The Journal of Open Source Software

EcoLogits: Evaluating the Environmental Impacts of Generative AI

Samuel Rincé [1] **and Adrien Banse** [1,2]

1 GenAI Impact **2** ICTEAM, UCLouvain, Belgium

DOI: 10.21105/joss.07471

Software

- Review
- Repository
- Archive

Editor: Fei Tao

Reviewers:

- @jaywonchung
- @Zeitsperre

Submitted: 24 September 2024
Published: 09 July 2025

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

Recent advancements in *Generative AI (GenAI)*, particularly Large Language Models (LLMs), have led to significant increases in computational power and energy demands. These models are not only becoming larger and more complex but are also being widely deployed across various platforms, reaching millions of users and developers. This widespread use has escalated the overall energy consumption and environmental impact associated with these technologies.

To address these concerns, we introduce EcoLogits, an open-source Python package designed to evaluate the environmental footprint of using GenAI models. EcoLogits estimates energy consumption and multi-criteria environmental impacts of GenAI API requests using a bottom-up Life Cycle Assessment methodology, considering both direct energy use and hardware production impacts.

Statement of need

While the recent developments in GenAI offer tools of unprecedented quality, there is a general trend of increased energy consumption associated with these technologies (Vries, 2023). It is a well-known fact that training large neural networks yield non-negligible carbon emissions (Lacoste et al., 2019; Patterson et al., 2021); Some industry actors communicate about the emissions of training their own models, such as Meta that estimated carbon emissions of Llama models (AI@Meta, 2024; Touvron, Lavril, et al., 2023; Touvron, Martin, et al., 2023).

The major difference in recent years is that GenAI models are not only large models but are also widely used by millions of people through web application such as OpenAI's ChatGPT or via APIs. The literature agrees on the fact that the share of carbon emissions from the serving of GenAI models (inference) is increasing (Desislavov et al., 2023; Patterson et al., 2021), although these proportions vary among providers, with AWS reporting 80-90%, Google 60% and Meta 33% (Luccioni et al., 2024).

Moreover, while energy consumption is important, it alone is insufficient for a comprehensive evaluation of the environmental impacts of training and inference. It is crucial to include embodied impacts that come from hardware production. For LLMs such as BLOOM 176B, accounting for the embodied impacts increased the total carbon emissions by 22% during the training phase of the model (Luccioni et al., 2022). Environmental impacts evaluations of AI models should be more transparent and follow *Life Cycle Assessment (LCA)* methodologies to be complete (Ligozat et al., 2022).

In parallel to the large deployment of energy-consuming models in the industry, there is a growing need for tools that automatically approximate their environmental impacts. While tools like CodeCarbon or Zeus[1] (Courty et al., 2024; You et al., 2023) can track energy

[1]See https://github.com/mlco2/codecarbon and https://github.com/ml-energy/zeus.

consumption and carbon emissions for locally-deployed models, estimating impacts becomes necessary when using external APIs from GenAI service providers. EcoLogits is the first Python library to approximate the environmental impacts of GenAI models inference via Python client calls. By adding `EcoLogits.init()` to their codebase, users can access an estimation of the energy consumption and the environmental impacts of their requests.

Methodology

The **Attributional Life Cycle Assessment (A-LCA)** methodology defined by the ISO 14044 standard (Boustead, 1996; Hunt et al., 1996; Klöpffer, 1997) is used to properly estimate potential environmental impacts of products, processes, projects, or services. In EcoLogits, we apply this methodology to estimate the impacts of a request to a GenAI inference service. The methodology encompasses multi-phase and multi-criteria environmental impacts.

In our approach we consider the usage phase accounting for the direct energy consumption (Ilyas Moutawwakil, 2023a, 2023b) and the embodied phase accounting for the hardware production, including raw material extraction, manufacturing, and transportation (Simon et al., 2025). We do not include the end-of-life phase for lack of data and transparency on e-waste collection and recycling (Bordage et al., 2021; Forti et al., 2020). EcoLogits reports an estimation of the direct energy consumption in addition to other environmental impact criteria such as *Global Warming Potential* (Stocker et al., 2013), *Abiotic Depletion Potential* (Oers et al., 2020) and *Primary Energy Consumption* (Frischknecht et al., 2015). Assessing multiple impact criteria helps prevent pollution shifting that can result from optimizing for a single criterion.

