Regulation of AI - Basics, Sustainability, Futures

Prof. Dr. Philipp Hacker, LL.M. (Yale)
Chair for Law and Ethics of the Digital Society
European New School of Digital Studies
European University Viadrina
Frankfurt (Oder)
Germany

Structure

I. Foundations
II. AI and Climate
III. The Future
IV. Summary

Overview based on

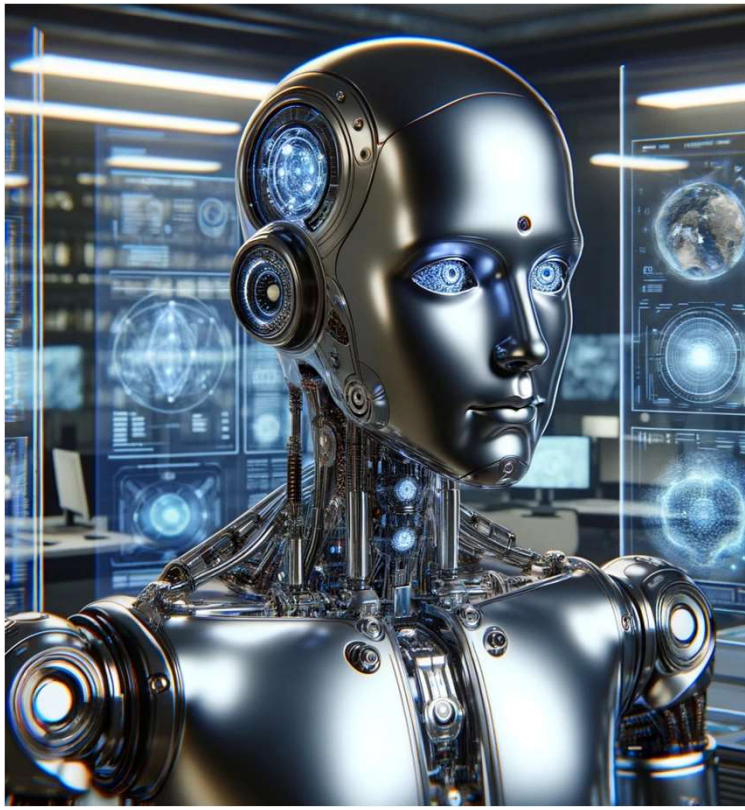
- Philipp Hacker, Andreas Engel, Marco Mauer, **Regulating ChatGPT and other Large Generative Models**, (2023) ACM Conference on Fairness, Accountability and Transparency (FAccT '23) 1112-1123
- Philipp Hacker, **The European AI Liability Directives**, (2023) 51 Computer Law & Security Review, Art. 105871
- Philipp Hacker, **Sustainable AI Regulation**, (2024) 61 Common Market Law Review 345
- Open access for all papers: http://arxiv.org/a/hacker_p_1

ens

Part I.
Foundations

I. AI Act: Definition of AI

- Art. 3(1): **machine-based** system
  - with varying degrees of **autonomy** and that may show adaptiveness after deployment,
  - that **infers**
  - from the **input** it receives,
  - how to generate **outputs**
    - ➢ such as predictions, content, recommendations or decisions,
  - that can **influence** physical or virtual environments.



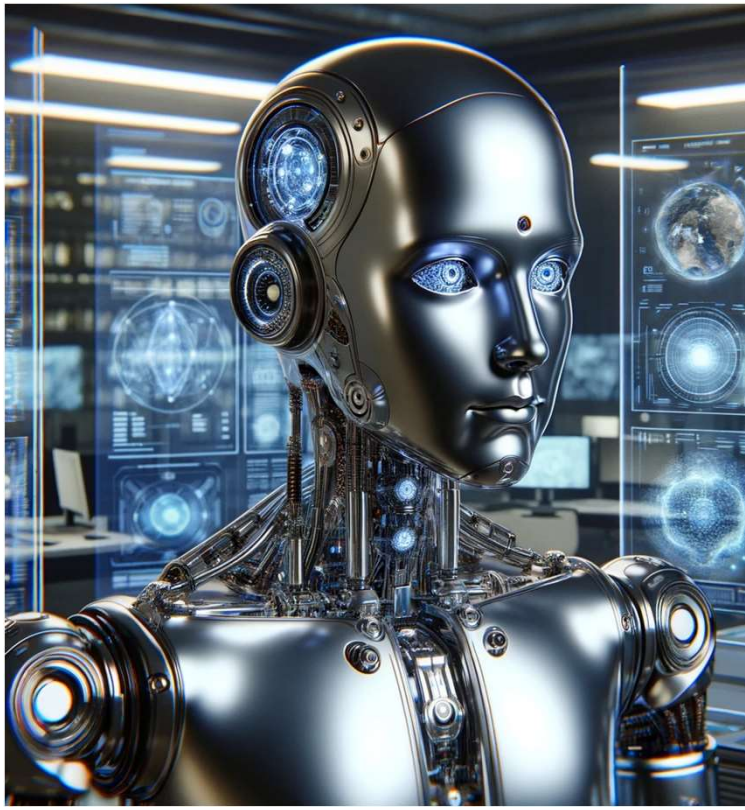
DALL-E 3, "picture of AI, photorealistic", 01/31/2024

I. AI Act: Definition of AI

Case: Excel Auto-sum?

