à un besoin, un consommateur va acheter non pas un ou plusieurs équipements, mais un service. Cette notion d’économie de fonctionnalité s’applique aux services numériques, le paradigme

cloud [27] en étant une illustration pertinente. Si l'économie de fonctionnalité porte depuis longtemps des ambitions de modèle plus soutenable environnementalement [28], force est de constater que, bien que portant des promesses similaires, le cloud a accompagné la croissance du numérique. À l'échelle de l'UE, la proportion d'entreprises utilisant des services basés sur le cloud a presque doublé en 5 ans, passant de 23,9% en 2018 à 45,2% en 2023 [29].

Ce paradigme cloud de plus en plus dominant concrétise la notion de services. Il matérialise aussi une contradiction de l'économie de fonctionnalité : l'effet rebond ou paradoxe de Jevons [30]. En résumé, augmenter l'efficacité d'utilisation d'une ressource diminue le coût du service utilisant cette ressource et peut amener à augmenter la consommation globale de cette ressource. On peut ainsi voir dans le cloud un facteur ayant contribué à l'accroissement de l'empreinte des datacenters au sein du numérique. Sans forcément chercher à le modéliser directement, l'effet rebond invite à considérer l'usage dans le coût du service.

Cette notion d'usage, présente dans le titre de ce manuscrit, est critique pour discuter de l'empreinte du numérique selon plusieurs scénarios comme vu précédemment. L'échelle service permet d'intégrer cette notion. Plutôt que de traiter indépendamment chaque brique technologique, le service permet de questionner plus largement les technologies. Il questionne le besoin et l'usage, qu'il peut alors intégrer dans l'arbitrage de choix de conceptions, afin de doser de la manière la plus éclairée possible le *pharmakon* qu'est le numérique dans notre situation.

1.3 Questions de recherches

L'ambition d'inscrire l'informatique dans un avenir environnementalement soutenable, alors même que l'usage des TIC continue de croître et de se complexifier, répond à plusieurs contraintes. Il y a d'abord la multiplicité des enjeux environnementaux au-delà du réchauffement climatique qui appelle à une approche multicritère. Ensuite vient la nécessité de prendre en compte l'ensemble du cycle de vie des TIC et surtout de ne pas considérer qu'une seule sous-partie des technologies informatiques, que ce soit les terminaux ou les datacenters. Toutes ces contraintes cherchent à permettre une vision plus juste de l'empreinte environnementale du numérique, mais aussi à orienter des choix technologiques dans la construction de services numériques. L'approche holistique recherchée vise à éviter les transferts d'impacts aux différents niveaux représentés par les contraintes.

— L'outil qui semble répondre à ces contraintes est l'ACV de services numériques.

Toutefois, bien que largement utilisée dans d'autres secteurs de l'industrie, l'ACV possède un historique plus restreint dans le domaine des TIC, encore plus lorsqu'on souhaite l'appliquer à des services. De plus, certaines organisations utilisent quotidiennement plusieurs dizaines de services numériques. Le nombre de services et leurs évolutions techniques rapides, font craindre qu'un outil comme l'ACV, pouvant nécessiter un certain temps afin d'être appliqué correctement, demande trop de temps ou de moyens pour de tels ensembles de services. Un tel passage à l'échelle, la création d'une ACV accélérée et assistée d'un ensemble de services, pourrait être notre question de recherche. Malheureusement, un tel travail suppose une certaine maturité de l'ACV de services numériques qui, comme nous l'exposerons dans les chapitres 2 et 3, n'a pas été constatée en l'état.

De la question "Comment passer à l'échelle l'ACV de service numérique", la question est devenue "Comment contribuer à l'ACV de service numérique". En effet, l'imaturité de ce champ scientifique [31] se manifeste sous plusieurs aspects qui seront adressés par les contributions. Il y a les difficultés d'accès aux données d'impact de cycle de vie (ICV) des équipements et infrastructures informatiques. Il y a un manque de méthodologie unifiée. Les différentes itérations ont fait face à plusieurs difficultés, au premier desquelles le haut degré de mutualisation des TIC. Cette mutualisation, e.g. parallélisation des tâches, propre aux équipements informatiques, rend significativement plus difficile d'allouer à un service une juste part de l'empreinte des équipements utilisés.

Il y aura dans ce manuscrit une contribution à l'ACV de service numérique. Cette contribution, même si centrale, n'est pas la seule. Les contributions se situent à 3 niveaux :

- Contribution méthodologique à l'ACV : adaptation de l'ACV à des cas généraux et spécifiques de services numériques
- Contribution méthodologique à l'éco-conception de service numérique : proposition de grilles de lecture pour la conception et évaluation de services numériques selon des critères environnementaux
- Contribution expérimentale : application des contributions méthodologiques précédentes à des services existants, en fonctionnement

Ces trois types de contributions se déploient dans le cadre d'un travail voulu interdisciplinaire entre la recherche en informatique d'une part, et d'autre part les domaines des sciences de la soutenabilité comme l'évaluation environnementale et l'économie circulaire.

1.4 Organisation de ce manuscrit

Après cette introduction, le présent manuscrit de thèse est organisé de la manière suivante :

Le **Chapitre 2** présente une revue des travaux ayant cherché à intégrer les enjeux environnementaux dans la recherche en informatique. Ce chapitre est construit de manière chronologique aussi bien pour la partie sur la recherche en informatique que pour celle consacrée à l'ACV.

Le **Chapitre 3** se concentre lui sur l'état de l'art concernant l'ACV de service numérique. Le corpus étudié est plus restreint, les rapports complets et ouverts étant rares. Ainsi, cette partie propose, en plus d'une revue des travaux, une analyse plus poussée des logiques communes à ces documents. En conséquent, le chapitre introduit un modèle simplifié plus haut niveau pour les ACV de service numérique.

Le **Chapitre 4** propose lui une méthode d'ACV de service numérique généraliste qui est celle utilisée dans les chapitres suivants pour les études de cas. Cette méthode fixe et synthétise une certaine vision de l'ACV attributionnelle de service numérique pour laquelle chaque étude de cas apporte un complément dans son chapitre dédié.

Le **Chapitre 5** comporte la première étude de cas concernant un unique service basé sur un modèle d'intelligence artificielle (IA) générative, Stable Diffusion, capable de générer des images à partir de texte. Cette partie contribue à appliquer la méthodologie du chapitre précédant tout en apportant un éclairage sur le contexte émergent de l'IA générative (Gen-IA).

Le **Chapitre 6** passe à l'échelle supérieure en étudiant non pas un unique service mais un ensemble de services numériques hébergés par une même entité. L'étude de cas porte sur les services proposés par l'association Picasoft qui sont majoritairement des alternatives libres à des nombreux services numériques communs, e.g. streaming vidéo, messagerie instantanée, hébergement de fichiers, édition partagée de contenu, etc.

Enfin, le **Chapitre 7** conclut le travail exposé dans ce manuscrit. Il ouvre sur une discussion plus large, en écho à l'introduction, et liste les possibles travaux futures découlant de cette thèse.

1.5 Note complémentaire : contexte scientifique de la thèse

N. B. : Cette section de l'introduction n'est pas à proprement parler nécessaire pour appréhender la suite de ce manuscrit de thèse. Toutefois, considérant l'évolution du champ pendant le travail de thèse derrière ce manuscrit, m'inspirant d'autres manuscrits de thèse, j'ai jugé enrichissant d'exposer ici le contexte scientifique entourant ce travail de thèse. Cette note est placée en introduction, similairement à une ACV où les acteurs influençant l'étude sont décrits avant l'exposition du contenu principal.

Tout d'abord, si l'ambition de cette thèse est bien d'être un pont entre la recherche en informatique et celle en soutenabilité et économie circulaire, il est peut-être plus intéressant de la placer dans une communauté naissante telle que décrit dans le tableau 1.5 issue de [32]. En reprenant les champs identifiés, cette thèse se situe entre le *Green IT* et *ICT for sustainability*, ou *TIC pour la soutenabilité*.

J'identifie par ailleurs, comme les thèses proches récemment soutenues [33–35], les conférences comme *ICT4S* comme représentant significativement cette communauté. En termes de lieu de publications, la communauté peine encore à établir les moyens adéquats de partage et promotion des travaux. Il y a d'une part des espaces qui, légitimement ou non, manquent de reconnaissance. Les conférences *ICT4S* ou *Computing within Limits* ne sont pas reconnues par les systèmes de classement de conférences, donnant lieu à une reconnaissance académique plus limitée. On notera aussi plusieurs travaux du domaine publiés dans le journal *Sustainability*, de l'éditeur MDPI, qui a fait l'objet de controverses [36, 37] sur sa qualité. D'autre part, les conférences déjà établies en informatique, semblent afficher dans leurs appels à contribution une ouverture encore mince, souvent restreinte aux enjeux de consommation d'électricité, aux enjeux de soutenabilité. Ainsi, l'espace des possibles est finalement encore assez réduit. Il est peut-être difficile de reproduire, pour un champ scientifique si jeune, le fonctionnement des conférences et journaux

| Nom du domaine | Méthodes principales | Contribution au développement durable |
|---------------------------------------|--|--|
| Informatique environnementale | Systèmes d'information, modélisation et simulation, traitement de données spatiales | Modéliser l'environnement Comprendre les systèmes complexes Partager les données et construire un consensus |
| Soutenabilité computationnelle | Modélisation, optimisation, raisonnement par contraintes, apprentissage automatique, etc. | Aide à la décision pour la gestion de ressources naturelles Trouver un compromis entre les objectifs contradictoires |
| IHM soutenable | Méthodes IHM empiriques, design, méthodes d'autres domaines | Longévité des dispositifs, promouvoir les modes de vie et comportements écoresponsables |
| <i>Green IT</i> | Gestion de l'informatique Ingénierie du logiciel | Réduire l'impact environnemental du matériel et du logiciel (le <i>Green by IT</i> étant couvert par d'autres domaines) |
| <i>ICT for sustainability</i> (ICT4S) | Méthodes d'évaluation environnementale (e.g. ACV), méthodes empiriques (incluant les sciences sociales), construction de scénarios, modélisation et simulation | Réduire les impacts (flux matériels et énergétiques) induits par le numérique Permettre des schémas soutenables de consommation et de production Comprendre et utiliser le numérique comme une technologie transformationnelle |

TABLE 1.1 – Les domaines scientifiques qui lient numérique et environnement, tableau issu de [32], traduction issu de [33]

classiques, ne serait-ce dû qu'au manque d'experts identifiés. Une reconnaissance encore plus compliquée à obtenir que le champ est jeune et ses frontières encore floues et peu connues d'autres experts.

Au sujet des frontières et dénominations floues, la notion de *numérique responsable* est parfois convoquée, avec toute fois encore une absence de consensus sur ce que la notion de *responsable* englobe. Est-ce que le *numérique responsable* inclut, en plus de la volonté de contrôler et réduire ses impacts environnementaux, une préoccupation pour l'accessibilité, l'inclusion sociale ou encore la protection de la vie privée ?

Ce flou sémantique est, encore une fois, caractéristique d'un champ scientifique à la fois jeune et dynamique. Ce dynamisme a d'ailleurs pour conséquence une activité de production de travaux très importante ces dernières années. Une manière intéressante d'illustrer ce fait est la Figure 1.5 ci-dessous. Près de la moitié des travaux cités dans ce manuscrit sont parus il y a moins de trois ans.

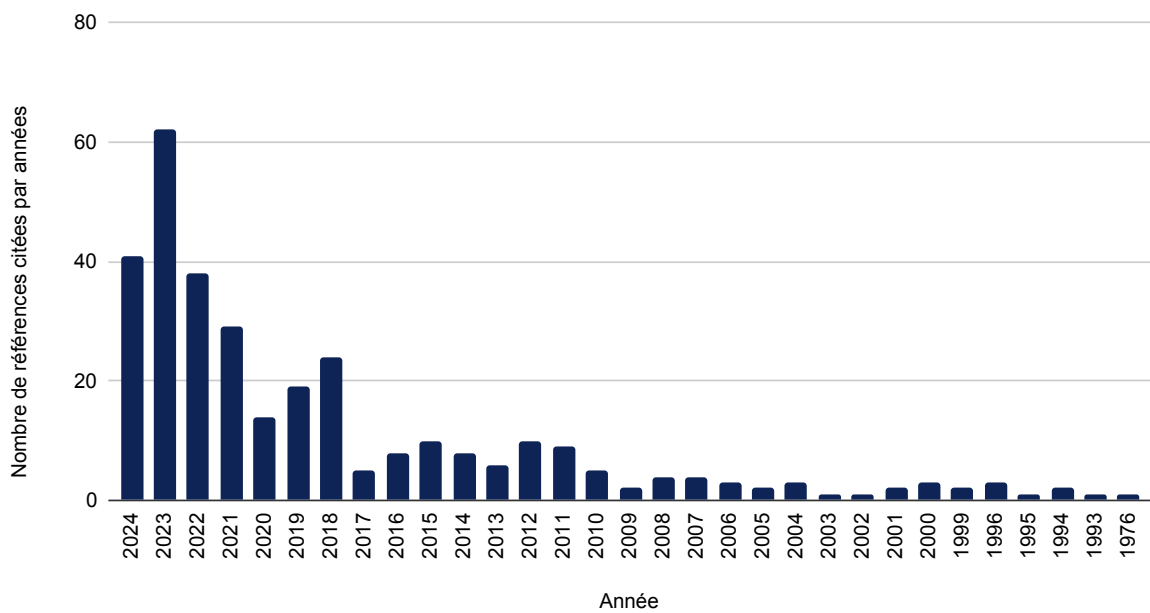


FIGURE 1.5 – Distribution temporelle des travaux cités dans ce manuscrit

La forte activité du domaine a pu rendre parfois difficile le choix de fixer certaines approches. D’autant plus que cette forte vitalité sur le sujet est le fait aussi bien de la communauté académique, que d’entreprises ou d’agences étatiques, avec des travaux à l’intersection d’une ou plusieurs de ces catégories.

Cette thèse est d’ailleurs représentative de cette convergence autour de l’enjeu de la mesure des impacts environnementaux du numérique. Elle est hébergée, via le dispositif CIFRE de l’association nationale de la recherche et de la technologie (ANRT), au sein d’un cabinet de conseil privé : OCTO Technology. Un cabinet de conseil localisé à Paris, faisant partie d’une entreprise de conseil internationale plus grande, dont certains des travaux et leurs critiques sont cités ici : Accenture. Cette position particulière me permet de témoigner ici que la dynamique scientifique autour du *Green IT* et des *ICT for sustainability* n’est pas limitée au monde académique. Bien que s’exprimant différemment, les initiatives des individus et organisations du secteur privé sont bien existantes.

Par exemple, dans le contexte de mon entreprise d’accueil, l’idée de ce projet de thèse a émergé suite au bilan carbone 2019 [38] de l’entreprise. Bilan durant lequel est apparu l’aspect très limité et insatisfaisant du bilan carbone, surtout en termes de périmètre de mesure, celui-ci n’incluant pas de base l’impact des produits vendus par l’entreprise. Mais surtout, c’est l’incapacité à mesurer l’empreinte des services numériques livrés par l’entreprise, éco-conçus ou non, qui a déclenché une prise de conscience de la nécessité d’un travail scientifique.

Pour conclure cette section et plus largement l’introduction, il est nécessaire de parler d’une des communautés les plus importantes pour ce travail de thèse : l’association Boavizta. Fondée en 2021 d’abord comme un collectif, aujourd’hui comme une association loi 1901, le groupe vise l’objectif suivant [39] :

“Permettre aux organisations publiques et privées de réellement se transformer

et se responsabiliser (de façon choisie ou régulée), et d'aligner leur numérique à une stratégie compatible avec les limites planétaires. Devenir une association interprofessionnelle, dont les communs deviennent des références internationales, qui contribuent à accélérer la réduction des impacts environnementaux directs et indirects du numérique et à infléchir les trajectoires de ces impacts. Matérialiser l'impact environnemental des transformations numériques grâce à la création, standardisation et propagation des communs dans une approche systémique, auditable, normalisée, automatisée et autonomisante.”

Ces objectifs font écho aux constats, préconisations et contributions de cette thèse. La communauté Boavizta est d'ailleurs, comme cette thèse, à l'intersection entre les communautés académique et industrielle. Les études de cas ont utilisé les communs Boavizta et contribué en retour à leur construction. La vision de la mesure comme un commun ouvert et transparent est sans doute une des solutions les plus vertueuses pour créer de la confiance envers la mesure des impacts environnementaux du numérique et par extension envers les politiques cherchant à réduire ces impacts. Une confiance d'autant plus critique que, contrairement à d'autres impacts directement mesurables, comme la consommation d'électricité dans une certaine mesure, l'ACV ne produit par définition que des estimations d'impacts potentiels. La complexité de l'ACV par rapport à une mesure directe, complexité qui permet de créer des connaissances critiques, le recours à des modèles et hypothèses, doit nécessairement se faire au prix d'une plus grande exigence de transparence.

Informatique et environnement : une prise en compte progressive

*La surprise que nous apporte
l'Anthropocène, c'est qu'il nous
force à passer d'un ancien
régime climatique à un nouveau
régime, au sens aussi scientifique
que politique de ce terme.*

Bruno Latour

Un des objectifs de cette thèse est de produire un travail scientifique à l'intersection entre la recherche en informatique et la recherche en évaluation environnementale. À cette fin, le présent chapitre décrit l'évolution de l'état de l'art dans chacune de ces deux disciplines jusqu'à aujourd'hui. D'une part, en informatique, les naissances successives de thématiques de recherche se rapprochant toujours plus d'une approche plus globale des enjeux environnementaux. D'autre part, l'émergence de l'ACV parmi les méthodes d'évaluation multicritères dans certains domaines d'industries et sa capacité à couvrir de plus en plus de secteurs. La finalité de ce chapitre est de présenter les développements scientifiques ayant conduit à poser la question de l'ACV de services numériques.

2.1 Informatique : considérer l'impact

2.1.1 Premiers travaux sur l'énergie

Parmi les thématiques de recherche en informatique actives pour réduire les impacts environnementaux liés au numérique, l'efficacité énergétique est probablement celle qui bénéficie du plus ancien héritage. Initialement loin de considérations environnementales, la communauté de recherche sur les systèmes embarqués voit naturellement la consommation d'électricité comme une contrainte. L'objectif est ici un objectif d'efficacité. On cherche à optimiser la dépense en électricité nécessaire aux fonctionnements des composants informatiques [40, 41]. Même si certains travaux appliquent presque une logique de

services, observant le coût énergétique d'applications mobiles [42], le but reste la diminution de la consommation électrique du terminal utilisateur, pas celle du service. Il est à noter, même si cela s'éloigne de l'informatique à proprement parler, qu'en parallèle de ces gains d'efficacité des systèmes, des avancées soutenues ont aussi été effectuées au sujet des batteries de ces systèmes. Ainsi, au cours des dernières décennies, on observe un progrès des batteries en termes d'efficacité, i.e. densité énergétique, mais surtout en termes de capacité pour répondre constamment aux besoins croissants des systèmes [43].

En dehors de l'informatique embarquée, les premières grandes concentrations de ressources informatiques génèrent un intérêt pour l'efficacité énergétique et plus largement le dimensionnement énergétique. La construction des premières grandes infrastructures numériques, d'abord de télécommunications [44], puis de calculs [45], amène cet enjeu. Il est d'ailleurs intéressant de voir comment des concepts comme le *Dynamic Voltage Scaling* (DVS) sont repris, du domaine des systèmes embarqués [46] vers le domaine des datacenters [47]. Concept emblématique, originaire de cette période, la notion de *Power Usage Effectiveness* (PUE)¹ est publiée en 2007 [48]. L'article en question expose consciemment les mécaniques d'escalade conjointe de l'efficacité et de la demande dans les datacenters. La concentration de nombreux équipements au même endroit change le modèle économique, l'achat de matériels n'est plus l'unique déterminant face au coût croissant de l'infrastructure et de l'énergie².

Initiées au début des années 2000, ces recherches de l'efficacité énergétique des grandes infrastructures numériques se développent au cours de la décennie 2010, aussi bien dans les parties réseaux [50] que datacenters [51, 52], et qui se poursuivent encore aujourd'hui [53]. On peut voir dans ces gains la tendance décrite par la *loi de Koomey*, qui observait de 1946 à 2009 un doublement de l'efficacité énergétique en moyenne tous les 18 mois [54]. Le Tableau 2.1 de l'IEA [55] synthétise les tendances ainsi décrites. En effet, la charge de travail augmente bien plus que le coût énergétique lié, pour les réseaux comme les datacenters, illustrant des gains d'efficacité certains.

Toutefois, ce tableau illustre aussi une partie du problème de la relation entre numérique et environnement. Il est indéniable que la consommation totale d'électricité des infrastructures décrites a augmenté. Est-ce que les gains d'efficacité ont permis de contenir cette croissance, ou bien en sont-ils une des causes directes ? Il n'y a pas nécessairement de réponse binaire à cette question, mais le phénomène observé semble correspondre à un effet rebond [30]. Les gains d'efficacité énergétique auraient permis de nouveaux usages du numérique qui auparavant n'aurait pas été viables, économiquement ou physiquement parlant. Les gains d'efficacité énergétique des systèmes embarqués sont à l'origine de l'explosion du nombre d'objets connectés et de smartphones, qui sinon n'auraient fonctionnellement pas pu exister. De même pour l'efficacité des infrastructures de calculs et communications, un usage destructeur pour l'environnement [56–58] comme le minage de crypto-monnaie est extrêmement dépendant d'un coût de calcul bas pour exister [59].

En conclusion, regarder l'efficacité énergétique comme un moyen de réduire l'empreinte du numérique s'avère au mieux incomplet, au pire incorrect. Les développements observés dans cette section reflètent une vision de l'économie d'énergie restreinte aux systèmes

1. Efficacité de l'utilisation de l'énergie.

2. Fait amusant avec le recul, montrant de manière certaine l'importance prise par le coût de l'infrastructure : Le cabinet Gartner prédit en 2006, qu'en 2008 50% des datacenters américains manqueront de capacité de refroidissement ou de fourniture d'énergie et ne pourront plus répondre à la demande [49].

Global trends in digital and energy indicators, 2015-2022

| | 2015 | 2022 | Change |
|---|-------------|-------------|-------------|
| Internet users | 3 billion | 5.3 billion | +78% |
| Internet traffic | 0.6 ZB | 4.4 ZB | +600% |
| Data centre workloads | 180 million | 800 million | +340% |
| Data centre energy use (excluding crypto) | 200 TWh | 240-340 TWh | +20-70% |
| Crypto mining energy use | 4 TWh | 100-150 TWh | +2300-3500% |
| Data transmission network energy use | 220 TWh | 260-360 TWh | +18-64% |

Sources: Internet users [ITU (2023)]; internet traffic [IEA analysis based on Cisco (2015); TeleGeography (2022); Telegeography (2023); Cisco (2019), Cisco Visual Networking Index]; data centre workloads [Cisco (2018), Cisco Global Cloud Index]; data centre energy use [IEA analysis based on Malmodin & Lundén (2018); ITU (2020); Masanet et al. (2020); Malmodin (2020); Hintemann & Hinterholzer (2022); Malmodin et al. (2023)]; cryptocurrency mining energy use [IEA analysis based on Cambridge Centre for Alternative Finance (2023); Gallersdörfer, Klaaßen and Stoll (2020); McDonald (2022)]; data transmission network energy use [Malmodin & Lundén (2018); Malmodin (2020); ITU (2020); Coroama (2021); GSMA (2022); GSMA (2023); Malmodin et al. (2023)].

FIGURE 2.1 – Synthèse issue d'un rapport de l'IEA concernant la demande en énergies des infrastructures de télécommunication et les datacenters de 2015 à 2022 [55]

étudiés, qu'ils soient un système embarqué ou une infrastructure distribuée. Même dans ce dernier cas, la consommation d'électricité ne prend que peu, voir pas, en compte la société dans laquelle s'insère cette technologie, et encore moins les enjeux environnementaux. Il n'est pas ici question de représenter les recherches sur l'efficacité énergétique comme nuisibles, mais de souligner qu'elles sont à elles seules insuffisantes face aux enjeux actuels. Néanmoins, l'efficacité et l'économie d'énergie ont parfois, dans certaines communautés de recherche en informatique, été la première étape vers une réflexion plus globale sur les impacts du numérique³.

Dernier point, il est raisonnable de : prendre du recul sur la loi de Koomey, aussi bien constater, comme pour la loi de Moore, l'épuisement de la tendance [60], reconnu déjà par Koomey lui-même [61], et de questionner cette posture optimiste et ses conséquences [62]. La consommation d'électricité du numérique dont il a été question dans cette partie n'est qu'une part de l'empreinte environnementale du numérique, ce n'est aussi qu'une part de

3. À titre d'exemple, le groupe CNRS ÉcoInfo a pu avoir à ses débuts un tropisme vers les enjeux d'efficacité énergétique (<https://ecoinfo.cnrs.fr/2009/>) avant que d'autres thématiques prennent ou reprennent de l'importance.

l’empreinte énergétique. Les travaux précédemment présentés ne se focalisaient que sur la **consommation d’électricité de la phase d’usage** du numérique, étudiant souvent leur sujet indépendamment du reste de l’écosystème numérique.

2.1.2 L’émergence du "Green IT"

Formé en 1988, le Groupe d’experts intergouvernemental sur l’évolution du climat (GIEC) publie en 1990 son premier rapport [63]. En 1997, le protocole de Kyoto est adopté, il entre en application en 2005. Cet accord international vise à réduire les émissions de GES de 5% entre 2008 et 2012, au niveau mondial. L’objectif de l’atténuation du réchauffement climatique est ensuite poursuivi par l’accord de Paris en 2015, mis en application dès 2016. Ce contexte, de prise de conscience et de lutte contre les émissions de GES, influence le secteur du numérique. Par intérêt scientifique ou par incitation économique⁴, la question du calcul des émissions du secteur est posée, de même que celle des moyens pour réduire ces émissions.

Entre 2007 et 2020 sont publiées plusieurs études qui illustrent dans le secteur des TIC l’enjeu des émissions de GES. La Figure 2.2 présente les différentes estimations produites durant ces années. Reprenant les conclusions et l’approche de [21], la régression linéaire montre la tendance à l’augmentation, mais celle-ci n’est pas forcément linéaire comme le montre la Figure 2.3 compilant différentes hypothèses de croissance. L’enjeu des émissions carbone du numérique s’accompagne des premières controverses sur la taille et la répartition de l’empreinte [23], comme le montre la dispersion des résultats de la Figure 2.2. Controverse difficilement évitable au vu du peu de maturité de la communauté sur les questions de cycle de vie complet ou sur le périmètre du numérique.

L’enjeu des émissions de GES du secteur étant posé, on observe le développement de ce que l’on pourrait qualifier de seconde vague de GreenIT, dans le prolongement de la première⁵. Elle se caractérise par la prise en compte d’enjeux extérieurs aux systèmes étudiés, comme l’enjeu du réchauffement climatique. La première vague, même si elle traitait parfois d’infrastructures distribuées à large échelle, traitait son sujet comme un système fermé. Qu’il s’agisse d’un smartphone ou d’un datacenter, il convenait d’optimiser la ressource énergie, de rester dans les frontières du système.

Cette ouverture à l’enjeu carbone se fait néanmoins clairement dans le prolongement des précédents travaux. Ainsi, l’indicateur *PUE* est complété par l’indicateur *Carbon Usage Effectiveness*⁶ (*CUE*) en 2010 [68]. Ce ratio d’efficacité est la somme des émissions de GES⁷ issues de la consommation d’électricité du datacenter divisée par la consommation totale d’électricité du datacenter. La focalisation choisie par les auteurs du *CUE* sur la consommation d’électricité issue de sources à forte émission de GES est assez représentative de plusieurs travaux restant focalisés sur l’empreinte carbone de la phase d’usage du numérique.

Si la notion de coût carbone, du datacenter, des services ou des workloads, est intégrée, cela reste dans le cadre d’une politique d’optimisation de l’utilisation d’énergie dites

4. Par exemple, le système d’échange de quotas d’émission de l’Union européenne est lancé en 2005, ce marché de "quotas carbone" a pour objectif d’être harmonisé avec les autres depuis l’accord de Paris.

5. Qui serait celle décrite dans la précédente section et correspond décrite par Mingay en 2007 [67].

6. Efficacité de l’utilisation du carbone.

7. en équivalent CO_2 .

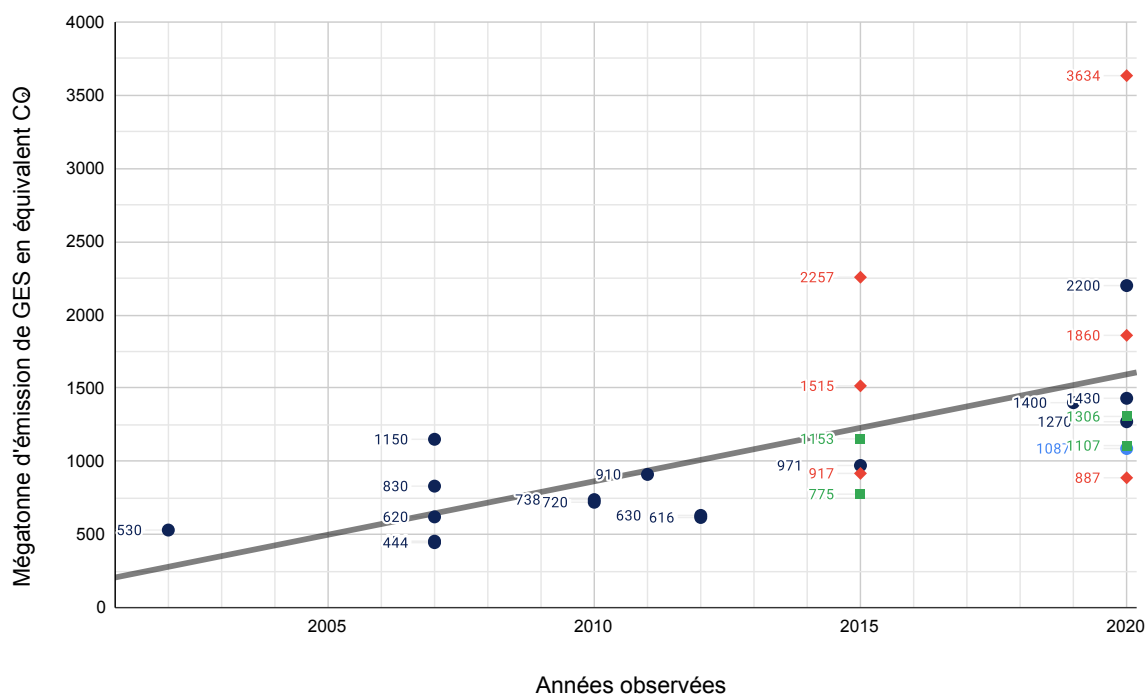


FIGURE 2.2 – Estimations de l'empreinte carbone du secteur des TIC issus de différentes études académiques et industrielles selon l'année observée. Pour une même année, les ■ représentent les valeurs minimales et maximales d'une même étude, de même les ♦ représentent les scénarios minimaux, maximaux et attendus. Graphique repris de [21], augmenté par les données de [64], [65] et [66]

décarbonée comme celle issue de la production solaire, éolienne ou hydraulique. Ainsi, un des premiers travaux sur le sujet [69] exprime très clairement des objectifs financiers connectés au marché du crédit carbone. Cette vision reprend des techniques d'optimisation déjà existantes comme la migration, temporelle ou géographique, de workload pour créer des approches dites "opportunistes" [70, 71]. Ces algorithmes opportunistes cherchent à déplacer des workloads, géographiquement comme temporellement, vers des nœuds de calcul connectés à des unités de production d'électricité solaire ou éolienne. Dans ce cadre, le défi principal est en effet d'adapter la consommation du datacenter, qui dépend de la demande qu'il doit satisfaire avec le niveau de *Service Level Agreement* (SLA)⁸ adéquat, à la production d'énergie décarbonée. Cette énergie provient souvent de sources non-pilotables comme le solaire ou l'éolien, i.e. qui ne contrôlent pas quand elles produisent de l'électricité. Durant cette décennie, un champ de recherche est donc ouvert pour optimiser ces paramètres, la thèse suivante étant une bonne représentation de cette approche [72].

En dehors des optimisations algorithmiques, mais en restant focalisé sur l'enjeu d'adapter le numérique aux énergies renouvelables non-pilotables, cette période est aussi marquée par construction de plateformes expérimentales directement auto-alimentées. Le projet

8. La SLA est l'engagement contractuel du datacenter à répondre à la demande selon certains critères de qualité de service.

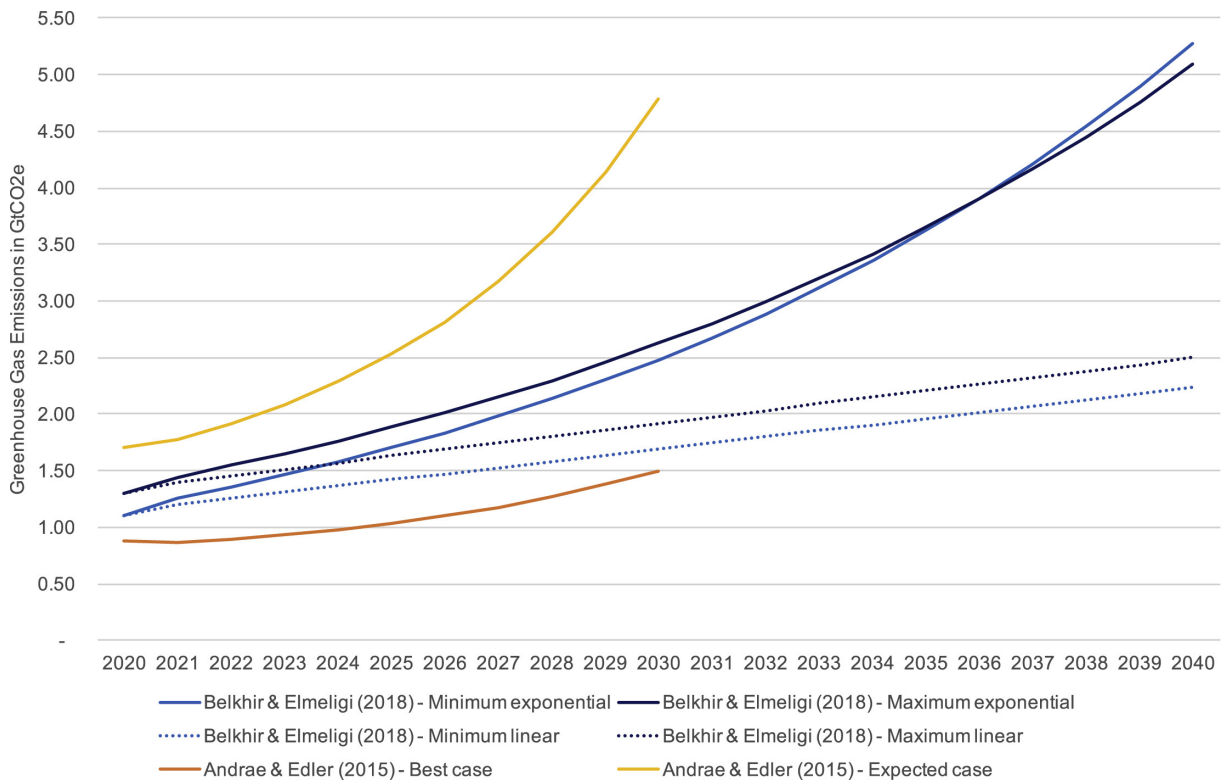


FIGURE 2.3 – Comparaison des différentes projections d’émissions de GES pour le secteur des TIC selon [64] et [65], graphique issu de [21]

Parasol [73] est un datacenter auto-alimenté par un panneau solaire, tandis qu’à plus large échelle, le projet GreenStar [74] met en place une infrastructure cloud et réseaux plus large, elle aussi auto-alimentée par une série d’unités de production d’électricité.

Les limites de ce cadrage sur les émissions carbone de la phase d’usage sont multiples. Tout d’abord, les stratégies d’optimisation de la charge dans les datacenters, suivant la production d’énergie renouvelable non-pilotable peuvent être redondantes avec celles déjà mises en place par les gestionnaires de réseaux électriques ou pire entrer en contradiction avec ces stratégies. Les réseaux électriques mettent déjà en place des stratégies afin d’intégrer les productions d’énergies renouvelables dans leur mix. La "perte" d’électricité renouvelable produite, mais non utilisée, est souvent marginale du moment que les unités de production sont connectées au réseau [75, 76]. On pourrait alors voir l’implémentation incontrôlée de ces solutions comme un risque pour la stabilité du réseau électrique. Le déploiement non-supervisé de grande capacité de calcul a pu dans certains pays mettre en péril la stabilité du réseau [59]. Si ce dernier exemple concerne le minage de cryptomonnaies, il n’est pas exclu d’arriver à des situations similaires tant les réseaux électriques de certains pays comme l’Irlande [77] ou le Danemark sont devenus dépendants de l’industrie datacenter [78]⁹. On pourrait d’autant plus regretter les conséquences de ces stratégies opportunistes que celles-ci pourraient inciter la création d’usages artificiels [79] mettant en avant l’utilisation d’énergie "verte" qui sinon serait "perdue".

9. En 2024 l’IEA s’attend à ce que la part des datacenters dans la consommation d’électricité au niveau national atteigne en 2026 près de 20% au Danemark et plus de 30% en Irlande.

Une autre faiblesse de ces stratégies, qui n'est pas intrinsèque, est l'utilisation dans le calcul du coût carbone soit d'un facteur d'émission reflétant l'empreinte environnementale moyenne du mix électrique locale, soit l'empreinte correspondant au contrat passé avec le distributeur d'électricité. La première option est dite *location-based*, elle prend l'empreinte carbone du mix électrique locale où est utilisée l'électricité, c'est un critère géographique. La seconde option, dite *market-based*, elle prend l'empreinte carbone du mix électrique que le fournisseur d'électricité s'engage à fournir. On peut avoir ainsi des contrats de fourniture d'électricité "garanti 100% énergies vertes" ou décarbonées, c'est une pratique courante dans l'industrie datacenter. Il s'agit dans ce cas d'un critère contractuel. Ce choix de facteur d'émission, *location-based* ou *market-based*, influence significativement l'intérêt environnemental ou économique de la stratégie. Malgré certains défenseurs, l'approche *market-based* est fortement critiquée [80], notamment par la communauté ACV [81], comme ne représentant pas la réalité des impacts environnementaux¹⁰ et n'ayant pas forcément les effets incitatifs désirés à large échelle. Ce fait, couplé aux approches précédemment citées qui admettent très clairement la recherche d'optimisation par rapport au marché carbone, i.e. dans une approche *market-based*, rend très limité le gain environnemental. Si les projets auto-alimentés précédemment cités échappent à cette critique, les cas d'usage de ces infrastructures auto-alimentées restent à trouver dans le cadre d'un projet de réduction de l'empreinte du numérique.

Au-delà de ces problèmes, la principale limite intrinsèque de ce que nous avons pu décrire ici comme le *carbon-aware computing* est justement cette approche mono-critère autour des émissions de GES. Les critiques contre la *carbon tunnel vision* [82] sont nombreuses¹¹, d'autant plus lorsque l'on observe uniquement l'empreinte carbone de la phase d'usage. Dans ce cadre, des équipements utilisant des batteries sont jugés au même niveau que ceux connectés au réseau électrique, cachant le coût de fabrication des batteries. Plus grave, ce cadre peut inciter à localiser des datacenters dans des régions où l'empreinte carbone de l'électricité est faible, mais où le stress hydrique est fort. C'est par exemple ce qu'il se passe récemment dans le sud de l'Espagne [83] où l'entreprise Meta qui a pris un engagement d'utiliser à 100% de l'énergie renouvelable, va construire un datacenter. Cet exemple a sans doute pu avoir un écho certain dans la presse due au fait que c'est un projet de construction immobilier assez visible, mais pourrait-on avoir la même vigilance face au déplacement plus discret de workload ?

Plus généralement, le *carbon-aware computing* est aussi souvent limité à une vision étroite de l'empreinte carbone de part sa non-prise en compte de l'empreinte carbone liée à la fabrication des équipements. On peut voir dans ce défaut un marqueur de la période où s'est développée cette approche. Comme le décrit bien [84] durant le début du XXI^e siècle, en termes de distribution entre les phases du cycle de vie, l'empreinte carbone du numérique est progressivement passée de la phase d'usage à la phase fabrication. Sans chercher à évaluer l'action du *carbon-aware computing*, on peut constater que se focaliser sur la phase d'usage avait plus de sens à l'époque de ces recherches.

Il n'en reste pas moins que même en ne gardant qu'une approche purement carbone,

10. "Absence de relation causale entre les impacts réels et les règles actuelles relatives aux instruments de marché" [81].

11. Même si certains observateurs pointent que souvent la baisse de l'empreinte carbone est corrélée à la baisse globale de l'empreinte, ce qui est vrai, ce n'est pas une règle générale particulièrement dans des situations où l'empreinte du mix électrique est très décarbonée.

l'absence de cycle de vie complet biaise cette recherche dans ses objectifs de réduire la contribution du numérique au réchauffement climatique. Cette approche n'a pas réussi à tracer un chemin différent que ceux de la croissance de l'empreinte, comme illustré par la Figure 2.3. Comme le conclut [85] les optimisations n'ont pas suffi et ne peuvent être le seul levier, il faut viser une sobriété dans l'usage comme l'innovation pour utiliser le plein potentiel des gains d'efficacité afin de répondre aux enjeux environnementaux.

En conclusion, de manière similaire à la précédente section, il n'est pas question ici de dévaloriser ces recherches, mais de souligner leurs limites face aux ambitions de contrôler et réduire l'empreinte du numérique. Et de même que pour l'efficacité énergétique, l'efficacité carbone a pu ouvrir la voie à des recherches prenant plus globalement en compte l'empreinte du numérique. Par exemple, parti de ce paradigme d'utilisation des énergies renouvelables non pilotables, le projet Datazero [86] a intégré par la suite l'impact carbone du cycle de vie complet dans des stratégies d'optimisation [87]. De même, comme le pointe [88], très critique sur les bénéfices du *carbon-aware computing*, si certains pays font le choix de continuer d'augmenter la part d'énergie renouvelable non-pilotable, les déplacements de workload pourraient en coordination avec les gestionnaires de réseaux être un outil pertinent.

2.1.3 Plus que de l'électricité, plus que du carbone

Il est évident que comme toute activité humaine, le numérique affecte l'environnement de plusieurs manières, bien au-delà du réchauffement climatique. Contrairement aux deux précédentes sections, la présente section représente les travaux cherchant à inscrire pleinement l'activité informatique au sein de l'environnement et de ses multiples problématiques.

Comme nous le verrons dans les sections suivantes, les premières approches multicritères prenant en compte tout le cycle de vie des équipements numériques sont arrivées il y a déjà plus de 10 ans. Ainsi, en 2012 sont publiés les premiers travaux sur l'ACV d'un datacenter [89], la même année sont aussi publiées les premières recommandations méthodologiques de l'*International Telecommunication Union* (ITU)¹² [25]. D'autres contributions émergent dès le début des années 2010 [90,91] mais sont souvent extérieures à la communauté informatique, elles sont le produit d'autres champs scientifiques comme le design, l'ingénierie, ou viennent de cabinet de conseil en environnement.

Il n'y a pas encore d'équivalent à ce qu'est le *carbon-aware computing* pour une communauté cherchant à prendre en compte dans la pratique de l'informatique l'ensemble du cycle de vie des équipements et un plus large éventail d'impacts environnementaux. La communauté appelée *Sustainable ICT* et annoncée en 2015 [32] n'émerge que lentement à la fin de la décennie 2010 et au début de la décennie 2020. Ainsi, l'équivalent des études de la Figure 2.2 ne sort que fin 2021 [18]. Initiés en 2019 [66] les travaux amenant la Figure 2.4 sont une manifestation de l'intérêt croissant de la communauté informatique pour l'ACV multicritère.

La distribution des impacts affichée par la Figure 2.4, bien que dominée par la partie relative aux terminaux utilisateurs, invite à considérer le problème de manière globale, aucun des niveaux n'ayant de raison d'être sans les autres. Sans toutefois produire de

12. Union Internationale des télécommunications.

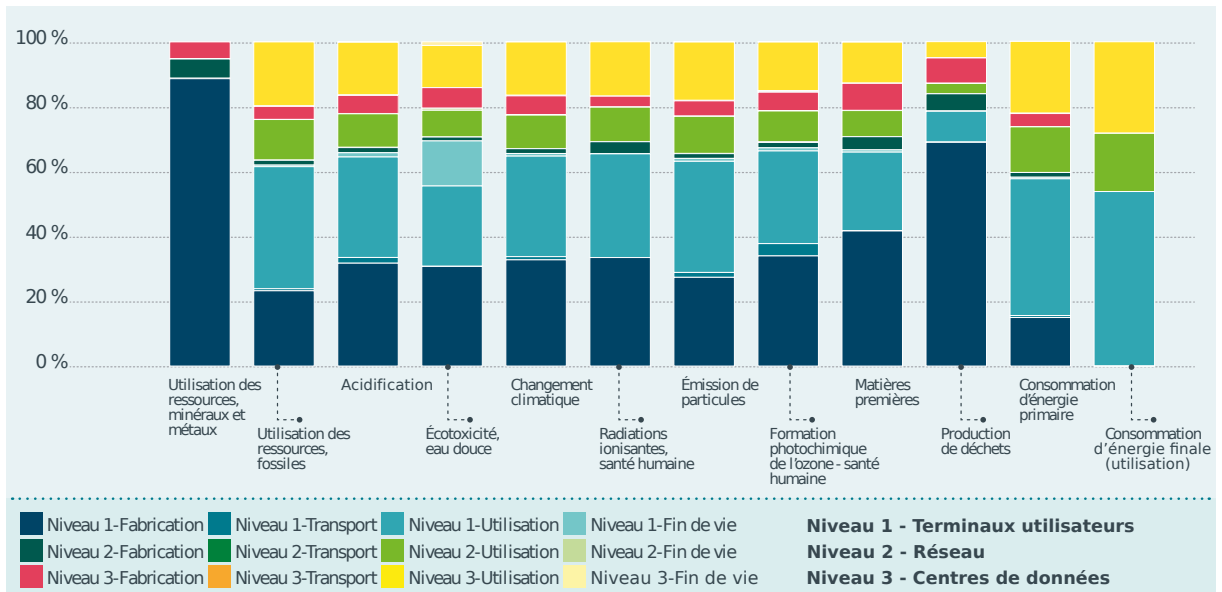


FIGURE 2.4 – Distribution de l'impact environnemental du numérique en 2019, par phase de cycle de vie, par niveau, graphique issu de l'étude [66] (modifié)

résultats quantitatifs, l'article [92] est un des premiers à souligner comment le design du service, notamment dans ses choix d'interaction avec l'utilisateur, influence l'ensemble du service et ses impacts environnementaux. Il est souligné l'interaction entre les biens de consommation individuelle que sont les terminaux et les infrastructures mutualisées que sont le réseau et le Cloud. Dans cette même lignée de l'angle "design de service", [93] propose un framework d'évaluation de projet informatique soutenable¹³. Bien que ne considérant que l'empreinte énergétique de la phase d'usage, [94] prend en compte l'ensemble du service publicitaire en ligne. Dans la même ligne, l'article [95] utilise des éléments issus de la méthode ACV pour quantifier les émissions de CO_2eq . liées à la phase d'usage du service.

Cette première série de travaux a la caractéristique de rester souvent à un assez haut niveau d'abstraction, en réalisant ses estimations dans une approche *top-down* ou en se concentrant sur l'élaboration de frameworks généraux. On peut expliquer ces particularités par la difficulté d'accès à des données primaires, aussi bien dans les bases d'Impact de Cycle de Vie (ICV) que dans les services numériques eux-mêmes. Les articles [96] et [97] soulignent aussi bien le potentiel de l'ACV comme outil pour transformer radicalement l'industrie numérique, que les difficultés faisant obstacle à son usage. Pour en citer quelques-unes : la difficulté d'accès aux données ICV, le manque de transparence des acteurs de l'industrie sur le fonctionnement exact des infrastructures de communication et de calcul, ou encore la difficulté d'établir des périmètres d'études clairs. Les données ICV sont d'autant plus critiques que, pour certaines catégories d'impacts environnementaux comme le réchauffement climatique, l'empreinte carbone du numérique s'est tendanciellement déplacée de la phase d'usage à la phase de fabrication [84]. Si certains obstacles persistent toujours, comme la très difficile inclusion de la phase de fin de

13. Sustainable Computing Evaluation Framework (SCEF).

vie des équipements [98, 99], des progrès significatifs sont observés au début des années 2020. Notamment sur l’empreinte de la fabrication des équipements où des initiatives pour créer des données ICV sont apparues [100, 101], indépendamment des fournisseurs classiques de données ICV.

On peut observer des avancées sur des segments spécifiques du service numérique. Par exemple, un travail est fait pour éclaircir l’empreinte du transfert de données via les infrastructures réseaux [102]. Le réseau a une consommation d’énergie constante en temps normal, le coût en électricité de la phase d’usage est ainsi difficile à évaluer, car il n’y a pas de causalité directe entre la transmission d’un gigaoctet de données et une dépense énergétique. La solution est alors souvent de prendre une valeur moyennée, aussi bien pour l’impact de la phase d’usage que pour l’empreinte incarnée¹⁴ du réseau. Cette option est souvent préférée pour sa simplicité qui ne demande pas de connaître précisément la configuration du réseau. D’autres travaux poursuivent néanmoins des évaluations plus fines sur le coût de transport des données dans une logique ACV [103, 104]. D’autres recherches pointent au contraire les limites de l’approche ACV attributionnelle, où une action, transférer un volume de données à travers le réseau, est associée à un coût direct. Suite à ces critiques, une approche inspirée de l’ACV conséquentielle est préférée [105], en cherchant à évaluer les conséquences sur la croissance de l’infrastructure de services transférant de grands volumes de données.

La "communauté datacenter" voit aussi apparaître, progressivement, une prise en compte plus large des enjeux environnementaux. À la suite des recherches précédentes en efficacité énergétique et carbone, d’autres métriques d’efficacité sont proposées [106]. Une peut particulièrement retenir l’attention, la *Water Usage Effectiveness*¹⁵, qui, diffusée avec peu d’impact en 2011 [107], a été intégré institutionnellement 10 ans plus tard, sous forme de standard ISO [108] et de réglementation au niveau de l’Union européenne [109]. Cette progression de l’enjeu de consommation de l’eau, illustre l’intérêt de la communauté, industrielle et académique, pour ce sujet [110] et les usages liés [111]. Autre nouvelle perspective sur l’efficacité des datacenters, le renouvellement des équipements fait écho à l’enjeu du coût environnemental de la phase de fabrication. Certaines études proposent donc de prendre en compte et d’optimiser ce renouvellement du matériel afin de réduire ce segment de l’empreinte [53, 112].

Ces avancées sur les parties réseaux et datacenter vont de concert, et permettent d’une certaine manière, l’étude de cas d’usage plus précis du numérique. On peut par exemple observer dans la continuité des recherches en efficacité sur le *storage as a service*, de plus en plus de rapports ACV sur les équipements de stockage de données [113], ainsi que des approches tentant d’inclure le cycle de vie complet des équipements pour ce sous-secteur du numérique [114].

Autre sous-domaine, remis en perspective par l’approche ACV, l’Intelligence Artificielle (IA) voit son empreinte reconsidérée par l’ajout de nouveaux critères environnementaux [111] ou l’approche par cycle de vie complet [115, 116]. Cette nouvelle perspective permet d’aller au-delà des aspects logiciels de l’IA et porter la recherche sur l’impact du déploiement matériel de ces technologies, que ce soit sur un Cloud classique, des smartphones ou des objets connectés [117, 118]. Une contribution de cette thèse sera d’ailleurs

14. Désigné comme *embodied* en anglais, c’est une manière simplifiée de désigner tout l’ensemble des coûts environnementaux qui ne relève pas de l’usage direct comme la phase de fabrication et de fin vie.

15. Efficacité de l’utilisation de l’eau.

sur l'application d'une méthode d'ACV de service numérique à une technologie basée sur de l'IA au chapitre 5 de ce manuscrit.

Pour finir ce tour d'horizon, parlons brièvement des ACVs de services numériques. Si les premières études mentionnées précédemment restaient sur l'impact de l'usage, souvent en mono-critère, le début de la décennie 2020 a vu la publication de plusieurs ACV, à l'échelle service, prenant en compte le cycle de vie complet des équipements et produisant des résultats quantitatifs sur plusieurs catégories d'impacts environnementaux. Ces études portaient sur divers services, blockchain [119], visioconférence [120] ou messagerie instantanée [121]. Ces travaux seront analysés plus spécifiquement en détail dans le chapitre 3. Il faut retenir que cette approche de l'empreinte du numérique par le service, permet de mieux refléter l'empreinte du déploiement et de l'usage du service. Associés par exemple à des études reflétant le coût de développement de ces services [122], les ACV de services numériques fournissent des données précieuses pour juger l'empreinte environnementale totale du déploiement d'un service. Ces données peuvent ainsi servir à comparer des services entre eux, mais aussi avec des alternatives non-numériques [33, 123, 124]. Par conséquent, il est même possible de pousser les analyses, sur la base de ces résultats d'impacts directs, vers l'estimation d'impacts indirects [125], des estimations qui permettraient de mieux prévoir un possible effet rebond [14, 126], permettant ainsi de dépasser certaines limites du *carbon-aware computing*. Toutefois, pour mener au mieux cet exercice, il sera nécessaire de dépasser le cadre de l'ACV attributionnelle pour s'inspirer de l'ACV conséquentielle comme l'ont déjà fait certains travaux [127].

En conclusion, on observe bien l'apparition d'un ensemble de travaux, ancrés dans la discipline informatique, qui intègrent dans leur démarche des éléments issus de l'ACV comme la prise en compte du cycle de vie complet des équipements, le multicritère, une approche réflexive considérant l'objet étude, service ou bien, comme interagissant avec l'environnement. Cette vision cherchant à établir la place que prend le numérique dans les limites planétaires, s'inscrit dans un processus de réduction de l'empreinte du numérique qui souhaite prendre en compte comme levier aussi bien l'efficacité technique que le changement des usages qui sera peut-être nécessaire pour un numérique restant dans les limites planétaires [128].

2.2 L'analyse de cycle de vie

Dans de nombreuses publications scientifiques, l'analyse de cycle de vie (ACV) est présentée comme une méthode d'évaluation des impacts environnementaux basée sur les normes ISO 14040 [129] et 14044 [130], éventuellement complété par certains standards sectoriels. Cette introduction donne souvent à l'ACV l'image d'un processus fixe, statique, d'une rigidité telle qu'on pourrait sans doute facilement l'automatiser. Ces jugements sont cependant grandement à nuancer [131]. Pour cela, la notion d'ACV sera abordée sous un angle historique afin de mieux comprendre sa définition, ses limites et surtout sa nature de sujet de recherche vivant, à l'évolution constante. La mise en valeur de cet aspect rendra plus claire l'utilisation qui est faite de l'ACV dans le cadre des services numériques et comment cette thèse s'inscrit dans la recherche en évaluation environnementale.

2.2.1 Perspective chronologique et définition

Il est d'usage de faire remonter l'histoire de l'ACV à une première étude interne effectuée en 1969 au sein de la Coca-Cola Company [132]. Cette étude, portant sur le choix du type de contenant, avait la caractéristique de faire un bilan comparatif de l'ensemble des ressources, flux et impacts environnementaux pour chaque solution. À cette époque, cette méthode ne s'appelait d'ailleurs pas encore ACV, cette dénomination ne s'ancrant que 20 ans plus tard dans les années 90. Durant ces deux décennies, la méthode appelée REPA¹⁶, va peu à peu s'enrichir d'apports, en lien avec les préoccupations environnementales de chaque époque. Les années 70 vont consolider l'intérêt pour l'efficacité énergétique [133], dû aux chocs pétroliers, les années 80 insisteront sur la production de déchets, cette instabilité de la méthode motive la constitution de l'ACV en norme [134].

Les années 90 sont un moment d'affirmation pour la méthode ACV. En 1996 le *International Journal of LCA* est fondé, donnant lieu à un espace académique autour de la méthode. Surtout, de 1997 à 2006 sont publiés successivement les normes ISO de la série 14040, de la 14040 à la 14044. En 2007, une branche indépendante de l'ACV est créée, l'analyse sociale du cycle de vie (AsCV) [135]. Cette méthode désire aller au-delà de l'estimation quantitative d'impacts environnementaux en explorant avec les méthodes des sciences humaines et sociales les coûts sociaux d'un produit ou d'un processus.

Au fil des années, l'ACV gagne en popularité, comme l'atteste la Figure 2.5, depuis divers secteurs, voir Figure 2.7. Cette notoriété croissante de l'ACV, comme outil entre monde industriel et académique, est indissociable de son adoption croissante comme norme européenne. Après des travaux préliminaires, la commission européenne adopte en 2013 une série de recommandations [136] prélude de l'adoption de l'ACV comme norme de déclaration des impacts environnementaux des produits (PEF)¹⁷ [137] et des organisations (OEF)¹⁸ [138].

La progression de l'usage de l'ACV participe à la stabilisation de sa définition. Les quatre principales phases telles que décrites par la norme ISO sont représentées sur la Figure 2.6, les flèches entre les étapes sont bidirectionnelles afin de représenter la nature itérative de la méthode, des aller-retours entre les phases pouvant être nécessaires. Brièvement, ces quatre phases peuvent être décrites comme :

1. Cadrage de l'étude, quels types de résultats cherche-t-on à obtenir, quel est le périmètre géographique et temporel de l'étude, quelles sont les frontières du système? Cette phase comprend la définition des unités fonctionnelles (UF) évaluées durant l'étude. Les UF sont une manière de quantifier la satisfaction du besoin apportée par le produit ou le service étudié.
2. L'analyse des inventaires de cycles de vie (ICV) et flux élémentaires utilisés par l'objet d'étude. Il s'agit de comptabiliser les flux, comme l'électricité, nécessaires à la réalisation de l'UF. De même, on comptabilise dans cet inventaire les produits et services nécessaires à l'UF en intégrant leurs données ICV dans l'inventaire. Ces données ICV viennent généralement de bases de données compilant des ACV précédemment réalisées.

16. Ressource and Environmental Profil Analysis.

17. *Product Environmental Footprint*.

18. *Organisation Environmental Footprint*.

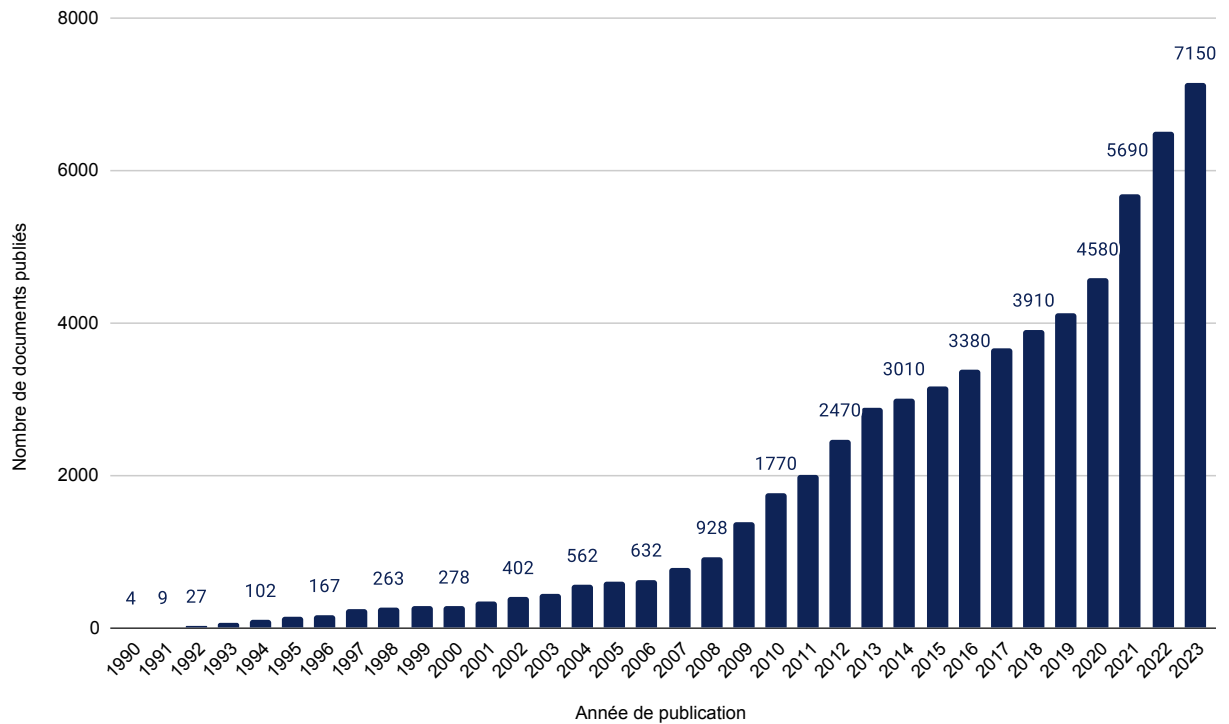


FIGURE 2.5 – Évolution du nombre de publications scientifiques mentionnant l'ACV chaque année. Statistiques réalisées en observant le nombre de documents retournés par une requête Google Scholar contenant : “life cycle analysis” OR “life cycle assessment” OR “Life cycle analysis” OR “Life cycle assessment”

3. L'analyse des impacts des cycles de vie (AiCV) classe et regroupe les impacts environnementaux découlant de l'inventaire précédent. Elle va choisir les catégories d'impacts environnementaux à évaluer. Elle va produire les résultats bruts dits caractérisés.
4. Phase d'analyse réflexive sur l'ACV. Est-ce que les résultats obtenus et la démarche effectuée durant l'ACV est cohérente avec les objectifs établis ? Si des hypothèses ont été prises, l'on pourra voir l'incertitude que créent ces paramètres en observant l'influence de leur variation sur les résultats finaux lors d'analyses de sensibilités (AS).