To estimate the impacts of a GenAI service we use a **bottom-up modeling approach**, i.e. we individually assess the environmental impacts of components that frame the service, and then aggregate and allocate these impacts to a request. The scope of our methodology encompasses cloud compute primitives and data center extra equipment such as cooling (see Figure 1). We focus our study on high-performance GPU-accelerated cloud instances since they are commonly used for GenAI inference tasks. We do not account for other impacts related to model training, networking transfers, or end-user devices.

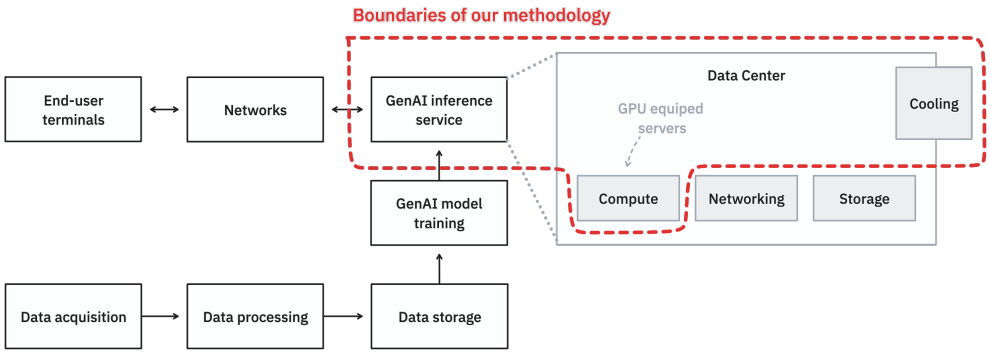

Figure 1: Scope and boundaries of our methodology focused on compute primitives

Discussions

Our approach for estimating the environmental impacts of GenAI inference aligns with industry standards and focuses on real-world deployments. As the field is constantly evolving, we are committed to refining our methodology to keep up with new model architectures, optimization techniques, and hardware innovations. Our tool enables GenAI practitioners to evaluate and

compare these models based on their environmental footprints. In the future, we plan to expand our scope to include more stages of the model life cycle and new modalities.

One challenge we face is dealing with proprietary models from companies that do not share details about their models and infrastructure. In these cases, we have to make educated guesses, such as estimating the number of parameters in models and which hardware is being used in production. We document these assumptions carefully and update them regularly as new models are released.

We wish to highlight deeper concerns about the impact assessment of GenAI. Our approach, while robust, does not capture the full range of environmental impacts associated with GenAI technologies. To get a complete picture, indirect impacts and potential rebound effects must also be considered. Pursuing "Green AI" means more than just making models smaller and more efficient; it requires a holistic view of the environmental changes driven by this technology, as well as its broader societal and service-level implications.

Acknowledgements

We extend our gratitude to Geneviève Fleury, Yoann Couble, and Caroline Jean-Pierre for their thorough review of this paper and the methodology. Additionally, we appreciate the contributions of Fabrice Gentot, Flavien Henrion, Luc Berton, Ryane Hamami, Sakina Ouisrani, Vinh Nguyen, and Andrea Leylavergne to this project. Their collaborative efforts have greatly supported our research.

References

AI@Meta. (2024). *Llama 3 model card*. https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md

Bordage, F., Montenay, L. de, Benqassem, S., Delmas-Orgelet, J., Domon, F., Prunel, D., Vateau, C., & Lees Perasso, E. (2021). *Digital technologies in Europe: an environmental life cycle approach*. https://www.greenit.fr/wp-content/uploads/2021/12/EU-Study-LCA-7-DEC-EN.pdf

Boustead, I. (1996). LCA — how it came about: — the beginning in the U.K. *The International Journal of Life Cycle Assessment*, *1*(3), 147–150. https://doi.org/10.1007/BF02978943

Courty, B., Schmidt, V., Luccioni, S., Goyal-Kamal, MarionCoutarel, Feld, B., Lecourt, J., LiamConnell, Saboni, A., Inimaz, supatomic, Léval, M., Blanche, L., Cruveiller, A., ouminasara, Zhao, F., Joshi, A., Bogroff, A., Lavoreille, H. de, … MinervaBooks. (2024). *mlco2/codecarbon: v2.4.1* (Version v2.4.1). Zenodo. https://doi.org/10.5281/zenodo.11171501

