

Quantifying the Environmental Impacts of LLMs in DH through Carbon Reporting

Sarah Lang*

2026-02-03

DHd AG GKI Webinar, <https://agki-dh.github.io/>

*Max Planck Institute for the History of Science (Berlin)

Is LLM Usage Environmentally Problematic?

Is (personal/academic) LLM usage environmentally problematic?

- AI-based search is more energy-intensive than traditional search.
- Many search engines now include AI by default.
- Some argue personal AI usage has negligible environmental impact (Masley 2025).
- We need environmental reporting but it also requires points of comparison.
- **Main Argument:** Carbon reporting should be standardized in (DH) research practices.



References

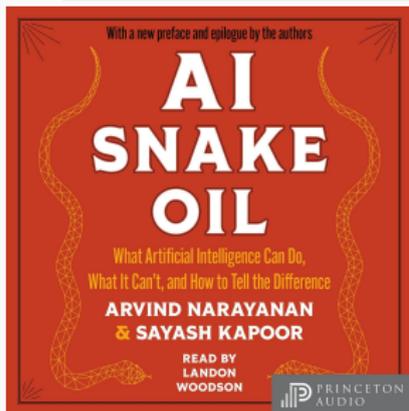
<https://www.nytimes.com/2025/09/26/opinion/ai-quartz-mining-hurricane-helene.html>

<https://www.youtube.com/watch?v=OJWcZLx45pw>

Environmental Impacts in Digital Humanities

LLMs and the Environment

- Early interest in environmental impacts via DHd's Greening DH and the *DH and the Climate Crisis* manifesto (Baillot et al. 2021).
- Sustainability in archiving already discussed since 2019 (Pendergrass et al. 2019).
- Surge in LLM use and public interest since 2023 with key papers referencing environmental aspects (e.g. *Stochastic Parrots* paper = Bender et al. 2021).
- But the perceived LLM revolution is mainly in usability and perception, not tech itself (Narayanan and Kapoor 2024).



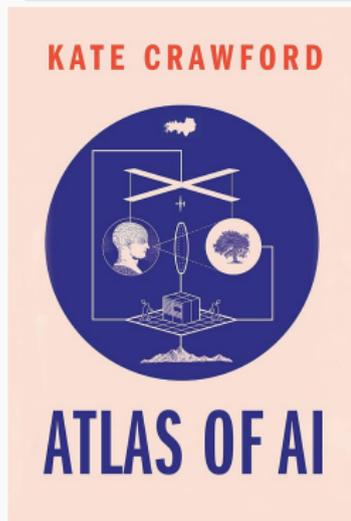
Learn more at

<https://dhd-greening.github.io/>

Critical Perspectives: Kate Crawford's *Atlas of AI*

AI Criticism: Earth, Water, Energy

- **Kate Crawford:** advocate advancing critical perspectives on AI (*ImageNet Roulette, Anatomy of an AI System, ...*).
- **Recent Work:** NY Times piece mentions Spruce Pine (NC, US) mining and hurricane Helene impacts → essential extraction site for rare earths (High Purity Quartz) in AI supply chain.



Water and Energy: Guesstimates and Uncertainty

AI, Water and the Challenges of Reporting

- **Hank Green** explains complexities of AI water usage.
- Multiple estimates are all kind of valid, yet differ widely.
- The term “**guesstimate**” reflects the uncertainty in impact metrics.
- Query impacts are difficult to track due to many hidden variables, even if all required information were publicly available.



Why is Everyone So Wrong About AI Water Use??



Watch on Youtube

https://www.youtube.com/watch?v=H_c6MWk7PQc

Key Points

- Carbon reporting is gaining attention in some LLM contexts.
- The BLOOM model is built with this in mind (Luccioni, Viguier, and Ligozat 2023).
- **DH Community:** Mixed or lacking response in Computational Humanities (CHR) conference – rejected CHR paper, reviewers indifferent.
- Does advocating for carbon reporting not matter in DH?



