

Côté obscur de l'IA : quels bénéfices réels de l'IA pour faire face aux crises environnementales ?

Anne-Laure Ligozat

Green Days 2023



D'où je parle

Enseignante-chercheuse en informatique



Recherche: impacts environnementaux du numérique

Membre de



et



POUR UNE INFORMATIQUE ÉCO-RESPONSABLE

IA ?

Intelligence artificielle



Apprentissage automatique
(*machine learning*)



Apprentissage profond
(*deep learning*)



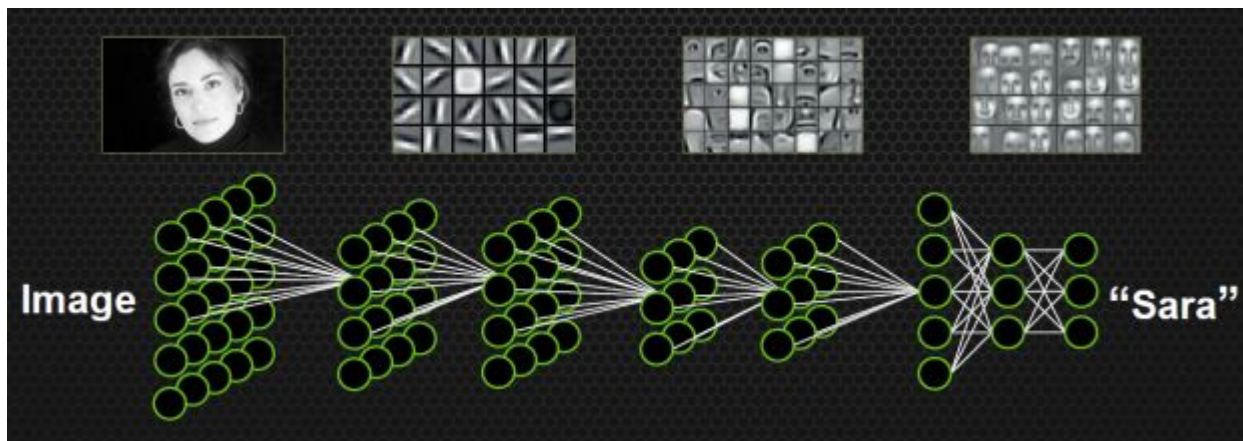
Pourquoi l'IA ?

Réseaux de neurones profonds

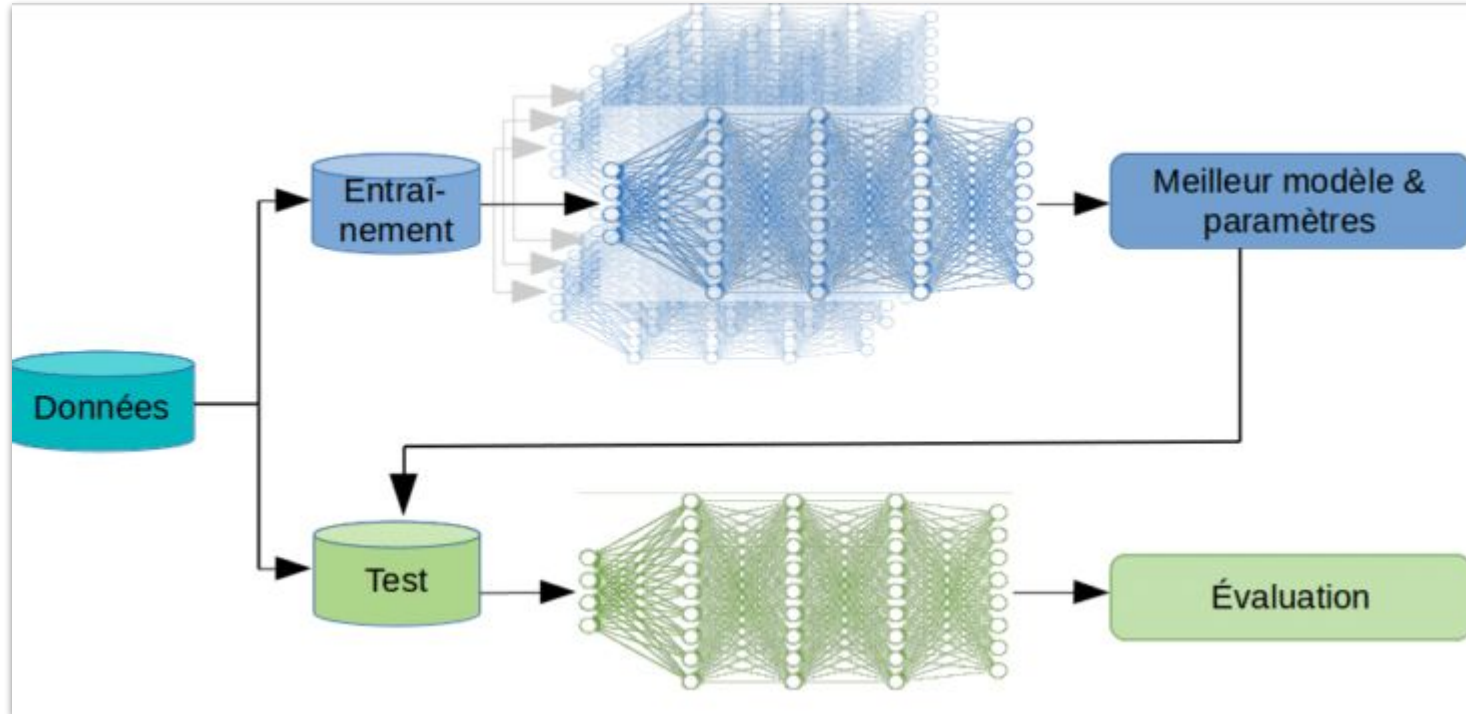
développements logiciels et matériels, notamment calcul sur GPU

→ rendu possible l'entraînement de ces modèles

- en un temps raisonnable
- sur du matériel relativement basique



Processus d'apprentissage profond supervisé



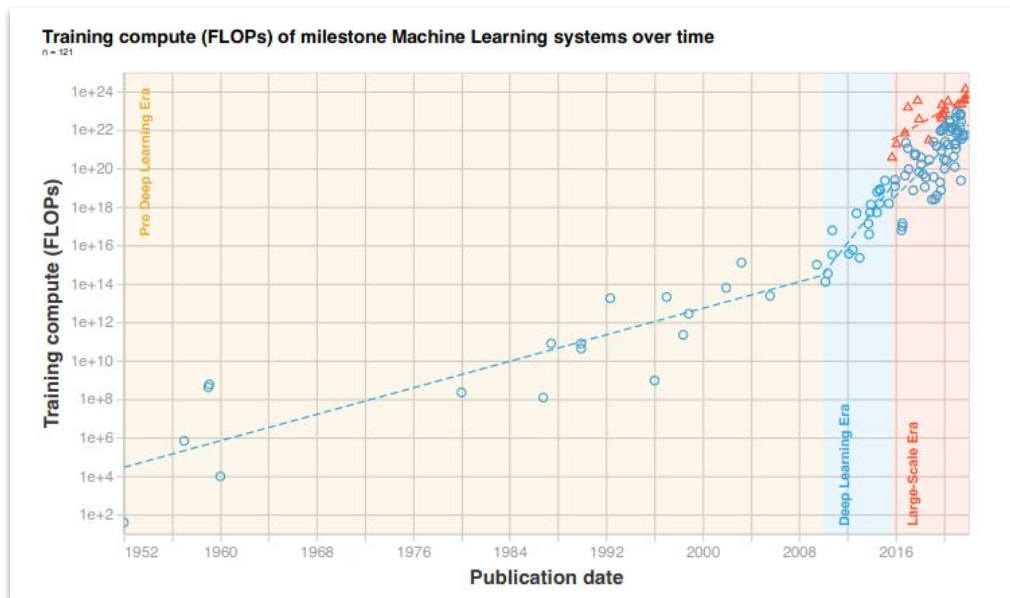
Pourquoi s'intéresser aux impacts de l'IA ?

impacts environnementaux
potentiellement importants :

- grande quantité de **données**
- ressources de **calcul**

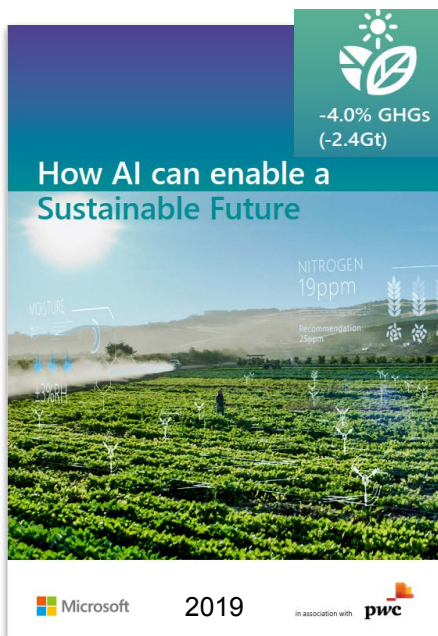
souvent présentée comme **solution**

... sans prendre en compte ses
impacts négatifs

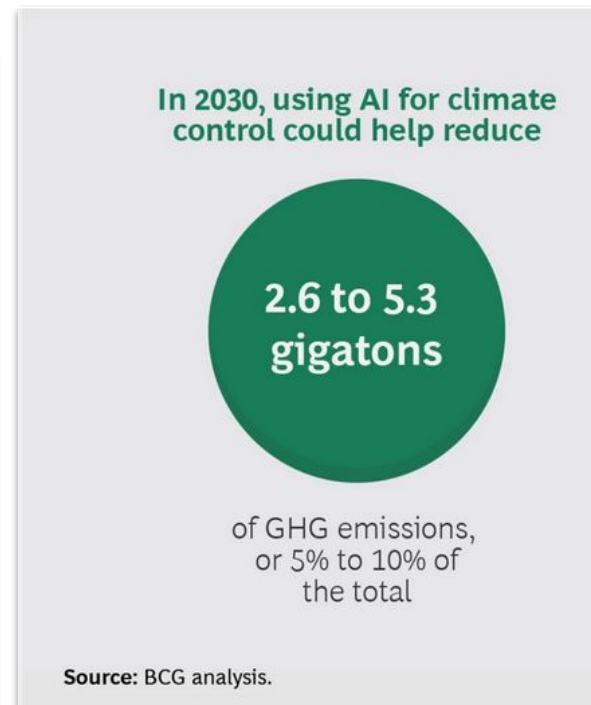


(Sevilla et al., 2022)

L'IA comme solution aux problèmes environnementaux ?



