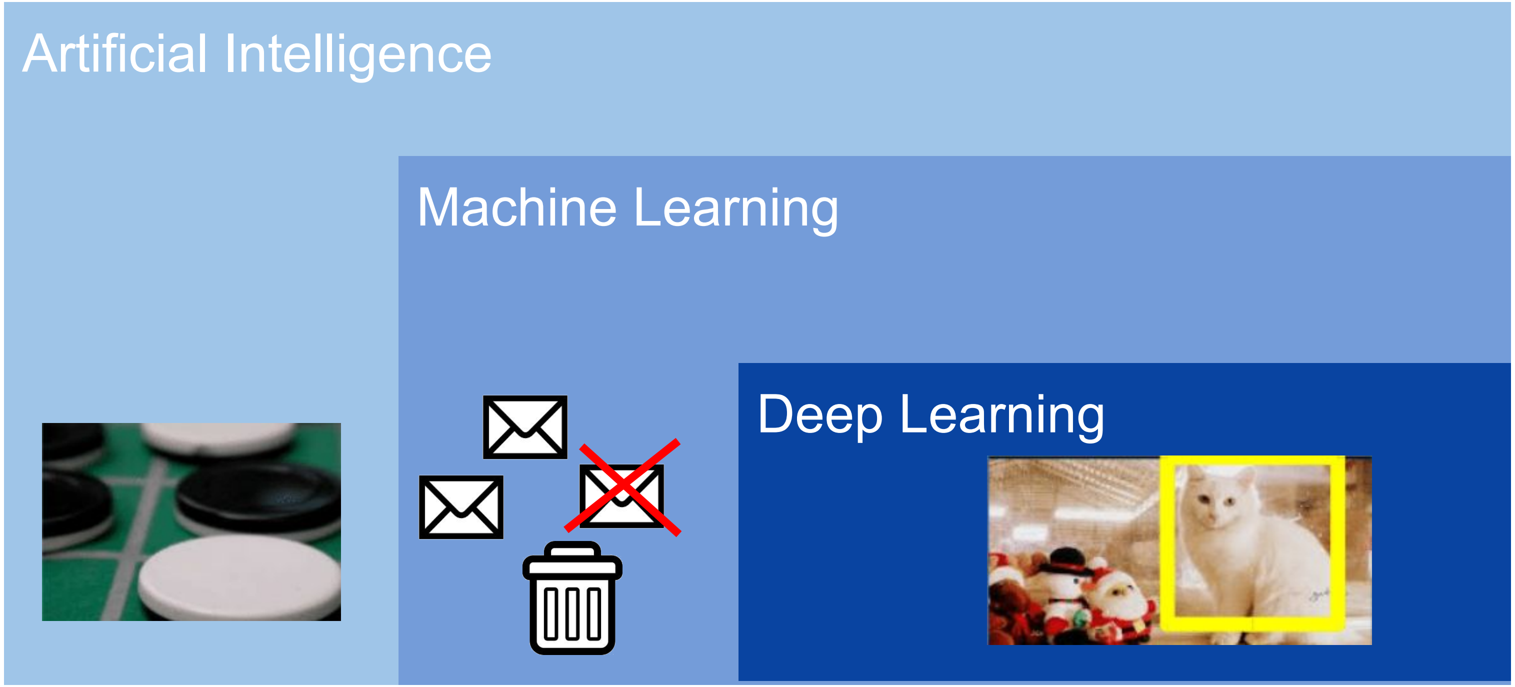
Environmental impacts of Artificial Intelligence

Cours Intelligence Artificielle et Environnement
Master MVA

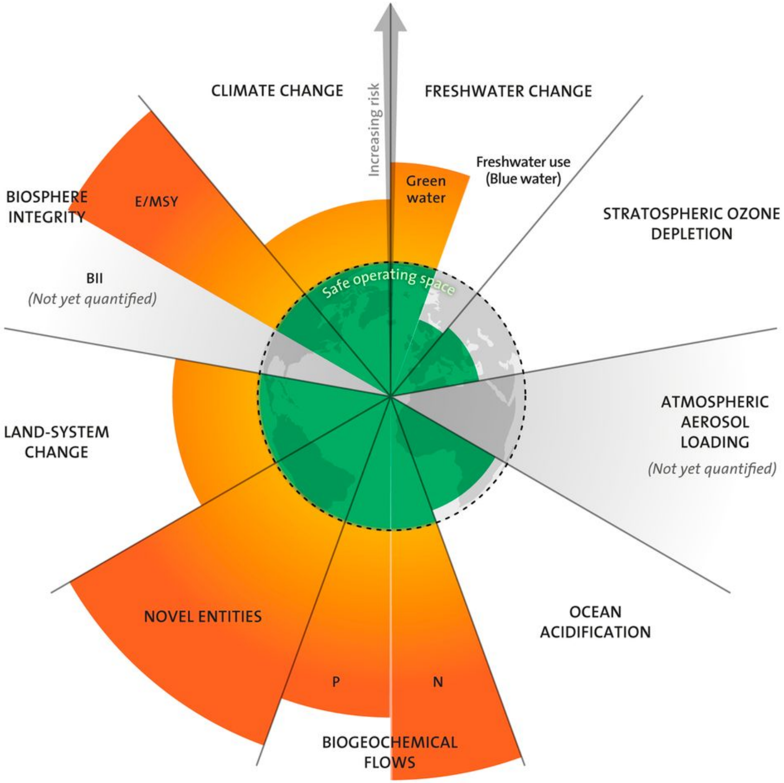
Anne-Laure Ligozat

AI?



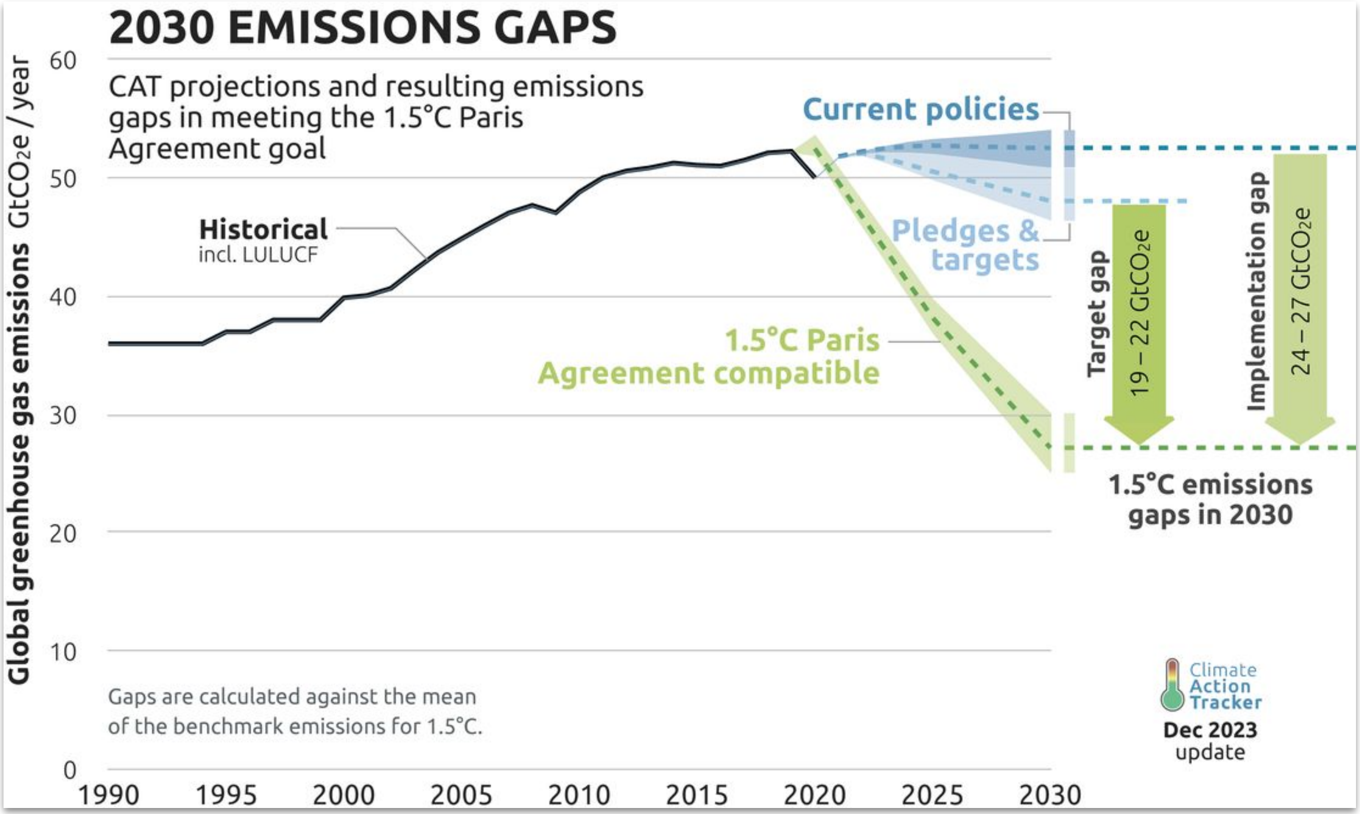
Context

Environmental context



Planetary boundaries, source: Wikipedia

Environmental context



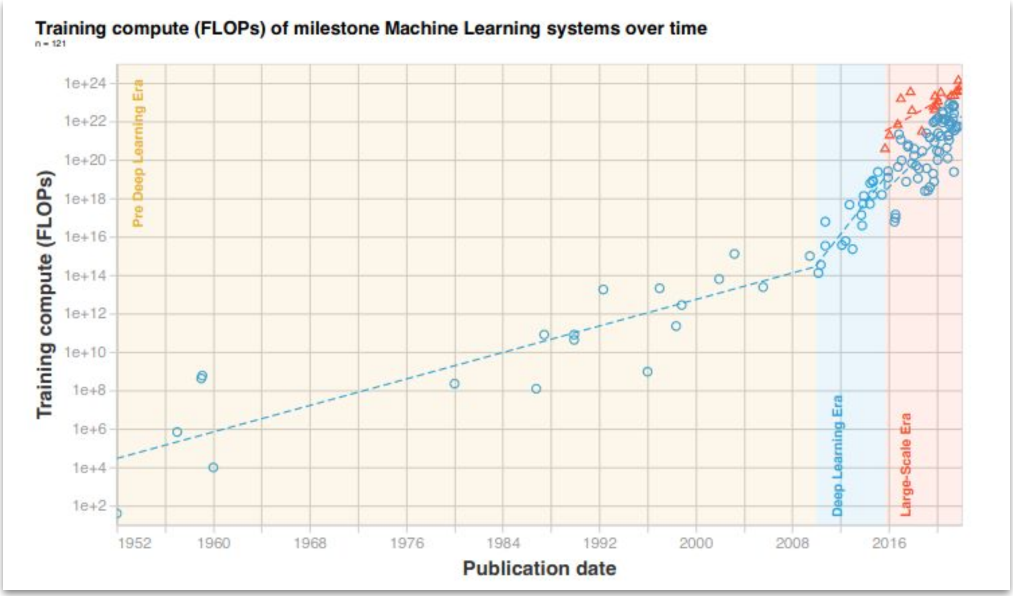
Why AI?

potential high environmental impacts:

- massive **data**
- **computation** demand

often presented as a **solution**

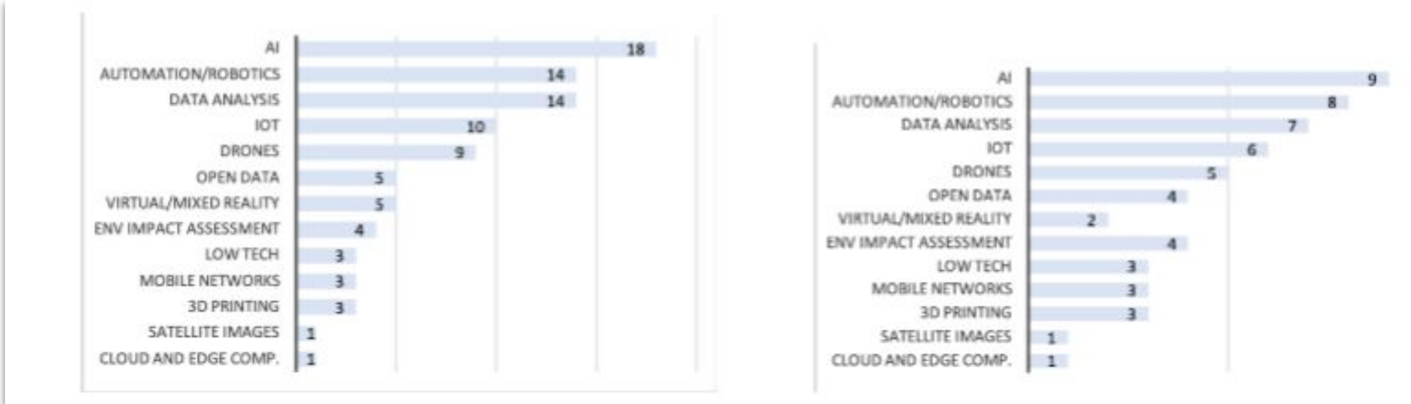
… without considering its negative impacts



(Sevilla et al., 2022)

AI as a solution?

Work on prospective studies (Bugeau & Ligozat, 2023)

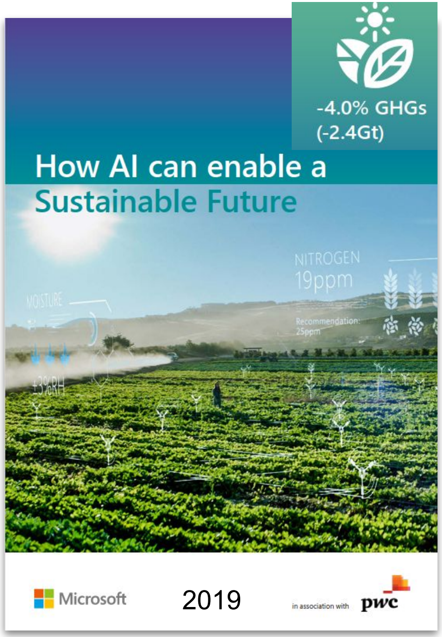


a) Digital technologies by scenario

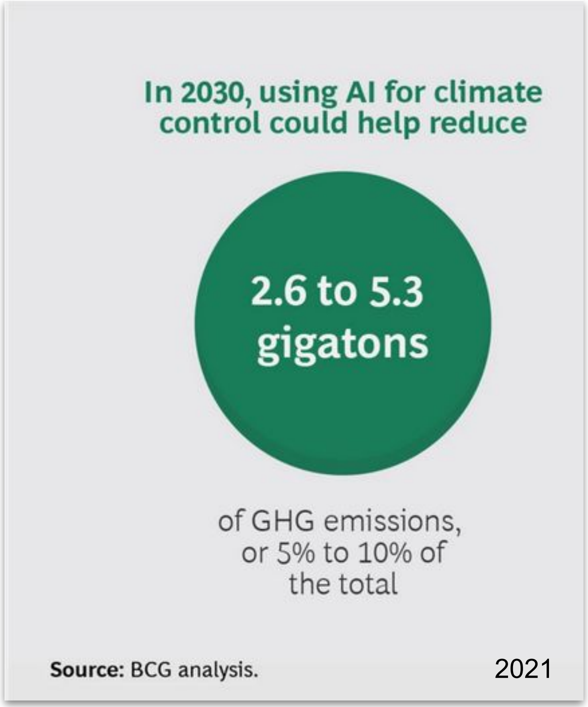
b) Digital technologies by studies

IPCC 2022	*DDC* 2020
Ademe 2022	*SNBC* 2020
negaWatt 2021	*RTE* 2022
EU green deal 2019	*Shift* 2020
Eionet 2022	*France 2072* 2018
Arup 2019	*D&A* 2022
	CNIL 2021
	Digit. Challenge 2022

AI as a solution?

