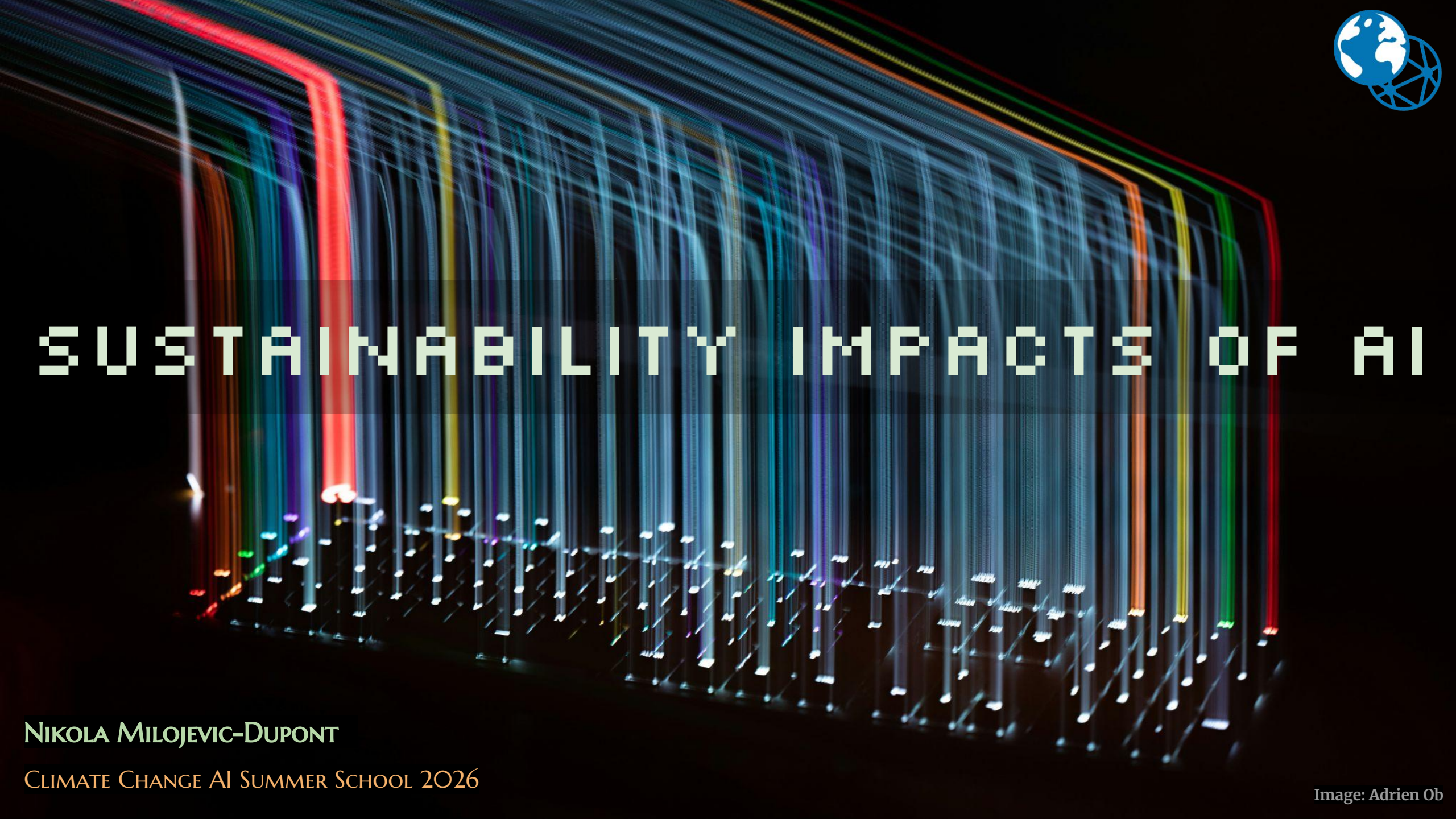


SUSTAINABILITY IMPACTS OF AI

Nikola Milojevic-Dupont

Climate Change AI Summer School 2026

Image: Adrien Ob

Acknowledgments

Thanks to Priya Donti and Shaolei Ren for inputs & feedback on this lecture, and Lynn Kaack for previous materials included!

Callback: Multifaceted relationship Between ML and Climate

Impacts from ML computation & hardware

ML applications in climate change mitigation

ML applications that increase emissions

ML's system-level impacts

Poll one

Which one of four impact categories do you believe may have the most substantial impact on climate currently?

- **Impacts from ML computation & hardware**
- **ML applications in climate change mitigation**
- **ML applications that increase emissions**
- **ML's system-level impacts**

Life-Cycle Assessment

Whole value chain from **cradle** (raw material extraction), through use phase, to **grave** (disposal)

Assessing the **many** environmental impacts

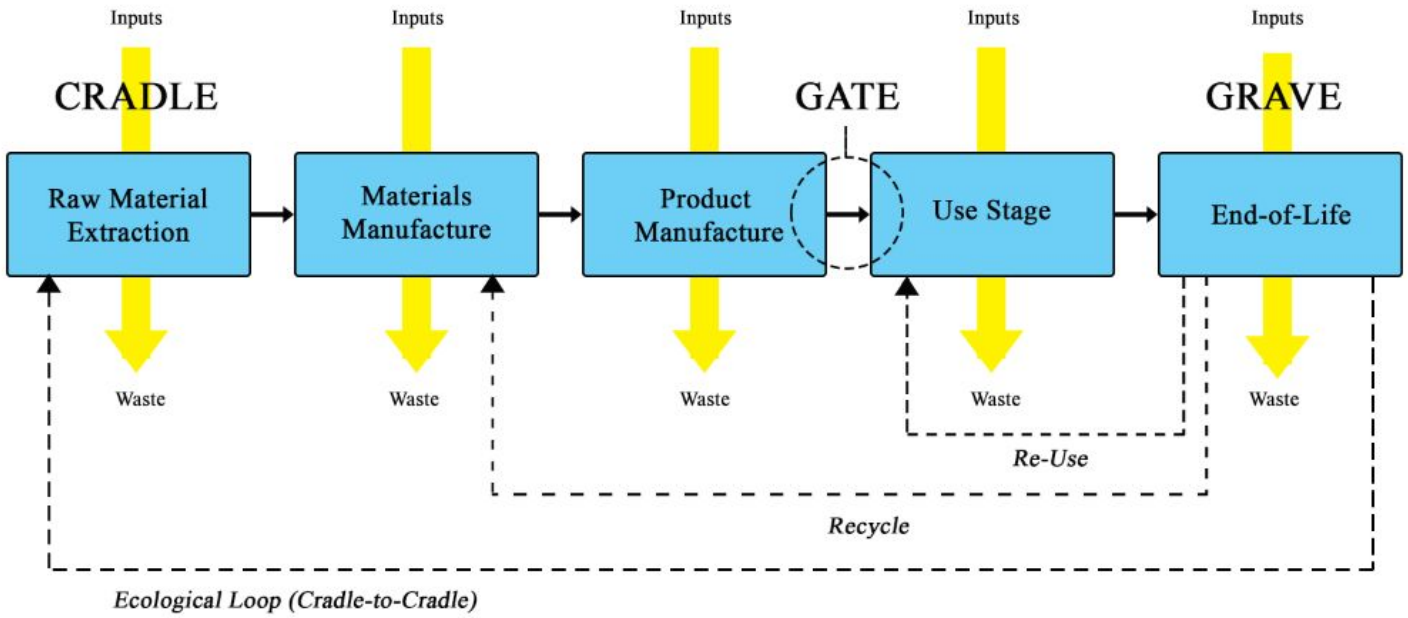


Image: Wikipedia

Inventory Analysis: GHG Protocol

- Global standardized frameworks to measure GHG emissions from private and public sector operations, value chains and mitigation actions
- **Scope 1:** Emissions from direct use of fossil fuels, e.g. for heat, fleet vehicles
- **Scope 2:** Emissions from purchased electricity and heat
- **Scope 3:** Emissions from value chains including customers
- Also extended to water

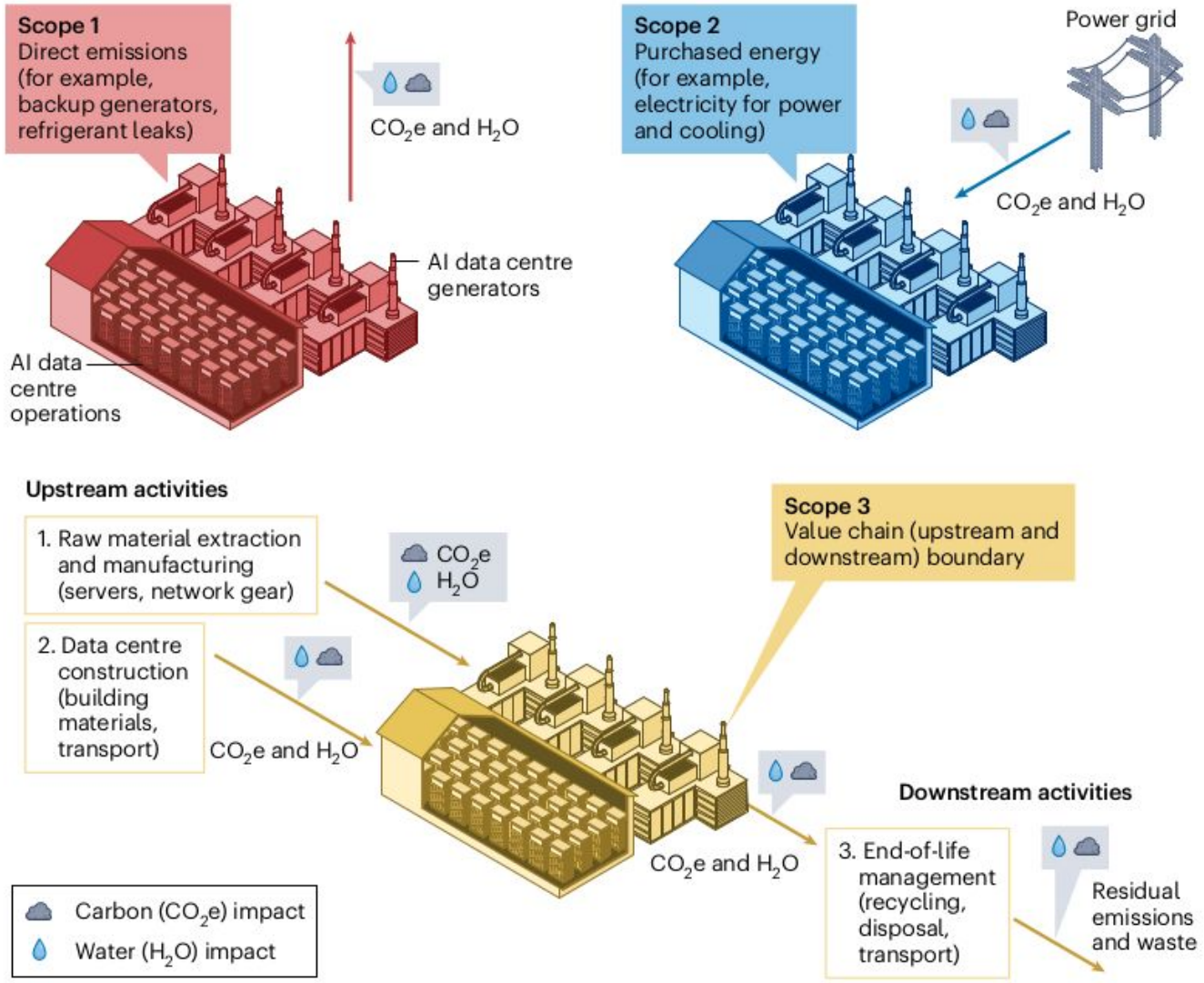
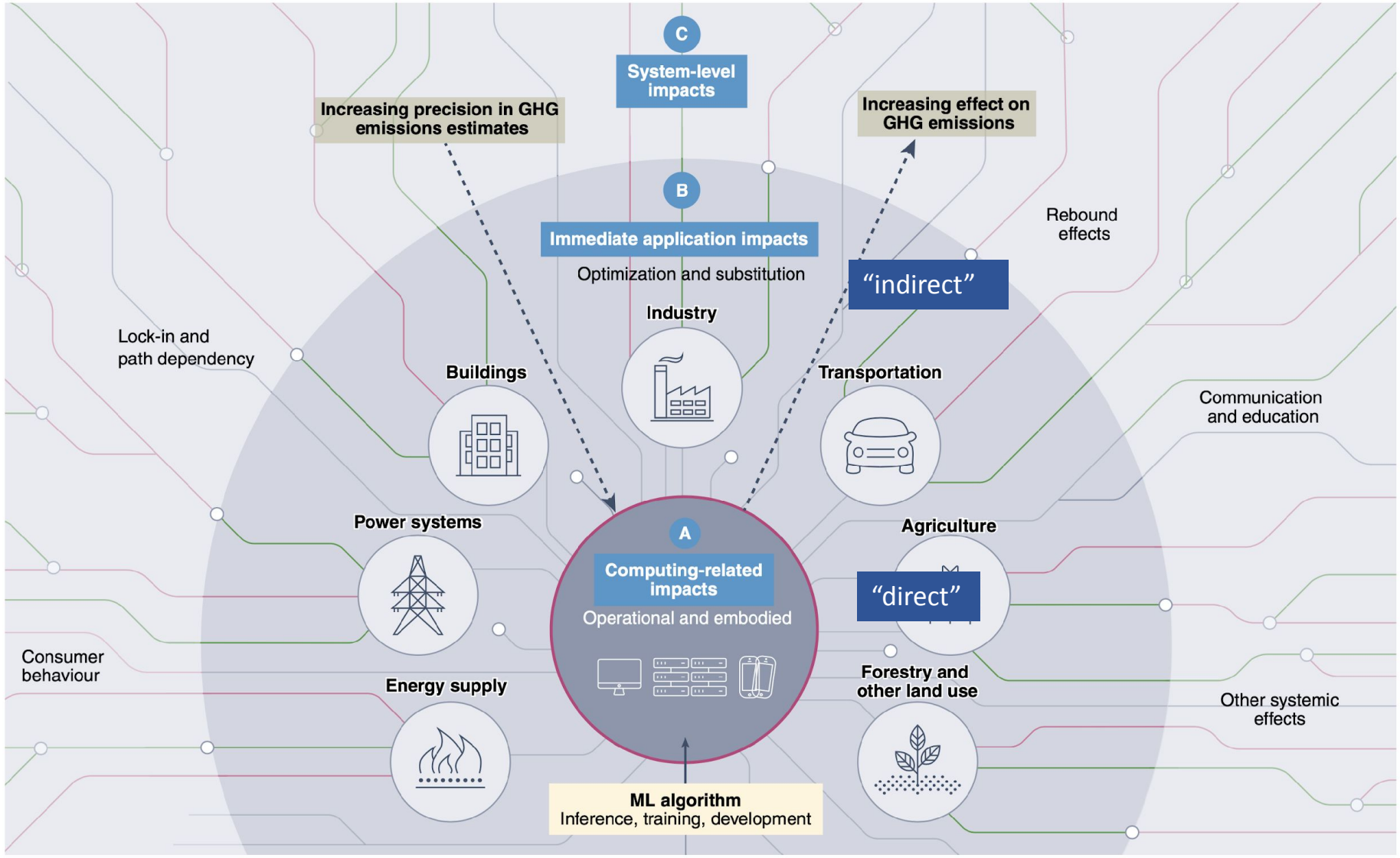


Image: Chien et al, 2026 in Nature Reviews Clean Technology

How AI Affects Greenhouse Gas Emissions



Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Nature Climate Change, 1-10.

Lecture outline

Computing-related impacts

Application impacts

System level impacts

Overarching considerations

Regulating energy and resource usage of AI

Lecture outline

Computing-related impacts

Application impacts

System level impacts

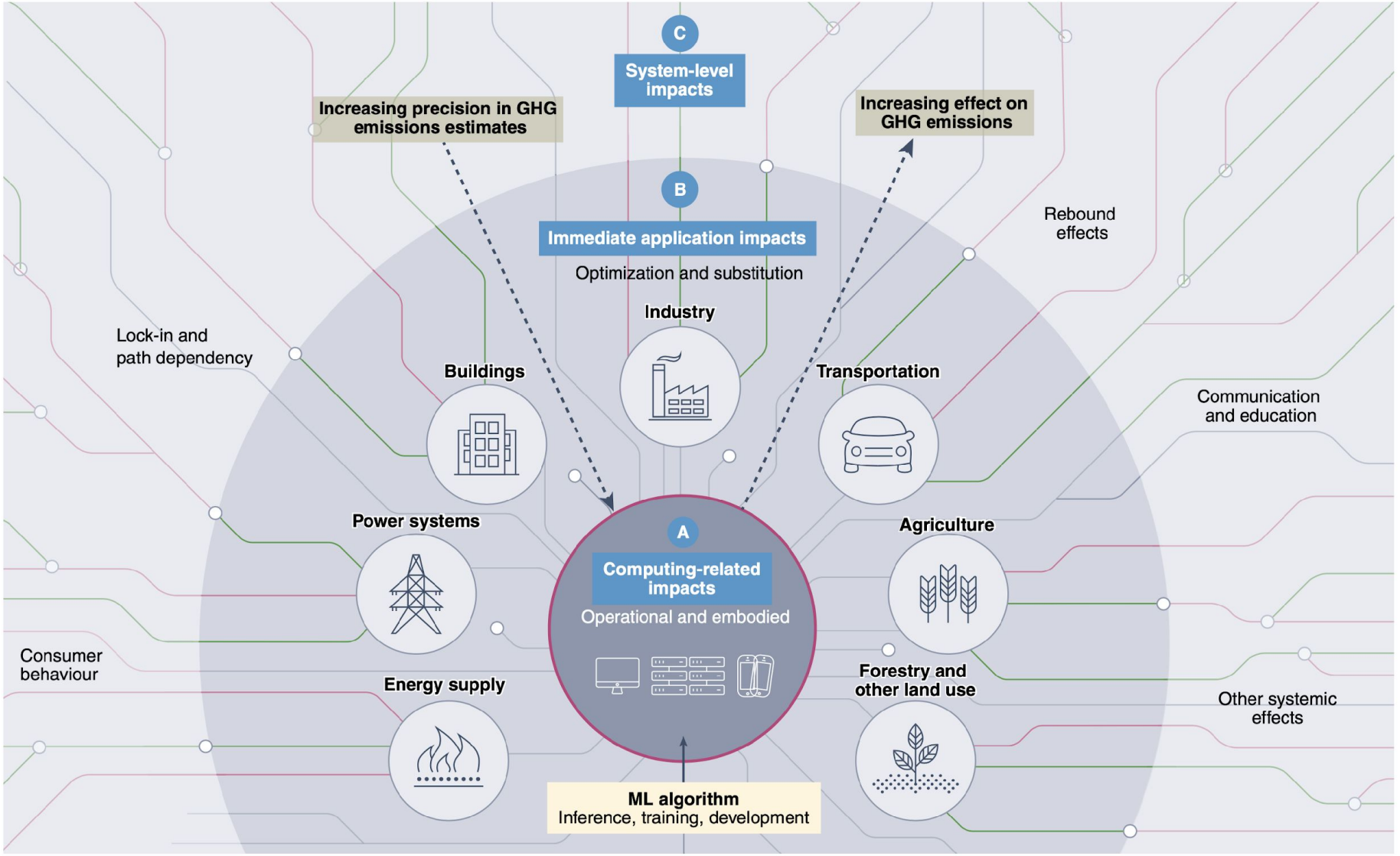
Overarching considerations

Regulating energy and resource usage of AI

Training, Inference

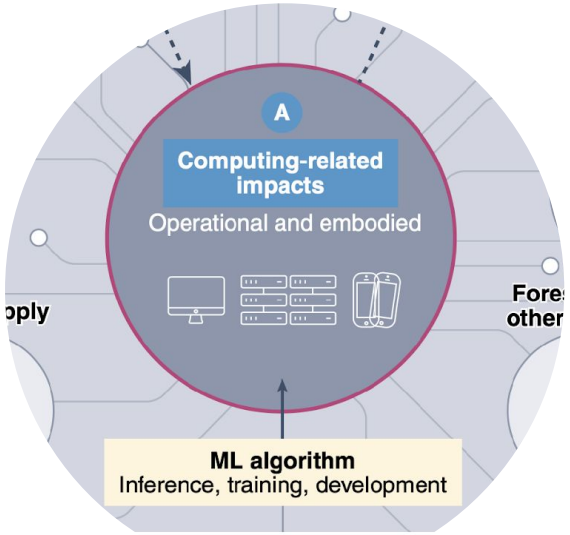
Impacts differentiated by tasks and models

How AI Affects Greenhouse Gas Emissions



Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Nature Climate Change, 1-10.