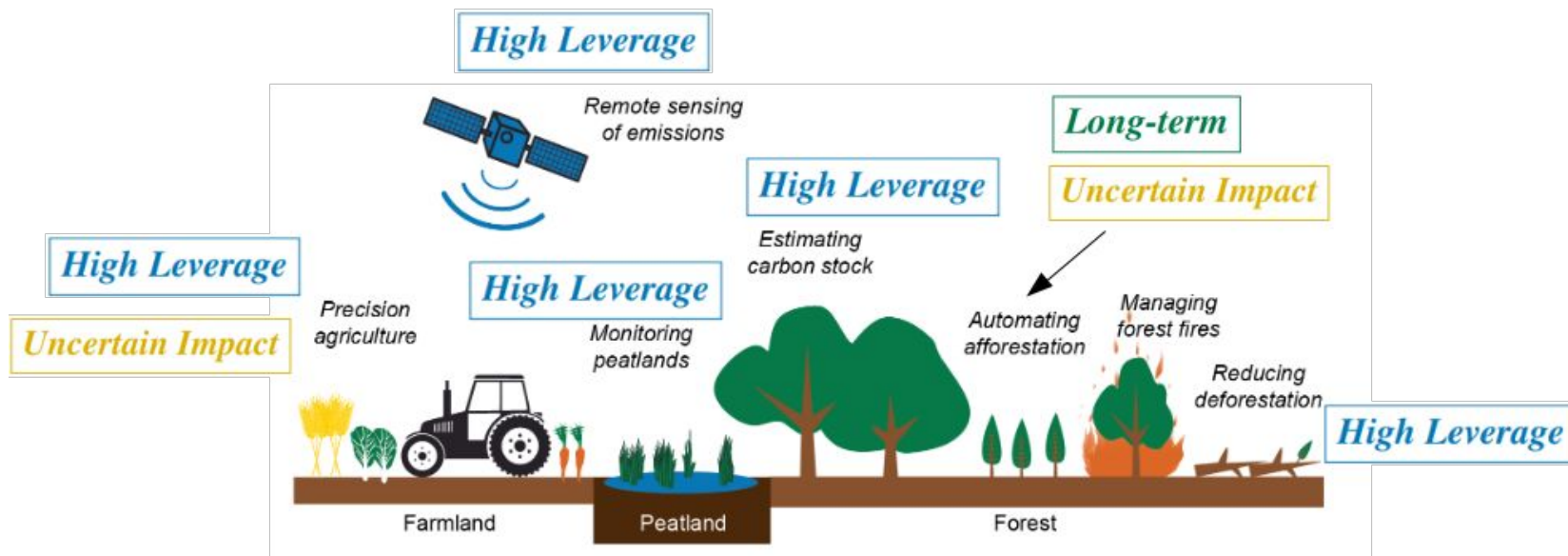
«AI can enable our future systems to be more productive for the economy and for nature. This supports the proposition that we can use AI to help 'decouple' economic growth from GHG emissions.»



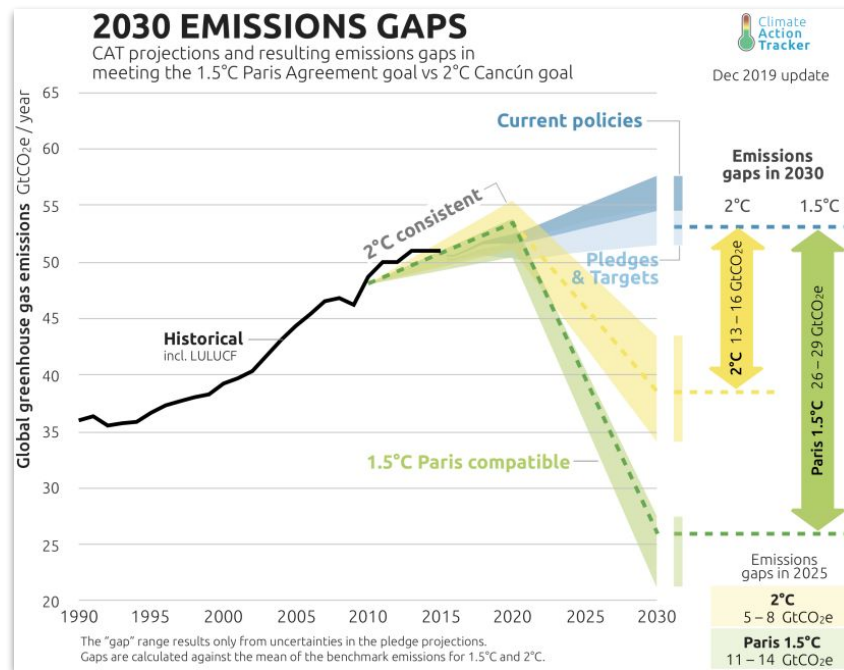
Applications de l'IA à des problématiques environnementales

«Tackling Climate Change with Machine Learning» (Rolnick et al., 2019)

Farms & Forests



Crises environnementales



en informatique, beaucoup de recherches déconnectées de ces problèmes



«Le poids du clic», La revue dessinée

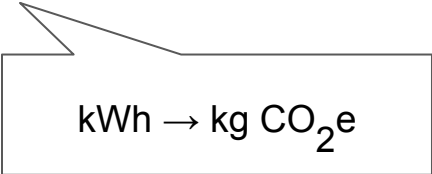
Comment mesurer ?

Approche bottom-up

sur mon serveur de calcul, ce qui va consommer de l'énergie supplémentaire du fait de l'exécution de mon programme d'IA c'est :

- processeur
- processeur graphique
- mémoire...

=> empreinte₁ = \sum (conso_{ressource}) x facteur d'émission élec

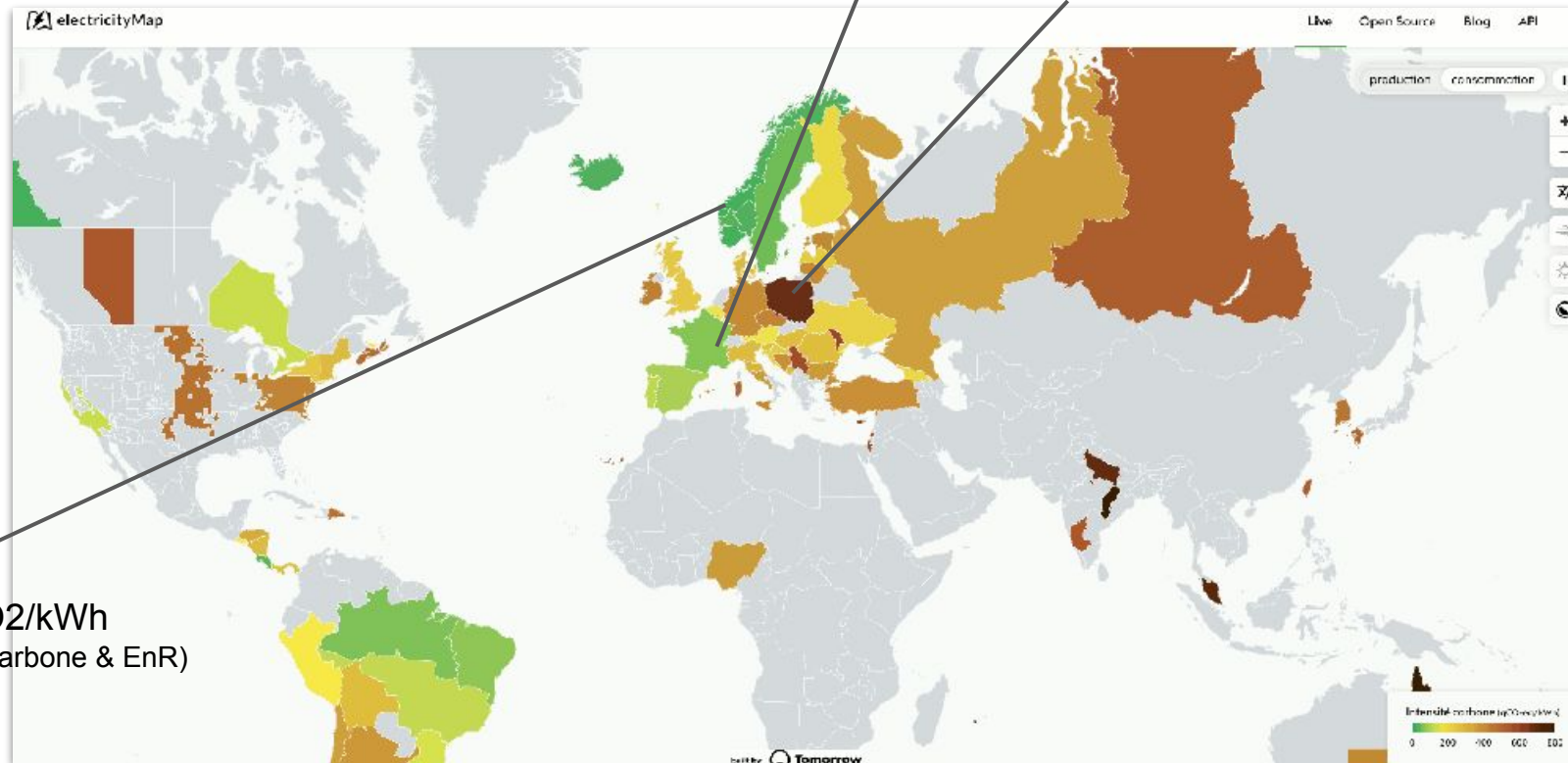


kWh → kg CO₂e

Facteur d'émission électricité

France : 101g éqCO₂/kWh
(86% bas carbone, 13% EnR)