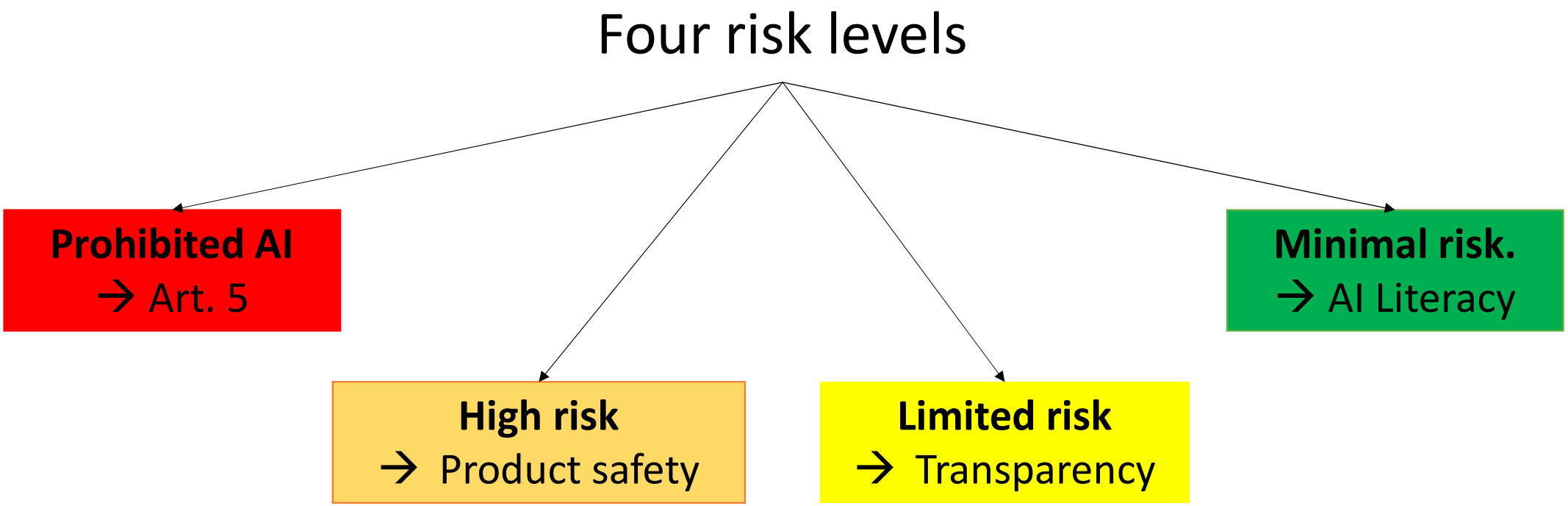
- **inference**: goes beyond **basic data processing** (Rec. 12)
- Not AI: Systems based on **rules** defined solely by **natural persons** for the automatic execution of operations (Rec. 12)

→ AI = machine learning, **broadly understood**



DALL-E 3, "picture of AI, photorealistic", 01/31/2024

Structure of the AI Act



General-purpose AI systems

AI Act: Limited Risk

Limited risk:

- Art. 52: **Disclosure** of the fact that AI is being used when AI interacts with humans
  - Also for deep fakes (Art. 52(3))
  - For **AI text production**: only for text to inform the public on matters of public interest (Art. 52(3))
    - ➢ Exception: editorial review and assumption of responsibility
- **Labeling** of AI-generated content (e.g. watermarks)