How AI Affects Greenhouse Gas Emissions



Impacts from ML Computation & Hardware

Operational emissions from energy consumed during computation





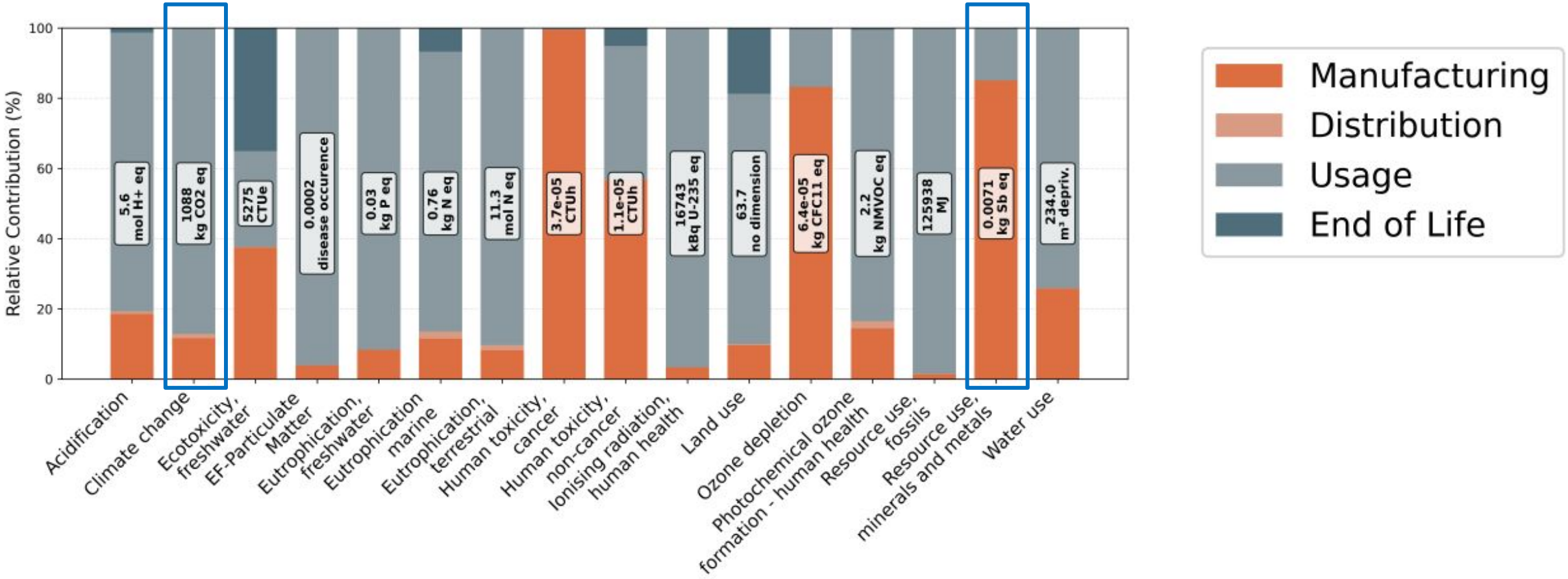
Embodied emissions from production and end-of-life of hardware







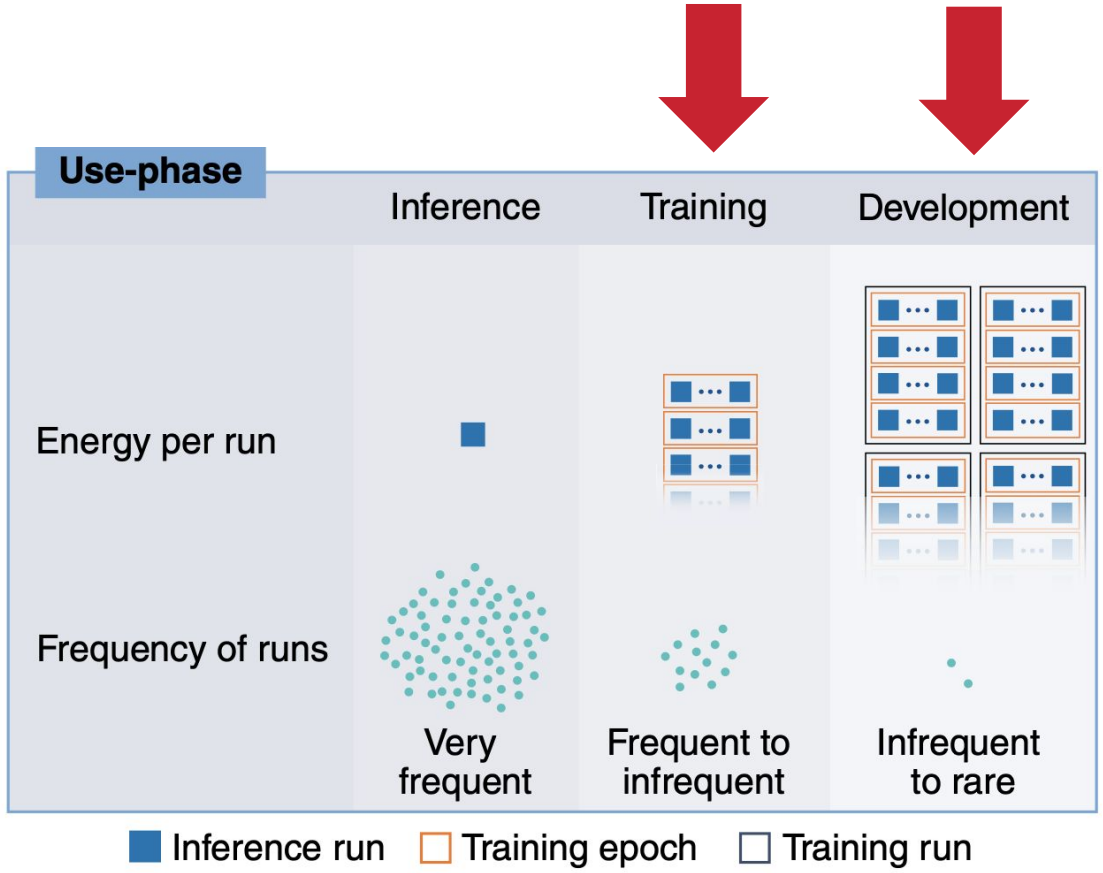
Sustainability impacts of a GPU card



- Operational emissions >> Embodied emissions
- Importance of mining

Source: arxiv.org/abs/2509.00093

Computing -Related Emissions from ML: Development & Training



Reminder:

AI is not just “one thing”

Models have different sizes



Model Size Influences the Energy Demand for Training

~ 2 days of powering Berlin

Parameters and electricity consumption evaluations for several leading LLMs.

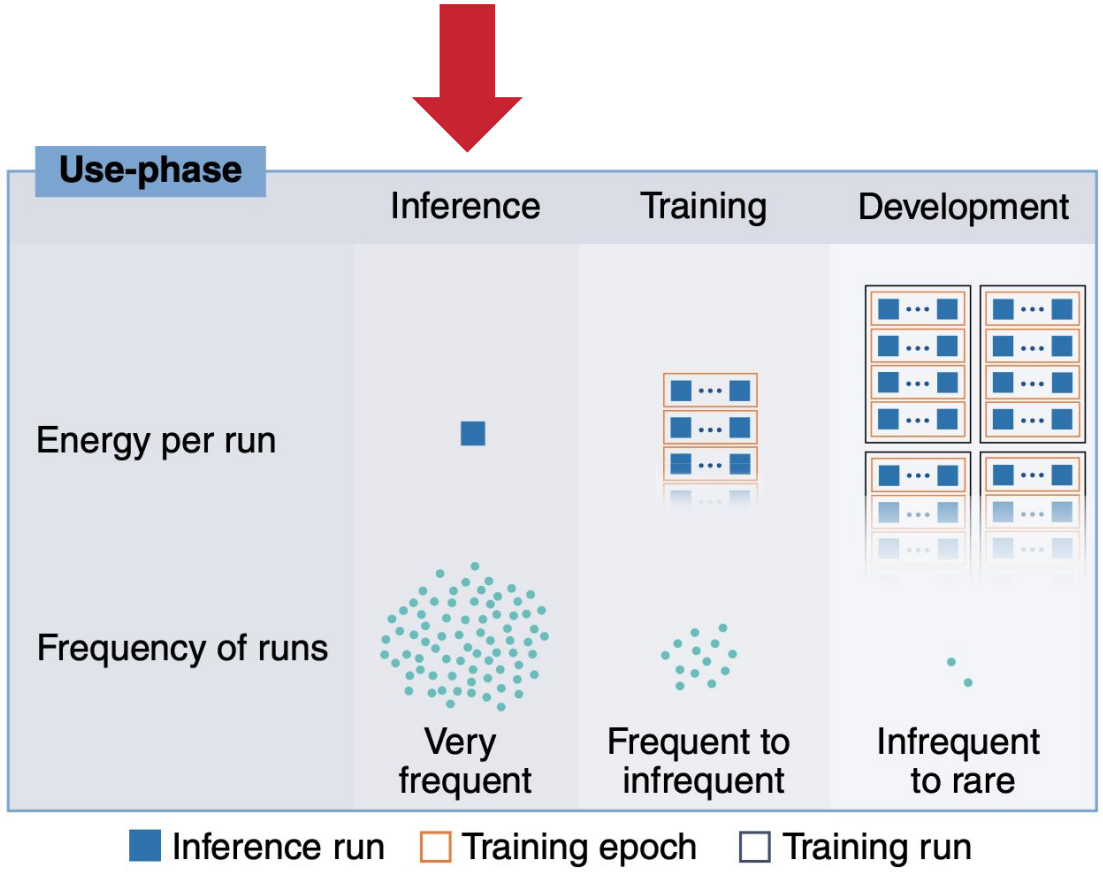
Developer	Model	Released year	Number of parameters (billion)	Accelerator type	Electricity consumption (MWh)	CO_2 emissions (ton)	Literature
Google	BERT	2018	0.1	GPU V100	1.5	0.7	[13,14]
	GLaM	2021	1162	TPU v4	456	40	[19]
	PaLM	2022	540	TPU v4	3436	271.4	[38]
OpenAI	GPT-2	2019	1.5	TPU v3	1.7	0.7	[14,39]
	GPT-3	2020	175	GPU V100	1287	552	[32,40]
	GPT-4	2023	1800	GPU A100	51,772 to 62,318	12,456 to 14,994	[41,42]
Meta	OPT	2022	175	GPU A100	324	75	[43]
	LLaMA 2	2023	7	GPU A100	74	31.2	[44]
			13		147	62.4	[44]
			70		688	291.4	[44]
	LLaMA 3	2024	1	GPU H100	/	107	[45,46]
			3		/	133	[45,46]
			8		/	390 to 420	[45,46]
			70		/	1900 to 2040	[45,46]
			405		/	8930	[45,46]
BigScience	BLOOM	2022	1	GPU A100	159	9.1	[18]
			104		267	15.2	[18]
			176		433	24.7	[18]
DeepMind	Gopher	2021	280	TPU v3	1151	380	[13,47]
				/	1066	352	[4]
TII	Falcon	2023	40	GPU A100	163[b]	8 to 90[b]	[48–50]
			180		2,940[b]	147 to 1,617[b]	[48–50]
Mistral AI	Mixtral 8 × 7B	2024	47[a]	GPU A100	232[b]	6.3[b]	[51–53]
DeepSeek	DeepSeek-V3	2024	671[a]	GPU H800	1,025[b]	550[b]	[54,55]

[a] These values denote the total number of parameters, with different quantities activated depending on the model architecture.

[b] The electricity consumption data of some LLMs lack direct values from open-published literature. Thus, these data are estimated, and the estimation process is provided in Appendix Table A1.

Source: Zhenya Ji and Ming Jiang 2026, https://doi.org/10.1016/j.rser.2025.116159.

Computing-Related Emissions from ML: Inference



Inference Emissions Are Growing

- Both Meta and Google report that inference emissions outweigh those of training
- Depending on use, training or inference may dominate
- Consumers need transparency to trade off between models



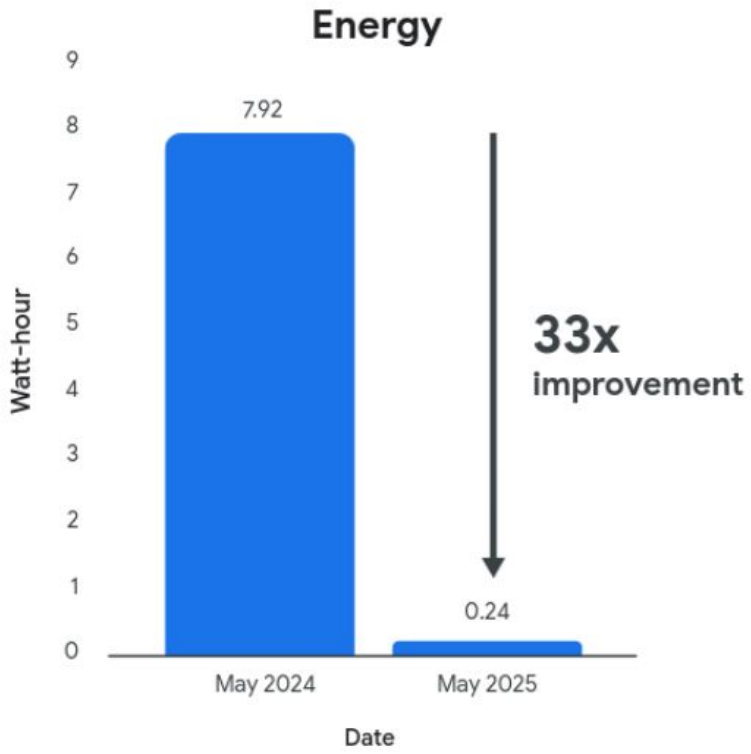
Model		Request freq.	Energy (kWh)	# inf. to equal training
Llama 3.1	8B	∞	0.011	276B
		8	0.051	59.5B
		1	0.333	9.12B
OLMo 2	7B	∞	0.018	20.9B
		8	0.049	7.68B
		1	0.348	1.05B
OLMo 2	13B	∞	0.033	22.1B
		8	0.057	12.8B
		1	0.386	1.89B

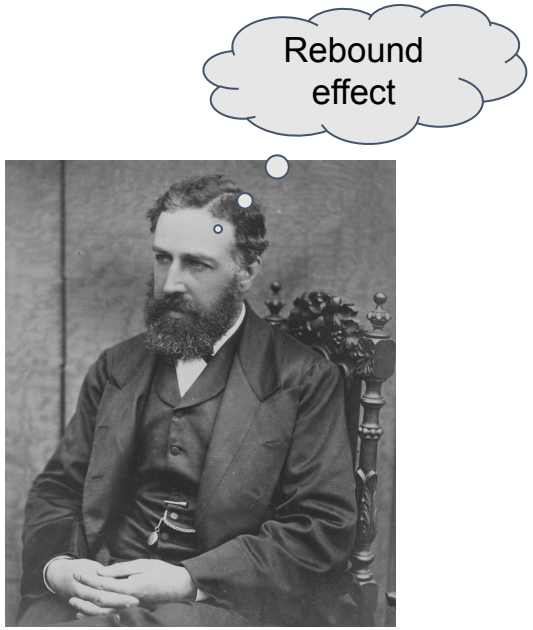
Jacob Morrison, Clara Na, Jared Fernandez, Tim Dettmers, Emma Strubell, Jesse Dodge. Holistically Evaluating the Environmental Impact of Creating Language Models. ICLR 2025 (Spotlight).

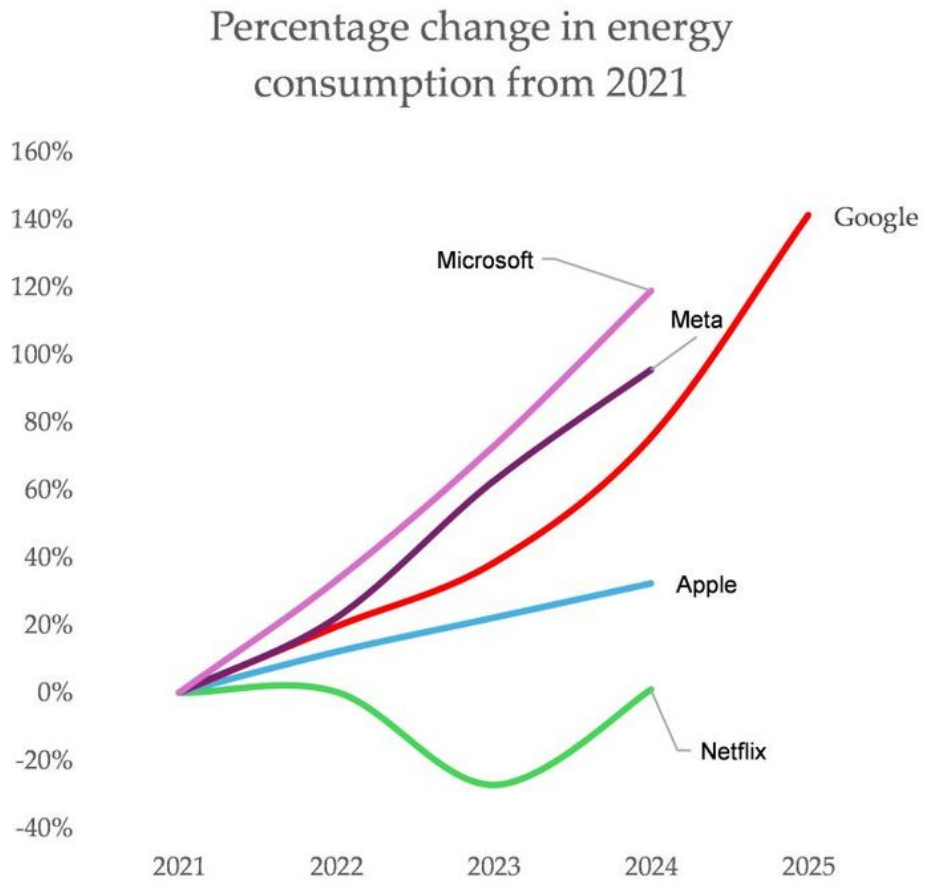
Higher Efficiency … Largely Outweighed by Usage Growth

Gemini efficiency improvements

Energy and carbon efficiency improvements per median Gemini Apps text prompt from May 2024 to May 2025.