Pologne : 927g éqCO₂/kWh
(13% bas carbone, 13% EnR)



Norvège :
22g éqCO₂/kWh
(100% bas carbone & EnR)

Influence du facteur d'émission

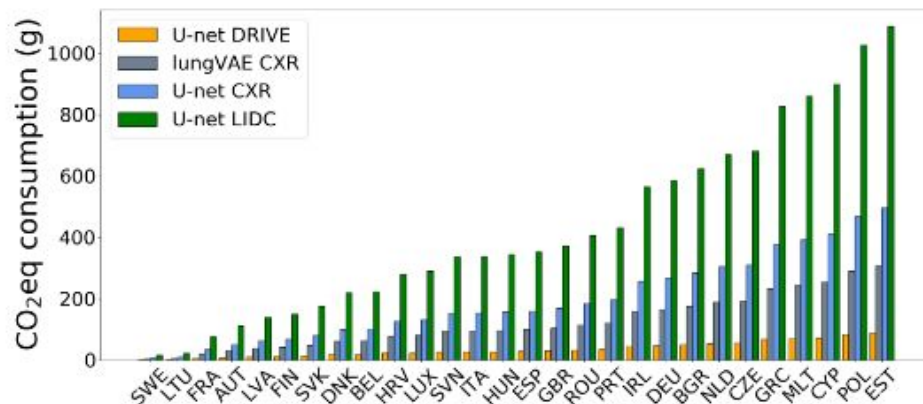


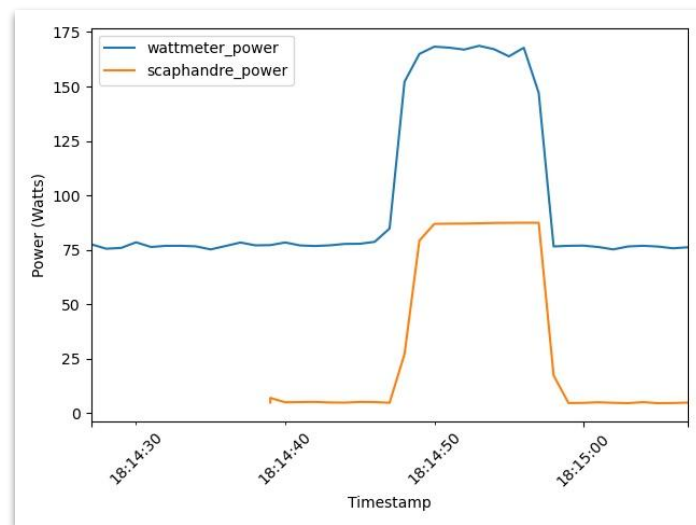
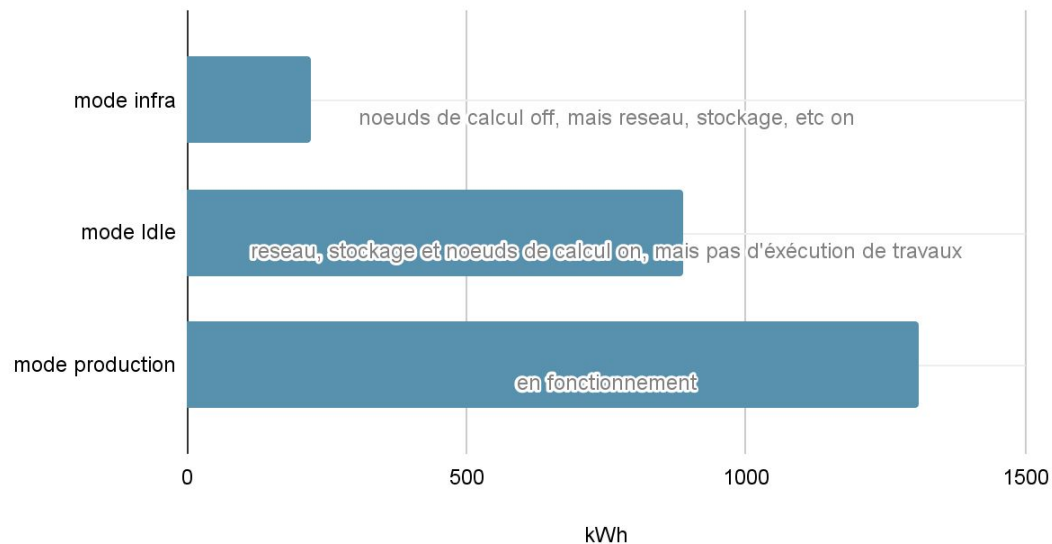
Figure 4. Estimated carbon emissions (gCO₂eq) of training our models (see [Appendix B](#)) in different EU-28 countries. The calculations are based on the average carbon intensities from 2016 (see [Figure 8](#) in Appendix).

Source : (Anthony et al., 2020)

Mesure-t-on la totalité des impacts dus
à l'exécution du programme ?

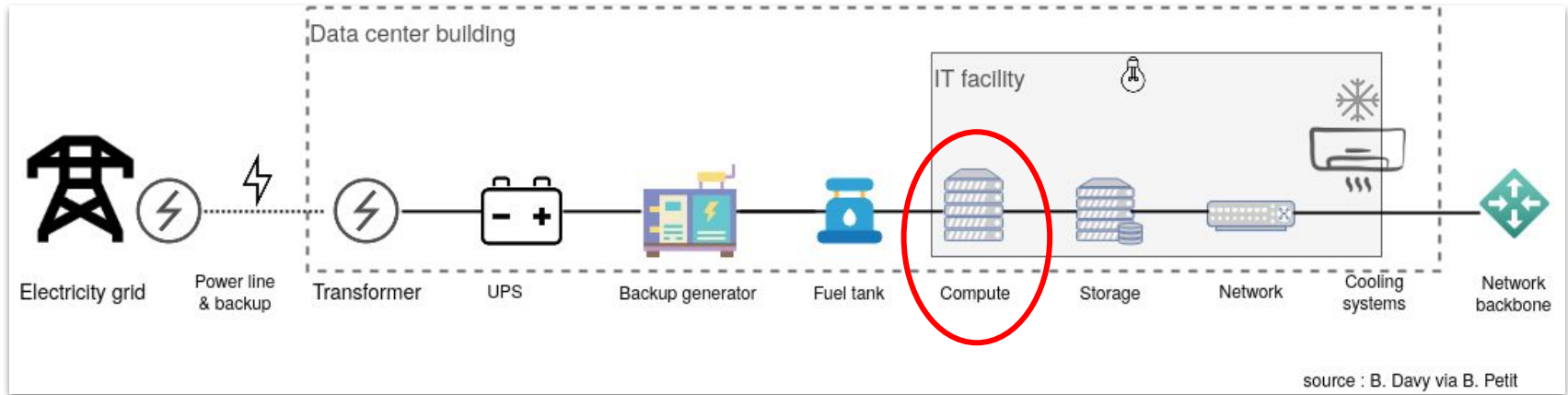
Consommation d'électricité dans un ordinateur

Consommation électrique Jean-Zay



Évaluer l'empreinte carbone d'un service d'IA

Quels équipements ?



Efficacité énergétique du centre de calcul

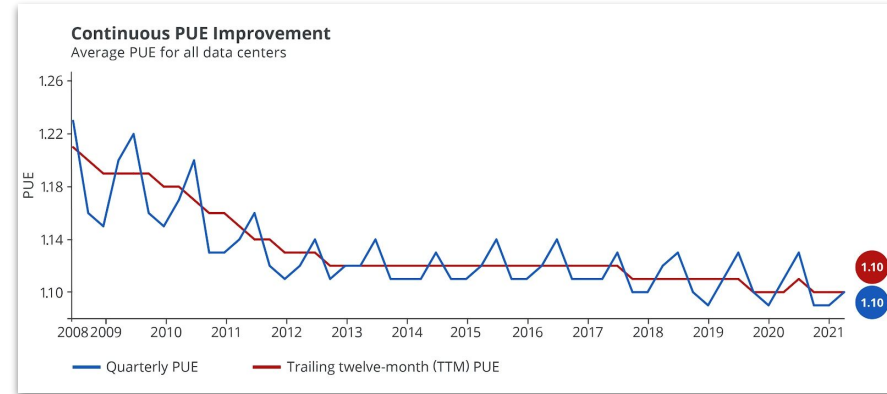
$$\text{PUE} = \frac{\text{consommation d'énergie du centre}}{\text{consommation des équipements informatiques}}$$

Autres équipements

refroidissement
éclairage
pertes élec...