Les 4 principales caractéristiques définissant l'ACV selon [131] peuvent être synthétisées ainsi :

- Avoir une perspective sur tout le cycle de vie, i.e. considérer au maximum l'ensemble des étapes du cycle de vie de l'objet d'étude et pas seulement la phase d'usage ou la phase de fabrication.
- Couvrir un large spectre de problèmes environnementaux, i.e. inclure au possible plusieurs catégories d'impacts environnementaux dans l'étude, pas seulement par exemple l'empreinte carbone ou le coût énergétique.
- Produire des résultats quantitatifs, produire des estimations pour chaque catégorie d'impacts qui pourront être comparés à d'autres études quantitatives.

- Avoir une approche scientifique, i.e. rechercher la fiabilité issue de la méthode scientifique, utiliser des hypothèses cohérentes avec l'état de l'art, prôner rigueur et transparence dans la méthode.

Il est important de noter qu'en pratique, du moins dans les ACVs détaillées au chapitre suivant ou dans les travaux de la section 2.1.3 ou même 2.2.3, ces 4 caractéristiques sont rarement toutes réunies. L'ACV pouvant être vue plus comme un moyen qu'une fin, il n'est pas forcément pertinent de considérer une étude comme relevant ou non de l'ACV de manière binaire. De nombreux auteurs préfèrent présenter leurs travaux comme "inspirés de l'ACV" pour cette raison. Toutefois, d'un point de vue scientifique, il est pertinent de rattacher à l'ACV, et de critiquer comme tel, une étude partageant plusieurs caractéristiques propres à l'ACV tel que présentée ici. D'autant plus que l'ACV a connu un nombre galopant de variantes au sein même de sa communauté [139], le plus souvent dans un objectif de simplification [140]¹⁹.

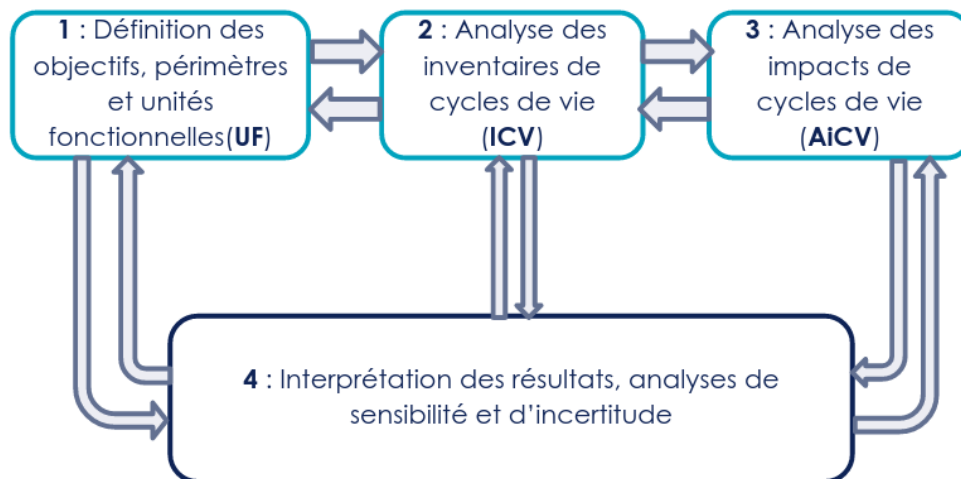


FIGURE 2.6 – Description des principales phases d'une ACV

En conclusion, l'hégémonie actuelle de l'ACV n'était peut-être pas inévitable, mais elle ne s'est pas faite par défaut. D'autres méthodes ont pu "concurrer" l'ACV comme outil. Par exemple, l'*Environmental Effect Analysis* (EEA) a pu être défendu comme meilleur que l'ACV sous certains aspects [141]. En dehors d'outils de design et de framework avec un haut niveau d'abstraction, l'ACV semble n'avoir laissé que peu de concurrents sur la quantification d'impacts environnementaux prenant en compte tout le cycle de vie. Expliquer son succès est complexe, mais il est possible que cette flexibilité, cette capacité d'adaptation et d'évolution ait joué un rôle crucial. L'ACV est finalement devenue, comme le décrit [142], un outil dominant, non plus seulement pour la recherche ou la création de rapport, mais aussi pour la création de politiques de soutenabilité environnementales.

19. L'article analyse 477 études simplifiant la méthode ACV majoritairement motivées par le manque de données.

2.2.2 Considérer les TIC puis le numérique

Depuis ces premiers travaux sur les produits d'emballages [132, 133], l'ACV s'est appliquée à d'autres secteurs. La Figure 2.7 montre à la fin des années 90 la répartition des sujets d'études ACV dans des groupes de pays européens. On peut déduire, de par la faible proportion dédiée aux "équipements électriques et électroniques" qui inclut entre autres les TIC, que ce secteur n'a pas connu un développement aussi rapide de la pratique de l'ACV que d'autres. Cette différence peut simplement s'expliquer par le poids économique plus important de certains secteurs, mais reste le constat, indépendamment des causes, que le secteur des TIC est peu représenté dans les travaux ACV au début des années 2000.

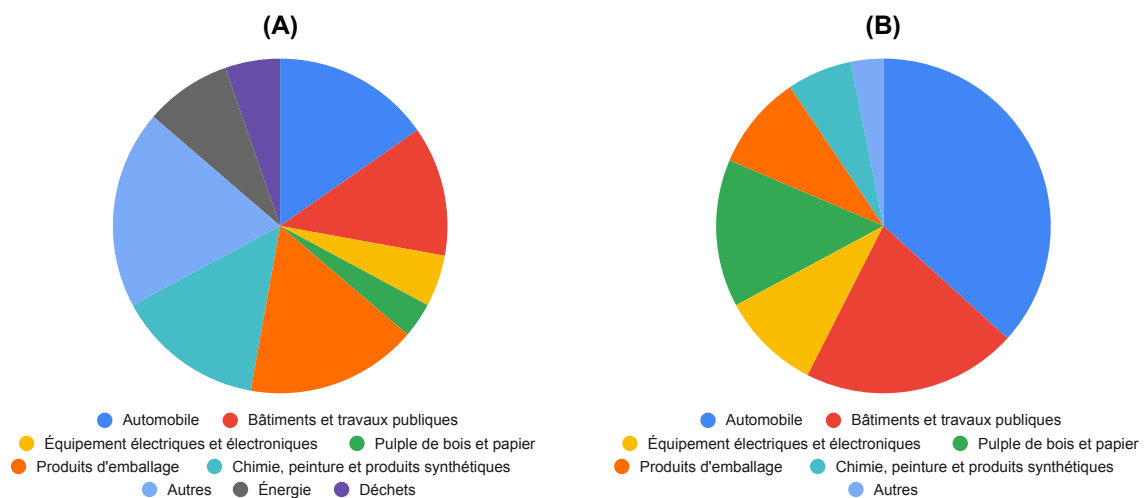


FIGURE 2.7 – Répartition des sujets d'ACV par secteur selon deux études européennes, catégories remaniées pour faciliter la comparaison - (A) Répartition par sujet de 621 ACVs, provenant de Suisse, d'Allemagne, d'Italie et de Suède, données issues de [143] en 2000 - (B) Répartition par sujet de 414 ACVs, provenant de Norvège, de Suède, de Finlande et du Danemark, données issues de [144] en 1999

La question du numérique et des TIC est abordée par la communauté proche de l'ACV, i.e. les chercheurs en évaluations environnementales et développement durable, sous l'angle de ses conséquences au niveau structurel. Dans un article paru en 2002, [145] avertit sur le potentiel effet rebond que le numérique, ou plutôt la numérisation de la société, pourrait produire. Il est déjà identifié que la numérisation peut permettre d'optimiser des processus existants mais que cette optimisation, malgré une réduction directe de l'impact environnemental, entraînerait au final une augmentation globale net de l'impact environnemental du fait de l'augmentation de la demande. Ainsi, le numérique est bien identifié comme faisant partie du phénomène de grande accélération des activités humaines et de leurs impacts environnementaux²⁰. En 2004 est publié [146], qui se préoccupe de l'augmentation du nombre d'équipements TIC mais aussi des infrastructures

20. voir 1, plus spécifiquement [1].

et services liés, mais surtout [147] fait figure de premier rapport important sur l'impact environnemental des TIC. Ces 2 travaux exposent : le coût environnemental des TIC, les économies qu'ils pourraient permettre mais aussi l'effet rebond qu'ils engendreraient. Les études concluent sur la possibilité pour les TIC de permettre des gains supérieurs à leurs coûts environnementaux. Toutefois, les deux études s'accordent sur le grand niveau d'incertitude liée à l'effet rebond et, déjà, le manque de données environnementales pour ce secteur. Elles se focalisent en grande partie sur la numérisation de l'industrie, i.e. l'intensification de l'usage des TIC dans les activités pré-existantes, et s'engagent dans le narratif autour de la "dématisation". Dans leur considération de l'effet rebond, il n'y a que le risque de l'intensification de la demande d'un processus. Les modèles utilisés ne pouvaient pas prendre en compte la potentielle création, via le numérique, de nouveaux besoins artificiels.

Détail important, ces études pointent une problématique qui persiste aujourd'hui : les frontières du secteur TIC ou du secteur numérique. Une chaîne de montage dans une usine automobile utilisant des robots finement connectés et coordonnés par un serveur fait-elle partie du secteur des TIC ou de l'automobile ? Doit-on parler de services TIC, services à base de TIC ou de services numériques²¹ ? La difficulté à délimiter les frontières de ce ou ces secteurs dure encore aujourd'hui du fait de la superposition partielle de ces catégories [148, 149], entraînant certains auteurs vers des termes légèrement redondant comme "Service numérique basé sur des TIC"²² [150]. Dans cette section, il sera plus question des TIC que du numérique. En effet, beaucoup des travaux issus de la communauté ACV ont commencé par l'étude des TIC, puis du numérique, qui sera abordé dans la section suivante plus spécifiquement.

On observe ainsi un intérêt de la communauté ACV partant des équipements électroniques, vers les équipements TIC, puis vers les TIC comme infrastructure, puis comme service. Ainsi, [151] décrit les premières études ACV dans le secteur des télécommunications, études ayant d'abord eu pour sujet les composants électroniques puis les terminaux utilisateurs et le réseau en lui-même. Cette hiérarchie "objet TIC" puis "réseaux de TIC" puis "service TIC", chaque niveau englobant le précédent, se retrouve par la suite dans les recommandations de l'ITU [25, 152] pour l'application de l'ACV aux produits TIC, réseaux de TIC et services TIC.

L'intérêt pour l'ACV des TIC vient souvent d'autres secteurs voisins, e.g. de l'électronique, aux téléphones, aux infrastructures numériques. Pareillement, l'industrie du papier de la pulpe de bois, présente tôt dans le monde ACV du fait du poids de cette industrie dans les pays nordiques, s'est intéressée aux conséquences de la digitalisation des supports de lecture. L'industrie du livre et plus largement du papier s'intéresse naturellement à l'évolution de l'usage de la lecture, du papier vers le numérique, du livre vers les premiers e-books [153]. De manière similaire, le secteur s'intéresse aussi aux conséquences de la vente en ligne de livres par rapport aux circuits de distribution classique [154]. Si ces premiers travaux sur l'impact des TIC s'inscrivent dans la dynamique autour des études sur les services TIC au début des années 2010, le sujet de la dématérialisation des supports de lecture restera un sujet actif [155, 156] étant un exemple typique d'étude des conséquences de numérisation d'un usage [157].

Il y a en effet, au début des années 2010, une dynamique certaine autour des ACV

21. *ICT service* ou *Digital service* en anglais.

22. "*digital Information and Communication Technology (ICT) solutions*"

de TIC mais surtout de réseaux et services TIC. Parallèlement à la publication des premières recommandations de l'ITU, en 2012 [25] puis 2014 [152], pour adapter l'ACV au bien, réseaux et services TIC²³, plusieurs études abordent des sujets plus au cœur de l'informatique. Selon [159], qui promeut l'ACV comme la meilleure méthodologie pour étudier les impacts environnementaux des TIC, ces normes sont une opportunité de gagner en maturité pour le domaine. Les résultats d'ACV étant très sensibles aux choix méthodologiques, la standardisation permet de réduire cette incertitude tout en renforçant l'expertise de la communauté.

Pour commencer, les travaux de Whitehead sur l'ACV de datacenters [89,160] sont une première pierre importante pour représenter des enjeux spécifiques du numérique au sein de la communauté ACV. Pareillement, les travaux issus des opérateurs télécoms [161,162] poursuivent les progrès méthodologiques sur la base de données d'opérateurs, même si les études restent parfois en mono-critère ou restreintes à la phase d'usage. Enfin, [163] propose non seulement l'ACV d'un ordinateur classique mais surtout pour la comparer à celle d'un ordinateur minimaliste connecté à un serveur distant comme un terminal. Cela introduit l'enjeu de comparer un objet personnel à un service. Les ACV de service TIC centré sur des usages spécifiquement numériques apparaissent, comme le mail [91]. À propos des services TIC, le travail d'Arushanyan [90,164] expose un panorama sur les ACV de biens et services TIC, résumé par les Figure 2.8a et 2.8b. Son travail présente toutefois une époque où l'ACV reste majoritairement centrée sur les équipements avec encore de nombreuses études quantifiant seulement l'impact de la phase d'usage en termes d'énergie ou d'empreinte carbone.

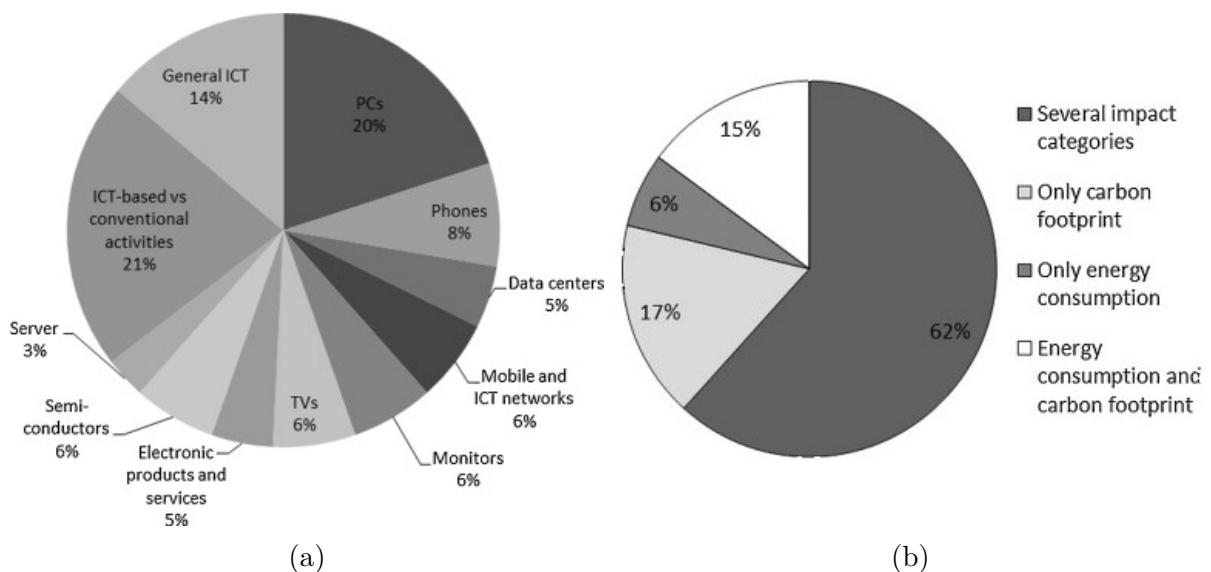


FIGURE 2.8 – Graphiques issu de l'étude de 65 articles d'ACV de biens ou service TIC par [164] en 2014

L'article [164] conclut sur les tendances générales, assez représentatives des travaux présentés dans cette section. Les phases du cycle de vie les plus impactantes sont la fabrication et l'usage, même s'il existe de nombreux cas où l'un dépasse l'autre et que

23. Recommandations adaptées aux normes européennes par l'ETSI [158].

cet équilibre est très sensible à la composition du mix électrique utilisé durant la phase d'usage. L'empreinte énergétique et carbone sont les plus étudiés, amenant une sous-représentation des autres catégories comme la raréfaction de ressources abiotiques, le volume de déchets ou la pollution de l'eau, ces deux dernières catégories étant plus liées à la phase de fin de vie, elles-mêmes sous-représentées. Les difficultés majeures sont le manque de données mais aussi le manque de standardisation des études associées souvent à leur caractère non-transparent.

En conclusion, le début des années 2010 est marqué par une dynamique certaine autour des impacts environnementaux du numérique, au-delà de la consommation d'électricité et des émissions de GES. Par exemple, c'est en 2012 que sort en France le livre [165], écrit par des membres du groupe EcoInfo, qui expose déjà à cette époque l'ensemble du problème environnemental que pose le numérique. Sans doute marqué par les promesses de la dématérialisation, l'accélération de la numérisation de la société, la seconde bulle internet avec le lancement des grands réseaux sociaux, l'arrivée sur le marché des smartphones, l'intérêt de la communauté ACV s'est significativement porté sur le numérique sous l'étiquette TIC. Toutefois, ces avancées ne semblent pas pénétrer directement la recherche en informatique ou l'industrie du numérique. Le cruel manque de données et les difficultés à engager les acteurs du secteur TIC dans la démarche ACV empêchent de créer directement une communauté ACV numérique. L'écart de plusieurs années, entre cette section et la suivante, représente sans doute le temps pour que les précieuses données ICV soient produites et que suffisamment d'acteurs soient sensibilisés et formés à ces enjeux. Dans ce travail, la progression comme nous l'avons vu de l'ACV de manière générale a joué. Par exemple en France, les lois REEN [166] et AGECE [167], en 2021 et 2020, participent grandement à générer, en partie par la contrainte juridique, un intérêt pour ces enjeux. Toutefois, ces avancées juridiques sont lentes et une certaine discontinuité entre les travaux peut être observée. Les travaux, issus de la communauté ACV, sur les TIC et service-TIC auront permis d'ouvrir la voie à l'ACV appliquée au secteur numérique et ses services.

2.2.3 Une difficile intégration des services numériques

Comme expliqué durant la section précédente, la différence entre d'une part le secteur TIC et les services TIC, et d'autre part le secteur numérique et les services numériques, ne fait pas l'objet de définitions admises. Particulièrement dans le contexte français, où le terme "service TIC" est presque totalement absent, la désignation de "service numérique" l'ayant presque totalement effacé [168]. Il est généralement admis que le numérique est un sous-ensemble inclus dans les TIC. De ce postulat, il faudrait ensuite définir ce qui fait qu'entre deux services TIC, l'un est un service numérique et l'autre non. On pourrait alors définir le service numérique par la centralité de l'utilisateur comme acteur. Une grande partie de la littérature, plus spécifiquement celle de la communauté *Systèmes d'information*, voit avant tout le service numérique comme la numérisation d'un service pré-existant [169–171]. Cette vision s'inscrit dans la grille de lecture de la "transition numérique"²⁴ où les services numériques ne sont que le résultat de cette supposée "déma-

24. Entre guillemets pour souligner les implicites contradictoires portés par cette désignation. Une transition est censée aller d'un état A vers un état B. Or jusqu'à présent, la transition numérique a certes changé des usages, mais surtout en a créé sans cesse de nouveaux, avec une intensité croissante, laissant

térialisation" [172]. Ce manuscrit se placera dans une définition inspirée de [173] et [171]. L'article [173] définit simplement le service numérique comme :

« des services, qui sont obtenus et/ou arrangés par le biais d'une transaction numérique (informations, modules logiciels ou biens de consommation) sur le protocole Internet (IP)»

L'article continue ensuite en comparant avant tout les services numériques et non-numériques. Les caractéristiques retenues sont que la transaction numérique est au cœur de la raison d'être du service, qu'il y a une "relation différente à l'intangibilité" et l'utilisation d'un ensemble d'équipements matériels. Ces deux derniers points soulignent le paradoxe du concept de "dématérialisation". La dématérialisation que propose un service numérique rend le service intangible, surtout pour l'utilisateur, voir le fournisseur du service. Ce rapport complexe entre interactions matérielles et dématérialisées est présent dans [171] qui souhaite proposer une définition des "plateformes numériques" comme Uber ou Airbnb. Il définit ce qu'il nomme des écosystèmes numériques comme l'association d'écosystèmes de biens matériels avec un service numérique. Le service numérique ayant au cœur de son activité la nécessité d'utiliser une plateforme numérique.

Partant de ces définitions, la définition de service numérique sera ici : un service TIC, coordonnant dans une plateforme un ensemble d'équipements et d'infrastructures numériques dans le but de satisfaire un besoin, centré sur le traitement, le stockage ou la communication de données numériques. Les services numériques que cite [173] dans sa Table 1 sont des bons exemples de services numériques²⁵. Ce ne sont pas des services pré-existants et améliorés par l'usage de TIC mais des services centrés sur le numérique, utilisant des TIC.

Ce cadrage a son importance pour comprendre la transition non triviale qu'il y a des services TIC aux services numériques. Les entreprises proposant ces services numériques ne fabriquent souvent pas de biens matériels spécifiques, elles se reposent sur des infrastructures existantes, elles se contentent de créer, louer et vendre des services numériques. Dans ce contexte, les acteurs des services numériques sont souvent dans une position où ils ne peuvent pas produire eux-mêmes leurs propres données comme le faisait dans la section précédente les grandes compagnies télécoms. En plus de ne pas produire de données, ils ne développent pas non plus de compétences ACV en leur sein. La position de client de ces fournisseurs de services numériques, sans négliger leur vision de la "dématérialisation", crée une distance supplémentaire entre ces acteurs et les enjeux environnementaux de leur activité.

Lorsque les enjeux environnementaux finissent par émerger dans ces secteurs, cela pour de multiples raisons juridiques, politiques, administratives ou encore des initiatives personnelles, les compétences d'ACV adaptées aux services numériques sont dispersées. On se repose alors sur une communauté qui est encore plus plurielle que pour les travaux des sections précédentes. Il y a un mélange d'acteurs : recherche académique, cabinets de conseil de taille allant de la PME à des multinationales, industriels, ESN et agences publiques.

Pour cette communauté hétéroclite, le premier défi reste les données, avec là aussi

sans réponse la nature de l'état B d'arrivée supposément visé.

25. **e.g.** : Youtube, Wikipedia, Facebook, Salesforces mais aussi Myspace, Secondlife et Craigslist, car cette liste date bien de 2007.

plusieurs approches disparates. Il y a une tension entre des données fermées, à l'accès payant, comme CEDaCI [174] et NegaOctet [175], et la construction de bases ouvertes comme Boavizta [101]. L'hétérogénéité de la communauté, pénalise sans doute l'émergence d'une base de données faisant consensus. Un problème qui sera presque systématiquement soulevé dans les études suivantes.

Les études ACV sur les services numériques sont rares, et les rapports complets disponibles encore plus, et ils seront analysés en détails au prochain chapitre. Mentionnons donc ici les travaux publiés comme cet ensemble d'études comparant des services numériques culturels et leur équivalent non numérique [157] ou encore cet article [176] sur un ensemble de services numériques communs. Ce dernier article, comme [177], même s'il utilise le terme "service TIC", souligne pour les services numériques l'importance du comportement utilisateur et des hypothèses faites à son sujet. Autre production utilisant la méthode ACV, [178] estime l'empreinte énergétique de 10 usages numériques. Elle souligne aussi l'hétérogénéité de la littérature, comme illustré dans Figure 2.9, ainsi que ses manquements en termes de transparence et exhaustivité des évaluations.

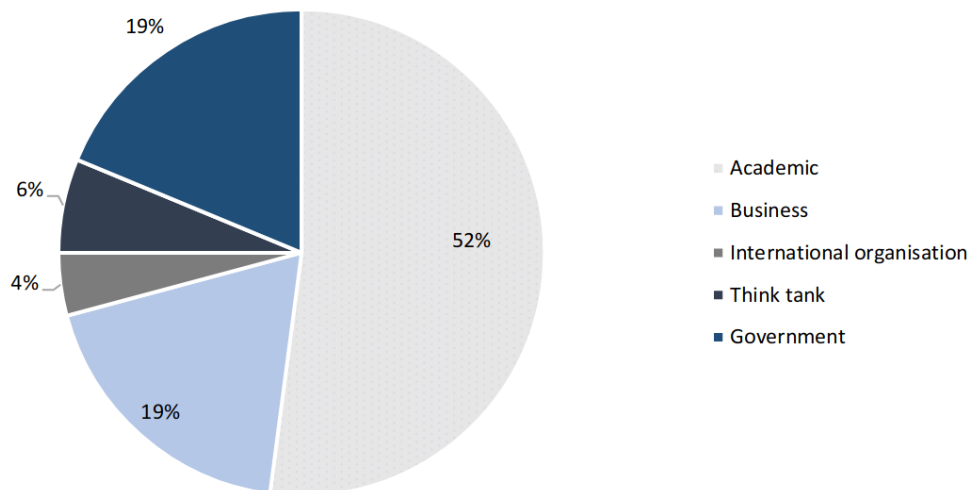


FIGURE 2.9 – Sources des 84 documents consultés dans la revue de littérature de [178] concernant l'empreinte énergétique de services numériques

Pour finir, il est important de mentionner, comme noté en section 2.1.3, qu'il existe une dynamique dans la communauté informatique pour réduire les impacts du numérique qui est en attente de résultats issus de l'ACV pour définir leurs objectifs et méthodes. Cela se traduit par la création par des institutions de normes d'éco-conception à destination des acteurs du numérique. Premier exemple, le Référentiel général pour l'IA frugale de l'Afnor [179], qui intègre des travaux ACV dans sa construction. Deuxièmement, le Référentiel Général de l'Écoconception des Services Numériques (RGESN) de l'ARCEP [180], souhaite aussi intégrer des connaissances et méthodologies issues de l'ACV de services numériques, même si cela reste pour le moment limité et parfois contradictoire [181, 182].

En conclusion, il y a bien d'une part dans la communauté informatique une volonté et un besoin d'introduire des évaluations environnementales plus complètes des technologies informatiques afin de répondre aux enjeux environnementaux, notamment grâce à l'ACV. D'autre part, dans le domaine de l'évaluation environnementale, l'ACV s'est

imposée comme un outil de référence qui a su évoluer et s'adapter à différents contextes comme celui des TIC. Néanmoins, il y a actuellement un blocage, un manque à combler, entre une communauté informatique qui cherche à utiliser l'ACV pour concevoir des services numériques plus responsables et une communauté ACV qui cherche à s'adapter aux services numériques.

Analyse de cycle de vie de services numériques : potentiel et limites

*Tous les modèles sont faux
mais certains sont utiles.*

Georges Box

Parmi les différents outils et méthodes existantes pour étudier l'impact environnemental de l'usage du numérique, le choix a été fait dans cette thèse d'utiliser la méthode ACV. Cette méthode semble être, comme nous l'avons vu au chapitre précédent, d'une part l'outil le plus propice à la prise en compte globale des impacts environnementaux du numérique, d'autre part un outil déjà largement adopté dans le domaine des sciences de la soutenabilité, déjà utilisé pour évaluer d'autres secteurs d'activités, cherchant à intégrer les services numériques à ses applications. Il a été constaté une attente pour l'application de l'ACV au numérique, mais plus spécifiquement, afin d'évaluer l'impact des usages, à l'ACV de services numériques. Le présent chapitre est consacré à l'analyse critique de l'état de l'art de l'ACV de services numériques. Dans un premier temps, il sera question de l'intérêt et du potentiel de l'ACV de services numériques. Dans un second temps, les rares rapports complets d'ACV de services numériques seront analysés et leurs limites étudiées. Enfin, une grille de lecture de ces rapports sera proposée, ainsi que les conclusions tirées de cet état de l'art vis-à-vis des contributions de cette thèse.

3.1 Potentiel de l'ACV de services numériques

3.1.1 Évaluation environnementale complète

L'une des premières raisons à même d'expliquer le choix de l'ACV de services numériques comme méthode pour étudier l'impact environnemental du numérique est l'exhaustivité de l'approche. Si l'exhaustivité n'est pas toujours absolument nécessaire, notamment pour répondre à des objectifs restreints, elle reste indispensable dès qu'il est question de dépasser un problème restreint. C'est une sécurité, à même de garantir une approche nécessairement rigoureuse face à la complexité des enjeux environnementaux.

En effet, en comparaison des approches décrites précédemment, voir le chapitre 2, l'ACV de services numériques permet de gagner en profondeur sur la question de l'empreinte environnementale du numérique selon plusieurs axes :

- Sur le cycle de vie : en allant au-delà de la phase d'usage, en incluant la phase de fabrication des équipements et infrastructures, mais aussi la phase de fin de vie [99] et la phase de transport, qui peuvent être critiques [183]
- Sur le nombre d'impacts environnementaux évalués : même en prenant une définition large du concept de catégorie d'impacts environnementaux étudiés, beaucoup d'études se limitent à évaluer la consommation en électricité. Cette consommation est par simplification souvent présentée comme l'empreinte énergétique, alors que celle-ci sous la forme de demande cumulée en énergie primaire [184] couvre un périmètre plus large. Les autres catégories d'impacts renvoient à des modèles d'enjeux environnementaux comme le réchauffement climatique et l'empreinte carbone [185] ou encore la raréfaction des métaux et minéraux et l'empreinte en équivalent antimoine [186]. D'autres catégories moins connues comme les rayonnements ionisants [187] peuvent être significatives pour le numérique [18], d'autant plus dans le cadre français. Le standard PEF reconnaît plus de 10 catégories d'impacts environnementaux [188].
- Sur l'axe service : la majorité des usages du numérique sont en réalité des services. Il est de nos jours presque impossible d'utiliser un terminal informatique sans faire usage d'infrastructures réseaux et datacenters. Il est donc pertinent de ne pas concentrer l'étude de l'impact environnementale du numérique uniquement sur une partie du service. L'ADEME a d'ailleurs fait le choix de prendre comme socle de différents usages du numérique (accès réseaux, cloud, IA, site internet,...) l'échelle service [189]¹, selon l'architecture en trois parties, tel que décrite par la figure 3.1.

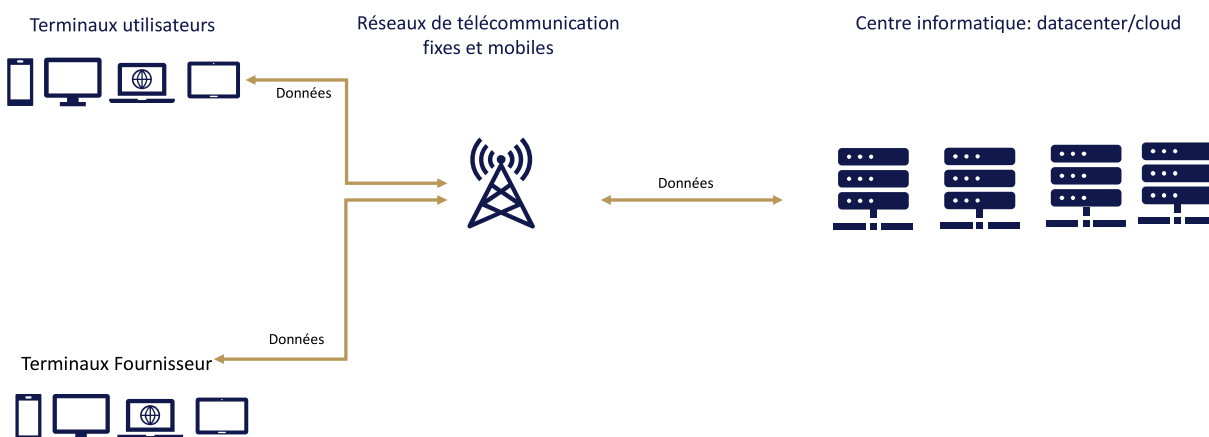


FIGURE 3.1 – Exemple schématique de l'architecture service en trois parties, terminaux, réseaux, datacenters, issu de [189]

1. Plus précisément, le référentiel de catégorie de produit (RCP) de l'ADEME pour les services numériques sert de RCP "mère" pour les RCP "filles" comme l'accès réseaux, l'IA, la blockchain, l'usage cloud ou les sites web,...

Ces trois gains de profondeur sont caractéristiques de l'ACV de services numériques. Ils présentent plusieurs intérêts.

Tout d'abord, la contrainte juridique progressive sur les évaluations d'impacts environnementaux des produits, services et organisations, comme la PEF [137] et l'OEF [138], est une incitation croissante. La conformité à ces standards passe de plus en plus par une évaluation complète des activités, y compris à l'échelle service.

L'intérêt principal d'une vision élargie des impacts environnementaux du numérique est principalement d'éviter la non-prise en compte d'impacts inhérents à l'activité étudiée. Ce "risque" de ne pas prendre en compte une partie des impacts est d'autant plus fort en informatique où le discours dominant sur le narratif de "dématérialisation" incite presque à rester sur une vision restreinte du problème, voire une absence de problème. Par exemple, une vision restreinte à l'empreinte carbone de la phase d'usage du numérique peut donner en France l'illusion d'un secteur numérique décarboné, due au mix électrique français lui-même fortement décarboné.

Derrière le risque d'invisibiliser ainsi certains impacts, se trouve aussi le risque d'un transfert d'impact. Un transfert d'impact se comprend comme le changement de nature d'une partie de l'empreinte environnementale, classiquement de manière à, volontairement ou non, exclure cette partie de l'empreinte du périmètre de l'évaluation. Le périmètre d'une ACV numérique multicritère, i.e. prenant en compte un large panel de catégorie d'impacts, permet d'éviter de nombreuses situations de transferts d'impacts. Des transferts d'impacts dans le cadre des services numériques pourraient prendre les formes suivantes, e.g. selon les trois axes précédemment exposés :

- **Transferts entre phases du cycle de vie** : le numérique a tendanciellement vu son empreinte carbone être transférée de sa phase d'usage à sa phase de fabrication [84]. L'efficacité énergétique et le recours croissant à des appareils alimentés par batterie ont baissé la demande en électricité, mais alourdit l'empreinte de la phase de fabrication ainsi que celle de la phase de fin de vie [99].
- **Transferts entre catégories d'impact** : le *carbon-aware computing*, comme décrit à la section 2.1.2, peut être vu comme un transfert d'impact. Exécuter un calcul dans un datacenter situé dans une région au mix électrique décarboné peut amener à le déplacer dans une zone de stress hydrique, sensible à une consommation d'eau supplémentaire du datacenter.
- **Transferts entre parties du service** : plusieurs services essayent de déplacer "l'effort" d'une partie du service à une autre. Premier exemple, les mécanismes de cache web stockent des données, au plus près de l'utilisateur, les répliquant en local afin de diminuer le nombre de transferts de données du serveur aux clients finaux, diminuant ainsi théoriquement la consommation d'énergie [190]. Ce mécanisme, théoriquement, permet donc de diminuer une charge de travail côté datacenters et réseaux vers les terminaux utilisateurs². Il s'agit d'un compromis entre l'utilisation d'infrastructures et la présence³ d'espace de stockage proche du client. Plusieurs études mettent ainsi en avant des économies d'énergie, mais de rares études comme [191] exposent clairement l'existence d'un compromis entre consommation électrique du stockage et consommation électrique du réseau. Autre exemple, l'apprentissage fédéré peut être

2. Ou pour être précis des sous-parties du réseau plus proche ou des nœuds edge.

3. mais aussi l'utilisation certes moins coûteuse.

utilisé pour entraîner un modèle d'IA sur des terminaux de manière décentralisée plutôt que sur un serveur centralisé, espérant ainsi une économie d'énergie [192]. Ces résultats sont cependant à nuancer, premièrement dû au fait que la plupart des études ne prennent en compte que la phase d'usage, mais même sans cela l'utilisation de terminaux plutôt que de serveurs spécialisés pénalise l'apprentissage fédéré. Un article de Google [118] revendique un coût énergétique 12 fois supérieur de l'entraînement sur smartphone par rapport à la même tâche réalisée sur un serveur.

Ce dernier point, concernant l'axe service, illustre la pertinence de l'ACV de services numériques. De nombreuses comparaisons entre différents équipements, terminaux ou serveurs de différentes tailles, et infrastructures, réseaux ou systèmes distribués, se font uniquement sur le critère de consommation électrique. Or, les empreintes de fabrication de ces différentes architectures sont significativement différentes les unes des autres et peuvent être décorréliées des consommations d'énergie associées. La comparaison entre des services numériques reposant sur une architecture centralisée ou distribuée est d'autant plus difficile qu'elle repose sur de nombreuses hypothèses concernant notamment le taux d'utilisation des équipements⁴. C'est un exercice d'autant plus difficile que l'empreinte environnementale n'est souvent qu'une considération secondaire dans le choix de ces architectures et donc peu étudiée. Pour reprendre les exemples cités plus haut, l'enjeu premier du *cache* reste la performance, tandis que pour l'apprentissage fédéré, l'enjeu fort est sans doute la protection des données personnelles.

En conclusion, comme elle l'a permis pour d'autres secteurs d'activité, l'ACV permet d'atteindre un niveau de compréhension plus haut des enjeux environnementaux. L'approche multicritère, en cycle de vie complet sur tout le service, permet une exhaustivité, un gain de qualité certain dans la compréhension des compromis derrière des choix techniques critiques. La capacité à évaluer de manière plus juste les impacts environnementaux est cruciale afin d'effectuer de meilleurs choix d'éco-conception des services numériques. Des progrès sont attendus de la part de l'ACV de services numériques afin d'améliorer les outils existants [193], le risque que les outils d'évaluation simplifiés donnent des incitations contre-productives existe. Des ACV ont déjà permis d'établir des résultats contre-intuitifs, par exemple [194] qui a comparé l'empreinte d'un système de capteurs connectés alimentés par batterie ou par un dispositif sans fil, conclut à une empreinte réduite pour le système alimenté par batteries⁵. Déjà en 2014, l'article [155], comparant la lecture sur support papier et la consommation de médias par services numériques, constatait la nécessité d'ACV complète à une échelle service comme définit ici. Il observait dans sa revue de littérature, la tendance des services numériques à déplacer leur empreinte de la phase d'usage vers les phases de fabrication et de fin de vie, mais aussi à augmenter leurs recours aux infrastructures réseaux et datacenters, dont les empreintes restent encore plus difficiles à étudier. Cette nécessaire meilleure connaissance de l'empreinte des services numériques, quantitativement et selon les trois axes, permet donc de mieux contrôler et *in fine* réduire l'empreinte des services et des organisations les utilisant, comme il en sera question à la

4. Que ce soit côté datacenters où les données sur le taux d'utilisation active des serveurs sont rares, de même pour une architecture distribuée, certains scénarios considéreront que les équipements sont pré-existant au service, i.e. ne sont pas construits spécifiquement pour cet usage, et donc ont un coût de fabrication "négligeable".

5. "As these findings may seem counterintuitive because primary batteries are usually associated with a high environmental impact, this work calls for the systematic use of LCA in the context of IoT." [194]

section 3.1.2 ci-dessous.

3.1.2 Éco-conception de services numériques

Une des principales motivations de l'ACV de services numériques est son utilisation dans le cadre de l'éco-conception des services numériques. L'éco-conception se définit comme l'intégration des aspects environnementaux dès la conception et dans le développement des biens et services dans le but de réduire les impacts environnementaux tout au long du cycle de vie [195]. Cette pratique entraîne logiquement une demande d'études ACV de produits et services, que ce soit, pendant la phase de conception du service, *ex ante*, comme durant la phase d'usage du service. Le référentiel d'éco-conception des services numériques (RGESN) de l'ARCEP et de l'ARCOM [180] mentionne et recommande l'usage de l'ACV de services numériques. Plus spécifiquement, dans le cadre de la limitation ou réduction de l'empreinte du service, ou de celle de l'hébergeur du service, l'ACV de services numériques sert à établir des indicateurs environnementaux fiables qui doivent servir ensuite à établir des objectifs chiffrés.

Le rôle de tels indicateurs est d'aider les développeurs et gestionnaires de projet à : évaluer l'empreinte du projet à l'avance, effectuer des choix techniques, mettre en place une amélioration continue de l'aspect environnemental du projet en phase avec l'évolution de celui-ci durant le développement puis la mise en production. Actuellement, une grande partie des outils proposant des indicateurs environnementaux à destinations des développeurs sont limités, comme le montre [193] spécifiquement pour l'un d'entre eux [196].

Les outils d'aide à l'éco-conception les plus courants [197–201] fonctionnent majoritairement par analyse du code d'une page web ou d'un projet logiciel complet. Ces logiciels s'apparentent alors à la famille des outils automatiques d'analyse statique de programmes [202]. S'appuyant sur des éléments décontextualisés, i.e. principalement le code affiché côté client, ces outils, qui bien qu'ayant un périmètre différent de l'ACV de services numériques poursuivent les mêmes buts, peuvent donner des résultats décorrélés de l'empreinte des systèmes étudiés. Actuellement, les indicateurs produits par ces outils ne répondent pas aux exigences du RGESN en ce qui concerne le contrôle et le suivi de l'empreinte d'un service numérique, notamment en termes de périmètre et de transparence des méthodologies et des données primaires. L'ACV de services numériques permettrait d'améliorer les modèles sur lesquels ces services reposent et pourrait servir d'exemple à une évaluation automatique de bout en bout d'un service numérique. Durant cette thèse, une méthodologie d'ACV de services numériques accélérées a été utilisée pour évaluer *ex post* un ensemble de services numériques livré à des clients par l'entreprise hébergeant la thèse. Les résultats de cette ACV de services numériques ont pu être comparés pour les services accessibles publiquement avec les résultats donnés par cet ensemble d'outils [197–201] afin de comparer les outils entre eux mais aussi d'observer d'éventuelles présences, ou absences franches, de corrélation entre résultats ACV d'une part et indicateurs d'outils automatiques d'autre part. Les premiers résultats et éléments de méthode sont présents en Annexe 7.3 du présent manuscrit.

3.1.3 Une échelle service activable pour les organisations

Une des applications potentielles de l'ACV de services numériques se trouve non pas à l'échelle du service même, mais à l'échelle de l'organisation⁶ ou de son système d'information (SI). En effet, une manière d'étudier l'empreinte du numérique au sein d'une organisation pourrait être au travers des usages du numérique dans l'organisation. Cet usage du numérique dans l'organisation peut être présenté comme l'ensemble des services numériques utilisés, ou l'usage de l'ensemble des services numériques hébergés ou permis par le SI.

Une telle approche offrirait une solution originale aux limites des actuelles approches d'évaluation et de réduction de l'empreinte du numérique au sein d'une organisation. En effet, le "Guide de bonnes pratiques numérique responsable pour les organisations" [204] met en avant, dans son volet "Mesure et évaluation", des outils d'évaluation d'empreinte du SI comme EcoDiag [205] ou WeNR [206]. Ces outils proposent de construire l'empreinte numérique de l'organisation, comme étant l'empreinte du SI, par inventaire de l'ensemble des équipements numériques et l'estimation de leur consommation d'électricité. Les méthodologies de ces outils se basent sur le périmètre des bilans carbone d'entreprises [207], à savoir les 3 *Scopes* d'émissions de GES. Une des premières limites de ces méthodologies, dans ce cadre de lecture, est la très difficile prise en compte du *Scope 3* dédié aux émissions indirectes, notamment l'utilisation d'infrastructures de télécommunication ou le recours à l'hébergement en datacenter. Ces méthodologies ne prennent pas en compte ces coûts environnementaux, ou alors très partiellement [208]⁷.

Premièrement, ces méthodologies portent le risque, surtout lorsqu'elles sont automatisées [208] dans un but d'amélioration continue, de créer des incitations délétères, par transfert d'impact entre les niveaux des services utilisés, comme expliqué dans la section 3.1.1. Ce risque est d'autant plus présent qu'il pourrait trouver appui dans un discours préexistant qui néglige l'empreinte environnementale de ces choix techniques et donc encourage l'utilisation des infrastructures de communication [13] et de calculs [209]⁸.

Deuxièmement, même si ces méthodologies parvenaient à inclure l'empreinte environnementale de l'usage des réseaux et datacenters par les organisations, certaines limites resteraient présentes, dues au choix de la méthode par inventaire d'équipement et flux. Dans ce cadre, les politiques de l'organisation pour réduire l'empreinte du SI se retrouvent vite limitées à :

1. Agir sur sa politique d'achat : Acheter moins souvent des équipements, i.e. rallonger la durée de vie des équipements, acheter mieux, i.e. acheter des produits ayant

6. Organisation au sens large [203], incluant : entreprise privée, administration publique, association, syndicat, etc.

7. Plus précisément, EcoDiag considère ces parties hors périmètre due à la complexité d'évaluation, la méthodologie de WeNR considère la consommation électrique des équipements de datacenter, mais laisse le flou sur la prise en compte, si c'est le cas, de l'empreinte de fabrication et ne considère pas le coût des infrastructures réseaux, seulement des équipements réseaux de l'organisation.

8. Les références [13] et [209] sont caractéristiques du discours incitant à l'utilisation des infrastructures. L'utilisation préférentielle de grandes infrastructures mutualisées n'est pas condamnable en soit, elle peut même sans doute être bénéfique, le risque réside premièrement dans l'incertitude sur le coût environnemental de tels choix, incertitude d'autant plus grande compte tenue du manque de transparence de certains acteurs dominants, deuxièmement dans une forme d'effet rebond, les grandes infrastructures permettant plus facilement, et surtout presque instantanément, un usage plus intense des ressources. Pour une analyse plus nuancer sur le cas des réseaux de télécommunications [210].

comparativement moins d'impacts environnementaux pour des qualités équivalentes, changer des fournisseurs de biens et services selon leur empreinte affichée.

2. Chercher l'efficacité : Eco-concevoir les services, diminuer les pertes, augmenter l'efficacité des services numériques, agir sur le PUE des équipements.

La limite ici est que la nature hautement mutualisée des équipements numériques rend très difficile la décroissance du nombre d'équipements et des ressources utilisées. En effet, diminuer les ressources informatiques, moins de serveurs, moins d'électricité, risque d'impacter de manière très incertaine les usages du SI. Or ce sont bien les usages qui sont la raison d'être du SI. L'ACV de l'ensemble des services numériques du SI permet, au-delà de sa précision, d'obtenir un bilan environnemental du SI sur la base des usages et non sur la base du stock d'équipements. Un tel bilan par l'ACV des services numériques permettrait de réellement questionner chaque usage du numérique et rendrait applicable une politique de réduction forte de l'empreinte du SI s'appuyant sur une décroissance du SI. Une décroissance du SI qui serait sans un tel bilan très hasardeuse, surtout à grande échelle, tant un seul serveur peut assurer des usages très hétérogènes compte tenu des nombreux services qu'il héberge. Ce bilan permettrait d'établir la correspondance, via l'ACV des services numériques, entre les besoins et les ressources déployées, facilitant un questionnement des besoins en question, pouvant amener à une véritable politique de sobriété numérique [128]. Cette approche holistique de l'empreinte du numérique est nécessaire. Baser la réduction des impacts environnementaux uniquement sur les politiques d'achats de matériel peut se révéler insuffisant [211] sans politique globale incluant le changement des usages du numérique.

3.1.4 Interopérabilité avec les autres secteurs

Une dernière application critique de l'ACV de services numériques est son usage dans le cadre d'études comparatives entre des solutions basées sur des services numériques et des solutions sans usage de numérique. Cette situation tire parti du fait que l'ACV est utilisée pour évaluer l'empreinte de services non numériques, l'ACV de services numériques permet une interopérabilité avec les autres évaluations environnementales en dehors du secteur numérique. Ce potentiel est rapidement identifié par les acteurs ancrés dans les méthodes d'évaluation environnementale. Par exemple, [32] et [125] citent l'ACV comme méthode caractéristique et nécessaire à la pratique de la soutenabilité dans les TIC, ils relèvent aussi que l'ACV est la méthode prédominante pour l'évaluation nette des impacts d'un service précis⁹. Les résultats d'ACV de services numériques sont en effet les données primaires qu'utilisent des études comme [123] pour étudier le bilan environnemental net de services numériques. Si d'ailleurs, les articles précédemment cités parlent de TIC ou de solutions à base de TIC cela correspond à la définition admise ici de services numériques. Possiblement dues aux difficultés inhérentes à un nouvel exercice, certains acteurs comme [155] pointent le manque de données et la complexité des services TIC comme un grand facteur d'incertitude pouvant minorer l'intérêt de telles études comparatives.

Toutefois, des progrès ont pu être faits, tant dans la stabilisation de la méthodologie

9. L'on entend par bilan ou évaluation nette des impacts environnementaux d'un service numérique, les impacts environnementaux directs du service auxquels on soustrait les impacts environnementaux "évités" grâce au service, pour plus détails [33].

que la robustesse des données permettant des études comme [157] qui comparent l’empreinte de services culturels avec leur version numérique¹⁰. L’étude reste cependant prudente, l’ACV de services numériques ne permet pas de comparer unilatéralement un service numérisé et son équivalent classique, elle permet de comparer des scénarios d’usages entre eux. Ces scénarios d’usages dépendent évidemment de la technologie qu’ils utilisent mais aussi d’hypothèses comme le comportement des utilisateurs par exemple. Ce dernier point est souligné par [14] qui met en avant l’utilisation nécessaire de l’ACV mais aussi ses limites propres, comme les données d’Inventaire de Cycle de Vie (ICV) encore une fois, mais aussi l’attention qui doit être portée aux scénarios qui restent des déterminants dominants dans ces analyses. D’autant plus que, pour reprendre la typologie d’impact de [32, 123] ou de modèles plus avancés comme [212], l’ACV de services numériques reste pour le moment un moyen d’évaluation uniquement des impacts environnementaux directs. L’ensemble des effets de second ordre ou même les effets structurels ne sont pas inclus dans l’ACV de services numériques. C’est pourquoi, il pourrait être défendu que l’ACV de services numériques est déjà à un niveau satisfaisant en comparaison des impacts non-couverts par cette méthodologie. Impacts qui devraient donc être couverts en priorité afin d’améliorer les évaluations comparatives ou les bilans nets des impacts environnementaux d’un service numérique. Cependant, comme rappelé dans l’état de l’art de [33], de nombreuses études ne prennent toujours pas en compte l’ensemble des effets directs comme devrait l’être une ACV de services numériques. Il y a encore une grande marge de progression à faire sur la qualité et l’exhaustivité des résultats d’ACV de services numériques, quand bien même, ils sont limités aux seuls effets directs. Cette qualité d’évaluation des impacts de premier ordre des services numériques est d’autant plus critique que les manquements et incertitudes dans ce premier niveau d’évaluations peuvent se propager dans les résultats des impacts de second et troisième ordre. Ce constat souligne donc l’importance et la nécessité de perfectionnement des ACVs de services numériques.

3.2 Les limites actuelles de l’ACV de services numériques

Après avoir exposé les qualités et potentiels de l’ACV de services numériques, un état de l’art plus détaillé des ACV accessibles est requis. Comme expliqué dans la section 1.5, l’état de l’art de services numériques s’est révélé difficile du fait de l’activité scientifique sur le sujet mais aussi de l’ambiguïté entre présence d’ACV de services numériques dans la littérature et présence de méthodologies et rapports ouverts dans la littérature. Cette présente section sera donc souvent dans la délicate situation¹¹ d’essayer de montrer l’absence de certaines qualités ou enjeux dans une littérature peu transparente. Les limites décrites ne sont donc pas absolument présentes chez tous les pratiquants de l’ACV de services numériques, mais correspondent aux manquements que l’on peut observer dans la littérature si l’on cherche à étudier l’empreinte du numérique par les ACVs de services numériques.

10. Par exemple, livre contre e-book, jeux-vidéo en tant que produit ou en tant que service, musique format CD ou en streaming, etc.

11. L’absence de preuve n’étant pas fondamentalement la preuve de l’absence.

3.2.1 Manque de maturité globale du champ scientifique

Il a été détaillé à la section 2.2.1 l'historique et les critères définissant l'ACV, montrant le caractère souple de sa définition et soulignant une pratique encore en évolution. Les caractéristiques peuvent être brièvement se résumer à :

1. Étudier le cycle de vie
2. Inclure un large panel de catégories d'impacts
3. Estimer des résultats quantitatifs
4. Adopter une approche scientifique

Ainsi que, dans le cadre de notre objectif d'étudier l'état de l'art des ACVs de *services* numériques, la caractéristique suivante :

5. Considérer l'ensemble du service numérique nécessaire à l'usage et non uniquement une sous-partie

Un premier regard rapide sur les rares, et légèrement datées, revues de plusieurs ACV de services numériques ou assimilables donne à voir peu de cas ayant ces quatre caractéristiques¹². Par exemple, dans [123] en 2016, les services numériques sont dans leur écrasante majorité étudiés sous l'angle de leur empreinte énergétique. Les périmètres des services numériques étudiés sont flous et ne comportent pas toujours l'ensemble du service. La majorité des services numériques sont étudiés sous l'angle du volume de données transportées. Ce volume est ensuite utilisé dans une approche descendante comme clé d'allocation pour les éventuels parties datacenters et équipements. La revue [164] de 2014 inclut peu de services numériques. Une des rares véritables ACV d'un service numérique compilé dans cette revue étudie le mail [91]. Cependant, le modèle utilisé pour évaluer l'empreinte de l'infrastructure réseaux n'est pas disponible et l'allocation des équipements datacenters se fait sur la base du volume de données stockées et la durée de stockage. Cette dernière clé d'allocation, en plus d'être inadaptée à la prise en compte de la phase d'usage de la partie datacenters et insensible à la fréquence d'accès de la ressource stockée, ne spécifie pas dans quelles proportions ces deux variables forment une clé d'allocation. D'autres revues, plus récentes, sont par définition en mono-critère sans forcément inclure tout le cycle de vie comme [178]. Cette dernière revue récente¹³, tout comme [213]¹⁴, constatent dans le cadre des services numériques un manque de littérature scientifique, plus particulièrement concernant les caractéristiques de leur usage, de nombreuses études ayant une approche descendante qui rend très incertaine l'empreinte d'actions individuelles. Pour finir, même dans l'ACV traitant de l'usage de services numériques de manière détaillée comme [157], une grande partie de la méthode est en "boîte noire" pour des raisons de droits sur les données ICV. Enfin, même avec un accès aux données ICV, certaines ACV comme la précédente repose sur des données ICV de bases de données comme Nega-Octet [175] et ont recours à des résultats d'ACV de services numériques¹⁵ issus de cette base de données. Malheureusement, les résultats proposés pour ces services basiques étant aussi en "boîte noire", reposant sur des scénarios définis très spécifiques, ils sont difficilement exploitables.

12. Et cela, sans vraiment chercher à évaluer la présence de la quatrième caractéristique, qui sauf cas exceptionnel pourrait être plus subjective.

13. 2023.

14. 2024.

15. Envoyer un mail, transférer un fichier, stocker des données dans le cloud, etc.

Cet enjeu des données a déjà été répété de nombreuses fois dans ce manuscrit ¹⁶ mais il est caractéristique d'un constat de manque de maturité de ce champ scientifique sous plusieurs aspects. L'accès aux données ICV par exemple. La base ICV NégaOctet [175] revendique d'être la seule source conforme aux exigences des normes ISO et au standard PEF. Malheureusement, malgré les limites évoquées précédemment, il est regrettable de constater que la base n'aura été maintenue que du 1^{er} février 2022 au 30 juin 2024, soit 29 mois. Les autres bases ICV pouvant être utilisées, offrent pour les TIC des résultats pouvant être inconsistants [214]. Cette difficulté d'accès à un des outils les plus basiques de l'ACV de services numériques, est à la fois cause et conséquence d'un manque de maturité du champ et de la difficile constitution d'une communauté de pratiquants de l'ACV de services numériques. L'existence, ou dans ce cas l'absence d'existence, de communauté de pratiquants est tout sauf anecdotique. La revue par les pairs est un processus aussi crucial dans le monde académique que dans le processus ACV professionnel. Les faits suivants sont à la fois causes et conséquences. Il y a peu de publications, encore moins en données et méthodologies de calculs ouvertes. Il est difficile de trouver des pairs pour effectuer des revues et gagner en qualité. Les études se limitent souvent à du mono-critère. Beaucoup d'études sont sur des études de cas théoriques, sans données de terrain, avec des méthodes de calculs en approche descendante.

Un dernier exemple de l'immaturité de l'ACV de services numériques en tant que champ scientifique : la difficulté de constituer un état de l'art de qualité pour un service numérique précis. L'un des rares rapports ACV ouvert, qui sera étudié à la section 3.2.2, [120] évalue, via l'ACV, l'empreinte carbone d'une visioconférence de 60 minutes et effectue l'état de l'art de la question ¹⁷. Les autrices constatent premièrement des écarts impressionnant allant jusqu'à des facteurs de plusieurs milliers entre les résultats. Parmi les 5 sources, il est décidé de se concentrer sur [215] et [216] considérés comme les seules sources universitaires du panel. Premier problème, le second article [216] ne fait que reprendre et re-traiter des résultats affichés du premier [215]. La seule source se trouve être [215], un article publié dans un journal de référence du domaine de l'évaluation environnementale ¹⁸. Malheureusement, cet article fait reposer son évaluation sur une base extrêmement fragile, un facteur d'impact unique de dépense énergétique pour un volume de données échangées. Les différentes estimations des services reposent uniquement sur le transfert de données par les infrastructures de télécommunications, un facteur d'impact simplifié pour les datacenters et l'empreinte carbone des mix électriques des pays concernés. Les tailles de volumes de données échangées, sur lesquels reposent tous les calculs, ne sont pas des mesures, mais sont des sources secondaires, majoritairement des articles de blog. Les approches d'évaluation des services numériques basées uniquement sur l'empreinte énergétique de la partie réseaux sont régulièrement critiqués, même l'évaluation de l'empreinte énergétique directe des réseaux avec comme unique facteur le volume de données transférées est considérée comme fallacieuse [217]. En dernier lieu, même en considérant toutes ces imperfections, les fourchettes de résultats établies par [215] pour une visioconférence de 60 minutes n'ont qu'une faible intersection avec les résultats de l'ACV. D'autant plus que cette intersection se fait entre la valeur maximale de [215] et la valeur minimum de l'ACV [120]. De plus, il faut considérer qu'il s'agit ici d'une empreinte carbone et que [215]

16. et le sera sans doute encore par la suite.

17. Tableau 19 du rapport.

18. *Resources, Conservation & Recycling* (RCR).

considère dans ses scénarios maximisant des locations où le mix électrique est fortement carboné, tandis que [120] considère le mix électrique français, très décarboné à l'échelle du monde. Il devient alors difficile de considérer qu'un résultat a été conforté ou consolidé dans cet état de l'art.

En conclusion, comme le dernier exemple autour de l'article paru dans le journal *RCR* nous le montre, l'immaturation du champ scientifique de l'ACV de services numériques est double. Il y a un manque de maturité de la communauté scientifique ACV étudiant les services numériques et un manque de maturité de la communauté informatique cherchant à appliquer l'ACV sur les services numériques. Dans la prochaine section, il sera donc question de constater la difficulté de construire des ACV de services numériques, sur la base de l'analyse de récents rapports à la méthodologie et aux données ouvertes.