Desislavov, R., Martínez-Plumed, F., & Hernández-Orallo, J. (2023). Trends in AI inference energy consumption: Beyond the performance-vs-parameter laws of deep learning. *Sustainable Computing: Informatics and Systems*, *38*, 100857. https://doi.org/10.1016/j.suscom.2023.100857

Forti, V., Baldé, C., Kuehr, R., & Bel, G. (2020). *The global e-waste monitor 2020. Quantities, flows, and the circular economy potential*. United Nations University (UNU)/United Nations Institute for Training; Research (UNITAR) – co-hosted SCYCLE Programme, International Telecommunication Union (ITU) & International Solid Waste Association (ISWA). ISBN: 978-92-808-9114-0

Frischknecht, R., Wyss, F., Büsser Knöpfel, S., Lützkendorf, T., & Balouktsi, M. (2015). Cumulative energy demand in LCA: The energy harvested approach. *The International Journal of Life Cycle Assessment*, *20*(7), 957–969. https://doi.org/10.1007/s11367-015-0897-4

Hunt, R. G., Franklin, W. E., & Hunt, R. G. (1996). LCA — how it came about: — personal reflections on the origin and the development of LCA in the USA. *The International Journal of Life Cycle Assessment*, *1*(1), 4–7. https://doi.org/10.1007/BF02978624

Ilyas Moutawwakil, R. P. (2023a). *LLM-perf leaderboard*. https://huggingface.co/spaces/optimum/llm-perf-leaderboard; Hugging Face.

Ilyas Moutawwakil, R. P. (2023b). *Optimum-benchmark: A framework for benchmarking the performance of transformers models with different hardwares, backends and optimizations*. Hugging Face.

Klöpffer, W. (1997). Life cycle assessment. *Environmental Science and Pollution Research*, *4*(4), 223–228. https://doi.org/10.1007/BF02986351

Lacoste, A., Luccioni, A., Schmidt, V., & Dandres, T. (2019). *Quantifying the Carbon Emissions of Machine Learning*. arXiv. https://doi.org/10.48550/arXiv.1910.09700

Ligozat, A.-L., Lefevre, J., Bugeau, A., & Combaz, J. (2022). Unraveling the hidden environmental impacts of AI solutions for environment life cycle assessment of AI solutions. *Sustainability*, *14*(9). https://doi.org/10.3390/su14095172

Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 85–99. https://doi.org/10.1145/3630106.3658542

Luccioni, A. S., Viguier, S., & Ligozat, A.-L. (2022). *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model*. arXiv. https://doi.org/10.48550/arXiv.2211.02001

Oers, L. van, Guinée, J. B., & Heijungs, R. (2020). Abiotic resource depletion potentials (ADPs) for elements revisited—updating ultimate reserve estimates and introducing time series for production data. *The International Journal of Life Cycle Assessment*, *25*(2), 294–308. https://doi.org/10.1007/s11367-019-01683-x

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M., & Dean, J. (2021). *Carbon Emissions and Large Neural Network Training*. arXiv. https://doi.org/10.48550/arXiv.2104.10350

Simon, T., Ekchajzer, D., Berthelot, A., Fourboul, E., Rince, S., & Rouvoy, R. (2025). BoaviztAPI: A bottom-up model to assess the environmental impacts of cloud services. *SIGENERGY Energy Inform. Rev.*, *4*(5), 84–90. https://doi.org/10.1145/3727200.3727213

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P. M. (2013). *Climate change 2013: The physical science basis*. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)(Cambridge University Press, Cambridge, United Kingdom; New York, NY, USA.).

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models*. arXiv. https://doi.org/10.48550/arXiv.2302.13971

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Ferrer, C. C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., … Scialom, T. (2023). *Llama 2: Open Foundation and Fine-Tuned Chat Models*. arXiv. https://doi.org/10.48550/arXiv.2307.09288

Vries, A. de. (2023). The growing energy footprint of artificial intelligence. *Joule*, *7*(10), 2191–2194. https://doi.org/10.1016/j.joule.2023.09.004

You, J., Chung, J.-W., & Chowdhury, M. (2023). Zeus: Understanding and optimizing GPU

energy consumption of DNN training. *USENIX NSDI*.