Équipements
informatiques

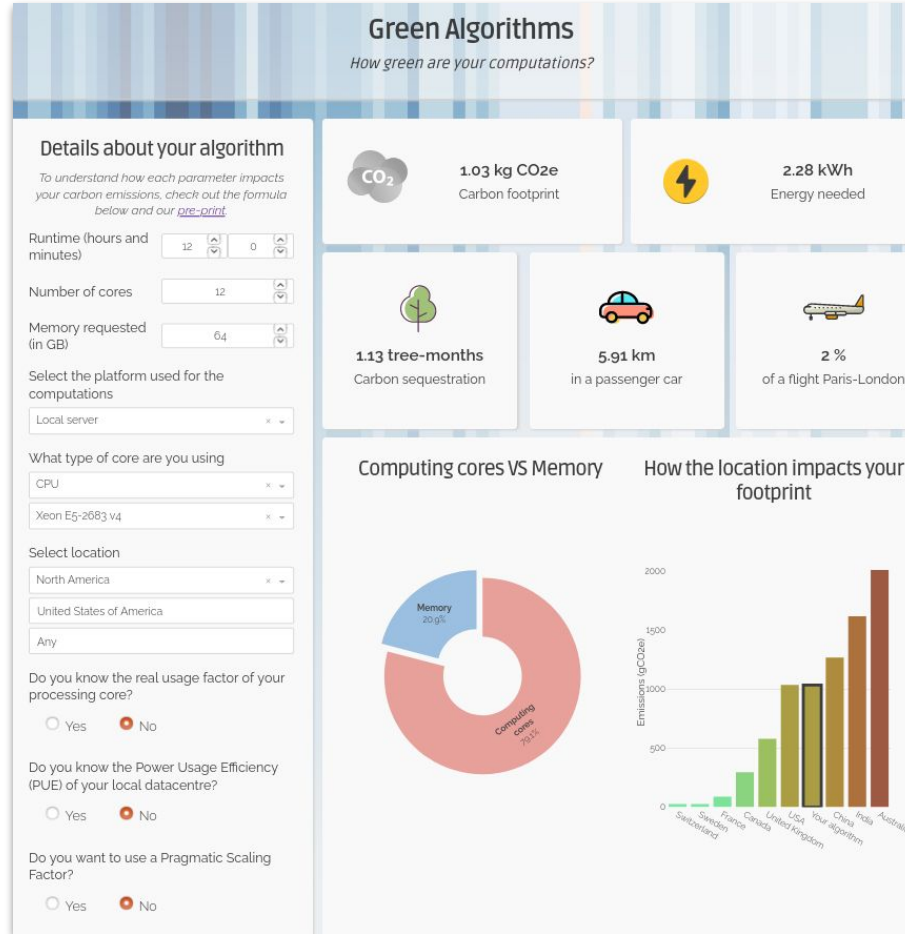
serveurs,
stockage...



PUE Google

$$\Rightarrow \text{empreinte}_2 = \text{empreinte}_1 \times \text{PUE}$$

Outils



Mais

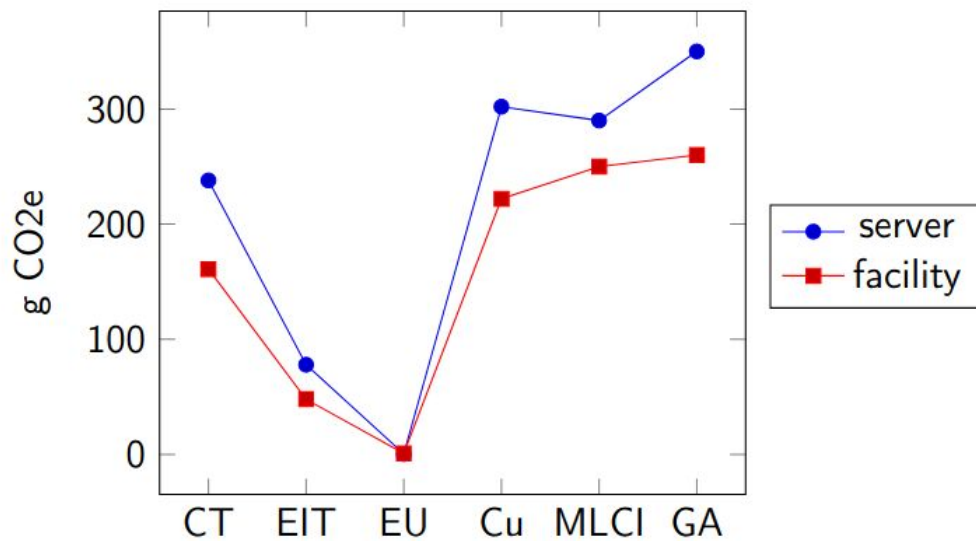


Figure – Nested NER Model on the French Broadcast News Corpus

Source : (Bannour et al, 2021)

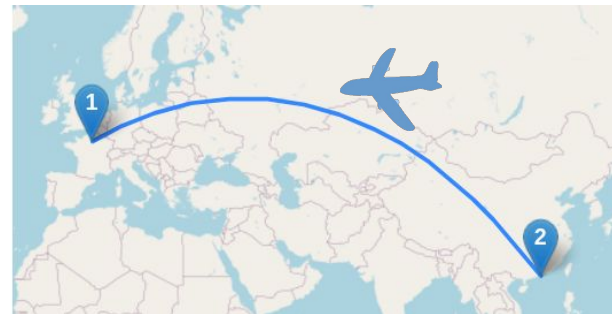
Exemple d'empreinte carbone (Strubell et al, 2019)

4 modèles de TAL/NLP état de l'art

mesure logicielle de la consommation

Résultats des entraînements

- quelques jours à quelques semaines
- émissions: entre 18kg CO₂e et 284 000 kg CO₂e
- modèle le plus utilisé : 652 kg CO₂e, soit
 - un aller Paris-Hong Kong en avion
 - ou 2 500km en voiture



envrion 58 GPU pendant 172 jours pour entraîner le modèle...

Précision vs CO2e (Parcollet et Ravanelli, 2021)

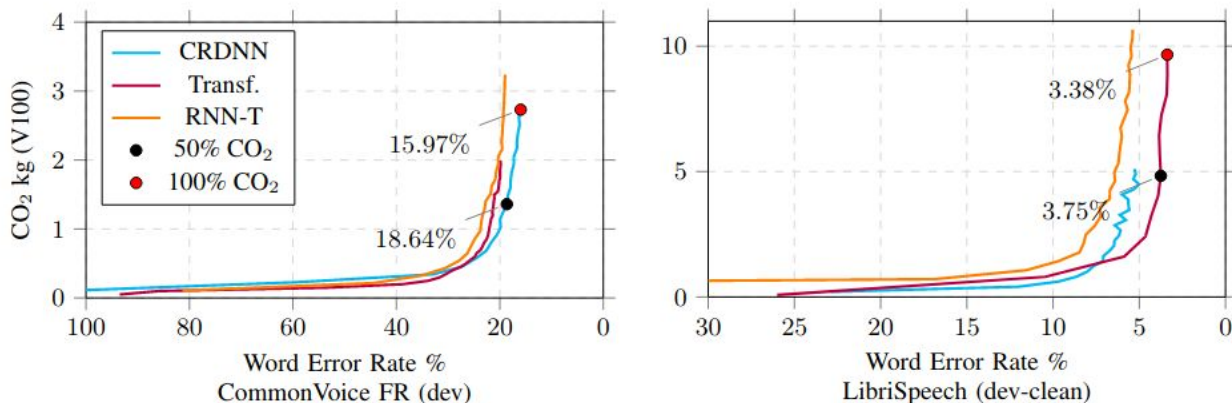
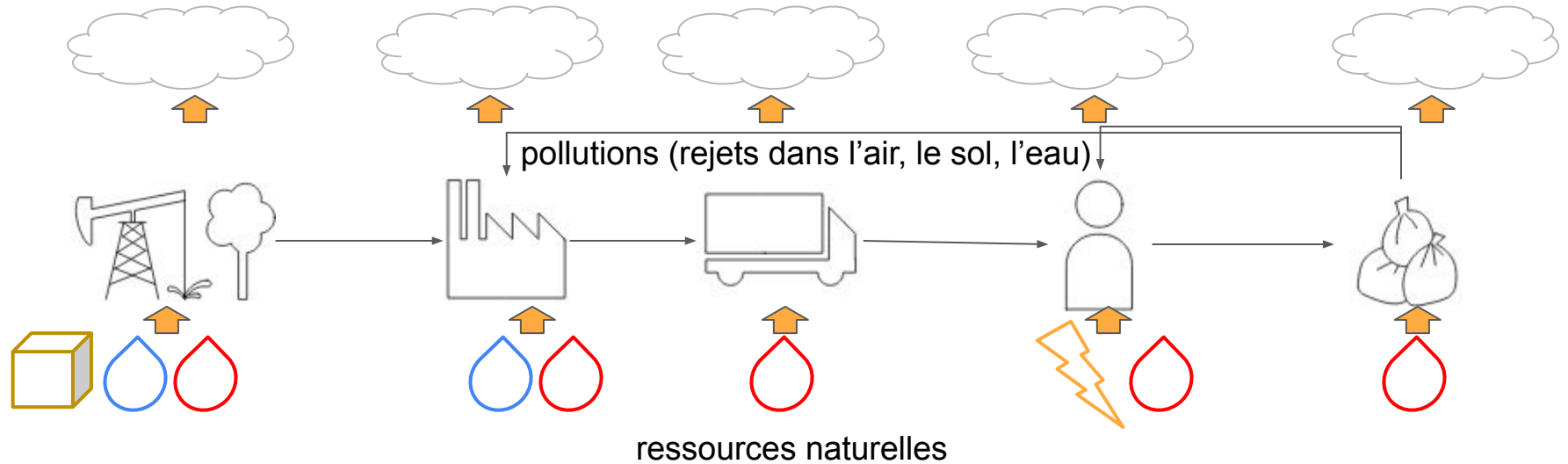


Figure 2: CO₂ emitted in kg (in France) by different E2E ASR models with respect to the word error rate (WER) on the dev sets of LibriSpeech and CommonVoice. The curves exhibit an exponential trend as most of the training time is devoted to slightly reduce the WER. The black and red dots indicates the WER obtained with 50% and 100% of the emitted CO₂. On LibriSpeech, 50% of the carbon emissions have been dedicated to reach SOTA results with an improvement of 0.37%.