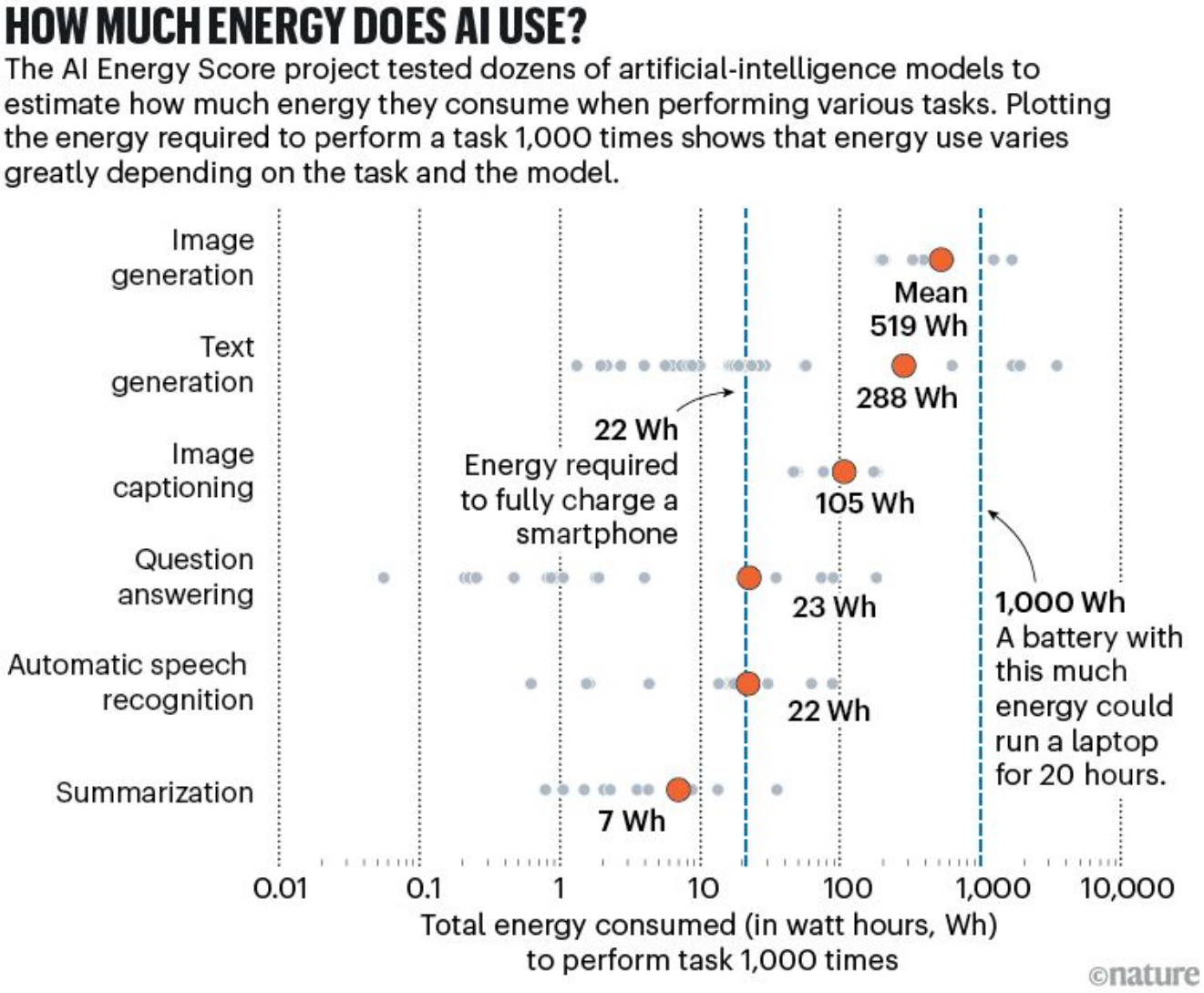






Sources: Google, Ketan Joshi

Large Differences in Inference Energy Needs Across Tasks



Sources: https://www.nature.com/articles/d41586-025-00616-z based on Luccioni et al., Hugging Face Energy Score

Poll Two

What is the order of magnitude between

one **prompt**

and one **8-sec medium-definition video**?

a) x 10, b) x 100, c) x 1,000, d) x 10,000

Sustainable AI Group

Lights, Camera, Carbon: Architectural Scaling Laws for Video Generation Energy Consumption

Nidhal Jegham, Boris Gamazaychikov, Sasha Luccioni

Sora 2.0 Pro is gone.
Here's what it cost to run.

OpenAI discontinued Sora in March 2026. We estimated its energy footprint from architecture and API timing alone, with no access to model weights required.

Per 8-second, 720p video generated

418.5 Wh

≈ 1,700× the energy of one Gemini text prompt (0.24Wh)

602.6 GWh

estimated if 4M users generated 2 videos/day for 6 months

~113,000

US households' electricity use during the same period, equivalent

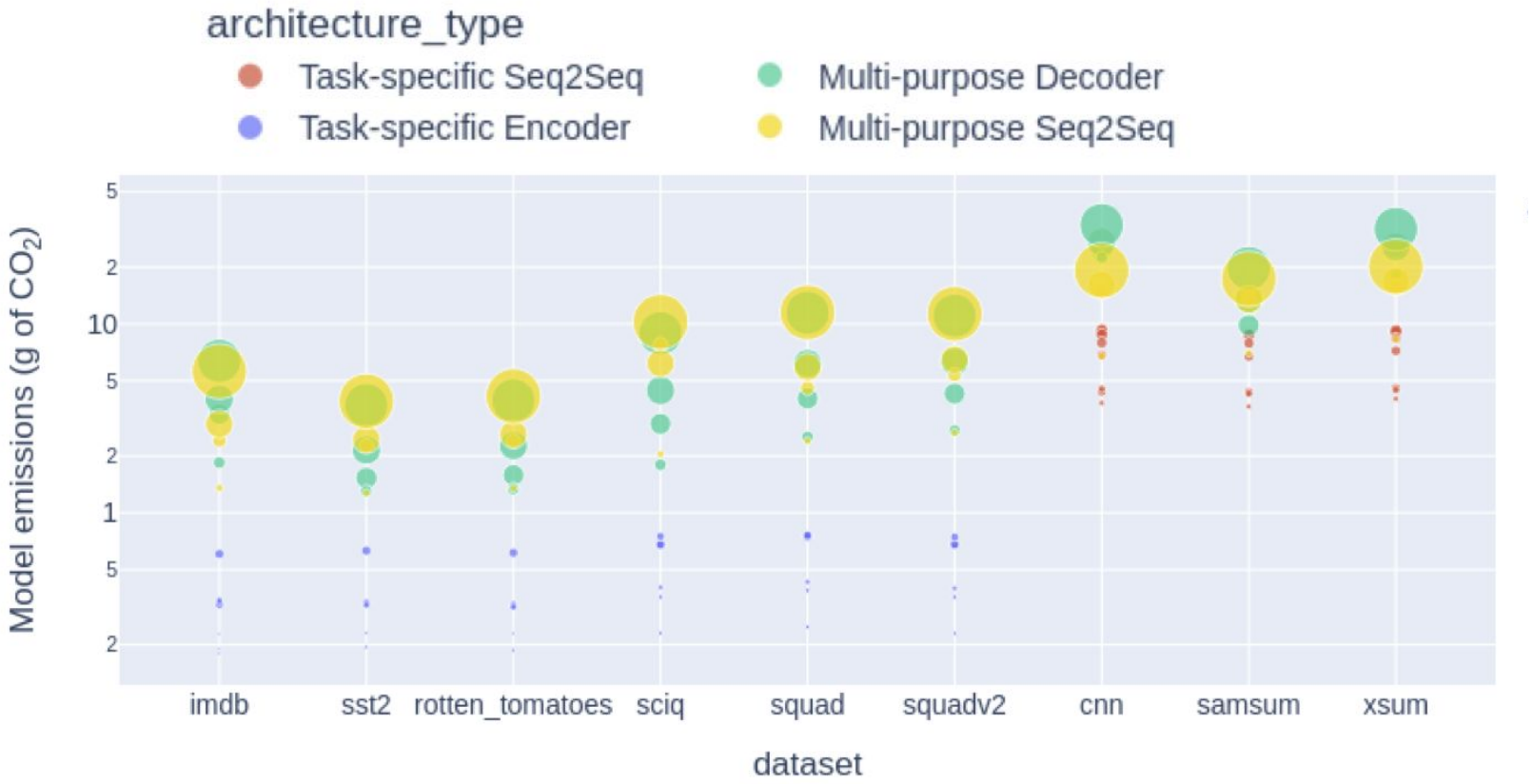
AI with Little Energy Use Is Common & Possible

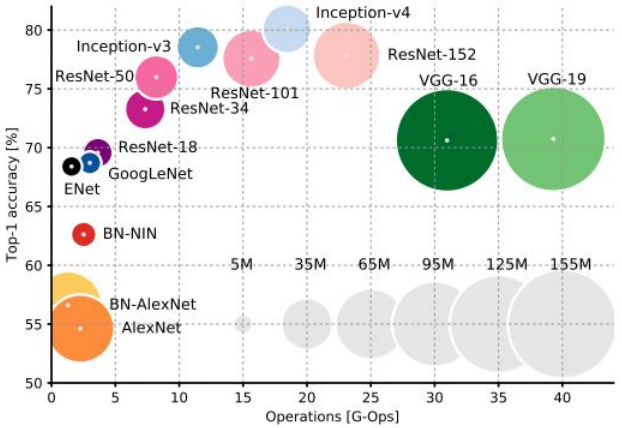
→ FRUGAL AI

Common Myth	Strategic Reality	Takeaways
Frugal AI = Cheap, Low Quality	Frugal AI prioritizes efficient, high-impact solutions, not cutting corners.	Reframe cost-saving as value-maximization through smarter, leaner AI investments.
Low-tech can't scale	Simple tools (WhatsApp, SMS, low-code AI) power platforms used by millions globally.	Embrace context-appropriate scaling – not every solution needs enterprise-grade infrastructure.
Only big tech can do AI	AI is increasingly democratized; startups and small teams build effective, lean models.	Tap into ecosystem partnerships and pilot frugal solutions for niche or underserved markets.
Frugal = Outdated Tech	Frugal AI leverages modern techniques (edge AI, model compression, efficient models).	Integrate Frugal AI metrics (TCO, energy score) into procurement and solution architecture.

Frugal AI: Driving Sustainable, Inclusive, and Scalable Innovation whitepaper, 2025

Using General-Purpose Models for Inference Is Often Inefficient Compared to Task-Specific Models





More parameters

→ more energy

Luccioni, Alexandra Sasha, Jernite, Yacine, and Strubell, Emma. "Power Hungry Processing: ⚡Watts ⚡ Driving the Cost of AI Deployment?" *arXiv preprint arXiv:2311.16863 (2023)*

More Data ≠ Always Better Performance

Table 6.1
Overall test set performance of local and global models (the best prediction performance on each local test set is highlighted in bold font).

Model	Trained on	Training epochs	Training samples	Training effort	Test RMSE (↓) on		
					Stanford	SIRTA	DEWA
SUNSET	Stanford	21.2 ± 6.1	125,876	2.67 ± 0.77M	2.68	136.96	155.82
	SIRTA	9.6 ± 1.8	438,172	4.21 ± 0.81M	6.53	96.42	219.79
	DEWA	11.8 ± 2.2	85,278	1.01 ± 0.19M	4.65	160.28	**89.49**
	Stanford+SIRTA	28.6 ± 5.6	564,048	16.13 ± 3.15M	**2.61**	**96.34**	129.21
ConvLSTM	Stanford	11.7 ± 2.7	125,876	1.47 ± 0.34M	2.62	113.05	N/A
	SIRTA	5.6 ± 2.6	438,172	2.49 ± 1.13M	4.39	96.46	N/A
	Stanford+SIRTA	6.0 ± 1.7	564,048	3.36 ± 0.99M	**2.62**	**95.53**	N/A

Paper also shows that **fine-tuning a global model on DEWA data** (all or big enough subset) yields **similar performance** for DEWA compared to the **local DEWA model** (but with higher total training effort)

Nie, Yuhao, et al. "Sky image-based solar forecasting using deep learning with heterogeneous multi-location data: Dataset fusion versus transfer learning." *Applied Energy* 369 (2024): 123467.

Stanford = Stanford University (Stanford, CA, USA)
SIRTA = SIRTA Atmospheric Observatory (Palaiseau, France)
DEWA = Dubai Electricity and Water Authority (Dubai, UAE)

Towards Local Models

Julien Chaumond • 2nd
CTO at Hugging Face
2mo •

Connect

This is where we are right now. And it feels pretty magical 🧚

Qwen3.6 27B running inside of Pi coding agent via Llama.cpp on the MacBook Pro

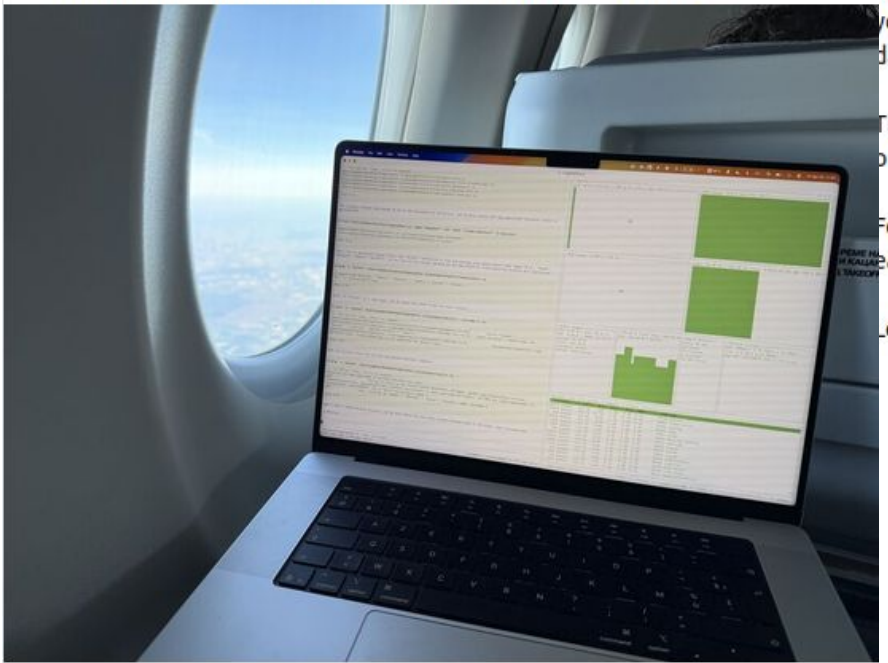
For non-trivial tasks on the **Hugging Face** codebases, this feels very, very close to hitting the latest Opus in Claude Code, or whatever shiny monopolistic closed source API of the day is.

In full airplane mode.

Most people haven't realized this yet.

If you have, it means you have a huge headstart to what I call the second revolution of AI.

Powerful local models for efficiency, security, privacy, sovereignty 🔥



Clem Delangue 🤗 • Following
Co-founder & CEO at Hugging Face
1w •

A study from **Stanford University** showed that 71.3% of chatgpt queries could be accurately answered by a local model. I suspect a major part of enterprise AI workloads could be run locally too for free (compared to the massive costs of frontier API cost).

Also, it reduces the risk of these workloads being taken away from you because you own the models instead of renting them - which sounds like a good idea these days haha.

That's why we're introducing the ability for everyone to filter AI models on **Hugging Face** based on your local hardware.

For me, there are 800k+ public models that fit on my M5 24GB and that I can use easily thanks to llamacpp.

Let's go local AI!

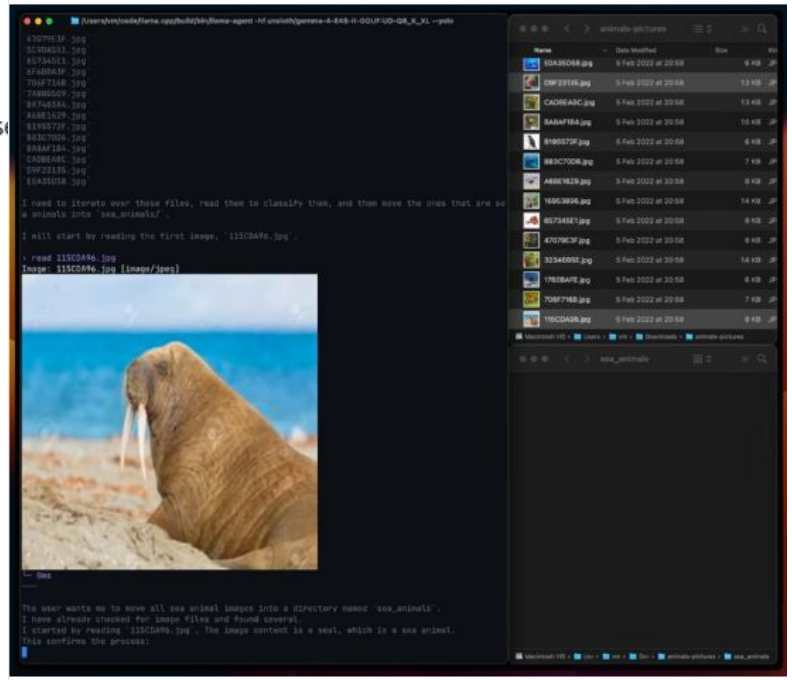
Victor Mustar • 2nd
Head of Product at Hugging Face
3mo •

+ Follow

I just watched the latest Google Gemma 4 open model work on my laptop. No cloud. No API key. No catch.

I pointed Gemma 4 E4B at a folder of mixed photos and asked it to sort the sea animals into their own directory. It opened each image, identified what it saw, and moved the files. All locally, in one agentic session, on a model small enough to fit in 5GB of RAM.

Three years ago this needed a data center. Today it runs while you're offline on a plane.

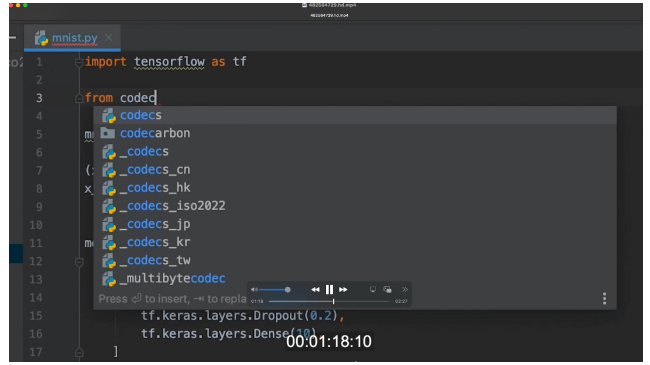


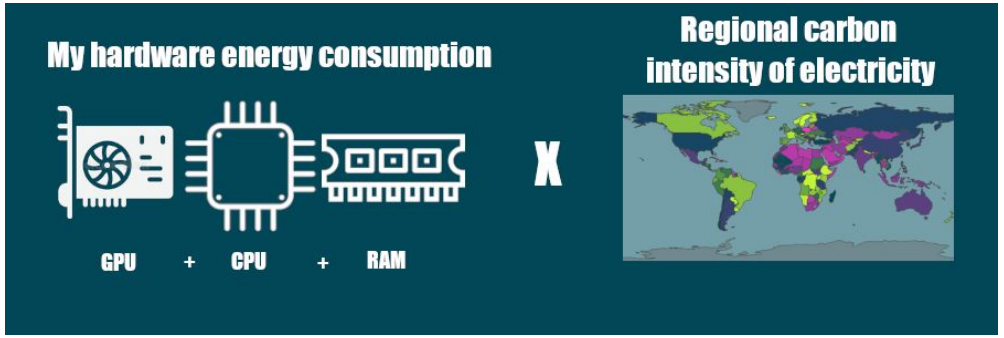
Frugality in AI development

- Frugal AI Challenge: Encourage academic and industry actors to keep efficiency in mind when deploying AI models
- Tracking energy consumption **and** performance
- Incentivize frugality in AI deployment while also addressing real-world challenges.
- Tasks:
  - Detecting climate disinformation
  - Classifying regions at risk of wildfires,
  - Detecting illegal deforestation.