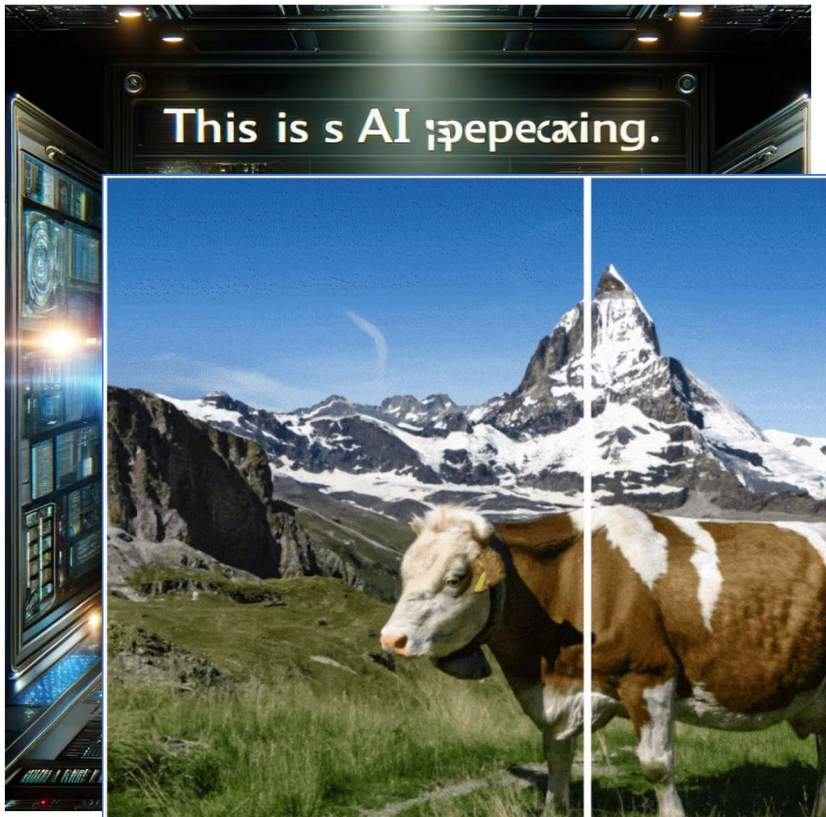
DALL-E 3, "This is AI speaking", 11/13/23

Melissa Heikkilä, Google DeepMind has launched a watermarking tool for AI-generated images, MIT Technology Review (Aug. 29, 2023)

AI Act: Limited Risk

Limited risk:

- Art. 52: **Disclosure** of the fact that AI is being used when AI interacts with humans
  - Also for deep fakes (Art. 52(3))
  - For **AI text production**: only for text to inform the public on matters of public interest (Art. 52(3))
    - ➢ Exception: editorial review and assumption of responsibility
- **Labeling** of AI-generated content (e.g. watermarks)



DALL-E 3, "This is AI speaking", 11/13/23

Melissa Heikkilä, Google DeepMind has launched a watermarking tool for AI-generated images, MIT Technology Review (Aug. 29, 2023)

High-risk AI systems

(Annexes II A, III AI Act):

- Facial recognition/biometrics
- Education
- Employment
- Medical AI
- Creditworthiness check
- Insurance (life, health)
- Social benefits, migration, asylum
- Judiciary
- Elections

Not included:

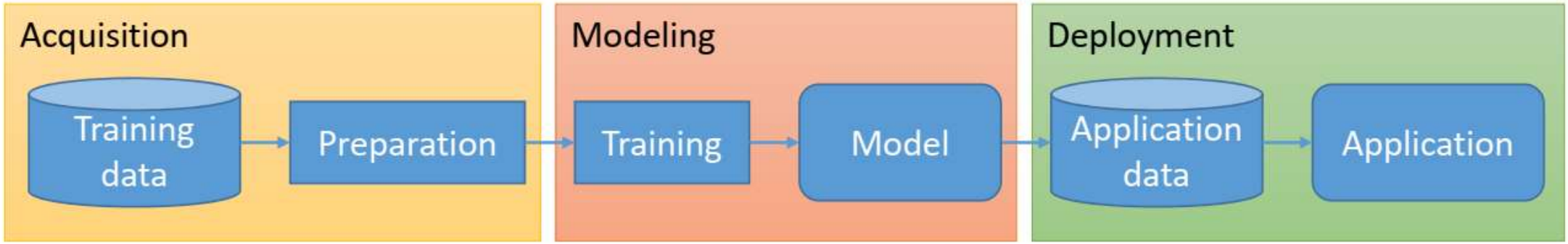
- E-Commerce
- Search engines

→ Digital Markets Act
Markets Act (DMA)

Important rules for high-risk AI

ML pipeline:

- Art. 11-13 AEOI: **Transparency & doc.**
- Art. 14 AIA: **human in the loop**



- Art. 10 AIA: Training data
  - **Correctness**
  - **Representativeness**
  - **Minimizing bias**
- Art. 15 AIA: Performance
  - **Accuracy**
  - Robustness
  - **IT security**

Risk and **quality management**

FMs - The tiered approach

All Foundation models, Art. 52c
- Transparency & copyright

A I S A F E T Y

FMs with system. Risks
In addition:
- **Risk management**
- **Red Teaming**
- **Incident Reporting**
- **Cybersecurity**

Systemic risks:
Assumption: 10^25 FLOPs
→ approx. GPT-4
- Compute
- Data
- Effects