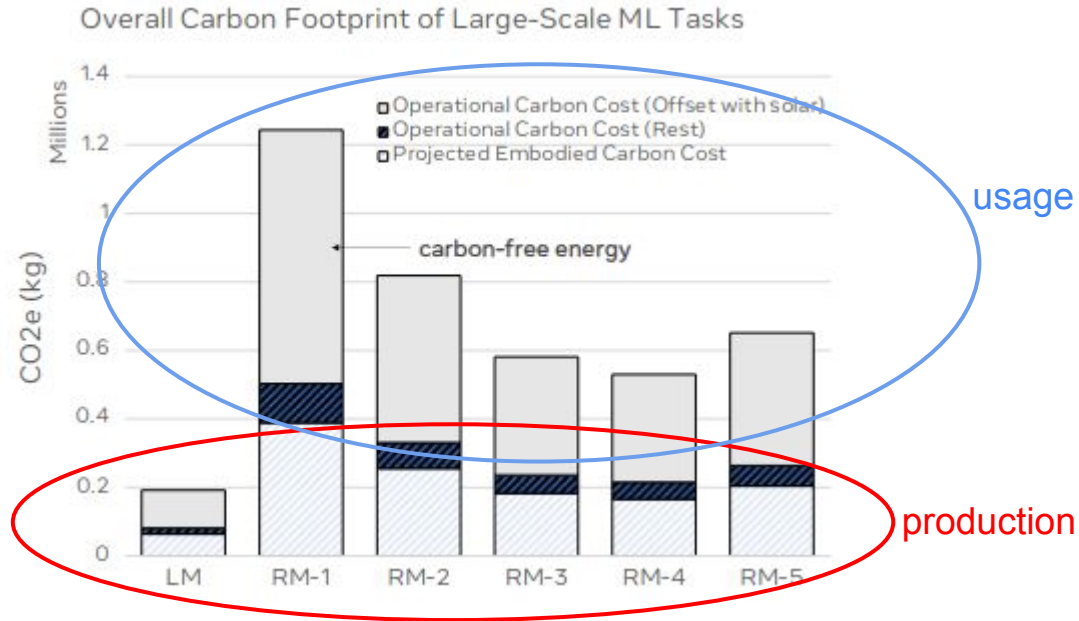
Cycle de vie des équipements et du modèle ?

Impacts environnementaux liés au cycle de vie des équipements



Énergie grise (production) vs opérationnelle (usage)

Étude Facebook



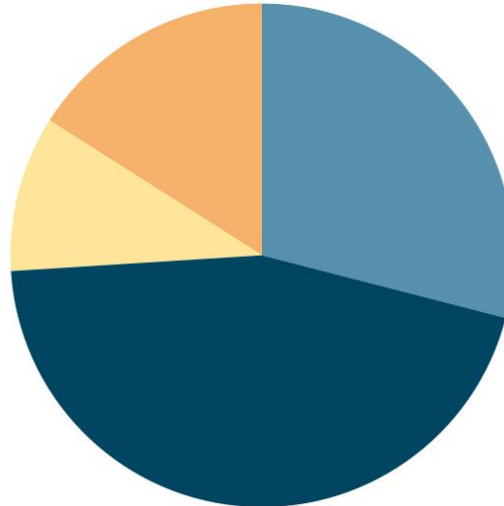
Source : (Wu et al., 2021)

Approche top down : exemple du GRICAD

Quid de la fin de vie ?

Empreinte carbone due aux serveurs

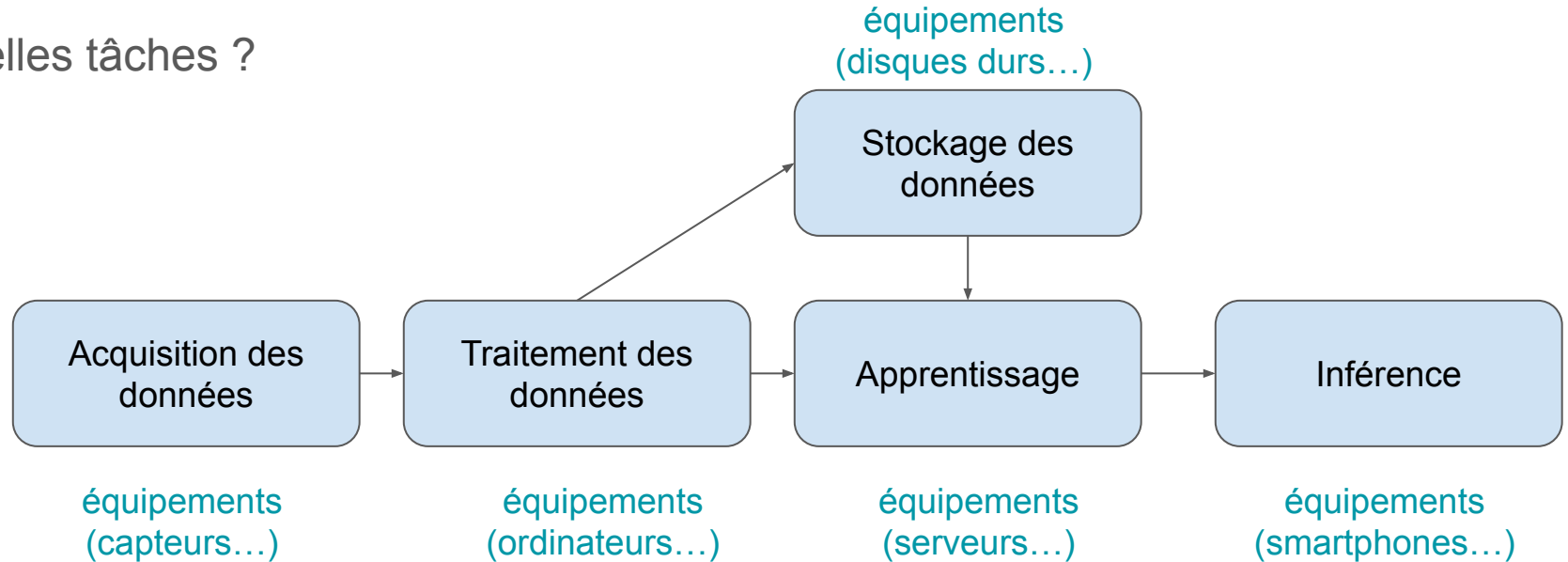
- Serveurs de calcul - production
- Serveurs de calcul - usage
- Autres serveurs - production
- Autres serveurs - usage



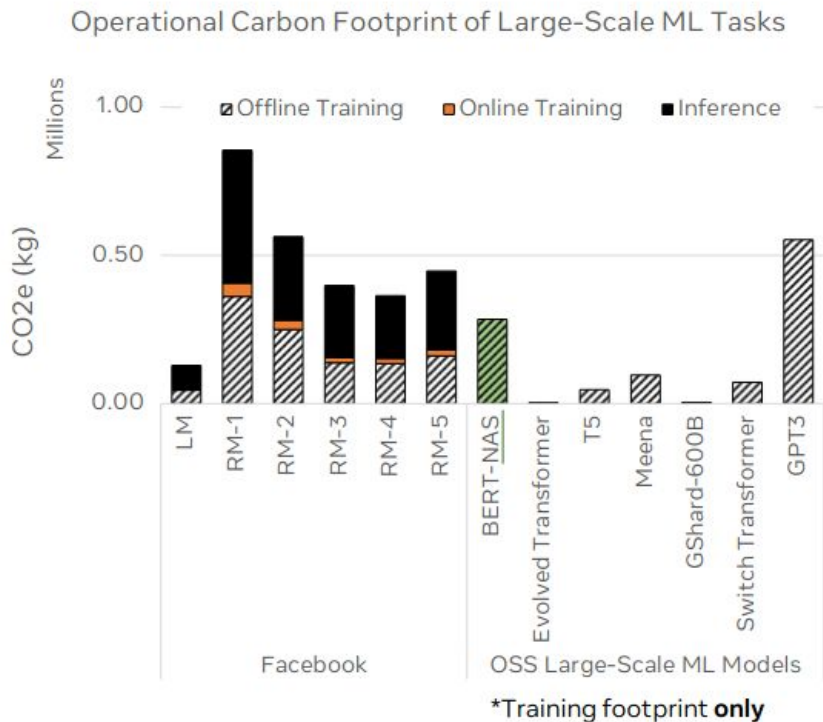
Source : (Berthoud et al., 2020)

Évaluer l'empreinte carbone d'un service d'IA

Quelles tâches ?

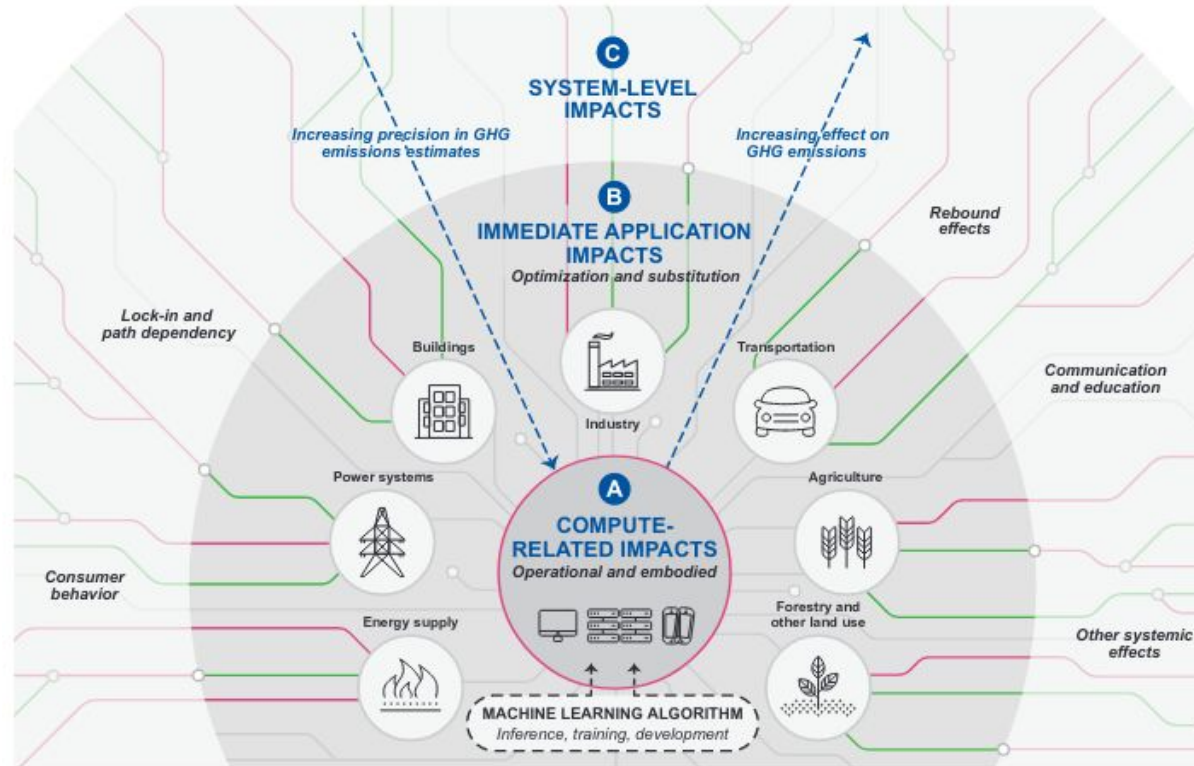


Entraînement vs inférence



Source : (Wu et al., 2021)

Impacts de 1er, 2e et 3e ordre de l'IA



Impacts indirects

optimiser le trafic
automobile ?



moins de consommation de
carburant

priorité aux systèmes avec
impacts significatifs ?

utilisation de nouveaux objets
connectés, capteurs...