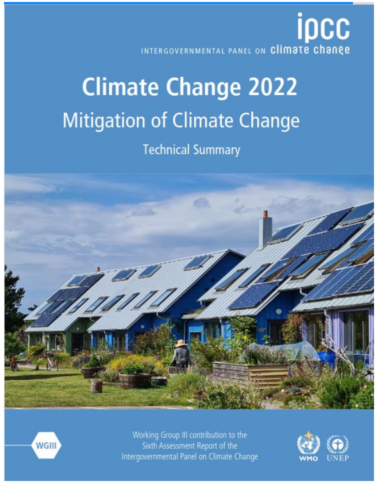


«AI can enable our future systems to be more productive for the economy and for nature. This supports the proposition that we can use AI to help 'decouple' economic growth from GHG emissions.»



AI as a solution?

«Artificial Intelligence (AI) can be deployed for a wide range of applications to promote the goals of the European Green Deal. However, adverse environmental impacts of AI could jeopardise the attainment of these goals.»



(....) artificial intelligence can improve energy management in all sectors, increase energy efficiency, and promote the adoption of many low-emission technologies, including decentralised renewable energy, while creating economic opportunities. However, some of these climate change mitigation gains can be reduced or counterbalanced by growth in demand for goods and services due to the use of digital devices.

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

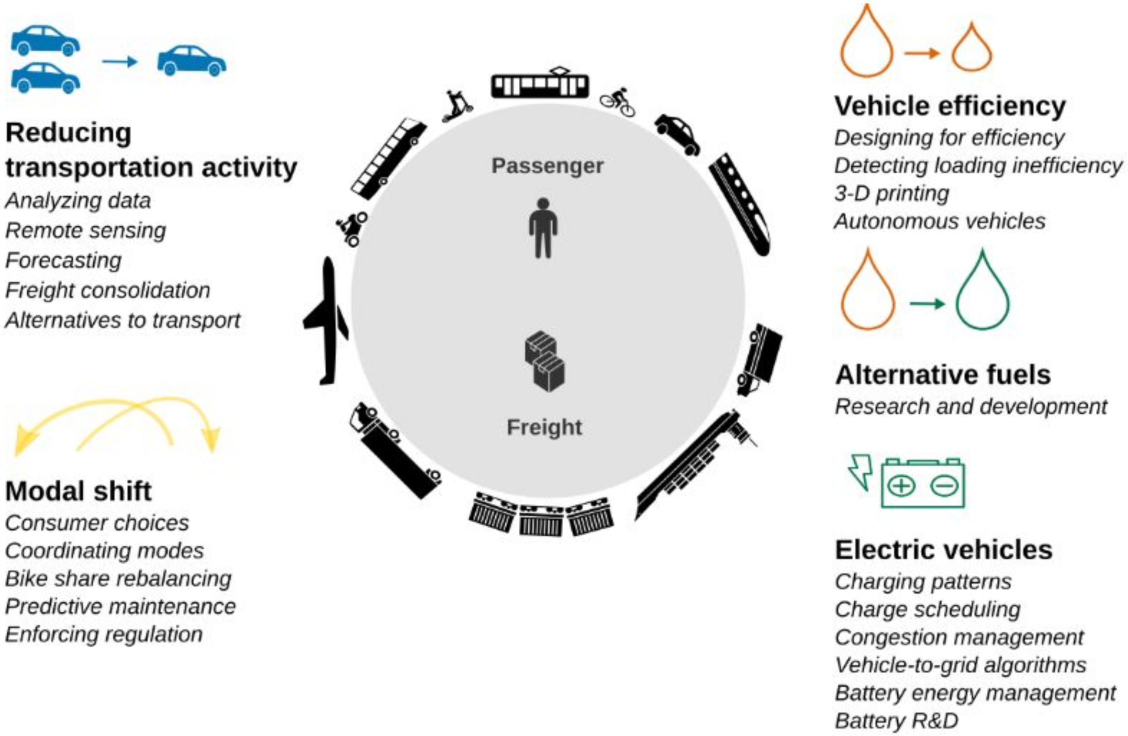
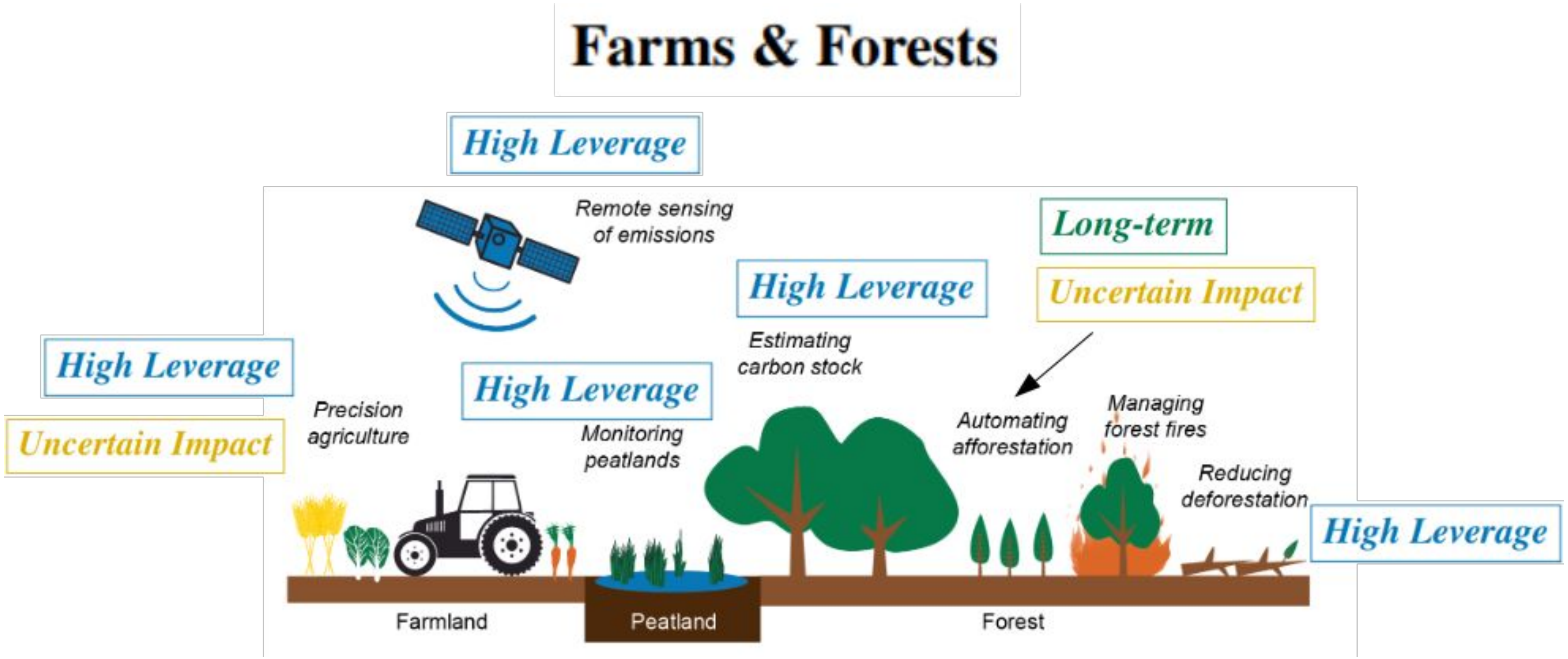


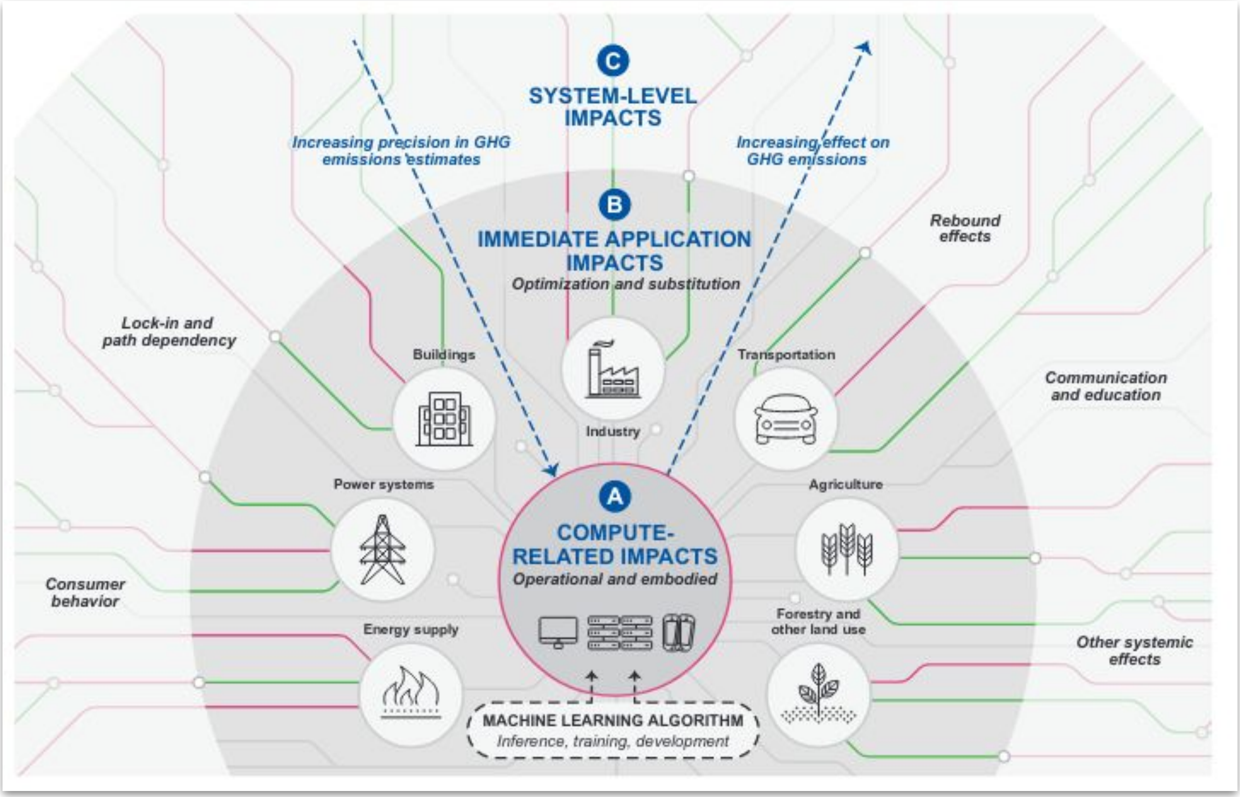
Figure 2: Selected strategies to mitigate GHG emissions from transportation using machine learning.

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



Environmental impacts of AI

First, second and third order impacts of AI



(Kaack et al., 2021)

First-order impacts

Bottom-up approach

on a server, what is the additional energy use due to the AI program running:

- processor
- GPU
- memory…

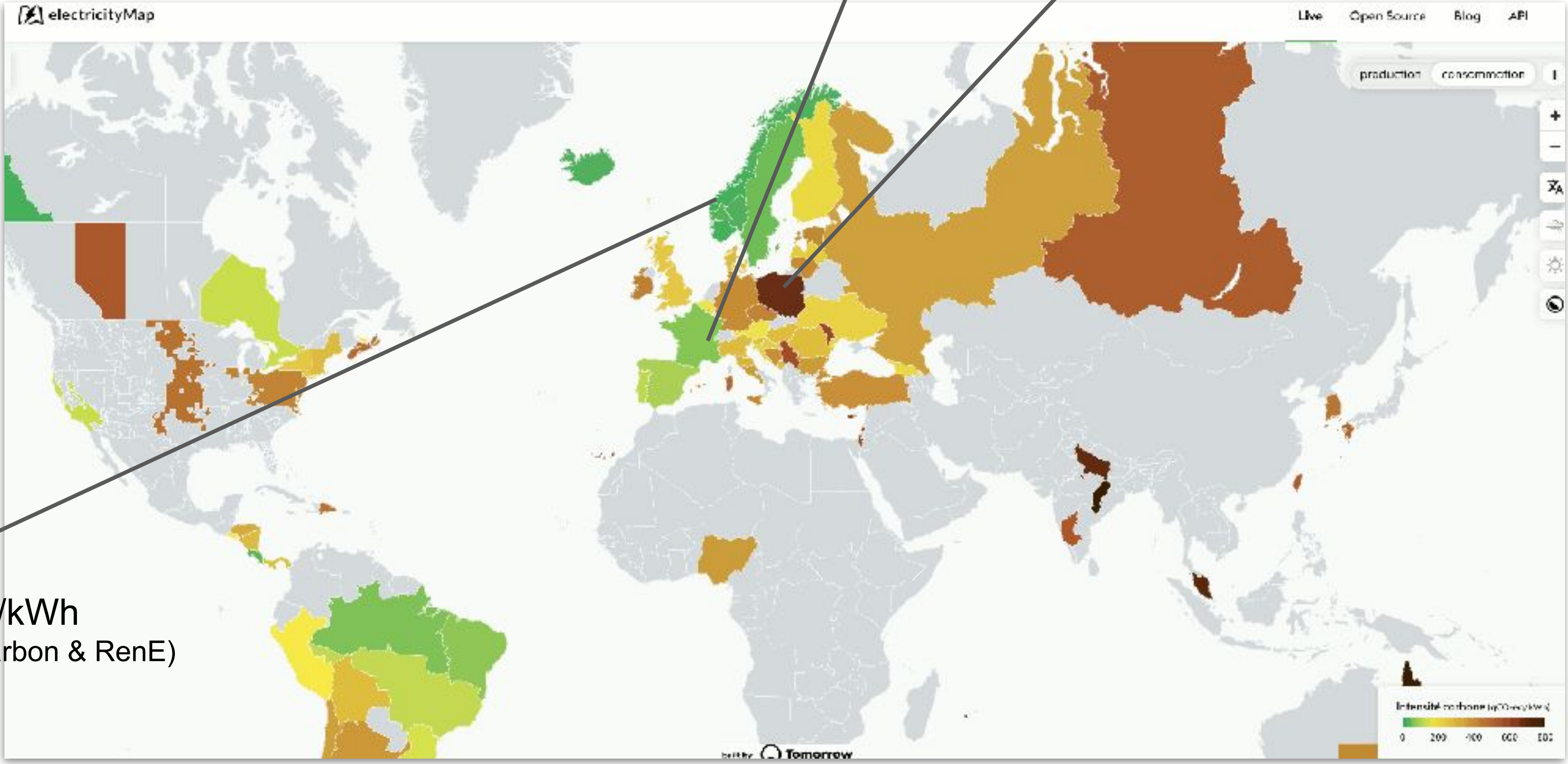
$$\Rightarrow \text{footprint}_1 = \sum(\text{use}_{\text{resource}}) \times \text{electricity carbon intensity}$$

kWh → kg CO_2e

Carbon intensity of electricity

France: 101g CO2e/kWh
(86% low carbon, 13% RenE)