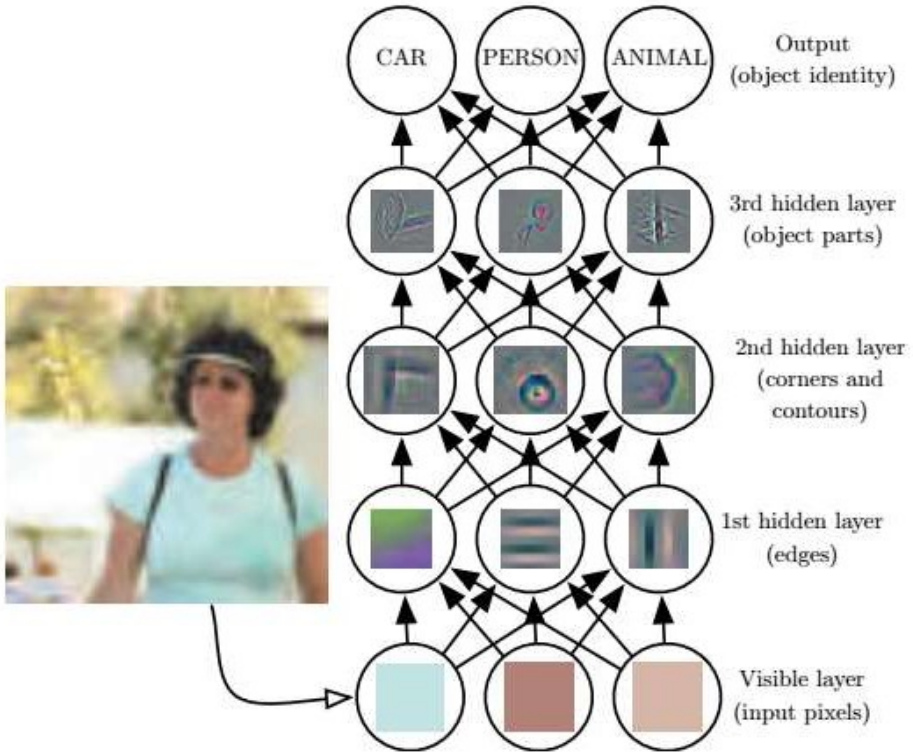
Fine-tuning as a significant change, Art. 28(1)?

Important rules for AI users

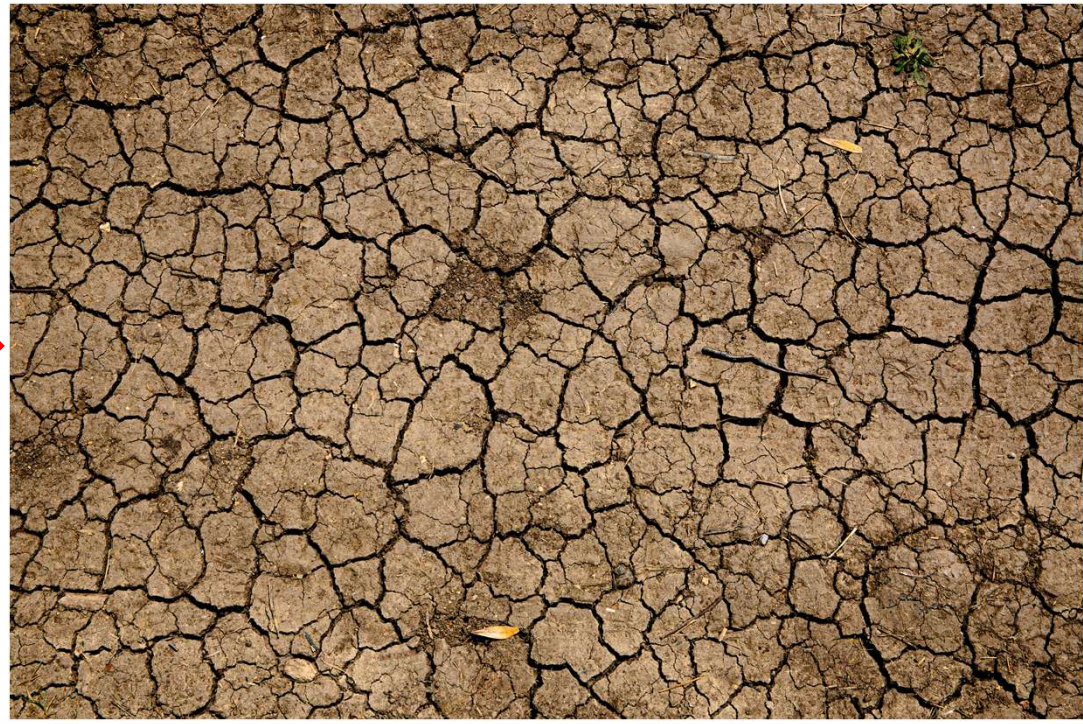
- Monitoring and supervision
- Data governance (input)
- Information and documentation
- Fundamental Rights Impact Assessment