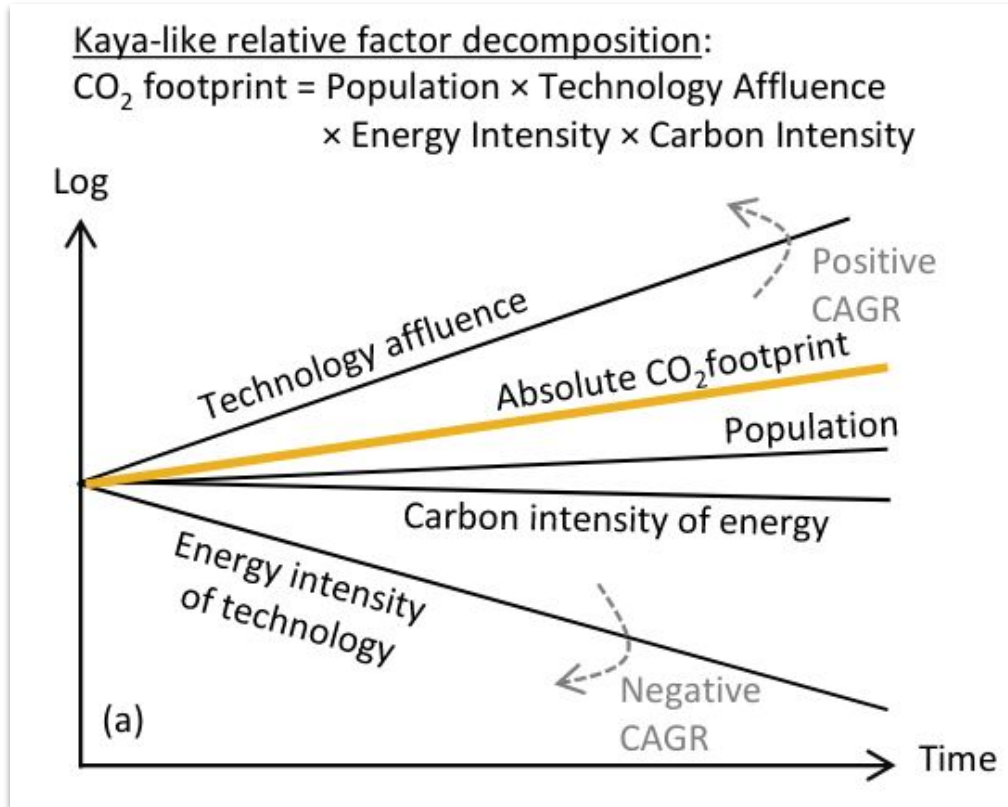
effet rebond

trafic plus fluide => gain de temps
=> éloignement du domicile =>
étalement urbain

dépendance de chemin

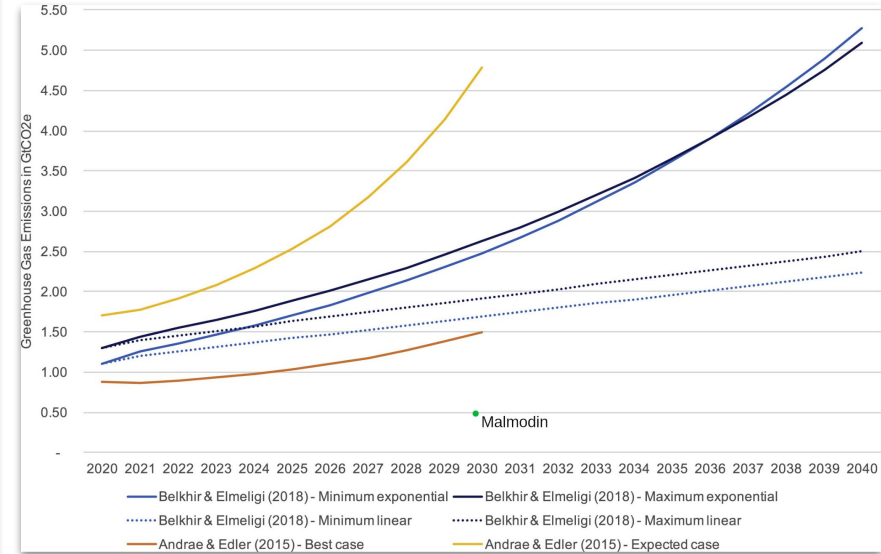
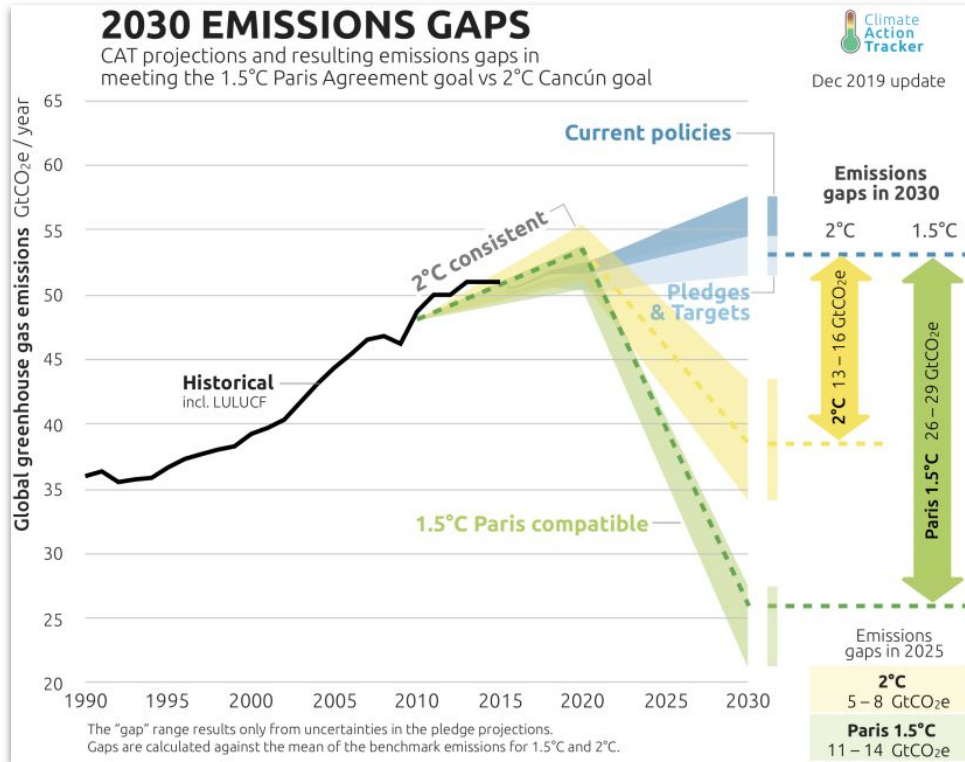
prolonge système actuel, vs
transports en commun, mobilités
actives...

Effet rebond sectoriel



source : Bol, D., Pirson, T., & Dekimpe, R. (2021). *Moore's Law and ICT Innovation in the Anthropocene*. In *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE.

Empreinte du “numérique”



(Freitag et al, 2021)