3.2.2 Critique des rares études accessibles

Compte tenu des écarts observés entre les études et des constats réguliers de manque de standardisation entre les études, il est intéressant d'essayer d'analyser plus en profondeur les études d'ACV de services numériques dont la méthodologie est ouverte. La présente section se concentrera sur le corpus restreint suivant :

- **Évaluation de l'empreinte carbone d'une visioconférence entre deux utilisateurs du service rendez-vous** : ACV d'un service de visioconférence, rapport de recherche CNRS-EcoInfo pour RENATER paru en mars 2022 [120]
- **Life Cycle Assessment of CO₂ Emissions of Online Music and Videos Streaming in Japan** : ACV des services de musique et vidéo en streaming au Japon, article de journal paru en novembre 2021 [218]
- **Study of the environmental impact of the Tezos blockchain** : ACV de la blockchain de crypto-monnaie Tezos, étude du cabinet de conseil *PricewaterhouseCoopers Advisory* paru en décembre 2021 [119]
- **Analyse de cycle de vie (ACV) d'un service de messagerie en ligne** : ACV du service de messagerie instantané Treebal, étude réalisée par un industriel et un académique¹⁹ paru en août 2022 [121]

Ces documents ont été choisis, premièrement, pour être tous un minimum ouvert sur leur méthodologie et calculs, voir sur leurs données, deuxièmement pour leur aspect restreint aux services numériques, leur parution encore relativement récente, la diversité de profils des auteurs et enfin, partiellement pour l'aspect pédagogique de certains défauts relevés. Les limites des études seront regroupées par thématiques dans le but plus d'illustrer les difficultés courantes de l'ACV de services numériques plus que d'individuellement critiquer les ACV en question.

Objectifs

Un des premiers aspects, loin d'être trivial, est l'objectif affiché par l'étude. Si des critiques peuvent être émises de manière générale sur la qualité ou la précision de l'ACV, beaucoup de ces critiques prennent leur importance s'il s'agit de relever une incohérence entre objectifs et choix méthodologiques durant l'ACV.

19. doctorant.

Toutes les études expriment naturellement l'objectif d'estimer l'empreinte du service étudié. Les ACV [120] et [121] sont clairement plus focalisées sur un objectif d'éco-conception d'un service préexistant afin de réduire ses émissions, [120] détaille même une focalisation sur les paramètres de l'usage du service. Le premier but affiché par [119] est la communication, le rapport étant destiné à être publié par l'entreprise développant la blockchain doit évaluer le coût de la blockchain au regard de sa fonction première : effectuer des transactions. Enfin, [218] reste sur un objectif moins bien défini, qui est d'étudier les émissions des services de streaming musique et vidéo dans le cadre conjoint des politiques publiques de numérisation et décarbonation à l'échelle du Japon. Derrière un objectif commun d'étudier l'empreinte du service, se trouve dans les ACV de services numériques des objectifs et niveaux de détails différents qui auront leur importance ultérieurement.

Périmètre

La définition du périmètre relève les premières incohérences avec les objectifs. Si toutes les études souhaitent effectuer une ACV de services numériques, des incohérences existent entre les Unités Fonctionnelles (UF) définies, les objectifs et les périmètres. Pour rappel, l'UF est une manière de quantifier la satisfaction du besoin auquel répond le produit ou le service étudié par l'ACV. Il est déjà à noter une différence sur l'inclusion de la partie terminaux utilisateurs. Seul [120] prend non seulement en compte l'empreinte du terminal, mais évalue le coût énergétique du service sur le terminal de l'utilisateur. Si cette partie peut sembler marginale, [120] montre au contraire un impact significatif, d'autant plus que comme les autres services étudiés ici, l'usage du service implique un grand nombre d'utilisateurs. L'étude [119] bien que souhaitant évaluer le coût total de la blockchain, ne considère que le coût des transactions à l'intérieur du réseau, mais pas celui des acteurs, utilisateurs via leurs terminaux ou systèmes automatiques sur serveur, aux origines de ces transactions²⁰.

Si [218] et [121] incluent l'empreinte des terminaux en réalisant une allocation identique pour la phase d'usage et de fabrication des terminaux, c'est donc en ignorant ce surcoût quand bien même ces deux services ont un nombre très élevé d'utilisateurs. La définition des UF laisse aussi place à des imperfections.

Premièrement, [121] décide d'évaluer l'empreinte du service de messagerie au travers d'une UF simple : "Envoi d'un message de 300 caractères à un utilisateur, en France"²¹. Si cette UF est représentative du service et permet de représenter le coût des parties terminaux et réseaux, elle devient plus ambiguë pour le coût de la partie datacenters qui assure d'autres actions nécessaires au fonctionnement du service, en dehors de la réalisation de cette UF²².

Deuxièmement, [218] ne définit pas d'UF et va donc poser les bases d'une modélisation où streaming de musique ou de vidéo sont deux usages équivalents. Les différences en termes de consommation de données seront tout de même considérées, mais aucune différence ne sera faite entre ces usages dans les parties terminaux ou datacenters, ce qui

20. Voir Figure 5 de [119].

21. Le scénario lié à cette UF est par la suite défini plus en détails.

22. Autres fonctionnalités qui seront traités en analyse de sensibilité mais en prenant en compte la même partie datacenter, laissant entendre que les équipements datacenters alloués pleinement à cette UF sont dédiés à d'autres usages dans une proportion inconnue.

mènera par la suite à des résultats discutables.

Enfin, en termes de catégorie d'impact étudiés, la moitié des études restent en monocritère empreinte carbone [120, 218]. On peut déjà voir ici une certaine forme de contradiction pour une étude à large échelle comme [218] qui souhaite modéliser l'empreinte carbone dans un contexte d'accroissement de la demande en services numériques. L'absence de prise en compte de l'impact en raréfaction des métaux, particulièrement dans le cas japonais [219], peut non seulement transférer une partie de l'empreinte carbone en dehors du périmètre national, mais négliger le fait que la rareté des métaux se traduit par l'augmentation de leur coût, notamment carbone, d'extraction. Si les études [119] et [121] couvrent plusieurs impacts environnementaux, aucun ne prend en compte l'impact en radiation [187] qui est pourtant significatif dans les usages du numérique dans les pays occidentaux [18], encore plus en France. Ces panels réduits d'impact sont imputables selon les auteurs à la difficulté d'accès aux données.

Allocations

L'allocation est un enjeu critique de l'ACV de services numériques dû à la diversité des équipements souvent hautement mutualisés avec d'autres services. Il sera fait la différence entre l'allocation de l'empreinte d'usage d'un équipement et l'allocation de l'empreinte de fabrication²³. Les allocations les plus fréquentes sont, en dehors des hypothèses d'équipements dédiés au service²⁴, les allocations basées sur le temps d'usage et celles basées sur le volume de données. Les logiques d'allocations pour la partie réseaux sont assez homogènes. Du reste, si la partie terminaux utilisateurs renvoie à des natures et usages différents selon les services, les parties datacenters, plus homogènes dans leurs utilisations par les différents services, font pourtant l'objet d'allocations assez différentes.

La **partie réseaux** est calculée via une allocation au volume de données transférées, via un facteur d'impact en cycle de vie complet pour [120, 121], via un facteur d'impact de consommation électrique multiplié par le facteur d'impact du mix électrique utilisé pour [119, 218].

La **partie terminaux utilisateurs**, est ignorée par [119] et allouée au temps d'usage du service sur le terminal par [121, 218]. Enfin, [120] alloue les équipements utilisateurs, équipements réseaux sur le site utilisateur compris, pour la phase d'usage au volume de données transférées pour l'UF divisé par le volume de données traitées par l'équipement pour toute sa durée d'usage. Pour ce qui est de l'allocation de la phase de fabrication, elle est faite au temps d'usage dédié à l'UF²⁵ divisé par la durée de vie totale de l'équipement. Cela crée un problème, car s'il est possible de défendre que durant la visioconférence, l'ordinateur est pleinement utilisé pour le service²⁶ cela devient plus questionnable pour les équipements réseaux lourds, caractéristiques des laboratoires où a été mené l'ACV, qui assure significativement plus qu'une visioconférence à la fois. Cette surévaluation de l'empreinte de fabrication d'équipements lourds, inclus dans la partie utilisateur, pourrait expliquer la très forte part de l'empreinte de la partie utilisateur dans l'empreinte de

23. qui comprend éventuellement l'empreinte des phases de transport et fin de vie.

24. qui donnent donc une allocation de 1.

25. ici 60 minutes.

26. ce qui pourrait être discuté compte tenu des usages parallèles d'une telle machine, même pendant une visioconférence.

l'UF²⁷. L'allocation temporelle suppose généralement un usage de l'équipement complètement dédié à l'UF et est souvent risquée dans le cadre de l'ACV de services numériques.

La **partie datacenters** fait elle l'objet d'allocations très diverses. Pour [120] il s'agit du même mécanisme décrit précédemment, avec comme nuance que l'ACV a effectué les allocations non pas sur des serveurs en entier, mais sur des machines virtuelles dédiées, minorant le problème de l'allocation temporelle décrite précédemment. Pour [119] les équipements traitant des données de la blockchain sont considérés comme dédiés aux services sauf dans le cas d'usage de location de serveur de datacenters où un facteur d'allocation arbitraire est défini par les auteurs²⁸. Dans le cas de [121], un serveur est dédié au service, son empreinte est allouée selon le nombre d'UFs traitées pendant une année de référence. Ce choix de modélisation suppose que les ressources affichées pour le datacenters sont uniquement dédiées à l'UF et non au service de messagerie en entier, ce qui n'est pas explicite dans le rapport ACV.

Pour finir, le cas de [218] est particulier. À l'échelle du Japon, les auteurs considèrent les émissions des datacenters, et les allouent, empreinte de fabrication et usage, selon les transferts de données transférées hors de ces datacenters. De même, les auteurs recueillent la consommation électrique moyenne de terminaux utilisateurs, le nombre d'heures moyennes sur une année de streaming audio et de streaming vidéo, le nombre moyen de données transférées par heure de streaming audio ou streaming vidéo. Toutes les émissions sont supposément faites selon un facteur d'impact GES émis par Mb de données transférées. Cependant, les auteurs arrivent à des résultats assez étonnants en contradiction avec l'état de l'art²⁹. Le streaming vidéo sur smartphone serait par exemple équivalent au streaming audio en termes d'émissions de GES, plus étonnant encore, le streaming de musique sur PC serait au moins 3 fois plus émetteur de GES que le streaming vidéo. Pour comprendre comment les auteurs en sont arrivés à ce qui ressemble à un contre-sens dans leur allocation, il faut remarquer que les résultats sont donnés en $kgCO_2/Mo$. Les auteurs ont calculé pour chaque service la somme des émissions GES des heures d'utilisation d'un terminal par an, avant de la diviser par la somme le volume total de données transférées par ce service pendant un an. Cela donne des résultats qui ne mettent pas en avant qu'une heure de streaming audio ne demande pas autant de données qu'une heure de streaming vidéo. Cela veut aussi dire que les auteurs considèrent que le terminal utilisateur, alloué selon le temps, est entièrement dédié à l'activité de streaming, ce qui cohérent pour le cas du streaming vidéo, moins pour le streaming audio. En incluant dans un modèle descendant les terminaux utilisateurs et en considérant deux usages différents comme équivalents, les auteurs ont produit un résultat contradictoire, qui illustre l'importance de définir correctement UF et clés d'allocations cohérentes.

Analyses de sensibilités et projection

Usuellement, les ACV de services numériques se terminent par des analyses de sensibilités et des conclusions au regard des objectifs de l'ACV. L'ACV de services numériques a

27. Même si les autrices reconnaissent aussi que l'empreinte d'utilisation de ces équipements réseaux côté utilisateur a pu être plus importante due à leur faible usage inhabituel pendant l'étude.

28. Les auteurs considèrent par exemple dans un cas que l'utilisateur hébergeant un nœud de blockchain sur un serveur n'utilise que 10% du serveur, sans plus de précisions sur l'établissement de ce pourcentage.

29. Table 6 de [218] contenant les résultats indiquant plus d'émissions carbonées pour le streaming audio que le streaming vidéo.

pour caractéristique, vraisemblablement lié à son secteur, de mettre une certaine emphase sur les scénarios de croissance. N'étant pas basé sur un service commercial, [120] reste sur ses objectifs d'évaluation des modalités d'usage. De son côté, [218] effectue une projection peu détaillée, multipliant le nombre d'utilisateurs. Cette projection ne semble pas tenir compte de la croissance de l'infrastructure, nécessaire à une telle augmentation de la charge. Cette limite est courante. Elle est due à l'utilisation d'une ACV attributionnelle à un cadre d'usage qui demanderait une ACV conséquentielle. L'exercice est donc périlleux et peut mener à des conclusions de qualité incertaine.

Par exemple, les résultats finaux de [119] mettent en avant une réduction du coût de la transaction au cours entre l'année 2020 et l'année 2021, avec cependant une augmentation globale de l'empreinte de la blockchain et de l'empreinte par hébergeur. Cela peut être dû au fait que d'une année à l'autre, la blockchain a changé le "coût" cryptographique des opérations en le diminuant. Il est aussi possible que cette différence soit due au fait que d'une année à l'autre, les données sur le nombre de transactions et d'hébergeurs sont mises à jour, le nombre de transactions a bondi et le nombre d'hébergeurs légèrement baissé. Mais les données concernant le nombre moyen de nœuds par hébergeur, leurs caractéristiques et activités, sont restées fixes car évaluées par un unique sondage manuel auprès des hébergeurs. La seule utilisation de ressources mises à jour est le volume d'échange de données sur les réseaux. Le modèle descendant, de l'empreinte de l'infrastructure blockchain divisé par le nombre d'opérations effectuées semble produire un phénomène où **l'activité varie mais l'infrastructure reste fixe**. Ce décalage est dû au problème de dimensionnement des ressources auquel cette étude n'a pas répondu. L'étude ne propose pas un modèle ascendant permettant d'associer un usage à un coût. Cette incertitude ne permet pas de tirer de conclusions nouvelles sur cette technologie qui malgré le fait qu'elle possède un protocole moins coûteux que d'autres blockchains produit les mêmes problèmes environnementaux [220,221] et s'inscrit dans la même dynamique de croissance globale de l'empreinte [222].

Une problématique similaire peut être observée chez [121]. Une des analyses de sensibilité propose une croissance de l'infrastructure vers une optimum. Il y est montré que l'augmentation du nombre de messages diminue l'empreinte de la partie datacenters par message. Il y est admis que l'augmentation peut se poursuivre par l'ajout de nœuds, chaque nouveau nœud pouvant soutenir une charge de 71500 messages par jour. Néanmoins, ce scénario de croissance conjointe du nombre de messages par jour et ajouts de nœuds, affiche une courbe lisse de diminution de l'empreinte datacenters par message³⁰. Ce genre de projection de croissance du service donnant à voir une évolution fluide de l'empreinte peut être expliqué par la présence de facteur d'impact dans la modélisation, comme l'illustre la Figure 3.2. La croissance d'un service, même avec les avantages d'un passage à l'échelle fluidifiée grâce au paradigme cloud, reste une perturbation pouvant briser des équilibres, diminuer l'efficacité d'économies d'échelle. Dans un contexte où le dimensionnement des ressources du service observé est déjà très incertain, projeter la croissance de ce service ne peut qu'amplifier ces incertitudes. De plus, ce scénario de croissance pâtit aussi du fait que le service est vu uniquement sous l'angle de l'addition d'UF dont

30. En détail, cette croissance lisse dans l'analyse de sensibilité de [121] est due à la fois à un choix de représentation des données et à ce qui semble être des imperfections dans la modélisation, comme la non-prise en compte de l'ajout du second nœud ou le fait que l'empreinte d'usage du datacenter ne varie pas malgré la montée en charge.

on ne sait pas dans quelle mesure elles sont représentatives du système. L'augmentation du nombre de messages simples, n'est pas une augmentation du nombre d'utilisateurs qui risquent de plus en plus de créer des groupes de discussions multipliant l'impact de leur message et des pièces jointes associées. Cette analyse de sensibilité, illustre encore une fois la difficulté autour de la notion de dimensionnement et de correspondance entre une UF simple et un coût.

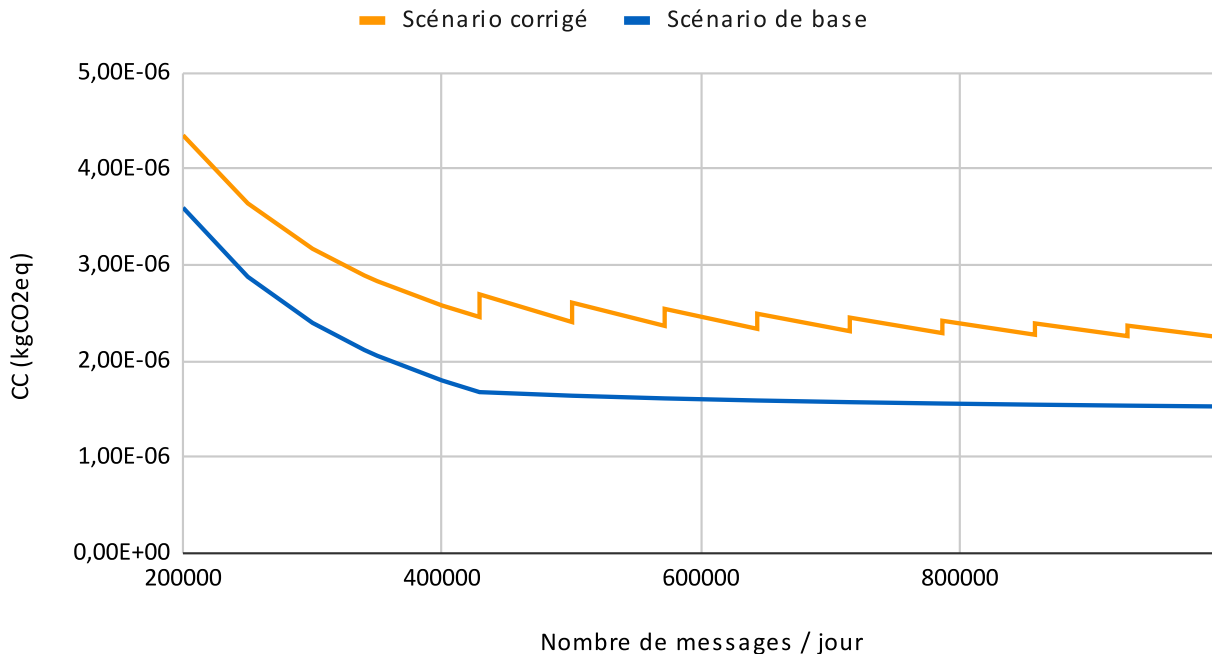


FIGURE 3.2 – Comparaison entre l'évolution de l'impact de la partie datacenters de [121] en fonction du nombre de messages par jour, avec en bleu une modélisation par facteur d'impact et en orange une modélisation prenant en compte la montée en charge de l'empreinte de l'usage de la partie datacenters et les ajouts de nouveaux nœuds, cas d'usage repris de [121]

3.3 Analyse de l'état de l'art

Quelles leçons tirer des deux dernières sections ? Comment concilier les avantages de l'ACV de services numériques avec les limites constatées ? Cette section propose premièrement, dans une démarche similaire à [217], une aide à la lecture critique d'ACV de services numériques, un ensemble de points d'attention à examiner lors d'une lecture. Deuxièmement, toujours dans une optique de synthétisation des apports de l'ACV de services numériques, une simplification des modèles pouvant aider à mettre en valeur les forces d'un service, d'une étude et les compromis exprimés par le modèle.

3.3.1 Comment lire une ACV de services numériques

Quels sont les points à particulièrement surveiller lors de la lecture d'une ACV d'un service numérique ? Au sein de rapport volumineux, bien plus épais que des articles aca-

démiques, quels sont usuellement les sujets critiques. Si certaines recommandations se recoupent avec des préconisations générales de revue critique d'ACV [223], le but de cette section est de rester dans le cadre de l'exercice d'ACV de services numériques et les recommandations sur l'ACV ou les études sont à lire ici comme des recommandations pour l'ACV de services numériques.

Correspondance entre objectifs, unité fonctionnelle et périmètre

La cohérence entre ces 3 paramètres est cruciale et est à observer tout au long de l'étude. Il est important qu'ils soient donc clairement définis en début d'étude.

Les **objectifs** peuvent classiquement être d'explorer l'empreinte du service, si c'est le cas, il peut être pertinent de préciser des objectifs plus précis comme l'impact des modalités d'usages pour [120], la localisation de l'empreinte pour le chapitre 5 ou encore la distribution de l'empreinte chapitre 6. L'objectif peut être de créer un résultat sur lequel communiquer, dans ce cas, il est recommandé de s'intégrer dans une démarche d'affichage environnementale qui doit être cohérente avec le contexte de communication qui peut être normé, comme en France [224], avec des recommandations spécifiques pour les services numériques [189]. Les cadres d'éco-conception et de comparaison devront porter une attention particulière aux scénarios mis en place, voir songer à une approche plus conséquente.

La définition des **UF** doit être claire et laisser le moins d'ambiguïtés possible. Souvent, 2 types se retrouvent, les UF basées sur un acte métier simple³¹ et celles basées sur un service³². La deuxième s'apparente à celle recommandée par le standard ITU [152] pour les services TIC, à savoir : "usage du service pendant un an"³³. Il est fréquent que dans un rapport ces deux UF soient présentes, mais qu'une seule soit réellement évaluée et que l'une ne soit que le produit ou la division de l'autre, ce qui peut amener des incohérences si ce point n'est pas éclairci. Il est important que les actes décrits par l'UF puissent être considérés comme offrant une qualité de service identique. Il est possible et conseillé de détailler cette qualité de service avec des indicateurs quantitatifs.

Enfin, pour le **périmètre**, l'architecture 3 tiers est une généralisation, l'important est d'inclure l'ensemble des équipements et infrastructures numériques nécessaires à l'usage. Il y a deux points fréquents de débat sur les frontières dans les ACV de services numériques : l'inclusion du développement du service et l'inclusion de la consommation du logiciel sur le terminal utilisateur. L'inclusion du coût de développement a fait l'objet de recherches [122], cette question peut se trancher au regard des objectifs. Le coût du développement, comme présenté dans [122], est par exemple tout à fait approprié pour une évaluation *ex ante*, dans un cadre comme l'aide à la prise de décision entre plusieurs projets ou parties de projet. Ce coût de développement peut être inclus d'une manière qui le confond avec le coût de maintenance comme dans [119, 120]. L'inclusion de la consommation électrique du service revient souvent, plus comme un enjeu technique, son inclusion demandant une modélisation plus fine et une bonne connaissance des terminaux clients du service [120]. Plus précisément, la question est celle de la prise en compte de la sur-

31. envoyer un message, générer une image, effectuer une visioconférence de 60 minutes, etc.

32. Permettre à un utilisateur d'accéder pendant un mois à un catalogue de musiques en streaming, héberger le service étudié pendant 1 an.

33. Ou plus prosaïquement : "Annual service use"[sic] [152].

consommation d'électricité du terminal pour assurer l'UF, la consommation médiane ou *idle* du terminal étant déjà incluse dans l'empreinte d'usage du terminal. Souvent, ce coût est écarté, considérant par exemple que le service est accessible par navigateur internet et que la consommation moyenne d'un terminal est assez représentative de cet usage, ou alors non pris en compte faute de connaissance assez fine sur les terminaux utilisateurs du service.

Clés d'allocation claire et cohérente avec les objectifs

Le haut degré de mutualisation des équipements et infrastructures numériques rend particulièrement sensible le choix des clés d'allocation, les équipements dédiés au seul service étant rares. Tout d'abord, l'empreinte de la phase d'usage des terminaux et datacenters peut être de plus en plus évaluée via des outils de mesure directe de la consommation électrique [225, 226]. L'évaluation de l'empreinte d'usage de la partie réseaux diffère selon l'usage du réseau filaire [217] ou mobile [227], mais doit généralement se reposer sur une modélisation.

Cependant, l'ensemble de l'empreinte d'usage du service ne peut pas toujours être mesurée et il reste à allouer l'empreinte de fabrication. En l'état, l'inclusion des phases de transports et fin de vie est soit faite dans l'empreinte de fabrication, soit non effective par manque de traçabilité des biens et connaissance du scénario de fin de vie³⁴ [99]. Pour allouer l'empreinte, i.e. estimer la part d'empreinte de l'équipement imputable au service, plusieurs types d'allocations existent.

L'**allocation temporelle**, temps d'usage de l'équipement par le service divisé par le temps total d'usage de l'équipement, est un choix fréquent. Il est d'abord à vérifier si l'allocation temporelle prend en compte le temps d'usage total de l'équipement, i.e. temps d'utilisation effective de l'équipement, ou sa durée de vie, le temps écoulé entre la mise en service et la fin d'usage de l'équipement. Comme montré par [121] ce choix peut avoir une forte influence. Il sera souvent préféré de prendre en compte le temps d'usage et faire le choix inverse devra être justifié³⁵. L'allocation temporelle suppose que l'équipement est entièrement dédié au service pendant le laps de temps sélectionné. C'est une hypothèse forte dont la véracité doit être surveillée. Elle peut convenir à un terminal nécessitant l'attention soutenue de l'utilisateur, moins à un serveur hébergeant plusieurs services ou gérant plusieurs utilisateurs en même temps.

L'**allocation au volume de données transférées**³⁶ est la seconde forme d'allocation fréquente en ACV de services numériques, mais présente aussi des limites. Elle possède l'avantage, comme l'allocation temporelle, de pouvoir être mesurée plutôt facilement. Ceci-dit, cette forme d'allocation suppose que le volume de données transférées est un indicateur représentatif de l'intensité d'usage d'un équipement. Cela se défend bien pour des équipements réseaux comme des routeurs, ou même pour certains serveurs dédiés uniquement aux transferts de flux média, mais cela devient moins pertinent pour un équipement aux usages hétérogènes en termes d'utilisation du réseau. Même dans le cadre

34. On ne sait pas si les équipements sont recyclés ou non collectés. Une hypothèse raisonnable pourrait être de considérer par défaut les équipements numériques comme non-collectés.

35. Par exemple, dans un cadre conséquentiel où le service viendrait utiliser un équipement, hors de son usage classique pendant une fenêtre de temps flexible où aucun autre usage ne peut être fait de l'équipement.

36. Souvent prononcé "Allocation au Giga" parfois ramené au concept de "One byte model" [228].

du réseau, allouer l'empreinte d'usage des réseaux selon cette clé est contesté [217] mais peut se défendre pour l'empreinte de fabrication.

En conclusion, suite à ces deux exemples communs, il est recommandé de prêter une attention prioritaire aux méthodes d'allocations. Les méthodes d'allocations doivent permettre d'utiliser ou être remplacée par des mesures directes si possible pour la phase d'usage. Les clés d'allocation doivent pouvoir être un critère juste de partage de l'empreinte d'un équipement, ce qui suppose de poser la question de l'usage fondamental de l'équipement en question. Est-ce un équipement dont le travail est de donner un certain temps d'écran, transférer un certain volume de données ou stocker un certain volume de données ?

Dimensionnement, estimé par modèle ou constaté par observation

Plutôt associée à la partie datacenter³⁷, la problématique du dimensionnement dans les ACV de services numériques est particulièrement critique dès qu'il s'agit d'effectuer des projections sur l'avenir d'un service existant ou dès que la partie datacenters est considérée en boîte noire. Cela dépend aussi du positionnement des auteurs en termes de modélisation du système, de quels types de données directes partent-ils ? Est-ce que les auteurs sont positionnés en termes de données, côté client ou côté serveur ? Dans le cas d'une ACV partant des informations clients, il est à noter que souvent le dimensionnement des ressources datacenters correspond au pic d'activité des utilisateurs. Dans le cas contraire, si des informations fiables ont été recueillies sur le dimensionnement des ressources pour une certaine intensité d'usage³⁸, il est précieux de les mettre en valeur dans le rapport comme scénario de base, et de rester prudent dans les hypothèses sur la variation de cet usage sans changement dans l'infrastructure. En somme, comme nous le verrons dans la prochaine section, le dimensionnement des ressources est un point critique de la modélisation proposée par l'ACV de services numériques auquel il faut porter attention.

De la qualité variable des bases de données ICV

Bien que cette critique existe aussi en général dans le monde ACV, elle est particulièrement présente dans le cadre des ACV de services numériques et devient donc une composante de l'exercice. Il est important de nuancer et délimiter dans les résultats de l'ACV ce qui peut être influencé par une mauvaise qualité de données. Toutes les catégories d'impact ne sont pas soumises aux mêmes variations suivant la base de données [58, 214], la catégorie ADP pour la raréfaction des métaux et minéraux l'est particulièrement, la consommation d'eau aussi dans une moindre mesure. Dans le cas d'une unique source pour les données ICV, il est toujours possible d'observer préférentiellement la distribution des impacts que les résultats absolus. De même, une telle réserve sur la qualité des données peut inciter à présenter et à construire les résultats comme une borne basse, tel que fait au chapitre 5.

37. des enjeux similaires peuvent théoriquement exister pour les parties terminaux ou réseaux [105] mais cela reste moins courant.

38. quantifiable en nombre d'utilisateurs simultanés, nombre de requêtes, nombre d'UF complétés.

3.3.2 Proposition d'une modélisation simplifiée de l'ACV de services numériques

Dans la continuité des conseils précédemment exposés sur les points d'attention à avoir lors de la lecture d'une ACV de services numériques, voici une analyse portée plus spécifiquement sur la modélisation quantitative présente dans les ACV. C'est un modèle simplifié haut-niveau qui peut être construit avec comme base un rapport ACV et servir à mieux communiquer à son sujet.

L'ACV est le calcul des impacts environnementaux potentiels d'un service numérique. C'est un exercice comptable. Partant de l'ACV de services numériques, il sera question, d'une part, de réduire la modélisation du service étudié à quelques variables et, d'autre part, utiliser l'analyse de sensibilité pour réaliser des scénarios sur l'avenir du service. Dans cette section, l'étude ACV sur l'application de messagerie Treabal [121] sera prise comme exemple. Pour rappel, l'UF étudiée dans l'ACV de Treabal est "Envoi d'un message de 300 caractères à un utilisateur, en France".

Un service numérique est l'utilisation organisée de différents équipements et infrastructures pour répondre à un besoin. Soit E , E est l'ensemble des équipements et infrastructures nécessaires pour réaliser une certaine activité A . A , l'activité, représente le nombre d'Unités Fonctionnelles (UF) satisfaites par le service. Il sera considéré que les UF comptabilisées dans A sont identiques, soit que le service réalise chaque UF de manière identique, soit que A est la somme d'UF moyennes identiques.

Soit enfin la fonction $I(E, A)$, c'est-à-dire l'impact environnemental en fonction de l'ensemble des équipements et de l'activité (type et quantité) exercée grâce à ces équipements. La fonction $I(E, A)$ donne comme résultat un vecteur d'impacts, chaque valeur du vecteur correspondant à une catégorie d'impact différent (émission de carbone, utilisation de ressources en minéraux et métaux, consommation d'énergie, etc.).

E , A et $I(E, A)$ peuvent être divisés par sous-parties du service étudié, correspondant par exemple à l'architecture classique en 3 tiers, terminaux, réseaux et datacenter. Par exemple, ici, dans le cas de Treabal, $E_{Terminaux}$ est constitué d'un smartphone moyen. Similairement, $A_{Réseaux} = 1$ serait l'action de transmettre à travers les réseaux 2 fois 300 octets (de l'émetteur au serveur puis du serveur au récepteur), 60% du temps en moyenne via le réseau mobile et le reste du temps via le réseau filaire classique.

Enfin, le résultat, étudiant l'impact pour une unique UF issue de l'ACV de Treabal, peut être représenté par :

$$I(E, 1) = I_{Terminaux}(E, 1) + I_{Réseaux}(E, 1) + I_{Datacenters}(E, 1) \quad (3.1)$$

$$I(E, 1) = (CC, ADP, PE) = (3.43e^{-4}, 1.62e^{-8}, 8.06e^{-3}) \quad (3.2)$$

Où CC, ADP et PE représentent les catégories d'impact utilisées dans l'étude. CC est le changement climatique exprimé en kg CO_2 eq. ADP est l'épuisement des ressources abiotiques, exprimé en kg Sb eq. PE est l'énergie primaire utilisée, exprimée en MJ. Compte tenu de la simplification précédente, il devient possible de pousser la modélisation plus loin en posant les ratios suivants :

- $Rm = \frac{A}{E}$: Le ratio E pour A , ou l'efficacité matérielle d'une activité à réaliser, c'est-à-dire quel est l'ensemble E d'équipements nécessaires pour réaliser une certaine quantité d'UF décrite par A ? Ce ratio peut être vu comme un indicateur de dimensionnement des ressources pour une certaine activité.

- $Re = \frac{I(E,A)}{E}$: Le ratio E pour $I(E, A)$, ou l'efficacité environnementale des équipements, c'est-à-dire quelle est l'influence des équipements sur les impacts? Comme E est un ensemble hétérogène de matériel, ce ratio est plus utile si l'on considère tout l'ensemble ou une partie moyenne hypothétique de celui-ci.
- $Ra = \frac{I(E,A)}{A}$: Le ratio A pour $I(E, A)$, ou l'efficacité environnementale de l'acte unitaire, qui est finalement celle souvent présentée par les études s'intéressant à l'impact d'une seule action du service.

L'objectif de ces ratios est de résumer chaque choix particulier concernant la gestion ou l'éco-conception du service à une influence sur ces 3 ratios ou leurs sous-parties. Par exemple, les choix potentiels explorables en analyse de sensibilité peuvent être :

- La mise en place côté datacenters de technique de consolidation ou mutualisation de machines virtuelles [229, 230] est un levier principalement de type Rm qui agit sur le nombre de ressources à mobiliser pour réaliser l'UF. L'achat de matériel plus performant rentrerait aussi dans cette catégorie. Dans l'étude Treebal, le ratio Rm est représenté par l'une des hypothèses qui est de considérer qu'au-delà de 429 000 messages par jour, un nœud de centre de données est ajouté à $E_{DataCenters}$ pouvant gérer jusqu'à 71500 messages/jour supplémentaires. Cette hypothèse de dimensionnement est un indicateur crucial à relever dans un rapport ACV.
- L'utilisation d'équipements avec une meilleure efficacité énergétique est un levier permettant de jouer sur Re . L'augmentation de la durée de vie moyenne du smartphone, considérée comme hypothèse dans l'ACV de Treebal, réduit l'empreinte de $E_{EndUser}$ et donc de $Re_{EndUser}$.
- Limiter les fonctionnalités mises en place pour satisfaire l'UF ou apporter des modifications à la conception du service modifiera Ra . Ce ratio représente un levier qui peut être actionné par le design du service. Si l'UF de Treebal était "Usage du service pendant un an", enlever la fonctionnalité "compteur d'arbres plantés grâce à Treebal" pourrait par exemple être une économie qui permet de toujours de satisfaire l'UF. Remettre en cause la manière dont l'UF est satisfaite [231] ou limiter l'intensité d'usage est un levier de type Ra .

Il convient de noter que les ratios sont valables pour une qualité de service (QoS) fixe. L'objectif de ce cadre est d'essayer de faire la distinction entre les différents types d'efficacité. Les résultats numériques de l'étude ACV sont maintenant réduits à une liste réduite de variables pouvant servir de support à des Analyses de Sensibilité (AS). La modélisation de A pourrait être par exemple décrite plus finement par un ensemble de variables continues : le temps X passé sur le terminal de l'utilisateur, la quantité Y d'octets transférés sur le réseau, le nombre Z de données traitées dans le centre de données et ainsi de suite.

Comme il a été vu précédemment, il est courant dans la conclusion des études ACV d'avoir des éléments sur le futur du service, un simple commentaire ou une projection, étayée ou non par une analyse de sensibilité. L'hypothèse étudiée pour répondre à cette question est alors celle de la croissance du service³⁹. La modélisation des services proposée par les ACV est souvent basée sur l'observation du système pendant une période de temps définie et une certaine activité. C'est souvent sur la base de cette activité, et le

39. l'inverse n'est jamais retenu à ce sujet.

dimensionnement des ressources observé, que l’empreinte d’un UF unitaire est calculée en divisant l’empreinte totale d’une partie ou de la totalité du service par le nombre d’UF traitées [119, 121, 218]. Ce type de modélisation présente des AS de croissance de l’activité. En d’autres termes, Ra est observé lorsque A augmente, ce qui nous amène mécaniquement à une diminution de Ra , négligeant la croissance d’infrastructures [105].

Ce problème d’infrastructure fixe pour activité variable peut être mieux mis en évidence par cette simplification du modèle issu d’un rapport d’ACV. Il est d’ailleurs plus fréquent d’observer ce biais lorsque la fonction $I(E, A)$ est calculée sur la base de facteur d’impact, e.g. d’allocation unique comme pour le réseau, que d’inventaire précis du nombre d’équipements et de leur intensité d’usage. Le caractère difficile du dimensionnement de ressources datacenters $Rm_{DataCenters}$ pour assurer un service est une caractéristique typique de l’ACV de services numériques par rapport à d’autres ACV. En outre, une des qualités fondamentales de l’ACV pour l’éco-conception de services est le questionnement du besoin au travers de la définition de l’UF qui permet d’utiliser Ra comme un indicateur de performance sortant des conceptions classiques de services numériques.

3.4 Conclusion : motivation des contributions de cette thèse

Il a été mis en perspective dans le chapitre 2 les différents travaux de recherche ayant conduit à l’émergence d’un champ de la discipline informatique s’inscrivant dans la philosophie des TIC soutenables [32], qui pourrait être nommé informatique écoresponsable ou encore informatique frugale . Devant l’importance de la méthode ACV, et plus particulièrement de l’ACV de services numériques, il a été exposé dans ce chapitre les avantages et limites de cette méthode au début des travaux de thèse, ainsi que les points méthodologiques les plus critiques dans l’application actuelle de cette méthode.

Partant des constats des chapitres 2 et 3, les contributions de thèses ont été posées en conséquence, tel que montré sur la figure 3.3. La motivation fondamentale de la thèse s’inscrit dans la question de la place du numérique au sein des limites planétaires [232]. Pour répondre à cette question, l’ACV de services numériques paraît être un des outils adéquats. Cependant, cet outil est présentement imparfait. Au travers d’études de cas, aux chapitres 5 et 6, et s’appuyant sur une méthodologie présentée au chapitre 4, ce manuscrit propose 3 types de contributions :

- Contribution contextuelle aux connaissances sur le numérique soutenable par la production de connaissances sur l’empreinte de différents services numériques.
- Contribution méthodologique à l’éco-conception des services numériques en proposant de nouveaux indicateurs et grilles de lecture de la soutenabilité des services numériques.
- Contribution méthodologique à l’ACV de services numériques en elle-même, par le travail d’état de l’art et d’application de la méthode à des cas d’usages concrets.

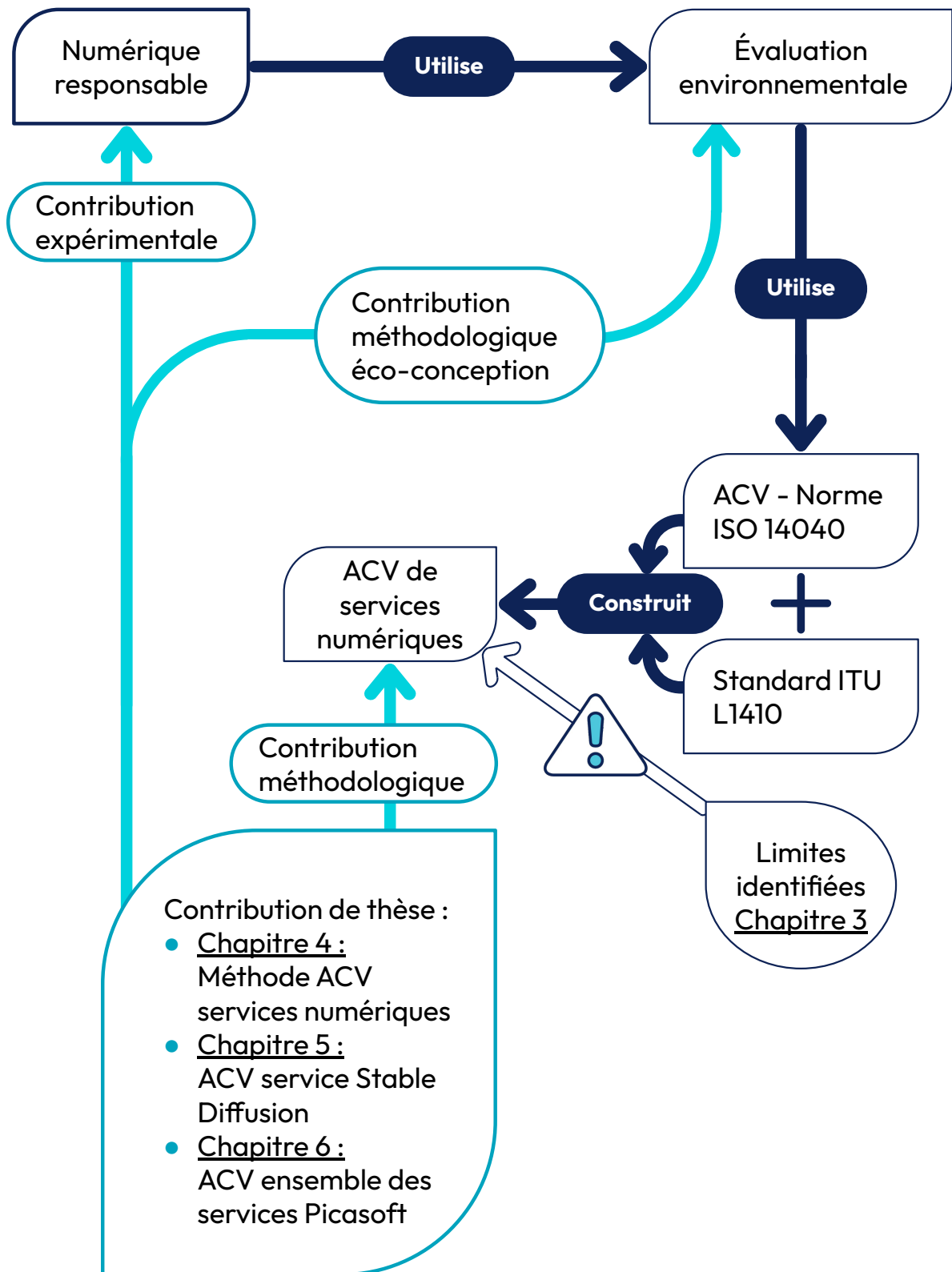


FIGURE 3.3 – Carte heuristique des contributions de la thèse

ACV de service numérique

[à propos de l'ACV] The exact definition is the subject of some controversy, and has been the subject of several recent workshops and studies.

Robert U. Ayres

Le but de ce chapitre, au-delà de la mise en valeur d'une contribution du manuscrit, est de présenter la méthode qui a guidé les travaux effectués durant cette thèse, dont deux sont présentés respectivement dans les chapitres 5 et 6.

La méthode d'ACV de services numériques présentée ici est inspirée de [120], mais aussi [121] et de la formation ACV de services numériques de Julie Orgelet et Étienne Lees-Perasso, suivant les recommandations de [152] pour les services TIC ainsi que [233]¹. Toutefois, elle est aussi l'aboutissement actuel de toutes les ACV de services numériques précédemment accessibles. Elle est une synthèse, un carcan général, à adapter à chacune des situations, en fonction des objectifs, mais aussi de la nature et qualité des données accessibles. Cette méthodologie sera sous certains angles une *simplification* de la méthode ACV générale, une spécification à l'usage des praticants de l'informatique afin d'obtenir une meilleure connaissance de leur sujet d'étude. Cette meilleure connaissance du sujet est celle des impacts environnementaux majoritairement, certes, mais l'apport de l'ACV, la "pensée en cycle de vie" ou "l'approche scientifique", caractéristiques de l'ACV [131], apportent une prise de recul, demandent un approfondissement des connaissances sur le sujet qui va au-delà d'un simple calcul des impacts environnementaux.

Il est utile de rappeler, en introduction du présent chapitre, que si la méthodologie est présentée dans un ordre pouvant paraître chronologique allant de choix généraux aux calculs de résultats quantitatifs, le processus d'une ACV n'est pas aussi linéaire. Comme exposée par la figure 2.6, l'ACV est un processus itératif comportant des allers-retours d'améliorations entre les différentes parties décrites. Ce chapitre est composé de deux sections majeures, les sections 4.1 et 4.2. La première se concentre le cadrage de l'étude,

1. Particulièrement le chapitre 37 *LCA Cookbook*, synthèse de nombreux chapitres du livre précédemment cités dans ce manuscrit.

les choix d'objectifs, de périmètres, mais aussi les choix de bases de données et d'outils de mesure. Cette section a donc un aspect plus prospectif, c'est un état de l'art et une boîte à outils. La seconde section se concentre pour sa part sur la phase d'inventaire et se concentre sur les méthodes de calculs et les problématiques d'allocations. Les sections 4.3 et 4.4 se concentreront sur le traitement et l'analyse des résultats, tandis que la section 4.5 conclura sur les limites et apports de la méthodologie proposée dans ce chapitre.

4.1 Une approche attributionnelle du service numérique : échelle et objectifs

La méthodologie d'ACV de services numériques proposée ici se définit plutôt comme attributionnelle, bien que les connaissances présentées ici puissent être utiles dans un cadre conséquentiel. Le choix de la démarche attributionnelle dans ce travail de recherche est celui de la pratique dominante², dans un contexte où l'objectif reste l'application de l'ACV aux services numériques. En effet, l'approche conséquentielle reste caractérisée par un degré de maturité méthodologique moindre [235, 236] qui, ajoutée à l'immaturité constatée durant les chapitres précédents de l'ACV de service numérique, risqueraient de ne pas produire des résultats scientifiquement satisfaisants. L'approche conséquentielle n'est bien sûr pas à proscrire, des progrès sont observés [237], des applications en services numériques apparaissent [127], notamment dans le domaine d'évaluation des impacts de 2nd et 3e ordre [150, 238].

4.1.1 Objectifs et périmètre

Cette section est celle du cadrage d'une étude ACV. Pourquoi et comment évalue-t-on une unité fonctionnelle ? Le but de ces paramètres est la cohérence. C'est au regard de la cohérence avec les objectifs de l'étude, au regard de la conformité avec le périmètre du service numérique étudié, que les qualités de l'étude seront jugées.

Objectifs

Il est possible de résumer la définition des objectifs de l'ACV de services numériques à 3 enjeux principaux :

- **Applications et finalités recherchées** - Il s'agit de définir le but fondamental de l'ACV. Est-ce qu'il s'agit d'explorer l'empreinte, d'observer la répartition de l'empreinte, entre catégories d'impacts ou parties du service ? Est-ce qu'il s'agit d'un enjeu d'éco-conception de service [180], pour un service en cours de développement ou pour un service existant ? Est-ce que le but de l'étude est la communication, l'affichage environnemental, un rapport d'impact s'inscrivant dans un standard comme [189] ? Est-ce que l'étude sera une ACV comparative ? Est-ce qu'il s'agira d'une comparaison entre des implémentations différentes d'un même service numérique, entre des services numériques répondant à une même Unité Fonctionnelle

2. La section 9.2.3 de [234] consacrée à la modélisation conséquentielle se termine par le constat que l'approche attributionnelle comporte moins d'incertitudes et est en moyenne plus correcte que l'approche conséquentielle, même pour des études dont le but est d'étudier différents scénarios de prise de décision.

(UF), ou la comparaison entre des services numériques et des solutions non numériques [33] ? Chaque type d'objectif s'inscrit dans le prolongement d'un état de l'art qui doit être examiné, notamment afin d'explicitier les choix de périmètres et d'UF qui s'inscriront ou non dans les pratiques pré-existantes.

- **Limites identifiées** - L'étude comporte, de par la nature de ses objectifs ou de son périmètre, des limites, notamment aux applications potentielles de l'étude. La limite principale est généralement l'accès aux données, plus précisément l'accès à des données d'Impact du Cycle de Vie (ICV) des équipements. Ces données ICV peuvent être limitées, en variété d'équipements couverts, e.g. ne comporter des données ICV que pour un seul type de serveur, ou en nombre de catégories d'impacts, e.g. seulement l'empreinte carbone ou quelques catégories d'impacts. Les limites identifiées peuvent être aussi, fréquemment dans le cadre des services numériques, des limites d'accès aux données d'activités du système, voire une incapacité à décrire la réalité physique du système³. Il est pertinent de préciser le niveau d'inconnu pour chaque niveau du service. Est-ce qu'il y a des inconnus sur les activités des utilisateurs, leur type de terminaux ? Est-ce qu'il y a des inconnus sur le type de réseaux utilisés, la localisation de l'activité réseau ? Y a-t-il des inconnus sur l'activité des serveurs de datacenters, leur configuration ou leur nombre ? Il est alors nécessaire d'explicitier si certaines inconnues sont comblées par des hypothèses ou non, et les limites que ces choix génèrent.
- **Par qui ? Pour qui ? Pourquoi ?** - Il est d'usage dans les ACV, hors du monde académique, dans un souci de transparence, d'exposer clairement le public cible, le commanditaire de l'étude ACV et ses motivations. Cet enjeu est traité dans le monde académique, que cela soit dans les déclarations d'intérêt avant publication ou la partie *Acknowledgment* en fin d'article. Le but est ici la transparence en termes de motivations, acteurs impliqués et financement. Il est loin d'être inutile de rappeler cet enjeu compte tenu des nombreuses controverses autour de l'empreinte environnementale du numérique dont il a été question durant les chapitres précédents, controverses s'appuyant parfois sur la présence d'acteurs proéminents du numérique dans ces sujets [21, 23, 24].

Périmètre du service numérique

La question du périmètre est particulièrement sensible dans le cadre des services numériques, à grande échelle [21], mais aussi dans le cadre des services numériques [22]. Il existe des raisons propres au numérique qui expliquent les difficultés autour de la notion de périmètre. Les difficultés en question peuvent aussi être en partie expliquées par le manque de maturité du champ disciplinaire, un manque qui pourra être lentement comblé par la pratique et publications d'études traitant le sujet rigoureusement. Les enjeux suivants sont à surveiller pour la question du périmètre :

- **Périmètre temporel** - Le standard pour les services TIC de l'ITU recommande, au travers de l'UF "usage annuel du service", le périmètre d'un an. Ce qu'il est cri-

3. L'on m'a déjà déclaré que le stockage de données côté datacenters d'un service n'avait pas d'impact, car le stockage était un S3 d'AWS dit *serverless*, sans serveur donc sans impact. En réalité, bien sûr un stockage probablement distribué mais dans une configuration opaque, peut-être même pour le fournisseur de stockage.

tique de retenir est que l'intensité d'usage d'un service numérique, son architecture ainsi que les allocations effectuées peuvent varier au cours d'une journée ou d'une semaine. Le choix d'un périmètre temporel d'un an est une recommandation générale plutôt sûre. Néanmoins, le critère principal d'établissement de ce périmètre doit rester la représentativité de la période observée par rapport à ce qui est considéré comme l'activité moyenne du service. Est-ce que la période permet d'établir des moyennes représentatives du fonctionnement du service ? Cet enjeu est particulièrement présent par exemple pour un service interactif à l'échelle d'un seul pays, où la population d'utilisateur dort et travail au même moment. À l'échelle internationale, le week-end ou l'été influencent l'activité, dans des cas plus extrêmes, une année elle-même peut être contestée comme période, comme ce fut le cas des années marquées par la pandémie de COVID-19.

- **Catégories d'impacts environnementaux** - Si la prise en compte d'un large nombre de catégories d'impacts environnementaux est considérée comme une caractéristique de l'ACV [131], il reste possible de nuancer. Il est considéré comme légitime, si l'état de l'art le justifie, d'exclure de l'étude des catégories d'impacts en raison de leurs faibles importances ou de leurs fortes corrélations avec d'autres catégories restant dans le périmètre. Il reste difficile d'accéder à des données ICV pour le numérique couvrant un large panel d'impacts comme la norme PEF et la qualité de ces bases reste critiquée [214]. De plus, une étude ou une méthodologie intégrant des données uniquement en format PCF⁴ peut très bien être, si elle est ouverte, recalculée avec des données incluant d'autres catégories d'impacts. Malgré ces nuances à noter, et le fait qu'il existe déjà des grandes tendances dans les catégories d'impacts du numérique, il devrait rester recommandé, dans le cadre des services numériques, d'étudier, si possible, un maximum de catégories d'impacts environnementaux, afin de continuer de renforcer les connaissances sur la répartition des impacts du numérique par catégories.
- **Périmètre du service numérique** - La structure en 3 parties décrite par la figure 3.1 est communément admise pour les services numériques, à l'exception de services à l'architecture distribuée [119]. Les avantages de cette approche ont déjà été discutés précédemment, les enjeux de périmètres spécifiques seront détaillés pour chaque partie :
 - ◇ **Terminaux** - Comme vu précédemment, la question de la surconsommation d'électricité du terminal, due au service, est à traiter. Il est plus important de connaître le type d'équipement (smartphone, ordinateur portable, ordinateur fixe, etc.) que le modèle précis. À noter, la question des accessoires, ou périphériques de certains terminaux peut avoir son importance. Par exemple, si l'ensemble des utilisateurs utilisent un poste fixe avec écran externe supplémentaire, ou bien si le service induit l'usage d'un périphérique spécifique comme une webcam [120].
 - ◇ **Réseaux** - Le type de réseaux et le mix électrique de la région d'utilisation sont les informations les plus critiques pour cette partie. Une attention particulière doit être faite sur la question de l'inclusion ou non des équipements réseaux personnels comme les box internet. De nombreuses études incluent les box

4. Product Carbon Footprint, format d'empreinte carbone dans le même périmètre que la PEF.

internet, mais pas toutes. Il est nécessaire de bien identifier les limites du facteur d'impact généralement utilisé pour calculer le coût de la partie réseaux.

- ◇ **Datacenters** - La principale difficulté concernant le périmètre de la partie datacenters est l'inclusion, et l'allocation, de tous les équipements non-informatiques en dehors du serveur. Même dans le cadre d'un serveur dédié, cette question de périmètre est pointue. Comment attribuer l'empreinte de l'armoire qui contient le serveur, des ventilateurs qui refroidissent le couloir ou même du bâtiment lui-même ? Le PUE permet de prendre en compte dans l'empreinte d'usage un périmètre assez large. L'inclusion de l'empreinte d'usage reste beaucoup plus périlleuse et devrait être claire.
- **Ressources humaines** - La création et l'entretien d'un service numérique demandent souvent l'activité de personnels utilisant des équipements numériques pouvant être évalués [122]. Ces coûts sont en effet induits par le service, mais incluent des activités, comme les trajets domicile-travail ou le chauffage du lieu de travail, qui peuvent être jugées comme hors du périmètre d'un service numérique. Le standard ITU pour les services TIC [152] affirme inclure le coût de développement du logiciel comme un équipement, donc incluant non seulement les déplacements domicile-travail, mais aussi, si on observe le périmètre dédié aux équipements, tout l'environnement de travail des personnes participant au développement⁵. Il est nécessaire de constater qu'inclure l'ensemble de ces coûts est périlleux, d'autant plus que ces coûts dilués dans le service peuvent être inférieurs à un critère de coupure⁶. En somme, l'inclusion ou non de ce coût devrait être spécifiée, et le cas échéant, précisée dans quelle mesure.
- **Unité fonctionnelle** - L'aboutissement de la définition du périmètre et des objectifs, l'UF doit formuler de manière exhaustive l'action évaluée. La norme ITU pour les services TIC recommande "usage annuel du service". Le but doit rester la cohérence entre l'UF et les périmètres et objectifs définis. Les unités fonctionnelles peuvent inclure des actes simples composant le service mais il faut bien identifier les limites de telles approches⁷ comme traité dans le chapitre 3. Enfin, la définition de l'UF d'un service numérique peut être enrichie par la définition d'indicateurs quantitatifs établissant la qualité de service attendue pour qu'une UF soit considérée comme "complétée".

4.1.2 Données et outils

Cette section est pensée comme un panorama pratique des différents outils disponibles pour mener à bien une ACV de services numériques. L'absence d'outils spécifiques à l'ACV comme les logiciels dédiés⁸ peut être notée. Ces outils n'ont pas été utilisés durant la thèse, pour des raisons comme : le coût d'apprentissage lourd, le manque de souplesse sur

5. Le chauffage, l'air conditionné, l'éclairage, les portes automatiques, les ascenseurs, etc.

6. Par exemple dans [119] le coût du développement représente au maximum, selon la catégorie d'impact choisie, 2,25% de l'empreinte du service, mais cette étude ne prend en compte que la consommation des ordinateurs portables des développeurs.

7. Est-ce qu'il est correct de modéliser l'empreinte du service uniquement comme la somme d'actes unitaires ?

8. Des logiciels comme OpenLCA, SimaPro et Ecosphera pour ne citer que les plus connus.

le choix des bases de données ou encore la difficulté d'effectuer un travail collaboratif avec ces outils. De nombreuses critiques précédemment mentionnées se recoupent avec cette analyse [212]. Certains projets semblent apporter de nouvelles approches⁹, plus proches du développement logiciel, mais ces outils sont encore balbutiants.

Bases de données d'inventaires de cycle de vie

Voici, ci-dessous, une liste des principales bases de données ICV utilisables dans le cadre d'ACV de services numériques :

- **Les bases ICV généralistes** comme Sphera-LCA, EIME-CODDE ou Ecoinvent sont largement répandues dans le milieu ACV, souvent directement intégrées dans des logiciels ACV. Leur principal avantage est leur popularité dans le milieu professionnel ACV, la conformité des données ICV aux standards de qualité attendus et la polyvalence de ces bases, notamment lorsqu'il s'agit d'intégrer des éléments non numériques. Les défauts sont le coût d'accès à ces outils qui peut être prohibitif, financièrement, mais aussi d'un point de vue formation nécessaire à leur utilisation. Enfin, comme mentionné précédemment, ces bases ICV peuvent être sujettes à des variations importantes sur des catégories d'impacts critiques pour le numérique comme l'extraction de métaux et minéraux ou l'usage d'eau [58, 214].
- **La base Empreinte de l'ADEME** est la base de données publique et générique de facteurs d'émission et de jeux de données d'inventaire résultant de la base Carbone et de la base Impact. La première étant la base de données pour le bilan carbone des organisations, la deuxième est une base de données ICV conforme aux exigences PEF [137] et OEF [138]. Cette dernière a l'avantage d'être gratuite et divisée en section permettant de facilement trouver les données adéquates. Les bases de données sont néanmoins difficiles à trouver sur un site internet aux dysfonctionnements fréquents. Néanmoins, une fois obtenues, ces bases de données ICV offrent un bon compromis, accessible, entre les bases classiques et les bases spécialisées comme NégaOctet, tout en offrant un niveau de qualité de données conséquent. À noter que les données ICV concernant le numérique sont majoritairement issues de la base de données ICV NégaOctet.
- **NégaOctet [175]** est la première base de données ICV spécialisée dans le numérique et conforme aux exigences PEF. La spécificité de NégaOctet est de comprendre, en plus de données ICV sur des équipements, des données sur des services numériques comme "Envoyer un mail" ou "Exécuter une requête sur un moteur de recherche" mais aussi le précieux service réseaux "Transport d'1 Go de données". L'apport de NégaOctet à l'ACV de services numériques est indéniable. Cependant, outre son coût d'accès qui peut être prohibitif¹⁰, le plus gros défaut de NégaOctet est l'opacité dans le calcul des impacts environnementaux, voire dans le choix des hypothèses effectuées. Ainsi, beaucoup de données ICV sont cadrées par des hypothèses très contraignantes¹¹ dont l'impact sur le résultat n'est pas quantifiable. Le fait que

9. Des projets comme "LCA as a code" ou la bibliothèque logicielle "Brightway".

10. Bien qu'il reste comparable à certaines bases de données ICV généralistes.

11. Un ordinateur fixe dans NégaOctet possède 1 CPU, 10GB de RAM, un disque HDD de 1173 GB et un disque SSD de 442 GB. Cela est bien sûr le résultat d'une moyenne mais cela produit une machine très exotique, difficile à inclure dans un modèle.

le modèle de calcul ne soit pas ouvert donne des données très dépendantes des hypothèses choisies, et cela encore plus pour les services. Pour utiliser les données de cette base, il est donc nécessaire de se conformer à des hypothèses restreintes, ce qui rend difficile toute utilisation de la base. Dernier point, comme précédemment évoqué, malheureusement, la base de données NégaOctet a discrètement annoncé la fin de sa commercialisation et de son développement pour le 30 juin 2024.