Part II.
AI and Climate

The Next Frontier in AI Policy: Sustainability

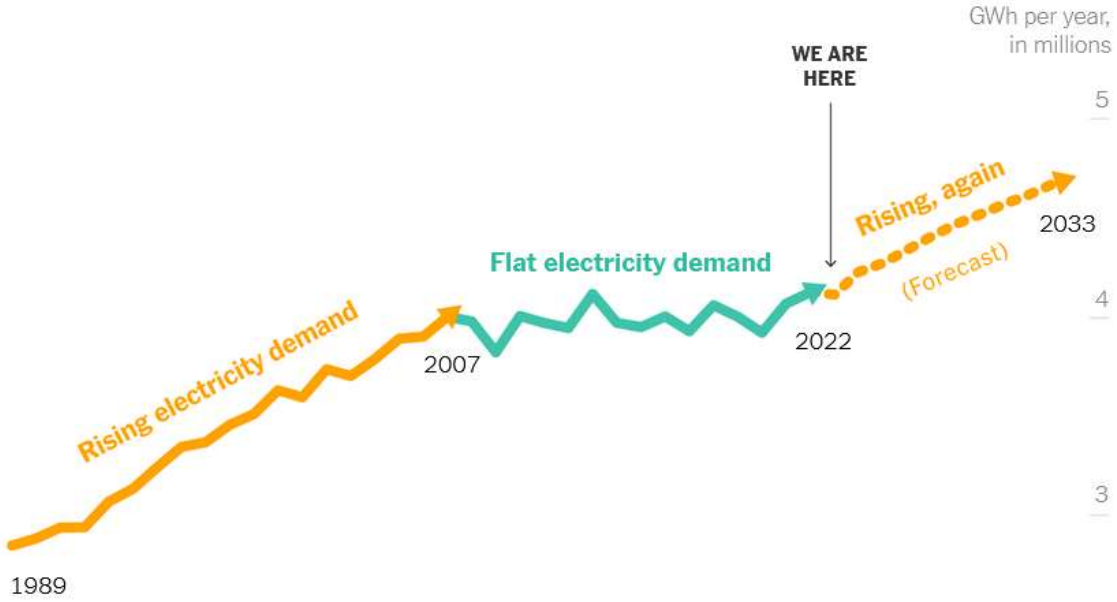


?