Autres indicateurs environnementaux

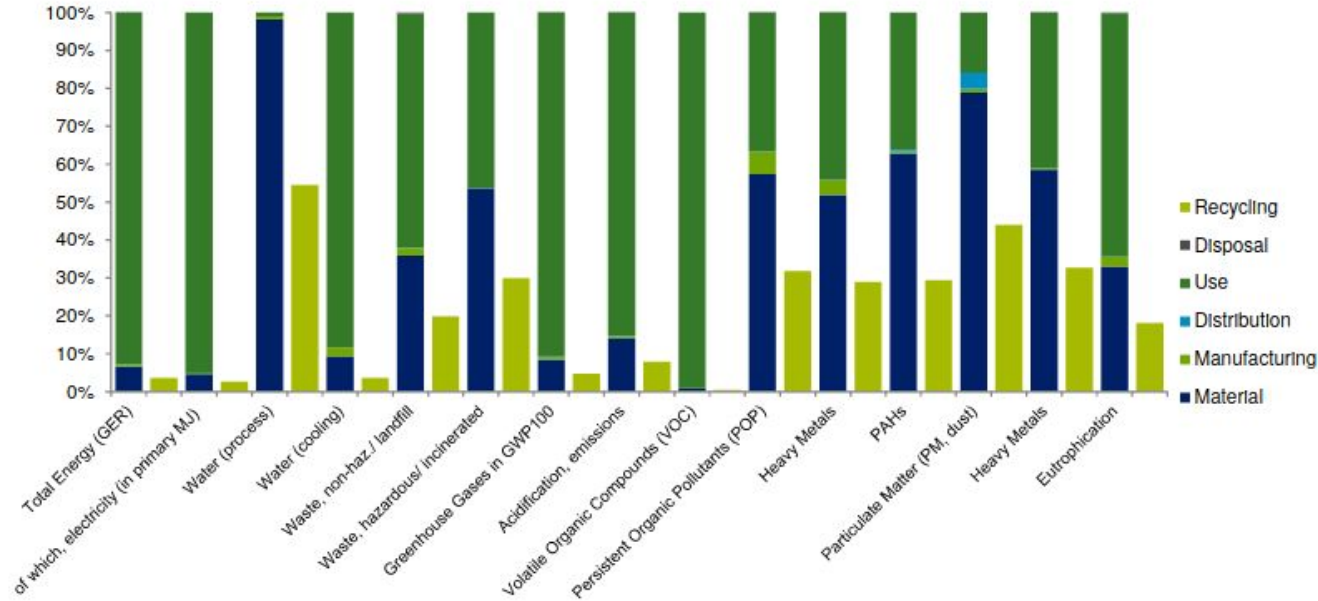
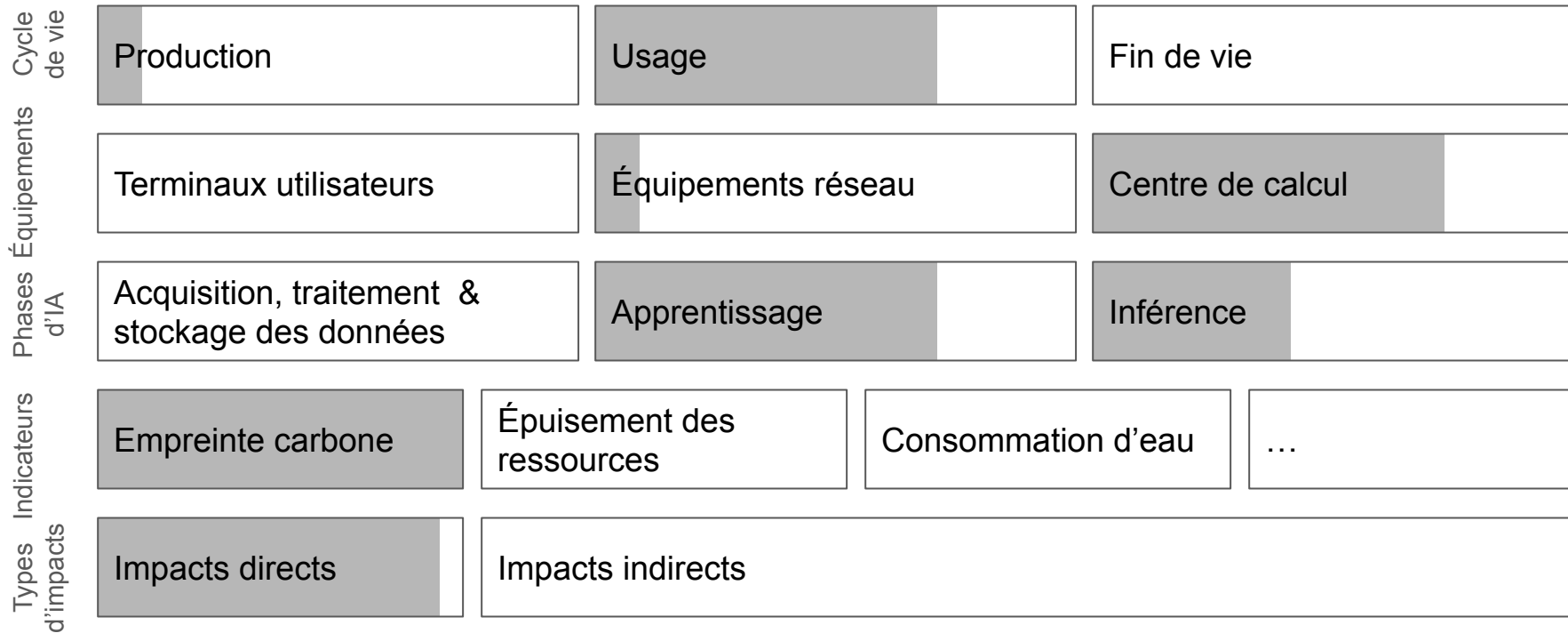


Figure 3: Distribution of BC-1 environmental impacts by life cycle phase²²

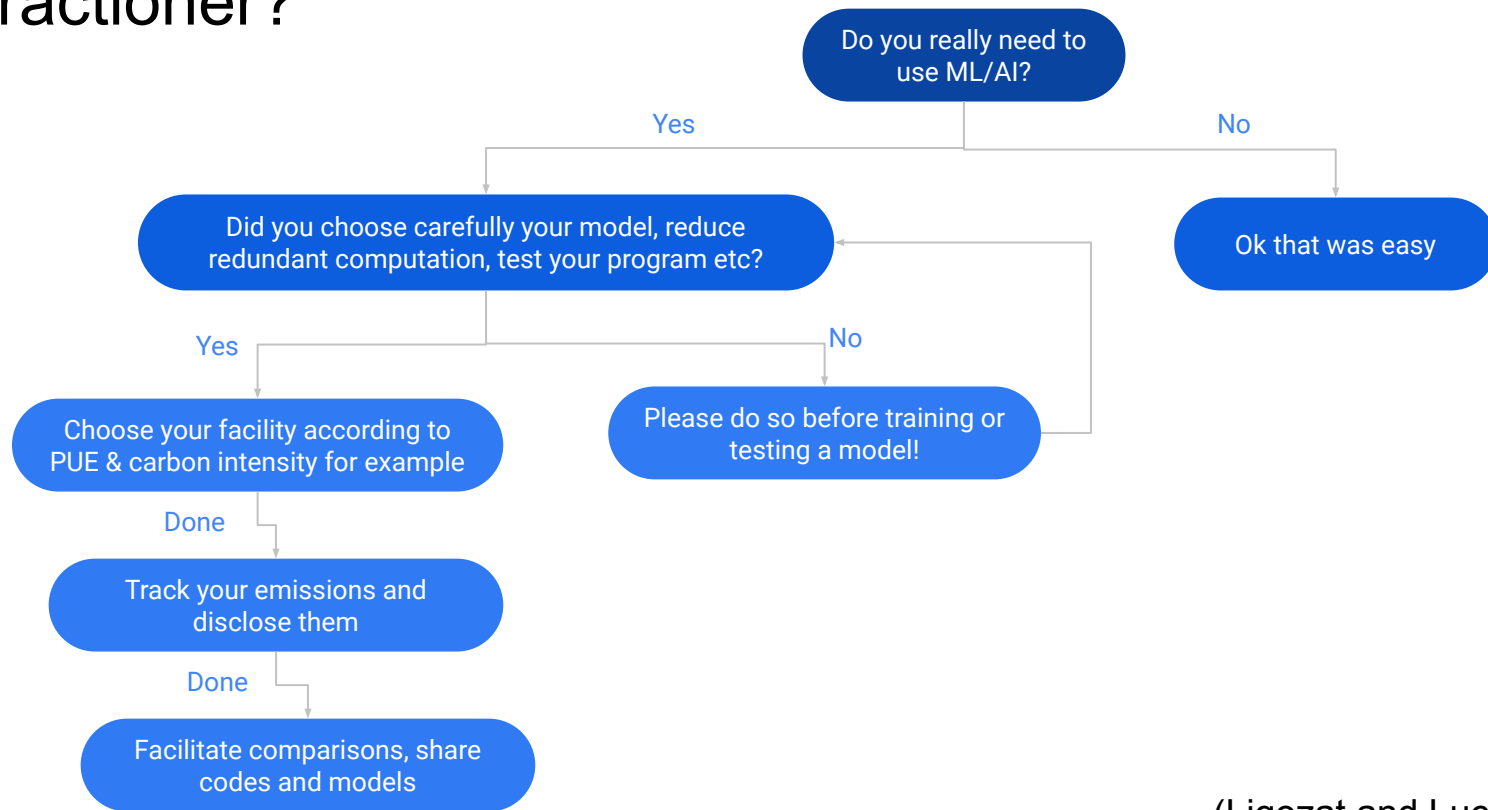
(serveur rack 2012)

Source : [Commission Européenne](#), 2015

Ce qu'on calcule actuellement



What can I do (to reduce my carbon footprint) as a ML/AI practitioner?



(Ligozat and Luccioni, 2021)

Climate performance model card (Hershcovich et al, 2022)

Minimum card

| Information | Unit |
|--|----------|
| 1. Is the resulting model publicly available? | Yes/No |
| 2. How much time does the training of the final model take? | Time |
| 3. How much time did all experiments take (incl. hyperparameter search)? | Time |
| 4. What was the energy consumption (GPU/CPU)? | Watt |
| 5. At which geo location were the computations performed? | Location |

Extended card

| | |
|---|-----------------------------|
| 6. What was the energy mix at the geo location? | gCO ₂ eq/ kWh |
| 7. How much CO ₂ eq was emitted to train the final model? | kg |
| 8. How much CO ₂ eq was emitted for all experiments? | kg |
| 9. What is the average CO ₂ eq emission for the inference of one sample? | kg |
| 10. Which positive environmental impact can be expected from this work? | Notes |
| 11. Comments | Notes |

Critères d'évaluation environnementale pour projets IA

- Impacts des équipements numériques
 - fabrication et fin de vie
 - usage : calcul, données
- Justification de la méthode
 - nécessité de l'IA
 - résilience
- Impacts dus aux changements sociétaux
 - scénario de référence
 - effets indirects potentiels



**Proposition de document de cadrage Évaluation
environnementale de projets impliquant des méthodes
d'IA**

Laurent Lefèvre, Anne-Laure Ligozat, Denis Trystram, Sylvain Bouveret,
Aurélien Bugcau, Jacques Combaz, Emmanuelle Frenoux, Gaël Guennebaud,
Julien Lefèvre, Jean-Philippe Nicolai

IA responsable ?