- **Le projet CEDaCI** ¹² [174] n'est pas à proprement parler une base de données ICV accessible, mais plutôt un outil difficilement utilisable dû à une démarche d'opacité encore plus poussée que NégaOctet. Le projet, centré sur la partie datacenters, implique des spécialistes reconnus du monde académique, des industriels et des spécialistes en ACV. Il a permis la réalisation de plusieurs ACV d'équipements datacenters et surtout les rares ACV sociales d'équipements numériques, le tout inscrit dans une démarche d'économie circulaire incluant la fin de vie. Malheureusement, les données brutes ne sont pas accessibles, le livrable est un outil comparant des équipements selon une notation *ad-hoc* permettant seulement des comparaisons.
- **ACT** ¹³ est un outil de modélisation de l'empreinte carbone d'équipements numériques, dans une démarche issue de la communauté architecture [100]. Bien qu'il ne s'agisse pas d'une base ICV standardisée, ni d'un outil multicritère, i.e. prenant en compte plusieurs catégories d'impacts environnementaux, sa démarche est notable. De plus, comme expliqué précédemment, bien que cela ne soit pas trivial, il est envisageable d'utiliser une modélisation mono-critère comme base pour une modélisation multicritère. Le socle méthodologique offert par ACT bien que centré sur les équipements, pourrait être un outil pertinent et très polyvalent pour intégrer finement l'empreinte d'équipement dans une ACV de services numériques.
- **BoaviztAPI** [101] et plus largement l'ensemble des projets portés par l'association Boavizta, s'ancre comme ACV dans une transparence, en filiation directe avec la philosophie de l'open source répandu en informatique. Dans le cadre du projet plus large, pour contrôler et réduire l'empreinte environnementale du numérique, le but est de considérer la mesure des impacts environnementaux comme un commun [239]. La concrétisation de cela est la mise à disposition d'outils et données ICV récoltés de diverses sources pour les proposer comme aide à la modélisation d'impacts environnementaux. La principale faiblesse des outils Boavizta est l'hétérogénéité des données en termes de sources et qualités. La principale force reste la souplesse totale dans l'utilisation des données et du modèle proposé qui permet de calculer finement l'empreinte multicritère de serveurs ou instances cloud par exemple. Enfin, par le caractère participatif et ouvert de l'outil, celui-ci permet une vraie dissémination des connaissances et compétences parmi les utilisateurs, participant à un gain de maturité globale du secteur sur le sujet de la mesure environnementale.

Pour finir, il peut être fait mention du nouveau projet français de base de données ICV publique pour le numérique : "Alt-Impact" [240]. Ce programme, finissant en 2025, affiche parmi ses différents objectifs, "L'accès à une base de données publiques pour l'évaluation des impacts environnementaux du numérique", avec potentiellement jusqu'à presque 1 mil-

12. Circular Economy for the Data Centre Industry

13. Architectural Carbon Modeling Tool

lion d'euros de budget consacré à ce point¹⁴. On peut donc espérer l'apparition prochaine d'une base ICV numérique qui saura dépasser les limites précédemment évoquées.

Outils d'observation du service numérique

En dehors de toutes considérations environnementales, la capacité à pouvoir surveiller l'activité d'un ou plusieurs services numériques, à les *monitorer*, est considérée positivement en informatique. Cependant, certains outils plus que d'autres permettent de livrer des données pertinentes sur le service dans le contexte d'une ACV. Ci-dessous se trouve donc une liste d'outils, nécessitant ou non un accès administrateur au service, permettant de recueillir des données aidant aux calculs des impacts du service :

- **Mesure directe de consommation électrique** - La mesure de la consommation électrique des équipements et infrastructures est un enjeu dont l'intérêt dépasse la seule ACV de services numériques. Outre les dispositifs expérimentaux permettant l'usage de wattmètres directement sur les équipements, de nombreux logiciels sont apparus ces dernières années permettant de "mesurer" la consommation électrique d'équipements, ou plutôt de composants, via ce qui est appelé des "wattmètres logiciels" [225, 226]. Ces outils de mesure directe sont généralement déployés pour évaluer la consommation de la partie datacenters, même si leurs usages peuvent être adaptés aux terminaux utilisateurs [241]. La mesure directe de la consommation des infrastructures réseaux est généralement impossible et se fait par modèle.
- **Modèle d'estimation de consommation électrique** - Malgré la qualité des données fournies, les wattmètres, même logiciels, demandent un accès physique ou au moins administrateur aux équipements pour une installation qui peut demander du temps. Il peut être alors envisagé d'utiliser des modèles estimant la consommation d'électricité en fonction d'autres paramètres. Ces estimations par modèles sont moins précises, mais peuvent, sur une durée de temps suffisamment longue, perdre toute différence significative avec les données issues de mesures directes. On peut observer :
 - ◇ La consommation énergétique du réseau est à estimer avec prudence dans le cas du réseau filaire où la consommation d'électricité n'est pas forcément directement corrélée avec le volume de données transférées. Certains modèles proposent donc de se baser plutôt sur la capacité du lien réseau utilisé [217]. Les modèles filaires peuvent être complétés avec d'autres approches dans le cadre des réseaux mobiles [227].
 - ◇ Pour la consommation de serveur, il est possible de se baser sur des modèles utilisant la consommation électrique nominale de l'équipement à laquelle s'ajoute une surconsommation dépendant de la charge de travail estimée [101]. La surconsommation est estimée via des benchmarks de composant déjà existants ou réalisées pour l'occasion, donnant une estimation composite de mesure directe et de modèle pour combler les limites de l'outil de mesure¹⁵.
- **Monitoring partie datacenters** - Il peut être critique, pour réaliser des allocations pertinentes ou pour alimenter les modèles précédemment mentionnés, d'obtenir

14. 398 285 € de frais fixes et 500 000 € de frais variables.

15. Voir le projet Energizta pour aller plus loin dans cette approche composite :<https://github.com/Boavizta/Energizta>.

des données d'activités du service. Cela est réalisable en utilisant des outils assez communs de gestion de serveurs et VM :

- ◇ De manière triviale, l'accès au volume de mémoire occupé par le service s'obtient facilement. Connaître le disque précis où sont stockées les données peut être pertinent, mais avec la gestion par volume logique de la mémoire (LVM) cela peut devenir indiscernable.
- ◇ Il est utile de connaître l'utilisation des ressources CPU et RAM faite par les VM ou les conteneurs docker. De nombreuses commandes dérivant souvent de la commande *top* permettent de relever ces métriques. Des agrégateurs de données comme VictoriaMetrics [242] ou dockerstats [243] peuvent être utilisés avec des tableaux de bord comme Grafana [244] pour générer des statistiques. Un équilibre est à trouver entre la précision des statistiques créées et le coût de ces outils de mesure [245].
- ◇ L'activité d'un service web peut être surveillée via le relevé de ses logs HTTP. Il est possible de connaître le nombre de visites d'un service, le système utilisé par les clients, leur localisation, le volume de données échangées ou encore le nombre de requêtes HTTP ayant échoué. Certains logiciels comme GoAccess [246] permettent de collecter et de visualiser efficacement ces logs.
- **Évaluation externe de l'activité d'un service web publique** - En l'absence d'accès aux logs HTTP d'un service web ou, même dans ce cas, en l'absence de collecte de données utilisateurs via des cookies, une solution originale peut être trouvée. Des services en ligne comme HypeStats [247] et Similarweb [248] proposent des estimations du nombre de visiteurs, nombre de pages chargées et de durée moyenne de visites d'un site web. Ces outils appartiennent à la famille des *Search engine optimizer*¹⁶ (SEO) qui se basent sur des ensembles de données de navigation publiques et privées pour estimer l'activité d'un site [249]. Les estimations des SEO peuvent être sous-estimées [249,250] mais sont un des rares moyens d'obtenir ces informations de manière externe au service.
- **Mesure flux réseaux** - Des outils connus de capture du trafic réseaux comme Wireshark permettent d'estimer la taille de flux ou contenus transférés par le service sur le réseau. De manière plus triviale, côté utilisateur, le mode développeur de navigateurs comme Firefox permet d'évaluer la taille de flux réseaux associés à un service.

4.2 Analyse de l'inventaire du cycle de vie du service numérique

L'analyse de l'inventaire du cycle de vie (AICV) est parfois considéré comme le cœur de l'étude ACV. Cette partie est en effet centrale pour répertorier l'ensemble des équipements, flux et infrastructures utilisés ainsi que l'intensité de leur utilisation. Cette section se concentrera sur la modélisation du service numérique selon ses 3 parties principales :

16. **trad.** Outils d'optimisation pour les moteurs de recherche

Terminaux utilisateurs, Réseaux et Datacenters. Les outils et jeux de données précédemment présentés serviront à alimenter le modèle décrit en données ICV ou indicateurs d'activité pour les allocations.

À des fins de simplification, l'AICV du service numérique sera considérée sous l'angle du formalisme CAPEX et OPEX, inspiré de la démarche de [84]. Le terme CAPEX, pour *capital expenditures*, désigne un coût payé une fois, une immobilisation au sens comptable. Le terme OPEX, pour *operational expenditures*, désigne lui un coût payé de manière plus ou moins récurrente, comme un consommable dont la quantité achetée dépend de l'intensité de la consommation. Déclinés dans le contexte de l'ACV de services numériques, CAPEX et OPEX désignent chacun une partie de l'empreinte environnementale. L'empreinte CAPEX représente la somme de l'empreinte des phases de fabrication, transport et fin de vie des équipements¹⁷. Cette empreinte, dans le cadre du service, représente un impact déjà existant, l'empreinte CAPEX du service sera surtout influencée par l'allocation faite. De la même manière qu'un coût CAPEX est lissé sur un ensemble d'année, on parle d'amortissement du coût des machines, l'empreinte CAPEX existe déjà et est attribuée au service suivant l'allocation choisie. L'empreinte OPEX du service représente, elle, l'empreinte d'usage du service, à savoir, presque exclusivement, la consommation d'électricité. Cette empreinte, comme les coûts OPEX, est dépendante directement de la consommation du service. L'impact environnemental du service S , $I(S)$ ou simplement I , est un vecteur colonne où chaque composante représente une catégorie d'impact tel que :

$$I(S) = I(S)_{CAPEX} + I(S)_{OPEX} \quad (4.1)$$

$$I(S) = I_Terminaux(S) + I_Réseaux(S) + I_Datacenters(S) \quad (4.2)$$

$$I(S) = I_Terminaux_{CAPEX}(S) + I_Terminaux_{OPEX}(S) + I_Réseaux_{CAPEX}(S) + I_Réseaux_{OPEX}(S) + I_Datacenters_{CAPEX}(S) + I_Datacenters_{OPEX}(S) \quad (4.3)$$

Où l'impact du service $I(S)$, qui sera par la suite abrégé en I le calcul du service S étant implicite, est divisible selon les parties CAPEX et OPEX, ainsi que selon les 3 parties du service. De manière générale, l'impact de chaque partie selon CAPEX ou OPEX peut être calculé ainsi :

$$I_{CAPEX} = \sum_e E_e \times A_e \quad (4.4)$$

$$I_{OPEX} = \sum_e C_e \times EGM_g \quad (4.5)$$

Où E_e est l'empreinte de fabrication d'un équipement e allouée à l'impact du service selon une clé d'allocation A_e , C_e est la consommation électrique de l'équipement et EGM_g ¹⁸ est l'empreinte du mix électrique utilisé selon sa localisation géographique g .

17. Bien souvent due aux données utilisées, uniquement l'empreinte de fabrication.

18. Pour *Electricity Grid Mix* la composition du mix du réseau électrique.

4.2.1 Clés d'allocation

Sur ces bases, les prochaines sous-sections proposeront différentes approches pour le calcul de l'impact de chaque partie, les clés d'allocation, ainsi que des points critiques à prendre en compte dans la modélisation. Les clés d'allocation d'équipements e pour S reposent généralement sur le principe suivant :

$$A_e = \frac{\text{Quantité de ressource utilisée par } e \text{ pour } S}{\text{Quantité totale de ressource que peut dépenser } e \text{ pendant } T} \quad (4.6)$$

Dans ce cas, la ressource peut être du temps, un certain volume de données à transférer depuis e , un pourcentage de sa charge de travail, etc. L'important est aussi de noter T la période de temps correspondant à l'empreinte E_e et à la quantité de ressource disponible pendant cette période. Dans de nombreux cas, la période T correspond à la durée de vie de l'équipement et E_e représente donc l'empreinte CAPEX pour le cycle de vie complet. La période T peut correspondre à une fraction de la durée de vie totale de l'équipement du moment que E_e est divisé de même. Ce détail permet une certaine souplesse dans le calcul de la quantité de ressources totales à allouer qui pourra être utile, tout en s'accordant facilement avec l'UF standard des services TIC "Usage annuel du service".

La ressource choisie pour calculer la clé d'allocation d'un équipement devrait avoir les deux qualités suivantes :

1. Être un indicateur **représentatif de l'usage** de l'équipement e
2. **Éviter la sur-allocation**, i.e. que la somme de toutes les clés d'allocation de e dépasse le total de 1

Si la première qualité peut être difficile à juger, de nombreux équipements numériques étant multi-usages, il est dur de trouver un unique indicateur de leur usage, la deuxième qualité est plus facilement contrôlable.

4.2.2 Terminaux utilisateurs

Contexte

La catégorie des terminaux inclut les équipements utilisés par les utilisateurs finaux des services numériques. Cela inclut : les téléviseurs, consoles de jeux, les téléphones fixes, les imprimantes, les écrans, les objets connectés comme les montres connectées ou autres assistants domestiques, etc. Surtout, les terminaux sont souvent représentés dans les ACV de services numériques par les ordinateurs (portable ou fixe), les smartphones et les tablettes, qui, contrairement aux équipements précédemment cités, sont caractérisés par une polyvalence d'usage qui va rendre leurs allocations difficiles. À noter que ces terminaux polyvalents font appel souvent à des périphériques (clavier, écran, webcam, souris, etc.) à prendre en compte parfois avec des allocations différentes. Dans tous les cas, l'identification du type de terminal, à minima mobile ou fixe, est importante pour estimer l'empreinte. Cette identification permettra de partager l'ensemble des terminaux entre une partie utilisant des terminaux mobiles (plutôt des smartphones sauf cas particulier) et une autre utilisant des terminaux fixes (plutôt des ordinateurs portables sauf cas particulier).

CAPEX

Les clés d'allocation les plus répandues pour les terminaux sont les allocations **temporelles**, au temps d'usage, et les allocations au volume de données échangées, dites allocations **au Giga**. Il est possible de les décrire tel que :

$$A_Terminaux_e = \frac{\text{Temps d'usage de l'équipement e pour le service}}{\text{Temps total d'usage de l'équipement e}} \quad (4.7)$$

ou

$$A_Terminaux_e = \frac{\text{Données transférées de l'équipement e pour le service}}{\text{Données transférées depuis l'équipement e pour toute sa vie}} \quad (4.8)$$

Le calcul de ces deux clés requiert :

1. Une mesure du temps d'usage ou du volume de données transférées du terminal dans le but de satisfaire l'UF.
2. Une connaissance, généralement une hypothèse, sur le temps d'usage total¹⁹ ou le volume total de données transférées par l'équipement.

Si le premier point est mesurable, considérant les outils précédemment mentionnés, le deuxième point demande généralement une hypothèse sur l'avenir de l'équipement. Combien d'années d'usage ? Quelle intensité d'usage, en temps ou données transférées, en moyenne ? Des enquêtes statistiques, parfois issues des bases de données ICV, peuvent être utilisées, mais les hypothèses choisies seront souvent à questionner en analyse de sensibilité.

Par rapport aux critères attendus pour les clés d'allocation, pour de nombreux terminaux, ces clés sont satisfaisantes. Le problème vient dans le cas, majoritaire dans le cadre des services numériques, de l'utilisation de smartphone, ordinateurs et tablettes. Si le temps peut bien représenter l'usage d'un tel équipement, cela suppose que pendant le temps d'usage mesuré, le terminal est uniquement consacré à la réalisation de l'UF ce qui peut être difficile à affirmer. Une part significative d'utilisateurs utilise un service de streaming pendant d'autres tâches avec le même terminal, ou encore reste en ligne sur un service de messagerie, les exemples sont nombreux. D'un autre côté, si le transfert de données peut difficilement être utilisé par deux services en même temps, entraînant une sur-allocation, il n'est pas évident de résumer un terminal comme un smartphone ou un ordinateur à sa capacité à échanger des données. Même avec des applications presque toutes devenues des services connectés, est-ce que le volume de données transférées est représentatif de l'usage du terminal fait par le service ? Dans une situation où un employé travaillerait toute la journée sur un logiciel professionnel en écoutant de la musique, il est probable qu'une allocation à la donnée alloue une plus grande part de l'empreinte au service de streaming qu'au logiciel de travail. Est-ce juste ? La question reste ouverte.

D'autres types d'allocations pourraient exister, par exemple une allocation en fonction de la surconsommation d'électricité, i.e. la consommation totale du terminal à laquelle on soustrait la consommation *idle*. En règle générale, l'allocation temporelle est recommandable lorsque le terminal offre surtout un "temps d'écran disponible". C'est le cas dans un service qui requiert l'attention totale ou du moins principale de l'utilisateur, limitant les risques d'usages parallèles et donc de sur-allocation.

19. Rappelons que le temps d'usage total n'est pas équivalent à la durée de vie de l'équipement. Il s'agit de la durée de vie à laquelle on soustrait les périodes où l'équipement est inutilisé.

OPEX

Dans le cadre de l'impact opérationnel des terminaux décrit comme :

$$I_Terminaux_{OPEX} = \sum_e C_e \times EGM_g \quad (4.9)$$

L'enjeu de prime abord est de calculer C_e pour chaque équipement e . L'obtention de cette donnée se heurte rapidement à des problèmes de moyens. Premièrement, mesurer cette consommation sur un grand ensemble de terminaux clients demanderait des moyens importants. Deuxièmement, l'obtention de C_e pour un équipement e pendant qu'il traite un UF du service expose, comme vu précédemment, à une problématique de sur-allocation. En effet, la consommation d'un terminal peut assurer plusieurs services en même temps. Dans ce cas, C_e devient non plus la consommation électrique de l'équipement e mais la surconsommation de l'équipement directement imputable au service $C_{e,S}$ différent de $C_{e,idle}$. Par conséquent, la consommation du terminal est divisée en deux parties tel que :

$$I_e_{OPEX} = C_{e,S} \times EGM_g + C_{e,idle} \times EGM_g \times A_{e,S} \quad (4.10)$$

Ainsi, la consommation $C_{e,idle}$ est traitée comme un coût CAPEX devant être partagé avec les autres usages du service. Cette consommation relève plus de l'usage général de l'équipement que de l'utilisation directe du service sur celui-ci. Reste la problématique de $C_{e,S}$, difficile à mesurer ; elle peut être établie par modèle et benchmark avant d'être généralisée à l'ensemble des terminaux.

Deux situations courantes, et assez opposées en termes de précisions des données, sont d'une part l'absence totale de données sur $C_{e,S}$, d'autre part une donnée fiable sur $C_{e,S}$ et $C_{e,idle}$. Dans la première situation, il est possible de simplifier l'empreinte des terminaux de la manière suivante :

$$I_Terminaux = \sum_e A_e \times (E_e + C_{e,moy} \times EGM_g) \quad (4.11)$$

Dans cette situation, s'il y a accès à $C_{e,moy}$, la consommation moyenne de l'équipement e , pour un T en accord avec E_e , il est possible de calculer l'empreinte CAPEX et OPEX de l'équipement en une fois. Dans la deuxième situation, l'empreinte OPEX du terminal est calculée normalement mais la clé d'allocation CAPEX peut être calculée comme suit :

$$A_Terminaux_e = \frac{C_{e,S}}{\sum_Z C_{e,Z}} \quad (4.12)$$

Où S est un service assuré par le terminal e qui assure un ensemble de services Z . Dans cette situation, on considère que la surconsommation d'électricité est un bon indicateur de l'usage de l'équipement. La difficulté à connaître aussi finement l'activité de la partie terminaux, surtout quand le nombre d'équipements est élevé, fait que cette méthode sera plutôt utilisée pour un nombre d'équipements faible ou pour un ensemble d'équipement très homogènes.

4.2.3 Réseaux

En dehors des ACV de services numériques considérant uniquement la consommation d'électricité du service, il est rare de calculer séparément l'empreinte OPEX et CAPEX

de la partie réseaux. Cette section sera donc divisée entre, un cas général de calcul de l’empreinte réseaux et les recommandations pour les cas particuliers.

Cas général

La majorité des cas peut se résoudre, avec la bonne donnée ICV, par l’équation suivante :

$$I_Réseaux = V \times (E_{mobile} \times A_Réseaux_{mobile} + E_{fixe} \times A_Réseaux_{fixe}) \quad (4.13)$$

Où V est le volume de données transférées, généralement exprimé en Giga, E l’empreinte du transport selon le type de réseau utilisé, généralement pour un Giga transporté, alloué selon la répartition du trafic entre réseaux fixes et mobiles tel que la somme de $A_Réseaux_{mobile}$ et $A_Réseaux_{fixe}$ est égale à 1. Si le jeu de données ICV le permet en différenciant l’empreinte CAPEX et OPEX, le calcul s’écrit :

$$\begin{aligned} I_Réseaux = V \times \\ (A_Réseaux_{mobile} \times (E_CAPEX_{mobile} + C_{mobile} \times EMG_g) \\ + A_Réseaux_{fixe} \times (E_CAPEX_{fixe} + C_{fixe} \times EMG_g)) \end{aligned} \quad (4.14)$$

Où C est la consommation d’électricité pour un Giga du volume V transporté et g la zone géographique où a lieu, au moins majoritairement, le transport de données. Cette approche a l’avantage d’être simple d’utilisation et de pouvoir inclure l’influence du mix énergétique local. La première limite est bien sûr le cas des communications internationales, qui dans une certaine mesure peut être résolue par l’usage d’un mix électrique moyen entre les pays concernés, mais n’a pas de solution globalement satisfaisante. Deuxième limite, cette approche établit dans son modèle une relation entre données transportées et consommation d’électricité. Dans le cadre de cette approche, il s’agit plus d’une allocation sur la consommation moyenne du réseau, estimée de la même manière que l’empreinte CAPEX du transport d’un Giga. Toutefois, ces critiques restent et le point suivant aborde brièvement les approches concurrentes.

Cas particulier

Le sujet du détail du coût réseau a été évoqué plusieurs fois dans ce manuscrit. Les applications des chapitres 5 et 6 resteront sur le cas général. Pour aller plus loin, voici quelques démarches plus exhaustives.

La première démarche consiste à intégrer chaque équipement réseaux. De manière analogue à [120] et [102], cette démarche inclut tous les équipements et leurs empreintes CAPEX comme OPEX avec une allocation au volume de données transportées. Dans cette situation, où l’usage unique des équipements réseaux est le transport de données et où généralement le volume total de données transporté pendant T est connu, cette allocation semble adéquate. Cette allocation est d’ailleurs adéquate pour inclure des équipements réseaux spécifiques au service, comme des équipements de sécurité tel que vu dans [119].

Enfin, plusieurs fois cité pour critiquer les power-model classiques, [217] propose, en plus de modéliser chaque sous-partie du réseau : le réseau de cœur, le réseau d’accès et

l'équipement sur site. Dans cette approche, ce qui compte n'est pas le volume transporté, mais la capacité des liaisons réseaux mobilisées. Dans une situation où un certain degré de liberté sur l'infrastructure réseau existe, ou du moins peut être considéré, il peut être intéressant d'implémenter dans une ACV de services numériques ce type de modèle pour la partie réseau fixe, complété avec celui de [227] pour la partie réseaux mobiles.

4.2.4 Datacenters

Contexte

La situation la plus simple pour estimer l'impact de la partie datacenters est le cas d'un serveur, sur site²⁰, dédié entièrement au service. Dans ce cas, on peut, similairement aux terminaux, réaliser sur le serveur une allocation temporelle. Toutefois, si la partie serveur est dénommée datacenters, c'est bien que cette fonction reste la plupart du temps externalisée chez un ou des fournisseurs de solutions d'hébergements, dans un contexte Cloud de virtualisation, voire de conteneurisation. L'environnement datacenters implique donc plusieurs conséquences, qui l'éloigne du cas basique où le service est hébergé grâce à un unique équipement serveur dédié à cette tâche :

1. Une concentration des serveurs au même endroit nécessitant l'usage d'équipements spécifiques. Ces équipements sont ceux qui assurent la gestion du refroidissement des équipements, de l'accès au réseau, l'alimentation en électricité, la sécurité ou encore le rangement des serveurs en armoires.
2. Une spécialisation des fonctions serveurs qui fait que de nombreux services ne sont pas simplement sur un unique serveur dans un datacenter. De nombreuses architectures de services comprennent l'utilisation de plusieurs équipements; serveurs spécialisés dans les fonctions de stockage, gestion de site web, calcul haute performance, entraînement de modèles d'IA ou même inférence à partir de modèles d'IA générative.
3. Une décomposition des ressources demandée par le biais de la virtualisation. Même si les réservations de serveurs dédiés, dite *bare-metal*, restent significatives, la majorité de l'usage de ressources datacenters se fait au travers de la réservation d'une ou plusieurs machines virtuelles (VM). Ces VM possèdent une partie des ressources (RAM, stockage et calcul) de machines physiques. Cette virtualisation peut permettre à l'architecture du service de varier rapidement en taille et nombre de VM. Ces VM peuvent être elles-mêmes décomposées en conteneurs qui peuvent se partager dynamiquement les ressources disponibles.

On considère donc souvent pour cette partie une série d'inclusions. La partie datacenters contient des serveurs, qui contiennent des VM, qui peuvent contenir des conteneurs sur lesquels sont hébergés le service. Comme il sera détaillé dans les deux parties suivantes, le calcul de l'empreinte de cette structure, à la fois mouvante et arborescente, peut être un challenge certain, surtout en considérant l'opacité de fonctionnement de nombreux datacenters ainsi que la difficulté de dimensionnement de l'hébergement précédemment évoquée.

20. Comprendre un serveur localisé dans l'organisation gérant le service, sans aménagement particulier pour celui-ci autre qu'un accès à internet et à de l'électricité.

CAPEX

Estimer l’empreinte CAPEX de la partie datacenters commence par identifier les ressources ou les serveurs dédiés (serveurs et VM) qui pourront être alloués temporellement. Il est ensuite envisageable, s’ils ont pu être inclus dans le périmètre de l’étude, d’inclure les équipements réseaux spécifiques du datacenter qui peuvent être alloués au volume de données transférées pour le service. De même, si les informations et données ICV sont disponibles, l’empreinte CAPEX des équipements de support du datacenter peut être allouée au service. Elle est modélisable comme un coût supplémentaire ajouté à l’empreinte CAPEX de chaque serveur. Ce coût est obtenu en divisant l’empreinte CAPEX des équipements supports par le nombre de serveurs qui bénéficie de ces équipements pendant la durée d’observation T , une allocation au serveur en somme.

Enfin, il est à prendre en compte dans un contexte Cloud, où l’hébergement est fourni par un tiers garantissant un accès rapide, voir automatique, à un plus grand nombre de ressources en vue de répondre à une activité croissante soudaine du service, le coût de cette disponibilité. En effet, dans ce contexte, la disponibilité permanente d’un grand nombre de serveurs toujours prêt à répondre rapidement à une nouvelle demande suppose que le datacenter ne loue pas en permanence toutes ses ressources. Ainsi, louer une ressource à un datacenter implique hériter de l’empreinte de cette ressource, mais aussi d’une part des autres ressources non utilisées. Les datacenters représentent parfois cela par la métrique du taux d’utilisation actif des serveurs, i.e. le pourcentage de temps pendant lequel le serveur est réellement utilisé dans sa vie. Les chiffres sont rares [251] mais cruciaux car ils influencent grandement les clés d’allocation en réduisant le total de ressources que l’équipement va partager dans son cycle de vie. Il est nécessaire d’appliquer ce taux d’utilisation actif tua , compris entre 0 et 1, de la manière suivante dans le calcul des clés d’allocation, par exemple pour un serveur e :

$$A_Serveur_e = \frac{\text{Temps d'usage de l'équipement } e \text{ pour le service}}{\text{Temps total d'usage de l'équipement } e \times tua} \quad (4.15)$$

L’enjeu principal de l’empreinte de la partie datacenters est souvent l’empreinte de la ou des VM utilisées par le service. Il est nécessaire de connaître les ressources allouées à la VM, en termes de nombre de vCPU, RAM et volume de stockage pour ensuite allouer une partie de l’empreinte du serveur à la VM. Dans l’idéal, le serveur est connu, sinon une hypothèse sur le type de serveur utilisé en moyenne par le fournisseur de la solution d’hébergement peut être effectué. L’important est que dans la modélisation, si on a connaissance de l’ensemble de l’activité du serveur, la somme des $A_{VM,e}$ du serveur e doit être de 1.

Plus généralement, s’il reste des ressources serveur non dédiées à additionner dans l’empreinte CAPEX, il s’agit alors de trouver l’optimum en termes de disponibilité et qualité des données, qu’il s’agisse des données ICV comme des données d’activité. Le but est d’avoir un compromis entre une valeur E_e la plus fiable possible et une clé d’allocation $A_Datacenters_e$ calculée sur la base de données d’activité fiables. On peut ainsi estimer au mieux l’usage des ressources utilisées par le service au sein des ressources disponibles.

Un dernier bon exemple d’une part de l’empreinte CAPEX datacenters difficile à allouer est la partie stockage. Hors équipement dédié, on peut basiquement faire une allocation à l’espace de stockage utilisé pendant la période T . Si le service occupe, pendant

la période T , 50% de l'espace disque du serveur de stockage, il récupère 50% de l'empreinte de l'équipement. Mais il serait aussi possible d'allouer selon le trafic lecture et écriture des disques représentés par le service par rapport au trafic sur la période T . Une donnée d'usage qu'il est raisonnablement possible de récupérer. Mais donc quelle clé d'allocation choisir, le volume occupé ou le trafic lecture et écriture ? La question à se poser dans cette situation pour choisir entre ces deux clés serait "Lequel de ces deux indicateurs détermine vraiment le choix de l'équipement ?". Où plutôt, lequel de ces deux indicateurs influence le plus le dimensionnement de l'équipement dans ce contexte ? Il est bien sûr possible de choisir, comme compromis, une moyenne de ces deux indicateurs, au prix d'une surcouche de complexité, mais ce choix doit être fait consciencieusement en phase avec le contexte de l'étude.

OPEX

L'enjeu principal de l'estimation de l'empreinte OPEX de la partie datacenters du service est de faire correspondre l'ensemble des ressources informatiques identifiées précédemment à une consommation électrique cohérente avec les ressources mobilisées pour le service et son activité. Différents modèles et outils permettent une plus ou moins grande finesse dans l'estimation de la consommation du service. Une finesse qui est à considérer comme nécessaire ou non suivant les objectifs de l'étude.

L'important pour cette partie reste d'inclure la consommation *idle* des ressources informatiques mobilisées, qu'il s'agisse de serveurs, de VM ou autres équipements spécialisés. Cette inclusion peut être facile dans le cadre d'équipement dédié mais devient plus difficile, par exemple, pour une VM. Dans cette situation, il peut être difficile d'estimer l'électricité nécessaire au serveur en période d'inactivité totale ou d'inactivité de la VM. Le service hébergé sur la VM a besoin que le serveur reste allumé en *idle* mais n'est pas forcément responsable de toute la consommation *idle* du serveur. Dans cette situation, une allocation sur la consommation *idle* du serveur est nécessaire. Comme expliqué précédemment, les critères d'allocations sont à établir au mieux selon les données disponibles. La solution la plus simple peut être d'appliquer à la consommation *idle* du serveur la même clé d'allocation que pour l'empreinte CAPEX du serveur.

Autre point à inclure obligatoirement dans le calcul de l'empreinte OPEX datacenters est le *PUE*. Il peut être connu ou pris par hypothèse, mais doit être inclus comme facteur augmentant les consommations mesurées de chaque serveur e , tel que :

$$I_Datacenters_{OPEX} = \sum_e C_e \times EGM_g \times PUE \quad (4.16)$$

4.3 Évaluation des impacts

La phase d'analyse des impacts des cycles de vie (AICV)²¹ est une partie critique du processus ACV classique [252] bien que pour plusieurs raisons, sa complexité est réduite dans le cadre de l'ACV de service numérique. En effet, l'ACV de services numériques se repose presque exclusivement sur des données ICV de produits ou services déjà évalués. Le premier enjeu de la phase d'AICV, qui est de sélectionner les catégories d'impacts à

21. En anglais : *Life cycle impact assessment*

évaluer et les méthodes d'AICV à utiliser, est donc préliminairement décidé par les ACV ayant produit les résultats ICV utilisés. Par exemple, si les données ICV utilisées sont en format PEF [137], le choix naturel est de prendre en compte pour les impacts du service l'ensemble des catégories d'impacts environnementaux incluses dans le format PEF.

Cette phase a moins d'enjeu dans le cadre des services numériques, due au faible nombre de flux élémentaires inclus dans les ACV. Toutefois, il ne faut pas négliger les problématiques concernant les flux d'électricité et d'eau dans ce contexte. Comme évoqué à la section 2.1.2 l'AICV dans le cas de l'électricité comprend un débat récurrent entre une attribution des impacts basé sur le mix du réseau électrique dans la région de consommation de l'électricité, dit *location-based*, ou basée sur les garanties d'origine de l'électricité fournie par le distributeur d'électricité utilisée²², dit *market-based*. Ce débat actuel existe dans d'autres secteurs d'application de l'ACV [81] mais est critique pour les services numériques où les impacts de la consommation d'électricité peuvent représenter l'intégralité de l'empreinte OPEX. Comme l'expose [81] le choix du *market-based* peut être pertinent dans certains contextes, mais doit être effectué prudemment. Dans le cadre des études de cas de cette thèse, le choix a toujours été celui du *location-based* qui semble par défaut plus robuste méthodologiquement et ne s'est jamais heurté au contexte des études de cas. Pour ce qui est de l'eau, le modèle utilisé majoritairement en AICV, recommandé par le standard PEF, est celui du volume pondéré par la rareté [253], même s'il existe toujours des débats avec la méthode concurrente en volume absolu d'eau prélevée [254]. Dans le cas de l'ACV de services numériques, les impacts en eau sont, pour le moment, issus de l'empreinte CAPEX et donc la phase d'AICV suit la méthodologie des données ICV utilisées. Néanmoins, une étude ACV prenant en compte la consommation d'eau, par exemple d'un datacenter, comme partie de l'empreinte OPEX du service devra se poser la question de l'eau pour son AICV.

Suite à ces considérations, l'AICV se clôt par 3 étapes dont seule la première est obligatoire :

1. La caractérisation des résultats
2. La normalisation des résultats
3. La pondération des résultats

La caractérisation des résultats correspond aux résultats bruts de l'ACV où chaque impact est exprimé dans sa propre unité de mesure. Cela correspond au vecteur colonne $I(S)$, une série de valeurs qui ne peuvent pas être comparées. La normalisation correspond à diviser par un vecteur colonne de même taille, où chaque composante représentant une catégorie d'impact est bien divisée par un résultat représentant la même catégorie. Par exemple, le cas de normalisation le plus fréquent est celui par "impact moyen d'un habitant du monde". Cet impact caractérisé représente l'ensemble des impacts environnementaux d'une personne moyenne sur un an. Le choix du facteur de normalisation peut faire l'objet d'une sélection particulière au regard du contexte, le plus important est qu'une fois normés, les impacts des différentes catégories peuvent être comparés, car, pour suivre l'exemple, tous exprimés en "impact moyen d'un habitant du monde". La pondération consiste ensuite à pondérer ou agréger les catégories selon des critères arbitraires, comme par exemple ceux recommandés par le JRC pour la PEF [255]. Ces critères peuvent servir

22. Par exemple : Un distributeur d'électricité peut s'engager contractuellement à ne délivrer à son client que de l'électricité provenant de sources d'énergies dites "vertes", "décarbonés" ou "renouvelables"

à représenter, à partir des résultats ACV, le risque, un coût monétaire, la distance par rapport à des objectifs fixés ou encore l'enjeu des limites planétaires. Les résultats pondérés sont des interprétations, calculées à des fins de communications pour des publics spécifiques.

Les étapes de normalisation et pondération sont encore rares en ACV de services numériques, encore une fois à cause du manque de maturité du champ. Ces étapes optionnelles servent à mieux communiquer sur des résultats, mais doivent être maniées avec transparence et prudence afin de ne pas être mésinterprétées. Il reste que la normalisation propose un bon compromis entre facilité d'explication et apport d'information en permettant d'exposer les ordres de grandeurs des impacts les uns par rapport aux autres. Il est enfin à noter, certains efforts pour intégrer les limites planétaires dans les ACV de TIC [256].

4.4 Analyse des résultats

L'analyse des résultats est le moment d'observer la concordance entre résultats, méthodes et objectifs au sein de l'étude. Il s'agit d'identifier les limites de l'étude, de les mettre en valeur via des analyses de sensibilités. Bien souvent, les hypothèses les plus susceptibles de créer des analyses de sensibilités pertinentes sont celle ayant été utilisées pour calculer les clés d'allocation. On parle ici des hypothèses influençant le dénominateur de l'équation (4.6) comme la durée de vie ou la somme totale de travail ; cela inclut aussi des facteurs comme le taux d'utilisation actif des serveurs comme vu dans l'équation (4.15).

Il est aussi intéressant d'être particulièrement attentif aux limites répandues dans l'ACV de services numériques, comme celle décrites dans le chapitre 3. En particulier, les projections sur le futur du service ou la variation de l'activité de celui-ci doit prendre en compte le dimensionnement des infrastructures.

Enfin, il peut être intéressant de terminer l'étude du service par une ouverture sur les potentiels impacts non pris en compte par l'approche attributionnelle de celle-ci. On parlera ici des éventuels impacts de 2nd et 3e ordre. Il est intéressant pour cela de se référer à des frameworks pré-existants et plus larges comme celui de Hilty [32], le SCEF [93] ou le plus exhaustif [212]. Ces méthodes d'évaluation de la soutenabilité de produits ou services numériques permettent, en fin d'étude ACV sur un service numérique, d'élargir les perspectives et de poser la question "Est-ce que les problèmes environnementaux prédominants sont ceux évalués dans l'ACV?". Même si ces questions n'ont pas de réponses directes, elles peuvent servir de piste à un travail futur des personnes ayant pratiqué l'ACV ou d'autres souhaitant reprendre ses résultats pour une étude plus large.

4.5 Conclusion et limites

Dans ce chapitre, nous avons présenté un cadre méthodologique général afin d'aider à la réalisation d'ACV attributionnelles de services numériques. Cette contribution tient à aider le lecteur par la synthèse faite des outils et approches disponibles afin d'enrichir la recherche en informatique environnementalement soutenable, par la pratique de l'ACV. Bien évidemment, la méthodologie exposée dans le présent chapitre a ses limites propres,

indépendamment des applications qui en seront faites et nécessiteront des adaptations au sujet d'étude.

Tout d'abord, la méthode hérite des limitations propres à l'ACV attributionnelle [236]. Ce type d'approche ACV est censé en effet être centré sur des situations comme la recherche de pistes d'amélioration dans l'empreinte du produit ou l'estimation, à des fins de communications, de l'empreinte du produit. Ce type de méthode se concentre sur une petite échelle. Plusieurs travaux comme [257] mettent en avant qu'extrapoler les résultats d'ACV attributionnelle peut amener à des prises de décisions biaisées par le cadre initial. En effet, l'approche attributionnelle considère le système étudié, les infrastructures et terminaux disponibles dans nos cas, comme statiques, préexistants au service évalué. Ce cadre, à l'inverse de l'ACV conséquentielle, n'est pas fait pour la prise de décision à grande échelle, comme le démontre [257], même une ACV comparative ne suffit pas à appuyer définitivement l'intérêt environnemental d'un produit ou service par rapport à un autre. Le débat entre ACV attributionnelle et conséquentielle est vaste et toujours actif [258, 259]. Notre méthodologie en effet partage les mêmes risques identifiés au chapitre 3 au sujet des scénarios de croissance de l'activité ou de l'infrastructure. Comme toute ACV attributionnelle, les analyses proposant des modifications par rapport à ce qui a été observé doivent être prises avec prudence et ne doivent pas seules soutenir un choix d'orientation pour le service. Comme le souligne [258], cela ne veut pas dire pour autant que les résultats d'ACV attributionnelle sont dépourvus d'intérêt. Au contraire, cette focalisation sur la responsabilité de l'empreinte du produit permet une approche fine sur la localisation des impacts environnementaux dans le cycle de vie ou dans les différentes parties du service. Cette enquête sur le fonctionnement du service que propose la méthode dans le but de rechercher des améliorations applicables afin de réduire l'empreinte d'un service ou d'un système d'information hébergeant un ou des services, correspond à l'ambition de proposer cette méthode ACV comme outil de recherche en informatique soutenable.

Une autre limite de cette méthode est liée à la simplification qu'elle propose. Elle se concentre beaucoup sur la phase d'inventaire et d'analyse de l'inventaire du cycle de vie, ainsi que sur la partie allocation. La méthode fait appel à une simplification en étiquetant d'une part les impacts de la phase d'usage comme impacts ou empreinte OPEX, regroupant tous les autres sous l'appellation CAPEX. Ce regroupement masque des dynamiques potentiellement hétérogènes entre phases de transport, de création et de fin de vie. Si la phase de création correspond bien à l'appellation CAPEX, i.e. un coût "payé" en une fois, les phases de transports et fin de vie peuvent avoir des coûts plus variables, notamment si l'on inclut des équipements reconditionnés [183]. Le choix de ce regroupement, effaçant potentiellement les spécificités de ces phases, est justifié par la quantité moindre de données aux sujets de ces dernières, ainsi que par le fait que l'échelle service amène à moins considérer ces détails que le ferait une ACV d'un équipement, se concentrant sur une échelle plus haute.

Enfin, cette méthode ne propose pas d'UF générique ou unique, autre que celle proposée par l'ITU [152] à savoir "une année de service" qui semble par défaut être la meilleure échelle d'étude d'un système tel qu'un service numérique. Cela témoigne en partie la difficulté voire l'impossibilité d'une réponse honnête et exhaustive à la question : "Quel est le coût individuel de telle action avec ce service numérique". Les UF centrées sur un acte simple offrent une base d'analyse moins solide que les UF à l'échelle du service, même si

elles peuvent être utilisées à des fins de communication, avec les réserves nécessaires.

En conclusion, comme rappelé au début de ce chapitre, le but ici n'était pas de mettre en valeur une contribution purement originale, mais d'exposer la synthèse générique et applicable de l'ACV de services numériques telle que je l'ai observée puis utilisées pour notamment les travaux présentés dans les deux chapitres suivants 5 et 6.

Dans ces chapitres, la méthodologie présentée ici sera un socle sur lequel des spécificités vont être ajoutées pour mieux s'adapter à des sous-types de services numériques, comme les services d'IA générative, ou bien à des situations d'études particulières, comme l'hébergement par une même entité d'un ensemble de services numériques hétérogènes.

Étude de cas d'un service unique : Stable Diffusion

*L'IA n'est ni artificielle ni intelligente.
Elle est à la fois incarnée et matérielle,
faite de ressources naturelles, de
carburant, de labeur humain,
d'infrastructures, de logistique,
d'histoires et de classifications.*

Kate Crawford - Atlas of AI

L'intelligence artificielle (IA) représente probablement une nouvelle étape de la transformation numérique due à ses nombreuses applications. Malheureusement, en accélérant la croissance du numérique, l'IA contribue aussi à la croissance des multiples dommages environnementaux causés par son secteur. Une tendance récente dans les modèles, dits d'IA générative (Gen-AI), tels que ChatGPT a particulièrement attiré l'attention, car leurs entraînements nécessitent l'utilisation intensive d'un nombre massif d'équipements et ressources informatiques spécialisées. La question de la durabilité des technologies TIC doit inclure cette nouvelle technologie et ses applications. Notamment en estimant son impact environnemental.

Dans ce chapitre, la méthodologie d'ACV de services numériques précédemment présentée sera appliquée à ce contexte. En calculant l'ensemble des coûts environnementaux d'un service Gen-AI de bout en bout, il sera montré comment élargir notre vision de l'impact de ces technologies. Il sera ainsi montré que c'est bien en tant que service que la Gen-AI engendre un ensemble hétérogène d'impacts environnementaux. La fonction d'inférence de ces modèles est accessible sur le web sous forme de services, et leur utilisation mobilise en outre les terminaux, les réseaux et les datacenters, des utilisateurs finaux. Il est aussi montré, grâce à l'approche ACV des services numériques que la décarbonation des sources d'électricité de ces services ne suffira pas à résoudre le problème de leur durabilité, en raison de leur consommation d'énergie et de métaux rares. Cette consommation posera inévitablement la question de la faisabilité dans un monde aux ressources finies.

Cette adaptation de la méthodologie d'ACV de services numériques se focalise sur un service basé sur un modèle d'IA générative pour donner une méthode générique d'éva-

luation *ex-ante* des impacts environnementaux d'une telle solution. Cette adaptation est supportée par l'étude d'un cas d'usage sur le service *Stable Diffusion* [260] capable de générer des images à partir de texte. Ce cas d'utilisation est basé sur une observation expérimentale de la consommation d'énergie de l'entraînement et de l'inférence de Stable Diffusion. Grâce à différentes analyses de sensibilité, différents scénarios sont explorés et permettent une réflexion sur les impacts environnementaux des services d'IA générative.

Ce chapitre reprend majoritairement deux articles, un publié [261] et un à paraître, tous deux inclus dans le sujet de la présente thèse.

5.1 Contexte IA et environnement

L'apparition successive de logiciels de génération d'images à l'été 2022, puis d'agents conversationnels performants comme Chat-GPT quelques mois plus tard, tous deux basés sur l'intelligence artificielle (IA), a popularisé le terme d'IA générative (Gen-AI). Ce terme désigne désormais un secteur qui revendique une croissance aussi rapide que significative [262], intégrant ses fonctionnalités dans de nombreux services existants, recherche en ligne, développement de logiciels, retouche d'images, etc. Cette croissance, pouvant être vue comme une nouvelle phase de digitalisation [263], s'additionne à celle préexistante du secteur numérique, amplifiant les problématiques environnementales précédemment décrites dans ce manuscrit.

Si l'on peut espérer que certaines applications numériques contribueront à relever les défis de la durabilité en réduisant l'impact des activités humaines sur l'environnement, il reste difficile de mesurer l'impact positif ou négatif de la technologie numérique sur l'environnement. Différents travaux ont étudié cet enjeu des impacts environnementaux "nets" d'une solution numérique, soulignant toutes les grandes incertitudes entourant ces exercices [14, 124, 177]. La question de la durabilité de l'IA ne peut être abordée scientifiquement sans moyens d'en évaluer les dommages environnementaux. En outre, il est nécessaire d'évaluer les applications réelles de cette technologie et pas seulement des sous-parties. Car c'est bien, le déploiement de la Gen-AI en tant que service, disponible en ligne à partir de n'importe quel appareil, comme Chat-GPT ou Stable-Diffusion, qui soulève de nombreuses questions, même au-delà de la soutenabilité. C'est pourquoi le choix est ici fait d'utiliser l'ACV de service numérique, qui permet une évaluation environnementale de Stable Diffusion en tant que service, de bout en bout.

5.1.1 IA et Gen-AI au sein du numérique : définitions et enjeux

L'IA peut être définie comme la capacité d'un ordinateur à automatiser une tâche qui nécessiterait normalement un discernement humain [264]. L'apprentissage machine, ou *machine learning* (ML), et l'apprentissage profond, ou *deep learning* (DL), font référence aux IA qui apprennent à partir de données existantes. Ces logiciels d'IA sont généralement appelés modèles. L'IA générative, ou Gen-AI, est un sous-domaine de l'intelligence artificielle dont l'objectif est de générer de nouveaux contenus, par exemple des images, des musiques ou encore des textes rédigés. Comme le souligne [265] dans son étude récente sur l'emploi de ce terme d'IA "générative", on peut faire remonter cette notion à la publication introduisant en 2014 les réseaux antagonistes génératifs [266], des algorithmes

d'apprentissage pour l'IA. Toutefois, c'est depuis la fin de l'année 2022 que ce terme gagne en notoriété et par conséquent en intérêt, bien au-delà de la recherche en IA. Cependant, la popularisation soudaine et importante de la notion d'IA générative ne doit pas occulter le long processus de recherche dans lequel s'inscrivent ces modèles [267].

Le développement de modèles d'IA générative nécessite la collecte de données et l'apprentissage à partir de ces mêmes données, ce qui comprend : premièrement, la sélection de la meilleure structure de modèle et du meilleur algorithme d'apprentissage pour la tâche donnée, deuxièmement, l'application de cet algorithme au modèle et aux données collectées. La première étape cache un processus de développement coûteux, car les modèles d'IA générative sont généralement composés de plusieurs modèles déjà développés. La deuxième étape s'appelle l'entraînement. Une fois qu'un modèle a atteint la qualité visée, il peut être utilisé sur de nouvelles données, ce que l'on appelle la phase d'inférence. En raison de la taille toujours croissante des bases de données et des modèles, l'entraînement peut nécessiter des centaines de processeurs graphiques (GPU)¹ fonctionnant en parallèle. À titre d'exemple, le modèle Stable Diffusion a été entraîné avec 258 GPU et 64 processeurs (CPU)². Les données sont regroupées en lots afin qu'elles soient efficacement traitées en parallèle. L'entraînement d'un lot correspond à une étape d'entraînement, utilisée comme métrique sous la dénomination de *step*³. Des centaines de milliers de *step* ont été nécessaires pour entraîner le modèle utilisé par Stable Diffusion, en utilisant un jeu de données d'images annotées *Laion2B* d'un volume de 84 TiB.

Il est à nouveau intéressant de noter que, si des modèles puissants existaient déjà avant 2022, c'est bien la disponibilité en ligne de ces modèles en tant que service, sous l'impulsion de Chat-GPT, qui est à l'origine de la popularité de la Gen-AI. La nouvelle utilisation massive de ces services et des infrastructures informatiques qui les soutiennent, avec leur forte demande en électricité et en équipements critiques, pose nécessairement la question des impacts environnementaux de tels services. Quel est l'impact environnemental de ces nouveaux usages numériques, de quelle nature et dans quel ordre de grandeur ?

5.1.2 L'empreinte environnementale de l'IA : un travail encore incomplet

Incluse dans les études globales sur l'impact environnemental du numérique, la croissance des coûts du machine learning [268] a fait émerger l'IA comme un enjeu de soutenabilité à part entière au sein du numérique en 2019 [269, 270]. Ces premières études se concentrent sur la consommation d'électricité pendant la phase d'entraînement de l'IA ainsi que sur les émissions de gaz à effet de serre (GES) associées à cette consommation. Sur cette lancée, plusieurs autres études ont suivi [271–274], sans pour autant s'éloigner du champ de la mesure de la consommation d'électricité pour l'entraînement des modèles d'IA. Ce choix peut s'expliquer en partie par le fait que, dans de nombreuses études, les modèles d'IA sont considérés plus comme des projets de recherche que des produits de consommation de masse, ces publications datant néanmoins généralement d'avant l'explosion de l'usage de la Gen-AI.

1. Pour *Graphics Processing Unit*

2. Pour *Central Processing Unit*

3. Il est donc bien à noter qu'une *step* n'est pas une métrique universelle, mais définit contextuellement par l'algorithme d'entraînement du modèle d'IA

Dans ce contexte initial, plusieurs méthodes ont été développées pour mesurer ou estimer la consommation d'électricité de la phase d'entraînement. La plus populaire est basée sur une donnée issue de la fabrication des GPU et CPU appelée *Thermal Design Power* (TDP) [275], exprimée en Watt, qui est une bonne estimation de la puissance maximale d'un composant⁴. La phase d'entraînement étant généralement très intensive, on peut supposer que les composants de calcul (généralement les GPU) fonctionnent à leur puissance maximale. Par conséquent, une estimation basée sur le TDP consiste à multiplier le TDP des GPU par la durée totale de l'entraînement. Cette méthode est assez simple, puisqu'elle ne nécessite que la connaissance du TDP et de la durée de l'entraînement, mais elle présente plusieurs limites. Tout d'abord, les GPU ne sont pas les seuls composants d'un serveur. Les CPU, la RAM, les commutateurs ou encore les ventilateurs doivent également être pris en compte.

Une méthode alternative consiste à mesurer la consommation d'électricité pendant l'exécution de la phase d'entraînement. Les wattmètres physiques sont les outils les plus précis et les plus complets, mais ils nécessitent également une intervention et donc un accès au matériel, ce qui est, dans la plupart des cas, difficile, voire impossible. La phase d'entraînement peut également être surveillée à l'aide de wattmètres logiciels (utilisant RAPL ou NVML), qui indiquent la consommation d'électricité d'une sélection de composants [225]. De nombreux outils ont été développés sur la base de ces études. Ces logiciels et bibliothèques sont plus précis qu'une estimation basée sur le TDP, mais présentent les mêmes limites quant à l'exhaustivité des composants pris en compte. Le principal inconvénient des wattmètres est la nécessité de les utiliser pendant l'exécution, sans possibilité d'estimation *ex ante* ou *ex post*. La littérature manque d'une méthodologie accessible, reproductible, fiable et précise pour la Gen-AI.

La popularisation de la Gen-AI a élargi le champ d'application de ces méthodologies, en englobant plus d'impacts directs. Conséquence directe de l'usage, les inférences faites à partir des modèles sont étudiées [262, 276–278], même si cela reste souvent uniquement à travers la mesure de la consommation d'électricité. Les études qui commencent à inclure le cycle de vie complet des équipements sont encore rares [115, 116], et se limitent à l'empreinte carbone de l'entraînement et de l'inférence. Ce périmètre actuel n'aborde pas entièrement toutes les questions concernant les impacts environnementaux de la Gen-AI. Comme le secteur numérique a une empreinte CAPEX, ou *embodied*, au moins aussi significative que son empreinte OPEX, ou *operational* [18, 84], il est essentiel d'inclure l'impact de l'ensemble du cycle de vie. Le secteur numérique a également un fort impact environnemental qui va au-delà des émissions de GES [18], comme détaillé dans les chapitres 1, 2 et 3. Enfin, la Gen-AI se repose et fonctionne grâce à l'écosystème numérique existant, sur lequel elle exerce logiquement une pression à la croissance. Elle nécessite des terminaux pour ses utilisateurs, e.g. des smartphones et des ordinateurs, ainsi que des réseaux de télécommunication pour être accessible depuis ses datacenters. Toutes ces ressources sont essentielles au déploiement de la Gen-AI telle qu'elle existe aujourd'hui et font donc partie de la problématique de la durabilité de l'IA. Il est donc attendu et nécessaire de continuer à étendre le champ d'étude de Gen-IA en tant que service accessible ayant des impacts

4. Plus précisément, le TDP représente le maximum de chaleur, généré par l'utilisation d'énergie d'une puce, que les capacités de refroidissement d'un système peuvent dissiper sans affecter le fonctionnement de la puce. Cette donnée permet ainsi d'estimer la consommation maximum que le composant s'impose par sécurité.

environnementaux multiples, afin de mieux discuter de sa durabilité.

5.2 Méthodologie : L'IA comme service

Les travaux se concentrent sur l'IA en tant que service, ce qui signifie que non seulement les coûts spécifiques de l'IA, c'est-à-dire les phases d'entraînement et d'inférence, sont intégrés, mais aussi l'infrastructure requise pour son utilisation en ligne, en tant que service, en supposant que le service soit accessible par l'intermédiaire d'une interface web, API ou site internet interactif. Comme le montre la Figure 5.1, un service Gen-AI est considéré ici comme la somme des terminaux des utilisateurs finaux, du réseau, de l'hébergement web, de l'inférence du modèle, de l'entraînement du modèle et de la gestion des données. Les flèches représentent les flux de données entre les différentes parties.

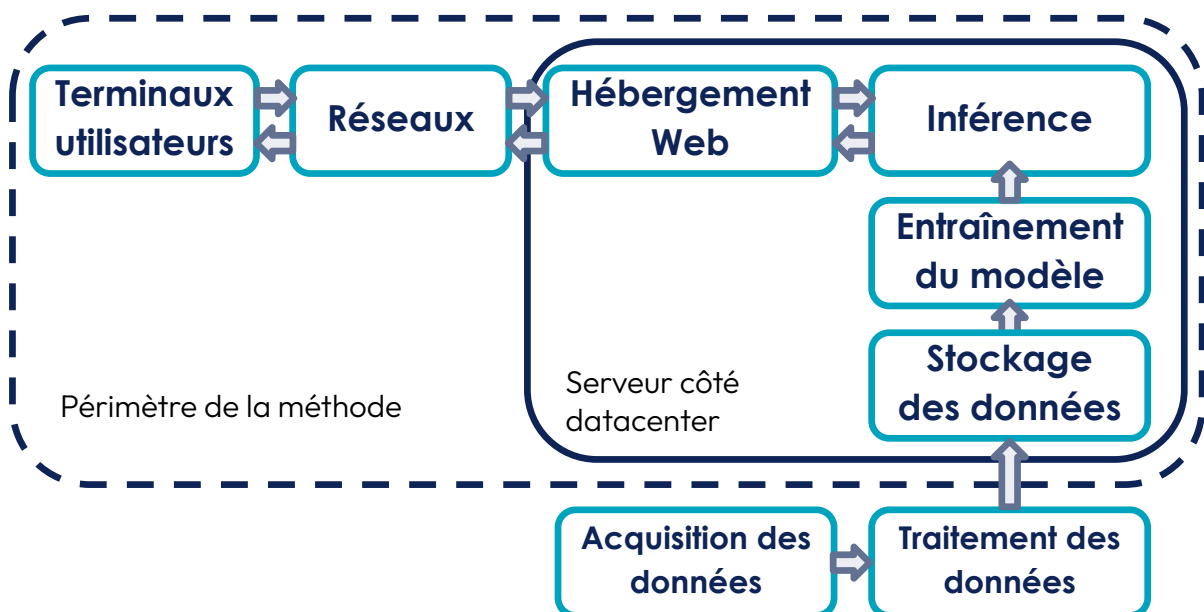


FIGURE 5.1 – Structure et périmètre considéré d'un service de Gen-AI pour l'étude

Les utilisateurs du service de Gen-AI accèdent au service à partir d'un terminal, en envoyant une requête qui est transférée sur les réseaux et gérée par un serveur web. Des serveurs spécialisés, équipés en GPU ou LPU, sont utilisés pour réaliser une inférence à partir du modèle et du contenu de la requête. Les résultats du processus d'inférence sont renvoyés aux utilisateurs par l'intermédiaire des réseaux. Le modèle a été préalablement entraîné au cours d'une phase spécifique d'entraînement, avec un accès aux données d'entraînement en plus de ses ressources de calcul. Les coûts environnementaux de la production et du traitement de ces volumineux jeux de données d'entraînement ne sont pas pris en compte dans ce travail, par manque de traçabilité sur des processus opaques. Il est difficile de faire ce que les auteurs de cet article [279] appellent "l'archéologie de grands jeux de données", malgré les enjeux de telles recherches.

Cette modélisation de l'activité d'un service Gen-AI et de l'impact environnemental qui en découle a été développée de manière à être modulaire et à pouvoir s'adapter à des

parcours utilisateurs plus spécifiques. Par exemple, un service hébergé et utilisé sur une machine personnelle pourrait supprimer les sections "Réseaux" et "Hébergement Web", ou un service utilisé sans interaction humaine par une autre machine via une API pourrait remanier la partie "Terminaux".

Cette méthodologie cherche à évaluer comme unité fonctionnelle (UF), en conformité avec [152], l'impact environnemental de l'usage du service pendant une année. Suivant la structure décrite dans la figure 5.1 et le tableau 5.1, les sections suivantes présentent la manière dont est calculé l'impact environnemental de chaque partie du service considéré. Les différentes parties sont calculées séparément afin de garder l'aspect modulaire voulu.