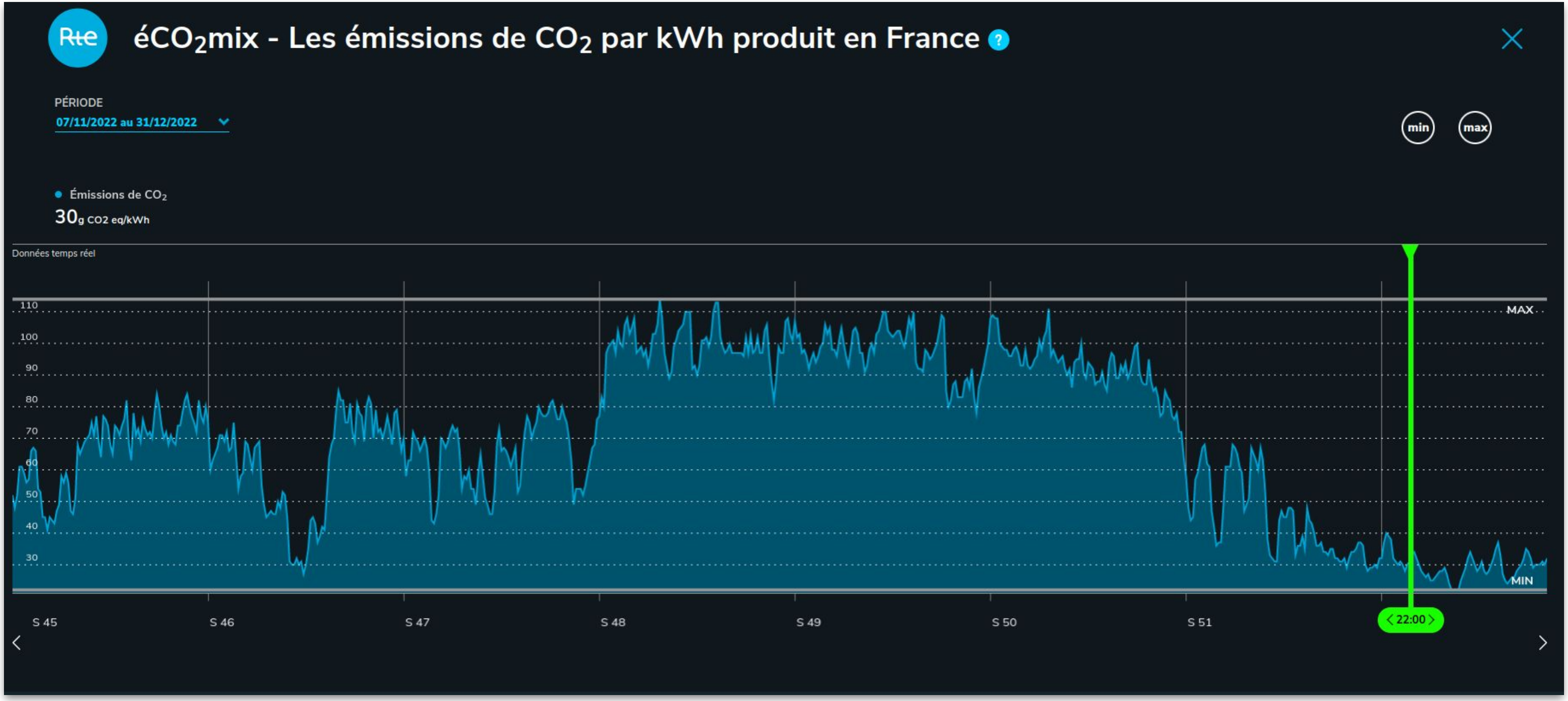
Poland: 927g CO2e/kWh
(13% low carbon, 13% RenE)



Norway:
22g CO2e/kWh
(100% low carbon & RenE)

source: electricityMap

Temporal evolution of the carbon intensity



Influence of the carbon intensity on the operational carbon footprint

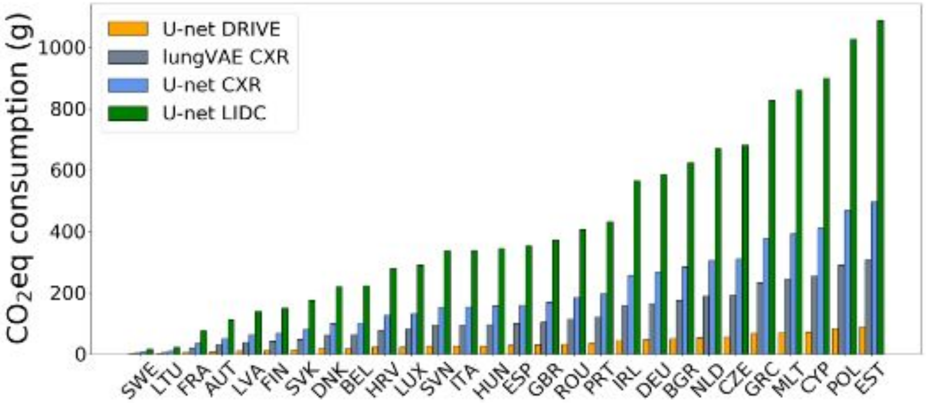


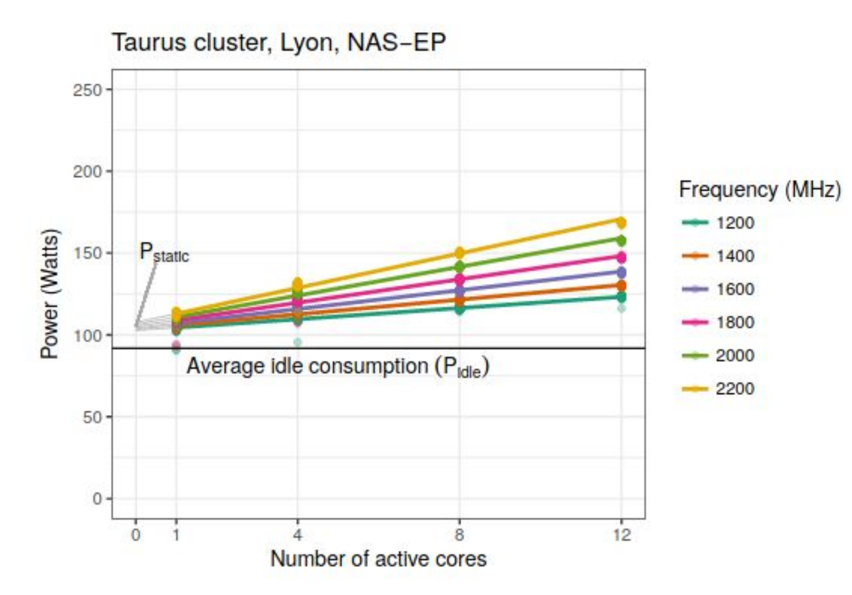
Figure 4. Estimated carbon emissions (gCO_2eq) of training our models (see Appendix B) in different EU-28 countries. The calculations are based on the average carbon intensities from 2016 (see Figure 8 in Appendix).

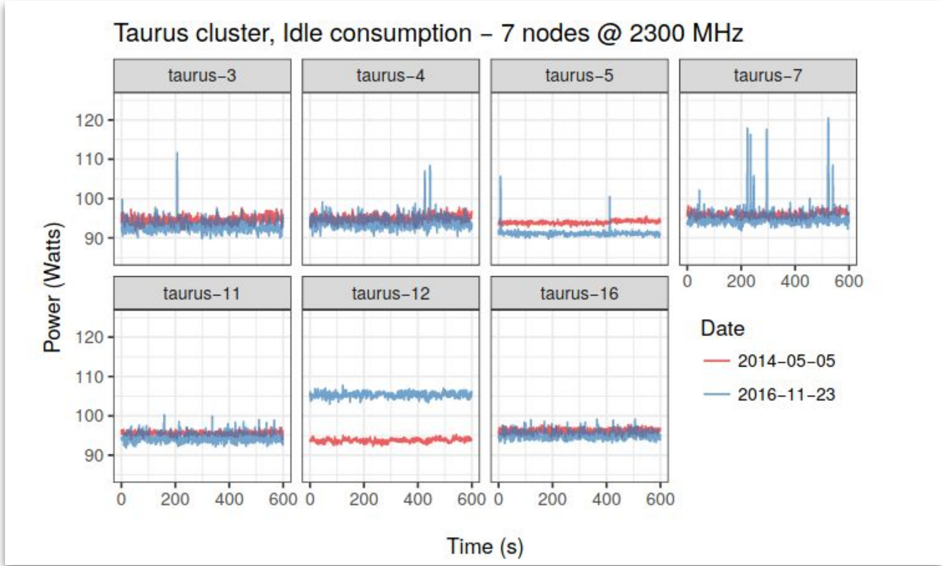
(Anthony et al., 2020)

Serveur energy use

not proportional to the charge

variation in time, with models…





Source: (Heinrich et al., 2017)

How to measure energy use?