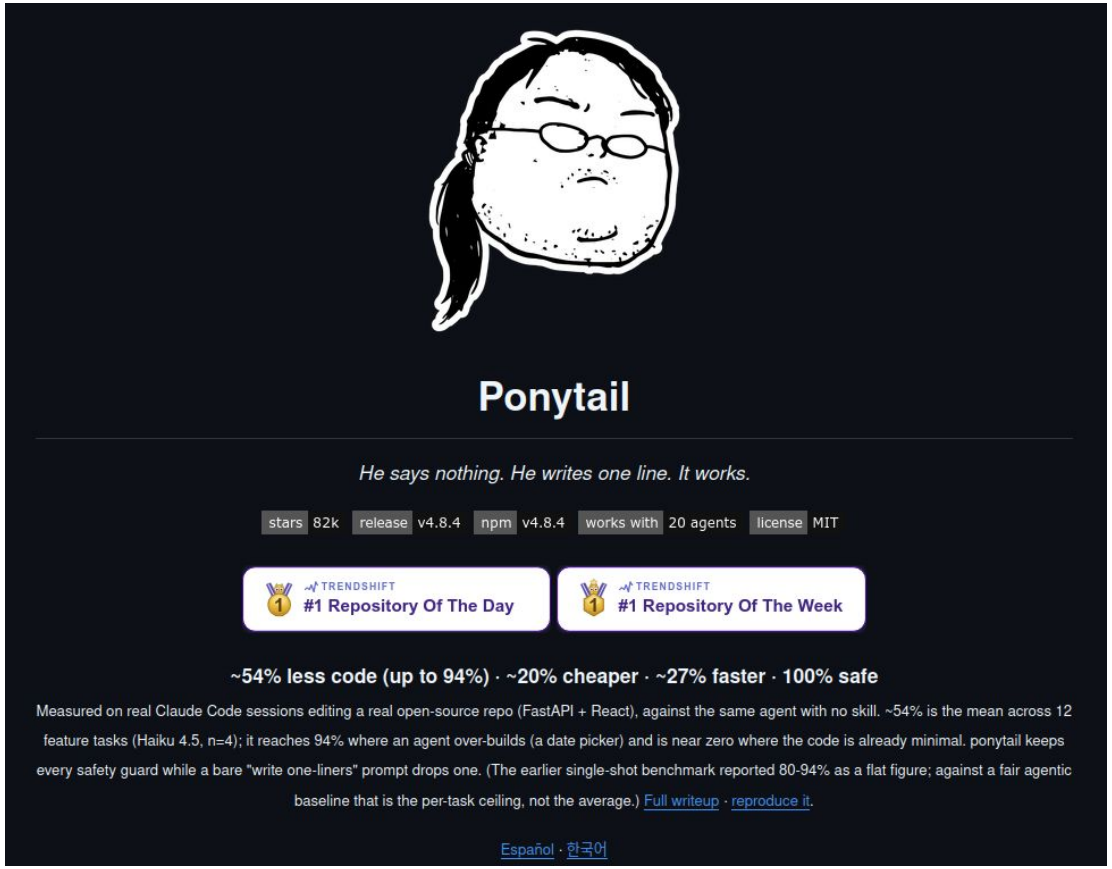
Code Carbon

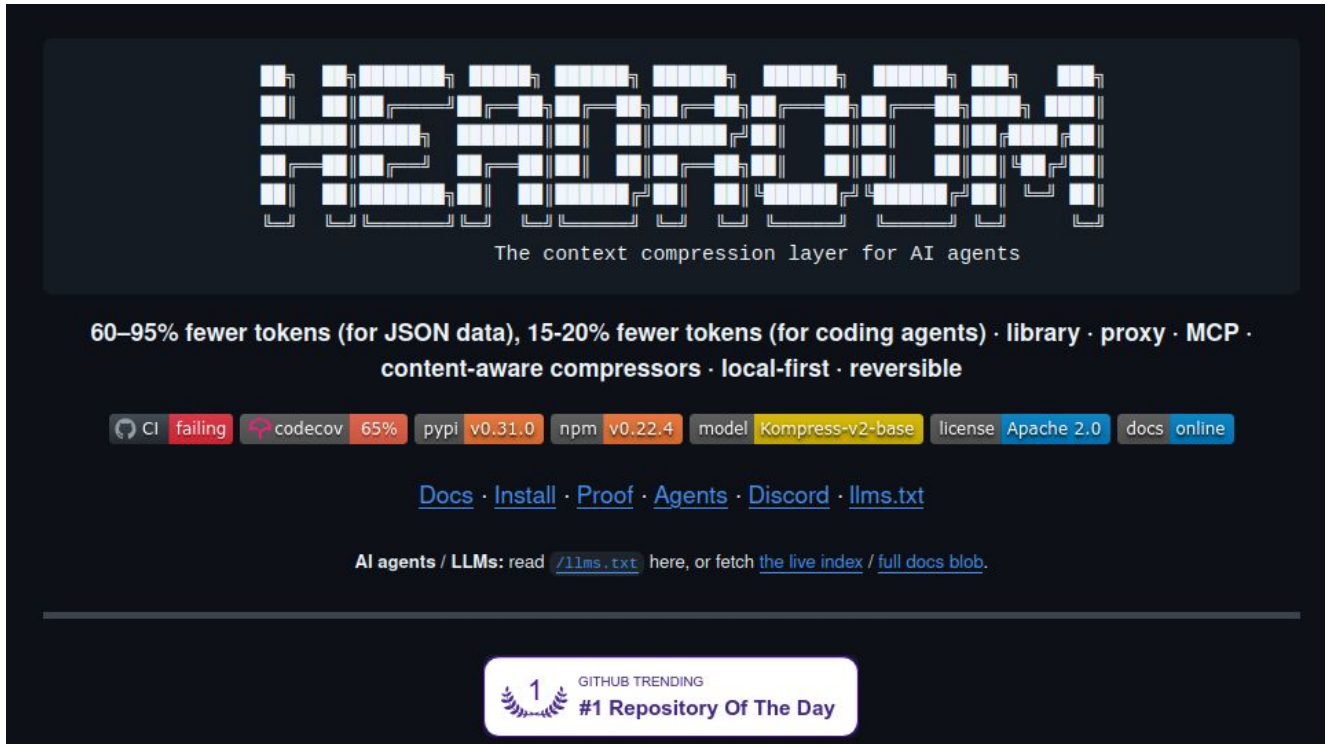




Layers **around the model** also matter

→ Escaping **"Token Maxxing"** via Skills, etc.





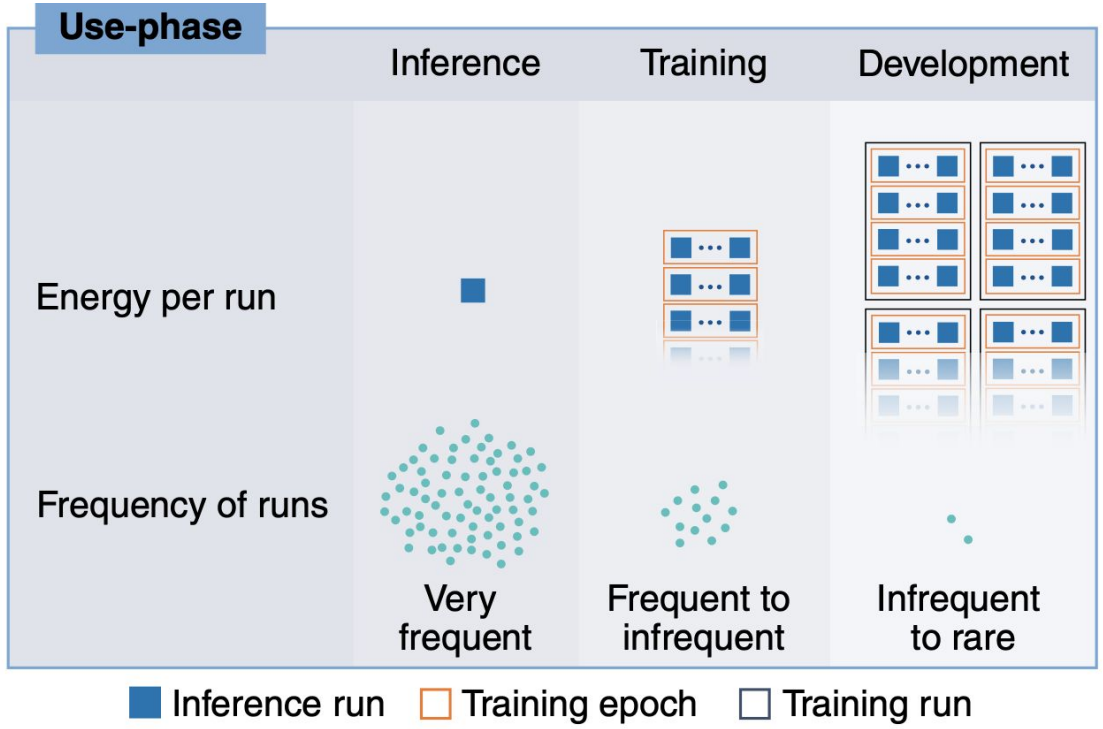
TinyML: Running AI on Microcontrollers

	Cloud AI (NVIDIA V100)		Mobile AI (iPhone 11)		Tiny AI (STM32F746)		ResNet-50	MobileNetV2	MobileNetV2 (int8)
Memory	16 GB	4× →	4 GB	3100× →	**320 kB**	← gap →	7.2 MB	6.8 MB	1.7 MB
Storage	TB~PB	1000× →	>64 GB	64000× →	**1 MB**	← gap →	102MB	13.6 MB	3.4 MB

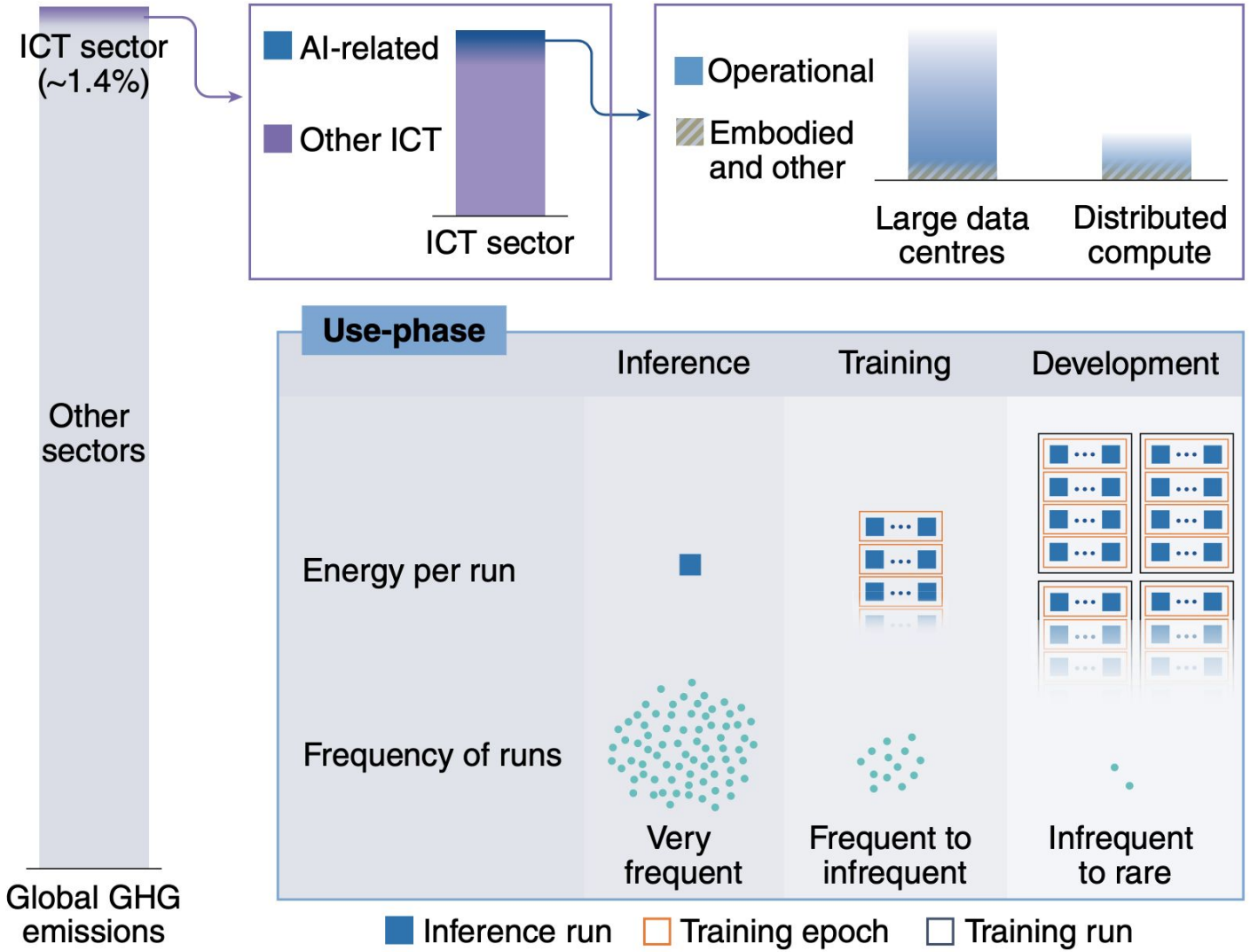


Lin et al 2024 Tiny Machine Learning: Progress and Futures; seedstudio.com

Moving from Bottom-up...

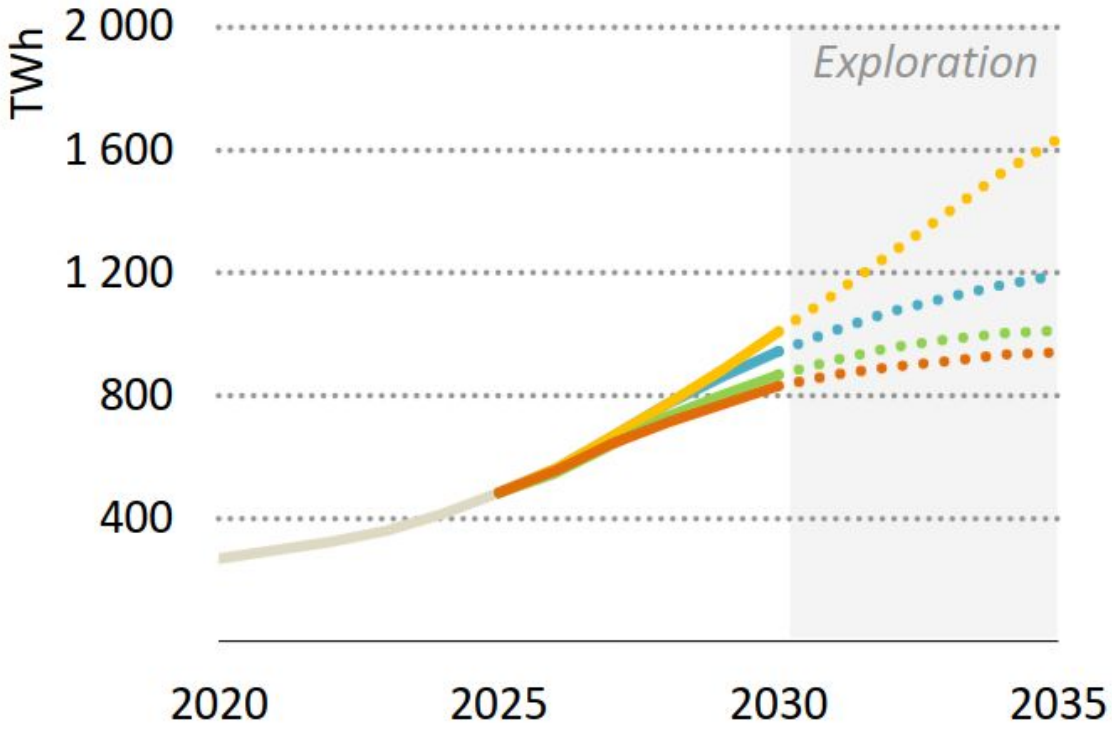


...to Top-Down Perspective



Electricity Demand is Rapidly Growing





6% increase

between 2010 and 2018

(despite 5.5x increase in compute instances)

250-500% increase

between 2020 and 2035

Rough estimates, based on:
Masanet, Eric, et al. "Recalibrating global data center energy-use estimates." Science (2020)

Source: IEA, “Energy and AI” 2026

Poll Three

Which **energy source** is expected to the **most rolled out** to **meet the increased demand** in electricity from AI?

a) renewables, b) coal, c) gas, d) nuclear

Additional Electricity Supply:

Carbon Intensity?

Competing uses?

Associated Impacts?