TABLE 5.1 – Définition des différentes variables utilisées pour les équations

| Notations communes | |
|----------------------------------|---|
| I | Impact |
| e | Équipement |
| C_e | Consommation moyenne d'électricité de e pendant sa durée de vie |
| F_e | Empreinte CAPEX de e : fabrication, transport, et fin de vie |
| s_e | Part des clients utilisant e pour compléter l'UF |
| EGM_g | Impact du mix électrique moyen dans la zone géographique g |
| PUE | Efficacité de la consommation d'énergie du datacenter |
| $a_e(t)$ | Allocation pour le temps d'usage de e par rapport à l'ensemble de sa durée d'usage (i.e. durée de vie fois pourcentage d'usage actif) |
| Notations terminaux utilisateurs | |
| t | Temps d'usage de l'équipement |
| Notations réseaux | |
| n | Localisation et type de réseaux, mobile ou combinaison de filaire et Wifi |
| d | Volume de données transférées à travers le réseau |
| IF_n | Impact environnemental du transfert d'un gigaoctet via un réseau de type n |
| Notations hébergement web | |
| v | Nombre de visites |
| $a_e(v)$ | Allocation au nombre de visites gérées par e par rapport au nombre total de visites |
| Notations inférence du modèle | |
| i | Inférence faite sur une machine grâce à un composant dédié comme un GPU |
| $C_{i,e}$ | Consommation d'électricité pour i réaliser avec e |
| Notations entraînement du modèle | |
| tr | Entraînement du modèle d'IA |
| $C_{tr,e}$ | Consommation d'électricité pour tr réaliser avec e |

5.2.1 Équipement d'accès - Terminaux utilisateurs

Les terminaux de l'utilisateur final sont définis comme l'ensemble des équipements directement utilisés par l'utilisateur pour satisfaire l'UF. Il peut s'agir d'un smartphone, d'un écran et d'un ordinateur de bureau, ou d'un autre équipement TIC utilisé pour réaliser l'UF. Pour chaque équipement e , en fonction de sa part d'utilisation s_e nécessaire à la réalisation de l'UF, par exemple 50% d'UF complété par smartphone et 50% par ordinateur portable, nous calculons son empreinte OPEX due à la consommation d'électricité, et le CAPEX due au reste de son cycle de vie. La consommation d'électricité C_e est la moyenne de sa consommation sur la durée de vie pour chaque type d'équipement e . L'impact de la phase d'utilisation OPEX est lié à la consommation d'électricité et la zone géographique g où elle est utilisée. La fonction EGM_g représente les impacts du mix électrique local. Le reste de l'impact du cycle de vie F_e correspond aux données de l'inventaire du cycle de vie (ICV) de l'équipement. Tous les impacts passent ensuite par la même clé d'allocation temporelle $a_e(t)$. La durée d'utilisation totale de l'équipement est définie comme la durée de vie multipliée par son taux d'utilisation active (TAU). Le TAU est le pourcentage de temps de la durée de vie de l'équipement pendant lequel il est activement utilisé, par opposition au moment où l'appareil est éteint ou en veille. Cette allocation temporelle suppose que pendant la réalisation de l'UF, le terminal est activement utilisé par l'utilisateur.

$$I_{Terminaux} = \sum_e s_e \times a_e(t) \times (F_e + C_e \times EGM_g) \quad (5.1)$$

5.2.2 Équipement d'accès - Réseaux

L'impact de la partie réseaux dépend du volume total de données transférées d nécessaire pour réaliser l'UF, et des types de réseaux utilisés n . Le facteur d'impact moyen IF représente de l'impact moyen du transfert d'un gigaoctet par l'intermédiaire d'un réseau fixe ou mobile. Comme précédemment, s_n représente la part d'UF réalisé par chaque type n , par exemple le réseau fixe européen ou le réseau mobile français. IF_n prend en compte les équipements réseaux de l'utilisateur final comme sa box internet par exemple.

$$I_{Réseaux} = \sum_n s_n \times d \times IF_n \quad (5.2)$$

5.2.3 Serveur de datacenters - Hébergement web

Cette section se concentre sur l'équipement datacenters utilisé pour accéder ensuite à la fonctionnalité d'inférence du modèle, c'est-à-dire le système qui reçoit et traite les requêtes des utilisateurs, ou même qui leur fournit une interface graphique via un site web, par exemple.

L'impact de cet hébergement web est calculé par la somme de l'utilisation s_e de chaque équipement e comme dans l'équation (5.1). Cependant, ici, la répartition ne se fait plus en fonction du temps, mais en fonction de l'utilisation. Les équipements de ce type sont capables de traiter plusieurs tâches en parallèle, ce qui expose une allocation basée sur le temps à un risque de sur-allocation. On suppose donc que l'équipement est dédié uniquement au service et à la satisfaction d'UF. L'allocation est basée sur le nombre total d'UF

auxquelles l'équipement a participé par rapport au nombre d'UF satisfaites par le service, ce qui correspond au nombre v de visites du service web. Le PUE est ajouté comme facteur à la consommation d'électricité pour tenir compte de l'efficacité énergétique du centre de données dans lequel se trouve l'hébergement web.

$$I_{HébergementWeb} = \sum_e s_e \times a_e(v) \times (F_e + C_e \times EGM_g \times PUE) \quad (5.3)$$

5.2.4 Serveur de datacenters - Inférence du modèle

L'inférence est prise en charge par un équipement disponible "à la demande", en l'occurrence un serveur équipé d'un GPU⁵ e . Le calcul d'une inférence i sur un appareil e consomme $C_{i,e}$ d'électricité et prend t de temps. Le coût de l'électricité est ensuite multiplié par le PUE du datacenter et le facteur d'impact local de l'électricité EGM_g .

Le reste de l'impact est calculé avec une allocation basée sur le temps $a_e(t)$ sur l'équipement e , en supposant que les GPU ne peuvent pas effectuer de tâches en parallèle. Comme pour les terminaux des utilisateurs finaux, l'allocation est proportionnelle à la durée de vie totale multipliée par le taux d'utilisation actif. Cela représente le fait que les équipements *à la demande* fournis par les datacenters ne sont pas toujours utilisés.

$$I_{Inférence} = \sum_i C_{i,e} \times EGM_g \times PUE + a_e(t) \times F_e \quad (5.4)$$

5.2.5 Serveur de datacenters - Entraînement du modèle

Cette section se concentre sur les serveurs utilisés en datacenters nécessaires pour effectuer un ou plusieurs tr d'entraînement pour le modèle d'IA. Comme précédemment, l'allocation temporelle est basée sur la durée de vie de l'équipement, la durée de l'entraînement et le taux d'utilisation active des serveurs utilisés.

Cependant, une attention particulière est accordée à l'estimation de la consommation d'électricité. La reproduction de l'ensemble de la phase d'entraînement étant trop coûteuse et la modélisation pure insatisfaisante, une solution intermédiaire est nécessaire. L'entraînement étant divisé en étapes à coût constant, la mesure de la consommation électrique de quelques étapes suffit à estimer le coût total de la phase d'entraînement [280], en supposant la connaissance de quelques éléments de l'entraînement original.

Enfin, l'impact de l'entraînement est un coût fixe nécessaire au lancement du service. Lors de l'évaluation du coût d'un service utilisant un modèle pendant sa durée de vie, l'ensemble du coût d'entraînement est pris en compte. Cependant, lorsqu'on évalue le coût d'une utilisation unique d'un service effectuant une inférence à l'aide du modèle, le coût de l'entraînement doit être alloué à chaque inférence au prorata du nombre total d'inférences réalisées à partir du modèle. Un exemple de cette répartition figure dans la section 5.3.1.

$$I_{Entraînement} = \sum_{tr} C_{tr,e} \times EGM_g \times PUE + a_e(t) \times F_e \quad (5.5)$$

5. De plus en plus se développe pour la tâche spécifique de l'inférence l'utilisation de *latency processing unit* (LPU) à la place de GPU, ce point est abordé plus loin dans le chapitre, la méthode présentée reste valable avec des LPU

5.3 Étude de cas : Stable Diffusion

Dans cette section, la méthodologie est appliquée et validée sur un service d'IA basé sur Stable Diffusion [281], un modèle open-source de *deep learning* permettant de générer des images à partir de texte. Stable Diffusion a été développé par des chercheurs du groupe CompVis de l'université Ludwig Maximilian de Munich et de Runway grâce à un don de puissance de calcul de Stability AI et à des données d'entraînement provenant d'organisations à but non lucratif. Stable Diffusion a été choisi parce que son modèle est en libre accès et que ses versions successives peuvent être téléchargées sur Hugging Face [282]. Le modèle est également disponible gratuitement en tant que service [260] depuis au moins août 2022. Sur la page web principale, les utilisateurs remplissent un formulaire en décrivant l'image souhaitée. Dans les paramètres par défaut, le service génère quatre petites images en format 512×512. Il existe plusieurs versions du modèle, créées par des entraînements successifs de v1-0 à v1-5. La première version du modèle utilisée par le service au début de l'étude en août 2022 était la v1-4, remplacée par le modèle v1-5 deux mois plus tard.

5.3.1 Cadrage de l'étude

Objectifs et périmètre

Le but de cette étude est de fournir une estimation quantitative de plusieurs impacts environnementaux du service Stable Diffusion et de son utilisation. L'ACV de Stable Diffusion considérera la méthodologie d'ACV de services numériques et montrera l'applicabilité de la méthode précédemment exposée. L'étude sera faite dans un cadre attributionnel. L'ACV cherchera à s'appuyer au maximum sur des données, mesures et résultats d'expérimentations directes, malgré le nonaccès aux données de la société opérant le service Stable Diffusion. Dans le cadre des hypothèses qui seront nécessairement prises, l'étude prendra, par défaut, les hypothèses les plus "favorables", celles qui sont le plus susceptible de diminuer le résultat final. L'ACV s'oriente donc sur un résultat unique, d'estimation basse, dit *single point low estimate*. Ce choix de viser l'estimation basse a été effectué dans une volonté de renforcer la crédibilité du résultat obtenu et de communiquer à son sujet comme "le coût environnemental est au moins de...". Cette décision s'effectue dans un cadre où certains acteurs encouragent unilatéralement l'usage de ces technologies [283] leur attribuant un coût significativement bas et critiquent les travaux estimant l'empreinte environnementale de l'IA comme exagérant de manière disproportionnée l'empreinte des technologies évaluées [118]. L'estimation quantitative du résultat final n'est aussi pas le seul objectif de l'étude, la nature de l'impact environnemental, sa répartition entre les différentes parties du service et les différentes catégories d'impact, est une information précieuse que l'étude cherche à obtenir.

Quant au périmètre de l'étude, elle couvre une année de service d'août 2022 à 2023. L'étude utilisera le périmètre décrit dans la figure 5.1 et cherchera à localiser géographiquement au mieux chaque partie du service.

Les catégories d'impacts évalués dans l'étude, issues des méthodologies composant la référence suivante [188], seront :

- L'épuisement des ressources abiotiques en minéraux et métaux (ADP)⁶, évaluant la contribution à la raréfaction des ressources disponibles due à leur extraction. Cet indicateur est estimé en kilo équivalent d'Antimoine extrait, kgSb eq, suivant la méthodologie suivante [186].
- Le potentiel de réchauffement climatique (GWP)⁷, évaluant la contribution au réchauffement climatique de la planète. Cet indicateur est estimé en kilo équivalent de dioxyde de carbone émis, kgCO₂ eq, suivant la méthodologie suivante [185].
- L'utilisation d'énergie primaire (PE)⁸, évaluant la demande cumulée en énergie primaire. Ce n'est pas à proprement parler une catégorie d'impact environnemental, mais c'est un indicateur pertinent de la demande en ressource. Il est exprimé en mégajoules, MJ, suivant la méthodologie suivante [184].

À des fins de lisibilité des figures et tables, les notations abrégées : ADP, GWP et PE, seront reprises dans les sections suivantes.

Unités fonctionnelles

Deux unités fonctionnelles (UF) différentes ont été considérées pour cette étude de cas, conformément à la recommandation de l'ITU pour les services TIC [152]. Tout d'abord, une vision *client* du service, **UF1**, représente l'impact moyen d'une personne visitant le site web et soumettant un message, générant quatre images. Deuxièmement, une vision *hébergeur* du service, **UF2**, considère le coût du service pour un an, couvrant les périodes d'activité des versions v1-4 et v1-5 du modèle (2 et 10 mois) avant qu'une nouvelle version ne soit proposée sur le site.

Le coût de **UF1** est calculée par allocation en divisant le coût d'entraînement du modèle d'IA par le nombre total potentiel d'inférences générées par cette version du modèle. Cette clé d'allocation utilise le nombre moyen de visiteurs sur le site web et la durée de vie moyenne d'un modèle, c'est-à-dire la moyenne du nombre de mois pendant lequel le modèle reste en ligne.

Outils de mesure et ressources

Nous avons répliqué l'inférence et une partie de l'entraînement sur des nœuds du cluster Sirius de la plateforme expérimentale de recherche appelée Grid'5000 [284]. Ce serveur a été choisi en raison de sa grande similarité avec les ressources utilisées par ses développeurs pour l'entraînement et l'inférence à partir du modèle Stable Diffusion. Sirius est un serveur Nvidia DGX A100 (8 GPU Nvidia A100-SXM4-40GB, 2 CPU AMD EPYC 7742). Bien qu'il existe de nombreuses méthodes pour estimer ou mesurer la consommation d'électricité d'une utilisation du serveur [225], nous nous sommes appuyés sur les wattmètres physiques de la société Omegawatt.

Nous avons estimé l'empreinte du serveur d'entraînement et d'inférence via l'outil de calcul d'empreinte de serveur de l'association Boavizta [101]. Plus précisément, pour estimer l'empreinte des GPU, nous avons utilisé une méthode plus approximative [285] normalement appliqué pour les CPU mais permettant d'obtenir une estimation basse,

6. *Abiotic Depletion Potential*

7. *Global Warming Potential*

8. *Primary Energy*

cohérente avec les objectifs de l'étude. Nous utilisons d'autres bases de données ICV, comme la base de données de l'ADEME, pour estimer l'empreinte du reste de l'équipement décrit dans la section 5.2. Nous nous sommes également appuyés sur des statistiques générales sur les TIC et l'utilisation du web [286] et sur des outils de mesure du trafic web tels que Similarweb [248] ou HypeStat [247] pour obtenir des données telles que le nombre mensuel de visites ou le temps de visite moyen du site web. Nous supposons qu'un serveur Amazon Web Services c6gd.metal était suffisant pour l'hébergement du site web. Nous avons dû estimer le taux d'utilisation active (AUR)⁹ des serveurs, car leur valeur n'est pas publiquement partagée par les fournisseurs de cloud. Nous avons choisi de fixer les serveurs d'entraînement et les serveurs d'inférence à 80% et 40%, respectivement. Ces AUR sont élevés par rapport à l'AUR moyen [251, 287] mais les GPU de ces serveurs sont des ressources très demandées, en particulier pour l'entraînement intensif, comme nous pouvons le voir par exemple dans [288] qui utilisent un datacenter orienté vers la recherche universitaire. Ce choix est examiné plus en détail dans la section 5.3.3.

5.3.2 Résultats

Les résultats bruts sont affichés dans le tableau 5.2 pour les UF1 et UF2. Les résultats pour le scénario d'héritage issu de l'analyse de sensibilité sont aussi inclus à la ligne "UF2 - bis". Ils seront plus amplement commentés à la section dédiée 5.3.3 mais seront utilisés pour l'ensemble des figures s'attardant sur la distribution afin d'utiliser toujours le même scénario.

TABLE 5.2 – Impact environnemental de Stable Diffusion pour l'UF1 et l'UF2

| Unité fonctionnelle | Épuisement des ressources abiotiques, minéraux et métaux ADP (kgSb eq) | Réchauffement climatique GWP (kgCO ₂ eq) | Énergie Primaire PE (MJ) |
|--------------------------------|--|---|--------------------------------|
| UF1 - Usage unique | $6.32e^{-08}$ | $7.78e^{-03}$ | $2.01e^{-01}$ |
| UF2 - Une année de service | $4.59e^{+00}$ | $3.81e^{+05}$ | $8.64e^{+06}$ |
| UF2 - bis scénario héritage | $4.77e^{+00}$ | $4.63e^{+05}$ | $1.00e^{+07}$ |

Dans le cadre de l'étude, nous avons avec nos outils [247, 248] pu estimer un nombre de visites équivalent en moyenne à environ 6,3 millions par mois. Nous avons fait l'hypothèse que la moitié des visites sur le site web occasionnait une génération d'images, correspondant ainsi à l'UF1.

9. *Active Utilization Rate*

Nous avons utilisé un mix électrique américain dans nos calculs pour les sections "Inférence" et "Entraînement". Pour les sections "Terminaux utilisateurs" et "Réseaux", nous avons calculé un mix électrique moyen basé sur les pays les plus représentés dans la population d'utilisateurs.

La figure 5.2 illustre l'UF2, pour une année de service, la répartition des impacts entre les différentes parties du service pour chaque catégorie d'impact. Pour plus de lisibilité, la partie "Hébergement Web" est incluse dans la partie "Inférence", et la partie "Stockage de données" dans la partie "Entraînement". Ces parties ont à la fois un impact négligeable et sont fortement corrélées à la partie dans laquelle elles ont été fusionnées. Pour produire nos résultats, nous nous sommes basés sur le nombre de visites mesurées et les caractéristiques de toutes les différentes phases (ou versions) d'entraînement du modèle, i.e pas seulement les derniers entraînements. Nous avons aussi mesuré la consommation d'électricité des inférences et estimé par échantillonnage le coût de l'entraînement. Nous avons également retravaillé nos données afin de pouvoir séparer les impacts OPEX des impacts CAPEX, comme le montre la figure 5.3.

Ces résultats montrent les impacts significatifs que peut avoir un unique service d'IA générative, et cela, pour chacune des trois catégories d'impact. Il peut être remarqué que les terminaux des utilisateurs représentent la majeure partie de l'impact dans l'ADP, ce qui était attendu puisque ces appareils contiennent généralement une batterie et un écran dont le coût de fabrication est important. La demande en énergie PE est dominée par le coût d'inférence, ce qui est cohérent avec les rapports des entreprises d'IA [276]. Une autre observation intéressante est que les impacts des réseaux et des terminaux des utilisateurs finaux ne sont pas négligeables, ce qui valide la nécessité de les inclure dans de telles évaluations. En conclusion, on peut observer que l'impact de ce service d'IA générative est assez dispersé, il n'y a pas une unique partie du service qui porterait toute l'empreinte environnementale. Même si les catégories n'ont pas une empreinte strictement équivalente, une réduction de l'empreinte d'un tel service passera forcément par un travail sur différentes parties du service.

5.3.3 Analyses de sensibilités

Cette section d'analyses de sensibilités explore l'impact des paramètres spécifiques à l'approche ACV de services numériques ou à la Gen-IA. Les cas triviaux ne seront pas traités, tels que la migration vers des zones ayant un mix électrique plus favorable ou des *PUE* plus performants.

Taux d'utilisation actif des serveurs

Tout d'abord, concentrons-nous sur le taux d'utilisation moyen des équipements du datacenter. Ce paramètre est crucial pour l'attribution de l'empreinte d'un équipement coûteux nécessaire à l'inférence et à l'entraînement comme l'est un serveur équipé en GPU. Un taux d'utilisation actif bas diminue la durée d'utilisation totale de l'équipement pour sa durée de vie augmentant la clé d'allocation associée. Moins un équipement est utilisé, plus chaque usage prend, par allocation, une part importante de son empreinte. Dans la figure 5.4, l'impact du taux d'utilisation moyen des serveurs d'inférence et d'entraînement peut être observé. On constate que cet impact est faible pour les taux supérieurs à 20%, ce

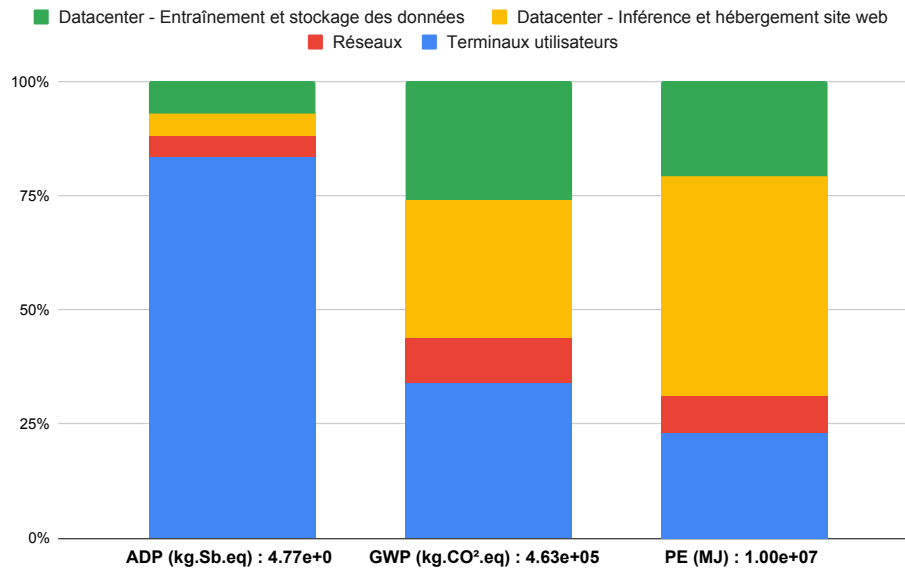


FIGURE 5.2 – Distribution des impacts environnementaux pour une année du service Stable Diffusion pour 75 millions de visites et 150 millions d’images générées

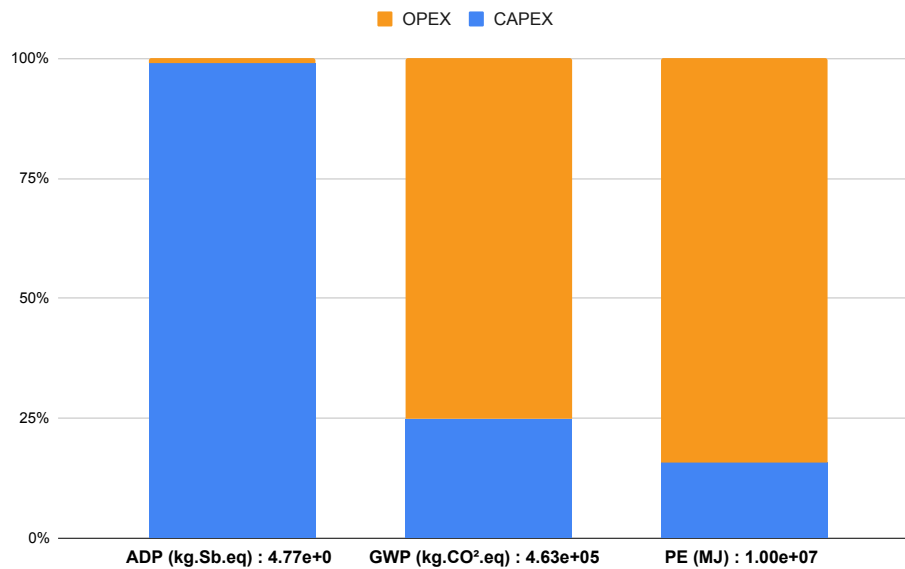


FIGURE 5.3 – Distribution des impacts environnementaux entre CAPEX et OPEX pour une année du service Stable Diffusion

qui était le cas pour l'étude. En dessous de 20%, l'impact est plus important pour le taux d'utilisation moyen des serveurs d'entraînement, ce qui était attendu puisque la formation nécessite un nombre important de ressources avec 32 nœuds de chacun 8 GPU. Si, dans les cas généraux, l'impact semble faible, il est difficile de faire des hypothèses solides sur ces paramètres. Le chiffre retenu de 80% pour l'entraînement provient d'un datacenter orienté recherche qui est beaucoup moins soumis à des contraintes de disponibilités comme le serait un datacenter commercial. Malheureusement, les rares études disponibles sur le sujet ne sont pas très optimistes [287], pointant entre 12% et 18% du taux d'utilisation moyen. Pour de tels taux, l'impact est important, notamment sur l'impact ADP. Cependant, ces études ne sont pas récentes, et compte tenu de l'amélioration rapide de la virtualisation et de l'ordonnancement dans les datacenters, on peut supposer que les AUR sont aujourd'hui supérieurs à 20% en moyenne mais cela ne tient pas compte de la spécificité des ressources demandées pour les services de Gen-IA, i.e. GPU et LPU, qui risquent de rendre plus difficile cet effort d'optimisation.

Inclusion des précédents entraînements du modèle

Deuxièmement, nous pouvons revoir la manière dont nous prenons en compte l'entraînement, nécessaire pour compléter l'UF2. Deux versions du modèle, v1-4 et v1-5, étaient disponibles au cours de la période évaluée, et donc, dans le scénario standard, la partie entraînement est l'addition de leurs deux phases entraînements. Cependant, ces versions du modèle avaient besoin de la phase d'entraînement des versions précédentes. Elles étaient nécessaires à l'achèvement de l'UF2 et, pour autant que nous le sachions, n'ont jamais été utilisées comme service à grande échelle. Nous comparons donc le scénario standard avec un scénario d'héritage complet consistant à inclure les coûts des entraînements de v1-0, v1-1 et v1-2, mais pas de v1-3, qui n'étaient pas directement nécessaires pour compléter l'entraînement de v1-4 et v1-5. Dans la figure 5.5 et le tableau 5.2, nous voyons que ce paramètre affecte significativement les catégories d'impacts GWP et le PE, même dans le cadre de l'UF2 où la phase l'entraînement a tendance à être plus négligeable que la phase d'inférence. Nous interprétons cela comme un avertissement de ne pas négliger l'impact des différents entraînements dans les services basés sur la Gen-AI.

Toutefois, il reste toujours difficile de définir la limite de ce périmètre d'inclusion. Les services d'IA générative sont basés sur de nombreuses précédentes recherches coûteuses qui ne sauraient être intégrées dans les ACV de services de Gen-AI. Il est nécessaire que ces entraînements aient un lien direct avec le service. Fixer le périmètre d'inclusion des entraînements est critique pour ce type de service, pour les entraînements en amont, mais aussi pour les entraînements en aval qui seront discutés section 5.4.4.

5.4 Discussions, limites et perspectives

Dans le cadre exprimé par la figure 3.3, le présent travail sur l'IA générative propose trois formes de contribution. Une contribution méthodologique section 5.2 et une contribution expérimentale section 5.3 ont été exposées précédemment. Cette section propose donc une contribution contextuelle, une réflexion sur comment traiter le sujet de l'IA générative au sein d'un numérique plus attentif aux enjeux environnementaux.

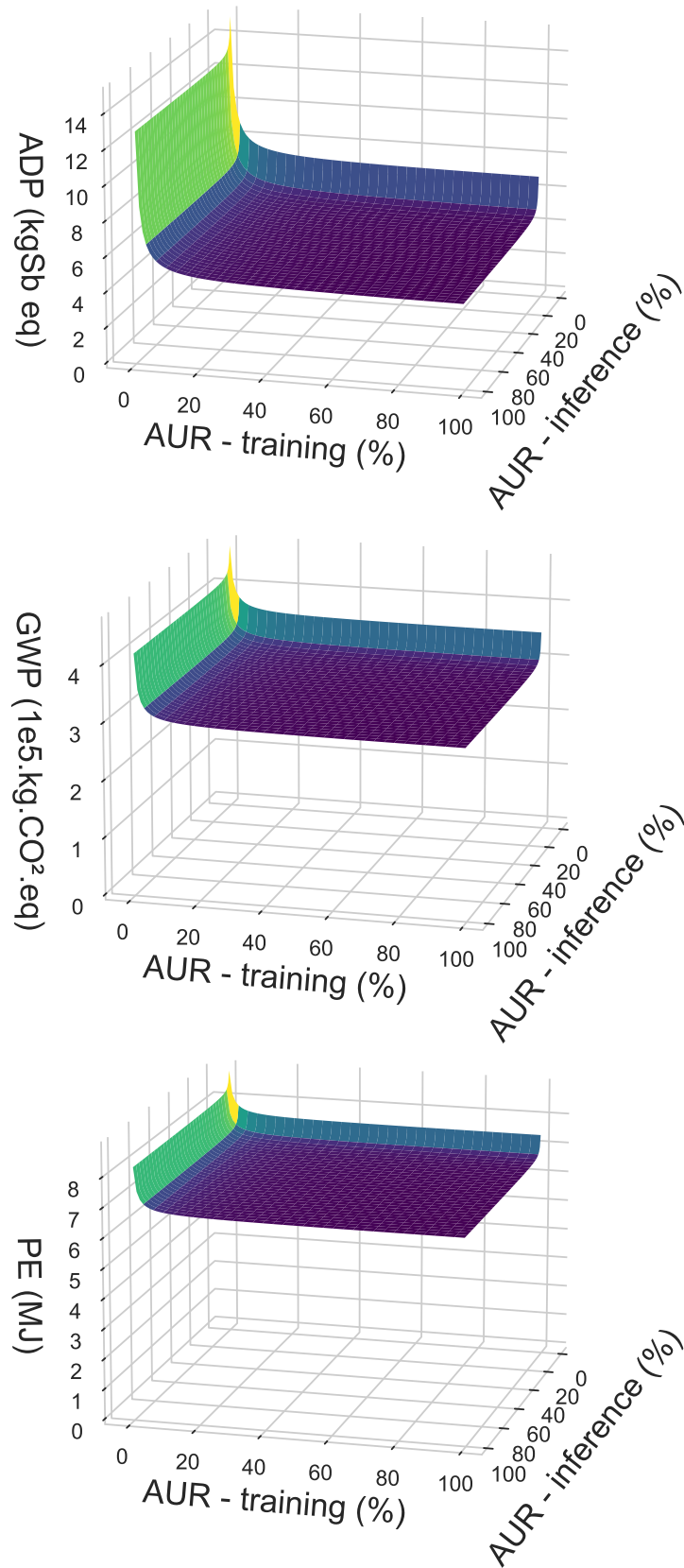


FIGURE 5.4 – Impact du taux d'utilisation actif des serveurs (AUR) du datacenter sur l'empreinte du service

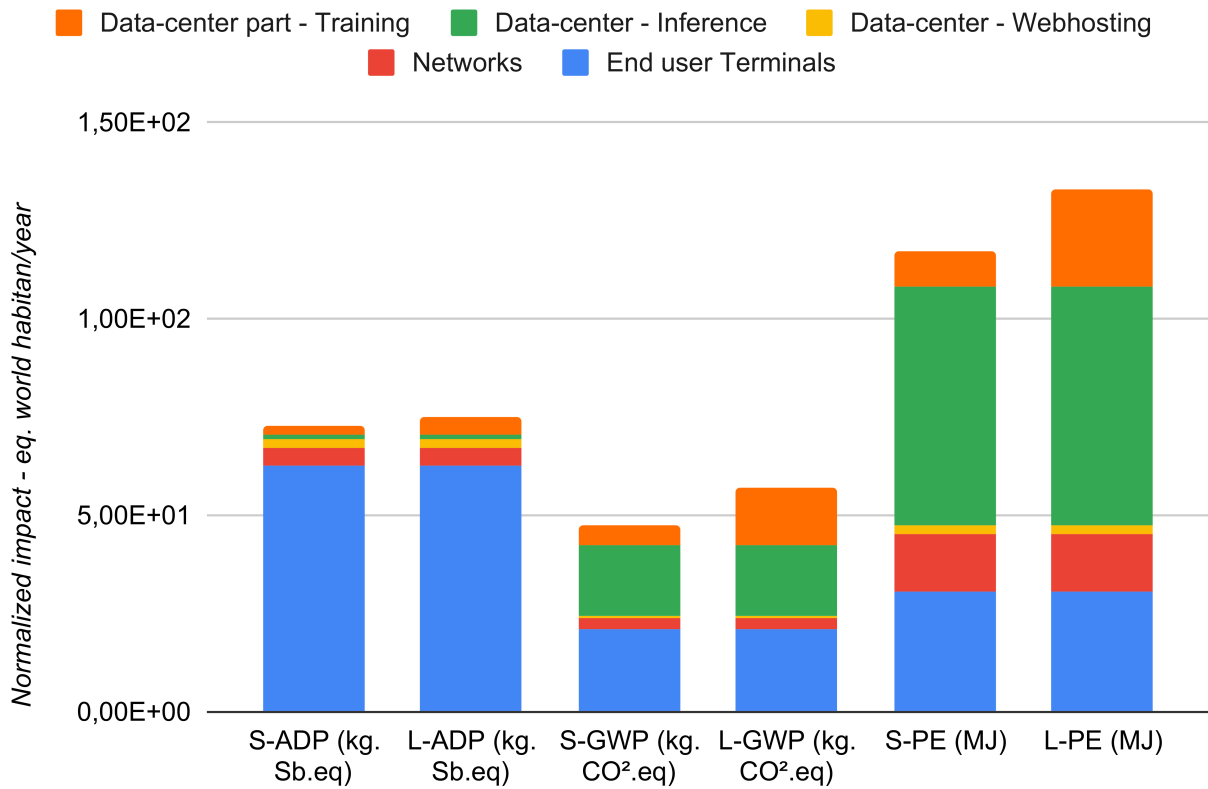


FIGURE 5.5 – Comparaison entre le scénario standard (S) le scénario d'héritage complet (L) sur la base des impacts normalisés, en équivalent habitant du monde, de l'UF2

5.4.1 L'enjeu de l'approche service dans un contexte comme celui de l'IA générative

L'usage de l'ACV de services numériques sur Stable Diffusion montre, au-delà du résultat brut important, la responsabilité de l'ensemble des parties du service dans l'empreinte de celui-ci.

Tout d'abord, la figure 5.2 montre que les terminaux et les réseaux représentent une part importante de l'impact du fonctionnement d'un service de Gen-AI : plus de 85% de l'impact ADP, plus de 30% de l'empreinte énergétique, et 45% de l'empreinte carbone. Cela valide la nécessité de les prendre en compte, d'autant plus que l'empreinte des réseaux et des terminaux croît avec le nombre d'utilisateurs, même si ces derniers n'utilisent pas de service en ligne, c'est-à-dire en ignorant le coût de la partie "réseaux". L'étude unique du coût de l'entraînement, ou de l'entraînement et de l'inférence ne suffit pas.

Deuxièmement, la vision multi-impact fournie par l'ACV nous montre que si la décarbonation de l'électricité consommée par les datacenters réduit l'impact sur le changement climatique, les émissions CAPEX de carbone et celles produites du côté de l'utilisateur restent significatives. L'empreinte énergétique est également une préoccupation à prendre en compte. La réduire nécessiterait d'importants gains d'efficacité, mais pourrait déclencher un effet de rebond [30]. Tout progrès dans l'efficacité énergétique de la Gen-AI pourrait bien conduire à une augmentation de son utilisation totale, comme c'est souvent le cas

dans le secteur numérique [126].

Enfin, la question de l'extraction des métaux pose un problème difficile à résoudre, déjà existant dans le numérique, mais pour lequel la Gen-AI n'apporte certainement pas d'améliorations. De plus, les services de Gen-AI stimulent la demande en GPU et autres équipements spécialisés [289], des ressources critiques dont l'empreinte est encore difficile à calculer. Dans notre étude, nous utilisons des valeurs sous-estimées pour l'empreinte des GPU, basées sur une méthode dédiée aux CPU [285].

La vision service pour les services de Gen-AI est d'autant plus critique quand il s'agit de services comme Stable Diffusion destinée à la consommation de masse. Si dans certains cas, un service basé sur l'IA peut, avec un usage raisonné, permettre des économies de ressources [290], le but de nombreuses technologies de Gen-AI est de créer des services s'additionnant à des usages préexistants. Cet ajout se fait avec la contrainte de proposer un service disponible en toute situation pour répondre à un nombre d'utilisateurs fluctuant, exerçant une pression sur l'ensemble des trois parties du service.

Dans le cadre de la réduction des impacts environnementaux de tels services, face aux limites de stratégies basées sur l'efficacité ou la décarbonation de l'électricité, un des principaux leviers restant semble être l'allongement de la durée de vie des équipements. Toutefois, la nature particulière des services de Gen-AI complexifie l'application d'un tel levier.

5.4.2 Obsolescence hardware et software

En effet, le principal levier de réduction de l'empreinte ADP et de l'empreinte CAPEX semble être l'allongement de la durée de vie des équipements. D'autant plus, lorsque l'on considère la tendance générale à l'augmentation des émissions de carbone CAPEX dans les TIC [84]. Cependant, les services de Gen-AI, avec leur importante consommation d'électricité [78, 262, 291], comme le montre la figure 5.3, pourraient indiquer un renversement clair des tendances, où la partie OPEX de l'empreinte redeviendrait prééminente. Cependant, cette conclusion n'a rien d'évident et dépend de la manière dont est considérée la séparation entre les impacts OPEX et les impacts CAPEX.

Si on considère à présent la figure 5.6, où, en reprenant la figure 5.3 précédente, sont représentés les impacts des parties "réseaux" et "entraînement du modèle" du service dans de nouvelles catégories distinctes. Dans la représentation précédente, la phase d'entraînement était principalement un impact OPEX en raison de sa forte consommation d'électricité. Cependant, dans le cas de la Gen-AI, il pourrait s'agir plutôt d'un impact CAPEX. La phase d'entraînement du modèle a déjà dépensé, en une fois, une grande quantité de ressources, de la même manière que la fabrication d'un équipement. Ce n'est pas un coût directement influencé et proportionnel à l'usage.

Il est donc possible de considérer la phase d'entraînement comme un dispositif permettant au service Gen-AI de fonctionner et comme faisant partie de son empreinte CAPEX, comme un coût de développement [122]. Il est alors possible de conclure qu'il est important d'inclure les modèles de Gen-AI dans l'action d'extension de la durée de cycle de vie des équipements. En d'autres termes, les entraînements à répétition peuvent rendre obsolètes les modèles précédents, gaspillant ainsi potentiellement des ressources, sans pour autant amener assurément un service significativement meilleur. Il est important de prolonger la durée de vie des équipements, de produits matériels, mais aussi celle de produits logiciels

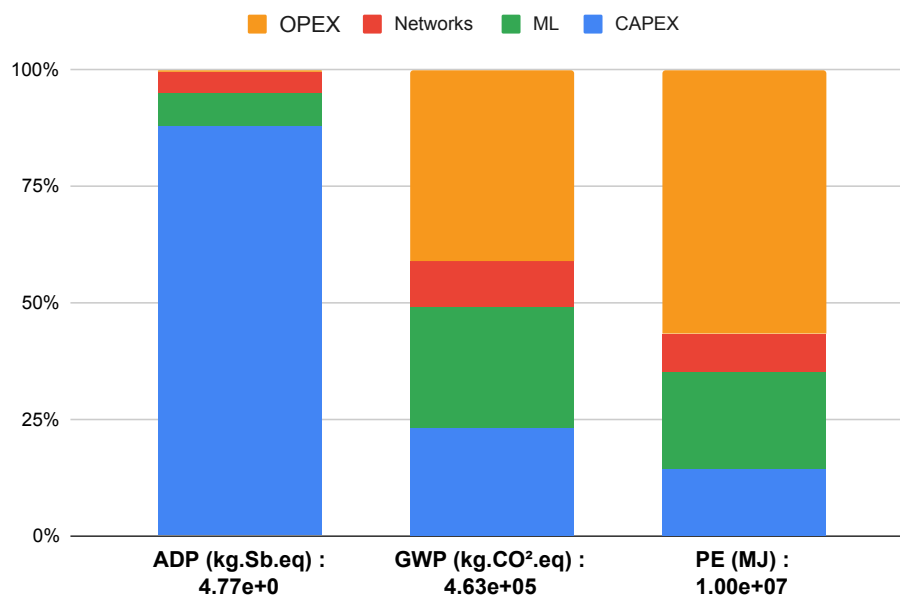


FIGURE 5.6 – Distribution des impacts environnementaux entre CAPEX, OPEX, réseaux et phase d'entraînement du modèle pour une année du service Stable Diffusion

comme les modèles d'IA.

Il est aussi possible de s'interroger sur la caractérisation de l'empreinte de la partie "réseaux". Certes, la majeure partie de leur empreinte provient de leur consommation d'électricité. Cependant, comme l'explique [105], les réseaux sont une infrastructure perpétuellement alimentée où le transfert de données pour un service ne génère pas en soi directement une consommation supplémentaire. En particulier pour les réseaux fixes, l'infrastructure a un coût de base qui couvre les besoins tant qu'ils restent dans les limites de sa capacité. La réduction du volume de données transférées par le service peut en fait avoir peu d'effet sur la consommation d'électricité [292], à moins qu'elle ne soit d'une ampleur telle qu'elle nécessiterait une montée en puissance de l'infrastructure [105] ou que la majeure partie du réseau ne soit mobile.

Ces deux derniers points ne visent pas à démontrer qu'une stratégie de réduction basée sur la réduction des flux de données et d'énergie pourrait être inefficace. Cependant, elle risquerait d'être au mieux insuffisante, ou pire, de contribuer à un effet rebond. En revanche, une approche basée sur l'utilisation parcimonieuse et efficace des ressources et des infrastructures existantes pourrait conduire à des gains plus durables. Il est peut-être plus intéressant d'optimiser les ressources existantes que de chercher à en créer de nouvelles, supposées plus efficaces.

5.4.3 Limites actuelles

La première limite de ce travail d'évaluation est, comme pour de nombreux travaux d'ACV dans le numérique, les hypothèses qui ont dû être posées par manque d'information sur le fonctionnement des services Gen-AI existants. Sans informations internes sur le

fonctionnement des services, il est naturellement plus difficile de modéliser correctement les impacts. Dans le cas présent, le service Stable Diffusion était disponible en open-source, tout comme son principal ensemble de données d'entraînement, et les informations sur les différentes phases d'entraînement, telles que le nombre d'étapes, étaient publiques. Plusieurs hypothèses ont néanmoins dû être posées, par exemple sur le dimensionnement de la partie "hébergement de site web". Bien que nous disposions d'indicateurs sur le trafic du site, nous n'avons pu trouver aucun moyen basé sur un modèle pour dimensionner le serveur web qui correspondrait à un tel besoin. C'est un exemple d'amélioration possible. Ces lacunes doivent être comblées, que ce soit par une meilleure modélisation des services basée sur des informations libres ou par une plus grande transparence de la part des fournisseurs de services.

La deuxième limite est que, même si ce travail représente l'une des premières approches multicritères de la Gen-AI, il est toujours important de prendre en compte les catégories d'impact environnemental qui n'ont pas pu être couvertes. Pour éviter de transférer des impacts des catégories mesurées vers des catégories non mesurées, il est nécessaire d'être exhaustifs. Par exemple, un mix électrique à dominante nucléaire pourrait transférer une partie de l'impact "réchauffement climatique" vers l'impact "rayonnements ionisants". Le transfert d'impact n'est pas fondamentalement mauvais, mais il doit rester transparent. De même, pour l'eau, un centre de données alimenté par des énergies renouvelables pourrait être un grand consommateur d'eau dans une région soumise au stress hydrique. L'augmentation du nombre de catégories d'impacts environnementaux observées pourrait conduire à des choix plus éclairés et plus légitimes.

La troisième limite est la question du partage de l'empreinte de la phase d'entraînement. Dans le cas de Stable Diffusion, le modèle étant librement accessible en ligne, il est vrai qu'il n'est pas possible de déterminer le nombre d'inférences réalisées grâce à lui. Les hypothèses ont été faites en se basant sur le trafic du site offrant le service. Comme les modèles ont été disponibles en tant que service sur le site pendant seulement un peu plus d'un an, il reste vraisemblable que le site ait concentré une grande partie de l'utilisation de ces modèles, encore plus aujourd'hui où certains pourraient les considérer comme obsolètes. La question de l'attribution de l'empreinte n'est pas triviale : si une organisation hébergeait un service d'IA générique utilisant un modèle qu'elle n'a pas entraîné, qui serait responsable de l'empreinte du modèle ? Certains pourraient même penser que les modèles les plus importants auraient un coût d'entraînement négligeable par rapport aux innombrables inférences qu'ils permettent dans de multiples services. Cela pourrait s'avérer exact, mais cela ne correspond pour l'instant pas à la situation d'entraînement répétitif de Stable Diffusion, ni à d'autres usages qui seront détaillés à la section suivante.

Enfin, la principale limite est peut-être le cadre attributionnel de l'ACV réalisée. Dans le cas des services de Gen-AI, proposant de nouveaux usages, avec parfois de nouveaux équipements [289], le cadre conséquentiel pourrait être plus à même de représenter plus spécifiquement l'excroissance d'empreinte que la Gen-AI représente pour le numérique.

5.4.4 Perspectives

L'un des mécanismes des services Gen-AI qui n'a pas encore pu être intégré est le *fine-tuning* [293]. Commun aux différents services de Gen-AI [267, 294], le *fine-tuning* correspond à tout entraînement supplémentaire du modèle entre sa phase principale d'en-

entraînement et son utilisation par le client final. Le fine-tuning peut être effectué au niveau d'une entreprise décidant d'intégrer un service de Gen-AI, qui entraînerait alors le modèle sur ses propres données. Il peut également être réalisé au niveau de l'utilisateur final, voire au niveau d'une session d'utilisation de l'utilisateur final. Ces actions mobilisent des ressources similaires à celles utilisées pour la phase d'entraînement. Il sera important d'intégrer ces coûts dans un modèle d'évaluation des services de Gen-AI et de trouver un moyen de mesurer ou d'estimer la fréquence et l'intensité de ces entraînements supplémentaires. Ces coûts pourraient modifier la façon dont est répartie l'empreinte de la Gen-AI entre l'inférence et la phase d'entraînement. Comme l'illustre la figure 5.7, derrière un simple usage du service peuvent se trouver différents coûts d'entraînement du modèle à allouer à l'usage. L'intégration de ces coûts de fine-tuning est d'autant plus nécessaire que ces actions sont souvent entreprises dans le but d'optimiser la capacité d'inférence à partir du modèle, parfois en termes de consommation d'électricité, une optimisation qui pourrait être donc questionnée.

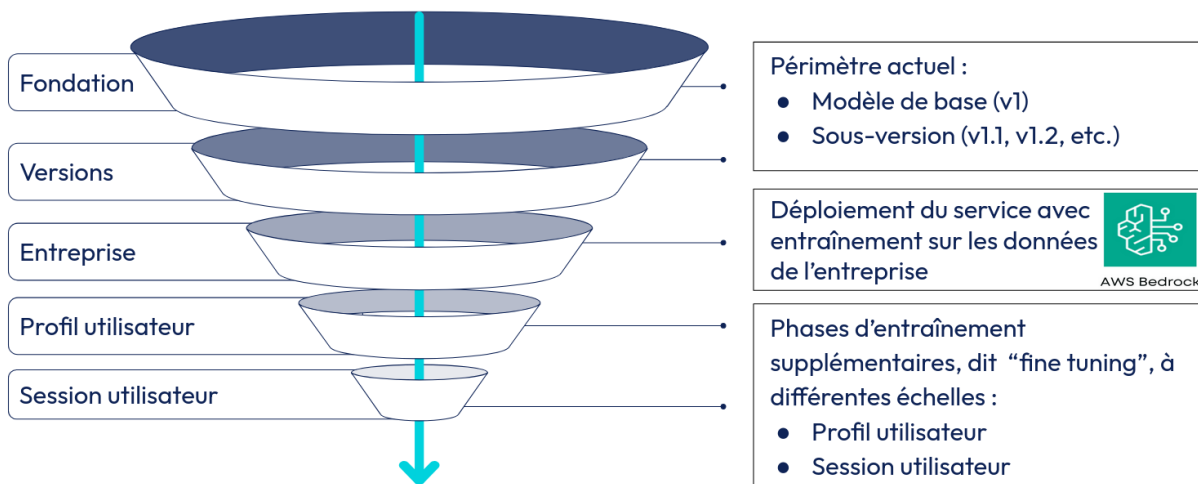


FIGURE 5.7 – Illustration des différents coûts d'entraînement potentiellement imbriqués derrière un unique usage d'un service d'IA générative

Une autre conséquence de la croissance des services de Gen-AI concerne les transformations qu'ils entraînent dans les datacenters. Les services de Gen-AI nécessitent l'utilisation massive de ressources spécifiques, comme certains GPU pour l'entraînement ou des GPU ou des LPU pour l'inférence. Cette nouvelle demande d'accès à un grand nombre de GPU en tant que service génère de nombreux défis organisationnels [295–297], en partie à cause de la difficulté de mutualiser et de virtualiser les GPU aussi facilement que les CPU. Il n'est pas possible pour un fournisseur de ressources cloud de partager aussi efficacement un GPU qu'un CPU. En conséquence, les datacenters risquent de devenir moins efficaces dans le partage des ressources. Cette baisse d'efficacité aura un coût environnemental, car elle devra probablement être compensée par la quantité pour répondre à la demande. Cette perte d'efficacité pourrait être accentuée par la spécialisation du hardware, précédemment les GPU servaient à la fois pour la phase d'entraînement et d'inférence, mais maintenant, à des fins de performances, des serveurs sont équipés avec des LPU [289] spécialisés uniquement pour la phase d'inférence. Une perte en flexibilité des équipements

induit le risque d'accélérer leur obsolescence, déjà hâtée par la pression, pour les datacenters, à proposer les équipements offrant les meilleures performances. L'hébergement de service de Gen-AI implique pour les datacenters de nombreux autres défis en termes d'organisations physiques des armoires de serveurs ou d'infrastructure électrique [298], les GPU consommant, pour une même surface dans un datacenter, significativement plus d'énergie. Comme mentionné précédemment, l'accès même à l'énergie pourrait devenir un enjeu pour une aussi soudaine et importante demande, nécessitant une adaptation du réseau électrique [291], voire l'implémentation de nouvelles capacités de production entièrement dédiées à ces datacenters [299]. Il serait alors intéressant, dans une approche conséquentialiste, de calculer le coût de cette croissance et spécialisation rapide des datacenters, largement imputable aux services de Gen-AI.

5.5 Conclusion

La mesure de l'impact environnemental de la Gen-AI peut être considérée comme l'une des bases de réflexion sur la soutenabilité d'une telle technologie. En ajoutant l'évaluation de l'échelle du service, le coût du cycle de vie complet des équipements et la prise en compte de différentes catégories d'impacts environnementaux, l'ACV de services numériques a permis de considérablement améliorer l'état de l'art des connaissances sur les impacts environnementaux de la Gen-AI qui était concentré sur des phases spécifiques comme l'entraînement ou l'inférence. En plus de souligner le coût important de cette technologie, ce chapitre a mis en évidence les différentes sources et les différents types d'impact sur l'environnement d'un service de Gen-AI en se basant sur l'étude réalisée sur Stable Diffusion. Cette connaissance détaillée de l'empreinte de la Gen-AI pourra permettre de développer des stratégies de réduction de son empreinte plus solides. La méthode présentée donne aussi les moyens de mieux prévoir, à l'avance, l'impact potentiel de l'entraînement du modèle et de l'utilisation du service. Dans une perspective de soutenabilité, cette capacité d'évaluation a priori est un outil essentiel d'aide à la décision. Cette connaissance des impacts environnementaux de la Gen-AI pourra être améliorée par la participation d'une communauté promouvant plus de transparence autour des services basés sur la Gen-AI. Cette transparence, dont les bénéfices dépassent les enjeux environnementaux, ne doit pas seulement se limiter au code et jeux de données utilisés, bien que cela soit crucial. Une évaluation plus profonde de l'impact d'une technologie de Gen-AI nécessite des données sur les moyens concrets utilisés pour la déployer et la faire fonctionner en tant que service. Un avis éclairé sur la soutenabilité de la Gen-AI dans notre société nécessitera une connaissance fiable de l'impact environnemental des services de Gen-AI.

Étude de cas d'un ensemble de services : le CHATONS Picasoft

Avec la digitalisation, les entreprises du secteur numérique deviennent des plateformes de services, plutôt que de simples vendeurs de matériel, ce qui leur permet de se placer en position centrale, et donc incontournable.

Fabrice Flipo

L'empreinte environnementale du secteur numérique est le plus souvent étudiée comme un tout ou, au mieux, comme l'addition de trois sous-secteurs principaux, les terminaux, les réseaux et les datacenters. Le présent manuscrit a mis en avant l'intérêt de l'usage de l'ACV pour approfondir la connaissance de cette empreinte, en étudiant les différentes catégories d'impacts et la perspective en cycle de vie. Plus précisément, l'ACV de services numériques permet de mieux approcher la corrélation entre usage du numérique et empreinte du numérique. Aucune des trois sous-parties précédemment citées n'existe de manière indépendante les unes des autres, elles sont reliées entre elles par les services numériques, autrement dit l'usage majoritaire du numérique, qui sont ainsi un cadre pertinent pour étudier la source des impacts environnementaux du numérique. Le principal problème est cependant que les services numériques reposent sur des équipements et des infrastructures fortement mutualisés, il est donc difficile d'évaluer le coût d'un service numérique. Dans le précédent chapitre, concernant les services d'IA générative, l'exercice fut facilité par l'évaluation d'un unique service, le fait que les équipements dédiés à l'IA ne traitent qu'une tâche à la fois et quelques hypothèses.

Ce chapitre propose une manière de dépasser cette limitation et d'évaluer, non pas un unique service, mais un ensemble de services numériques opérés par un même hébergeur. Similairement au chapitre précédent, il propose une contribution méthodologique à l'ACV d'un ensemble de services numérique, une contribution expérimentale en appliquant la méthode à un fournisseur de services numériques, et enfin une contribution contextuelle en proposant de nouvelles approches d'écoconception de services à partir de ses mesures. Ce chapitre adopte le point de vue du fournisseur de services sur la manière de calculer et

de réduire l’empreinte de ses services, et avec quelles données provenant de l’hébergement des services. C’est du point de vue de l’hébergeur, que l’empreinte est répartie entre les services, ce qui permet d’examiner précisément les impacts. Avec de telles données sur les impacts environnementaux des services hébergés, des stratégies pour prioriser les efforts d’optimisation et les politiques de décommissionnement sont proposées.

6.1 Contexte : Impact environnemental des services numériques

Le paradigme cloud [27], et l’industrie qui en découle, ont joué un rôle majeur dans l’établissement du monde numérique tel que nous le connaissons actuellement. La transition vers le cloud, encore en cours, transforme le secteur numérique d’une économie basée sur l’échange de logiciels et de biens matériels en une économie basée sur les services, aussi appelée économie de fonctionnalité [26]. Comme une extension des infrastructures de réseau, les infrastructures cloud ont été un moteur de croissance essentiel pour le secteur numérique. Malheureusement, au vu des enjeux environnementaux actuels, cette croissance soulève un certain nombre de questions.

S’il est vrai que les services numériques basés sur le cloud pourraient contribuer à réduire l’empreinte d’autres secteurs de l’activité humaine [300], cette affirmation reste sujette à caution [14]. Elle l’est d’autant plus dans un contexte où les services en question ont parfois été présentés comme une « dématérialisation », négligeant par principe leur propre impact environnemental [301]. Mais cette invisibilisation des coûts pourrait aussi être vue comme un manque de compétences, d’outils et de données pour évaluer l’impact environnemental des services basés sur divers équipements et infrastructures fortement mutualisés comme le cloud [96], rendant l’évaluation plus complexe que pour les autres biens¹.

Afin de répondre à ces enjeux d’évaluation de l’empreinte environnementale des services numériques, la section 3.1 a déjà exposé pleinement les avantages de l’ACV de services numériques par rapport aux autres approches préexistantes. Les principales limites peuvent néanmoins être résumées par les trois points suivants :

1. La focalisation sur la phase d’usage du numérique.
2. La prise en compte uniquement de la consommation électrique, et éventuellement de l’empreinte carbone de celle-ci, comme impact environnemental.
3. Le choix d’échelle soit trop haut, à l’échelle d’une organisation ou d’un secteur, ou trop bas à l’échelle d’un équipement ou d’un logiciel, mais rarement à l’échelle service.

Que ce soit pour rendre les services cloud plus durables en contrôlant ou en réduisant leur impact, ou pour juger des bénéfices environnementaux d’une solution basée sur le cloud, une meilleure mesure des impacts environnementaux, à l’échelle service, est nécessaire. À cette fin, la méthodologie proposée permet de calculer les impacts à l’échelle des services basés sur le cloud. Que ces services soient hébergés sur une instance physique, par

1. Voir les sections 2.1.3 et 2.2.3 pour plus de détails sur ces difficultés

exemple *bare metal*, ou virtuelle, la méthodologie permet de répartir l'impact de l'hébergement entre les services. Elle fournit également les outils nécessaires pour mesurer l'impact de bout en bout de chaque service à l'aide des informations relatives à l'hébergement. La méthode est conçue pour intégrer au mieux les données issues d'ACV d'équipements ou flux afin de fournir une estimation des divers impacts environnementaux de chaque service.

Ces évaluations individuelles de tous les services permettent ensuite de construire un tableau de bord environnemental permettant de prendre des décisions éclairées pour contrôler ou réduire l'empreinte des services. De cette manière, les services peuvent être priorisés en vue de leur optimisation, de leur réduction ou même de leur fermeture, dans le cadre d'une politique de sobriété numérique [128]. Enfin, ces évaluations environnementales des services permettront aux gestionnaires de rendre compte des impacts environnementaux de l'utilisation aux utilisateurs ou aux clients. En plus d'aider à se conformer aux normes pertinentes, ce rapport fournira des données utiles pour juger des avantages environnementaux ou non des solutions basées sur le cloud.

6.2 Méthodologie générale

Pour calculer l'empreinte au niveau du service, la méthodologie utilise des concepts de l'ACV, des données environnementales de l'ACV et des outils plus traditionnels de surveillance et de gestion des applications. La méthodologie est appliquée du point de vue du gestionnaire de services, à partir de l'instance cloud où les services sont hébergés. Le cœur de la méthodologie est la distribution des impacts environnementaux des équipements et de l'infrastructure parmi les services gérés par l'hébergeur. Elle est donc indifférente au type d'hébergement et au type de données environnementales introduites. Ce qui est important, c'est la répartition de l'empreinte et son influence au sein du périmètre choisi. Tout d'abord, il est important de présenter les outils utilisés pour fournir les données.

6.2.1 Outils mis en œuvre

ACV

L'ACV de services numériques telle que décrite au chapitre 4 sera la base de la méthode utilisée. Elle sera de nature attributionnelle, avec comme but de produire des estimations sur l'empreinte de chaque service numérique qui compose un ensemble de services opérés par un même hébergeur. L'ACV des services numériques utilisera pour ses clés d'allocations les données d'activités de différents outils décrits dans le paragraphe suivant.

Outils d'observabilité

Afin de mesurer l'utilisation des ressources de chaque service, des mesures en termes de pourcentage d'usage du CPU et de la mémoire vive (RAM) sont prélevées et calculées à intervalles réguliers. Ces mesures peuvent être enregistrées, par exemple, dans une base de données de séries temporelles VictoriaMetrics [242] en vue d'un traitement ultérieur. La fréquence d'interrogation des mesures d'usage des ressources est importante. Une valeur faible pour cet intervalle signifie une plus grande fidélité des valeurs collectées, car

l'utilisation de ressources peut parfois varier considérablement d'une seconde à l'autre. Mais une fréquence de lecture plus courte signifie également un coût plus élevé en termes de ressources de calcul et de stockage pour l'outil d'évaluation lui-même [245]. Un compromis doit être trouvé entre le coût et la précision de la mesure. Dans l'application de la méthode, si la période d'observation est suffisamment longue, un intervalle de relevé court n'est pas forcément nécessaire. La principale question reste la répartition de l'empreinte entre les services. Si un intervalle plus long peut potentiellement "manquer" plus de pics de consommation de CPU ou de RAM, on peut s'attendre, sur une période suffisamment longue, à ce que ces "manqués" soient répartis uniformément entre les services.

Lors de la récupération des métriques, nous devons les associer à chaque service que nous voulons étudier. L'utilisation de la conteneurisation (ex. Docker) permet d'atteindre cet objectif grâce à la technologie Linux cgroups. En effet, les cgroups peuvent être utilisés pour isoler un groupe de processus et permettre de rendre compte de l'utilisation des ressources de ce groupe de processus [243].

Afin de calculer l'empreinte d'un service par utilisateur, nous devons récupérer des informations sur l'utilisation du service. Tous les services sont accessibles via des proxys inversés. Ces composants émettent des journaux d'accès contenant, entre autres, la date et l'heure de l'accès, l'adresse IP du client accédant au service, son *user-agent*, le nom du service accédé, l'URL d'accès et la quantité de données transférées. L'analyse de ces logs nous permet de déduire le nombre d'accès, le nombre de visiteurs uniques et la quantité de données transférées. Nous utilisons GoAccess [246] pour effectuer ces analyses. Les journaux sont centralisés à l'aide de Loki [244] et analysés quotidiennement pour générer un rapport avec une répartition des opérations de traitement. GoAccess considère une combinaison unique d'adresse IP, de User-Agent et de date comme un seul visiteur. Ce nombre de « visiteurs » n'est évidemment pas une représentation exacte du nombre d'utilisateurs, mais il fournit un bon indicateur de l'affluence. Cette approximation est utile pour les services accessibles sans authentification. Elle peut bien sûr être remplacée par des données plus précises pour les services disposant de cookies utilisateurs plus fournis ou bien de logins et de comptes d'utilisateurs.

6.2.2 Distribuer l'empreinte de l'hébergement entre les services

La méthodologie calcule l'impact environnemental des solutions SaaS, que nous appellerons services, gérées par un seul fournisseur d'hébergement. L'évaluation est réalisée a posteriori pour une période de temps définie.