hardware

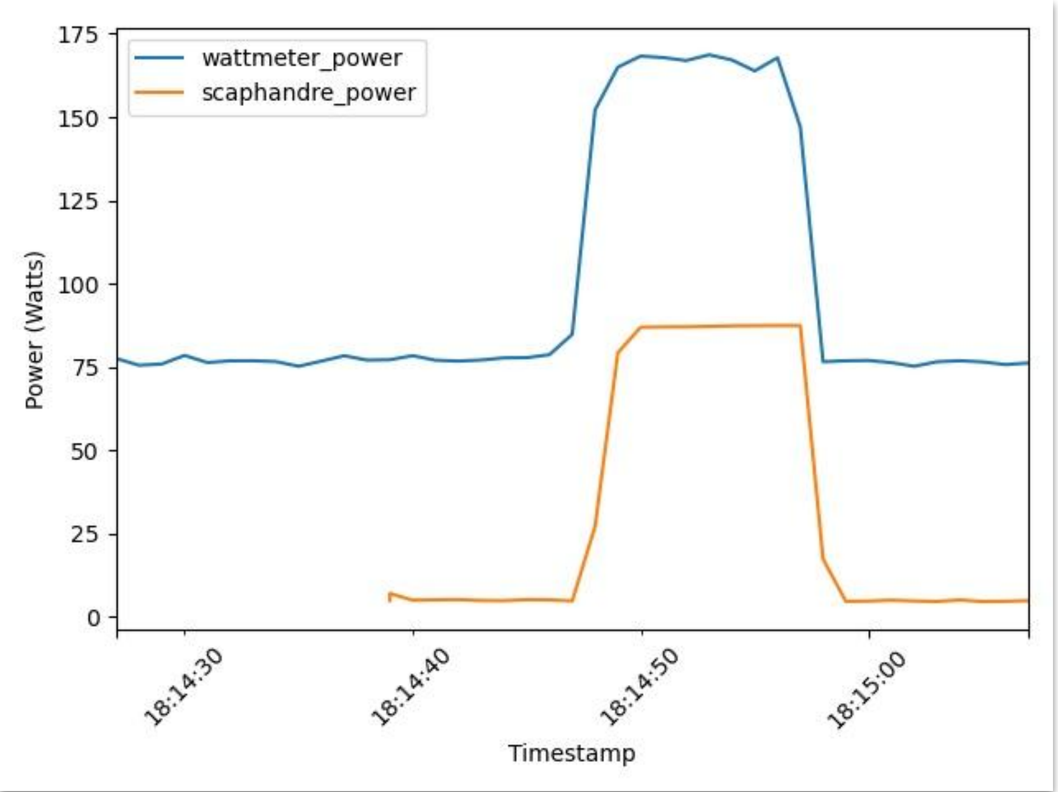
software

CODE CARBON

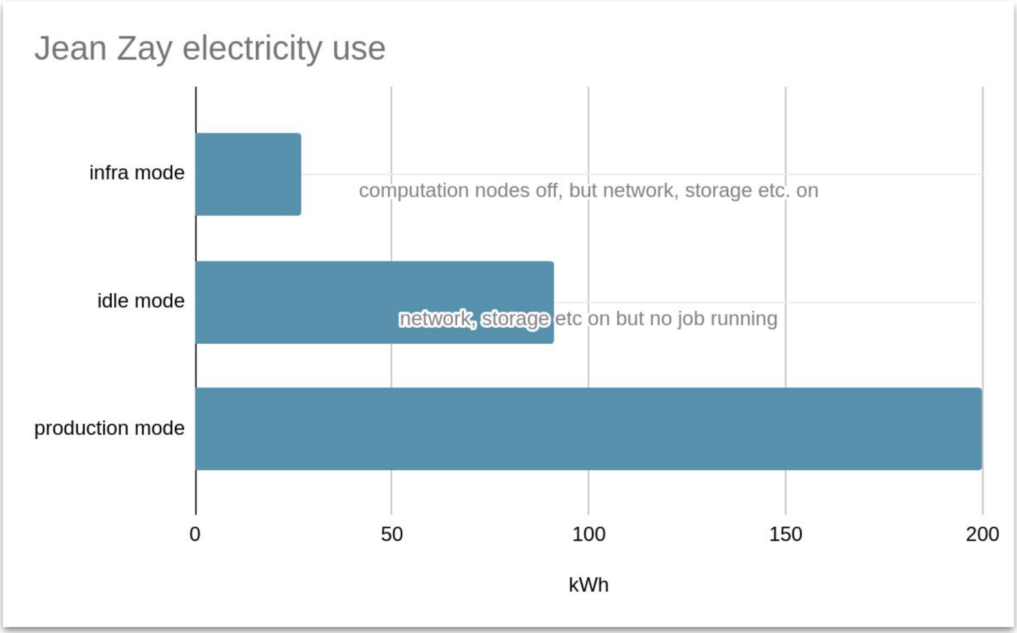
Green Algorithms

Towards environmentally sustainable computational science

Hardware vs software

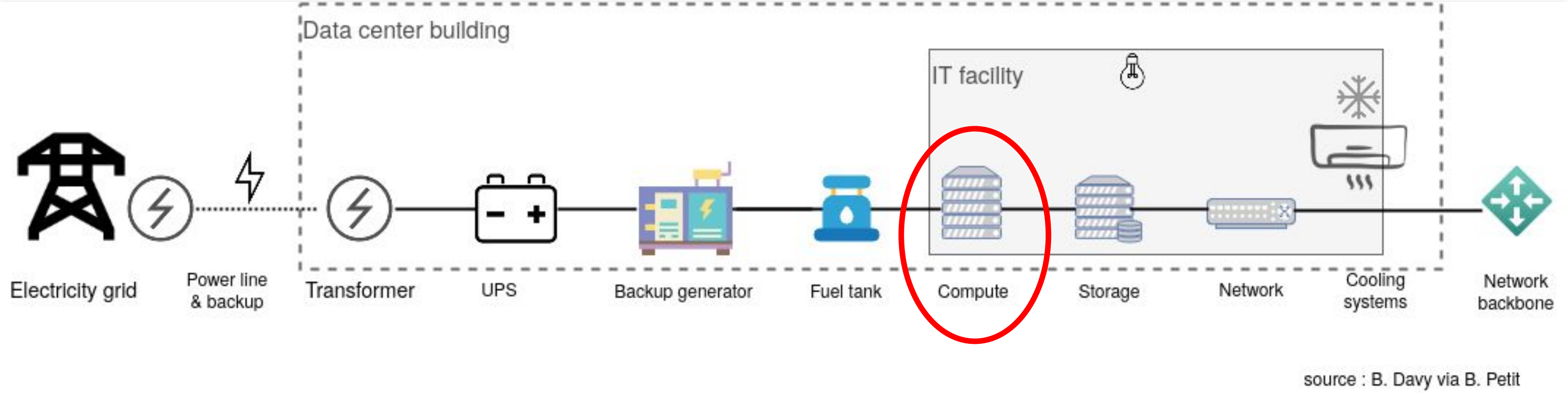


Electricity consumption in Jean Zay

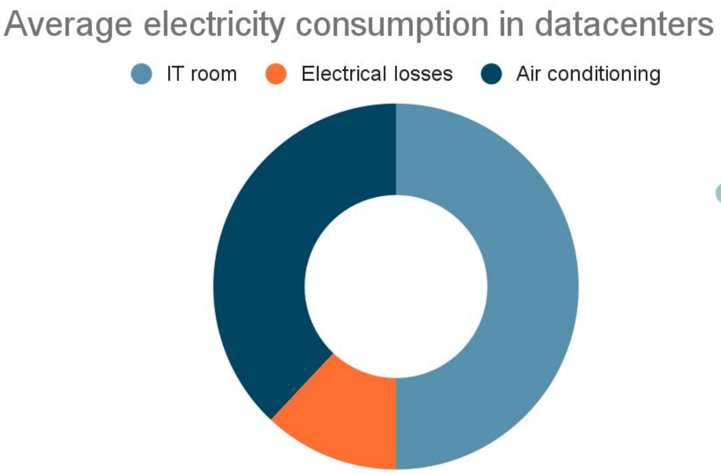


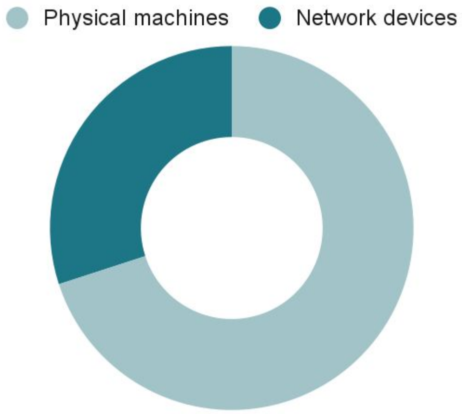
Evaluating the carbon footprint of an AI service

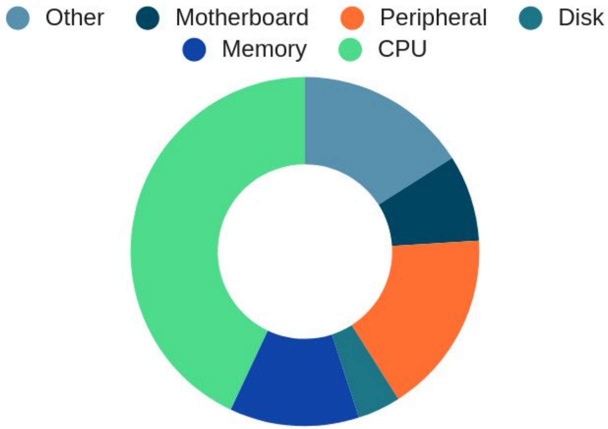
Which equipment?



Other energy use



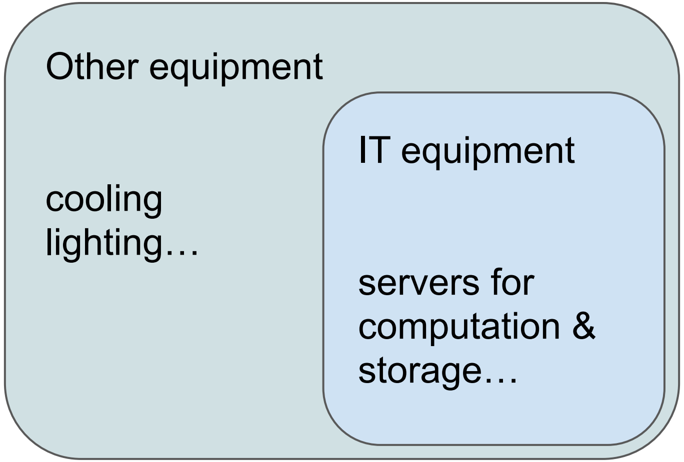




Source: (Guyon, 2018)

Efficiency of the facility

$$\text{PUE} = \frac{\text{total facility energy}}{\text{IT equipment energy}}$$

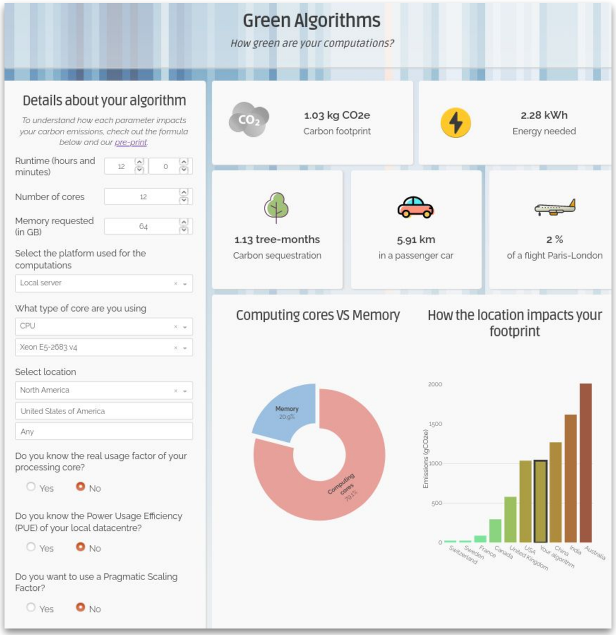


$\Rightarrow \text{footprint}_2 = \text{footprint}_1 \times \text{PUE}$

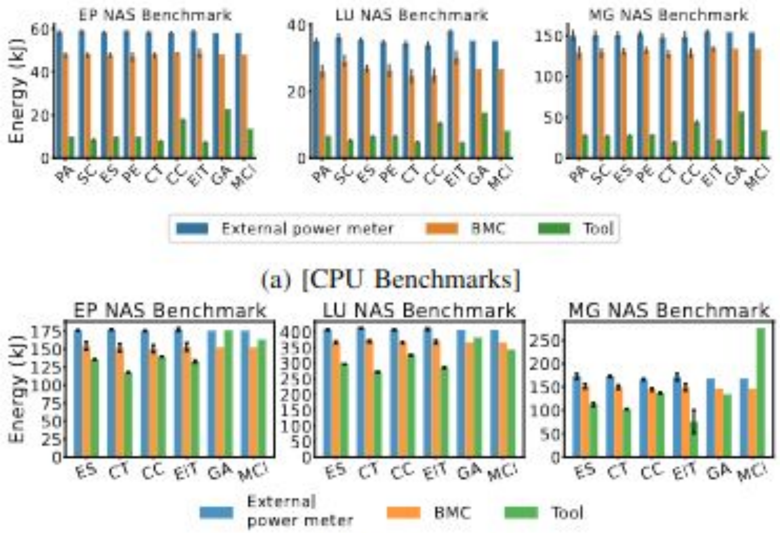
Tools for carbon footprint estimation

Many factors influence the carbon footprint of this phase

- model, data…
- energy efficiency of the data center
- carbon intensity of the electricity



Comparison of several tools



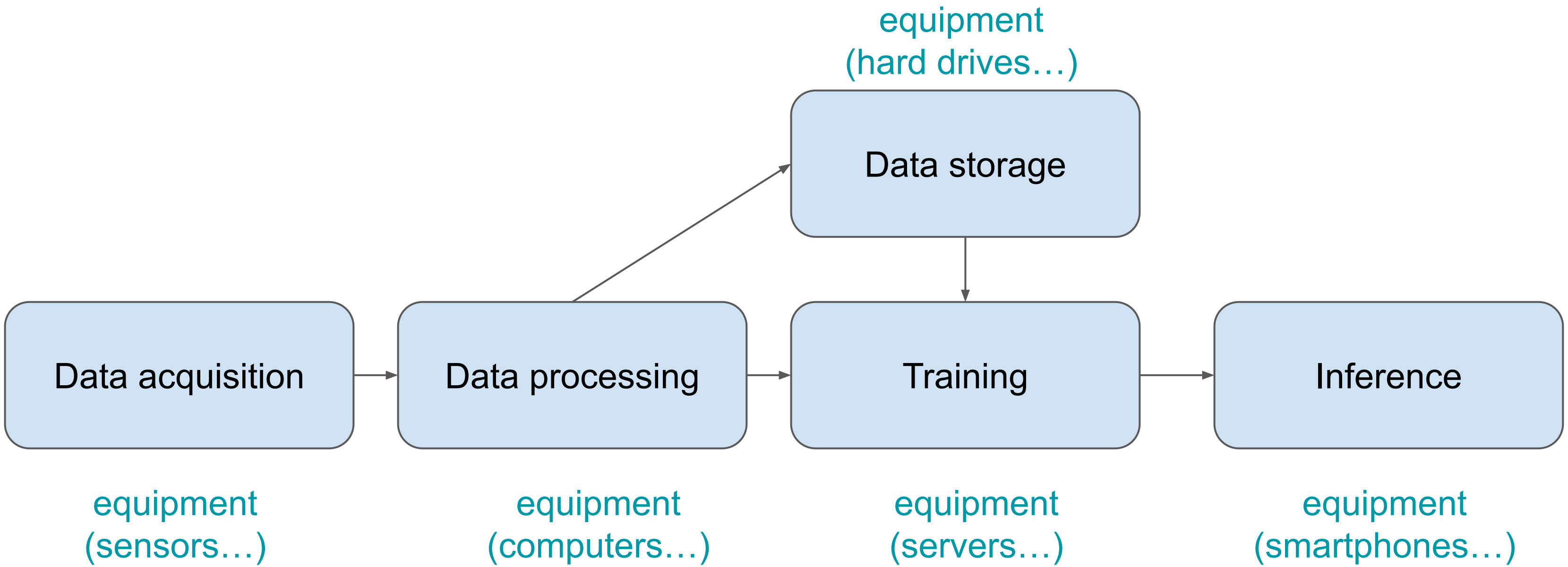
(a) [CPU Benchmarks]

(b) [GPU Benchmarks]

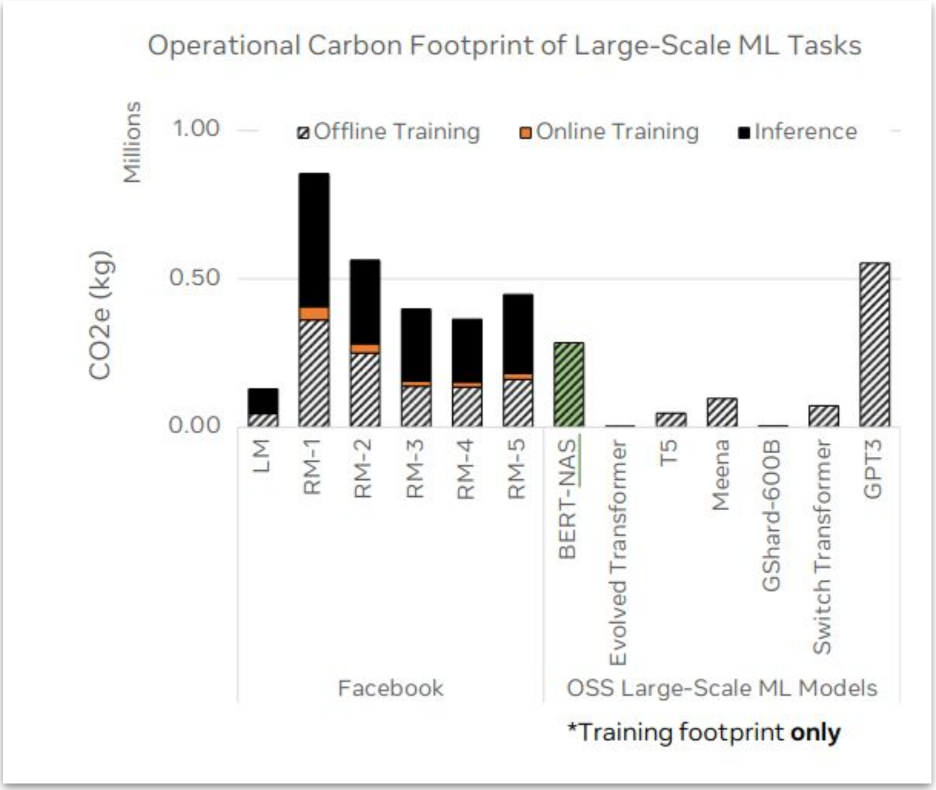
Fig. 2: Total energy consumed by the benchmarks as reported by the power meters. Tools: PowerAPI (PA), Scaphandre (SC), Energy Scope (ES), Perf (PE), Carbon Tracker (CT), Code Carbon (CC), Experiment Impact Tracker (EIT), Green Algorithm (GA), ML CO2 Impact (MCI)

source: (Jay et al., 2023)

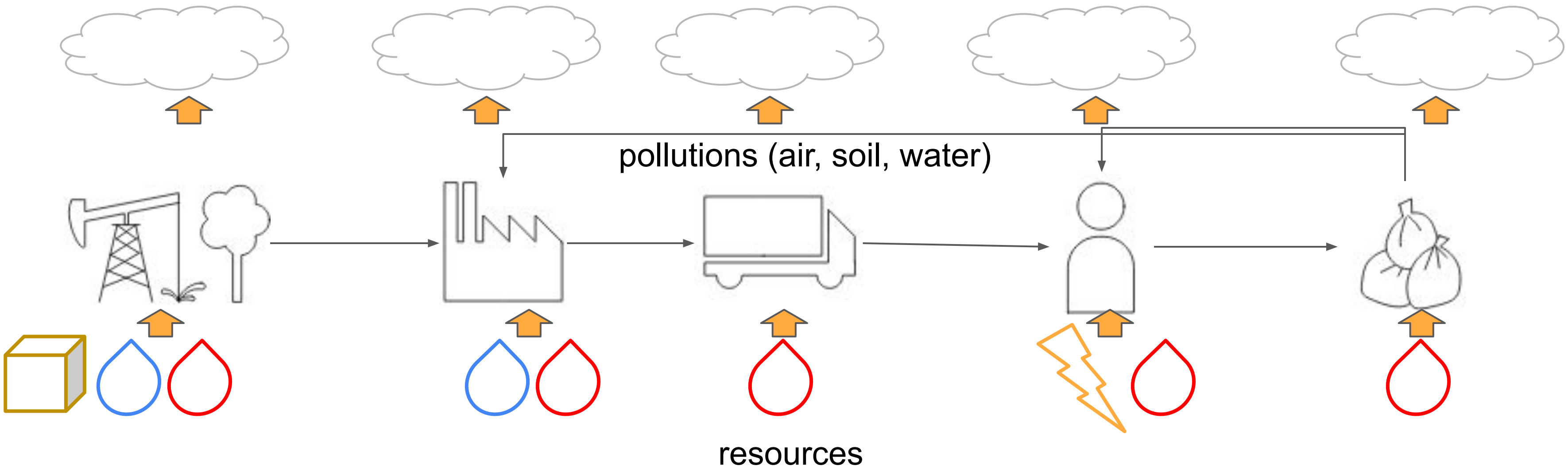
AI: which tasks?



