

Metrics for Responsible AI Projects

Data Intelligence Tokyo Meetup

Victor Steinborn

September 30, 2025

About me

- PhD in computer science
 - LMU Munich, Germany
 - Research focus: Responsible AI
- Industry Experience: Responsible AI Research and Software Development

Responsible AI

What is Responsible AI?

- Study of developing and deploying AI systems responsibly
- Some topics:
 - Fairness
 - Alignment
 - Explainability
 - Environment

Why Responsible AI?

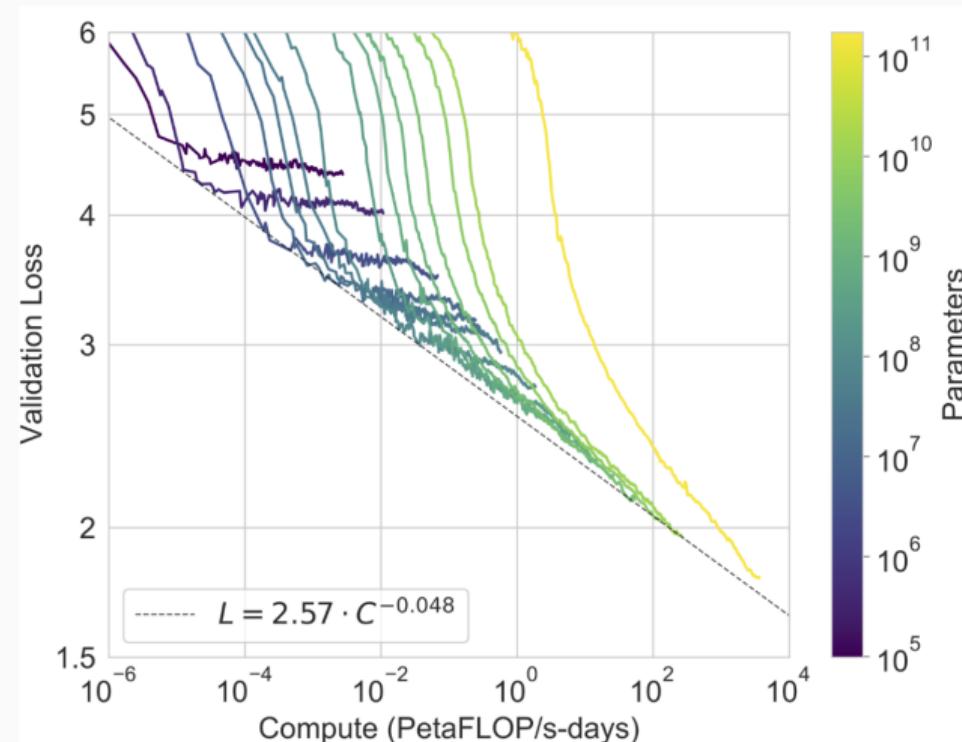
- An opportunity to think about social and environmental impact
- Framework to align metrics to (unforeseen) project constraints

ChatGPT Case Study

ChatGPT (Recap)

- ChatGPT first released November 2022 with GPT-3.5 [3]
- Focus on the predecessor model, GPT-3 [1]

GPT-3 Model Training [1]



\uparrow Parameters \rightarrow \downarrow Loss (\uparrow Training Performance) AND \uparrow Compute

Power

Definition of Power:

$$\mathcal{P} = \frac{dE}{dt}$$

- Training GPT-3 consumes about 1.287 *GWh* of energy [2]
- 1GWh is a huge amount of energy

How Much is 1 GWh? [5]

x100Mil



x1.3Mil



x294



x1.88Mil



For 1 Hour

Energy Consumption in AI Today [6]

≡ WIRED

SECURITY POLITICS THE BIG STORY BUSINESS SCIENCE CULTURE REVIEWS



ZOE SCHIFFER WILL KNIGHT LAUREN GOODE BUSINESS SEP 23, 2025 5:00 PM

OpenAI Teams Up With Oracle and SoftBank to Build 5 New Stargate Data Centers

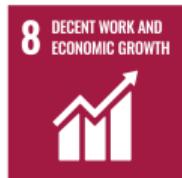
The new sites will boost Stargate's planned capacity to nearly 7 gigawatts—about equal to the output of seven large nuclear reactors.



What We Can Do

Guidelines [4]

SUSTAINABLE DEVELOPMENT GOALS



Questions?

Contact

Feel free to get in touch:



<https://vsteinborn.github.io/>

References i

-  T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei.

Language models are few-shot learners.

Jul 2020.

<http://arxiv.org/abs/2005.14165>.

References ii

-  A. S. Luccioni, S. Viguier, and A.-L. Ligozat.
Estimating the carbon footprint of BLOOM, a 176b parameter language model, Nov 2022.
<http://arxiv.org/abs/2211.02001>.
-  OpenAI.
Introducing ChatGPT, Nov 2022.
<https://openai.com/index/chatgpt/>.
-  United Nations.
United nations sustainable development goals, 2015.
<https://www.un.org/sustainabledevelopment/>.

References iii

-  US Department of Energy.
How much power is 1 gigawatt?, Aug 2024.
<https://www.energy.gov/eere/articles/how-much-power-1-gigawatt>.
-  Zoë Schiffer, Will Knight, and Lauren Goode.
OpenAI teams up with oracle and SoftBank to build 5 new stargate data centers, Sep 2025.
<https://www.wired.com/story/openai-oracle-softbank-data-center-stargate-us/>.