ENERGY

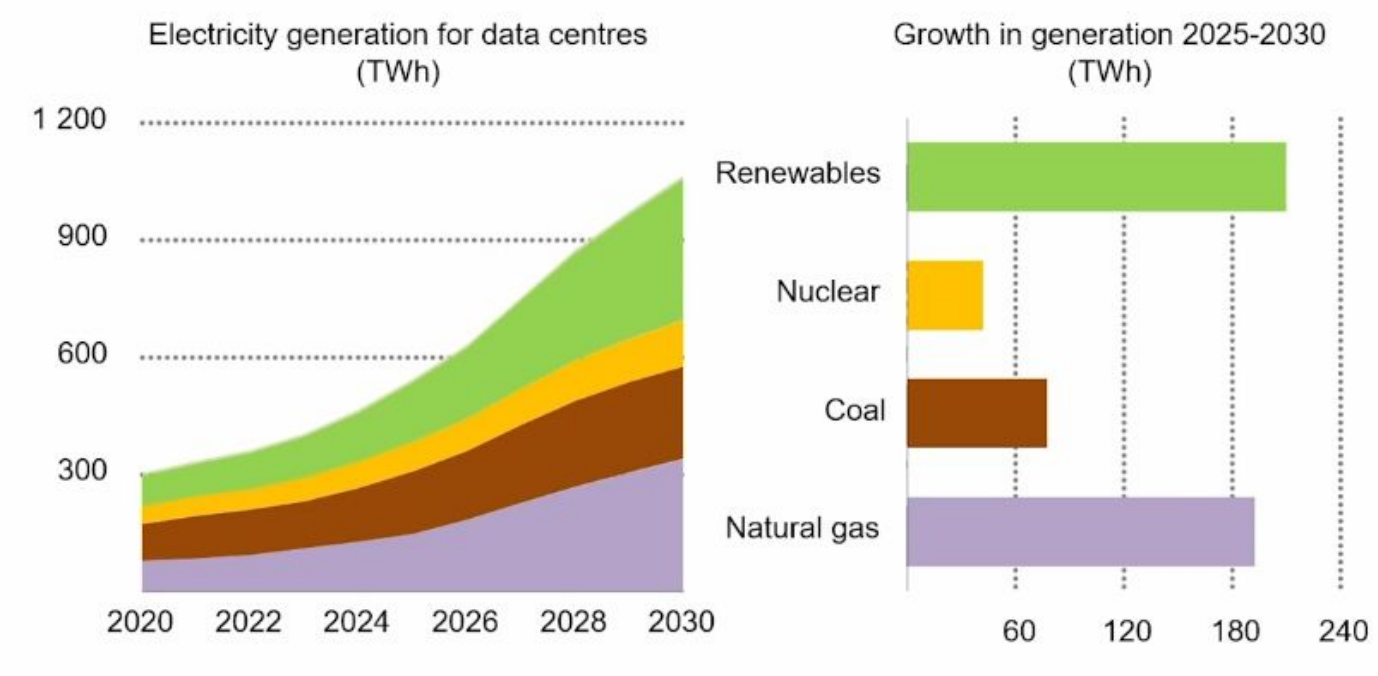
Chevron to fuel massive Microsoft data center in Texas using natural gas

PUBLISHED MON, JUN 22 2026•7:59 AM EDT | UPDATED MON, JUN 22 2026•3:19 PM EDT



COMPANIES | ENVIRONMENT

Frontier Raises Over $900 Million for Carbon Removal Investments from Google, Anthropic and Others



Source: IEA, “Energy and AI” 2026

GRID INTEGRATION OF DATA CENTERS: SPIKY POWER CONSUMPTION

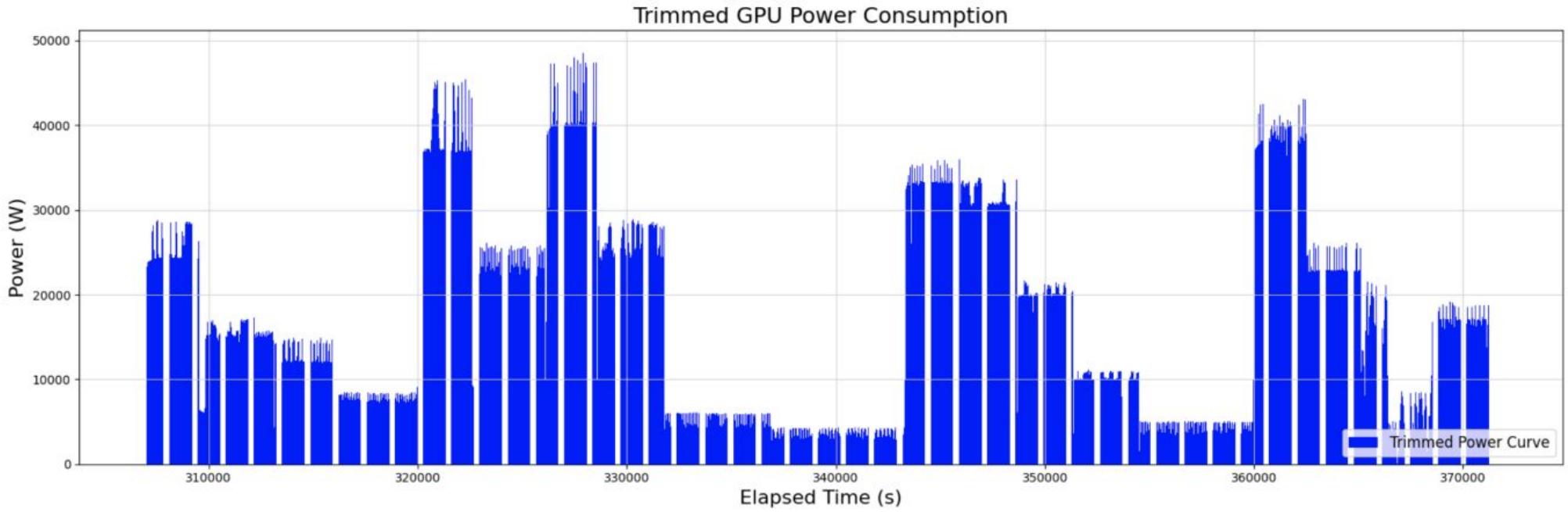


Fig. 6. Power Consumption of BERT in MIT Supercloud Dataset (Peak power consumption of approximately 48.70 kW, with an average consumption of 17.80 kW and a standard deviation of 12.39 kW. The job ran for an extended period of 4 days).

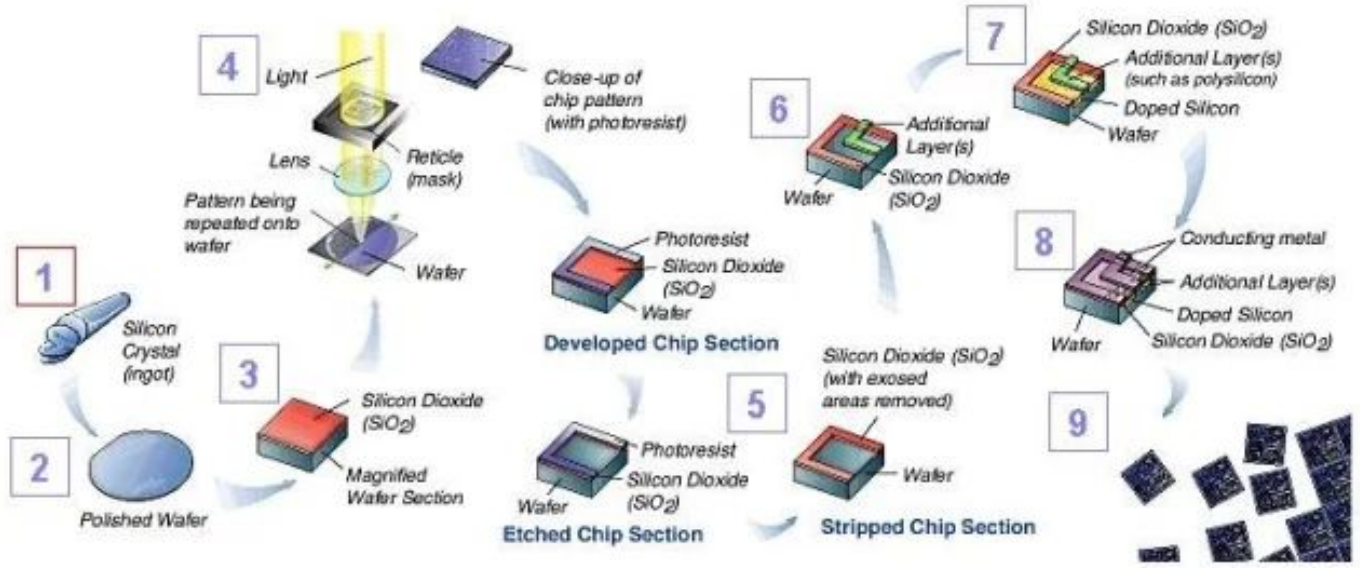
Li, Yuzhuo, et al. "The unseen AI disruptions for power grids: LLM-induced transients." *arXiv preprint arXiv:2409.11416* (2024).

Takeaways for practice (energy)

- Thinking about energy requirements of the models is important
- Not all models use a lot of energy! Big models vs small models 💻
- The consumption of the largest models doesn't discredit AI-for-climate as a whole
- Try to use a small / task-specific when possible: significant efficiency gains
- Think about which tasks are most impactful vs. potentially wasteful
- There are tools to compute the energy/GHG impacts of your project, e.g., Code Carbon

Non-GHG Sustainability Impacts: **Raw Materials**

Raw materials used to make computing chips come from 50 of our 90 naturally occurring elements here on Earth. For some, the stocks are getting rapidly depleted.



Non-GHG Sustainability Impacts:

Water

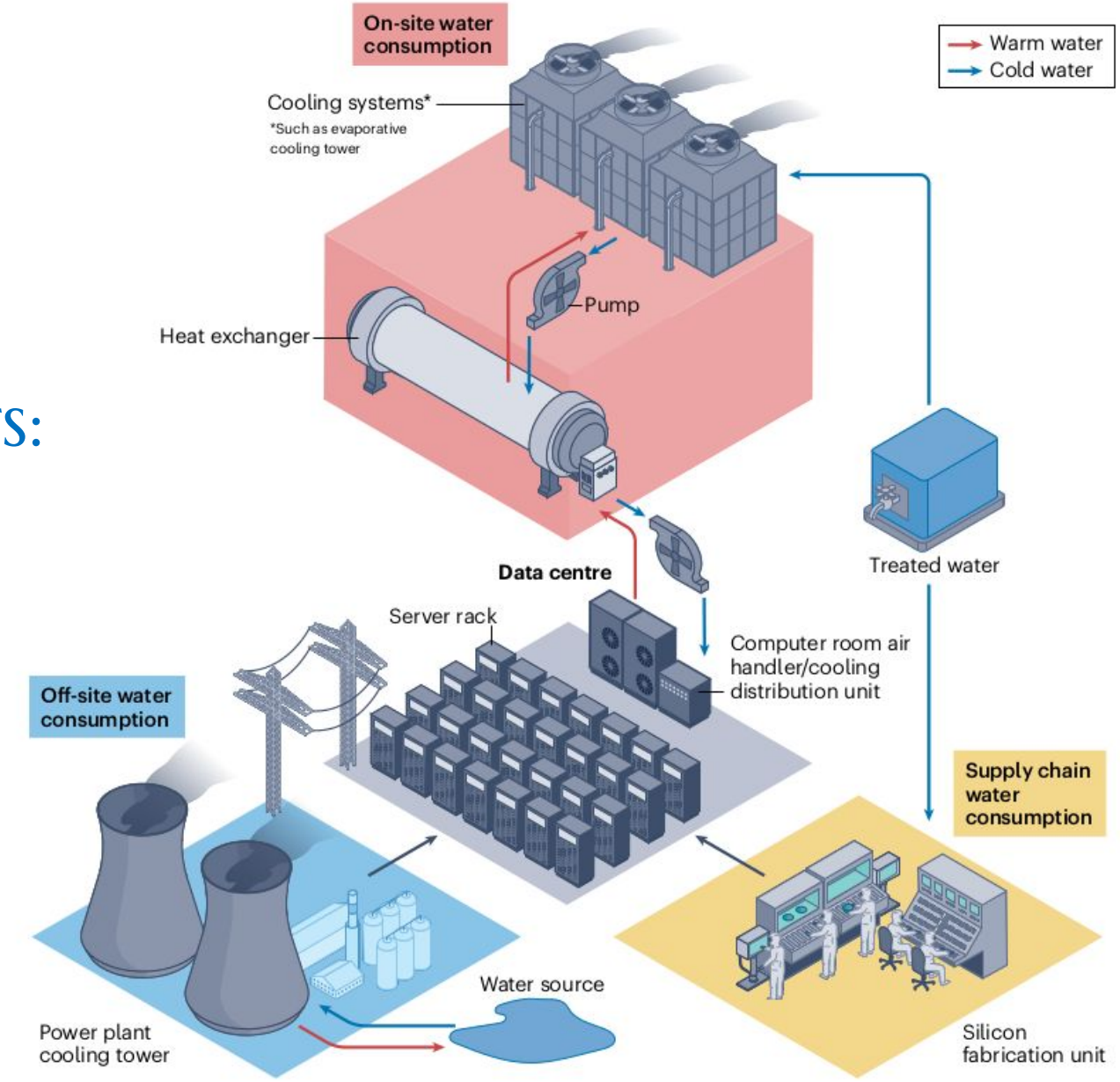


Image: Chien et al, 2026 in Nature Reviews Clean Technology

Poll Four

Is **direct** or **indirect** use of water for data centers larger?

a) direct, b) indirect, c) similar

Water: Differentiated Impacts

- **Water withdrawal** vs **Water consumption** vs **Water discharge**
- About 50% consumption, direct currently less than 0.5% of US water consumption

- **Indirect use >>> Direct use** (up to 80%)
- Projected global scope 1 + 2 water withdrawal in 2027:
  - 4.2–6.6 billion cubic meters (more than 4-6x Denmark or ½ U.K.)

- **Local impact/stress:** Problem often more **peak** than annual totals;
- Bloomberg News reports that about two-thirds of U.S. data centers built since 2022 are in high water-stress areas.



The Bottle: Annual Use. National totals seem modest and manageable on paper.



The Pipe: Peak Demand. Massive withdrawals on hot days strain local infrastructure.

Image: Shaolei Ren

Water: Solutions to Reduce Needs?

- "**Closed-loop**" liquid cooling (no evaporation) reduces water use to almost zero, but requires more energy (and can create more peaks)

- Increasingly possible to better **schedule** or **switch** between different cooling methods, but not necessarily for all parts of the data center

- **Indirect** consumption is more challenging. Depends on the choice of power technology (e.g., gas/nuclear >> solar) and on the local climate





Image: The Guardian, Wikimedia

When Impacts Are Worst… At the Same Time: **Heat waves**

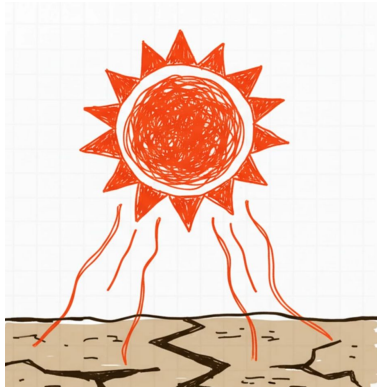
- **When** water consumption/withdrawal happens matters: worst during dry season / heat stress

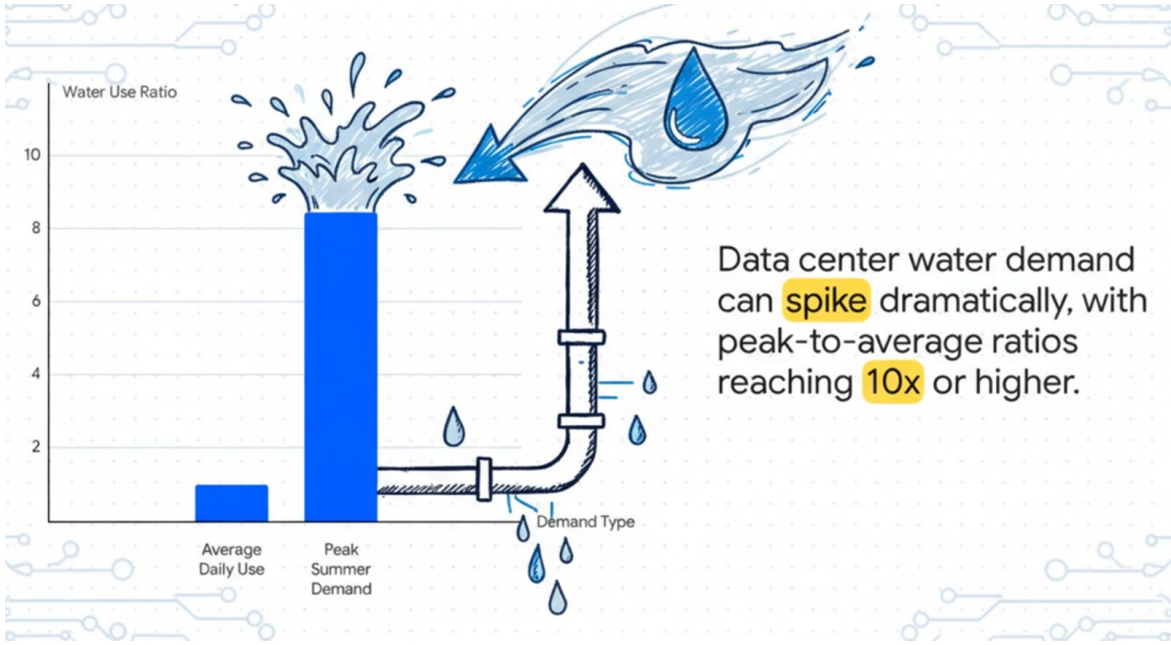
- Same for energy consumption: during heat waves AC consumption stress the grid

- **Dilemna**: to cool down data centers, one can substitute water-intensive cooling and energy-intensive cooling, but both may be under pressure

- Water can shave peak loads but resource allocation is an issue; as many communities don't have the surplus water supply
  → **need for resilience planning**







Images: Shaolei Ren, Wikimedia

SITUATED CASE STUDY: BRAZIL'S DATA CENTER EXPANSION

Massive ongoing wave of AI-driven infrastructure investments, exceeding R$25 billion

89% of Brazil's electricity matrix is **renewable**, with almost half hydro-based → **ideal for data centers?**

Increasing droughts are putting pressure of water relevant both to hydropower generation and cooling, while access to **water is already an issue** for communities

Persistent **energy poverty**, over 1.3 million people still lack reliable electricity



Example: New data centers in Ceará and Rio Grande do Sul are threatening the access to water for local communities