Goodfellow et al., Deep Learning, 2016, 6

The New York Times

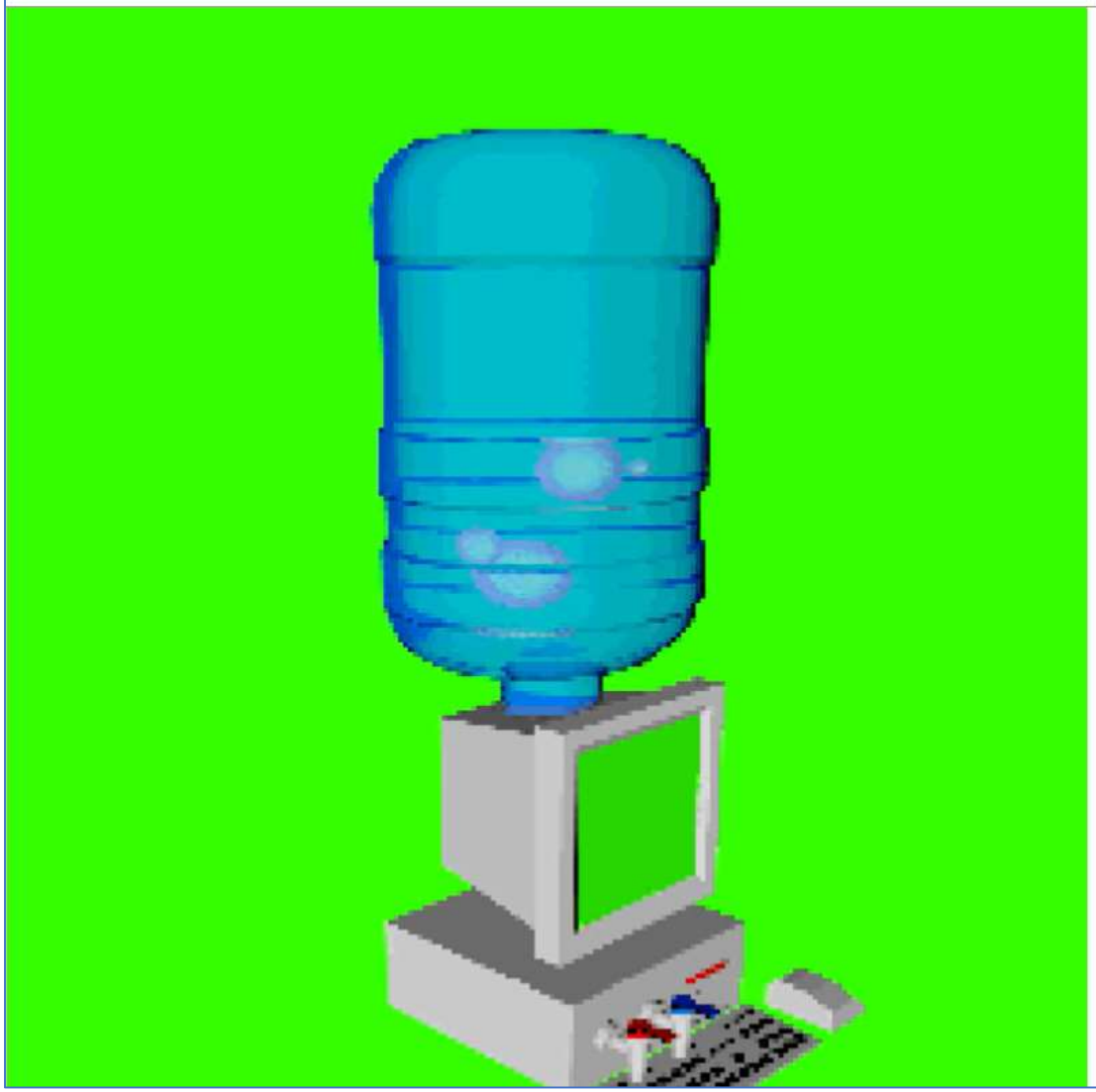


A New Surge in Power Use Is Threatening U.S. Climate Goals

A boom in data centers and factories is straining electric grids and propping up fossil fuels.

By **Brad Plumer** and **Nadja Popovich** March 14, 2024

The Atlantic



TECHNOLOGY

AI IS TAKING WATER FROM THE DESERT

New data centers are springing up every week. Can the Earth sustain them?

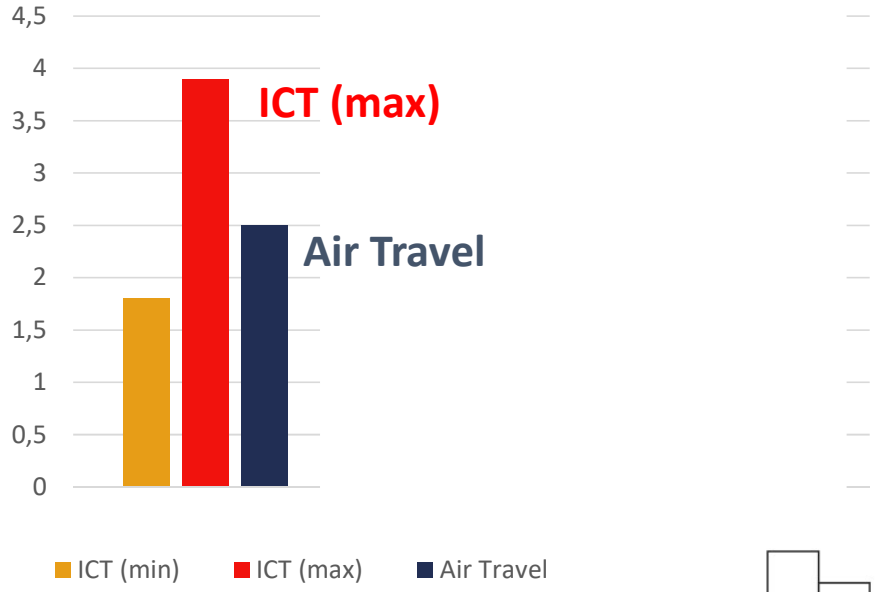
By Karen Hao

GHG Effects of ICT / AI

ICT / AI

- Global GHG emissions by ICT: **1.8 – 3.9 %** (Freitag et al., 2021)
- AI compute GHG costs: **skyrocketing** (Schwartz et al., 2020)
  - By 2027: GHG emissions of **Argentina** (de Vries, 2023)
  - One Stability XL image = one smart phone charge (Luccioni et al., 2024)
- GHG costs of AI applications
  - Ex.: oil & gas exploration (Kaack et al., 2022)
- Potential mitigation
  - But not GenAI