Étant donné une capacité d'hébergement H et un ensemble de services S . S peut être décomposé en plusieurs services, s_1, s_2, \dots, s_n , qui sont hébergés sur H . H peut être un serveur ou une machine virtuelle (VM), ou peut être décomposé en différentes parties h_1, h_2, \dots, h_m . Un serveur peut héberger plusieurs VM. L'important est que les ressources représentées par H soient dédiées uniquement aux services de S . Pour une période donnée, H a un impact environnemental I_{Host} . Ainsi :

$$I_{Host}(S) = \sum_n I_{Host}(s_n) \quad (6.1)$$

L'impact de I_{Host} peut correspondre à une ou plusieurs catégories d'impacts environnementaux, tels que le potentiel de réchauffement global (GWP) en kg d'équivalent

CO_2 , ou le potentiel d'épuisement des ressources abiotiques (ADP) en kg d'équivalent Sb. Reprenant la typologie du chapitre 4, I_Host peut être divisé en I_Host_{capex} et I_Host_{opex} , tel que l'équation (6.2) le décrit. Pour rappel, d'une part, CAPEX correspond à un coût unique, ici l'impact de la fabrication, du transport et de la fin de vie. D'autre part, l'OPEX correspond au coût de fonctionnement, au coût d'utilisation, par exemple la consommation d'électricité de l'équipement.

$$I_Host(S) = I_Host_{capex}(S) + I_Host_{opex}(S) \quad (6.2)$$

Empreinte CAPEX de l'hébergement

L'empreinte de H sera décomposée en 3 ressources, CPU, RAM et Stockage, plus une catégorie supplémentaire "Reste" qui représente ce qui n'entre pas dans les autres catégories², comme suit :

$$I_Host(S)_{capex} = I_Host(S)_{capex}^{CPU} + I_Host(S)_{capex}^{RAM} + I_Host(S)_{capex}^{Sto} + I_Host(S)_{capex}^{Rest} \quad (6.3)$$

Chaque service s_n se voit attribuer une part de l'empreinte de chaque pièce. Cette part est déterminée par les différentes clés d'allocation, telles que :

- $a_{CPU}(s_n, h_m)$ déterminé par le pourcentage moyen d'usage CPU de s_n au sein de h_m , proportionnellement à la somme de tous les pourcentages moyens d'utilisation CPU des s_n au sein de h_m .
- $a_{RAM}(s_n, h_m)$ déterminé par le pourcentage moyen d'usage RAM de s_n au sein de h_m , proportionnellement à la somme de tous les pourcentages moyens d'utilisation RAM des s_n au sein de h_m .
- $a_{Sto}(s_n, h_m)$ déterminé par le pourcentage moyen d'utilisation du stockage par s_n au sein de h_m , proportionnellement à la somme de tous les pourcentages moyens d'utilisation du stockage des s_n au sein de h_m .
- $a_{Rest}(s_n, h_m)$ déterminé pour chaque service s_n par la moyenne des trois clés d'allocation précédentes.

Ces clés d'allocation vérifient toutes, pour H , ou pour un h_m spécifique, que :

$$\sum_n a_{CPU}(s_n, h_m) = 1 \quad (6.4)$$

L'allocation du "reste" sert à allouer les parties de l'empreinte où l'utilisation n'est pas forcément mesurable. Par exemple, dans un serveur physique, l'empreinte du boîtier métallique ou de l'armoire. Cette clé d'allocation pourrait également être utilisée pour représenter une ressource telle que le CPU, la RAM ou le stockage dans le cas où l'accès à une mesure n'est pas possible pour une de ces ressources.

2. Par exemple des composants du serveur comme la carte mère, la carte réseau, le bloc d'alimentation ou le boîtier métallique.

L'empreinte CAPEX pour la partie hébergement du service peut donc s'écrire comme suit :

$$\begin{aligned}
 I_Host(s_n)_{capex} = & \\
 & \sum_m a_{CPU}(s_n, h_m) \times I_Host(s_n, h_m)_{capex}^{CPU} \\
 & + a_{RAM}(s_n, h_m) \times I_Host(s_n, h_m)_{capex}^{RAM} \\
 & + a_{Sto}(s_n, h_m) \times I_Host(s_n, h_m)_{capex}^{Sto} \\
 & + a_{Rest}(s_n, h_m) \times I_Host(s_n, h_m)_{capex}^{Rest}
 \end{aligned} \tag{6.5}$$

Empreinte OPEX de l'hébergement

L'empreinte OPEX de H sera considérée ici uniquement comme l'impact environnemental de la consommation d'électricité. Même si cela représente la majeure partie de l'impact OPEX, il s'agit d'une simplification qui laisse potentiellement de côté d'autres coûts OPEX, par exemple la consommation d'eau [110]. L'impact environnemental de H peut donc s'écrire comme suit :

$$I_Host_{opex}(S) = \sum_{s,h} C_s \times EGM_h \times PUE_h \tag{6.6}$$

Où C est la consommation d'électricité du service s et EGM le facteur d'impact environnemental associé à la consommation d'électricité du mix du réseau électrique dans la zone h .

Il y a plusieurs manières d'estimer la consommation de l'hébergement H . La borne basse est l'état de veille, dit *Idle*, et la borne haute une charge de 100% qui atteint le TDP des composants. Nous essayons alors d'estimer la consommation de H entre les deux en utilisant cette notion de pourcentage de charge :

Nous décomposons donc l'empreinte OPEX en deux sous-parties :

$$I_Host_{opex}(S) = I_Host_{opex}^{Idle}(S) + I_Host_{opex}^{Charge}(S) \tag{6.7}$$

Il existe plusieurs façons d'obtenir des estimations de ces valeurs. Elles peuvent être obtenues par des mesures directes [225], des analyses comparatives, des informations du fabricant, ou un mélange [101], comme utilisé dans la section 6.3. Pour chaque service, l'empreinte OPEX est répartie selon deux clés d'allocation, telles que :

- $a_{Idle}(s_n, h_m)$ est égale à $a_{Rest}(s_n, h_m)$, à défaut de meilleur indicateur pour partager ce coût, qui bénéficie à tous les services. En d'autres termes, de la même manière qu'ils ont tous besoin que le serveur ait un boîtier, considérant qu'ils ont un niveau de disponibilité identique entre eux, tous les services ont aussi besoin de la consommation *idle* du serveur pour fonctionner.
- $a_{Charge}(s_n, h_m)$ est égale à $a_{CPU}(s_n, h_m)$, considérant que le CPU est le meilleur indicateur de charge et donc de surconsommation d'électricité par rapport au stade *idle*.

Même si la consommation d'électricité est due à une charge de travail de forte intensité de RAM ou de disque, l'activité du CPU augmente également dans ces situations, ce qui en fait un bon indicateur de consommation d'électricité et non bien sûr la cause exclusive de la consommation. Ainsi, pour déterminer C_S^{Charge} , les données d'entrée ne se limitent pas à l'utilisation du CPU. L'empreinte OPEX d'un service s_n s'écrit donc comme suit :

$$\begin{aligned}
 I_Host(s_n)_{opex} = & \\
 & \sum_m a_{Idle}(s_n, h_m) \times I_Host(s_n, h_m)_{opex}^{CPU} \\
 & + a_{Charge}(s_n, h_m) \times I_Host(s_n, h_m)_{opex}^{RAM}
 \end{aligned} \tag{6.8}$$

6.2.3 Inclure les parties réseaux et terminaux utilisateurs

Reprenant la structure classique des services telle que montrée dans figure 3.1, maintenant que l'impact de l'hébergement de la partie datacenters a été pris en compte, une vision plus précise de l'empreinte du service implique de considérer l'empreinte des deux autres parties. Pour ce faire, les journaux d'accès seront utilisés pour attribuer à chaque service un impact basé sur son utilisation mesurée. Les données utilisées sont les suivantes :

- Nombre de visites par service
- Le système d'exploitation associé à chaque visite
- Le volume du trafic de données pour chaque service

Bien qu'elle utilise toujours des données mesurables, cette partie nécessite une plus grande confiance dans les hypothèses, en raison de la complexité et de l'incomplétude des métriques caractérisant les internautes utilisant les services web [302]. La définition d'une visite, suivant les données utilisées, peut varier, comme il sera détaillé dans le cas d'étude.

Réseaux

L'impact environnemental de l'utilisation du réseau, c'est-à-dire I_Net , peut être estimé assez facilement par des modèles d'impact par gigaoctet de données transférées. Certes, cette répartition par gigaoctet est critiquable [105, 176, 292], car le coût OPEX de transfert d'un gigaoctet n'est jamais identique. Néanmoins, cette répartition permet de partager équitablement I_Net entre les différents services et nécessite moins de données et de connaissances du réseau que d'autres méthodologies [103, 217]. L'impact de la partie réseaux d'un service s'écrit comme suit :

$$I_Net(s_n) = \sum_{type} p_{type} \times d_{n,type} \times IF_{type} \tag{6.9}$$

Où p_{type} est le pourcentage du trafic de données transférées par le type de réseau et $d_{n,type}$ le volume de données associées au service s_n et au type de réseau. La variable IF_{type} est un facteur d'impact, représentant l'impact environnemental moyen du transfert d'un gigaoctet sur le type de réseau. Le type de réseau comprend sa localisation et sa catégorie, par exemple s'il s'agit d'un réseau mobile ou fixe, les réseaux mobiles étant généralement

plus coûteux. Il peut être difficile d'obtenir le type de trafic associé à un service spécifique. Des hypothèses crédibles peuvent être formulées sur la base de la localisation des visites et du système d'exploitation utilisé pour ces visites. Il est raisonnable de considérer qu'une majorité du trafic provenant d'un système d'exploitation mobile est liée au trafic mobile et que le reste est du trafic standard.

Terminaux utilisateurs

L'impact des terminaux, c'est-à-dire I_Term , est le plus difficile à quantifier au niveau du service. Les équipements e utilisés pour une visite ont une empreinte I_{capex} et I_{opex} qui doit être attribuée. L'approche simpliste consiste à prendre la durée moyenne t_n d'une visite pour un service s_n et à construire une clé d'allocation temporelle. Un calcul idéal de l'empreinte donnerait, en développant les clés d'allocation temporelles, les équations suivantes pour **une visite** :

$$\begin{aligned} I_Term_{capex}(visit) &= \frac{t}{totalusetime_e} \times I_{capex}(e) \\ I_Term_{opex}(visit) &= C_e \times t \times EGM_n \end{aligned} \quad (6.10)$$

L'allocation de l'empreinte CAPEX se fait en utilisant la notion de temps d'utilisation total de l'équipement e . Elle diffère de la durée de vie en ne prenant en compte que la durée d'utilisation active. Cela implique que l'équipement e a un temps d'utilisation total correspondant à une partie seulement de sa durée de vie. Par exemple, un ordinateur professionnel peut être utilisé 8 heures par jour, ce qui fait que son temps d'utilisation total ne représente qu'un tiers de sa durée de vie. Cela influence considérablement I_Term_{capex} , d'autant plus si l'on considère que 8 heures est une utilisation quotidienne moyenne déjà élevée. Les équations (6.10) sont également fortement influencées par la durée de la visite. Une notion qui peut avoir du sens pour un service de streaming vidéo, mais moins pour quantifier l'intensité d'utilisation d'un chat en ligne où les utilisateurs sont connectés toute la journée. De plus, le type exact d'équipement e utilisé comme terminal par l'utilisateur final peut être difficile à détecter. Le système d'exploitation visité peut certainement identifier les utilisateurs de smartphones, mais ne peut pas faire la différence entre les utilisateurs d'ordinateurs portables et d'ordinateurs de bureau, des équipements dont l'empreinte est assez différente. Enfin, il est encore difficile de mesurer C_e , la consommation d'électricité pour une visite d'une durée t sur e , d'un point de vue serveur sans un suivi important des équipements des utilisateurs [241]. Nous proposons donc la simplification suivante pour l'empreinte des terminaux d'un service s_n :

$$I_Term(s_n) = \sum_e p_e \times a_e(t_n) \times (I_{capex}(e) + I_{opex}(e)) \quad (6.11)$$

Où p_e est le pourcentage moyen d'utilisation de e . Ici, $a_e(t_n)$ est une allocation unique basée sur le temps aussi bien pour la OPEX que CAPEX de l'empreinte, considérant $I_{opex}(e)$ comme l'ensemble de l'impact OPEX de e pour une durée de vie identique à celle utilisée pour déterminer $I_{capex}(e)$. Cette équation donne une estimation de l'impact sur les terminaux avec cette durée moyenne abstraite de visite t_n . Plus de détails et de discussions sur les difficultés d'intégration de cette partie sont disponibles dans la Section 6.4.3.

6.3 Étude de cas : Les services Picasoft

6.3.1 Contexte

Notre étude de cas nécessitait l'accès à un fournisseur de services, en mesure de nous fournir un accès à ses journaux d'accès et d'avoir une connaissance détaillée de son infrastructure d'hébergement. Nous devions appliquer la méthodologie à un ensemble de services hétérogènes utilisés par un public varié. Nous avons choisi le CHATONS³ Picasoft [303], car ses administrateurs étaient en mesure de nous fournir ces données et étaient gestionnaires et propriétaires de leurs serveurs, ce qui nous a permis d'estimer I_{Host} de manière fiable. Picasoft est un fournisseur local de services numériques avec un large panel de fonctions, adéquat pour une évaluation fine comme proposée par la méthodologie. Il était important pour la validation de la méthode d'avoir un nombre suffisant d'utilisateurs et de services sur les mêmes serveurs dont les données d'activités étaient accessibles.

Picasoft fournit gratuitement un ensemble de solutions SaaS open-source à divers utilisateurs, tel que décrits dans le tableau 6.1. Bien que certaines solutions SaaS fournissent un service spécifique, certains services sont un groupe de solutions SaaS. Il s'agit d'une vision fonctionnelle. Par exemple, Picasoft fournit un espace de travail numérique à trois associations différentes, dont elle-même. Si par exemple deux associations utilisent principalement Nextcloud, c'est chacune sur une instance différente, les applications ne sont accessibles qu'à un ensemble d'utilisateurs enregistrés, et ils l'utilisent en même temps que d'autres applications. Dernier exemple de groupe, nous avons défini l'administration du serveur comme un service. Nous avons inclus dans ce service toutes les applications nécessaires à tous les autres services. Cela inclut les métriques des systèmes utilisées par l'administrateur et cette étude, les processus d'arrière-plan et la documentation en ligne, disponible sous la forme d'un Wiki.

6.3.2 Objectif et périmètre

L'objectif de cette étude ACV est d'estimer la répartition de l'empreinte environnementale de Picasoft entre ses différents services. Le but est de valider la méthodologie présentée et d'explorer différemment la répartition de l'empreinte entre les services sous différentes méthodes. L'unité fonctionnelle observée sera le fonctionnement de l'ensemble des services pendant un mois.

Nous avons réalisé une étude sur un mois, du 17 novembre au 17 décembre 2023. Cette période permet de lisser l'effet potentiel de journées ou semaines anormales, en termes d'activités, et d'obtenir une période représentative d'une activité moyenne. Seules les procédures de sauvegarde des données ont été exclues, car ne s'intégrant pas de manière pertinente dans cette durée. Cette période ne contient pas, par exemple, de vacances scolaires susceptibles d'influencer le comportement des utilisateurs. Les métriques ont été enregistrées dans une base de données de séries temporelles VictoriaMetrics [242] pour ensuite être traitées. La fréquence de remontée des mesures de consommation des ressources a été fixée à 10 secondes. Comme expliqué dans la section 6.2.1, nous pensons

3. CHATONS pour : **C**ollectif des **H**ébergeurs **A**lternatifs, **T**ransparents, **O**uverts, **N**eutres et **S**olidaires, CHATONS est un collectif visant à "rassembler des structures proposant des services en ligne libres, éthiques et décentralisés".

TABLE 6.1 – Liste et description des services considérés pour l'étude

| Nom du service | <i>Nom solution SaaS</i> | <i>Description/Fonction</i> |
|-----------------------------|---------------------------------|---|
| podcast | Castopod | Plateforme de podcast |
| tube | Peertube | Plateforme de partage et de streaming vidéo |
| team | Mattermost | Salons de discussions textuel |
| pad | Etherpad | Éditeur de textes collaboratif en ligne |
| md | CodiMD | Éditeur markdown collaboratif en ligne |
| kanban | Wekan | Gestion de projets |
| voice | Mumble | Salons de discussions vocaux |
| drop | Lufi | Partage de fichiers volumineux |
| mobilizon | Mobilizon | Organisation d'évènements |
| board | Whiteboard | Tableau collaboratif en ligne |
| paste | Privatebin | Affichage texte en ligne |
| Other | N/A | Différents projets étudiants |
| mail | Postfix | Gestion du mail |
| Picasoft association | Nextcloud (principalement) | Groupe de services utilisés par l'association Picasoft |
| CET association | Nextcloud (principalement) | Groupe de services utilisés par l'association CET ^a |
| Solidaire association | Nextcloud (principalement) | Groupe de services utilisés par la section locale du syndicat Solidaires ^b |
| blog | Plume | Plateforme de blogs |
| Administration des serveurs | Groupe de services | Documentation wiki, logiciels d'administration du serveur, etc. |

^a : Association écologiste locale : <https://www.compiegne-en-transition.org/fr>

^b : Section syndicale étudiante : <https://solidaires-etudiant-e-s.org/site/syndicats/>

que la période d’observation totale est suffisamment longue pour que les relevés soient équilibrés entre les services.

Les résultats de l’étude seront toutefois influencés par le contexte strictement français et l’absence de mécanismes de sauvegarde des données. Cela impliquera une empreinte carbone très faible pour I_Net ⁴ et pour l’empreinte OPEX, en raison de la spécificité du mix électrique français, qui est fortement décarboné. Pour cette étude, comme pour l’étude de cas du précédent chapitre, nous avons choisi d’observer les trois catégories suivantes :

- L’épuisement des ressources abiotiques en minéraux et métaux (ADP)⁵, évaluant la contribution à la raréfaction des ressources disponibles due à leur extraction. Cet indicateur est estimé en kilo équivalent d’Antimoine extrait, kgSb eq, suivant la méthodologie suivante [186].
- Le potentiel de réchauffement climatique (GWP)⁶, évaluant la contribution au réchauffement climatique de la planète. Cet indicateur est estimé en kilo équivalent de dioxyde de carbone émis, kgCO₂ eq, suivant la méthodologie suivante [185].
- L’utilisation d’énergie primaire (PE)⁷, évaluant la demande cumulée en énergie primaire. Ce n’est pas à proprement parler une catégorie d’impact environnemental, mais c’est un indicateur pertinent de la demande en ressource. Il est exprimé en mégajoules, MJ, suivant la méthodologie suivante [184].

Nous avons choisi ces catégories pour des raisons de disponibilité des données et parce qu’elles couvrent la majeure partie de l’empreinte des TIC [18]. Le cadre méthodologique de cette étude de cas est résumé par la figure 6.1. Nous avons utilisé les données d’impact environnemental de Boavizta [101] pour I_Host et de Nega-Octet [175] pour I_Net et I_Term . L’impact est ensuite réparti entre les services en fonction des mesures collectées par GoAccess et VictoriaMetrics, comme expliqué dans la section 6.2.1. GoAccess collecte des logs HTTP capables de décrire le nombre de visites, l’origine des visites et le volume de données transférées par service. De son côté, VictoriaMetrics récupère les données de la commande `dockerstats` pour afficher l’utilisation du CPU et de la RAM par conteneur docker sur chaque VM. Ces données sont ensuite agrégées par service pour calculer a_{CPU} et a_{RAM} , a_{Sto} est directement observé avec VictoriaMetrics.

6.3.3 Résultats

La figure 6.2 montre la distribution de l’empreinte entre les services pour chaque catégorie d’impact. Plus spécifiquement, la figure 6.3 montre la relation entre les services et les sources d’impact I_Host_{capex} , I_Host_{opex} et I_Net , pour la catégorie GWP. Ces deux figures nous donnent une estimation de la quantité et de la distribution des impacts, ce qui permet quelques observations. Tout d’abord, sur figure 6.2, nous pouvons voir une certaine homogénéité de l’empreinte de chaque service quelle que soit la catégorie d’impact. Un service se voit attribuer un pourcentage similaire de l’empreinte dans chaque

4. Le processus de sauvegarde, exclue de cette étude, générant un important flux de données comparé à l’ensemble des autres services.

5. Pour *Abiotic Depletion Potential*

6. Pour *Global Warming Potential*

7. Pour *Primary Energy*

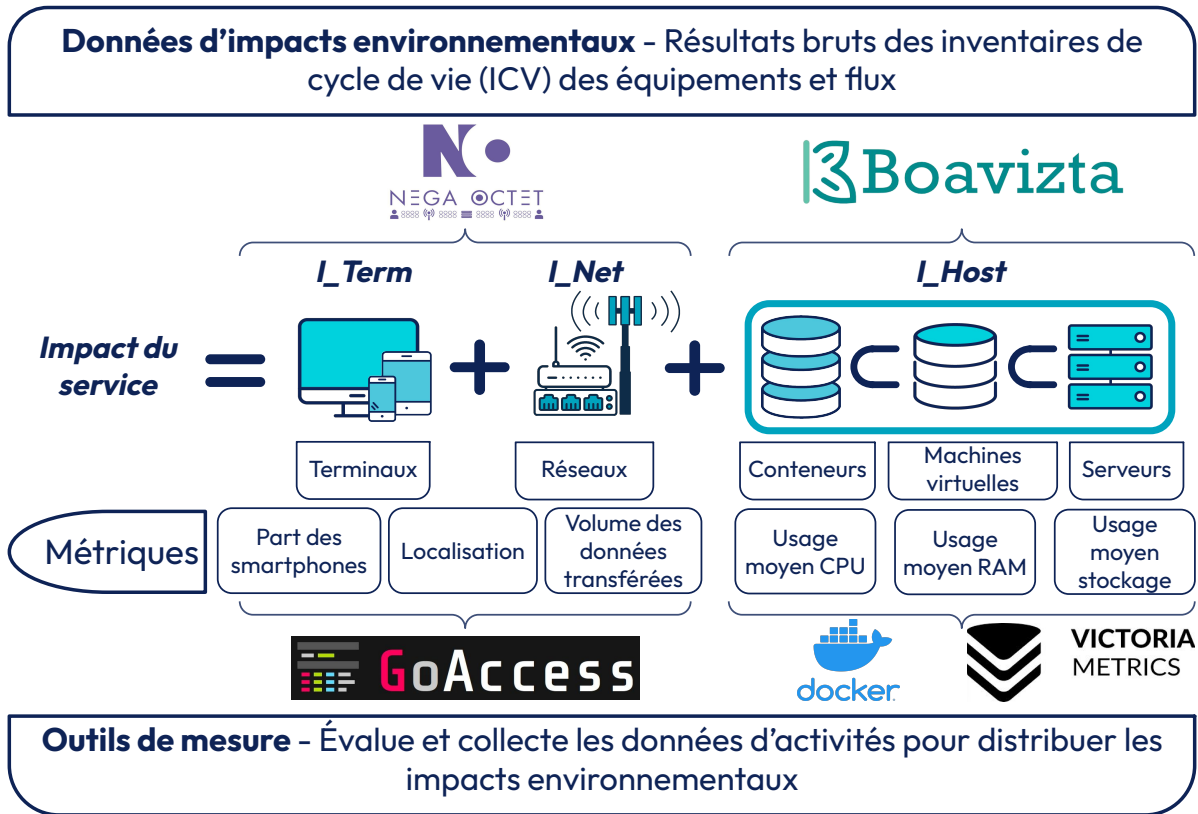


FIGURE 6.1 – Synthèse de la méthodologie, des outils et des données utilisés pour l'ACV de l'ensemble des services numériques de Picasoft

catégorie d'impact. Deuxièmement, comme présagé, dans la figure 6.3 la faible part de l'empreinte des parties $I_{Host_{opex}}$ et I_{Net} peut s'expliquer par le contexte français et le mix électrique fortement décarboné. Si ces premiers résultats nous permettent de rendre compte des impacts environnementaux de chaque service et de l'ensemble, ils ne fournissent pas nécessairement les outils nécessaires à l'application de stratégies de réduction de l'empreinte. Pour cela, nous présentons 3 compléments à la méthode de base afin de mieux prendre en compte la responsabilité de l'impact environnemental entre les différents services et parties des services.

6.4 De nouvelles stratégies de réduction de l'empreinte environnementale des services

Ces stratégies sont des extensions de la méthodologie précédemment proposée. Elles sont destinées à faciliter la mise en place de politiques de réduction de l'empreinte des services. Cette aide est fournie en donnant différentes façons de comparer les services entre eux, pour un gestionnaire, du point de vue de l'hébergement des services. Quel service doit être étudié en priorité d'un point de vue environnemental, qu'il s'agisse de l'optimiser, de réduire ses fonctionnalités, de le limiter ou de l'arrêter ?

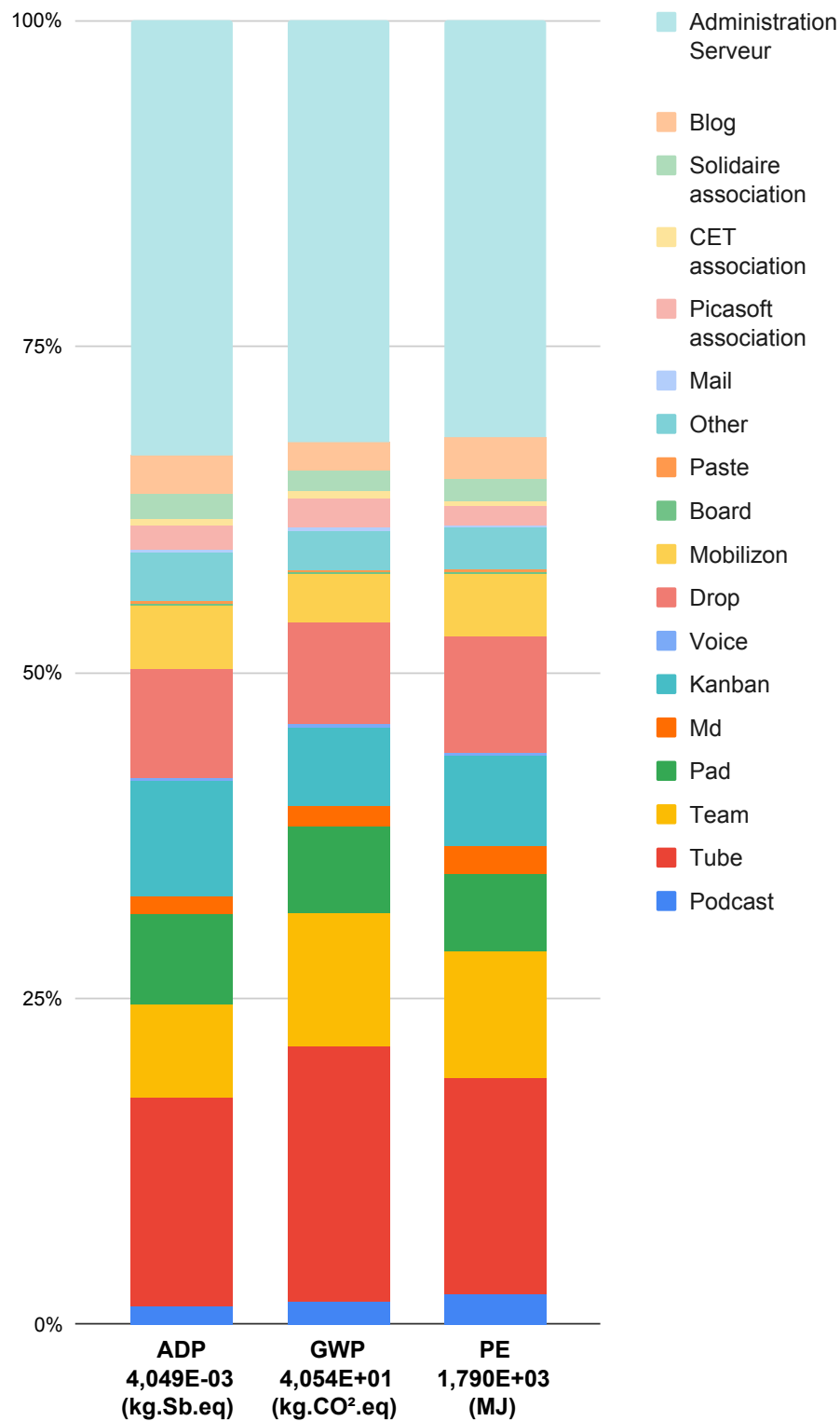


FIGURE 6.2 – Impacts environnementaux caractérisés de l’ensemble des services Picasoft pendant un mois, incluant la partie hébergement et réseaux

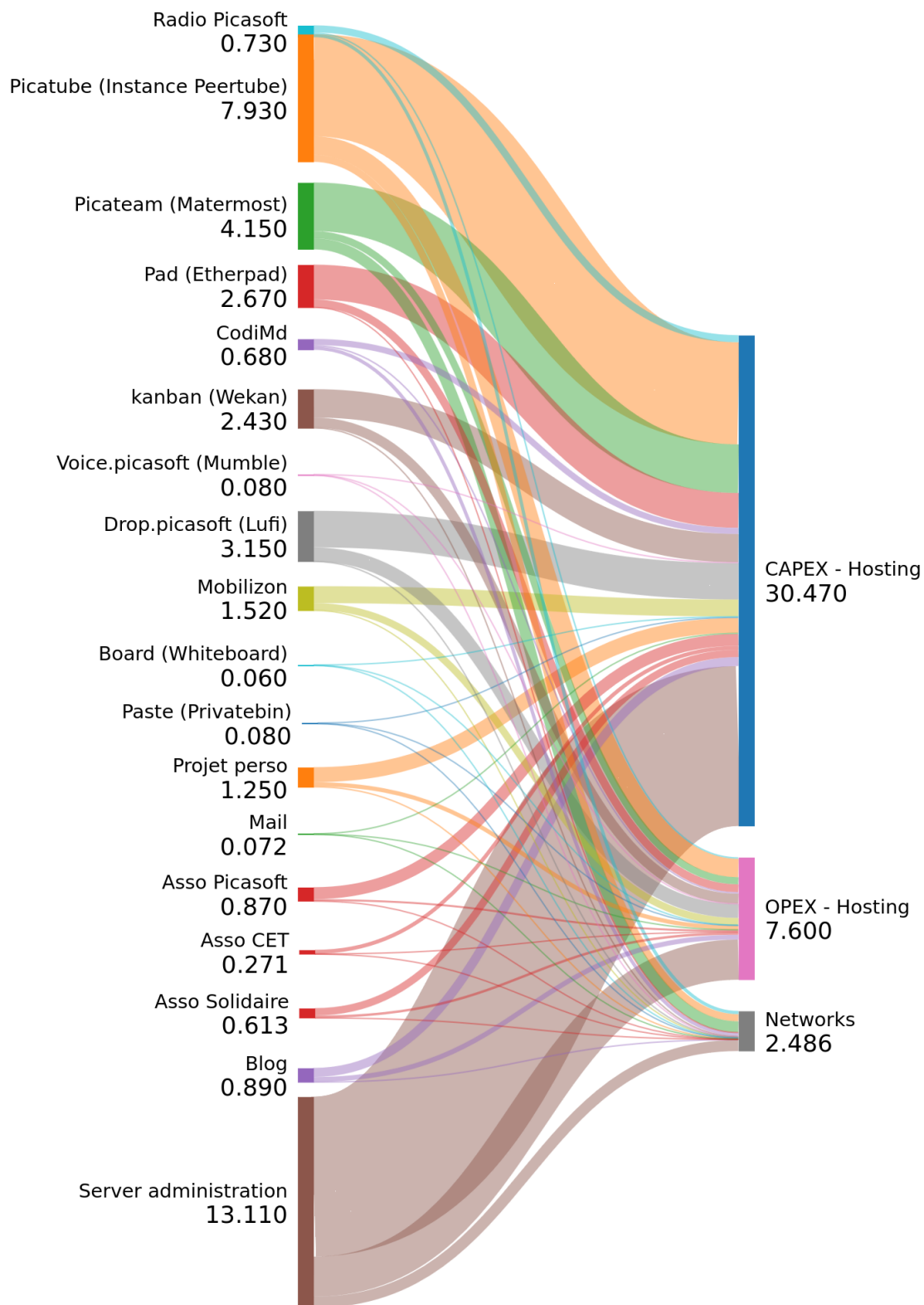


FIGURE 6.3 – Contribution au réchauffement climatique, exprimé en kg eq CO₂, de l'ensemble des services Picasoft pendant un mois, distribué entre trois parties du service

L’étude de cas Picasoft sera utilisée pour illustrer ces contributions, bien qu’il s’agisse d’un cas d’usage à plus petite échelle que d’autres déploiements de solutions SaaS.

6.4.1 Empreinte par visiteur comme indicateur d’efficacité

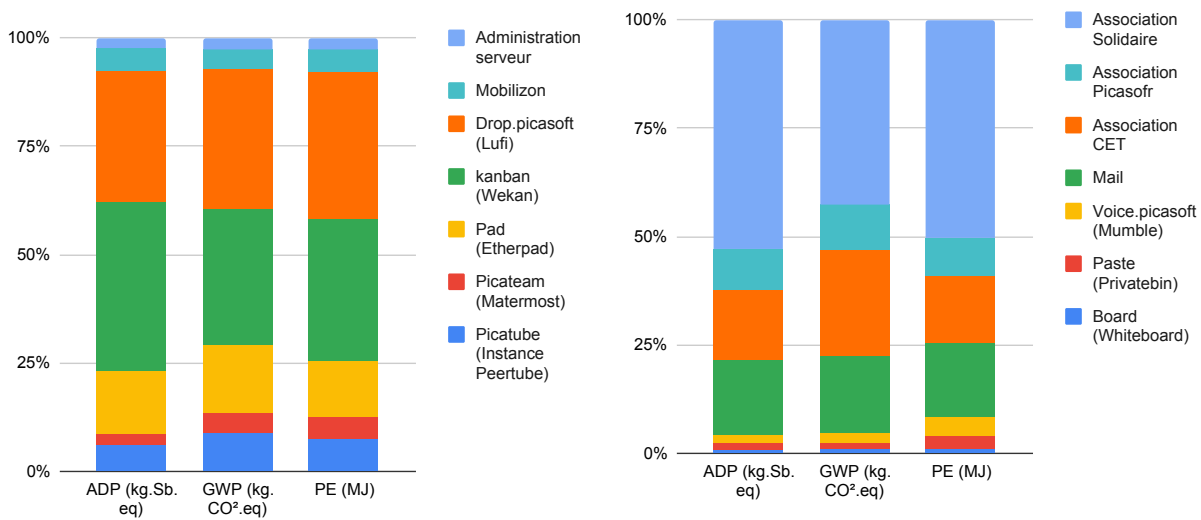
La première approche pour choisir le service à prioritairement considérer, pourrait être de se concentrer sur le service dont l’empreinte est la plus importante. Cependant, un service peut représenter une empreinte importante simplement parce qu’il est très utilisé. À l’inverse, nous proposons de comparer via les figures 6.4b et 6.4a l’empreinte de chaque service divisée par le nombre de visiteurs uniques⁸, par rapport à des services ayant des empreintes similaires. On peut alors constater des disparités en termes d’efficacité entre des services de même ordre de grandeur en termes d’impact.

L’empreinte divisée ici est la même que celle de la figure 6.2, incluant les parties hébergement et réseaux, mais excluant la partie terminaux. On constate que, malgré une empreinte importante, certains services semblent plus efficaces pour gérer un grand nombre de visites. Par exemple, dans le cas de l’administration des serveurs, même si certains logiciels (par exemple Loki ou Traefik) impliquent très peu de visites pour un coût élevé, ce service inclut aussi une documentation technique sous la forme d’un wiki en ligne. Ce wiki semble remplir sa fonction pour un faible coût par visite. À l’inverse, certains services semblent plus coûteux à cet égard. Il serait intéressant de comparer ces services avec d’autres solutions SaaS répondant à des besoins similaires.

6.4.2 Le coût de l’imprévisibilité

Dans la section 6.2.2, nous avons observé l’empreinte CAPEX distribuée selon des clés d’allocation calculées basées sur des moyennes. Si le calcul de la moyenne permet de partager facilement l’utilisation de ressources, telles que la RAM, entre les services, cela peut masquer des problèmes. L’efficacité, dans le contexte de l’empreinte CAPEX de l’hébergement, peut être considérée comme la capacité à effectuer le bon dimensionnement, par exemple quelle taille de VM pour quel service ? L’une des difficultés de cet exercice réside dans le fait que deux services peuvent avoir la même utilisation moyenne de RAM, mais l’un a un besoin stable et l’autre des pics réguliers d’utilisation très élevée. Dans cette situation, le second service pousse à un surdimensionnement des ressources d’hébergement et pourrait être ainsi considéré comme proportionnellement plus responsable que le premier du coût de l’hébergement. Pour représenter ce coût de l’instabilité de la consommation des ressources, nous avons recalculé les allocations, mais en utilisant la moyenne quadratique au lieu des moyennes simples. Ce faisant, nous avons attribué une plus grande part de l’impact aux services dont les rythmes d’utilisation des ressources sont irréguliers. Le tableau 6.2 montre la variation de l’empreinte générée par cette nouvelle distribution de l’empreinte CAPEX de l’hébergement. Les services dont l’utilisation des ressources est stable, et donc dont l’empreinte diminue avec ce nouveau calcul, sont représentés en bleu. Inversement, les services dont l’empreinte est plus importante en raison du nouveau calcul d’allocation sont indiqués en jaune.

8. Un visiteur unique est considéré dans les logs HTTP comme un client doté d’une unique adresse IP et visitant le site au moins une fois. Cette donnée ne prend pas en compte le nombre de visites au sein d’une même journée, ni leur durée.



(a) Distribution de l’empreinte des services représentant ensemble plus de 80% de l’empreinte de Picasoft où l’empreinte de chaque service a été divisée par le nombre de visiteurs associés (sans compter la partie terminaux utilisateurs)

(b) Distribution de l’empreinte des services représentant ensemble moins de 5% de l’empreinte de Picasoft où l’empreinte de chaque service a été divisée par le nombre de visiteurs associés (sans compter la partie terminaux utilisateurs)

FIGURE 6.4 – Empreinte par visiteurs de deux groupes de services

La partie jaune du tableau montre les services les plus influencés par leur intensité d’utilisation, mais pas nécessairement au même degré. En bleu, on voit certains services stables, comme l’administration du serveur, ou avec très peu d’utilisateurs, comme le service *Voice*. Cependant, même les services avec peu d’utilisateurs, comme les associations, montrent de fortes variations dans l’empreinte. Nous pouvons voir que l’impact d’ADP est le plus important, ce qui est normal puisque cette catégorie est principalement influencée par la partie CAPEX dont les clés d’allocation ont été modifiées, comprenant les impacts de fabrication.

6.4.3 Le coût de l’utilisateur : la grande inconnue restante

Si l’importance relative de la part des terminaux d’utilisateurs dans l’empreinte numérique globale est sujette à débat [21, 23], aucune estimation ne considère les terminaux comme une part négligeable, encore plus si l’on considère des évaluations multicritères [18, 176]. Les terminaux des utilisateurs finaux sont un élément indispensable des services numériques. Bien que leur gestion puisse incomber à l’utilisateur plutôt qu’au fournisseur de services, il serait anormal de ne pas prendre en compte l’empreinte des terminaux permettant l’accès aux services numériques. Si l’on veut calculer équitablement l’impact environnemental d’un service, il faut attribuer à ce service une part de l’empreinte de chaque terminal utilisé pour accéder au service.

Ce problème d’allocation est vaste et comporte des limites en termes de méthodologie et d’accès aux données. Un certain nombre d’études ont mis en évidence les incertitudes associées à l’intégration des terminaux d’utilisateurs dans le coût des services [90, 164, 177].

TABLE 6.2 – Variation de l’empreinte CAPEX de l’hébergement due au changement du calcul de distribution de l’empreinte de la méthode par moyenne simple à la méthode par racine quadratique

| Services | Variation pour chaque impact environnemental | | |
|--------------------------------|--|---|----------------------------|
| | <i>ADP</i> (<i>kg.Sb.eq</i>) | <i>GWP</i> (<i>kg.CO₂.eq</i>) | <i>PE</i> (<i>MJ</i>) |
| podcast | 15.42% | 6.35% | 6.93% |
| tube | 18.36% | 6.49% | 7.26% |
| team | 4.76% | 1.34% | 1.62% |
| pad | 19.45% | 10.98% | 11.62% |
| md | -11.86% | -9.21% | -9.55% |
| kanban | -20.54% | -11.57% | -12.25% |
| voice | -32.91% | -20.02% | -20.74% |
| Drop | -41.39% | -15.62% | -17.34% |
| mobilizon | 8.44% | 4.68% | 4.91% |
| board | 6.83% | 1.68% | 1.93% |
| paste | 32.31% | 20.99% | 21.79% |
| other | 9.13% | 5.96% | 6.13% |
| mail | 26.23% | 10.19% | 11.22% |
| Picasoft association | 23.86% | 8.37% | 9.65% |
| CET association | 7.60% | 1.93% | 2.37% |
| Solidaire association | -35.76% | -23.47% | -24.21% |
| blog | 6.92% | 5.27% | 5.37% |
| Administration des serveurs | -9.88% | -4.15% | -4.61% |

Comme le montre la section 6.2.3, $I_Term(S)$ dépend de la durée totale d'utilisation du terminal, c'est-à-dire que $I_Term(S)$ dépend à la fois de la durée de vie de l'équipement et de son utilisation moyenne par l'utilisateur. Ainsi, $I_Term(S)$ dépend à la fois du comportement de l'utilisateur pendant l'utilisation du service et en dehors de celle-ci.

Une dernière limite au calcul de l'empreinte des terminaux est l'hétérogénéité des services et des terminaux. Ainsi, l'utilisation d'un service peut induire une consommation d'énergie supplémentaire importante sur un terminal, la taille de cette consommation dépendant de la charge de travail du service côté client et l'importance relative de cette surconsommation dépendant de la consommation du terminal.

Le problème fondamental de cette hétérogénéité reste cependant le critère d'allocation, qui est actuellement par défaut : le temps. On peut allouer 3 heures de temps du terminal à un service vidéo qui a permis 3 heures de streaming, en supposant ainsi que le terminal soit entièrement mobilisé pendant ces 3 heures. Cependant, pour un service de messagerie instantanée fonctionnant toute la journée en arrière-plan, ou un service de podcast audio, cette allocation basée sur le temps perd de son sens. En effet, il ne suffit pas de savoir pendant combien de temps l'utilisateur utilise le service via son terminal pour attribuer une part équitable de l'empreinte à chaque service. En termes purement hypothétiques, il faudrait disposer du même degré d'information sur les terminaux que sur les serveurs en nuage, ce qui n'est pas possible dans ce contexte.

Cependant, dans la figure 6.5, nous avons choisi de représenter le coût carbone des terminaux en fonction des différents services. Nous avons fait l'hypothèse volontairement simpliste de représenter chaque visiteur par une unique minute d'utilisation d'un terminal utilisateur, quel que soit le service. Pour rappel, le terme "visite" fait référence à un "visiteur quotidien unique", qui peut avoir effectué plusieurs visites au service au cours d'une journée, en utilisant le même équipement, il s'agit donc d'une estimation très basse. Malgré les nombreuses limites que nous avons mentionnées, nous pensons qu'il est nécessaire de représenter et de prendre en compte l'impact des terminaux, tout en sachant que nos données sont limitées. En comparant les figures 6.3 et 6.5, il faut noter que même avec une hypothèse minimaliste sur l'utilisation des terminaux, leur impact carbone serait considérable. Il est également intéressant de noter que la prise en compte de l'impact des terminaux permet de mieux apprécier l'empreinte de certains services qui ne semblaient pas si coûteux. Le service de blog, par exemple, n'a jusqu'à présent pas fait l'objet de discussions, alors qu'il génère une utilisation importante des terminaux.

Comme pour les sous-sections 6.4.1 et 6.4.2, l'objectif ici n'est pas de quantifier l'impact exact, mais d'ajouter un autre niveau de compréhension de l'empreinte du service. Il est important d'estimer la taille de l'empreinte des terminaux entre les services afin de hiérarchiser les stratégies de réduction. Même s'il est difficile, du point de vue du gestionnaire de service, d'avoir un impact direct sur l'empreinte de la partie terminale, ce n'est pas impossible. L'éco-conception des services, leur capacité à être utilisés par le plus grand nombre de terminaux, par exemple en proposant des versions allégées des services, pourrait permettre de réduire la consommation d'électricité côté client du service. De plus, de tels efforts pourraient contribuer à ralentir le renouvellement des terminaux [304], à allonger leur durée de vie moyenne, à réduire leur empreinte et in fine celle des services associés.

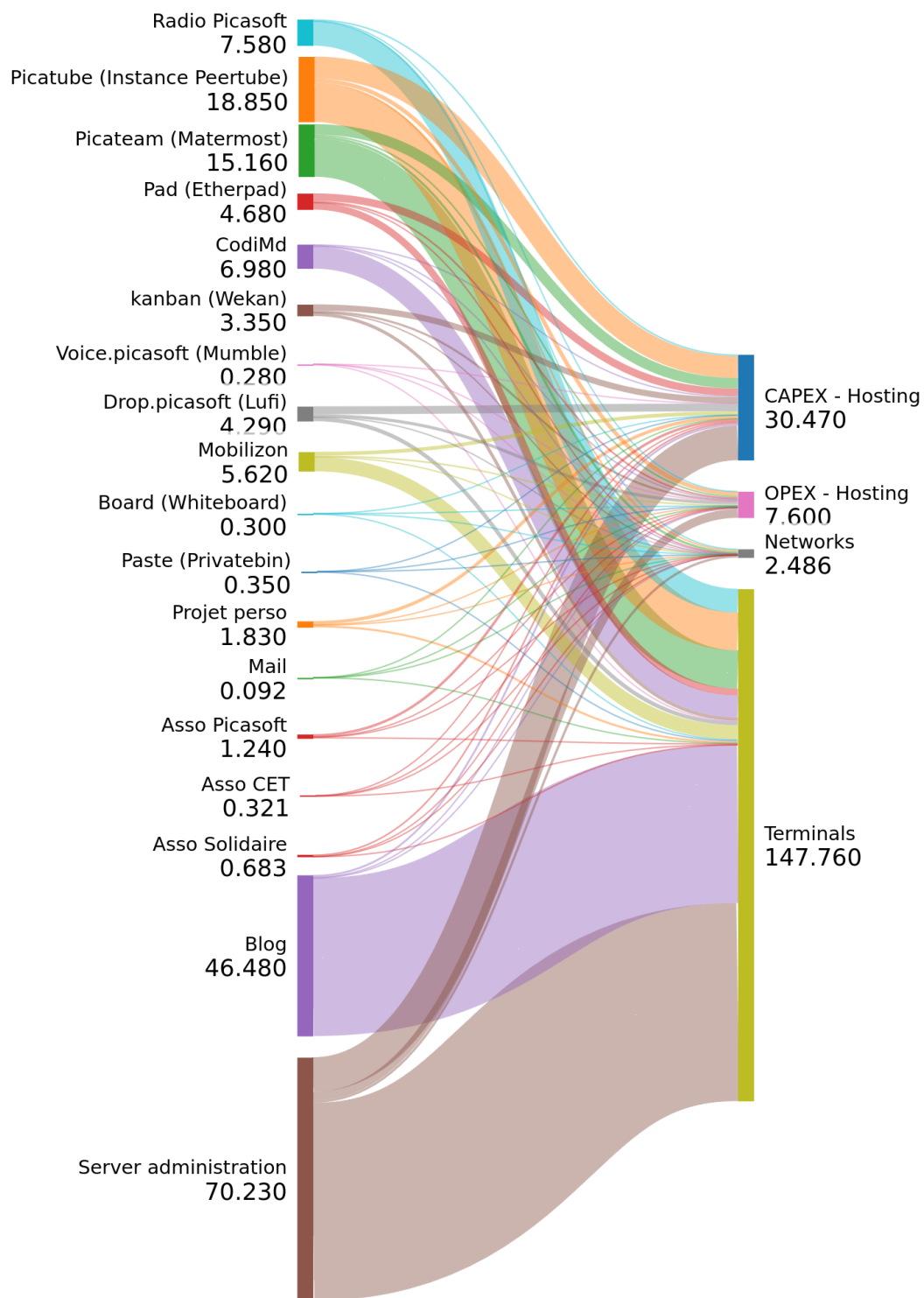


FIGURE 6.5 – Impact carbone, en kg eq CO₂, des services Picasoft pour un mois incluant les impacts des terminaux utilisateurs avec l’hypothèse d’une visite d’une minute par visiteur unique

6.5 Limites

La principale limite de cette étude, outre l'inclusion des terminaux évoquée ci-dessus, est l'absence de méthodologies préexistantes permettant une comparaison significative. Le cadre de cette méthodologie, à savoir l'évaluation multicritère d'un ensemble de services hébergés sur les mêmes serveurs, reste très rare. D'autres études et méthodologies existent sur la notion d'empreinte des services, mais elles sont souvent limitées en termes de catégories d'impact et ne sont pas aussi détaillées, se concentrant sur un seul service [178]. Les auteurs arrivent aux mêmes conclusions que [178] sur le manque de littérature scientifique transparente sur l'empreinte des services, et sur les avantages de l'ACV en tant qu'approche pour combler ces lacunes.

Une autre limite à la distribution fine des impacts proposée dans la méthodologie est l'allocation faite pour le stockage. Le problème consiste à attribuer à chaque service une part de l'empreinte de l'équipement de stockage. Cette part est bien sûr proportionnelle à la durée de l'étude et à la durée de vie du disque de stockage, mais une fois encore se pose le problème du partage de l'équipement entre les services. Comment évaluer la proportion des équipements de stockage par les différents services ? L'approche naïve choisie dans cet article est de regarder la part moyenne d'occupation des disques par rapport aux autres services. On suppose que le critère principal de dimensionnement du stockage, et donc de son encombrement, est sa capacité en termes de taille. Cependant, ce critère peut être remis en cause car la question de la fréquence d'accès a pu être déterminante. Finalement, le service qui "utilise" le plus le disque est-il celui qui a le plus grand volume de données stockées ou celui qui effectue le plus d'opérations de lecture et d'écriture ?

Ce document se concentre sur la distribution des impacts environnementaux, mais il dépend clairement des données d'entrée comme I_{capex} ou I_{opex} . Bien que des progrès aient été réalisés pour fournir une vue plus détaillée de l'empreinte OPEX [225, 305], c'est-à-dire de la consommation d'électricité, la qualité des données I_{capex} et leur continuité dans le temps est encore insatisfaisante⁹. À noter d'ailleurs, que cette étude ACV utilise deux bases ICV différentes, ce qui l'expose à de potentiels problèmes de cohérence des résultats au regard des données d'entrée. Ce choix est justifiable par le fait que l'étude ACV se concentre prioritairement sur la distribution des impacts, et non les résultats quantitatifs finaux, et par le fait que ces deux bases ICV s'appuient sur des périodes temporelles similaires, entre 2019 et 2022.

6.6 Discussions

La principale question de recherche qui a initié ces travaux est de savoir comment calculer l'empreinte de bout en bout d'un service numérique. En adoptant le point de vue d'un gestionnaire de service, nous avons détaillé la partie hébergement et choisi des modèles simplifiés pour les autres parties, voire incomplets pour la partie terminaux utilisateurs. Des approches plus détaillées pour la partie réseaux devraient être intégrées [102, 103], mais elles nécessitent un niveau de connaissance des infrastructures de télécommunication difficile à obtenir. De plus, en dehors de services spécifiques, la partie réseaux a tendance à avoir une empreinte plus faible. Enfin, cette partie offre potentiellement peu de leviers

9. Comme évoqué plusieurs fois dans ce manuscrit, section 3.3.1 et 4.1.2

d'éco-conception. Il n'est pas possible de modifier l'infrastructure mais seulement l'usage qui en est fait, avec de plus une connaissance limitée de son fonctionnement direct.

Les terminaux, en revanche, représentent une part importante de l'empreinte et pourraient donc, au prix d'un changement de comportement de l'utilisateur [177, 306], révéler des leviers cruciaux pour la réduction de l'empreinte des services. Ce comportement est largement influencé par l'implémentation et le design du service sur le terminal de l'utilisateur et influence en retour l'utilisation des équipements côté hébergeur [307]. Afin de mieux comprendre ces leviers, il est donc important de mieux comprendre ou modéliser l'utilisation faite des terminaux dans le contexte d'un service étudié.

Dans ce chapitre, nous nous sommes basés sur des données de première main : les journaux de requêtes HTTP. Ces données sont fiables et fournissent des informations pertinentes sur la proportion d'utilisateurs de smartphones, mais ne fournissent pas plus d'informations sur les sessions utilisateurs par service, par exemple la durée moyenne, le nombre de pages visitées, etc. À plus grande échelle, il peut être possible d'utiliser des outils d'analyse plus larges, tels que Similarweb [248] ou HypeStat [247], mais au prix d'une perte de précision et de fiabilité [249, 250]. Plutôt que la mesure directe, la simulation pourrait être un axe de recherche pertinent. La simulation d'une population d'utilisateurs et de leur utilisation, de manière similaire à [308], pourrait permettre de tester différentes implémentations du service, tant du côté du serveur que du côté de l'utilisateur. Cela pourrait contribuer à améliorer certains projets de modélisation de l'empreinte des services, tels que e-footprint [309].

6.7 Perspectives

Naturellement, une partie des travaux futurs portera sur l'extension de la méthodologie afin d'inclure davantage de catégories d'impacts environnementaux, une plus grande précision sur la consommation des équipements et un calcul plus détaillé de l'empreinte des terminaux. L'idéal serait de disposer d'un outil automatique pouvant être intégré dans un panneau de contrôle typique des administrateurs de services. Un tel outil permettrait de poursuivre l'étude de l'empreinte des services sur la base d'un plus grand nombre d'expériences en conditions réelles, ou au moins en simulation.

En dehors de cette méthodologie générale, il est apparu aux auteurs que la fonction de stockage, notamment en ce qui concerne les procédures de sauvegarde et de réplication, devrait être traitée séparément. Il serait intéressant d'étudier le stockage en tant que service, en utilisant et en adaptant notre méthodologie d'ACV aux services numériques. Il s'agit d'un domaine de recherche actif [310, 311] qu'il serait intéressant d'observer du point de vue de son coût environnemental. Le but pourrait être de produire un facteur d'impact fiable à la manière de ce qui est fait pour calculer l'empreinte de la partie réseaux.

Enfin, les services fournis par Picasoft sont des équivalents libres et gratuits des suites bureautiques bien connues des grandes entreprises. Il serait intéressant de comparer les coûts environnementaux de ces différentes solutions, fonctionnellement similaires mais dont les implémentations et les échelles sont radicalement différentes.

6.8 Conclusion

Dans cet article, nous avons présenté notre méthodologie pour attribuer un coût environnemental complet à un service hébergé dans le Cloud. L'une des principales contributions de notre méthodologie est la répartition de l'empreinte environnementale de l'hébergement entre différents services numériques hétérogènes. Cette répartition est basée sur l'utilisation par chaque service du processeur, de la mémoire vive et des ressources de stockage de l'hébergement. De cette manière, notre méthodologie permet d'attribuer à chaque service une part de l'empreinte OPEX et CAPEX de l'hébergement. En outre, nous avons proposé des méthodes pour inclure les impacts réseau et terminal des services. Nous avons ainsi produit une méthode de calcul du coût environnemental de chaque service, sur l'ensemble du cycle de vie et sur une base multicritère, dépassant les limites d'évaluation actuelles.

Pour illustrer la méthodologie et démontrer son applicabilité, nous avons produit une étude de cas sur le fournisseur de services numériques Picasoft. Nous avons pu observer tous les services en conditions réelles et établir une évaluation après un mois. Cette évaluation comprend les impacts environnementaux dans trois catégories, pour le fournisseur et pour chaque service. Outre cette capacité à rendre compte de l'impact de chaque service, le détail de l'évaluation nous a permis de reconsidérer l'empreinte de chaque service par rapport aux autres sous différents angles.

Nous avons donc proposé 3 nouvelles perspectives pour reconsidérer l'empreinte des services basés sur le cloud. Ces perspectives ont pour but d'aider à hiérarchiser les efforts entre les services dans le cadre d'une politique de réduction de l'empreinte. Quels services doivent être améliorés, réglementés ou fermés, en priorité, pour atteindre les objectifs environnementaux du fournisseur de services cloud ?

L'étape suivante et la principale limite de notre approche actuelle reste l'empreinte des terminaux des utilisateurs. Comment inclure, du point de vue du fournisseur de services, observant et gérant depuis le nuage, cette partie de l'impact du service qui peut être la principale ? Nous avons proposé une première approche, mais elle doit être améliorée pour s'adapter à l'hétérogénéité des services. Il faudra peut-être aller au-delà de l'allocation temporelle pour adopter une approche plus holistique. Cette prise en compte des terminaux permettra alors d'obtenir une estimation de bout en bout du coût environnemental de chaque service. La répétition de cas d'utilisation tels que le nôtre permettra ensuite de construire des modèles de prédiction d'impact fiables. Ainsi, grâce à l'observation fine et continue de l'impact environnemental des services, les gestionnaires de services basés sur le cloud pourront s'appuyer sur de nouvelles métriques fiables pour orienter les activités du cloud vers la durabilité.

Conclusion

La valeur ne porte donc pas écrit sur le front ce qu'elle est.

Karl Marx

Le titre de cette thèse renvoie à la question simple du coût environnemental d'un usage du numérique. Des questions comme "Quelle est l'empreinte d'un email?" ou "Les datacenters consomment-ils beaucoup?" parsèment la vie des personnes travaillant sur les impacts environnementaux du numérique. L'usage du numérique se fait majoritairement aujourd'hui au travers de services. Pour répondre à ces questions simples, mais si importantes, cette thèse a alors pris le parti d'étudier les différentes approches et a convergé vers l'utilisation l'ACV de services numériques comme principal outil et sujet de recherche afin d'étendre notre connaissance de l'empreinte de ces usages du numérique. Ce chapitre propose donc un retour sur les contributions apportées dans ce manuscrit, par et pour l'ACV de services numériques, ainsi qu'une discussion plus large sur l'empreinte environnementale du numérique et les enjeux de son évaluation. Enfin, ce manuscrit se clôturera par un ensemble de perspectives possibles à nos travaux de recherche.

7.1 Conclusion générale

Bien qu'il reste disciplinairement ancré en informatique, le travail de cette thèse se situe à l'intersection entre, d'une part, la communauté de recherche en informatique cherchant à mieux comprendre et réduire les impacts environnementaux du numérique, et d'autre part, la communauté de recherche issue de l'évaluation environnementale et de l'écologie industrielle. Le chapitre 2 a illustré, sous la forme d'une revue chronologique de l'état de l'art, la progressive convergence entre ces deux communautés de recherche qui a donné lieu à des travaux comme ceux que nous avons menés, que l'on pourrait inscrire dans de l'informatique soutenable [32]. Cette convergence autour de la recherche sur les moyens de mesurer, comprendre et réduire les impacts environnementaux du numérique amène à un usage de l'ACV de services numériques.

Le chapitre 3 s'est concentré sur le sujet spécifique de l'ACV de services numériques

par une revue de l'état de l'art. En effet, cette méthode possède différentes qualités pour l'informatique soutenable, au-delà de la recherche académique. Pour les résumer, une capacité d'évaluation quantitative plus exhaustive des impacts, sur les différentes phases du cycle de vie, les différentes catégories d'impacts environnementaux et les différentes parties du service numérique. Ces qualités créent un potentiel pour améliorer l'éco-conception des services numériques, réduire l'empreinte du numérique au sein des organisations ou encore permettent une évaluation des services numériques homogène avec d'autres types de biens et services. Ces qualités sont néanmoins à nuancer au regard des limites examinées. Le chapitre fait en effet le constat d'un manque global de maturité scientifique du domaine, dû majoritairement au manque de contenu, outils et résultats, accessibles. Ce manque de maturité est illustré par une revue critique de quatre ACV de services numériques parmi les rares accessibles ne serait-ce que partiellement. Les limites de ces travaux sont présentées et, en conséquence de cette revue, une analyse est présentée. L'analyse propose premièrement une aide à la lecture des rapports d'ACV de services numériques, deuxièmement un modèle simplifié de calcul d'impacts environnementaux qui peut servir à mieux comprendre ou communiquer sur un rapport d'ACV de services numériques en mettant en valeur des indicateurs d'optimisation du service à éco-concevoir ou critiquer. Le chapitre se termine au regard du constat sur le manque de maturité de l'ACV de services numériques par une volonté d'orienter la thèse autour de trois types de contributions :

- Contribution méthodologique à l'ACV : adaptation de l'ACV à des cas généraux et spécifiques de services numériques
- Contribution méthodologique à l'éco-conception de services numériques : proposition de grilles de lecture pour la conception et évaluation de services numériques selon des critères environnementaux
- Contribution expérimentale : application des contributions méthodologiques précédentes à des services existants, en fonctionnement

Une fois les motivations explicitées, les contributions des chapitres suivants s'expriment au sein de ces trois catégories de la manière qui suit.