Lecture outline

Computing-related impacts

Application impacts

System level impacts

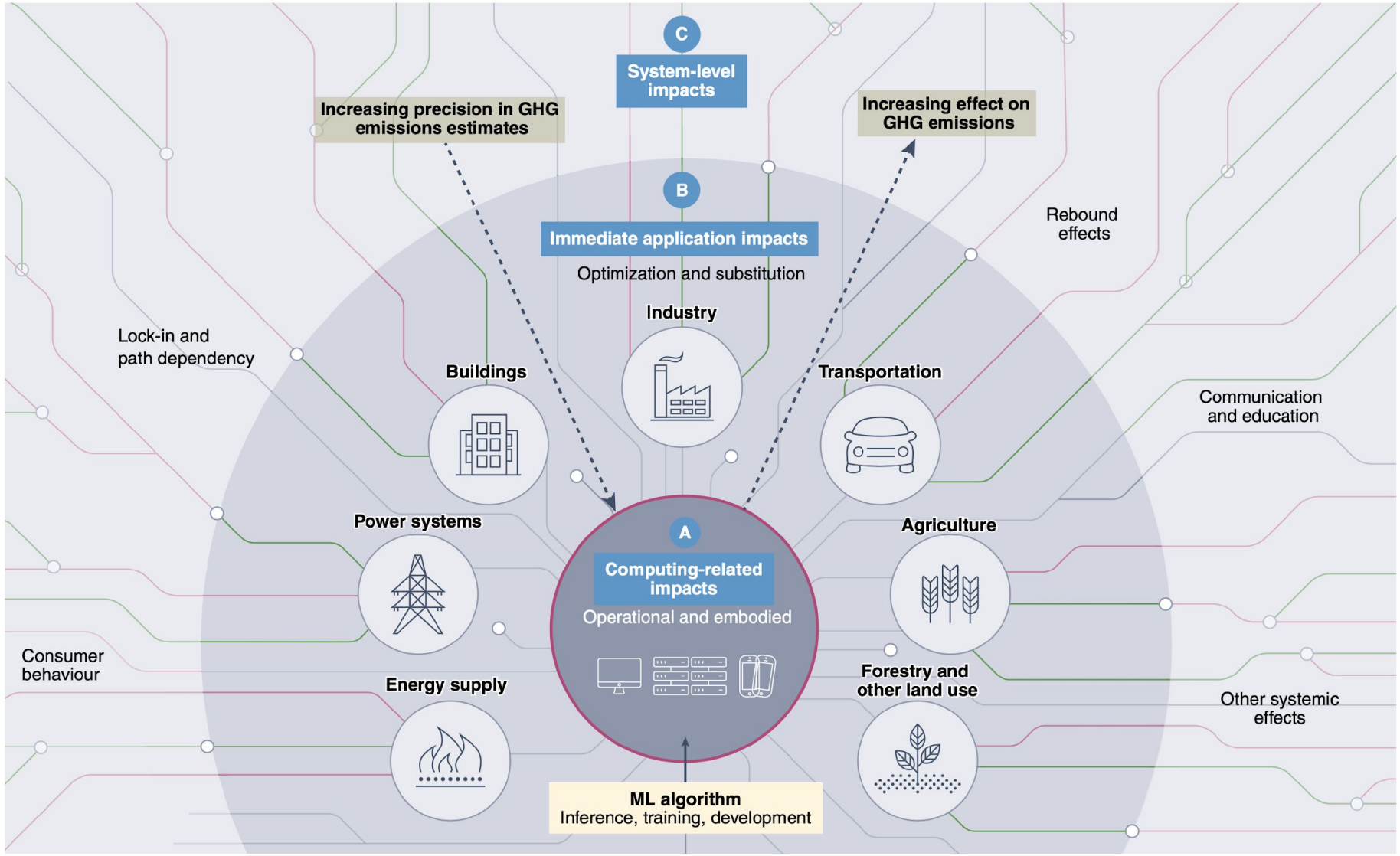
Overarching considerations

Regulating energy and resource usage of AI

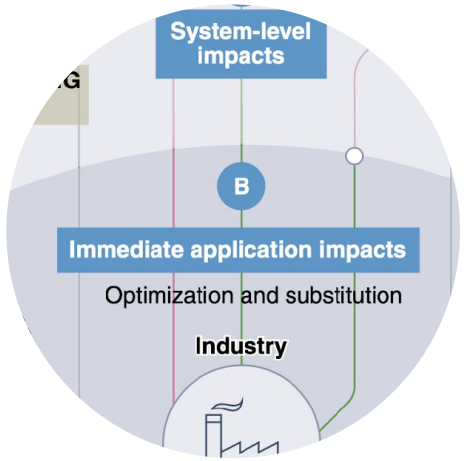
Positive impacts

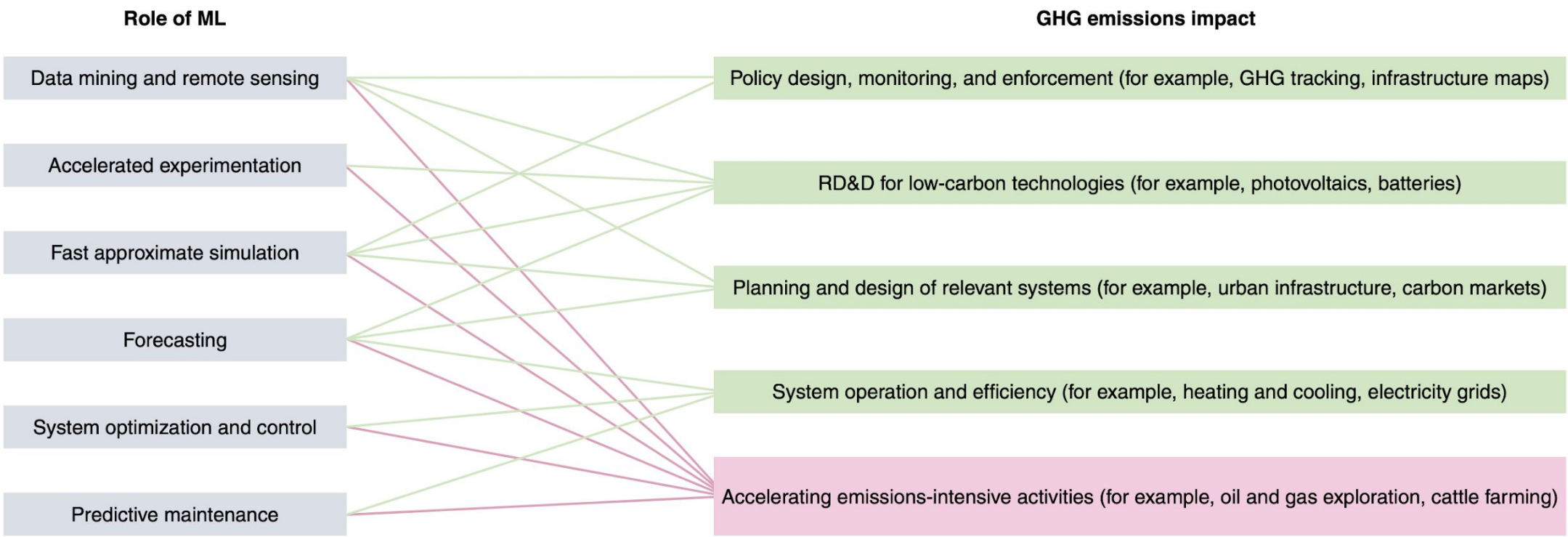
Negative impacts

How AI Affects Greenhouse Gas Emissions



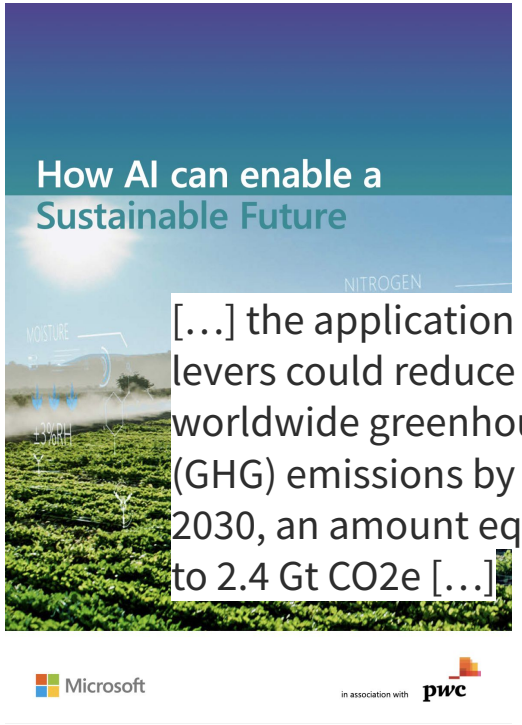
How AI Affects Greenhouse Gas Emissions





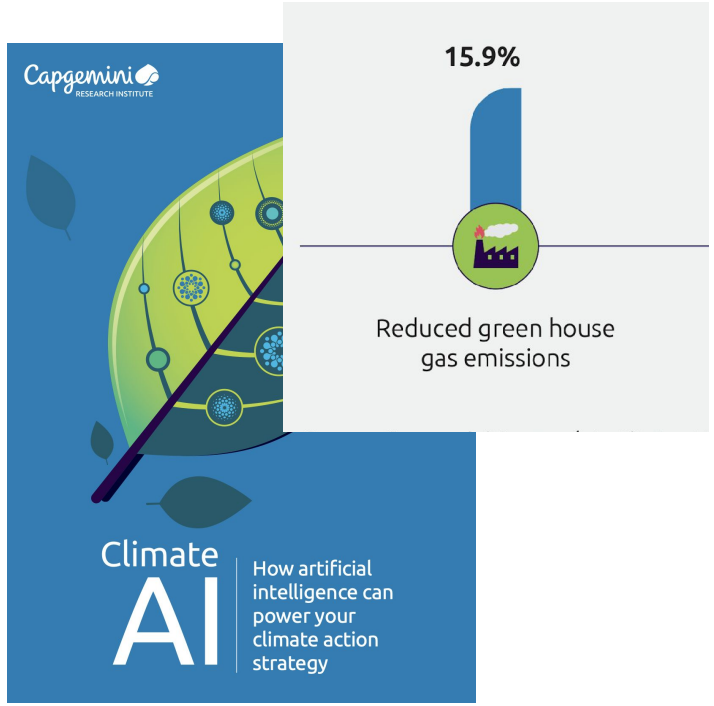
Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). *Nature Climate Change*, 1-10.

Estimates of the GHG Emissions Savings Potential are Unreliable

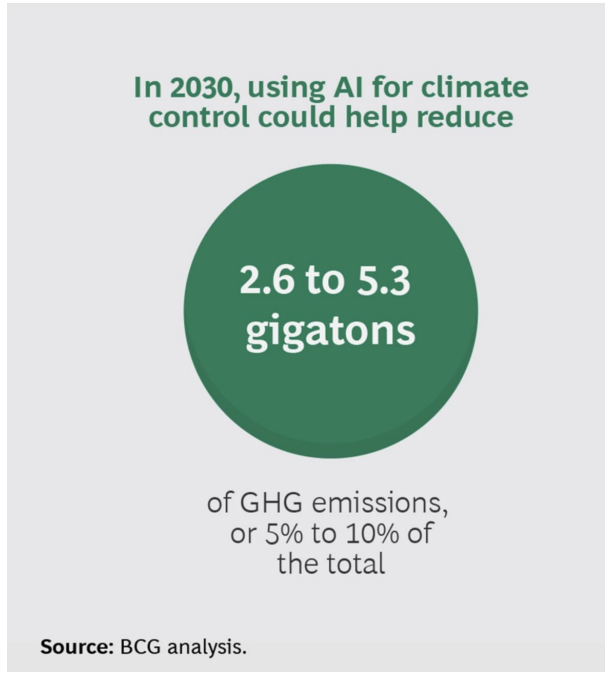


[…] the application of AI levers could reduce worldwide greenhouse gas (GHG) emissions by 4% in 2030, an amount equivalent to 2.4 Gt CO2e […]

2019



2020



2021

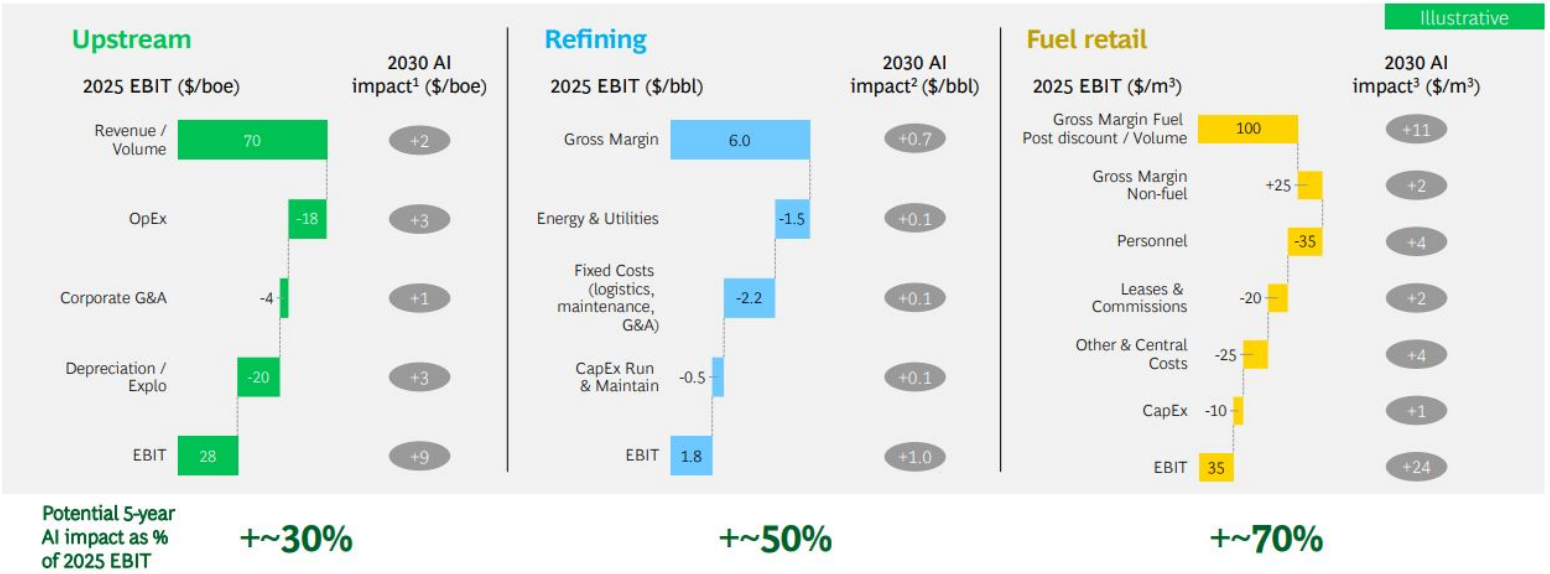
Estimates from consulting and tech firms often based on undisclosed modeling or extrapolation of stakeholder interviews → **vested interests likely overestimating the positive impacts**

AI Applications Increasing Emissions



AI use to accelerate **emissions-intensive industries**

Example: Oil and gas applications



AI Increasing Impacts Emissions (And Beyond): **Oil Drilling**

Upstream

Case study | A real-time prescriptive tool provides drilling field personnel with continuous AI guidance

Instructions' compliance from rig

RT sensor data from rig





Architecture and models



Customized screen



1

The **office** sets up the AI system and configures key technical inputs

2

- Rig sensor data is contextualized with planning information, DDRs, and any third-party software available
- AI adaptive models are executed and retrained in RT

Superior analytical performance with high adaptability

Real-time AI and physical models for one truth recommendations

Bespoke solution, not a black box

Fully tailored to client rules, org, and data

3

The **tool** assists at all moments by sending a recommendation of optimal parameters and time targets for each field activity

Simplified information

Easy-to-use interface showing only relevant data – no field engineer needed to interpret complex inputs

4

Field personnel follow recommendations, saving minutes from each activity

They also send any feedback as part of the continuous improvement process

End-to-end process taking < 5s to allow for immediate action from the field

Existing platform with total rig integration

10-20%

Capex reduction potential

Focus on optimizing ILT (invisible lost time)

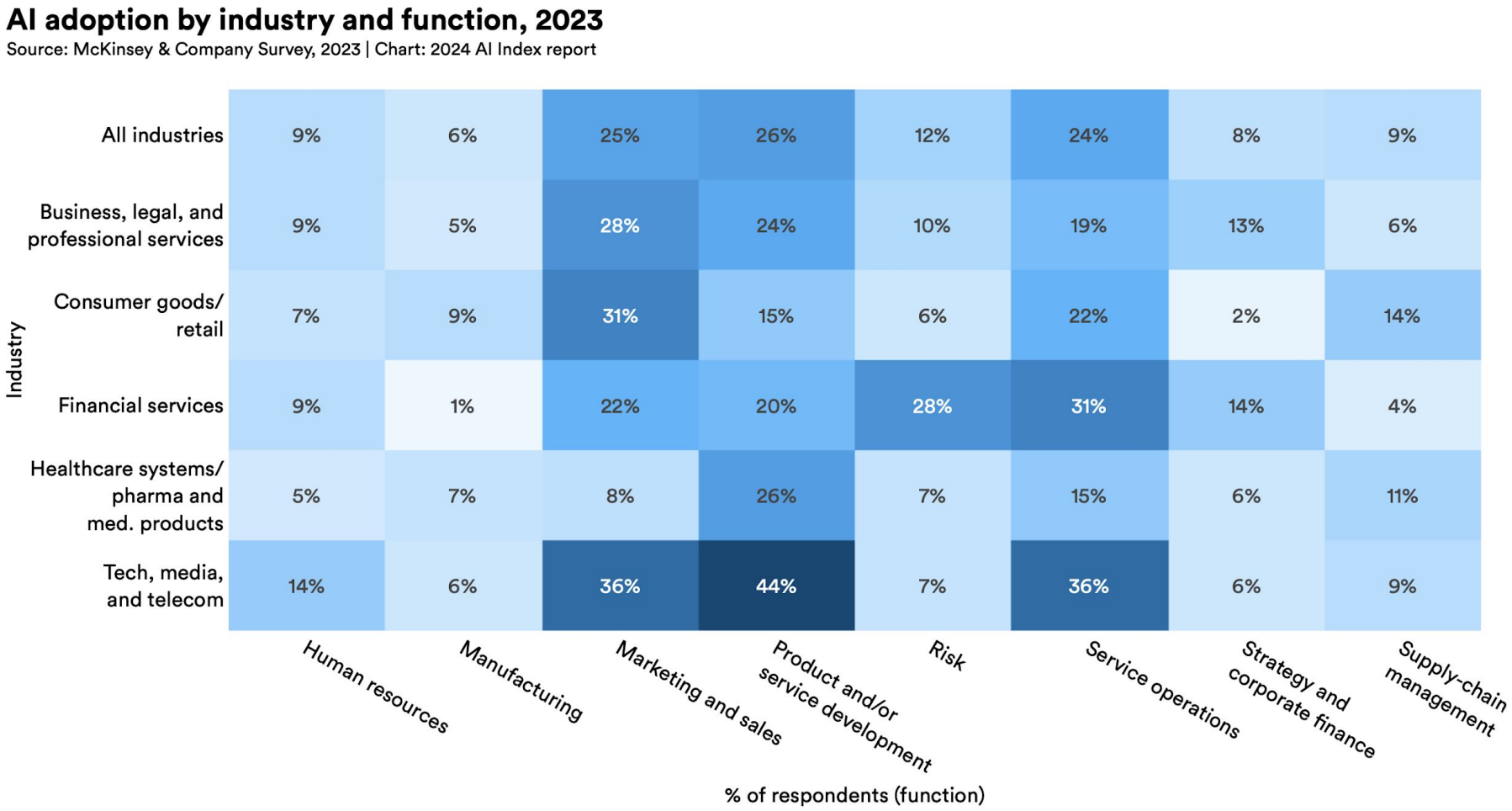
ILT represents +35% of the drilling time

Real-time prescriptive solution that connecting to any field service software

13

Copyright © 2025 by Boston Consulting Group. All rights reserved.

Most AI Use Across Society Is in Non-Climate Applications



So Far, AI-for-Climate is Rather Traditional AI

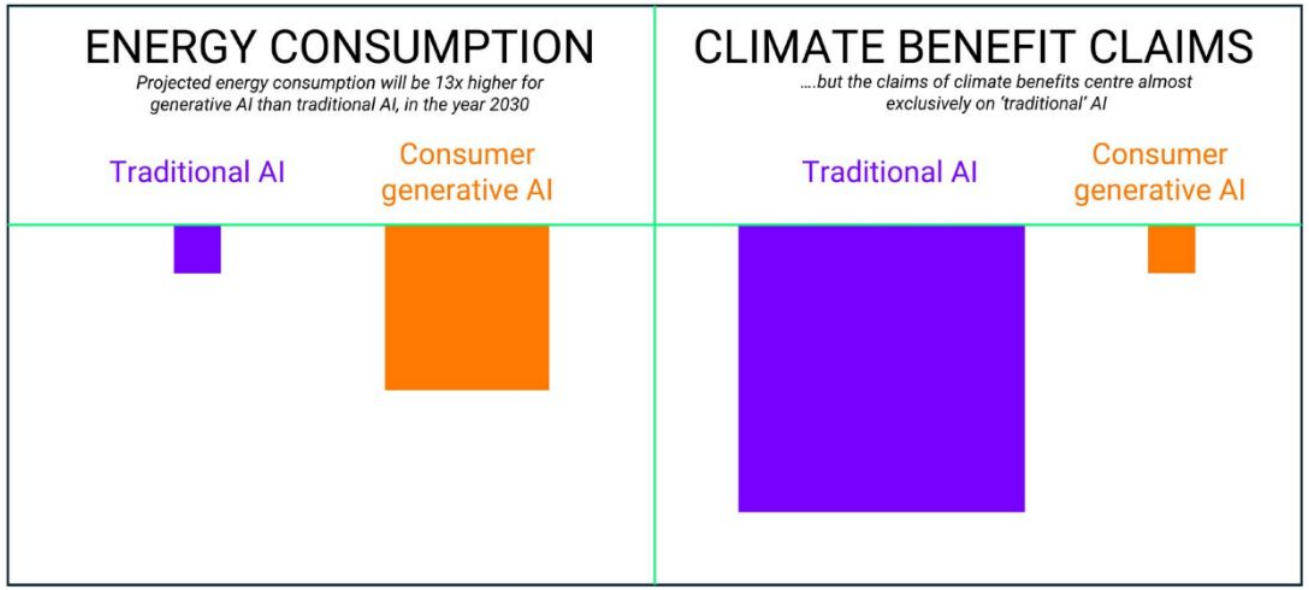


FIGURE 1 - CLIMATE BENEFIT versus DAMAGE
This graphic illustrates how most climate benefit claims relate to traditional AI, whereas most harm will relate to consumer generative AI. The first compares the projected electricity consumption in a central 2030 scenario, also represented by the area of each square. The second comparison shows the number of claims by AI type, represented by the area of each square.