Air Travel

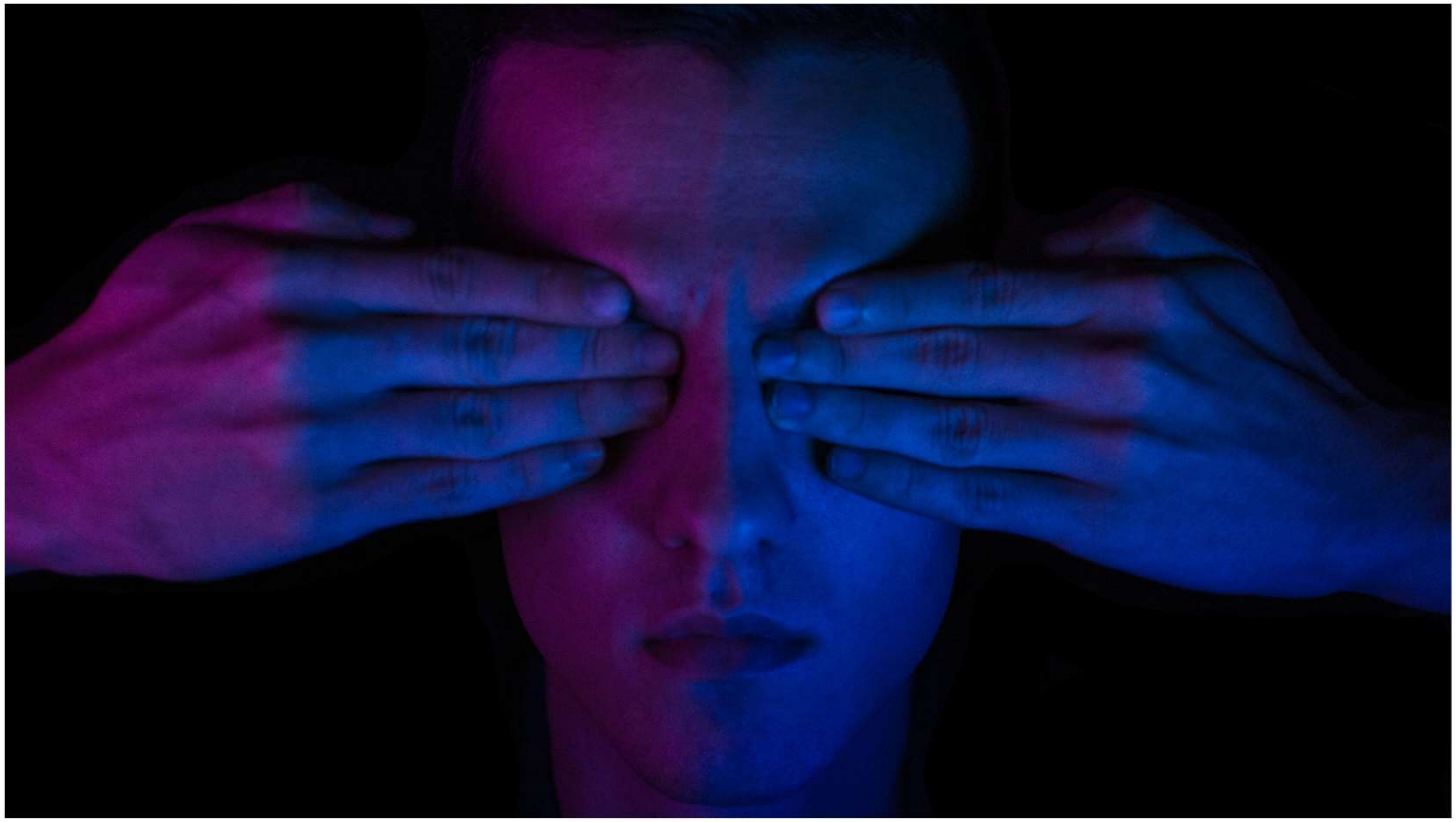


ens

More on this: Day 10 – Lynn Kaack: GHG Impact Assessment of AI - July 22, 2024

The Next Frontier: Sustainable AI

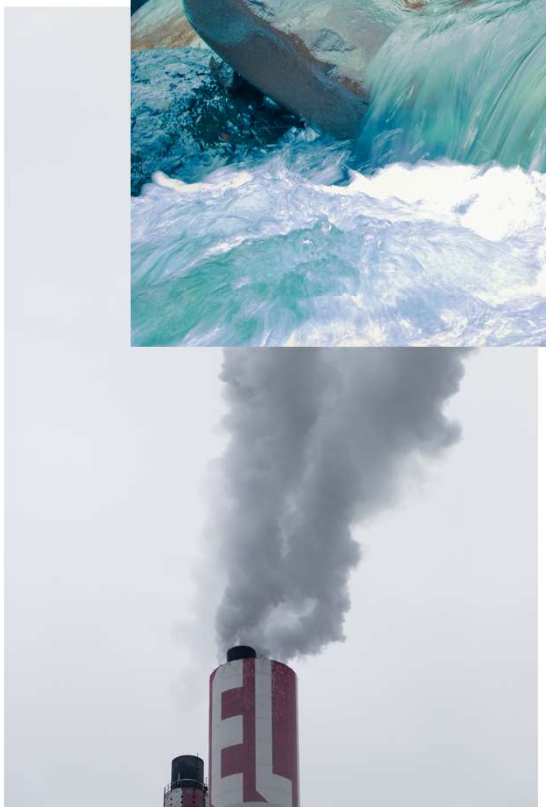
- **Blind spot** in AI regulation so far



Sustainable AI and Environmental Law

- Sustainable AI & Environmental Law:
  - ETS and EU WFD
  - AI not covered by ETS
    - ➢ Focus on traditional high-consumption sectors

→ Gaps in AI Regulation



What Does the EU AI Act Say?