Training vs inference (Wu et al., 2021)

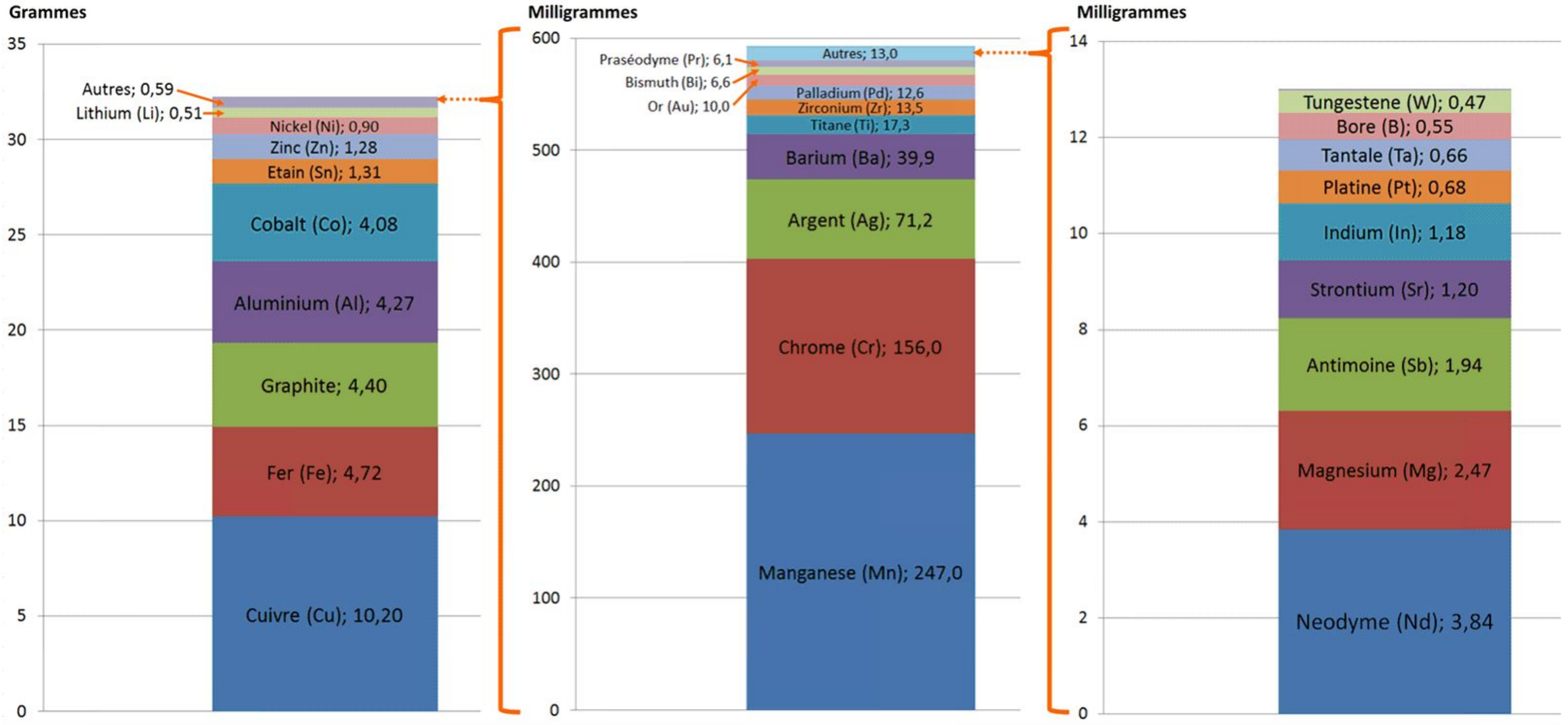


Life Cycle Assessment



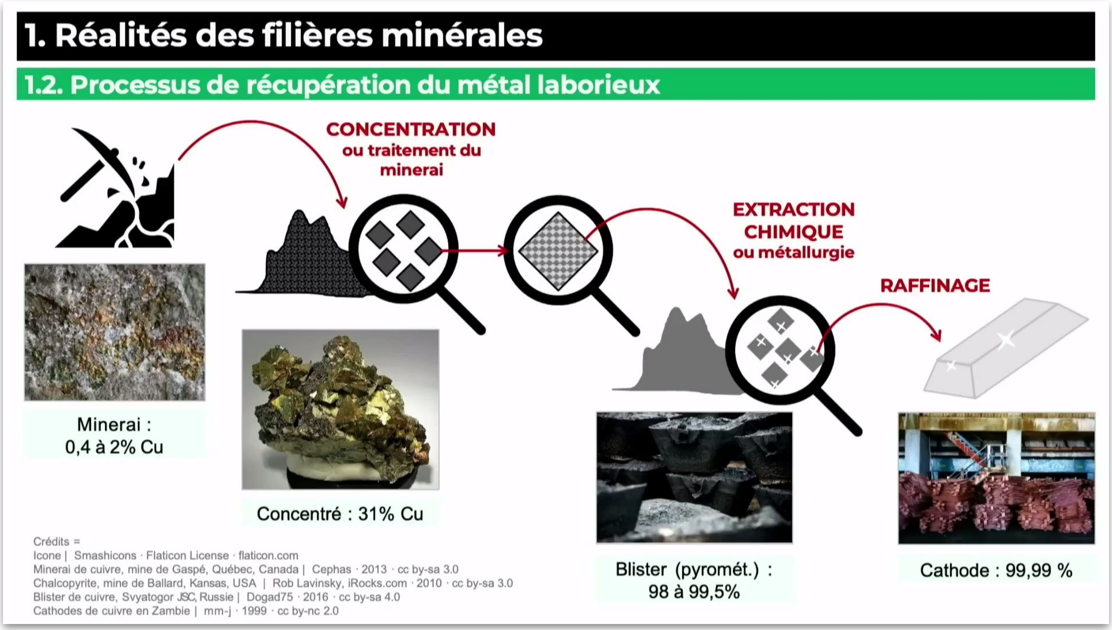
schema from Jacques Combaz

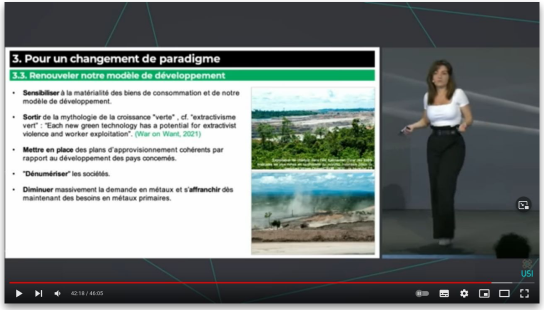
Composition of a smartphone



Source: [report from French Sénat on smartphones]

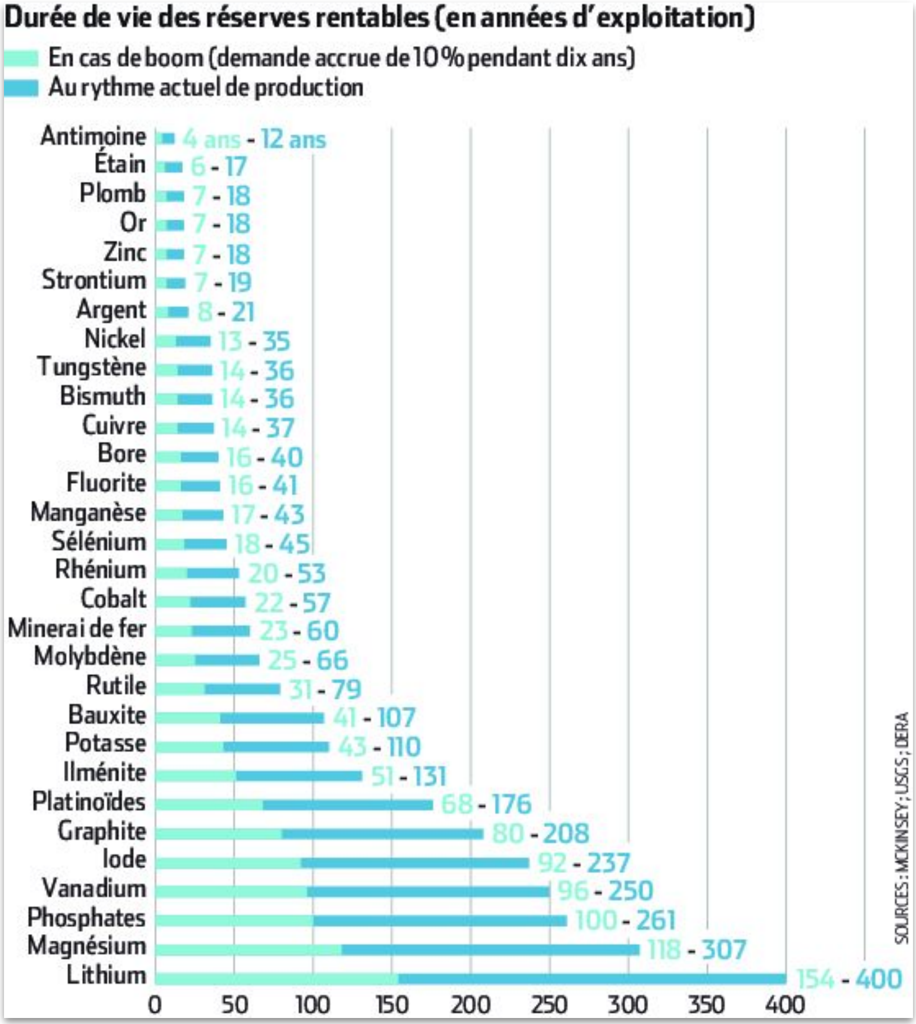
Metal recovery





[Ruée minière au XXIè siècle : jusqu'où les limites seront-elles repoussées ?]() - Aurore Stephant at USI

Raw material availability



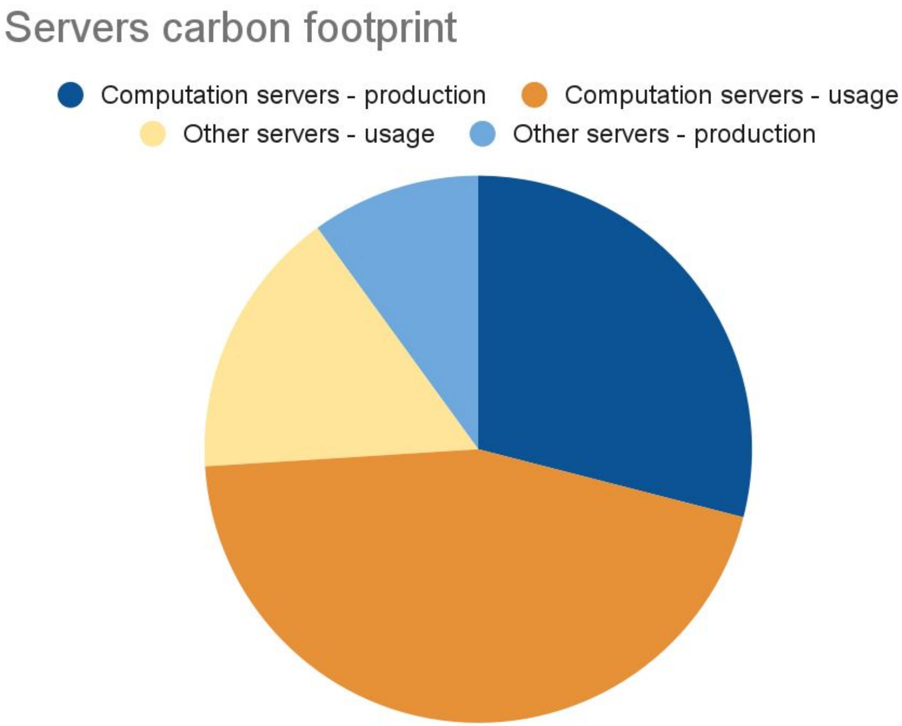
E-waste

Source : Global E-wasteMonitor 2020

Informal recycling

Dumping and processing of electronic waste in Agbogbloshie, Accra, Ghana

Top down approach at GRICAD



Source: (Berthoud et al., 2020)

Life cycle assessment of AI systems

(Luccioni et al, 2023)

a BigScience initiative

BLOOM

176B params · 59 languages · Open-access

Process	CO_2emissions (CO_2eq)	Percentage of total emissions
Embodied emissions	11.2 tonnes	22.2 %
Dynamic consumption	24.69 tonnes	48.9 %
Idle consumption	14.6 tonnes	28.9 %
Total	**50.5 tonnes**	**100.00%**

Table 3: Breakdown of CO_2 emissions from different sources of the BLOOM model life cycle

- Methodology for estimating the carbon footprint of the Jean Zay infrastructure
- Estimation of the carbon footprint
  - for training the model, including idle consumption & embodied emissions
  - for inference

Integrating life cycle aspects in environmental evaluation

Outil	Life cycle phase considered						Multiple impacts considered	Estimates consumption	GPU support
	Ext.	Man.	Tra.	Uti.		EoL.			
				Infra.	Dyn.				
Green Algorithms	✗	✗	✗	✓	✓	✗	✗	✓	✓
ML CO_2 Impact	✗	✗	✗	✗	✓	✗	✗	✓	✓
CarbonTracker	✗	✗	✗	✓	✓	✗	✗	✗	✓
CodeCarbon	✗	✗	✗	✓	✓	✗	✗	✗	✓
Boavizta	✓	✓	✗	✗	✗	✗	✓	-	✗

source: (Morand, 2023)