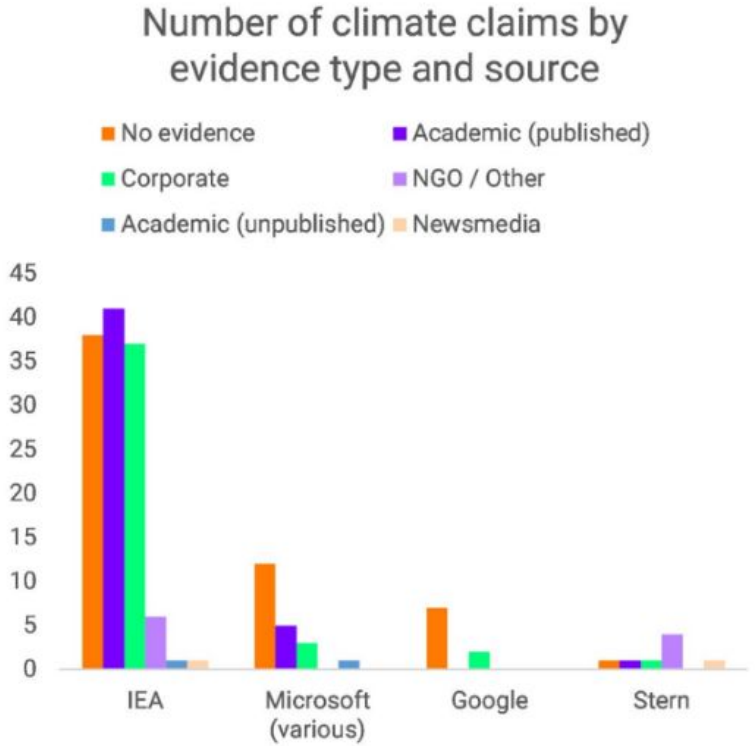


FIGURE 7 - TYPES OF EVIDENCE FOR CLIMATE CLAIMS
This graphic shows the distribution of evidence types by source. Corporate sources tended to rely on no evidence or corporate sites. The IEA's claims drew from a mix of academic, corporate and no evidence.

Images: Ketan Joshi, The AI Climate Hoax: Behind the Curtain of How Big Tech Greenwashes Impacts, 2026

Lecture Outline

Computing-related impacts

Application impacts

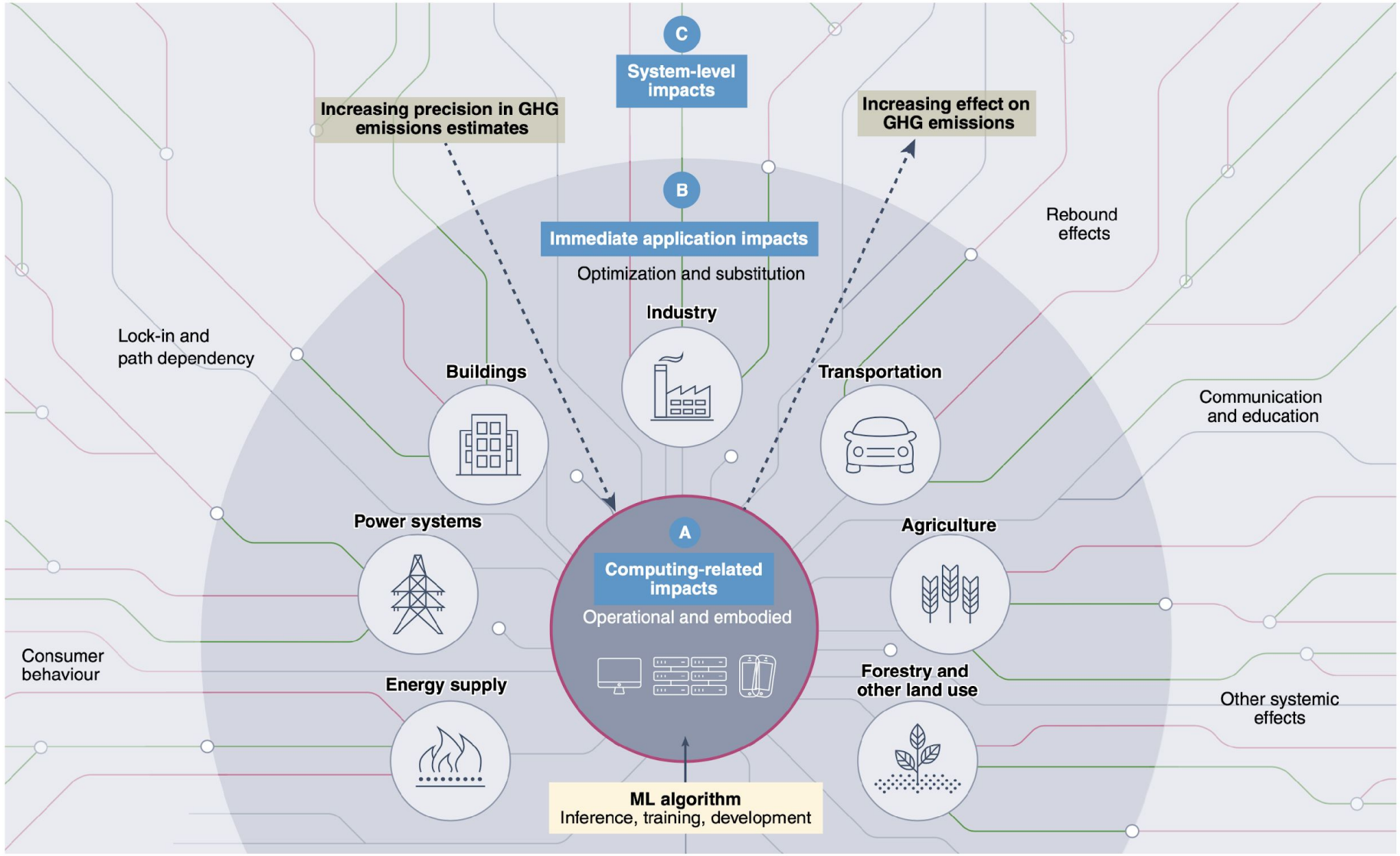
System level impacts

Overarching considerations

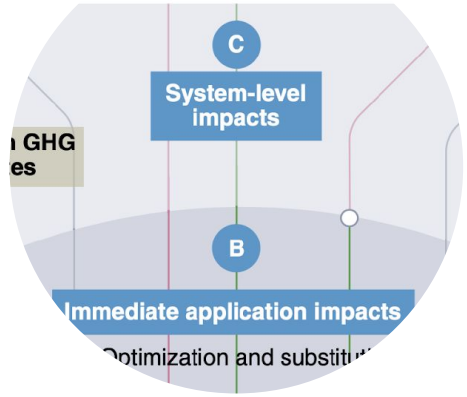
Regulating energy and resource usage of AI

- Rebound effects
- Societal power shifts

How AI Affects Greenhouse Gas Emissions



How AI Affects Greenhouse Gas Emissions



System-Level Impacts of AI Applications

Potential **rebound** and **lock-in** effects

- Autonomous vehicles, ridesharing



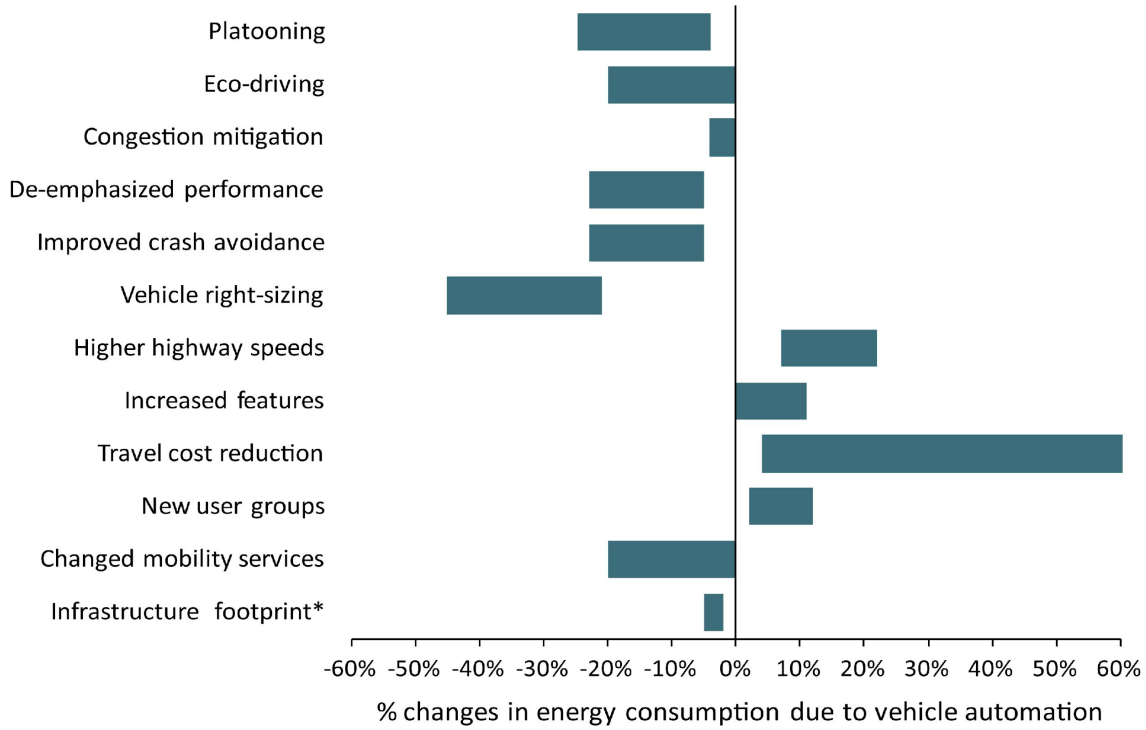


Image source: Wadud et al. 2016

System-Level Impacts of AI Applications

Potential **rebound** and **lock-in** effects

- Autonomous vehicles, ridesharing

Increasing **societal consumption**

- Personalized ads, on-demand delivery



Image credit: Megan_Rexazin_Conde / Pixabay.com

System-Level Impacts of AI Applications

Potential **rebound** and **lock-in** effects

- Autonomous vehicles, ridesharing

Increasing **societal consumption**

- Personalized ads, on-demand delivery

(Mis)information and **polarization**

- Content personalization/amplification



Image credit: Rose Wong / for NBC News

System-Level Impacts of AI Applications

Potential **rebound** and **lock-in** effects
- Autonomous vehicles, ridesharing

Increasing **societal consumption**
- Personalized ads, on-demand delivery

(Mis)information and **polarization**
- Content personalization/amplification

Inducing **societal power shifts**
due to access and agency

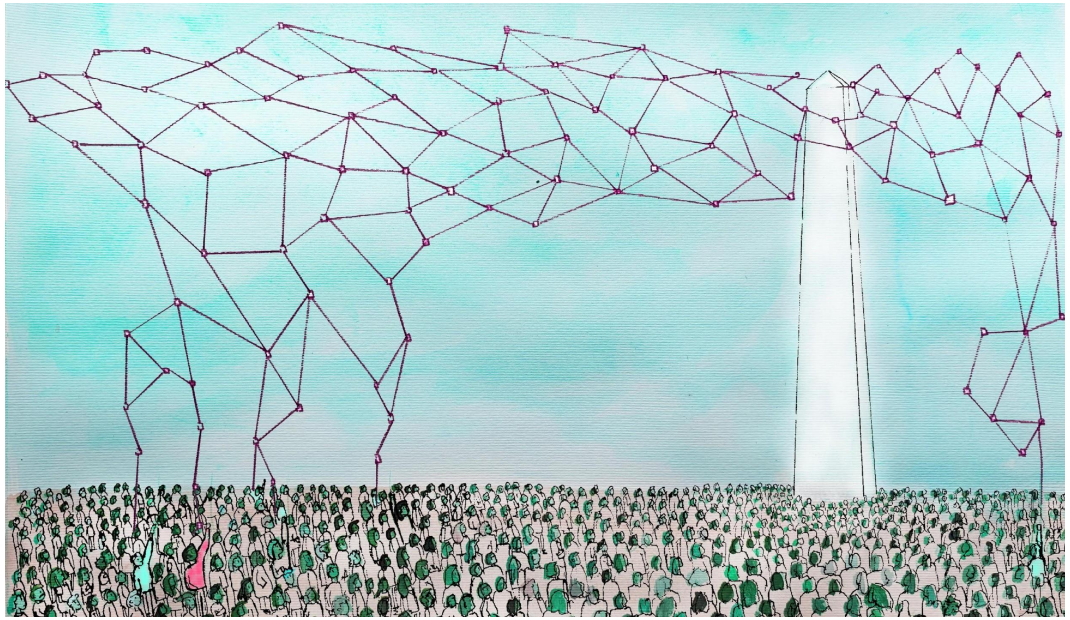


Image credit: Jamillah Knowles & We and AI / Better Images of AI / People and Ivory Tower AI / CC-BY 4.0

Poll Five

Share an example of a system-level impact of AI you believe has/will have a significant impact on climate, for one of the categories:

a) rebound effect, b) increased consumption,
c) mis-information and polarization, d) societal power shifts

Lecture outline

Computing-related impacts

Application impacts

System level impacts

Overarching considerations

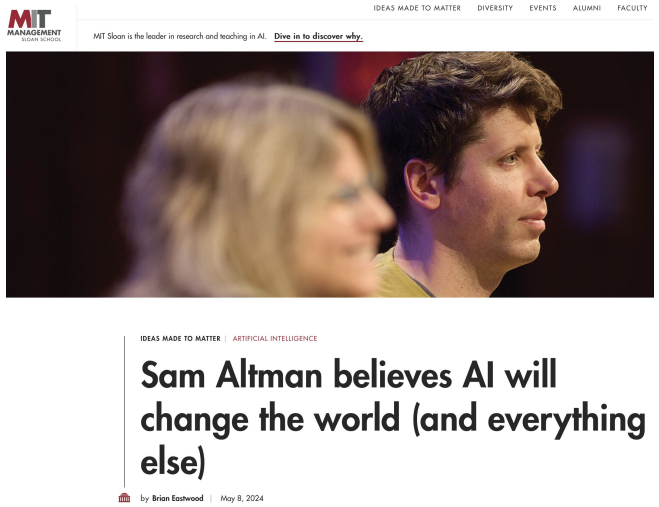
Regulating energy and resource usage of AI

Narratives

Estimates

The "GenAI Fallacy" : Using AI-for-Climate as Justification for Large Models

- To justify the energy consumption in datacenters for large models such as LLMs, AI-for-climate applications are a **welcome narrative** for tech companies
- But, most AI applied in climate-relevant sectors is not large models
- We can have AI-for-climate progress **without** large computing-related energy costs



"*If we have to spend even 1% of the world's electricity training powerful AI, and that AI helps us figure out how to get to non-carbon-based energy or to do carbon capture better, that would be a massive win*"

Misleading statement from Sam Altman

"Chronic **Potential** -itis"

- Concept from Vlad Coroamă: lot of discussion about potential…
- But there are significant barriers to AI deployment in climate-related domains
- Actors are presenting AI's savings potentials by constructing **best case adoption scenarios**
- Visualizations and summaries communicate **overconfidence** in reaching hypothetical estimates

Key messages from the 2025 Shell Energy Security Scenarios



1. In a world where artificial intelligence (AI) brings societal change and catalyses economic growth, demand for energy increases and the transition to low-carbon energy accelerates.

Source: https://www.shell.com/news-and-insights/scenarios/the-2025-energy-security-scenarios.html

Holistic Estimates?

Source	Type	Computing-related	Application-related (reductions)	Application-related (increases)	System-level
Stern et al. 2025	Journal article	Total data center increase	Power, meat and dairy, and light road vehicles	No	No
IEA 2025	Technical report	Total data center	Some energy applications	No	Only rebound
PWC 2025	Consulting report	Total data center increase	Economy-wide	No	No
Schneider Electric 2024	Technical report	AI	No	No	No
de Vries 2023	Journal article	AI	No	No	No

No application-related increases & no system-level effects assessed

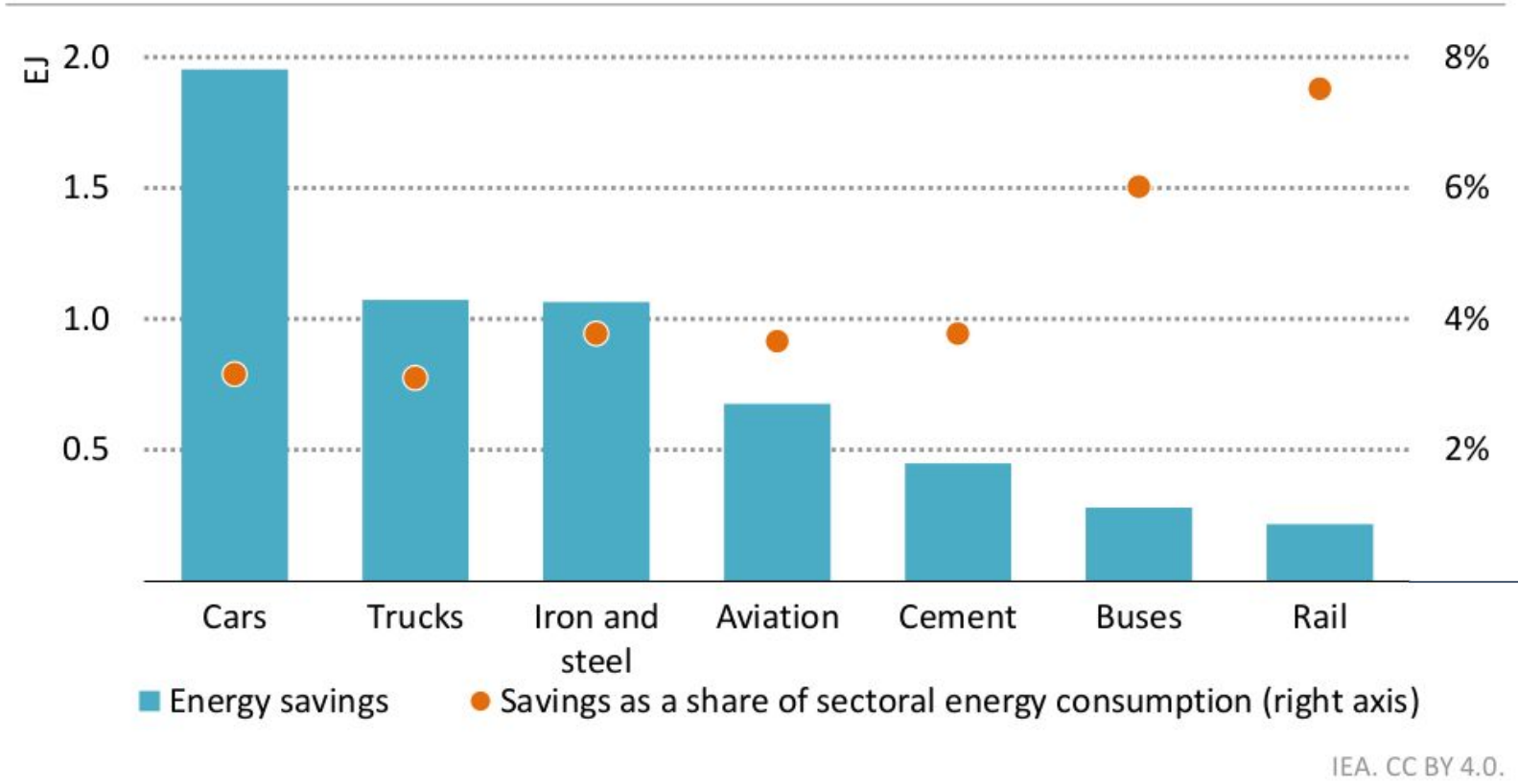
Danger of Fallacies in (Lack of) Comparisons Between Positive & Negative Impacts of AI

"*is AI's energy use worth the climate benefits*?"
is an **ill-posed** question

potential energy savings in certain sectors

≠ net climate benefit

≠ a major global accelerant in achieving global climate goals.