Contribution méthodologique à l'ACV de services numériques

Le chapitre 4 offre une synthèse méthodologique de l'ACV de services numériques. Ce cadre généraliste pour l'ACV de services numériques offre un ensemble de recommandations ainsi qu'un d'état de l'art des outils disponibles afin d'aider les personnes souhaitant réaliser des ACV de services numériques. Pour la partie inventaire des cycles de vie, un modèle simplifié des coûts environnementaux séparés entre OPEX et CAPEX est proposé. La méthodologie de ce chapitre sert de guide aux études de cas des chapitres suivants. Dans le chapitre 5, la méthodologie est adaptée, une méthode spécifique pour les services numériques basés sur l'IA générative est proposée, intégrant par exemple la spécificité du coût d'apprentissage du modèle d'IA. Cette méthode a comme limite d'être faite pour l'évaluation d'un unique service opéré par du matériel spécifique, là où beaucoup de services sont proposés par groupe, hébergés par un même acteur sur des serveurs mutualisés. Le chapitre 6 dépasse cette limite en proposant une méthode d'évaluation d'un ensemble de services, tout en évaluant l'empreinte individuelle de chaque service. Ce chapitre propose une méthode plus précise qui permet de finement distribuer l'empreinte

de l'hébergement aux différents services selon leur utilisation des ressources des serveurs qu'ils partagent.

Contribution méthodologique à l'éco-conception de services numériques

Comme exposé dans le chapitre 2, il existe un long passif de recherches sur les moyens de réduire l'empreinte environnementale du numérique. En conséquent, les outils capables de mesurer cette empreinte sont au centre de cette thématique. Le chapitre 3 expose plusieurs usages de l'ACV de service numérique réalisés dans un but d'éco-conception des services évalués. Le chapitre expose les limites actuelles de telles approches et propose, avec le chapitre 4, un cadre méthodologique pour l'ACV de services numériques. Parmi les différentes caractéristiques de l'ACV listées dans le chapitre 2, la plus critique à intégrer pour contribuer à l'éco-conception des services numériques est la perspective en cycle de vie. Pouvoir intégrer des impacts environnementaux qui ne sont pas issus de la phase d'usage du service numérique est cruciale. C'est pour cela que la méthode ACV proposée se focalise sur la différence entre coûts environnementaux OPEX et CAPEX. L'application de cette méthode à des cas d'étude aux chapitres 5 et 6 permet de mettre en valeur de nouveaux indicateurs spécifiques pour l'éco-conception des services. L'étude sur les services d'IA générative met en valeur les enjeux du taux moyens d'utilisation des serveurs disponibles à la demande ou l'intégration du coût des différentes phases d'entraînement dans l'empreinte du service. L'étude sur les services Picasoft propose d'observer le coût des services par rapport aux nombres de visites qu'ils gèrent ou encore comment mettre en valeur les services ayant des pics d'utilisation de ressources entraînant un surdimensionnement des ressources. Toutes ces connaissances méthodologiques permettent de mieux évaluer les potentiels gains d'éco-conception au regard de l'empreinte OPEX comme CAPEX de tout le service.

Contribution expérimentale

Le travail de thèse a mené à l'application de la méthode de bout en bout sur plusieurs cas d'étude contribuant à enrichir l'état de l'art de résultats quantitatifs en cycle de vie complet. Le chapitre 5 a notamment mis en valeur le coût environnemental certains des services basés sur l'IA générative, domaine dont la popularité a explosé ces dernières années. Cette estimation a permis de consolider les analyses faites sur le sujet quant à l'empreinte de telles technologies. Le chapitre 6 a lui contribué avec une étude sur l'empreinte d'un hébergeur locale de services numériques. Ce travail pourra servir de premier indicateur de comparaison pour d'autres hébergeurs locaux de services numériques et aussi de base de comparaison si un travail sur l'empreinte des suites bureautiques *as a service* venait à être accompli. Ce travail de thèse a permis aussi la production d'autres résultats d'ACV de services numériques [231] comme détaillé en annexe 7.3.

7.2 Discussion : de la mesure des impacts environnementaux

Dans un mouvement inverse à celui opéré pendant l'introduction, en desserrant la focale des ACV de services numériques, cette section de discussion se portera sur le sujet plus

large de la mesure des impacts environnementaux et ses implications pour le numérique. Tout d'abord : pourquoi avons-nous besoin de la mesure des impacts environnementaux en premier lieu ? Une réponse pourrait être qu'il s'agit là d'un outil, permettant d'apposer une valeur sur un bien ou service, une valeur ici généralement négative, un coût, mais une valeur qui sert alors à organiser l'activité économique, plus spécifiquement la production et gestion de la technosphère.

L'évaluation environnementale dépasse les enjeux environnementaux singuliers comme le réchauffement climatique ou la dégradation de la couche d'ozone. Comme vu à l'époque de Kapp dans les années 50 [9] celui-ci citait déjà de nombreuses problématiques, dont bon nombre aujourd'hui seraient désignées comme des problèmes environnementaux, mais qui étaient désignées comme des problématiques sociales. Car dans son ouvrage "Les coûts sociaux de l'entreprise privée" comme dans le domaine de l'évaluation environnementale, s'opèrent une démarche de valuation des biens et services différente de celle du système économique standard basé sur la valeur marchande.

Le débat sur la notion de "valeur" en économie est vaste [312], et, par son existence, l'évaluation environnementale questionne donc les valeurs marchandes établies qui pilotent principalement la production et gestion de la technosphère. La correspondance entre les valeurs produites par l'évaluation environnementale et les valeurs économiques classiques est un point de débat critique. Est-il possible d'avoir ce qui est appelé un découplage entre production de valeur économique et production de valeur négative, de coût environnemental [313] ? Le fait même d'associer coût économique et coût environnemental relève même pour certains d'un cadre dit "de soutenabilité faible" [314] où les ressources naturelles sont abondantes et surtout substituables par des ressources artificiellement produites.

Indépendamment de ces questions, la correspondance entre ces valeurs est souvent en pratique difficile. Les deux systèmes évoluent avec une certaine indépendance, dans le meilleur des cas en parallèle, mais parfois en opposition. Les coûts environnementaux semblent être établis et discutés toujours avec un retard par rapport au coût monétaire et c'est particulièrement vrai dans le secteur numérique. Le rythme d'innovation plus soutenu du numérique est souvent mis en avant comme l'une des difficultés d'intégration de celui-ci dans l'évaluation environnementale. Il y a aussi le problème du périmètre, le numérique n'étant pas toujours représenté comme un secteur distinct, mais parfois comme l'amélioration interne de secteurs d'industries.

Il y a cependant aussi un frein idéologique à l'intégration d'un coût environnemental, en addition ou en opposition, à un coût économique que le numérique amplifie. La marchandisation permettant d'attribuer une valeur économique, au sens marxiste du terme, est précisément la décorrélation, la décontextualisation, d'un objet par rapport à ses conditions de création et d'utilisation [315]. L'effacement de ses conditions d'extraction des matières premières et de ses conditions de production derrière la notion de valeur marchande, de prix monétaire, est ainsi le processus inverse proposé par l'évaluation environnementale. Ce phénomène est amplifié par le numérique, comme nous avons pu le voir dans le chapitre 2, avec le concept de "dématérialisation". Dans un tel contexte d'opposition, il est pertinent de poser la question de l'utilité de la mesure environnementale et de l'ACV pour les services numériques.

7.2.1 Que faire de l'ACV

Est-ce que l'ACV a la capacité, en l'état actuel, de rendre environnementalement plus vertueux les services numériques ? La contribution de l'ACV attributionnelle de services numériques, telle que pratiquée dans ce manuscrit, à l'émergence de futurs environnementalement viables peut se résumer en deux principaux points.

Un premier est l'aide à l'éco-conception de services numériques. Si la mesure fine des impacts environnementaux d'un service numérique peut aider à réduire son empreinte, dans le registre de l'éco-conception, la limite est mince entre tendre vers une forme de sobriété numérique [128, 316] et tendre vers une recherche de l'efficacité environnementale, de l'optimisation menant potentiellement à des effets rebonds [123, 126], répétant les limites du *carbon-aware computing*¹.

Deuxièmement, l'ACV de services numériques concerne plus largement la communication sur les impacts de ces services. Cette communication peut servir à de l'affichage environnemental [189], espérant ainsi créer une incitation à la consommation de services aux l'impact environnementaux moindres. Ces résultats peuvent servir à comparer des services, éclairer des prises de décisions, même si pour cela, une ACV conséquentielle pourrait être mieux adaptée. Enfin, en lien avec les précédents cas, l'ACV de services numériques permet de mieux insérer le numérique dans une forme de comptabilité environnementale [317]. Cette comptabilité, comme la classique, a pour but d'éclairer la prise de décision, de rendre compte de l'état actuel, de la solvabilité, et de garder trace des décisions. Il est alors important de rappeler que les modèles ne sont pas faits pour prendre des décisions, mais bien éclairer des choix. Dans le cadre des services numériques, au-delà des choix de conceptions, c'est aussi les usages, des modes de vies avec le numérique, qui sont à questionner, comme le titre [316]. Dans l'introduction, nous avons vu, au travers de la figure 1.3 et de ses quatre scénarios, qu'il s'agit bien plus qu'un enjeu technique, il y a un enjeu profondément politique. Les modèles d'impact environnementaux comme l'ACV de services numériques pourront contribuer à éclairer de telles décisions politiques, révélant à quel point l'enjeu de communication autour des ACV est critique.

Pour répondre aux exigences de communication autour des ACV, il est important de constater les limites sur cet enjeu. Il est notable que les ACV sont souvent réalisées avec une posture d'expert isolé, ne facilitant pas dans son processus la collaboration ouverte avec les non-experts, concernés et parties prenantes, comme le souligne [318]. Ce fonctionnement pourrait être amélioré, comme le suggère [35], afin d'être mieux adapté à une prise de décision collective plus en phase avec les principes de soutenabilité [319] et nécessaires à une prise de décision politique qui se voudrait plus démocratique.

Cette ouverture de l'ACV à un processus plus collectif, entrerait en résonance avec des enjeux plus actuels de l'écologie comme : les communs négatifs [320], l'organisation du renoncement à certaines technologies [321] ou encore la redirection écologique. Soutenir scientifiquement ces concepts est un usage futur possible de l'ACV. Dans [321], l'auteur explique que l'enquête, sur les enjeux environnementaux d'une technologie, peut être une forme de commun. Cette phrase fait écho au travail mené dans le chapitre 5 qui a été produit sans participation des acteurs opérant le service d'IA générative, mais qui a permis en étant réalisé comme une enquête d'un point de vue externe, d'alimenter le débat scientifique et sociétal autour des impacts de ces technologies.

1. Voir chapitre 2.

7.2.2 Que faire de l'évaluation environnementale

La quantification est présentée comme l'une des caractéristiques centrales de la méthode ACV [131]. Plus généralement, la quantification s'impose comme un prérequis de l'évaluation environnementale, s'insérant dans une dynamique où l'aide à la décision se fait sur la base de données chiffrées [322]. Il est donc important de conclure cette discussion sur les limites de l'évaluation et quantification des impacts environnementaux du numérique. Comme relevé au début de la section 7.2, la quantification des impacts environnementaux se déroule souvent dans un contexte antagoniste avec la quantification économique. La quantification alternative que représente l'évaluation environnementale peut rencontrer une opposition frontale comme la difficulté d'accès aux données, le manque de standards communs [81], le transfert d'impacts [323] ou encore l'inconsistance de périmètres. Toutefois, l'évaluation environnementale peut rencontrer une opposition plus insidieuse, évoquée via le concept de "soutenabilité faible" [314]. Le concept de soutenabilité faible repose sur plusieurs hypothèses [324], notamment celle de la "substitution des capitaux". Dans ce paradigme, les capitaux "naturels", e.g. des écosystèmes forestiers ou océaniques, peuvent voir leurs services écosystémiques remplacés, substitués, par des capitaux techniques comme des équipements ou des infrastructures techniques. En simplifiant, l'on pourrait alors donner un prix aux capitaux naturels équivalent à celui de la solution de substitution technologique jugée satisfaisante. Entre autres problèmes, l'association de coûts environnementaux et financiers peut créer des situations où les failles du modèle de coûts environnementaux sont exploitées par les bénéficiaires d'un modèle de coûts financiers s'adaptant impitoyablement plus vite que le premier [325]. Cette asymétrie est un risque encore plus présent pour le numérique compte tenu de la capacité d'adaptation rapide de ce secteur [326]. Pour finir, le risque principal de la quantification, comme le conclut [327], est que, plutôt que d'aider à la prise de décisions politiques de manière démocratique, la quantification produise une dépolitisation de la décision et un accaparement de la question par les experts, à rebours du potentiel espéré pour l'ACV à la fin de la section 7.2.1.

Une dernière réflexion, en lien avec l'introduction de ce manuscrit, la section 1.2 a amené le concept de *Pharmakon*, plus spécifiquement l'usage qu'en fait Stiegler². Ce concept pose un cadre où toute technologie numérique a des effets bénéfiques et des effets indésirables. La mesure des impacts environnementaux du numérique peut facilement s'inscrire dans un tel cadre, où d'une part serait mesuré via l'ACV les impacts environnementaux du service numérique et d'autre part les impacts "évités" grâce à ce service. Le problème du "juste dosage" posé par le *Pharmakon* a tendance à insister sur le problème des conditions d'usage des technologies plus que sur les conditions et motivations de la production des technologies [328]. Cette posture peut mener à une certaine forme de relativisme qui pourrait éloigner des questionnements plus globaux sur le numérique. Est-ce que le numérique est par essence un processus écocidaire [329] ? Est-ce que l'ensemble du numérique est une "technologie zombie"³ [330], ou existe-t-il une place pour un numérique soutenable à l'intérieur des limites planétaires ?

2. Défini tel que : <https://arsindustrialis.org/pharmakon>.

3. Terme issu du physicien belge José Halloy désignant une technologie surconsommatrice de ressources donc non soutenable sur le long terme, faisant d'elle une technologie sans avenir, déjà morte, mais dangereuse du fait de son incitation à augmenter son utilisation et ses impacts.

Ces approches techno-critique ne doivent cependant pas être la fin des réflexions sur la soutenabilité du numérique. Ce manuscrit défend en effet une approche d'évaluation environnementale proche du *Pharmakon* mais qui doit rester attentive aux limites exposées dans cette section de discussion. La critique à retenir, autre que le potentiel relativisme qu'il porte, est que le *Pharmakon* est une posture d'attente par rapport à la production de technologie numérique. L'évaluation environnementale dans son travail de contre valuation est dans la réaction. Comme pour les services d'IA générative, l'initiative est complètement laissée à des acteurs qui en profitent. En réponse à cela et à l'ultime section sur les travaux futurs qui suit, j'espère que les méthodes d'évaluation environnementales permettront plus que de constater les dégâts environnementaux des services numériques ou de minorer ces impacts. La conscience et les moyens d'estimation de l'empreinte environnementale des services numériques peut être le moyen de construire un numérique plus sobre et désirable pour la société.

7.3 Perspectives

Différentes pistes pour de futurs travaux de recherche ont déjà été évoquées au cours de ce manuscrit. Voici néanmoins les principales perspectives de recherche envisagées dans la continuité des travaux de cette thèse.

Tout d'abord, il serait pertinent de répliquer le travail réalisé d'adaptation de l'ACV attributionnelle aux services numériques avec cette fois l'approche d'ACV conséquentielle. Cela, en lien avec la fin de la discussion, permettrait en partie de regagner une forme d'initiative, de produire des connaissances sur les impacts potentiels d'un service numérique avant sa production. Une approche conséquentielle permettrait aussi d'observer les impacts du déploiement de services à large échelle et les conséquences que cela engendre sur les grandes infrastructures numériques, de communication, stockage et calculs.

De manière similaire, des travaux sur la simulation du comportement des utilisateurs et des infrastructures amélioreraient grandement la qualité d'ACV *ex ante*. Plus qu'améliorer les estimations d'impacts environnementaux, cela permettrait de mieux dimensionner les infrastructures, en étudiant aussi différents comportements d'utilisateurs et d'infrastructures comme l'envisage [34].

Autre enjeu environnemental, la consommation d'eau reste un enjeu difficile à cerner et traiter dans le cadre des services numériques. S'il est vrai que dans le contexte de la consommation d'eau le "où" et le "quand" sont déterminants, cela signifie-t-il que toute ACV incluant cet impact environnemental doit être restreinte à un cadre local? Les récentes avancées législatives européennes sur la consommation d'eau des datacenters [109], notamment l'utilisation de l'indicateur d'efficacité en eau [107], pourraient rendre disponible un important et précieux volume de données à exploiter scientifiquement. Cela pourrait permettre d'accélérer l'estimation de l'impact en eau d'un service de la même manière qu'il est trivial d'obtenir les impacts de l'électricité utilisée par ce service.

Sur le registre de l'accélération, la maturité progressive de l'ACV de services numériques à laquelle a participé ce manuscrit pourrait permettre de construire une méthodologie assez robuste et éprouvée pour être en partie automatisée. Loin de chercher une estimation fine des impacts à même d'être communiquée, l'évaluation assistée et accélérée de l'ensemble des services numérique d'un important système d'information pourrait se

concentrer sur la notion d'éco-conception. Le but serait de détecter à grande échelle les améliorations les plus rapides à déployer pour réduire l'empreinte du système d'information, créant ainsi des heuristiques d'éco-conception et éco-gestion des services numériques.

Enfin, le travail sur Picasoft, un fournisseur de services auto-hébergés sur un petit ensemble de serveurs, représentant une partie d'une fédération d'hébergeur de services numériques indépendants, semble être un terrain d'expérimentation prometteur. Notamment pour la question de l'impact environnemental de l'hébergement, quel est le mieux entre hébergement dans le cloud d'une grande organisation ou sur site en gestion locale ? Est-ce que les services numériques tels que proposés par ces petits hébergeurs locaux sont plus ou moins soutenables que ceux proposés par les géants de l'industrie numérique ? Les services étudiés via Picasoft représentent une grande part des usages numériques des citoyens. Est-ce que de tels services pourraient exister, s'ils étaient utilisés à grande échelle dans la société, dans le cadre d'un futur au sein des limites planétaires ?

Annexe 1 : ACV accélérée d'un ensemble de services numériques

Comme expliqué dans [38], l'entreprise hébergeant la thèse dont est issu le présent manuscrit, OCTO Technology a rencontré durant la réalisation de son bilan carbone un obstacle : le calcul des émissions des services numériques livrés par l'entreprise à ses clients. Pour répondre à cette problématique, un modèle simplifié de calcul des émissions d'un service numérique pour un an a été créé. Il s'inspire du modèle utilisé dans [231] pour calculer l'empreinte environnementale d'un site web statique.

Le modèle s'inspire de manière simplifiée de la démarche ACV, utilisant comme données d'entrée d'une part des données ICV. D'autre part, les données sur l'activité du service sont recueillies sur une base déclarative auprès des responsables de missions liées à la livraison des services numériques concernés.

Au final, le modèle ainsi créé avec les données récoltées permet bien plus que simplement répondre à la problématique originelle. Le modèle de calcul simplifié des impacts environnementaux des services numériques pour l'année 2023 permet d'explorer les enjeux suivants :

- Établir une première étude statistique large sur l'empreinte des services numériques à l'échelle de la France. Cette étude pourra fournir des bases en termes d'ordre de grandeur des impacts mais aussi donner des informations sur leur distribution fréquente, au sein du service ou entre les catégories d'impact.
- Voir quels sont les paramètres qui influencent le plus l'empreinte des services, par exemple : Le poids moyen des pages, le temps d'écran moyen, le nombre d'utilisateurs, le dimensionnement de la partie hébergement, etc.
- Il existe de nombreux outils de mesure plus ou moins automatique de l'empreinte d'une page ou d'un site web, voir d'un service. Comparer dans les cas où cela est possible le résultat de l'ACV accélérée et celui des différents outils pourra donner à voir s'il existe des corrélations plus ou moins fréquentes avec certains outils.

Ces différents aspects de ce qui sera une des premières études ACV multicritères d'un grand ensemble de services numériques permettront de créer une base de fondements statistiques, par exemple pour la rédaction de recommandations d'éco-conceptions.

Publications et présentations

Publié ou accepté

A. Berthelot. L'analyse de cycle de vie de services numériques sur systèmes distribués : définitions et enjeux. Compas'2022, Juillet 2022, Amiens, France. <https://inria.hal.science/hal-03941852v1>

A. Berthelot, E. Caron, M. Jay, L. Lefevre. 2024. Estimating the environmental impact of Generative-AI services using an LCA-based methodology. 31st CIRP Conference on Life Cycle Engineering, Juin 2024, Turin, Italie. <https://inria.hal.science/hal-04346102>

T. Simon, D. Ekchajzer, **A. Berthelot**, E. Fourboul, S. Rince, R. Rouvoy. BoaviztAPI : A Bottom-Up Model to Assess the Environmental Impacts of Cloud Services. HotCarbon'24, Juillet 2024, Santa Cruz, États-Unies. <https://inria.hal.science/hal-04621947v2>

A. Berthelot, E. Caron, M. Jay, L. Lefevre. 2024. The environmental impact of generative AI : measuring and understanding. Communications of the ACM. 2024. Parution prochaine

En cours de revue ou d'écriture

A. Berthelot, E. Caron, R. De Laage, L. Lefevre, A. Nicolas, Fine-grained methodology to assess environmental impact of a set of digital services. TechRxiv. 2024. <https://doi.org/10.36227/techrxiv.173273484.41918649/v1> En cours de revue

Présentations

Octobre 2024. *Life cycle assessment and computer science : the What and the How.* Intervention à l'invitation de l'école d'été ECO-ICT 2024, Fouesnant, France. <http://>

https://perso.ens-lyon.fr/laurent.lefevre/Eco-ICT2024_school/ecoict.html

Octobre 2024. *L'analyse de cycle de vie comme outil et domaine de recherche, contributions à la sobriété numérique*. Conférence à l'invitation des laboratoires IRIT et LAAS-CNRS dans le cadre de leur séminaire sur la sobriété numérique, Toulouse, France. <https://scalelite01.orsay.bbb.cnrs.fr/playback/presentation/2.3/75b66eeb18bd75d92dfaf61579cae5f2925a25ca-1727959975821>

Juin 2024. *Towards a multi-criteria evaluation of the environmental footprint of generative AI services*. Poster sélectionné, avec Mathilde Jay, pour la conférence *ICT4S 2024*, Stockholm, Suède. <https://inria.hal.science/hal-04586653v1>

Avril 2024. *Estimating and understanding the environmental impact of a generative AI service*. Présentation sélectionnée pour la conférence *GreenDays 2024*, Toulouse, France. http://avalon.ens-lyon.fr/greendays2023/GreenDays2023_Adrien_Berthelot.mp4

Mars 2024. *Countering the environmental impact of large digital infrastructures : levers, AI and life cycle*. Intervention à l'invitation de l'École Nationale Supérieure des Mines de Saint-Étienne pour le festival du numérique responsable, avec Laurent Lefevre et Mathilde Jay, Saint-Étienne, France.

Septembre 2023. *LCA of a showcase website, a typical digital service use case*. Présentation sélectionnée pour la conférence *The 11th International Conference on Life Cycle Management(LCM)*, sur la base de travaux réalisés en collaboration avec Eddy Caron, Laurent Lefèvre et Alexis Nicolas, Lille, France. <https://youtu.be/KWbK4i80rXg?t=523>

Mars 2023. *Digital service lifecycle analysis : definitions and challenges*. Présentation sélectionnée pour la conférence *GreenDays 2023*, Lyon, France. http://avalon.ens-lyon.fr/~lfevre/greendays2024/GreenDays2024_Berthelot.mp4

Septembre 2022. *Digital technology and the environment : false solutions and real problems*. Intervention à l'invitation de l'association *Confér'ENS*, association de l'ENS de Lyon pour la vulgarisation scientifique, Lyon, France.

Postface : Bilan environnemental de cette thèse

Inspiré par la démarche de [34]⁴ et d'autres d'examiner le bilan carbone de leur thèse, j'ai souhaité moi aussi réaliser l'exercice, bien moins trivial qu'il ne le semble. Dans une volonté d'approfondir l'exercice, j'ai souhaité utiliser la méthode d'évaluation des impacts environnementaux issue de [122]. Cette méthode inspirée de l'ACV est à l'origine destinée à l'évaluation du coût de développement des projets logiciels. Elle sera utile pour effectuer un bilan environnemental multicritère de ce que représente 3 années de travail. Toutefois, nous allons rapidement voir apparaître des problèmes de données et de périmètre qui feront naturellement écho aux problématiques rencontrées dans ce manuscrit. Nous effectuerons donc une seconde évaluation, un bilan carbone dans la continuité de celui réalisé par [34] afin de comparer les résultats produits par les deux méthodes.

Évaluation multicritère

La méthode décrite par [122] est accompagnée d'une application clé-en-main dont le code est disponible en ligne⁵. L'application utilise principalement quelques données issues de la base ICV de Néga-Octet [175], ainsi que quelques données issues de la base Empreinte de l'ADEME notamment pour les déplacements. J'ai renseigné dans le logiciel mes déplacements domicile-travail, mes déplacements professionnels (tous effectués en train), mon ordinateur portable de travail ainsi que l'écran supplémentaire que j'utilise et pour finir, j'ai tenté de représenter mon usage de services numériques professionnelles. Plusieurs hypothèses et choix de périmètre déterminants sont pris. Tout d'abord, faute de données, l'empreinte du bâtiment où je travaille est écartée, nous tenterons de l'inclure dans la prochaine section à la manière de [34]. Ensuite, j'ai tenté de partiellement représenter l'usage de mes services numériques par la modélisation de mon espace de stockage en ligne de 10 Go, ainsi que par l'usage pendant mes heures de travail de l'équivalent d'une petite VM dans un datacenter. Je prends aussi comme hypothèse que mon jury de thèse viendra en train. Les résultats produits donnent les figures 7.1 et 7.2.

On remarquera, via la figure 7.1, les caractéristiques usuelles du numérique en France, à savoir un mix électrique à la fois très décarboné, mais avec un fort impact en radiations

4. Page 132 en Postface.

5. <https://github.com/Orange-OpenSource/SoftwareLifecycleEnvImpact>

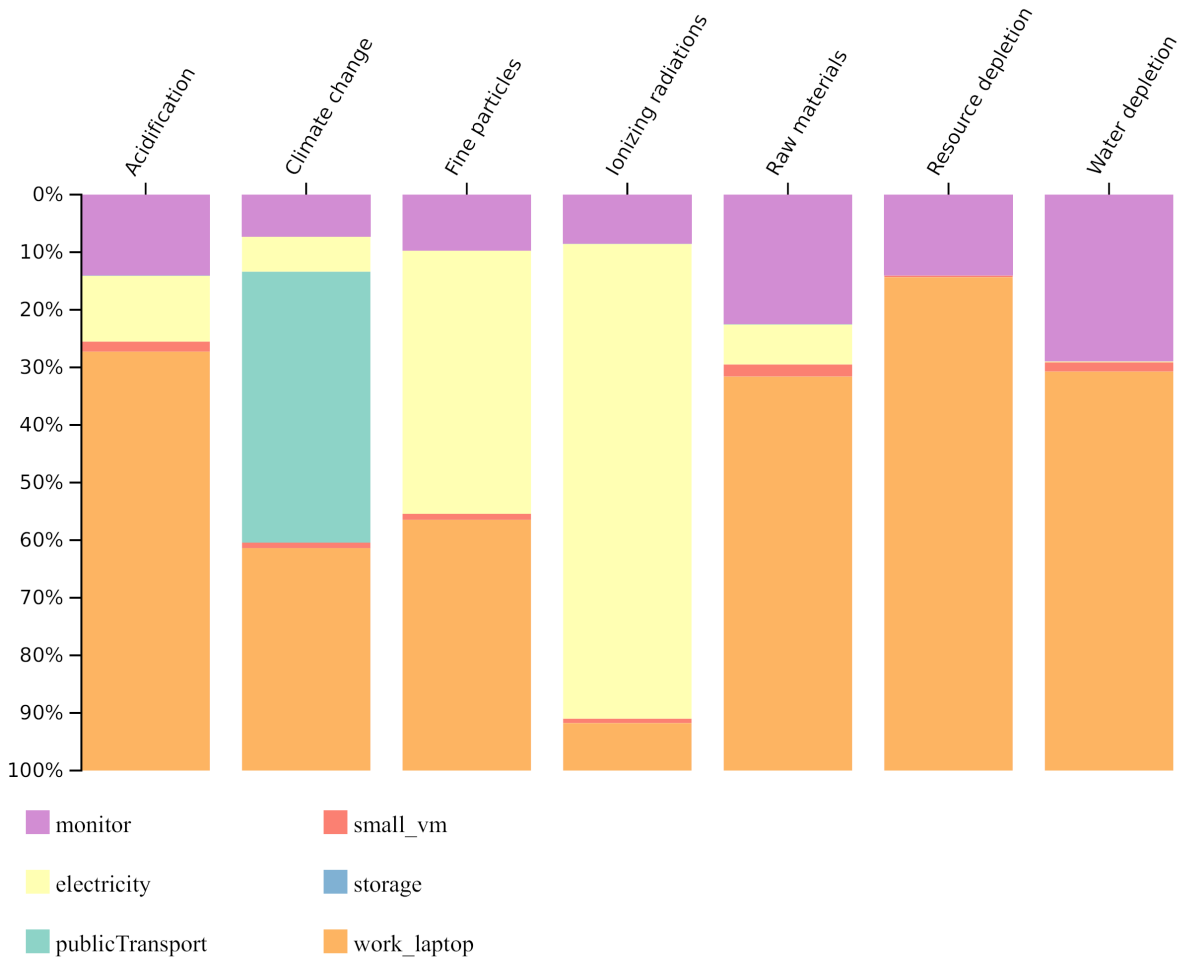


FIGURE 7.1 – Distribution des impacts environnementaux de 3 années de thèse relative entre les sources d’impacts

ionisantes. Les limites du jeu de données utilisé sont néanmoins rapidement visibles, à savoir que les données ADEME utilisées pour les déplacements ne possèdent que l’impact réchauffement climatique. On voit néanmoins dans ce partage, l’aspect non négligeable de l’ajout d’un écran en tant qu’outil de travail. Dans la figure 7.2, on observe un bilan de 287kg d’équivalent CO_2 pour 3 ans de thèse. On voit aussi la part importante liée aux déplacements professionnels, de mon laboratoire à Lyon jusqu’aux locaux d’Octo Technology à Paris, ainsi que les déplacements occasionnés par les conférences. Une part importante de l’empreinte carbone est portée par la seule empreinte de fabrication de l’ordinateur portable. Cette empreinte est sans doute plus faible compte tenu que la donnée Négatoc considère une durée de vie moyenne de 4 ans pour ses ordinateurs portables à usage professionnel mais que les pratiques à Octo tendent à faire évoluer cette durée de vie vers 6 ans⁶. Les résultats de la catégorie d’impact réchauffement climatique posent néanmoins question compte tenu du résultat obtenu de **287kg** de $CO_2eq.$, significativement plus faible que [34] qui affiche **2394kg**. D’autant que, par exemple, malgré le fait que les

6. Ayant personnellement reçu au début de mon contrat un ordinateur professionnel non-neuf, il est probable que ma machine ait déjà dépassé les 4 ans.

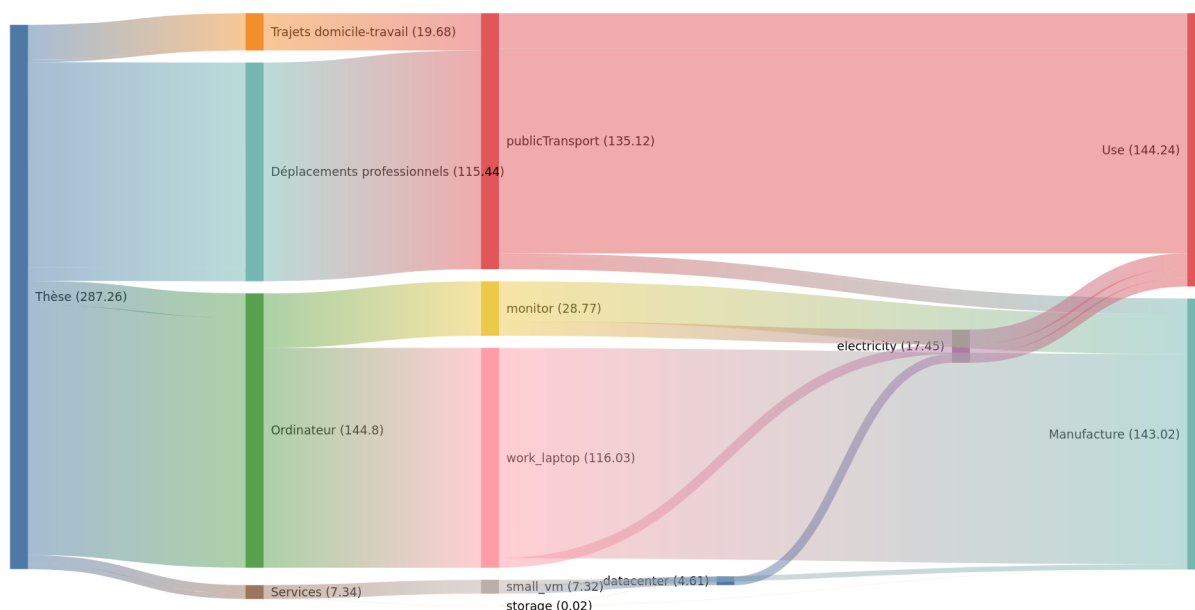


FIGURE 7.2 – Distribution et valeurs des différentes sources d’émissions pour l’impact contribution au réchauffement climatique

trajets domicile-travail aient été réalisés purement avec un vélo classique, l’empreinte des trajets domicile-travail est supérieure pour ce bilan. Y a-t-il un problème de données ?

Bilan carbone

Nous allons donc réaliser dans cette section un bilan carbone en imitant le calcul réalisé dans [34], s’inspirant lui-même de l’initiative Labo1point5⁷. Brièvement, les écarts semblent plus liés à une différence de périmètre que l’usage de données différentes. La majeure partie du bilan carbone de [34] étant due au bâtiment dans lequel il travaille. Les résultats produits en suivant la méthode donnent le bilan suivant de la figure 7.3.

Le détail des calculs peut être établi de la manière suivante :

- **Trajets domicile-travail** : On utilise l’outil disponible sur le site de Labo1point5⁸. Pour aller de mon domicile à mon lieu travail, je prends le tram sur 2km. Le simulateur donne, pour 3 années d’allers-retours 5 jours par semaine, un total de **8 ± 6 kg eq.CO₂**.
- **Déplacements professionnels** : On utilise à nouveau l’outil disponible sur le site de Labo1point5⁹. En cumulant l’ensemble de mes trajets en train, l’on obtient un total de **145 ± 79 kg eq.CO₂**.
- **Lieu de travail** : Je travaille majoritairement dans un laboratoire au sein de l’ENS de Lyon. Je n’ai pas d’information sur le bilan carbone de mon seul laboratoire, le LIP, je vais donc prendre comme référence le bilan carbone le plus récent de

7. <https://labo1point5.org/>

8. <https://apps.labos1point5.org/commutes-simulator>

9. <https://apps.labos1point5.org/travels-simulator>

l'école datant de 2021¹⁰. J'ai ainsi le bilan total de l'ENS de Lyon, je sélectionne les catégories à intégrer dans mon bilan carbone¹¹. Ces catégories représentent un total de 5659 tonnes de CO_2 , soit plus de 60% de l'empreinte annuelle de l'établissement. Pour m'en allouer une part, je vais diviser cette empreinte annuelle par la population de l'établissement. Selon les listes électorales disponibles cette population est de 3708 individus. Cela donne ainsi, divisé par 3708 et multiplié par 3 années de thèse un total de **4579 kg** de CO_2 .

- **Équipements informatiques et bureaux** : En utilisant cette catégorie disponible dans le bilan carbone de l'ENS ainsi que la même allocation, j'obtiens pour 3 ans **284 kg** de CO_2 .

On obtient un total d'environ **5 tonnes** équivalent CO_2 émis pour ces trois années de thèses, comme décrit par la figure 7.3.

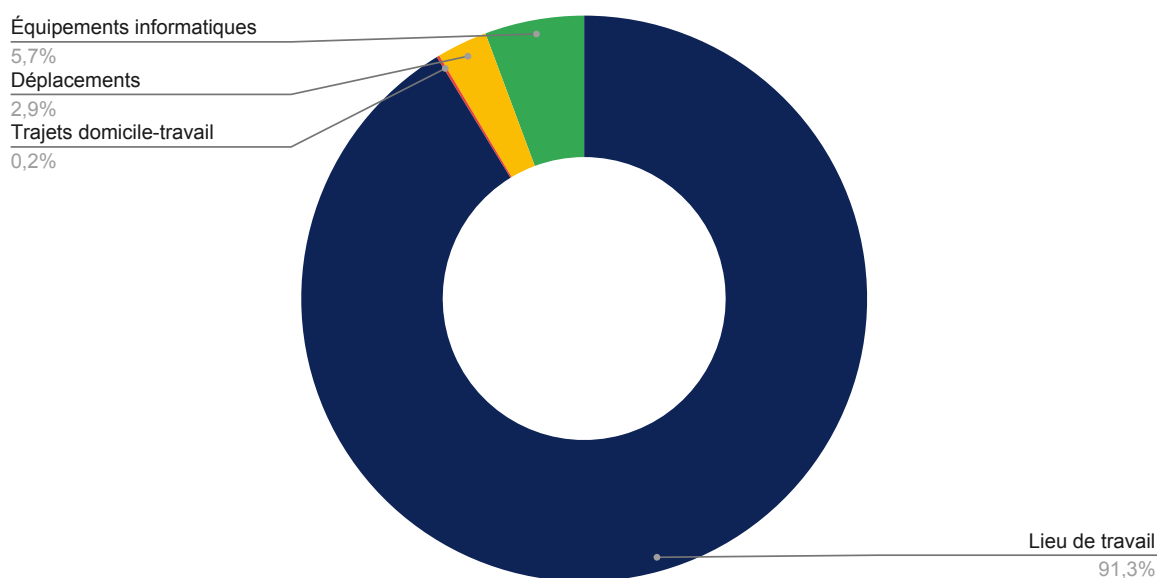


FIGURE 7.3 – Distribution du bilan carbone pour 3 ans de thèse, par catégories d'émissions

La répartition de l'empreinte carbone entre les différentes catégories souligne l'importance écrasante de la structure par rapport aux autres catégories. Bien que l'allocation réalisée, du bilan carbone de l'ENS de Lyon à celui de la thèse, soit critiquable, il est indéniable qu'un ensemble aussi important de bâtiments au chauffage majoritairement non-électrique ne peut qu'impacter lourdement le bilan carbone de toute activité s'y déroulant. Si ce fait peut sembler décourageant pour toute action individuelle, il est toutefois à noter que d'autres choix individuels de mobilités auraient pu sévèrement modifier

10. <https://www.ens-lyon.fr/lecole/transition-ecologique/reduction-de-lempreinte-carbone>

11. "Énergie des bâtiments et machines", "Locaux (amortissement construction)", "Équipements informatiques et bureaux", "Émissions directes de GES Déchets de bureau Utilisation véhicules établissement" et "Véhicules et machines opérés par l'entreprise (amortissement fabrication)".

le résultat. L'impact de trajet domicile-travail en voiture ou de l'avion aurait pu presque doubler le bilan carbone de cette thèse.

Conclusion

Cette section sur l'impact environnemental de cette thèse relève plus de l'expérience de pensée que de la contribution scientifique concrète (justifiant sa place ici en postface du présent manuscrit). L'exercice a de nombreuses limites, et malgré un aspect qui peut sembler convenu, cette auto-évaluation reste une expérience en grande partie exploratoire. Je n'ai pas eu besoin de faire de bilan carbone pour savoir que si je désirais significativement réduire mon empreinte carbone il fallait très certainement éviter l'avion. Toutefois, je ne pensais pas l'empreinte de mon lieu de travail si grande, et je ne pensais pas avoir en prenant le tram, une empreinte plus faible que si je prenais le vélo¹².

En conclusion, comme pour le bilan carbone, l'exercice d'évaluation de l'empreinte environnementale liée à une personne ou son activité n'est évidemment pas *La* solution mais fait selon moi partie de la solution. Il est évident que la mesure est insuffisante en elle-même mais elle est difficilement contournable, que cela soit pour prioriser les sujets, communiquer et convaincre à leur sujet ou investiguer sous un autre le fonctionnement d'un système.

12. Ce résultat m'a surpris car il est inférieur à celui obtenu par mon collègue qui effectue ses trajets domicile-travail en vélo (24 ± 18 kg de CO_2 pour ses trois années de thèse). Le simulateur utilise des facteurs d'impacts, retravaillés, issus de la base Carbone de l'ADEME (<https://apps.labos1point5.org/documentation>). Si la différence entre les deux modes de transport reste mince, elle illustre la capacité certaine des infrastructures comme les transports en commun à réaliser des économies d'échelle.

Bibliographie

- [1] J. R. McNeill and P. Engelke, *The Great Acceleration : An Environmental History of the Anthropocene Since 1945*. Harvard University Press, 2016.
- [2] S. L. Lewis and M. A. Maslin, “Defining the anthropocene,” *Nature*, vol. 519, no. 7542, pp. 171–180, 2015. Publisher : Nature Publishing Group.
- [3] S. Graham, M. Wary, F. Calcagni, M. Cisneros, C. de Luca, S. Gorostiza, O. Stedje Hanserud, G. Kallis, P. Kotsila, S. Leipold, J. Malumbres-Olarte, T. Partridge, A. Petit-Boix, A. Schaffartzik, G. Shokry, S. Tirado-Herrero, J. van den Bergh, and P. Ziveri, “An interdisciplinary framework for navigating social–climatic tipping points,” *People and Nature*, vol. 5, no. 5, pp. 1445–1456, 2023. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pan3.10516>.
- [4] V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, *IPCC Summary for Policymakers*, pp. 3–32. Cambridge, United Kingdom and New York, NY, USA : Cambridge University Press, 2021.
- [5] T. Wiedmann and J. Minx, “A definition of ‘carbon footprint’,” *Ecological economics research trends*, vol. 1, no. 2008, pp. 1–11, 2008.
- [6] I. P. on Climate Change (IPCC), *The Earth’s Energy Budget, Climate Feedbacks and Climate Sensitivity*, p. 923–1054. Cambridge University Press, 2023.
- [7] S. Godin-Beekmann, P. A. Newman, and I. Petropavlovskikh, “30th anniversary of the montreal protocol : From the safeguard of the ozone layer to the protection of the earth’s climate,” *Comptes Rendus. Géoscience*, vol. 350, no. 7, pp. 331–333, 2018.
- [8] K. Richardson, W. Steffen, W. Lucht, J. Bendtsen, S. E. Cornell, J. F. Donges, M. Drüke, I. Fetzer, G. Bala, W. von Bloh, G. Feulner, S. Fiedler, D. Gerten, T. Gleeson, M. Hofmann, W. Huiskamp, M. Kummu, C. Mohan, D. Nogués-Bravo, S. Petri, M. Porkka, S. Rahmstorf, S. Schaphoff, K. Thonicke, A. Tobian, V. Virkki, L. Wang-Erlandsson, L. Weber, and J. Rockström, “Earth beyond six of nine planetary boundaries,” *Science Advances*, vol. 9, no. 37, p. eadh2458, 2023. Publisher : American Association for the Advancement of Science.
- [9] K. W. Kapp, *Les Coûts sociaux de l’entreprise privée*. Les petits matins, 2015. Originellement publié en 1950. Traduit de l’anglais en français par Blanche Bronstein-Vinaver.

- [10] B. Stiegler, “L’esprit des lumières à l’époque du philosophical engineering,” *Intellectica*, vol. 59, no. 1, pp. 29–40, 2013.
- [11] M. Hilbert, “Digital technology and social change : the digital transformation of society from a historical perspective,” *Dialogues in Clinical Neuroscience*, vol. 22, no. 2, pp. 189–194, 2020.
- [12] GeSI et Accenture Strategy, “SMARTer2030 – ICT solutions for 21st century challenges,” 2015.
- [13] GSMA, “The enablement effect - the impact of mobile communications technologies on carbon emission reductions,” 2019.
- [14] A. Rasoldier, J. Combaz, A. Girault, K. Marquet, and S. Quinton, “How realistic are claims about the benefits of using digital technologies for GHG emissions mitigation?,” in *Eighth Workshop on Computing within Limits 2022*, (245 Main St, Cambridge MA), p. 14, LIMITS, PubPub, 2022.
- [15] G. Roussilhe, “Que peut le numérique pour la transition écologique?,” 2021. <https://gauthierroussilhe.com/ressources/que-peut-le-numerique-pour-la-transition-ecologique>.
- [16] “Etude ADEME – Arcep sur l’empreinte environnementale du numérique en 2020, 2030 et 2050,” 2023. <https://www.arcep.fr/la-regulation/grands-dossiers-thematiques-transverses/lempreinte-environnementale-du-numerique/etude-ademe-arcep-empreinte-environnemental-numerique-2020-2030-2050.html>.
- [17] ADEME, “Prospective - transitions 2050 - rapport,” 2021.
- [18] Bordage, F., de Montenay, L., Benqassem, S., Delmas-Orgelet, J., Domon, F., Prunel, D., Vateau, C. et Lees Perasso, E., “Le numérique en europe : une approche des impacts environnementaux par l’analyse du cycle de vie,” Dec. 2021.
- [19] S. Sala, A. Cerutti, and R. Pant, “Development of a weighting approach for the environmental footprint,” *Publications Office of the European Union*, no. KJ-NA-28562-EN-C (print), KJ-NA-28562-EN-N (online), 2017. Luxembourg, 2017, ISBN 978-92-79-68041-0 (print).
- [20] The Climate Group on behalf of the Global eSustainability Initiative (GeSI), “SMART 2020 : Enabling the low carbon economy in the information age,” 2008. <https://gesi.org/research/smart-2020-enabling-the-low-carbon-economy-in-the-information-age>.
- [21] C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. S. Blair, and A. Friday, “The real climate and transformative impact of ICT : A critique of estimates, trends, and regulations,” *Patterns*, vol. 2, p. 100340, September 2021.
- [22] Comité d’experts techniques sur la mesure, présidé par Catherine Mancini, “Evaluation de l’impact environnemental des tic : analyse des écarts méthodologiques,” tech. rep., ADEME - ARCEP, Avril 2023. ISSN n°2258-3106 https://www.arcep.fr/uploads/tx_gspublication/impact-environnement-analyse-ecarts-metodologiques-secteurTIC_avril2023.pdf.
- [23] A. Pasek, H. Vaughan, and N. Starosielski, “The world wide web of carbon : Toward a relational footprinting of information and communications technology’s climate impacts,” *Big Data & Society*, vol. 10, no. 1, p. 20539517231158994, 2023.

- [24] G. Roussilhe, A.-L. Ligozat, and S. Quinton, "A long road ahead : a review of the state of knowledge of the environmental effects of digitization," *Current Opinion in Environmental Sustainability*, vol. 62, p. 101296, 2023.
- [25] ITU, "ITU 11410 : Methodology for environmental life cycle assessments of information and communication technology goods, networks and services," 2012. https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-L.1410-201203-S!!PDF-E&type=items.
- [26] T. S. Baines, H. W. Lightfoot, S. Evans, A. Neely, R. Greenough, J. Peppard, R. Roy, E. Shehab, A. Braganza, A. Tiwari, J. R. Alcock, J. P. Angus, M. Bastl, A. Cousens, P. Irving, M. Johnson, J. Kingston, H. Lockett, V. Martinez, P. Michele, D. Tranfield, I. M. Walton, and H. Wilson, "State-of-the-art in product-service systems," *Proceedings of the Institution of Mechanical Engineers, Part B : Journal of Engineering Manufacture*, vol. 221, no. 10, pp. 1543–1552, 2007.
- [27] M. Hogan, F. Liu, A. Sokol, and J. Tong, "NIST cloud computing standards roadmap : version 1.0," in *NIST Publication 500-291*, pp. NIST SP 500–291v1, National Institute of Standards and Technology, 2011. Edition : 0 Report Number : NIST SP 500-291v1.
- [28] N. Buclet, "Concevoir une nouvelle relation à la consommation : l'économie de fonctionnalité," *Annales des mines - Série Responsabilité et environnement*, pp. 57–66, July 2005.
- [29] Eurostat, "Utilisation des services de cloud computing, par classe de taille d'entreprise, par classe de taille d'entreprise," 2023. DOI :10.2908/isoc_cicce_use.
- [30] S. Sorrell, "Jevons' paradox revisited : The evidence for backfire from improved energy efficiency," *Energy Policy*, vol. 37, pp. 1456–1469, Apr. 2009.
- [31] P. Bourdieu, "Le champ scientifique," *Actes de la Recherche en Sciences Sociales*, vol. 2, no. 2, pp. 88–104, 1976.
- [32] L. M. Hilty and B. Aebischer, "ICT for sustainability : An emerging research field," in *ICT Innovations for Sustainability* (L. M. Hilty and B. Aebischer, eds.), Advances in Intelligent Systems and Computing, (Cham), pp. 3–36, Springer International Publishing, 2015.
- [33] A. Rasoldier, *Comment évaluer le potentiel d'une solution numérique face à l'urgence écologique ? Application aux plateformes de covoiturage régulier à l'échelle locale*. PhD thesis, Université Grenoble Alpes, 2024.
- [34] M. Madon, *Étude de comportements de sobriété dans les centres de calculs*. PhD thesis, Université Toulouse III - Paul Sabatier, 2024.
- [35] L. Grimal, *Permaingénierie : un cadre théorique de transition vers un paradigme de soutenabilité forte. Étude par les IHM*. PhD thesis, Université de technologie de Troyes, Mar. 2023.
- [36] "Publication costs – "we are on the edge of the abyss"." CNRS, 2023. <https://www.cnrs.fr/en/update/publication-costs-we-are-edge-abyss>.
- [37] E. Kincaid, "Elsevier's scopus to continue indexing mdpi's sustainability after reevaluation," 2024. <https://retractionwatch.com/2024/01/05/elseviers-scopus-to-continue-indexing-mdpis-sustainability-after-reevaluation/>.

- [38] A. Nicolas, “Octo a construit sa calculette pour estimer les émissions co2e de l’usage du numérique, mais ça ne fait pas tout !,” 2021. <https://blog.octo.com/bilan-carbone-octo-a-construit-sa-calculette-pour-estimer-les-emissions-co2-e-de-lusage-du-numerique-mais-ca-ne-fait-pas-tout>.
- [39] “Boavizta : Manifeste & vision.” <https://boavizta.org/manifeste-vision>.
- [40] M. Weiser, B. Welch, A. Demers, and S. Shenker, “Scheduling for reduced CPU energy,” in *First Symposium on Operating Systems Design and Implementation (OSDI 94)*, (Monterey, CA), USENIX Association, Nov. 1994.
- [41] M. Culbert, “Low power hardware for a high performance pda,” in *Proceedings of COMPCON ’94*, pp. 144–147, 1994.
- [42] J. Flinn and M. Satyanarayanan, “Energy-aware adaptation for mobile applications,” *SIGOPS Oper. Syst. Rev.*, vol. 33, p. 48–63, dec 1999.
- [43] Y. Liang, C.-Z. Zhao, H. Yuan, Y. Chen, W. Zhang, J.-Q. Huang, D. Yu, Y. Liu, M.-M. Titirici, Y.-L. Chueh, H. Yu, and Q. Zhang, “A review of rechargeable batteries for portable electronic devices,” *InfoMat*, vol. 1, no. 1, pp. 6–32, 2019.
- [44] L. M. Feeney, “An Energy Consumption Model for Performance Analysis of Routing Protocols for Mobile Ad Hoc Networks,” *Mobile Networks and Applications*, vol. 6, pp. 239–249, June 2001.
- [45] R. Bianchini and R. Rajamony, “Power and energy management for server systems,” *Computer*, vol. 37, no. 11, pp. 68–76, 2004.
- [46] M. T. Schmitz and B. M. Al-Hashimi, “Considering power variations of dvs processing elements for energy minimisation in distributed systems,” in *Proceedings of the 14th International Symposium on Systems Synthesis, ISSS ’01*, (New York, NY, USA), p. 250–255, Association for Computing Machinery, 2001.
- [47] K. H. Kim, R. Buyya, and J. Kim, “Power aware scheduling of bag-of-tasks applications with deadline constraints on dvs-enabled clusters,” in *Seventh IEEE International Symposium on Cluster Computing and the Grid (CCGrid ’07)*, pp. 541–548, 2007.
- [48] C. L. Belady and C. G. Malone, “Metrics and an infrastructure model to evaluate data center efficiency,” in *Proceedings of the ASME 2007 InterPACK Conference collocated with the ASME/JSME 2007 Thermal Engineering Heat Transfer Summer Conference*, vol. ASME 2007 InterPACK Conference, Volume 1, pp. 751–755, 07 2007.
- [49] G. Says, “Percent of data centers will have insufficient power and cooling capacity by 2008,” *Gartner Inc. press release*, 2006. Accessed 16 Jun. 2024 : <https://web.archive.org/web/20070522085017/http://www.gartner.com/it/page.jsp?id=499090>.
- [50] X. Cao, L. Liu, Y. Cheng, and X. Shen, “Towards energy-efficient wireless networking in the big data era : A survey,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 303–332, 2018.
- [51] G. Da Costa, M. D. de Assunção, J.-P. Gelas, Y. Georgiou, L. Lefèvre, A.-C. Orgerie, J.-M. Pierson, O. Richard, and A. Sayah, “Multi-facet approach to reduce energy consumption in clouds and grids : the green-net framework,” in *Proceedings of the 1st*

- International Conference on Energy-Efficient Computing and Networking*, e-Energy '10, (New York, NY, USA), p. 95–104, Association for Computing Machinery, 2010.
- [52] A.-C. Orgerie, M. D. d. Assuncao, and L. Lefevre, “A survey on techniques for improving the energy efficiency of large-scale distributed systems,” *ACM Comput. Surv.*, vol. 46, mar 2014.
- [53] V. Ostapenco, L. Lefèvre, A.-C. Orgerie, and B. Fichel, “Modeling, evaluating and orchestrating heterogeneous environmental leverages for large scale data centers management,” *International Journal of High Performance Computing Applications*, 2023.
- [54] J. Koomey, S. Berard, M. Sanchez, and H. Wong, “Implications of historical trends in the electrical efficiency of computing,” *Annals of the History of Computing, IEEE*, vol. 33, pp. 46 – 54, 04 2011.
- [55] “Tracking data centres and data transmission networks - energy.” International Energy Agency IEA, 2023. <https://www.iea.org/energy-system/building-s/data-centres-and-data-transmission-networks#tracking>.
- [56] A. De Vries, “Bitcoin’s energy consumption is underestimated : A market dynamics approach,” *Energy Research & Social Science*, vol. 70, p. 101721, 2020.
- [57] N. Lally, K. Kay, and J. Thatcher, “Computational parasites and hydropower : A political ecology of bitcoin mining on the columbia river,” *Environment and Planning E : Nature and Space*, vol. 5, no. 1, pp. 18–38, 2022.
- [58] L. Courtillat-Piazza, T. Pirson, L. Golard, and D. Bol, “A cradle-to-gate life cycle analysis of bitcoin mining equipment using sphera LCA andecoinvent databases,” 2024.
- [59] H. Estecahandy, “Geopolitics of cryptocurrency mining in kazakhstan,” *Central Asian Survey*, vol. 0, no. 0, pp. 1–19, 2024.
- [60] V. Mazauric, a. auffeves, o. ezratty, and S. Ciliberto, “Quid après les lois de moore et de koomey ?,” *Annales des Mines - Responsabilité et environnement*, vol. 110, pp. 41–45, 04 2023.
- [61] S. Naffziger and J. Koomey, “Energy efficiency of computing : What’s next,” *Retrieved December*, vol. 15, p. 2022, 2016.
- [62] L. Williams and N. Bergman, “Koomey’s law forevermore? a document analysis of the production and circulation of the promise of ‘green 5g’,” *Technological Forecasting and Social Change*, vol. 187, p. 122193, 2023.
- [63] J. Houghton, G. Jenkins, and J. Ephraums, *Climate Change 1990 The Science of Climate Change*. The Intergovernmental Panel on Climate Change IPCC - Working Group I, 1996.
- [64] L. Belkhir and A. Elmeligi, “Assessing ict global emissions footprint : Trends to 2040 & recommendations,” *Journal of Cleaner Production*, vol. 177, pp. 448–463, 2018.
- [65] A. S. Andrae and T. Edler, “On global electricity usage of communication technology : trends to 2030,” *Challenges*, vol. 6, no. 1, pp. 117–157, 2015.

- [66] F. Bordage, F. Abrikh, A. Annaïg, H. Ferreboeuf, and J. Orgelet. GreenIT.fr, 2019. [urlhttps://www.greenit.fr/etude-empreinte-environnementale-du-numerique-mondial/](https://www.greenit.fr/etude-empreinte-environnementale-du-numerique-mondial/).
- [67] S. Mingay, “Green it : the new industry shock wave,” *Gartner RAS Research Note G*, vol. 153703, no. 7, 2007.
- [68] Christian Belady, Dan Azevedo, Michael Patterson, Jack Pouchet, and Roger Tipley, “WP#32 - carbon usage effectiveness (CUE) : A green grid data center sustainability metric | the green grid,” 2010. <https://www.thegreengrid.org/en/resources/library-and-tools/241-Carbon-Usage-Effectiveness-%28CUE%29%3A-A-Green-Grid-Data-Center-Sustainability-Metric>.
- [69] K. Le, O. Bilgir, R. Bianchini, M. Martonosi, and T. Nguyen, “Managing the cost, energy consumption, and carbon footprint of internet services,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 38, pp. 357–358, 06 2010.
- [70] S. Akoush, R. Sohan, A. Rice, A. Moore, and A. Hopper, “Free lunch : Exploiting renewable energy for computing,” *13th Workshop on Hot Topics in Operating Systems (HotOS XIII)*, 01 2011.
- [71] N. Sharma, S. Barker, D. Irwin, and P. Shenoy, “Blink : managing server clusters on intermittent power,” in *Proceedings of the Sixteenth International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS XVI*, (New York, NY, USA), p. 185–198, Association for Computing Machinery, 2011.
- [72] L. Grange, *Datacenter management for on-site intermittent and uncertain renewable energy sources*. phdthesis, Université Paul Sabatier - Toulouse III, 10 2019-10-03.
- [73] I. n. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, “Parasol and greenswitch : managing datacenters powered by renewable energy,” *SIGARCH Comput. Archit. News*, vol. 41, p. 51–64, mar 2013.
- [74] K. K. Nguyen, M. Cheriet, M. Lemay, V. Reijs, A. Mackarel, and A. Pastrama, “Environmental-aware virtual data center network,” *Computer Networks*, vol. 56, no. 10, pp. 2538–2550, 2012. Green communication networks.
- [75] P. Fraser, “More of a good thing – is surplus renewable electricity an opportunity for early decarbonisation?.” International Energy Agency IEA, 2019. <https://www.iea.org/commentaries/more-of-a-good-thing-is-surplus-renewable-electricity-an-opportunity-for-early-decarbonisation>.
- [76] M. A. Vaziri Rad, A. Kasaeian, X. Niu, K. Zhang, and O. Mahian, “Excess electricity problem in off-grid hybrid renewable energy systems : A comprehensive review from challenges to prevalent solutions,” *Renewable Energy*, vol. 212, pp. 538–560, 2023.
- [77] D. Robinson, “Irish power crunch could be prompting aws to ration compute resources.” The Register, 2024. https://www.theregister.com/2024/04/09/aws_resource_restrictions/.
- [78] “Electricity 2024 - analysis and key findings. a report by the international energy agency.” IEA, 01 2024. <https://www.iea.org/reports/electricity-2024>.
- [79] A. Lal and F. You, “Climate concerns and the future of nonfungible tokens : Leveraging environmental benefits of the ethereum merge,” *Proceedings of the National Academy of Sciences*, vol. 120, no. 29, p. e2303109120, 2023.