- Applies to int‘l providers, too! (if they offer service in EU)
- **Transparency Requirements**:
  - Providers of foundation models

  → disclose **energy consumption**

    ➢ If actual energy consumption is unknown, estimates can be based on **computational resources** used.
  - Providers of high-risk AI system:

  → Disclose compute for development

What Does the EU AI Act Say?

- What‘s missing?
  - Open-source FMs
  - Non-high-risk AI systems
  - Inference

What Does the EU AI Act Say?

- **Assessment and Mitigation of Systemic Risks**:
  - Providers of **very large foundation models**, such as GPT-4,
  - and providers of **high-risk models**

  → mitigate risks to fundamental rights
    - ➢ Including environmental protection

An Emissions Trading Regime for AI

- Incentivizing GHG Emission Limits in AI:
  - ICT emissions comparable to commercial aviation
  - Gradually include AI and ICT in the EU Emissions Trading System

  → Set workable financial incentives for GHG reduction in AI

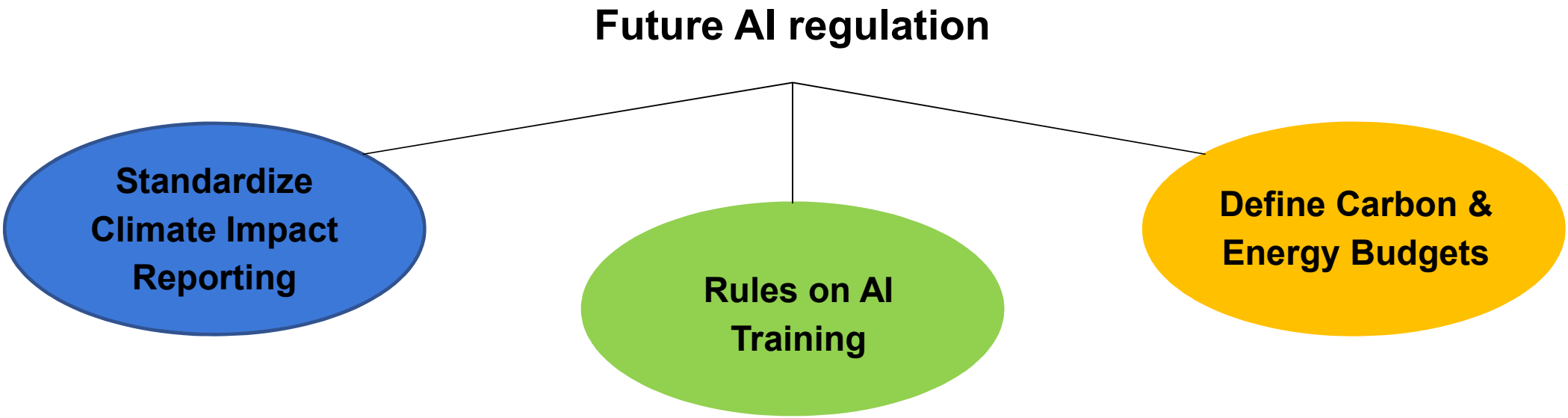
ETS

ETS ?

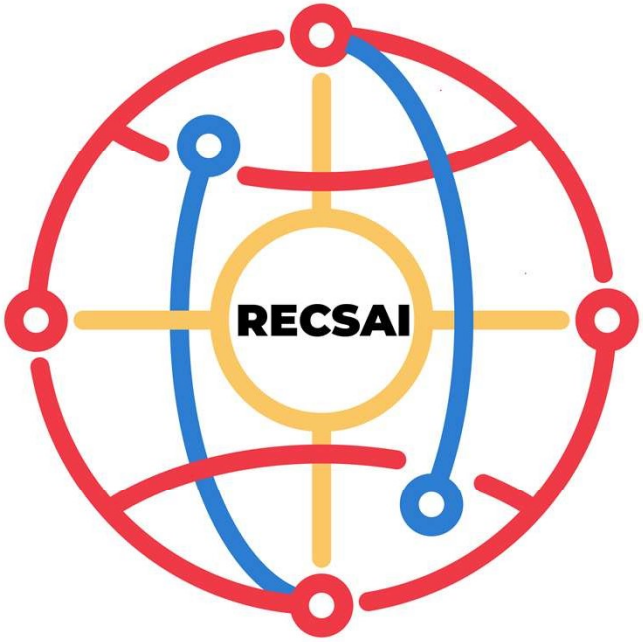
ens

Part III.
Future

Future Challenges



International dimension



International Expert
Consortium on AI

www.recsai.org

Part IV.
Conclusion

The future will be exciting.

Thanks!

hacker@europa-uni.de

LinkedIn:
https://www.linkedin.com/in/philipp-hacker-078940257/

ens