Widespread adoption of known AI applications has the potential to save 13.5 EJ across all end-use sectors in 2035, greater than Indonesia's annual energy consumption today

Note: This energy savings potential is estimated using existing AI applications only and does not factor in solutions that might emerge in the future. A detailed discussion of this approach is available in IEA (2025a).

Source: IEA 2026, Energy and AI

Widespread adoption of AI across industries

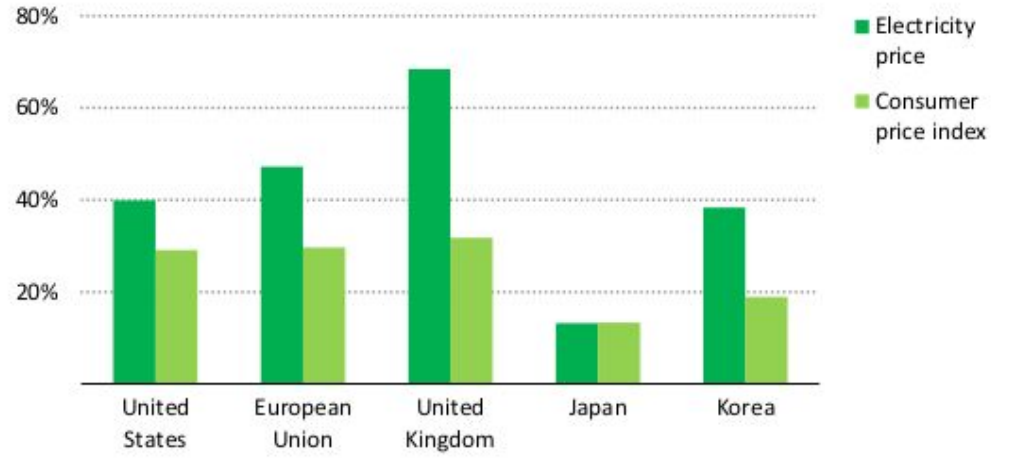
Greening the grid is not sufficient

- Unbridled AI growth is not something that can be handled by the Planet
- has social impacts already

The path to net zero emissions is narrow: staying on it requires immediate and massive deployment of all available clean and efficient energy technologies. In the net zero emissions pathway presented in this report, the world economy in 2030 is some 40% larger than today but uses 7% less energy. A major worldwide push to increase energy efficiency is an essential part of these efforts, resulting in the annual rate of energy intensity improvements averaging 4% to 2030 – about three-times the average rate achieved over the last two decades.

Source: IEA, "Net Zero by 2050: A Roadmap for the Global Energy Sector" (2021)

Figure 8.1 ▷ Change in nominal residential electricity price and consumer price index by country or region, January 2019-December 2025



Residential electricity prices have risen faster than headline inflation in several markets in recent years

Source: IEA, "Energy and AI" 2026

Lecture outline

Computing-related impacts

Application impacts

System level impacts

Overarching considerations

Regulating energy and resource usage of AI

Fostering AI-for-climate

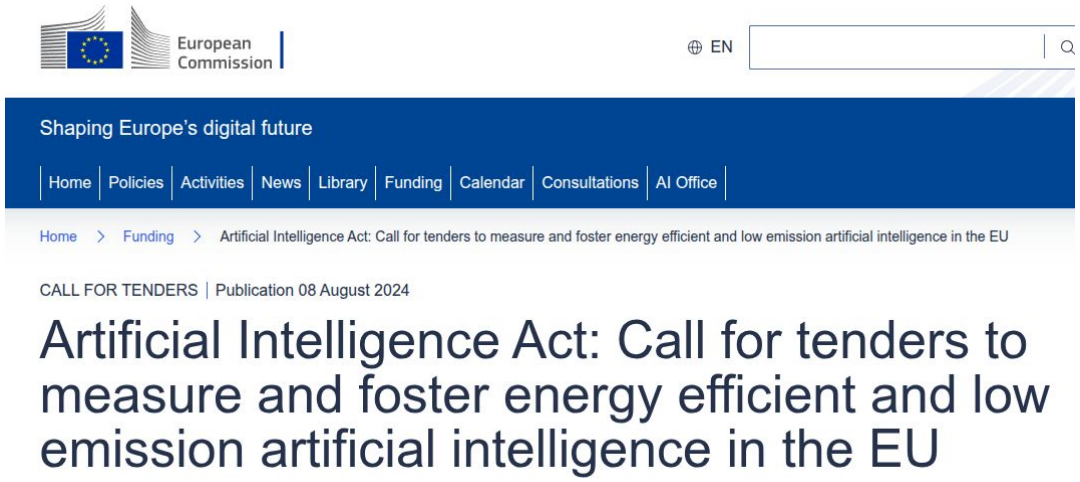
Requiring transparency

Callback: AI & Climate Policy and Regulation Lectures

TODO: add summary slides if i can get access to the other decks

Existing initiatives to **FOSTER APPLICATIONS**

- Dedicated governmental research and development funding (EU, Germany, Austria, Sweden, US, etc.)
- Large research and monitoring programs (Copernicus, Nasa Harvest, etc.)
- Some AI strategies and EU AI Act (2024) emphasize climate change as application area



Incentivizing

Frugal applications

afnor NORMALISATION

General framework for frugal AI

Measuring and reducing the environmental impact of AI

IA et environnement : 12 territoires sélectionnés pour concevoir l'IA frugale au service de la transition écologique

Publié le 13 septembre 2024 , par Olivier Devillers, pour Localtis

Environnement, Numérique

La Banque des Territoires a accueilli le 10 septembre 2024 les 8 nouveaux lauréats de l'appel à projets démonstrateur d'IA frugale pour la transition écologique des territoires, rejoignant les 4 territoires déjà labellisés. L'occasion de découvrir, avec des exemples variés et concrets, la notion de frugalité appliquée à l'intelligence artificielle.



Existing Initiatives Requiring Transparency

- The EU AI Act (2024)
  - Reporting on the compute resources and energy use of high-risk systems and large models

- Proposed US Artificial Intelligence Environmental Impacts Act (2024)
  - Require environmental impacts assessments (by the EPA and NIST)
  - Application- and computing-related impacts



Official Journal of the European Union — EN, L series

2024/1689 — 12.7.2024

REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 13 June 2024

laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

(Text with EEA relevance)

Annex IV: Technical Documentation
[for high-risk AI systems]
(c) […] the computational resources used to develop, train, test and validate the AI system;

Annex XI: Section 1 – Information to be provided by all providers of general-purpose AI models
(d) the computational resources used to train the model (e.g. number of floating point operations), training time, and other relevant details related to the training;
(e) known or estimated energy consumption of the model. With regard to point (e), where the energy consumption of the model is unknown, the energy consumption may be based on information about computational resources used.

Poll Six

The EU AI Act requires public disclosure for the energy use for general-purpose AI required for:

a) training, b) inference, c) both, d) none

Omissions matter

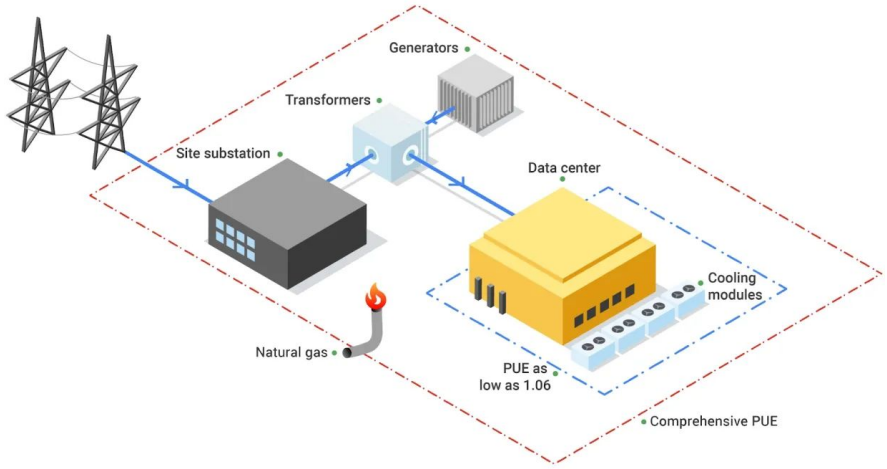


Example in the **EU AI act**:

- For **general-purpose AI**: significant gap as it only covers the energy used during the model's development phase, but **leaves out the inference phase**
- Where the Act does mandate the disclosure of energy consumption, this **information is restricted to authorities and is not accessible to downstream providers →** significantly reduces transparency and accountability, thereby weakening the potential for public oversight and market responses.
- AI Act fails to address the greenhouse gas (GHG) emissions generated **by** AI applications
- Water consumption not addressed
- Environment can in theory be regarded as a criterion for high-risk systems but it is not made explicit.

Regulating Data Centers

- Energy Efficiency Directive EU/2023/1791 ("EED") and the Commission Delegated Regulation EU/2024/1364
  - applies to all data centers in the EU with a power demand of at least 500kW, which includes small-sized data centers.
  - Data center operators are required to collect, make publicly available and report to a EU database information. The reporting is mandated on an annual basis.
  - The required data includes energy consumption, power utilization, temperature set points, waste heat utilisation, water usage and use of renewable energy



Power Usage Effectiveness
(Image: Google)

- German 2023 Energy Efficiency Act
  - requires data center operators to inform their customers on an annual basis on the energy consumption directly attributable to them

Need for Integrated Oversight for Datacenter Sitting

- Transform data-infrastructure governance from a purely energy-water challenge into a lever for equitable territorial planning and people's centred policy vision.
- Land use interaction: Siting has various implications given interaction with local zoning, property valuation, and ecosystem restoration priorities
- Ongoing discussion on compensation mechanisms, and community benefit frameworks needed
- **→ Need for integrated oversight across energy, environment, and digital-economy policy vision.**





Summary

📝 Sustainability impacts of AI span the whole AI lifecycle and include various impact categories

→ Look at computing, applications and system-level impacts

📝 Impacts of large models need to be better addressed via evidence-based action

📝 Yes to frugal AI-for-climate applications!

📝Estimates remain uncertain and outlooks/narratives should be cautious

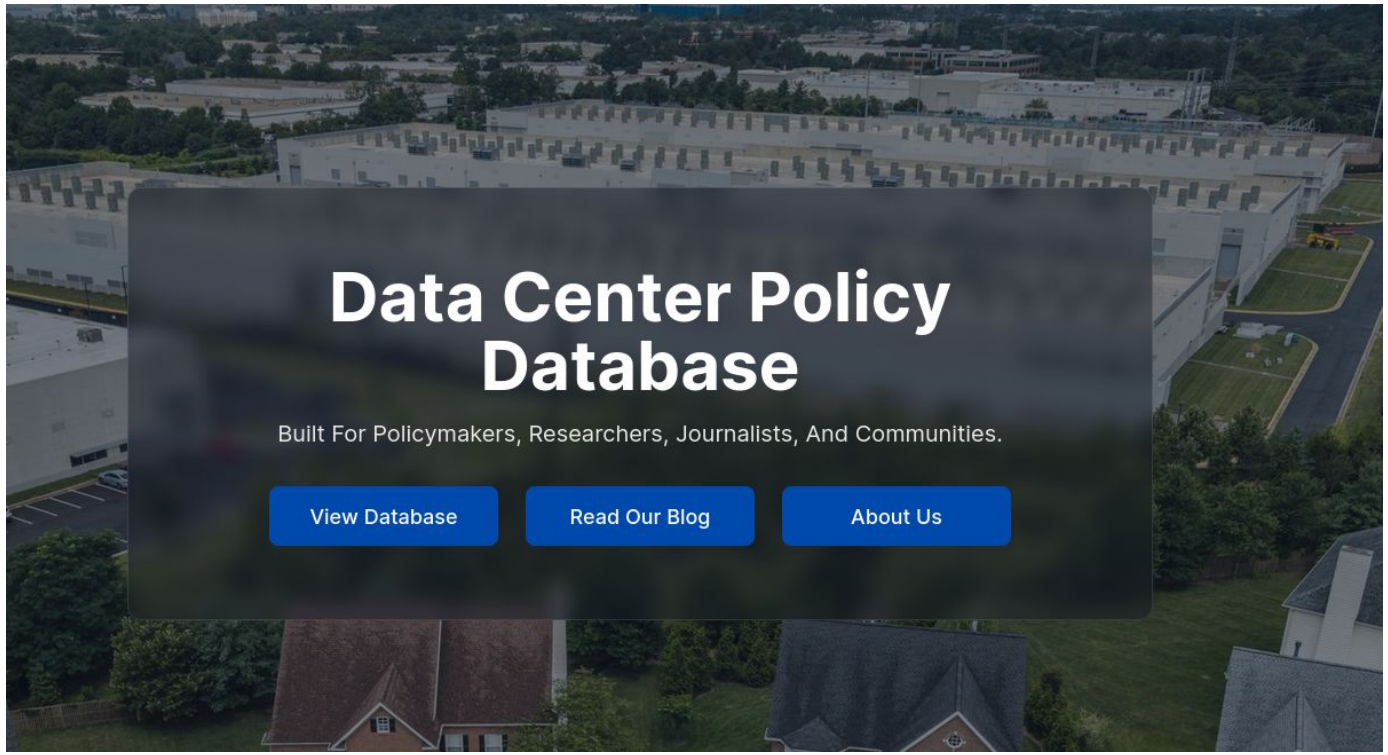
📝 Regulating AI, beyond market incentives, can lead to further transparency, which is key to public debate

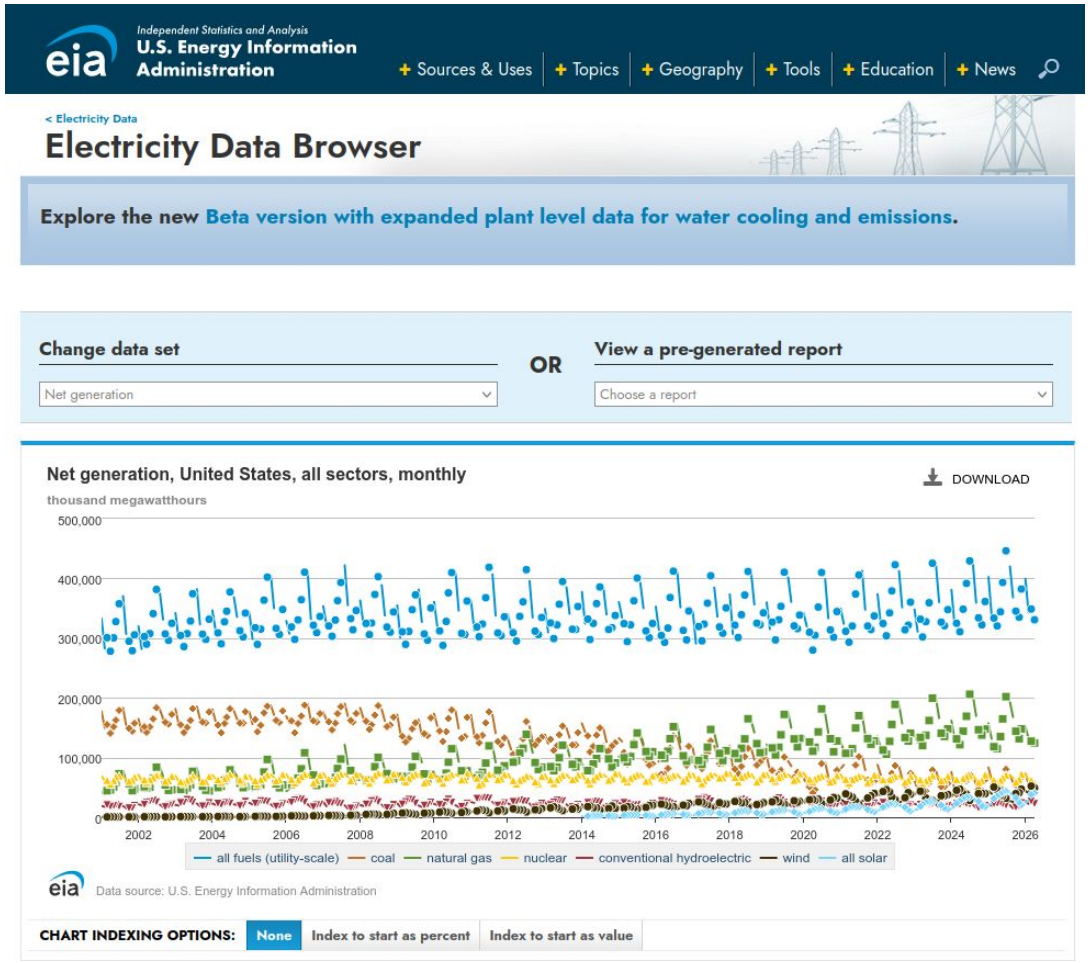
More Resources



https://www.sustainableaicoalition.org/inventory-of-key-reports-on-sustainable-ai/

DATA


Data Center Policy Database
Built For Policymakers, Researchers, Journalists, And Communities.
View Database
Read Our Blog
About Us


eia
Independent Statistics and Analysis
U.S. Energy Information Administration
+ Sources & Uses
+ Topics
+ Geography
+ Tools
+ Education
+ News
< Electricity Data
Electricity Data Browser
Explore the new Beta version with expanded plant level data for water cooling and emissions.
Change data set
Net generation
OR
View a pre-generated report
Choose a report
Net generation, United States, all sectors, monthly
DOWNLOAD
thousand megawatthours
500,000
400,000
300,000
200,000
100,000
0
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
2026
all fuels (utility-scale)
coal
natural gas
nuclear
conventional hydroelectric
wind
all solar
Data source: U.S. Energy Information Administration
CHART INDEXING OPTIONS:
None
Index to start as percent
Index to start as value

Q&A

Quiz One

Is the energy use for training or inference most important for the main commercial large language models for

a) Training, b) Inference?

Quiz Two

Are larger models always performing better than smaller models in terms of results for climate-relevant tasks?

a) True, b) False

Quiz Three

Do we have robust scientific estimates about the *future* net impact of AI on climate?

a) True, b) False

Quiz Four

Qualitatively,

is the *current* net impact of AI on climate:

a) Negative, b) Positive?

Quiz Five

Is using more renewable energy the solution to the increasing electricity demand of AI?

a) True, b) False

Pre-Readings

- Aligning artificial intelligence with climate change mitigation, Kaack et al 2022
  https://www.nature.com/articles/s41558-022-01377-7
- The AI Climate Hoax: Behind the Curtain of How Big Tech Greenwashes Impacts, Ketan Joshi, 2026
  https://drive.google.com/file/d/12l1W4W25b-_ff6yFNJABkfal9_9oevxe/view
- The Real Story on AI's Water Use–and How to Tackle It, Shaolei Ren & Amy Luers, 2025
  https://spectrum.ieee.org/ai-water-usage