Integrating life cycle aspects in environmental evaluation

	ADP	GWP	PE	Human toxicity	Water Consumption	...
Extraction	✓	✓	✓	✗	✗	✗
Manufacturing	✓	✓	✓	✗	✗	✗
Transport	✗	✗	✗	✗	✗	✗
Usage	✓	✓	✓	✗	✗	✗
End of Life	✗	✗	✗	✗	✗	✗

Modeling graphics card

Manufacturing impacts attribution

Infrastructure consumption

PARIS 2015

Putting impacts in perspective

source: (Morand, 2023)

Environmental impacts

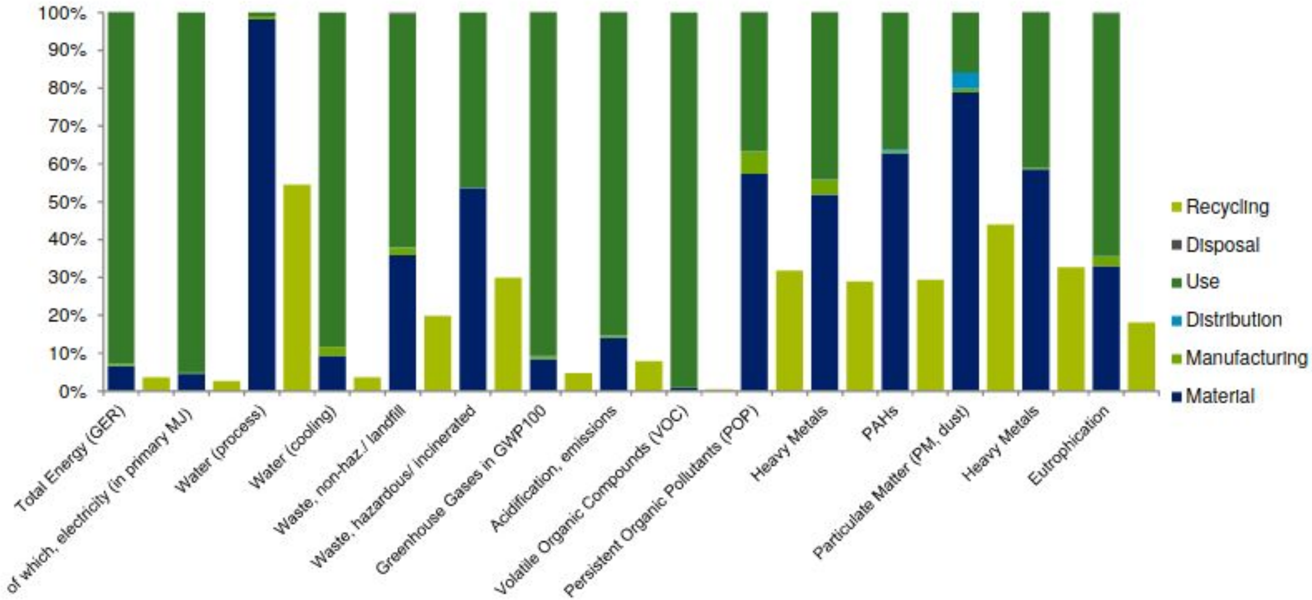


Figure 3: Distribution of BC-1 environmental impacts by life cycle phase[22]

Source: [European commission](), 2015

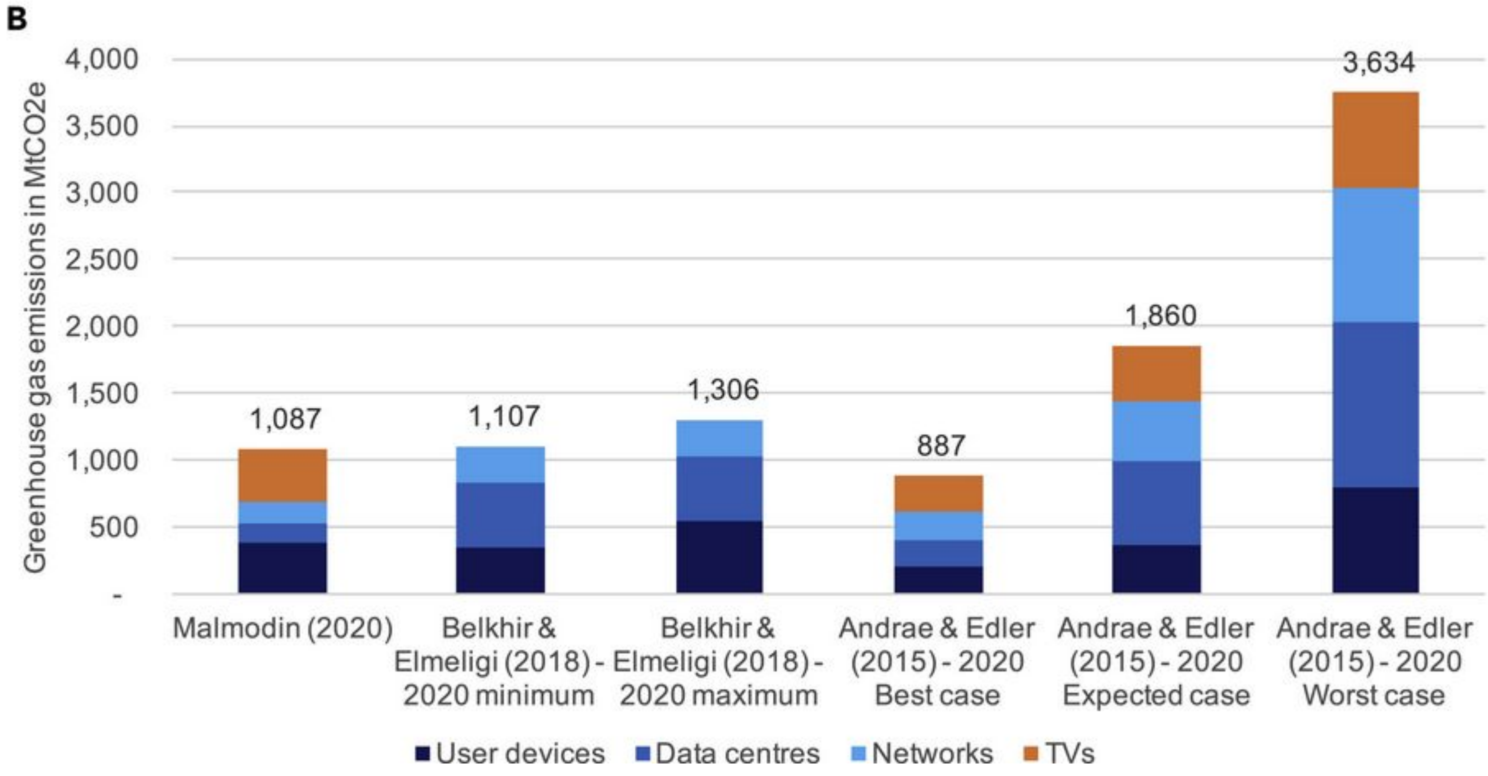
Results for BLOOM training

http://calculator.green-algorithms.org/

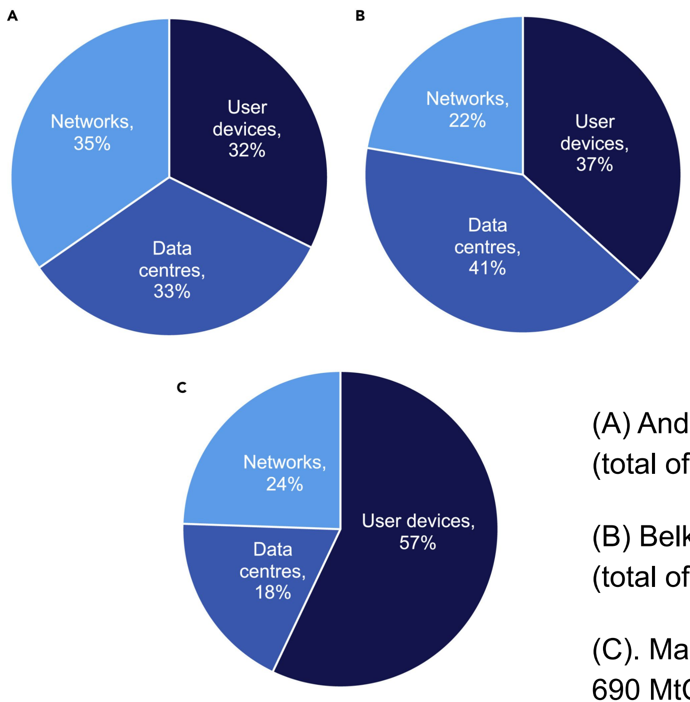
training BLOOM

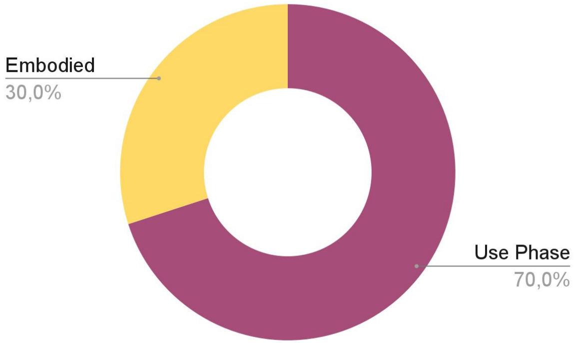
- GWP: 59tCO_2 eq
  - annual emissions of 59 person (PB_{GWP})
  - annual emissions of 29 person (SNBC)
- ADP: 1.2 kgSb eq
  - annual resource extraction of 38 person (PB_{ADP})
- PE: 9800000 MJ

Carbon footprint of ICT in 2020 (Freitag et al, 2021)



Proportional breakdown of ICT's carbon footprint, excluding TV (Freitag et al, 2021)



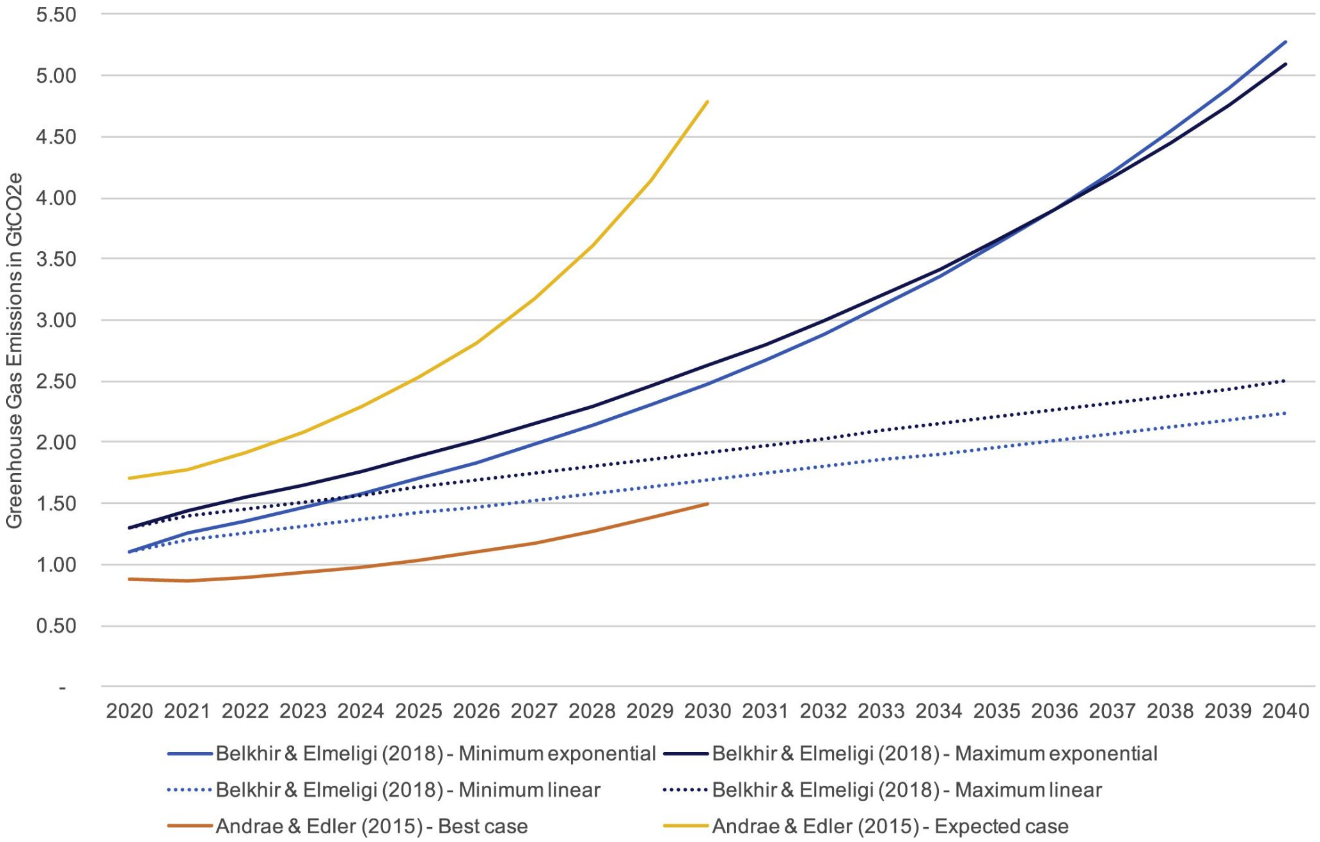


(A) Andrae and Edler (2015): 2020 best case (total of 623 $MtCO_2e$).

(B) Belkhir and Elmeligi (2018): 2020 average (total of 1,207 $MtCO_2e$).

(C). Malmodin (2020): 2020 estimate (total of 690 $MtCO_2e$).

Projections of ICT's GHG emissions from 2020 (Freitag et al, 2021)



Second and third-order impacts

Indirect impacts

direct impacts

indirect impacts

optimize traffic flow?

use of new connected objects, sensors…

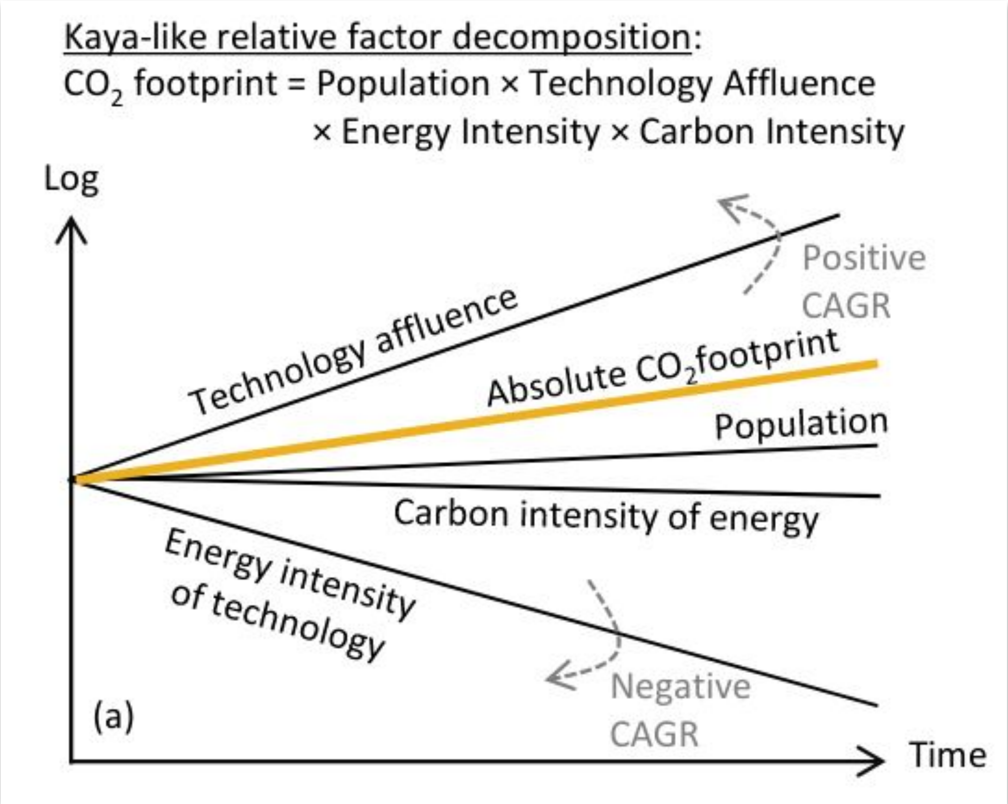
lower fuel consumption

rebound effect
smoother traffic flow => time savings => greater distance from home => urban sprawl

path dependency
prolongs current system, vs. public transport, active mobility...

priority to systems with significant impacts?