- [80] Q. Xu, W. Ricks, A. Manocha, N. Patankar, and J. D. Jenkins, “System-level impacts of voluntary carbon-free electricity procurement strategies,” *Joule*, vol. 8, no. 2, pp. 374–400, 2024.
- [81] Magali Palluau, Charlotte Hugrel, Yannick Le Guern, and Frédéric Croison, “MIX ELECTRIQUES GEOGRAPHIQUES VERSUS MIX ELECTRIQUES MARCHE : QUELLES RECOMMANDATIONS EN ACV ?,” Tech. Rep. 2022-01, Association SCORE LCA, 02 2024. https://scorelca.org/espace_mb/etudes/SCORE-2022-01-synthfr-8387ulsgmx-81.pdf.
- [82] Jan Konietzko, “Moving beyond carbon tunnel vision with a sustainability data strategy,” 02 2022. <https://www.cognizant.com/us/en/insights/insights-blog/moving-beyond-carbon-tunnel-vision-with-a-sustainability-data-strategy-codex7121>.
- [83] M. G. Pascual, “Artificial intelligence guzzles billions of liters of water.” *El Pais*, 11 2023. <https://english.elpais.com/technology/2023-11-15/artificial-intelligence-guzzles-billions-of-liters-of-water.html>.
- [84] U. Gupta, Y. G. Kim, S. Lee, J. Tse, H.-H. S. Lee, G.-Y. Wei, D. Brooks, and C.-J. Wu, “Chasing carbon : The elusive environmental footprint of computing,” in *2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, (Los Alamitos, CA, USA), pp. 854–867, IEEE Computer Society, 2021.
- [85] D. Bol, T. Pirson, and R. Dekimpe, “Moore’s law and ICT innovation in the anthropocene,” in *IEEE Design, Automation and Test in Europe Conference 2021*, 2021.
- [86] J.-M. Pierson, G. Baudic, S. Caux, B. Celik, G. Da Costa, L. Grange, M. Haddad, J. Lecuivre, J.-M. Nicod, L. Philippe, V. Rehn-Sonigo, R. Roche, G. Rostirolla, A. Sayah, P. Stolf, M.-T. Thi, and C. Varnier, “Datazero : Datacenter with zero emission and robust management using renewable energy,” *IEEE Access*, vol. 7, pp. 103209–103230, 2019.
- [87] M. Vasconcelos, D. Cordeiro, G. da Costa, F. Dufossé, J.-M. Nicod, and V. Rehn-Sonigo, “Optimal sizing of a globally distributed low carbon cloud federation,” in *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pp. 203–215, 2023.
- [88] T. Sukprasert, A. Souza, N. Bashir, D. Irwin, and P. Shenoy, “On the limitations of carbon-aware temporal and spatial workload shifting in the cloud,” in *Proceedings of the Nineteenth European Conference on Computer Systems*, EuroSys ’24, pp. 924–941, Association for Computing Machinery, 04 2024.
- [89] B. Whitehead, R. Tozer, A. Shah, B. Kosik, G. Maidment, D. Andrews, and A. Bsc, “The environmental burden of data centres – a screening LCA methodology,” in *CIBSE ASHRAE Technical Symposium, Imperial College, London UK*, (222 Balham High Road, London), p. 17, CIBSE, May 2012.
- [90] Y. Arushanyan, “LCA of ICT solutions : environmental impacts and challenges of assessment.” Licentiate thesis - KTH Royal Institute of Technology, 2013. <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-137199>.
- [91] L. Farrant and Y. Le Guern, “Which environmental impacts for ict ? - lca case study on electronic mail,” in *2012 Electronics Goes Green 2012+*, pp. 1–5, 2012.

- [92] C. Preist, D. Schien, and E. Blevis, “Understanding and mitigating the effects of device and cloud service design decisions on the environmental footprint of digital infrastructure,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI ’16, (New York, NY, USA), p. 1324–1337, Association for Computing Machinery, 2016.
- [93] A. Lundström and D. Pargman, “Developing a framework for evaluating the sustainability of computing projects,” in *Proceedings of the 2017 Workshop on Computing Within Limits*, LIMITS ’17, (New York, NY, USA), pp. 111–117, Association for Computing Machinery, 2017.
- [94] M. Pärssinen, M. Kotila, R. Cuevas, A. Phansalkar, and J. Manner, “Environmental impact assessment of online advertising,” *Environmental Impact Assessment Review*, vol. 73, pp. 177–200, 2018.
- [95] C. Preist, D. Schien, and P. Shabajee, “Evaluating sustainable interaction design of digital services : The case of youtube,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, (New York, NY, USA), p. 1–12, Association for Computing Machinery, 2019.
- [96] R. Itten, R. Hischer, A. Andrae, J. Bieser, L. Cabernard, A. Falke, H. Ferreboeuf, L. Hilty, R. Keller, E. Lees-Perasso, C. Preist, and M. Stucki, “Digital transformation—life cycle assessment of digital services, multifunctional devices and cloud computing,” *The International Journal of Life Cycle Assessment*, vol. 25, 2020.
- [97] T. Billstein, A. Björklund, and T. Rydberg, “Life cycle assessment of network traffic : A review of challenges and possible solutions,” *Sustainability*, vol. 13, p. 11155, Jan. 2021. Number : 20 Publisher : Multidisciplinary Digital Publishing Institute.
- [98] V. Forti, C. Baldé, R. Kuehr, and G. Bel, *The Global E-waste Monitor 2020. Quantities, flows, and the circular economy potential*. United Nations University (UNU)/United Nations Institute for Training and Research (UNITAR) – co-hosted SCYCLE Programme, International Telecommunication Union (ITU) & International Solid Waste Association (ISWA), 2020.
- [99] M. Ficher, T. Bauer, and A.-L. Ligozat, “A comprehensive review of the end-of-life modeling in LCAs of digital equipment.” working paper or preprint, Mar. 2024.
- [100] U. Gupta, M. Elgamal, G. Hills, G.-Y. Wei, H.-H. S. Lee, D. Brooks, and C.-J. Wu, “Act : designing sustainable computer systems with an architectural carbon modeling tool,” in *Proceedings of the 49th Annual International Symposium on Computer Architecture*, ISCA ’22, (New York, NY, USA), p. 784–799, Association for Computing Machinery, 2022.
- [101] T. Simon, D. Ekchajzer, A. Berthelot, E. Fourboul, S. Rince, and R. Rouvoy, “BoaviztAPI : a bottom-up model to assess the environmental impacts of cloud services,” in *HotCarbon’24. Workshop on Sustainable Computer Systems*, (Santa Cruz, United States), July 2024.
- [102] M. Ficher, F. Berthoud, A.-L. Ligozat, P. Sigonneau, M. Wisslé, and B. Tebbani, “Assessing the carbon footprint of the data transmission on a backbone network,” in *24th Conference on Innovation in Clouds, Internet and Networks*, 03 2021.

- [103] N. Ullrich, F. M. Piontek, C. Herrmann, A. Saraev, and T. Viere, “Estimating the resource intensity of the internet : A meta-model to account for cloud-based services in LCA,” *Procedia CIRP*, vol. 105, pp. 80–85, 2022.
- [104] S. Daniel, S. Paul J S, B. A. Huseyin, B. Luke, and K. Angeliki, “Assessing the carbon reduction potential for video streaming from short-term coding changes,” in *16th International Conference on Quality of Multimedia Experience (QoMEX)*, International Workshop on Quality of Multimedia Experience, QoMEX, (United States), Institute of Electrical and Electronics Engineers (IEEE), mar 2024. <https://qomex2024.itec.aau.at/>.
- [105] G. Guennebaud, A. Bugeau, and A. Dudouit, “Assessing VoD pressure on network power consumption,” in *ICT4S - International Conference on Information and Communications Technology for Sustainability*, (Rennes, France), June 2023.
- [106] V. D. Reddy, B. Setz, G. S. V. R. K. Rao, G. R. Gangadharan, and M. Aiello, “Metrics for sustainable data centers,” *IEEE Transactions on Sustainable Computing*, vol. 2, no. 3, pp. 290–303, 2017.
- [107] Michael Patterson, Dan Azevedo, Christian Belady, and Jack Pouchet, “WP#35 - water usage effectiveness (WUE) : A green grid data center sustainability metric | the green grid.” The Green Grid., 2011. <https://www.thegreengrid.org/en/resources/library-and-tools/238-Water-Usage-Effectiveness-%28WUE%29%3A-A-Green-Grid-Data-Center-Sustainability-Metric->.
- [108] ISO Central Secretary, “Information technology — Data centres key performance indicators - part 9 : Water usage effectiveness (WUE),” Standard ISO/IEC 30134-9 :2022, International Organization for Standardization, Geneva, CH, 2022.
- [109] Commission Européenne, “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,” 2024. https://eur-lex.europa.eu/eli/reg_del/2024/1364/oj.
- [110] D. Mytton, “Data centre water consumption,” *npj Clean Water*, vol. 4, no. 1, pp. 1–6, 2021. Publisher : Nature Publishing Group.
- [111] P. Li, J. Yang, M. A. Islam, and S. Ren, “Making AI less "thirsty" : Uncovering and addressing the secret water footprint of AI models,” 2023.
- [112] R. Bashroush, N. Rteil, R. Kenny, and A. Wynne, “Optimizing server refresh cycles : The case for circular economy with an aging moore’s law,” *IEEE Transactions on Sustainable Computing*, vol. 7, no. 1, pp. 189–200, 2022.
- [113] Julie Sinistore, Mukunth Natarajan, and Eric Christensen, “Comparative life cycle assessment of pure storage FlashArray//x70,” ISO 14044-CONFORMANT REPORT Version 1 12/2/2022, WSP USA, 12 2022. <https://www.purestorage.com/docs.html?item=/type/pdf/subtype/doc/path/content/dam/pdf/en/analyst-reports/protected/ar-flasharray-x70-comparative-life-cycle-assessment.pdf>.
- [114] G. Mersy and S. Krishnan, “Toward a life cycle assessment for the carbon footprint of data,” in *Proceedings of the 2nd Workshop on Sustainable Computer Systems*, HotCarbon ’23, (New York, NY, USA), Association for Computing Machinery, 2023.

- [115] A.-L. Ligozat, J. Lefevre, A. Bugeau, and J. Combaz, “Unraveling the hidden environmental impacts of AI solutions for environment life cycle assessment of AI solutions,” *Sustainability*, vol. 14, p. 5172, 2022.
- [116] A. S. Luccioni, S. Viguier, and A.-L. Ligozat, “Estimating the carbon footprint of BLOOM, a 176b parameter language model,” *Journal of Machine Learning Research*, vol. 24, no. 253, pp. 1–15, 2023.
- [117] S. Prakash, M. Stewart, C. Banbury, M. Mazumder, P. Warden, B. Plancher, and V. J. Reddi, “Is tinyml sustainable?,” *Commun. ACM*, vol. 66, p. 68–77, oct 2023.
- [118] D. Patterson, J. M. Gilbert, M. Gruteser, E. Robles, K. Sekar, Y. Wei, and T. Zhu, “Energy and emissions of machine learning on smartphones vs. the cloud,” *Communication of the ACM*, vol. 67, p. 86–97, jan 2024.
- [119] PricewaterhouseCoopers Advisory, “Study of the environmental impact of the tezos blockchain - life cycle assessment of the tezos blockchain protocol,” 12 2021. <https://tezos.com/2021-12-06-Tezos-LCA-Final.pdf>.
- [120] F. Berthoud and M. Ficher, “Évaluation de l’empreinte carbone d’une visioconférence entre deux utilisateurs du service rendez-vous,” techreport, CNRS - EcoInfo, 03 2022. <https://hal-cnrs.archives-ouvertes.fr/hal-03605638>.
- [121] G. Roussilhe and E. Fourboul, “Analyse de cycle de vie (ACV) d’un service de messagerie en ligne,” 08 2022. <https://gauthierroussilhe.com/ressources/analyse-de-cycle-de-vie-service-numerique>.
- [122] T. Simon, P. Rust, R. Rouvoy, and J. Penhoat, “Uncovering the environmental impact of software life cycle,” in *International Conference on Information and Communications Technology for Sustainability*, June 2023. <https://inria.hal.science/hal-04082263>.
- [123] N. C. Horner, A. Shehabi, and I. L. Azevedo, “Known unknowns : indirect energy effects of information and communication technology,” *Environmental Research Letters*, vol. 11, no. 10, p. 103001, 2016. Institution : Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States) Publisher : IOP Publishing.
- [124] J. C. T. Bieser, R. Hintemann, L. M. Hilty, and S. Beucker, “A review of assessments of the greenhouse gas footprint and abatement potential of information and communication technology,” *Environmental Impact Assessment Review*, vol. 99, p. 107033, March 2023.
- [125] J. C. T. Bieser and L. M. Hilty, “Indirect effects of the digital transformation on environmental sustainability : Methodological challenges in assessing the greenhouse gas abatement potential of ICT,” in *EPiC Series in Computing*, vol. 52, pp. 68–81, EasyChair, 2018. ISSN : 2398-7340.
- [126] V. C. Coroama and F. Mattern, “Digital rebound - why digitalization will not redeem us our environmental sins,” in *Proceedings of the 6th International Conference on ICT for Sustainability, ICT4S 2019, Lappeenranta, Finland, June 10-14, 2019*, (Aachen, Germany), CEUR-WS.org, 2019.
- [127] Hubblo and Iroco, “Réalisation d’un diagnostic éco-conception de mail - technologie et usages,” 2024. <https://blog.iroco.co/images/ACV/ACV-mail-Hubblo-Iroco-01032024.pdf>.

- [128] T. Santarius, J. C. T. Bieser, V. Frick, M. Höjer, M. Gossen, L. M. Hilty, E. Kern, J. Pohl, F. Rohde, and S. Lange, “Digital sufficiency : conceptual considerations for ICTs on a finite planet,” *Annals of Telecommunications*, 2022.
- [129] ISO/TC 207/SC 5, “Environmental management — life cycle assessment — principles and framework,” Standard ISO 14040 :2006, International Organization for Standardization, Geneva, CH, 07 2006. <https://www.iso.org/standard/37456.html>.
- [130] ISO/TC 207/SC 5, “Environmental management — life cycle assessment — requirements and guidelines,” Standard ISO 14044 :2006, International Organization for Standardization, Geneva, CH, 07 2006. <https://www.iso.org/standard/38498.html>.
- [131] A. Bjørn, M. Owsianiak, C. Molin, and A. Laurent, *Main Characteristics of LCA*, pp. 9–16. Cham : Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-56475-3_2.
- [132] R. G. Hunt, W. E. Franklin, and R. G. Hunt, “LCA — how it came about : — personal reflections on the origin and the development of LCA in the USA,” *The International Journal of Life Cycle Assessment*, vol. 1, no. 1, pp. 4–7, 1996.
- [133] I. Boustead, “LCA — how it came about : — the beginning in the u.k.,” *The International Journal of Life Cycle Assessment*, vol. 1, pp. 147–150, September 1996.
- [134] A. Bjørn, M. Owsianiak, C. Molin, and M. Z. Hauschild, *LCA History*, pp. 17–30. Cham : Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-56475-3_3.
- [135] C. Benoît, G. A. Norris, S. Valdivia, A. Ciroth, A. Moberg, U. Bos, S. Prakash, C. Ugaya, and T. Beck, “The guidelines for social life cycle assessment of products : just in time!,” *The International Journal of Life Cycle Assessment*, vol. 15, pp. 156–163, 01 2010.
- [136] E. Commission, “2013/179/EU : Commission recommendation of 9 april 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations text with EEA relevance,” 04 2013. <http://data.europa.eu/eli/reco/2013/179/oj/eng>.
- [137] S. Manfredi, K. Allacker, N. Pelletier, E. Schau, K. Wolf, R. Pant, and D. Pennington, “Comparing the european commission product environmental footprint method with other environmental accounting methods,” *The International Journal of Life Cycle Assessment*, vol. 20, 03 2015.
- [138] N. Pelletier, K. Allacker, R. Pant, and S. Manfredi, “The european commission organisation environmental footprint method : Comparison with other methods, and rationales for key requirements,” *The International Journal of Life Cycle Assessment*, 02 2013.
- [139] J. B. Guinée, S. Cucurachi, P. J. Henriksson, and R. Heijungs, “Digesting the alphabet soup of LCA,” *The International Journal of Life Cycle Assessment*, vol. 23, no. 7, pp. 1507–1511, 2018.

- [140] K. T. Gradin and A. Björklund, “The common understanding of simplification approaches in published LCA studies—a review and mapping,” *The International Journal of Life Cycle Assessment*, vol. 26, no. 1, pp. 50–63, 2021.
- [141] M. Lindahl, J. Tingström, and C. Jensen, “A comparison between the environmental effect analysis (eea) and the life cycle assessment (lca) methods - based on four case studies,” in *Life Cycle Planning, Design and Management for Eco-Products and Systems - 7th International Seminar on Life Cycle Engineering*, (Tokyo, Japan), 2000. <https://api.semanticscholar.org/CorpusID:107059408>.
- [142] M. Jegen, “Life cycle assessment : from industry to policy to politics,” *The International Journal of Life Cycle Assessment*, vol. 29, pp. 597–606, 04 2024.
- [143] P. Frankl and F. Rubik, *Application of LCA in general*, pp. 43–51. Berlin, Heidelberg : Springer Berlin Heidelberg, 2000. https://doi.org/10.1007/978-3-662-04127-7_3.
- [144] O. J. Hanssen, “Status of life cycle assessment (lca) activities in the nordic region,” *The International Journal of Life Cycle Assessment*, vol. 4, pp. 315–320, 1999.
- [145] A. Plepys, “The grey side of ict,” *Environmental Impact Assessment Review*, vol. 22, no. 5, pp. 509–523, 2002. Sustainability in the Information Society.
- [146] A. Köhler and L. Erdmann, “Expected environmental impacts of pervasive computing,” *Human and Ecological Risk Assessment : An International Journal*, vol. 10, no. 5, pp. 831–852, 2004. <https://doi.org/10.1080/10807030490513856>.
- [147] C. Casal, C. Wunnik, L. Sancho, J. Burgelman, P. Desruelle, Lorenz, Erdmann, Lorenz, Hilty, James, Goodman, and Peter, Arnfalk, “The future impact of ICT on environmental sustainability,” Tech. Rep. EUR 21384 EN, European Commission - Joint Research Centre - Institute for Prospective Technological Studies (IPTS), 08 2004. [http://publicationslist.org/data/lorenz.hilty/ref-47/2004%20IPTS%20Final%20Report%202004%20\(synthesis\).pdf](http://publicationslist.org/data/lorenz.hilty/ref-47/2004%20IPTS%20Final%20Report%202004%20(synthesis).pdf).
- [148] E. Strassner and J. Nicholson, “Measuring the digital economy in the united states,” *Statistical Journal of the IAOS*, vol. 36, pp. 647–655, 08 2020. doi : 10.3233/SJI-200666.
- [149] J. R. Nicholson, T. F. Howells III, and D. B. Wasshausen, “Measuring the US digital economy,” in *Advances in Business Statistics, Methods and Data Collection*, pp. 105–124, John Wiley & Sons, Ltd, 2023. Section : 6, doi : 10.1002/9781119672333.ch6.
- [150] A. Andrae, “Method for calculating the avoided impact of specific information and communication technology services,” *International Journal of Environmental Engineering and Development*, vol. 2, pp. 73–87, 03 2024.
- [151] W. Scharnhorst, “Life cycle assessment in the telecommunication industry : A review,” *The International Journal of Life Cycle Assessment*, vol. 13, pp. 75–86, 01 2008.
- [152] ITU, “ITU I1410 : Methodology for environmental life cycle assessments of information and communication technology goods, networks and services,” 2014. https://www.itu.int/rec/dologin_pub.asp?lang=f&id=T-REC-L.1410-201412-I!!PDF-E&type=items.

- [153] A. Moberg, C. Borggren, and G. Finnveden, “Books from an environmental perspective - part 2 : E-books as an alternative to paper books,” *The International Journal of Life Cycle Assessment*, vol. 16, pp. 238–246, 03 2011.
- [154] C. Borggren, A. Moberg, and G. Finnveden, “Books from an environmental perspective - part 1 : environmental impacts of paper books sold in traditional and internet bookshops,” *The International Journal of Life Cycle Assessment*, vol. 16, pp. 138–147, 02 2011.
- [155] J. G. Bull and R. A. Kozak, “Comparative life cycle assessments : The case of paper and digital media,” *Environmental Impact Assessment Review*, vol. 45, pp. 10–18, 2014.
- [156] E. Amasawa, T. Ihara, and K. Hanaki, “Role of e-reader adoption in life cycle greenhouse gas emissions of book reading activities,” *The International Journal of Life Cycle Assessment*, vol. 23, 09 2018.
- [157] MEYER Julia (ADEME), NICO Tom (I Care), BURGUBURU Alexis (I Care), RIGAL Margot (I Care), LIZON Benjamin (I Care), GENIN Léo (I Care), CATALAN Caroline (I Care), ADAM Isaure (I Care), “Environmental impact assessment of the digitalization of cultural services.” La librairie ADEME, 11 2022. <https://librairie.ademe.fr/dechets-economie-circulaire/5961-environmental-impact-assessment-of-the-digitalization-of-cultural-services.html>.
- [158] A. Anders and P. Marcello, “Environmental engineering (EE); methodology for environmental life cycle assessment (LCA) of information and communication technology (ICT) goods, networks and services,” Tech. Rep. RES/EE-EEPS14, ETSI, 12 2014. https://www.etsi.org/deliver/etsi_es/203100_203199/203199/01.03.00_50/es_203199v010300m.pdf.
- [159] F. Guldbrandsson and P. Bergmark, “Opportunities and limitations of using life cycle assessment methodology in the ICT sector,” in *2012 Electronics Goes Green 2012+*, pp. 1–6, 09 2012.
- [160] B. Whitehead, D. Andrews, and A. Shah, “The life cycle assessment of a uk data centre,” *The International Journal of Life Cycle Assessment*, vol. 20, no. 3, pp. 332–349, 2015. Electronic supplementary material The online version of this article (doi : 10.1007/s11367-014-0838-7) contains supplementary material, which is available to authorized users.
- [161] J. Malmodin, D. Lundén, M. Nilsson, and G. Andersson, “LCA of data transmission and IP core networks,” *2012 Electronics Goes Green 2012+*, pp. 1–6, 2012.
- [162] J. Malmodin, D. Lundén, A. Moberg, G. Andersson, and M. Nilsson, “Life cycle assessment of ict,” *Journal of Industrial Ecology*, vol. 18, no. 6, pp. 829–845, 2014.
- [163] D. Maga, M. Hiebel, and C. Knermann, “Comparison of two ict solutions : Desktop pc versus thin client computing,” *The International Journal of Life Cycle Assessment*, vol. 18, 05 2012.
- [164] Y. Arushanyan, E. Ekener-Petersen, and G. Finnveden, “Lessons learned – review of LCAs for ICT products and services,” *Computers in Industry*, vol. 65, no. 2, pp. 211–234, 2014.

- [165] P. Balin, A. Bohas, C. Charbuillet, and E. Drezet, *Impacts écologiques des Technologies de l'Information et de la Communication : Les faces cachées de l'immatérialité*. EDP Sciences, 2012. Publication Title : Impacts écologiques des Technologies de l'Information et de la Communication.
- [166] “LOI n° 2021-1485 du 15 novembre 2021 visant à réduire l’empreinte environnementale du numérique en France,” Nov. 2021. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044327272>.
- [167] “Article 13 - LOI n° 2020-105 du 10 février 2020 relative à la lutte contre le gaspillage et à l'économie circulaire (1) - Légifrance,” Feb. 2020. https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000041554658.
- [168] A. Moatti, “le numérique, adjectif substantivé,” *Le Debat*, vol. 3, no. 170, p. pp. 133, 2012.
- [169] S. Chowdhury and M. Åkesson, “A proposed conceptual framework for identifying the logic of digital services.,” in *Pacific Asia Conference on Information Systems*, p. 47, 01 2011.
- [170] G. Vial, “Understanding digital transformation : A review and a research agenda,” *The Journal of Strategic Information Systems*, vol. 28, pp. 118–144, 06 2019.
- [171] M. Koch, D. Krohmer, M. Naab, D. Rost, and M. Trapp, “A matter of definition : Criteria for digital ecosystems,” *Digital Business*, vol. 2, no. 2, p. 100027, 2022.
- [172] N. Ensmenger, “The environmental history of computing,” *Technology and Culture*, vol. 59, pp. S7–S33, 2018.
- [173] K. Williams, S. Chatterjee, and M. Rossi, *Design of Emerging Digital Services : A Taxonomy*, pp. 235–253. Boston, MA : Springer US, 06 2010.
- [174] K. Kerwin, D. Andrews, N. Adibi, B. Whitehead, C. Szablewski, and J. Chenadec, “Developing a circular economy for the data centre industry – how the CEDaCI project contributes to sustainable decision making,” *E3S Web of Conferences*, vol. 349, p. 06007, 2022. Publisher : EDP Sciences.
- [175] LCIE Bureau Veritas, APL Datacenter, DDemain, and GreenIT.fr, “Negaoctet,” 2022. <https://negaoctet.org/en/home/>.
- [176] L. R. Viana, M. Cheriet, K.-K. Nguyen, D. Marchenko, and J.-F. Boucher, “Sending fewer emails will not save the planet ! an approach to make environmental impacts of ICT tangible for canadian end users,” *Sustainable Production and Consumption*, vol. 34, pp. 453–466, 2022.
- [177] J. Pohl, V. Frick, M. Finkbeiner, and T. Santarius, “Assessing the environmental performance of ICT-based services : Does user behaviour make all the difference ?,” *Sustainable Production and Consumption*, vol. 31, pp. 828–838, 2022.
- [178] D.-G. for Energy (European Commission), Ramboll, Resilio, A. Louguet, M. Caspani, D. Pytel, A. Pirlot, M. Faura Rosendo, and H. Blanadet, *Assessment of the energy footprint of digital actions and services*. Publications Office of the European Union, 06 2023. doi :10.2833/478689.
- [179] L. Ismail, “Référentiel général pour l'IA frugale - mesurer et réduire l'impact environnemental de l'IA,” Tech. Rep. AFNOR SPEC 2314, Association française de

- normalisation (AFNOR), 06 2024. <https://www.boutique.afnor.org/fr-fr/norme/afnor-spec-2314/referentiel-general-pour-lia-frugale-mesurer-et-reduire-limpact-environneme/fa208976/421140>.
- [180] “Référentiel général de l'écoconception des services numériques,” 05 2024. <https://www.arcep.fr/mes-demarches-et-services/entreprises/fiches-pratiques/referentiel-general-ecoconception-services-numeriques.html>.
- [181] C. Youen, “Analyse critique de la nouvelle version du référentiel général d'Écoconception des services numériques (rgesn),” 06 2024. https://www.lewebvert.fr/blog/2024-06-03-rgesn-analyse-critique-3_3/.
- [182] Y. Lina and B. Nicolas, “Le référentiel général pour l'écoconception des services numériques (rgesn),” 06 2024. [https://blog.octo.com/octo-le-referentiel-general-pour-l'ecoconception-des-services-numeriques-\(rgesn\)](https://blog.octo.com/octo-le-referentiel-general-pour-l'ecoconception-des-services-numeriques-(rgesn)).
- [183] R. Pamminger, S. Glaser, and W. Wimmer, “Modelling of different circular end-of-use scenarios for smartphones,” *The International Journal of Life Cycle Assessment*, vol. 26, pp. 470–482, 03 2021.
- [184] R. Frischknecht, F. Wyss, S. Büsser Knöpfel, T. Lützkendorf, and M. Balouktsi, “Cumulative energy demand in LCA : the energy harvested approach,” *The International Journal of Life Cycle Assessment*, vol. 20, no. 7, pp. 957–969, 2015.
- [185] T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, “Climate change 2013 : The physical science basis,” Tech. Rep. IPCC :13, Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)(Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.), 2013. 1535 pp.
- [186] L. van Oers, J. B. Guinée, and R. Heijungs, “Abiotic resource depletion potentials (ADPs) for elements revisited—updating ultimate reserve estimates and introducing time series for production data,” *The International Journal of Life Cycle Assessment*, vol. 25, no. 2, pp. 294–308, 2020.
- [187] R. Frischknecht, A. Braunschweig, P. Hofstetter, and P. Suter, “Human health damages due to ionising radiation in life cycle impact assessment,” *Environmental Impact Assessment Review*, vol. 20, no. 2, pp. 159–189, 2000.
- [188] S. Fazio, F. Biganzoli, V. De Laurentiis, L. Zampori, S. Sala, and E. Diaconu, “Supporting information to the characterisation factors of recommended of life cycle impact assessment methods.” Publications Office of the European Union, 2018. <https://doi.org/10.2760/002447>.
- [189] BUREAU VERITAS, APL - datacenter, DDemain, and ADEME, “Référentiel méthodologique d'évaluation environnementale des services numériques,” July 2021. <https://librairie.ademe.fr/produire-autrement/6022-referentiel-methodologique-d-evaluation-environnementale-des-services-numeriques.html>.
- [190] J. Xu, K. Ota, and M. Dong, “Saving energy on the edge : In-memory caching for multi-tier heterogeneous networks,” *IEEE Communications Magazine*, vol. 56, no. 5, pp. 102–107, 2018.

- [191] M. Yan, C. A. Chan, W. Li, L. Lei, A. F. Gygax, and C.-L. I, “Assessing the energy consumption of proactive mobile edge caching in wireless networks,” *IEEE Access*, vol. 7, pp. 104394–104404, 2019.
- [192] S. Savazzi, V. Rampa, S. Kianoush, and M. Bennis, “An energy and carbon footprint analysis of distributed and federated learning,” *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 1, pp. 248–264, 2023.
- [193] C. Cérin, M. Jay, L. Lefèvre, and D. Trystram, “A methodology and a toolbox to explore dataset related to the environmental impact of http requests,” in *2023 IEEE International Conference on Big Data (BigData)*, pp. 3753–3762, 2023.
- [194] M. Gonzalez, P. Xu, R. Dekimpe, M. Schramme, I. Stupia, T. Pirson, and D. Bol, “Technical and ecological limits of 2.45-ghz wireless power transfer for battery-less sensors,” *IEEE Internet of Things Journal*, vol. 10, no. 17, pp. 15431–15442, 2023.
- [195] ISO/TC 207/SC 1, “Systèmes de management environnemental — lignes directrices pour intégrer l’éco-conception,” Standard ISO 14006 :2020, International Organization for Standardization, Geneva, CH, 01 2020. <https://www.iso.org/fr/standard/72644.html>.
- [196] F. Bordage, “Ecoindex,” 09 2022. <https://www.ecoindex.fr/>.
- [197] Spécinov, “Kastor - audit d’éco-socio-conception pour les sites web,” 06 2022. <https://web.archive.org/web/20220707133304/https://www.specinov.fr/bao/kastor-audit-ecoconception>.
- [198] Conserto, “Greentrackr,” 04 2024. <https://conserto.pro/greentrackr/>.
- [199] GreenIT.fr, “Greenit-analysis,” 2022. <https://github.com/cnumr/GreenIT-Analysis>.
- [200] Wholegrain Digital, “Website carbon calculator v3,” 2018. <https://www.websitcarbon.com/>.
- [201] Digital4better, “Fruggr.io,” 04 2021. <https://www.fruggr.io/fr/>.
- [202] C. Vassallo, S. Panichella, F. Palomba, S. Proksch, H. C. Gall, and A. Zaidman, “How developers engage with static analysis tools in different contexts,” *Empirical Software Engineering*, vol. 25, pp. 1419–1457, 2020.
- [203] J. G. March and H. A. Simon, *Organizations*. John wiley & sons, 1993.
- [204] R. Hanna and O. Joviado, “Guide de bonnes pratiques numérique responsable pour les organisations,” tech. rep., MiNumEco, mission interministérielle numérique éco-responsable, Direction interministérielle du numérique (DINUM), Ministère de la Transition écologique et de la Cohésion des territoires, Institut du Numérique Responsable (INR), Juin 2023. <https://ecoresponsable.numerique.gouv.fr/publications/bonnes-pratiques/>.
- [205] B. Montbroussous, J. Schaeffer, G. P. A. Moreau, F. Berthoud, and G. Feltin, “Calculer le bilan Carbone de votre parc informatique avec EcoDiag, un service EcoInfo,” in *Congrès JRES : Les Journées Réseaux de l’Enseignement et de la Recherche*, (Dijon, France), RENATER, Dec. 2019.
- [206] I. du numérique responsable, “Wenr,” Mars 2021. <https://wenr.isit-europe.org/fr/note-methodologique-wenr/>.

- [207] ISO/TC 207/SC 7, “Gaz à effet de serre - partie 1 : Spécifications et lignes directrices, au niveau des organismes, pour la quantification et la déclaration des émissions et des suppressions des gaz à effet de serre,” Standard ISO 14064-1 :2018, International Organization for Standardization, Geneva, CH, Décembre 2018. <https://www.iso.org/fr/standard/72644.html>.
- [208] G. Bourgeois, K. Bigot, V. Courboulay, and B. Duthil, “Tools for calculating the ict footprint of organisations : adaptation of a european study,” in *Information Systems* (M. Papadaki, P. Rupino da Cunha, M. Themistocleous, and K. Christodoulou, eds.), (Cham), pp. 277–290, Springer Nature Switzerland, 04 2023.
- [209] J. Malmmodin and D. Lundén, “The energy and carbon footprint of the global ict and e&m sectors 2010 - 2015,” in *ICT4S2018. 5th International Conference on Information and Communication Technology for Sustainability* (B. Penzenstadler, S. Easterbrook, C. Venters, and S. I. Ahmed, eds.), vol. 52 of *EPiC Series in Computing*, pp. 187–208, EasyChair, 2018.
- [210] J. C. T. Bieser, V. C. Coroamă, P. Bergmark, and M. Stürmer, “The greenhouse gas (ghg) reduction potential of ict : A critical review of telecommunication companies’ ghg enablement assessments,” *Journal of Industrial Ecology*, vol. n/a, no. n/a, 2024.
- [211] D. R. Cooper and T. G. Gutowski, “The environmental impacts of reuse : A review,” *Journal of Industrial Ecology*, vol. 21, no. 1, pp. 38–56, 2017.
- [212] L. Grimal, I. Di Loreto, N. Burger, and N. Troussier, “Design of an interdisciplinary evaluation method for multi-scaled sustainability of computer-based projects. a workbased on the sustainable computing evaluation framework (SCEF),” *LIMITS Workshop on Computing within Limits*, 2021.
- [213] Body of European Regulators for Electronic Communications (BEREC), “BEREC report on ICT sustainability for end-users : Empowering end-users through environmental transparency on digital products,” 06 2024. <https://www.berec.europa.eu/en/document-categories/berec/reports/berec-report-on-ict-sustainability-for-end-users-empowering-end-users-through-environmental-transparency-on-digital-products>.
- [214] A. Wattiez, K. Le Goff, and D. Bol, “Exploring the influence of database selection on the life cycle assessment of digital services.” Poster presented at the 2024 International Conference on ICT for Sustainability (ICT4S) at KTH Stockholm, Sweden. <https://dial.uclouvain.be/pr/boreal/en/object/boreal%3A290519>.
- [215] R. Obringer, B. Rachunok, D. Maia-Silva, M. Arbabzadeh, R. Nateghi, and K. Madani, “The overlooked environmental footprint of increasing internet use,” *Resources, Conservation and Recycling*, vol. 167, p. 105389, 2021.
- [216] P. University, “Turn off that camera during virtual meetings, environmental study says - simple tips to go green with your internet use during a pandemic,” 2021. <https://www.eurekalert.org/news-releases/539100>.
- [217] D. Mytton, D. Lundén, and J. Malmmodin, “Network energy use not directly proportional to data volume : The power model approach for more reliable network energy consumption calculations,” *Journal of Industrial Ecology*, vol. n/a, no. n/a, 2024. Online Version of Record before inclusion in an issue. <https://doi.org/10.1111/jiec.13512>.

- [218] T. Tabata and T. Y. Wang, “Life cycle assessment of co2 emissions of online music and videos streaming in japan,” *Applied Sciences*, vol. 11, no. 9, 2021.
- [219] K. Nansai, K. Nakajima, S. Kagawa, Y. Kondo, Y. Shigetomi, and S. Suh, “Global mining risk footprint of critical metals necessary for low-carbon technologies : The case of neodymium, cobalt, and platinum in japan,” *Environmental Science & Technology*, vol. 49, no. 4, pp. 2022–2031, 2015. PMID : 25622132.
- [220] A. de Vries and C. Stoll, “Bitcoin’s growing e-waste problem,” *Resources, Conservation and Recycling*, vol. 175, p. 105901, 2021.
- [221] A. De Vries, “Cryptocurrencies on the road to sustainability : Ethereum paving the way for bitcoin,” *Patterns*, vol. 4, no. 1, p. 100633, 2023.
- [222] P. O’Donovan and D. T. J. O’Sullivan, “A systematic analysis of real-world energy blockchain initiatives,” *Future Internet*, vol. 11, no. 8, 2019.
- [223] R. K. Rosenbaum and S. I. Olsen, *Critical Review*, pp. 335–347. Cham : Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-56475-3_13.
- [224] Afnor, “Principes généraux pour l’affichage environnemental des produits de grande consommation - partie 0 : principes généraux et cadre méthodologique,” Normes nationales et documents normatifs nationaux BP X30-323-0, Association française de normalisation (AFNOR), 06 2011. <https://www.boutique.afnor.org/fr-fr/norme/bp-x303230/principes-generaux-pour-laffichage-environnemental-des-produits-de-grande-c/fa170405/37615#AreasStoreProductsSummaryView>.
- [225] M. Jay, V. Ostapenco, L. Lefèvre, D. Trystram, A.-C. Orgerie, and B. Fichel, “An experimental comparison of software-based power meters : focus on CPU and GPU,” in *CCGrid 2023 - 23rd IEEE/ACM international symposium on cluster, cloud and internet computing*, p. 1, IEEE, 2023.
- [226] H. Hernan, F. Ravat, J. Song, P. Roose, and N. Vallès-Parlangeau, “FLOC : outil de mesure énergétique multi-composants,” in *42ème Congrès INformatique des ORganisations et Systèmes d’Information et de Décision (INFORSID 2024)*, (Nancy, France), p. à paraître, May 2024.
- [227] L. Golard, Y. Agram, F. Rottenberg, F. Quitin, D. Bol, and J. Louveaux, “A parametric power model of multi-band sub-6 GHz cellular base stations using on-site measurements,” in *IEEE International Symposium Personal, Indoor and Mobile Radio Communications*, 2024. <https://dial.uclouvain.be/pr/boreal/object/boreal:289070>.
- [228] H. Ferreboeuf, P. Bihouix, P. Fabre, D. Kaplan, L. Lefèvre, A. Monnin, O. Ridoux, S. Vaija, M. Vautier, X. Verne, A. Ducass, M. Efoui-Hess, and Z. Kahraman, “LEAN ICT - POUR UNE SOBRIÉTÉ NUMÉRIQUE,” tech. rep., The Shift Project, Octobre 2018. <https://theshiftproject.org/article/pour-une-sobriete-numerique-rapport-shift/>.
- [229] P. Jacquet, T. Ledoux, and R. Rouvoy, “ScroogeVM : Boosting Cloud Resource Utilization with Dynamic Oversubscription,” *IEEE Transactions on Sustainable Computing*, pp. 1–13, 2024.

- [230] S. Lambert, E. Caron, L. Lefèvre, and R. Grivel, “Étude et modélisation des mécanismes de surallocation de mémoire pour la consolidation de machines virtuelles,” in *Compas 2023 - Conférence francophone d’informatique en Parallélisme, Architecture et Système*, (Annecy, France), July 2023.
- [231] A. Berthelot, E. Caron, L. Lefèvre, and A. Nicolas, “Lca of a showcase website, a typical digital service use case.” Selected talk at the 11th International Conference on Life Cycle Management (LCM), Lille, France, Septembre 2023. <https://youtu.be/KWbK4i80rXg?t=523>.
- [232] J. Rockström, W. Steffen, K. Noone, A. Persson, F. S. Chapin, E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. A. Foley, “A safe operating space for humanity,” *Nature*, vol. 461, no. 7263, pp. 472–475, 2009. Number : 7263 Publisher : Nature Publishing Group.
- [233] M. Z. Hauschild, R. K. Rosenbaum, and S. I. Olsen, eds., *Life Cycle Assessment : Theory and Practice*. Springer International Publishing, 2018.
- [234] A. Bjørn, A. Moltesen, A. Laurent, M. Owsianiak, A. Corona, M. Birkved, and M. Z. Hauschild, *Life Cycle Inventory Analysis*, pp. 117–165. Cham : Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-56475-3_9.
- [235] A. Zamagni, J. Guinée, R. Heijungs, P. Masoni, and A. Raggi, “Lights and shadows in consequential LCA,” *The International Journal of Life Cycle Assessment*, vol. 17, pp. 904–918, 08 2012.
- [236] T. Ekvall, A. Azapagic, G. Finnveden, T. Rydberg, B. P. Weidema, and A. Zamagni, “Attributional and consequential LCA in the ILCD handbook,” *The International Journal of Life Cycle Assessment*, vol. 21, pp. 293–296, 03 2016.
- [237] T. Schaubroeck, S. Schaubroeck, R. Heijungs, A. Zamagni, M. Brandão, and E. Benetto, “Attributional & consequential life cycle assessment : Definitions, conceptual characteristics and modelling restrictions,” *Sustainability*, vol. 13, no. 13, 2021.
- [238] J. Pohl, L. M. Hilty, and M. Finkbeiner, “How LCA contributes to the environmental assessment of higher order effects of ICT application : A review of different approaches,” *Journal of Cleaner Production*, vol. 219, pp. 698–712, 05 2019.
- [239] Y. Benkler and H. Nissenbaum, “Commons-based peer production and virtue,” *Journal of political philosophy*, vol. 14, no. 4, 2006.
- [240] ADEME, “Convention de la mise en oeuvre du programme "alt impact",” Tech. Rep. PRO-INNO-62, Ministère de la Transition écologique et de la Cohésion des territoires - CEE, Juillet 2023. https://www.ecologie.gouv.fr/sites/default/files/documents/Convention_ALT%20IMPACT_sobri%C3%A9t%C3%A9_num%C3%A9rique.pdf.
- [241] L. Siffre, G. Breuil, A. Noureddine, and R. Pawlak, “Demeter : An Architecture for Long-Term Monitoring of Software Power Consumption,” in *17th European Conference on Software Architecture*, (Istanbul, Turkey), Sept. 2023.

- [242] “Victoria metrics v1.102.1.” <https://github.com/VictoriaMetrics/VictoriaMetrics>.
- [243] K. Kumar and M. Kurhekar, “Economically efficient virtualization over cloud using docker containers,” in *2016 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, pp. 95–100, 2016.
- [244] “Grafana loki v3.1.” <https://github.com/grafana/loki>.
- [245] H. J. Syed, A. Gani, R. W. Ahmad, M. K. Khan, and A. I. A. Ahmed, “Cloud monitoring : A review, taxonomy, and open research issues,” *Journal of Network and Computer Applications*, vol. 98, pp. 11–26, 2017.
- [246] “Goaccess v1.8.1.” <https://github.com/allinurl/goaccess>.
- [247] “Hypestat,” 2023. <https://hypestat.com/>.
- [248] “Similarweb,” 2023. <https://www.similarweb.com/>.
- [249] B. J. Jansen, S.-g. Jung, and J. Salminen, “Measuring user interactions with websites : A comparison of two industry standard analytics approaches using data of 86 websites,” *PLOS ONE*, vol. 17, no. 5, 2022. Publisher : Public Library of Science.
- [250] D. Prantl and M. Prantl, “Website traffic measurement and rankings : competitive intelligence tools examination,” *International Journal of Web Information Systems*, vol. 14, no. 4, pp. 423–437, 2018. Publisher : Emerald Publishing Limited.
- [251] A. Shehabi, S. Smith, D. Sartor, R. Brown, M. Herrlin, J. Koomey, E. Masanet, N. Horner, I. Azevedo, and W. Lintner, “United states data center energy usage report,” Tech. Rep. LBNL-1005775, Berkeley National Laboratory, Berkeley, California, June 2016. <https://escholarship.org/content/qt84p772fc/qt84p772fc.pdf>.
- [252] R. K. Rosenbaum, M. Z. Hauschild, A.-M. Boulay, P. Fantke, A. Laurent, M. Núñez, and M. Vieira, *Life cycle impact assessment*, pp. 167–270. Cham : Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-56475-3_10.
- [253] A.-M. Boulay, J. Bare, L. Benini, M. Berger, M. J. Lathuillière, A. Manzardo, M. Margni, M. Motoshita, M. Núñez, A. V. Pastor, *et al.*, “The wulca consensus characterization model for water scarcity footprints : assessing impacts of water consumption based on available water remaining (aware),” *The International Journal of Life Cycle Assessment*, vol. 23, pp. 368–378, 2018.
- [254] S. Pfister, L. Scherer, A.-M. Boulay, M. Motoshita, M. Núñez, M. Damiani, A. Manzardo, J. Huang, A. Link, J. Bunsen, and M. Berger, “Letter to the editor re : “the scarcity-weighted water footprint provides unreliable water sustainability scoring” by vanham and mekonnen, 2021,” *Science of The Total Environment*, vol. 825, p. 154108, 2022.
- [255] S. Sala, A. K. Cerutti, R. Pant, *et al.*, “Development of a weighting approach for the environmental footprint,” *Publications Office of the European Union : Luxembourg*, 2018.
- [256] P. Bergmark and G. Zachrisson, “Towards considering planetary boundaries in life cycle assessments of ict,” in *2022 International Conference on ICT for Sustainability (ICT4S)*, pp. 128–139, 2022.

- [257] R. J. Plevin, M. A. Delucchi, and F. Creutzig, “Using attributional life cycle assessment to estimate climate-change mitigation benefits misleads policy makers,” *Journal of Industrial Ecology*, vol. 18, no. 1, pp. 73–83, 2014.
- [258] M. Brander, R. L. Burritt, and K. L. Christ, “Coupling attributional and consequential life cycle assessment : A matter of social responsibility,” *Journal of Cleaner Production*, vol. 215, pp. 514–521, 2019.
- [259] T. Schaubroeck, “Relevance of attributional and consequential life cycle assessment for society and decision support,” *Frontiers in Sustainability*, vol. 4, p. 1063583, 2023.
- [260] SD, “Stable diffusion service,” 2022. <https://stablediffusionweb.com/#demo>.
- [261] A. Berthelot, E. Caron, M. Jay, and L. Lefèvre, “Estimating the environmental impact of Generative-AI services using an LCA-based methodology,” in *CIRP LCE 2024 - 31st Conference on Life Cycle Engineering*, (Turin, Italy), pp. 1–10, June 2024.
- [262] A. de Vries, “The growing energy footprint of artificial intelligence,” *Joule*, vol. 7, no. 10, pp. 2191–2194, 2023.
- [263] E. Brynjolfsson, D. Li, and L. R. Raymond, “Generative AI at work,” Apr. 2023.
- [264] S. J. Russell, P. Norvig, and E. Davis, *Artificial intelligence : a modern approach*. Prentice Hall series in artificial intelligence, Prentice Hall Upper Saddle River, New Jersey, third edition ed., 2010.
- [265] F. García-Peñalvo and A. Vázquez-Ingelmo, “What do we mean by GenAI? a systematic mapping of the evolution, trends, and techniques involved in generative AI,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 8, pp. 7–16, july 2023. Accepted : 2023-08-28T12 :17 :56Z Publisher : International Journal of Interactive Multimedia and Artificial Intelligence.
- [266] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, vol. 27, Curran Associates, Inc., 2014.
- [267] A. D. Cheok and E. Y. Zhang, “From turing to transformers : A comprehensive review and tutorial on the evolution and applications of generative transformer models,” 10 2023.
- [268] J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn, and P. Villalobos, “Compute trends across three eras of machine learning,” in *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, July 2022. ISSN : 2161-440.
- [269] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for deep learning in NLP,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (A. Korhonen, D. Traum, and L. Màrquez, eds.), (Florence, Italy), pp. 3645–3650, Association for Computational Linguistics, July 2019.
- [270] A. Lacoste, A. Luccioni, V. Schmidt, and T. Dandres, “Quantifying the Carbon Emissions of Machine Learning,” Nov. 2019. <http://arxiv.org/abs/1910.09700>.

- [271] Q. Cao, A. Balasubramanian, and N. Balasubramanian, “Towards accurate and reliable energy measurement of NLP models,” in *Proceedings of SustainNLP : Workshop on Simple and Efficient Natural Language Processing*, (Online), pp. 141–148, Association for Computational Linguistics, 2020.
- [272] P. Henderson, J. Hu, J. Romoff, E. Brunskill, D. Jurafsky, and J. Pineau, “Towards the systematic reporting of the energy and carbon footprints of machine learning,” *Journal of Machine Learning Research*, vol. 21, no. 248, pp. 1–43, 2020.
- [273] D. Patterson, J. Gonzalez, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. So, M. Texier, and J. Dean, “Carbon Emissions and Large Neural Network Training,” *ArXiv*, vol. abs/2104.10350, p. 22, 2021.
- [274] D. Patterson, J. Gonzalez, U. Hölzle, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. R. So, M. Texier, and J. Dean, “The carbon footprint of machine learning training will plateau, then shrink,” *Computer*, vol. 55, no. 7, pp. 18–28, 2022. Conference Name : Computer.
- [275] D. Ganapathy and E. J. Warner, “Defining thermal design power based on real-world usage models,” in *2008 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 1242–1246, 2008.
- [276] C.-J. Wu, R. Raghavendra, U. Gupta, B. Acun, N. Ardalani, K. Maeng, G. Chang, F. Behram, J. Huang, C. Bai, M. Gschwind, A. Gupta, M. Ott, A. Melnikov, S. Candido, D. Brooks, G. Chauhan, B. Lee, H.-H. Lee, and K. Hazelwood, *Sustainable AI : Environmental Implications, Challenges and Opportunities*. Arxiv, Oct. 2021.
- [277] A. Das and A. Modak, “The carbon footprint of machine learning models,” *IJERA*, vol. 3, pp. 246–249, 2023.
- [278] A. S. Luccioni, Y. Jernite, and E. Strubell, “Power hungry processing : Watts driving the cost of AI deployment ?,” 2023.
- [279] C. L. Ludec, M. Cornet, and A. A. Casilli, “The problem with annotation. human labour and outsourcing between france and madagascar,” *Big Data & Society*, vol. 10, no. 2, pp. 1–13, 2023.
- [280] L. Bouza, A. Bugeau, and L. Lannelongue, “How to estimate carbon footprint when training deep learning models? a guide and review,” *Environmental Research Communications*, vol. 5, p. 115014, Nov 2023.
- [281] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (Los Alamitos, CA, USA), pp. 10684–10695, IEEE Computer Society, June 2022.
- [282] HF, “Hugging Face runwayml stable diffusion repository,” 2016.
- [283] B. Tomlinson, R. W. Black, D. J. Patterson, and A. W. Torrance, “The carbon emissions of writing and illustrating are lower for ai than for humans,” *Scientific Reports*, vol. 14, no. 1, p. 3732, 2024.
- [284] F. Cappello, E. Caron, M. Dayde, F. Desprez, E. Jeannot, Y. Jegou, S. Lanteri, J. Leduc, N. Melab, G. Mornet, R. Namyst, P. Primet, and O. Richard, “Grid’5000 : a large scale, reconfigurable, controlable and monitorable Grid platform,” in *SC’05 : Proc. The 6th IEEE/ACM International Workshop on Grid Computing Grid’2005*,

- (Seattle, USA), pp. 99–106, IEEE/ACM, IEEE/ACM, Nov. 2005. hal number inria-00000284.
- [285] S. Rince, “GPU component manufacture impacts.” BoaviztAPI, 2023. <https://github.com/Boavizta/boaviztapi/issues/65>.
- [286] S. Kemp, “Digital 2023 : Global overview report.” DataReportal – Global Digital Insights, 01 2023. <https://datareportal.com/reports/digital-2023-global-overview-report>.
- [287] J. W. Pierre Delforge, “America’s data centers consuming and wasting growing amounts of energy,” tech. rep., NRDC, 02 2015.
- [288] C. Morand, A. Névéol, and A.-L. Ligozat, “MLCA : a tool for Machine Learning Life Cycle Assessment,” in *2024 International Conference on ICT for Sustainability (ICT4S)*, (Stockholm, Sweden), June 2024.
- [289] S. Moon, J.-H. Kim, J. Kim, S. Hong, J. Cha, M. Kim, S. Lim, G. Choi, D. Seo, J. Kim, H. Lee, H. Park, R. Ko, S. Choi, J. Park, J. Lee, and J.-Y. Kim, “Lpu : A latency-optimized and highly scalable processor for large language model inference,” *IEEE Micro*, pp. 1–12, 2024.
- [290] N. Hübner, J. Caspers, V. C. Coroamă, and M. Finkbeiner, “Machine-learning-based demand forecasting against food waste : Life cycle environmental impacts and benefits of a bakery case study,” *Journal of Industrial Ecology*, vol. n/a, no. n/a, p. 1–15, 2024. Online Version of Record before inclusion in an issue : <https://doi.org/10.1111/jiec.13528>.
- [291] L. Lin, R. Wijayawardana, V. Rao, H. Nguyen, E. W. GNIBGA, and A. A. Chien, “Exploding ai power use : an opportunity to rethink grid planning and management,” in *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems, e-Energy ’24*, (New York, NY, USA), p. 434–441, Association for Computing Machinery, 2024.
- [292] D. Schien, P. J. S. Shabajee, H. B. Akyol, L. Benson, and A. Katsenou, “Help, i shrunk my savings! assessing the carbon reduction potential for video streaming from short-term coding changes,” 2023.
- [293] R. Hidalgo, N. Salah, R. Chandra Jetty, A. Jetty, and A. S. Varde, “Personalizing text-to-image diffusion models by fine-tuning classification for ai applications,” in *Proceedings of SAI Intelligent Systems Conference*, pp. 642–658, Springer, 2023.
- [294] A. Bandi, P. V. S. R. Adapa, and Y. E. V. P. K. Kuchi, “The power of generative AI : A review of requirements, models, input–output formats, evaluation metrics, and challenges,” *Future Internet*, vol. 15, p. 260, aug 2023. Number : 8 Publisher : Multidisciplinary Digital Publishing Institute.
- [295] B. He, X. Zheng, Y. Chen, W. Li, Y. Zhou, X. Long, P. Zhang, X. Lu, L. Jiang, Q. Liu, D. Cai, and X. Zhang, “Dxpu : Large-scale disaggregated gpu pools in the datacenter,” *ACM Trans. Archit. Code Optim.*, vol. 20, dec 2023.
- [296] F. Filippini, J. Anselmi, D. Ardagna, and B. Gaujal, “A Stochastic Approach for Scheduling AI Training Jobs in GPU-based Systems,” *IEEE Transactions on Cloud Computing*, no. 01, pp. 1–17, 2023.

- [297] W. Gao, Q. Hu, Z. Ye, P. Sun, X. Wang, Y. Luo, T. Zhang, and Y. Wen, “Deep learning workload scheduling in GPU datacenters : Taxonomy, challenges and vision,” 2022.
- [298] D. Patel, D. Nishball, and J. Eliahou Ontiveros, “Ai datacenter energy dilemma - race for ai datacenter space.” SemiAnalysis, March 2024. <https://www.semianalysis.com/p/ai-datacenter-energy-dilemma-race>.
- [299] T. A. Hemphill, “Us ai data centers and deployment challenges for small modular reactors : proposed regulatory policy recommendations,” *Science and Public Policy*, p. scae040, 2024.
- [300] IPCC[P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)], *Climate Change 2022 : Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA : Cambridge University Press, 2022. doi : 10.1017/9781009157926.
- [301] B. Fix, “Dematerialization through services : Evaluating the evidence,” *BioPhysical Economics and Resource Quality*, vol. 4, no. 2, p. 6, 2019.
- [302] I. Bekavac and D. G. Praničević, “Web analytics tools and web metrics tools : An overview and comparative analysis,” *Croatian Operational Research Review*, vol. 6, pp. 373–386, 2015.
- [303] “Picasoft.” <https://picasoft.net/index.html>.
- [304] L. Mosesso, N. Maudet, E. Nano, T. Thibault, and A. Tabard, “Obsolescence paths : living with aging devices,” in *ICT4S 2023 - International Conference on Information and Communications Technology for Sustainability*, 2023.
- [305] M. Amaral, H. Chen, T. Chiba, R. Nakazawa, S. Choochotkaew, E. K. Lee, and T. Eilam, “Kepler : A framework to calculate the energy consumption of containerized applications,” in *2023 IEEE 16th International Conference on Cloud Computing (CLOUD)*, pp. 69–71, 2023. ISSN : 2159-6190.
- [306] A. Nouredine, M. D. Lodeiro, N. Bru, and R. Chbeir, “The impact of green feedback on users’ software usage,” *IEEE Transactions on Sustainable Computing*, vol. 8, no. 2, p. 280, 2023.
- [307] M. Madon, G. Da Costa, and J.-M. Pierson, “Replay with feedback : How does the performance of HPC system impact user submission behavior?,” *Future Generation Computer Systems*, vol. 155, pp. 66–79, 2024.
- [308] M. Madon, G. Da Costa, and J.-M. Pierson, “Characterization of different user behaviors for demand response in data centers,” in *Euro-Par 2022 : Parallel Processing* (J. Cano and P. Trinder, eds.), pp. 53–68, Springer International Publishing, 2022.
- [309] “GitHub - boavizta/e-footprint : A toolkit for exhaustively modeling the environmental impact of digital services..” <https://github.com/Boavizta/e-footprint>.
- [310] M. Séguéla, R. Mokadem, and J.-M. Pierson, “Dynamic energy and expenditure aware data replication strategy,” in *IEEE International Conference on Cloud Computing Technical Program (CLOUD 2022)*, IEEE, 2022.

- [311] R. D and G. A, “Survey on data replication in cloud systems,” *Web Intelligence*, vol. 22, no. 1, pp. 83–109, 2024. Publisher : IOS Press.
- [312] M. Mazzucato, *The value of everything : Making and taking in the global economy*. Hachette UK, 2018.
- [313] T. Parrique, J. Barth, F. Briens, J. Spangenberg, and A. Kraus-Polk, *Decoupling Debunked. Evidence and arguments against green growth as a sole strategy for sustainability. A study edited by the European Environment Bureau EEB*. European Environmental Bureau, 07 2019.
- [314] S. Dietz and E. Neumayer, “Weak and strong sustainability in the sea : Concepts and measurement,” *Ecological Economics*, vol. 61, no. 4, pp. 617–626, 2007. Special Issue on Environmental Accounting : Introducing the System of Integrated Environmental and Economic Accounting 2003.
- [315] K. Marx, *Le caractère fétiche de la marchandise et son secret*. Editions Allia, 1995. Originellement publié en juillet 1872, quatrième partie du chapitre 1, section 1, du Livre I du Capital. Traduit de l’allemand au français par Joseph Roy.
- [316] F. Flipo, *L’impératif de la sobriété numérique : l’enjeu des modes de vie*. Essais (Éd. matériologiques), Éditions Matériologiques, 2020.
- [317] J. Richard and A. Rambaud, *Révolution comptable : pour une entreprise écologique et sociale*. Les Éditions de l’Atelier / Éditions Ouvrières, 2020.
- [318] L. Grimal, I. di Loreto, and N. Troussier, “Exploratory research on cultural computing for human-computer interaction. the case of sustainability.,” *Proceedings of the Design Society*, vol. 1, p. 1471–1480, 2021.
- [319] J. Norton, B. Penzenstadler, and B. Tomlinson, “Implications of grassroots sustainable agriculture community values on the design of information systems,” *Proc. ACM Hum.-Comput. Interact.*, vol. 3, nov 2019.
- [320] A. Monnin, “Planetary negative commons,” *Multitudes*, vol. 85, no. 4, pp. 117–125, 2021.
- [321] A. Monnin, *Politiser le renoncement*. Éditions divergences, 2023.
- [322] A. Mennicken and R. Salais, *The new politics of numbers : Utopia, evidence and democracy*. Springer Nature, 2022.
- [323] M. Grubb, N. D. Jordan, E. Hertwich, K. Neuhoff, K. Das, K. R. Bandyopadhyay, H. Van Asselt, M. Sato, R. Wang, W. A. Pizer, *et al.*, “Carbon leakage, consumption, and trade,” *Annual Review of Environment and Resources*, vol. 47, no. 1, pp. 753–795, 2022.
- [324] F.-D. Vivien, “Les modèles économiques de soutenabilité et le changement climatique,” vol. 6, no. 2, pp. 75–83. Publisher : La Découverte Section : Économie.
- [325] C. Aylett, “The value of a whale : on the illusions of green capitalism,” *International Affairs*, vol. 98, pp. 2163–2164, 11 2022.
- [326] J. Temple, “Google, amazon and the problem with big tech’s climate claims - how companies reach their emissions goals is more important than how fast.,” *MIT Technology Review*, 07 2024. <https://www.technologyreview.com/2024/07/17/1095019/google-amazon-and-the-problem-with-big-techs-climate-claims/>.

- [327] M. Di Fiore, M. Kuc-Czarnecka, S. Lo Piano, A. Puy, and A. Saltelli, “The challenge of quantification : an interdisciplinary reading,” *Minerva*, vol. 61, no. 1, pp. 53–70, 2023.
- [328] T. Lefort, “Les technologies numériques dans l’éducation : ni remède, ni poison. Éduquer à la technique au-delà du pharmakon,” *Tétralogiques*, no. 28, pp. 151–181, 2023. ISSN : 1634-0450.
- [329] R. Comber and E. Eriksson, “Computing as ecocide,” in *Computing within Limits*, LIMITS, 06 2023.
- [330] J. Halloy, N. Nova, and A. Monnin, “Au-delà du low tech : technologies zombies, soutenabilité et inventions,” *Revue Passerelle*, vol. 21, pp. 120–128, 09 2020.