- (Schwartz et al., 2020) proposent **Red** vs **Green** AI

-  (Dilhac et al., 2018)

< >
Déclaration de Montréal
IA responsable_
< / >

- les systèmes d'IA et équipements associés doivent viser la plus grande efficacité énergétique et minimiser l'empreinte carbone sur l'ensemble de leur cycle de vie, les impacts sur les écosystèmes et la biodiversité...
- **Rapport Villani (2018)**
 - (...) l'IA peut entraîner de nombreux effets rebonds. Ainsi, l'IA peut nous éviter de repenser nos modes de croissance, de consommation et de mesure des richesses produites, et nous amener à consommer tout autant, voire plus qu'auparavant



Réponse de Google à l'article de (Strubell et al., 2019)

The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink

David Patterson^{1,2}, Joseph Gonzalez², Urs Hölzle¹, Quoc Le¹, Chen Liang¹, Lluís-Miquel Munguia¹, Daniel Rothchild², David So¹, Maud Texier¹, and Jeff Dean¹

Bonnes pratiques proposées :

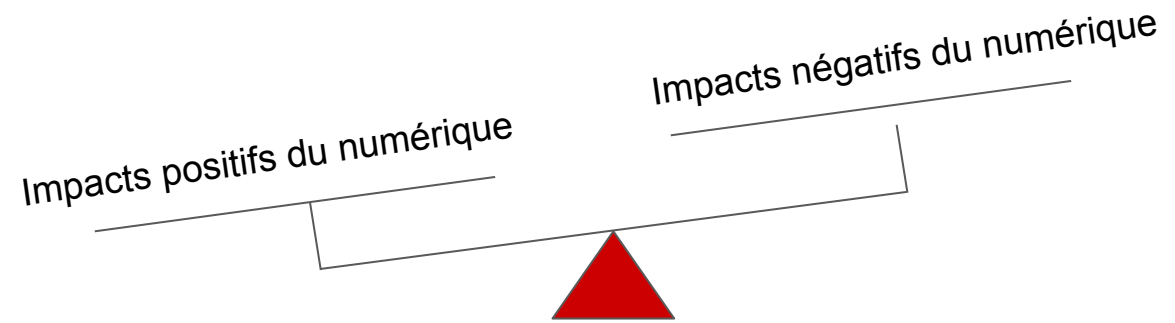
- Modèle efficient
- Processeurs optimisés pour ML
- Cloud pour efficacité énergétique
- Localisation avec mix électrique bas carbone

et pour finir «Google's renewable energy purchases further reduce the impact to zero»

mais :

- quid du cycle de vie ?
 - processeurs récents ⇒ empreinte carbone ↗
- quid de l'inférence ?
- énergie «carbon free» ou «net zero impact» ?
- empreinte carbone potentielle si tout optimisé, et non réelle

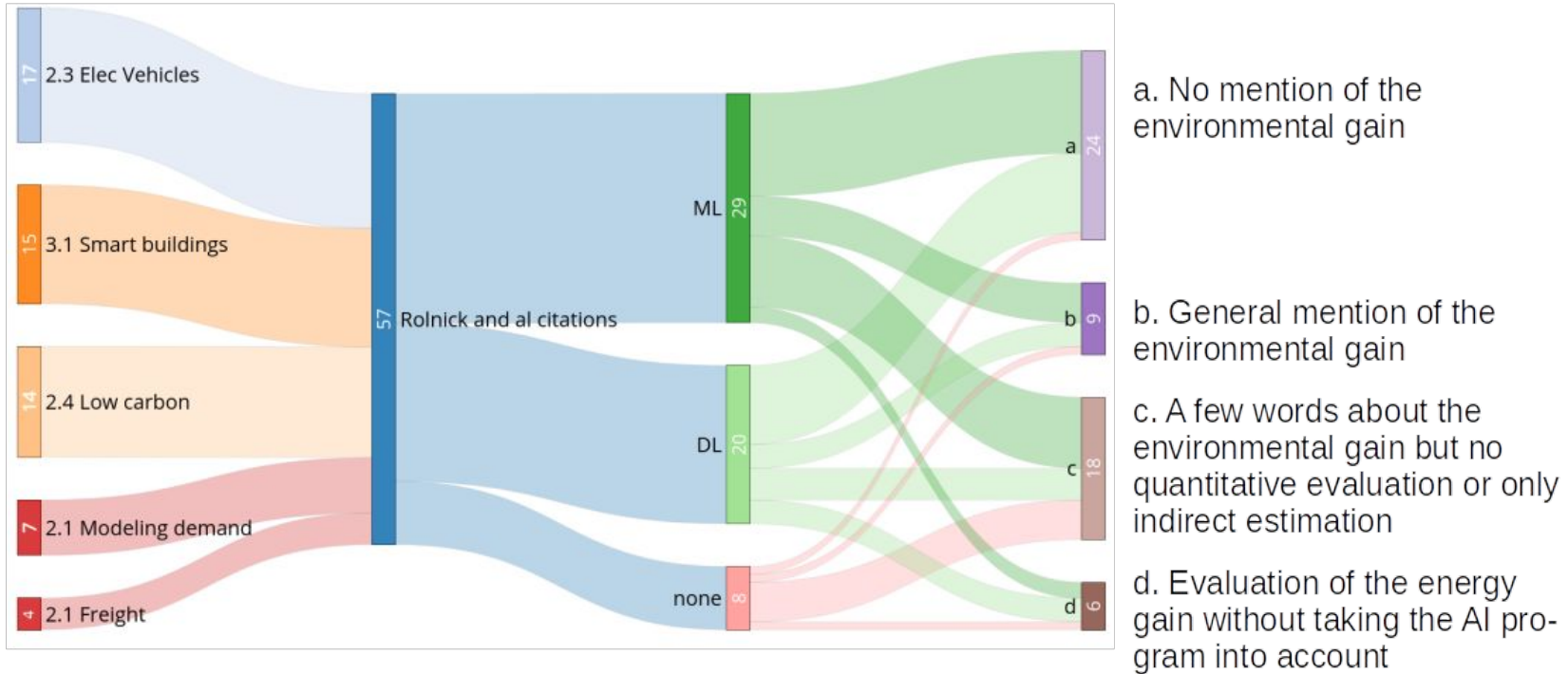
L'IA pour des applications environnementales



au moins avec des Analyses de Cycle de vie

en prenant en compte autant d'effets indirects que possible

Evaluations dans applications de (Rolnick et al., 2019)



Biais des études d'impact (Rasoldier et al., 2022)

Périmètre

- pas de prise en compte du cycle de vie : (Ligozat et al., 2021) pour l'IA
- pas de prise en compte des effets indirects : 5G

Hypothèses

- comparaison à quel scénario de référence ?

Déconnexion de scénarios globaux

- bénéfices minimales + incertitudes mal gérées
- incompatibilité entre les mesures

Discussion

Utilisation raisonnée de l'IA

Prendre en compte les impacts environnementaux impacts directs et indirects lors de déploiement de systèmes d'IA. Mais :

Complexité de l'évaluation

- difficultés méthodologiques, techniques, de mise à jour...

Limites de la quantification

- attention aux résultats hors contexte, aux "indicateurs d'efficacité"...

Réfléchir de façon systémique et interdisciplinaire

Merci !



Références

- Anthony, L. F. W., Kanding, B., & Selvan, R. (2020). Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models. *ArXiv:2007.03051 [Cs, Eess, Stat]*. <http://arxiv.org/abs/2007.03051>
- Berthoud, F.; Bzezniak, B.; Gibelin, N.; Laurens, M.; Bonamy, C.; Morel, M.; Schwindenhammer, X. Estimation de l'empreinte carbone d'une heure.coeur de calcul. Research report, UGA - Université Grenoble Alpes ; CNRS ; INP Grenoble ; INRIA, 2020. <https://hal.archives-ouvertes.fr/hal-02549565v4/>
- Dilhac, M.-A., Abrassart, C., Voarino, N., et al. (2018). Rapport de la déclaration de montréal pour un développement responsable de l'intelligence artificielle. <https://www.declarationmontreal-iaresponsable.com/la-declaration>
- Guyon, D. (2018). Supporting energy-awareness for cloud users. PhD thesis, Université Rennes 1.
- Heinrich, F. C., Cornebize, T., Degomme, A., Legrand, A., Carpen-Amarie, A., Hunold, S., Orgerie, A.-C., and Quinson, M. (2017). Predicting the energy-consumption of mpi applications at scale using only a single node. In 2017 IEEE International Conference on Cluster Computing (CLUSTER).
- Hershovich, D., Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2022). Towards Climate Awareness in NLP Research. arXiv:2205.05071 [cs]. <http://arxiv.org/abs/2205.05071>
- Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2021). *Aligning artificial intelligence with climate change mitigation*. <https://hal.archives-ouvertes.fr/hal-03368037>
- Ligozat, A.-L., Lefèvre, J., Bugeau, A., & Combaz, J. (2021). Unraveling the hidden environmental impacts of AI solutions for environment. *arXiv:2110.11822 [cs]*. <http://arxiv.org/abs/2110.11822> and Sustainability <https://www.mdpi.com/2071-1050/14/9/5172/htm>
- Ligozat, A.L.; Luccioni, A. A Practical Guide to Quantifying Carbon Emissions for Machine Learning researchers and practitioners. 462 Technical report, Bigscience project, LISN and MILA, 2021. <https://hal.archives-ouvertes.fr/hal-03376391>
- Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., ... (2019). Tackling Climate Change with Machine Learning. *ArXiv:1906.05433 [Cs, Stat]*. <http://arxiv.org/abs/1906.05433>
- Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54-63. <https://doi.org/10.1145/3381831>
- Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M., & Villalobos, P. (2022). Compute Trends Across Three Eras of Machine Learning. *arXiv:2202.05924 [cs]*. <http://arxiv.org/abs/2202.05924>
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. *ArXiv:1906.02243 [Cs]*. <http://arxiv.org/abs/1906.02243>
- Villani, Cédric. (2018). *For a meaningful artificial intelligence*. https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
- Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Behram, F. A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott, M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H.-H. S., ... Hazelwood, K. (2021). Sustainable AI: Environmental Implications, Challenges and Opportunities. *arXiv:2111.00364 [cs]*. <http://arxiv.org/abs/2111.00364>