Carbon footprint of the ICT sector(s)



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.

Structural effects

Our societies are dependent on digital technology

How do we adapt to climate change and resource depletion?

Case of storm Alex in the Alpes-Maritimes

Numerous communes in the valleys without water or electricity, without road or rail links, and without telephone communications (mobile, copper and fiber-optic sites having been affected).

source: Orange

Infrastructure resilience

Table 2 – Network infrastructure risk qualification test

CLIMATE RELATED HAZARDS		NETWOR K: Electricity transmission	Electricity distribution	Rail transport	Road transport	Fixed telecommunications	Mobile telecommunications
Trends	Increase in average temperature	green	green	orange	orange	orange	orange
Extremes	Heat waves, fires and drought	orange	orange	orange	orange	red	red
	Flooding, submersion, high water and landslides	orange	orange	red	red	orange	orange
	High winds and storms	green	orange	red	green	orange	orange

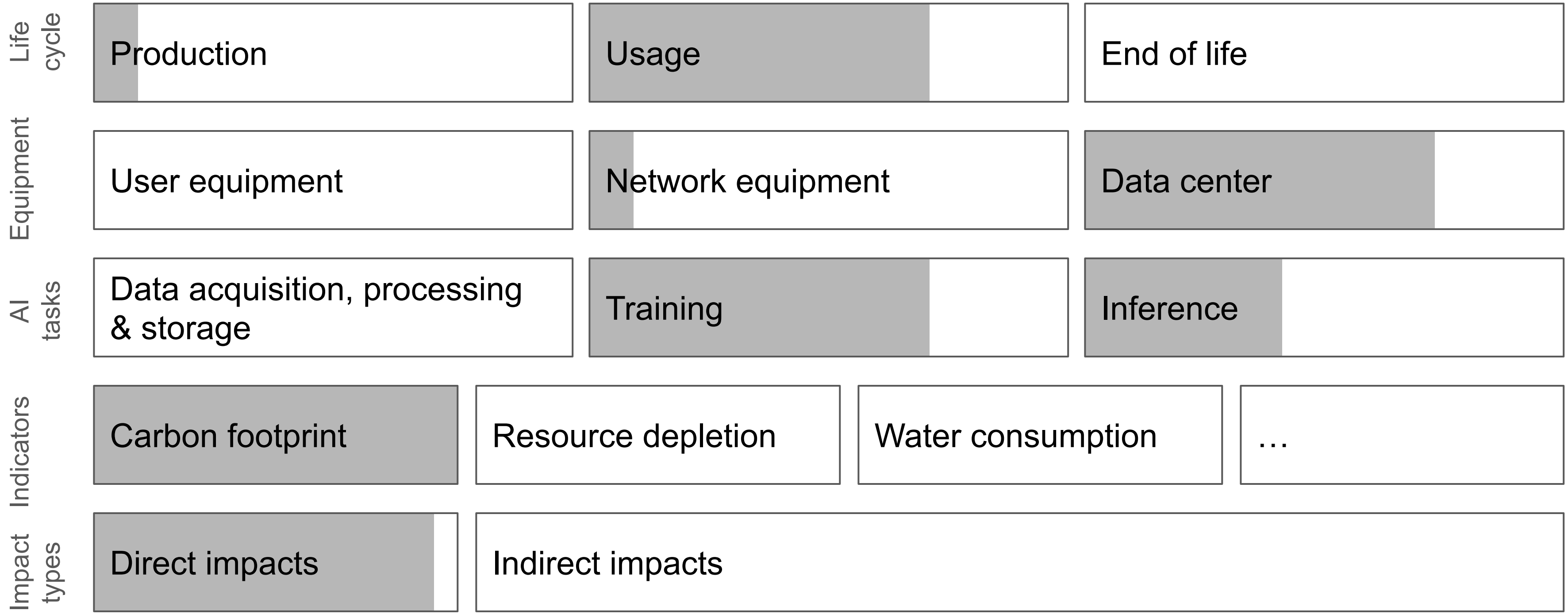
Note: the qualitative assessment is based on interviews conducted for the study (including RTE, Enedis, SNCF Réseau, Cerema and Vinci Autoroutes). The colour represents the intensity of the physical risk (green when vulnerability is limited, red when it is high).

Summary: The physical risk to transport infrastructure from high winds and storms is considered to be limited, and the increase in average temperature has been anticipated for electricity infrastructure (green boxes). Flooding poses risks of structural deformation or even failure of transport network infrastructures (red boxes). Heat waves pose significant risks to the operation of air-conditioning systems for strategic active equipment in telecommunications networks (boxes in red).

Source: France Stratégie

In ML/NLP?

What is presently assessed



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

Red AI

- improve accuracy rather than efficiency, through the use of massive computational power while disregarding the cost
  - even though relationship between model performance and model complexity is at best logarithmic
- yet valuable: contributes to what we know about pushing the boundaries of AI

but

⇒ allow for more equitable comparisons, eg reporting training curves

⇒ recognize Green AI work

Green AI

novel results encouraging a reduction in resources spent

Responsible AI?

- < > Déclaration de Montréal IA responsable_ </ > (Dilhac et al., 2018)
    - AI systems and associated equipment must aim for maximum energy efficiency and minimize the carbon footprint over their entire lifecycle, as well as impacts on ecosystems and biodiversity...
- Villani report (2018)
    - (...) AI can lead to numerous rebound effects. For example AI can prevent us from rethinking our modes of growth, consumption, and measurement of wealth produced, and instead to consume just as much as before, if not more.

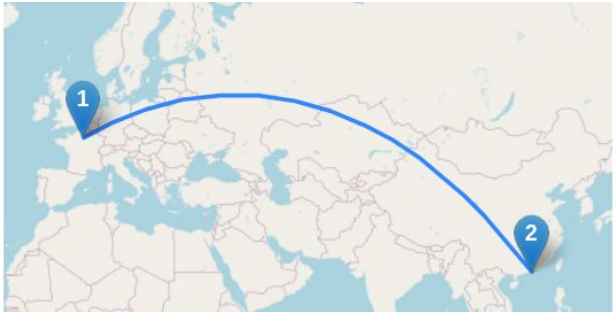
Environmental impacts of AI? (Strubell et al, 2019)

variety of state-of-the-art NLP models

software-based energy measurement

Training

- 12 hours to several weeks
- emissions: between 18kg CO_2e and 284 t CO_2e
- most used model: 652 kg CO_2e, or
  - one one-way flight from Paris to Hong Kong
  - or 2 500km by car



sum GPU time ~ 60 GPU during 6 months

Precision vs CO2e (Parcollet et Ravanelli, 2021)

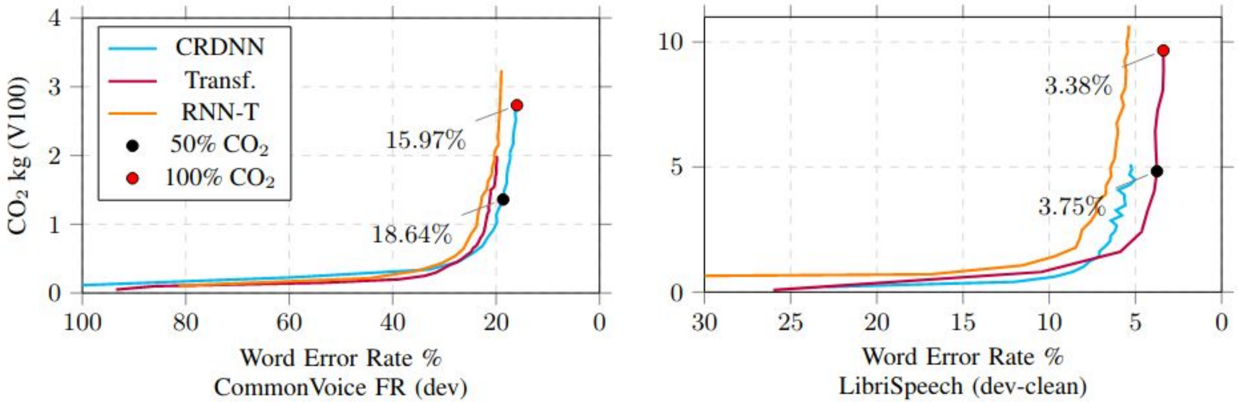


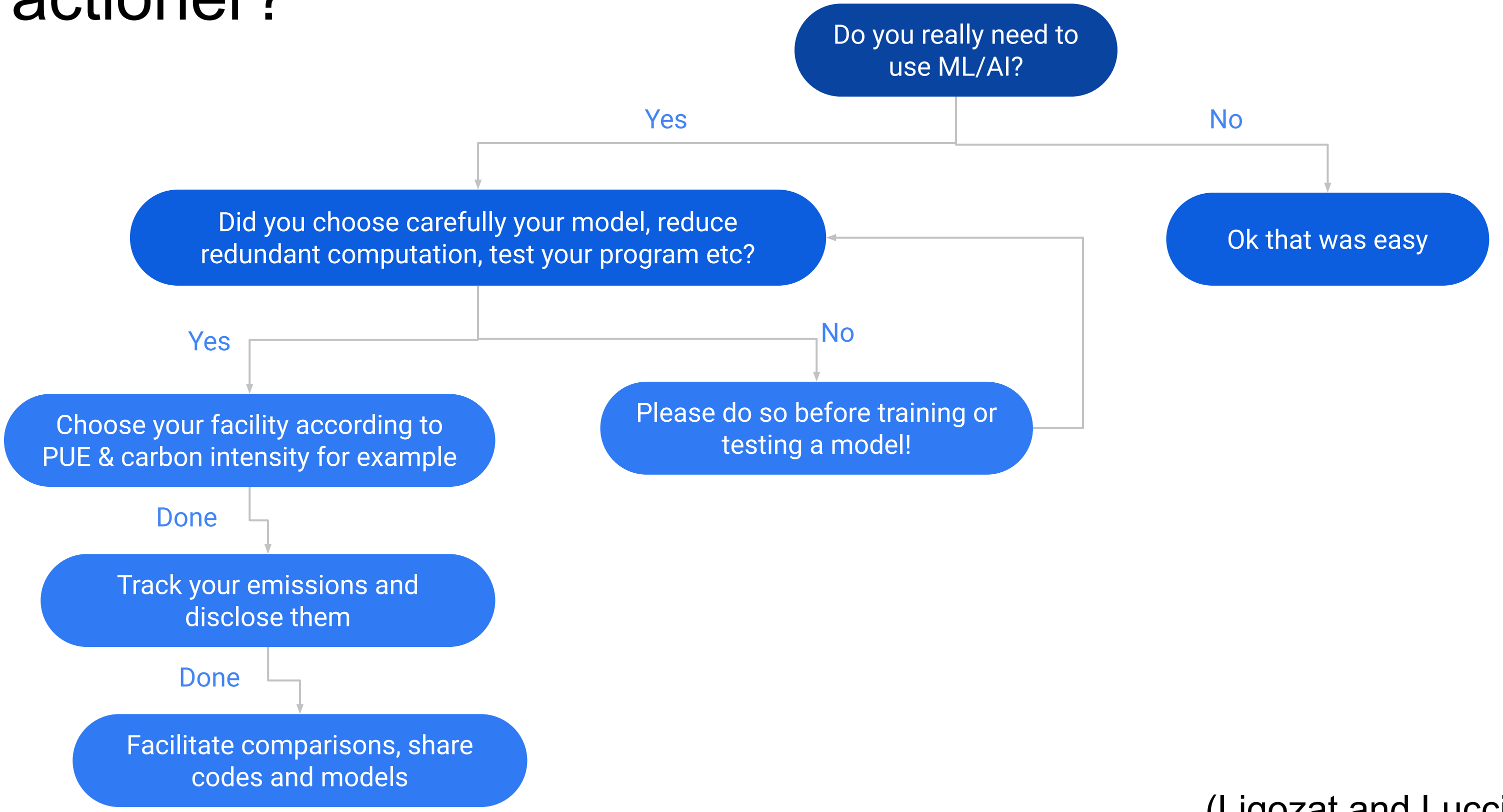
Figure 2: *CO_2 emitted in kg (in France) by different E2E ASR models with respect to the word error rate (WER) on the dev sets of LibriSpeech and CommonVoice. The curves exhibit an exponential trend as most of the training time is devoted to slightly reduce the WER. The black and red dots indicates the WER obtained with 50% and 100% of the emitted CO_2. On LibriSpeech, 50% of the carbon emissions have been dedicated to reach SOTA results with an improvement of 0.37%.*

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?	gCO2eq/ kWh
7. How much CO2eq was emitted to train the final model?	kg
8. How much CO2eq was emitted for all experiments?	kg
9. What is the average CO2eq emission for the inference of one sample?	kg
10. Which positive environmental impact can be expected from this work?	Notes
11. Comments	Notes

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



(Ligozat and Luccioni, 2021)

Google’s answer to (Strubell et al., 2019)

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

David Patterson[1,2], Joseph Gonzalez[2], Urs Hölzle[1], Quoc Le[1], Chen Liang[1], Lluis-Miquel Munguia[1], Daniel Rothchild[2], David So[1], Maud Texier[1], and Jeff Dean[1]

Best practices proposed:

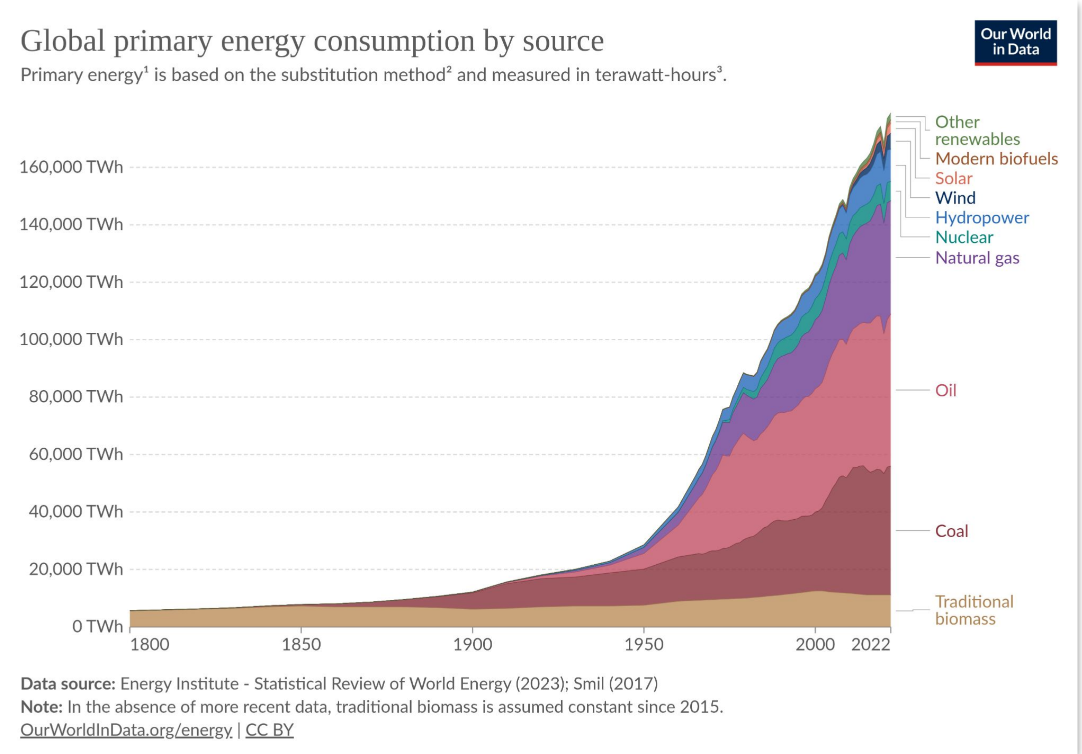
- Efficient ML model
- Processors optimized for ML training
- Cloud pour better energy efficiency
- Location with the “cleanest” energy

and «Google's renewable energy purchases further reduce the impact to zero»

but:

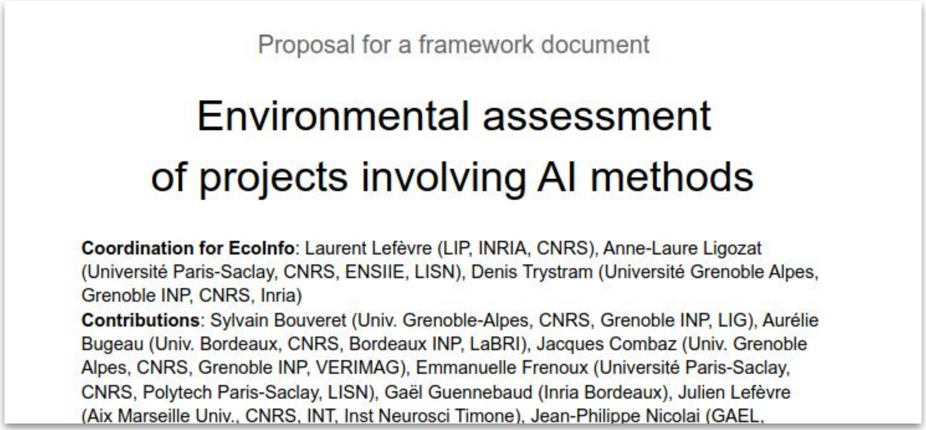
- what about the life cycle?
    - recent processors ⇒ carbon footprint ↗
- what about inference?
- «carbon free» energy and «net zero impact»?
- potential carbon footprint if everything optimized, but not actual one
- focus on carbon footprint

Decarbonization of energy?



Environmental assessment of projects involving AI methods

- Impacts of ICT equipment
  - material extraction, manufacturing, end of life
  - use: computation, data
- Justification of the AI method
  - nécessity of AI
  - resilience
- Impacts due to societal changes
  - reference scenario
  - potential indirect impacts



Proposal for a framework document

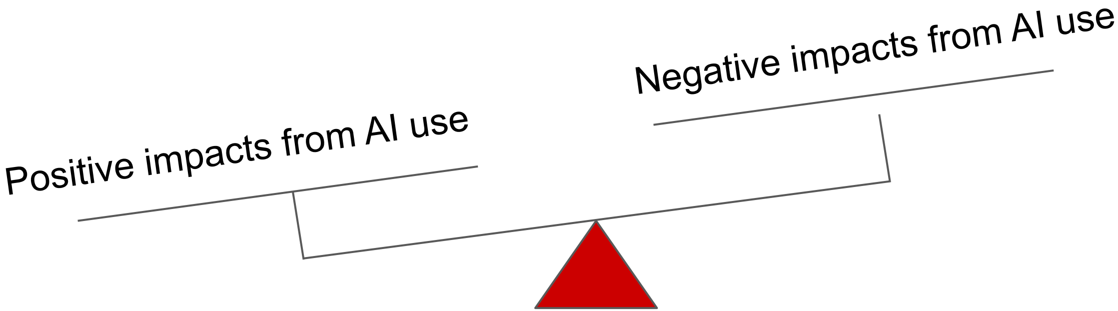
Environmental assessment
of projects involving AI methods

Coordination for EcoInfo: Laurent Lefèvre (LIP, INRIA, CNRS), Anne-Laure Ligozat (Université Paris-Saclay, CNRS, ENSIIE, LISN), Denis Trystram (Université Grenoble Alpes, Grenoble INP, CNRS, Inria)
Contributions: Sylvain Bouveret (Univ. Grenoble-Alpes, CNRS, Grenoble INP, LIG), Aurélie Bugeau (Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI), Jacques Combaz (Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG), Emmanuelle Frenoux (Université Paris-Saclay, CNRS, Polytech Paris-Saclay, LISN), Gaël Guennebaud (Inria Bordeaux), Julien Lefèvre (Aix Marseille Univ., CNRS, INT, Inst Neurosci Timone), Jean-Philippe Nicolai (GAEL,

https://hal.science/hal-03922093

Back to AI to tackle climate change

AI for environmental applications



at least with Life Cycle Assessment

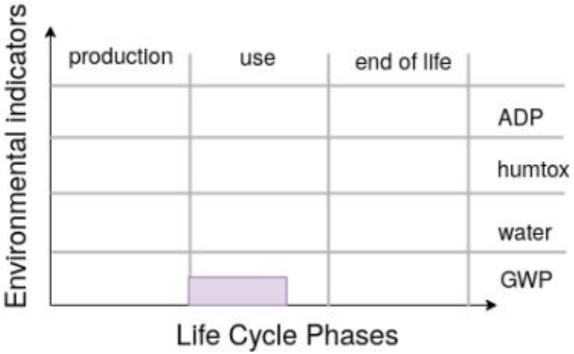
taking into account as many indirect effects as possible

Life cycle assessment of AI systems

(Ligozat et al, 2021)

Assessing the environmental impacts of an AI system should at least include a Life Cycle Assessment

How are AI for Green systems benefits assessed?

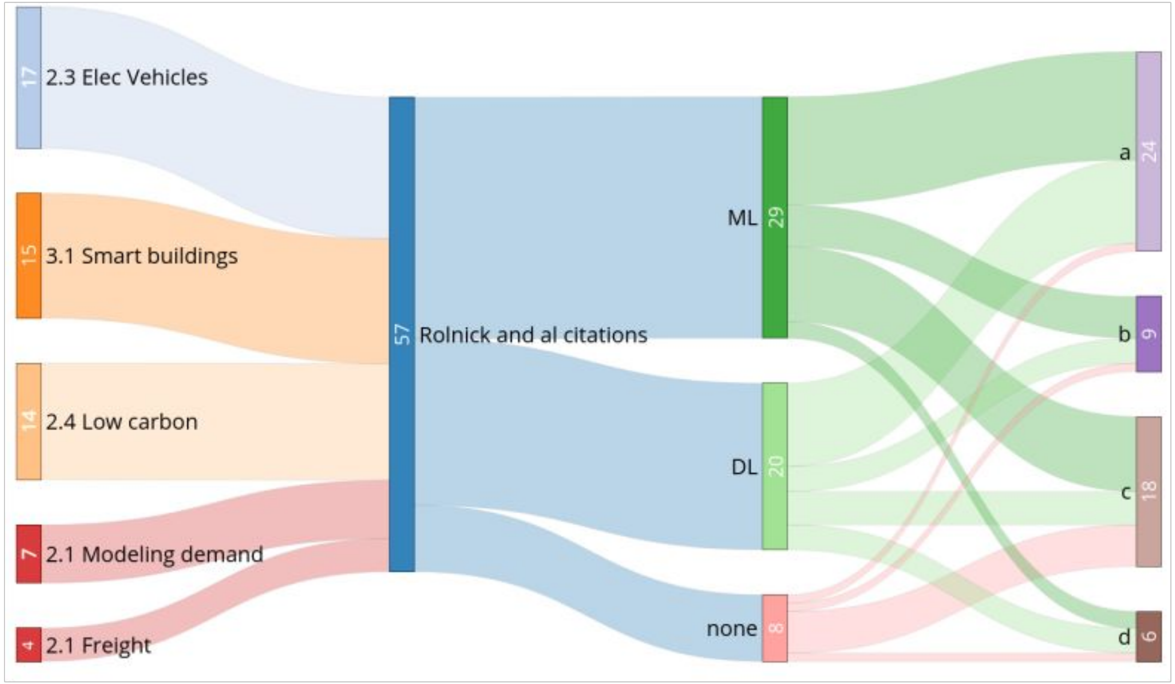


$$\Delta(M_2|M_1) = LCA(M_2) - LCA(M_1) \in \mathbb{R}^d \quad (1)$$

with:

- M_1 the reference application without using the AI service,
- M_2 the application enhanced by AI,
- $LCA(x)$ a quantification of d types of environmental impacts (e.g., GHG emissions, water footprint, etc.). The LCA methodology is described in Section 3.2. Note that $LCA(M_2)$ includes the impacts of the AI service itself, i.e., $LCA_{AI}(M_2)$.

Evaluations in (Rolnick et al., 2019)



a. No mention of the environmental gain

b. General mention of the environmental gain

c. A few words about the environmental gain but no quantitative evaluation or only indirect estimation

d. Evaluation of the energy gain without taking the AI program into account

Biases of impact studies (Rasoldier et al., 2022)

Perimeter

- life cycle not taken into account: (Ligozat et al., 2021) for AI
- indirect (2nd and 3rd order) not taken into account: 5G

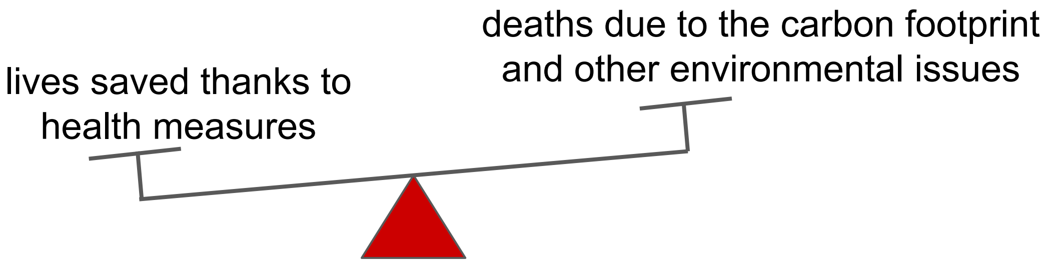
Hypotheses

- comparison to what reference scenario?

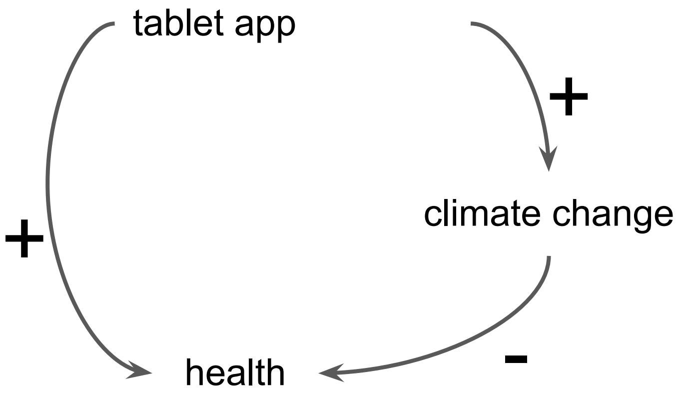
Disconnection from global scenarios

- minimal benefits + poorly managed uncertainties
- incompatibility between measures

Example in the health sector



[Conférence Comprendre et Agir -- Valérie d'Acremont]



Conclusion

- Comprehensive evaluation of the environmental impacts remains a WIP
- But tools for partial evaluation of 1st order impacts exist and can easily be used
- As well as guidelines for a discussion of 2nd and 3rd order impacts

- But be careful with partial indicators
- Need for discussion of the role of AI in a green transition

Références (1/2)

- 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.; Bzeznik, 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/
- Bol, D., Pirson, T., Dekimpe, R. (). Moore's Law and ICT Innovation in the Anthropocene. IEEE Design, Automation and Test in Europe Conference 2021. http://hdl.handle.net/2078.1/243578
- Bugeau, A., Ligozat, A.-L. (2023). How digital will the future be? Analysis of prospective scenarios. https://hal.science/hal-04362220v2
- 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
- Freitag, C., Berners-Lee, M., Widdicks, K., Knowles, B., Blair, G. S., Friday, A. (2022). The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. Patterns, Volume 3, Issue 8, 12 August 2022, Pages 100576. https://doi.org/10.1016/j.patter.2021.100340
- 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).
- Hershcovich, 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
- Mathilde Jay, Vladimir Ostapenco, Laurent Lefèvre, Denis Trystram, Anne-Cécile Orgerie, et al.. An experimental comparison of software-based power meters: focus on CPU and GPU. CCGrid 2023 - 23rd IEEE/ACM international symposium on cluster, cloud and internet computing, May 2023, Bangalore, India. pp.1-13. https://cnrs.hal.science/hal-04030223v2
- 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

Références (2/2)

- 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
- Luccioni, A. S., Viguier, S., Ligozat, A.-L. Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. Journal of Machine Learning Research, 2023, 24 (253). https://www.jmlr.org/papers/volume24/23-0069/23-0069.pdf
- Parcollet, T., Ravanelli, M.. The Energy and Carbon Footprint of Training End-to-End Speech Recognizers. 2021. https://hal.science/hal-03190119
- Rasoldier, A., Combaz, J., Girault, A., Marquet, K., Quinton, S. (2022). How realistic are claims about the benefits of using digital technologies for GHG emissions mitigation?. LIMITS 2022 - Eighth Workshop on Computing within Limits, Jun 2022, Virtual, France. https://hal.science/hal-03949261/
- 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