The screenshot shows the Zenodo interface for a working paper. At the top, the Zenodo logo is on the left, and a search bar, 'Communities', and 'My dashboard' are on the right. Below the header, there is a circular logo for 'DH AG Greening DH' and the text 'DH AG Greening DH'. The paper's publication date is 'Published December 11, 2025 | Version v1'. The title is 'Quantifying the Environmental Footprint of Curating Datasets with LLMs'. The authors listed are Lang, Sarah; Pitawanik, Wisyut; Belouin, Pascal; Sevink, Emma; Olszynko-Gryn, Jesse; Freeborn, Alfred; and Benson, Etienne. There is a 'Show affiliations' button. The abstract text is partially visible at the bottom.

zenodo Search records... Communities My dashboard

 DH AG Greening DH

Published December 11, 2025 | Version v1 Working paper Open

Quantifying the Environmental Footprint of Curating Datasets with LLMs

Lang, Sarah (Contact person)¹  Pitawanik, Wisyut (Researcher)¹  ;
Belouin, Pascal (Researcher)¹  Sevink, Emma (Data curator)²  ;
Olszynko-Gryn, Jesse (Researcher)¹  Freeborn, Alfred (Researcher)¹  ;
Benson, Etienne (Project leader)¹ 

Show affiliations

This study evaluates the environmental trade-offs of using large language models to curate cross-collection oral-history datasets in the Commoning Oral Histories of Knowledge (CORAL) project. Manual screening of 2,606 interviews was benchmarked against a workflow that tested four instruction-tuned LLMs and two prompt designs. Environmental impact was approximated using token-based inputs to EcoLogits, although implementing such assessments remains non-trivial. Ultimately, we conclude that the environmental impact of our project's use case could be considered moderate compared to common academic activities such as traveling to conferences. However, such impacts should be monitored closely, as they may vary significantly across different research setups and are likely to scale with larger datasets and broader adoption of LLMs in the field. Finally, the paper urges sufficiency-oriented practices and transparent carbon reporting in Computational Humanities research.

Learn more:

Lang, Sarah, Wisyut Pitawanik, Pascal Belouin, Emma Sevink, Jesse Olszynko-Gryn, Alfred Freeborn, and Etienne Benson. "Quantifying the Environmental Footprint of Curating Datasets with LLMs". Zenodo, December 11, 2025. <https://doi.org/10.5281/zenodo.17902822>. (Zenodo DHd AG Greening DH Community)

Carbon Reporting: Technical and Regulatory Considerations

It's challenging...

- U.S.-companies-based tools are not compliant with German privacy regulations. The local alternative provider GWDG seems to be working on a solution.
- Our paper attempts to provide practical guidance and derived metrics for local use. The methodology is (somewhat naively) based on input/output tokens and location of computation.
- Remember this is a guesstimate due to partial and indirect metrics! Be wary of spurious precision! Modeling theory teaches us that quantifying isn't inherently meaningful. Putting a number on it does not necessarily make it more objective!



The screenshot shows the Zenodo interface for a working paper. At the top, the Zenodo logo is on the left, and a search bar, 'Communities', and 'My dashboard' are on the right. Below the header is a circular logo for 'DH AG Greening DH'. The paper title is 'Quantifying the Environmental Footprint of Curating Datasets with LLMs', published on December 11, 2025, with version 1. The authors listed are Sarah Lang, Wishyut Pitawanik, Pascal Belouin, Emma Sevink, Jesse Olszynko-Gryn, Alfred Freeborn, and Etienne Benson. A 'Show affiliations' button is visible. The abstract text at the bottom of the screenshot reads: 'This study evaluates the environmental trade-offs of using large language models to curate cross-collection oral-history datasets in the Commoning Oral Histories of Knowledge (CORAL) project. Manual screening of 2,606 interviews was benchmarked against a workflow that tested four instruction-tuned LLMs and two prompt designs. Environmental impact was approximated using token...

Reference

Lang, Sarah, Wishyut Pitawanik, Pascal Belouin, Emma Sevink, Jesse Olszynko-Gryn, Alfred Freeborn, and Etienne Benson. "Quantifying the Environmental Footprint of Curating Datasets with LLMs". Zenodo, December 11, 2025. <https://doi.org/10.5281/zenodo.17902822>.^{7/11}
(Zenodo, DHd AG Greening DH Community)

The CORAL Project: Carbon Reporting in Practice

CORAL

- *Commoning Oral Histories of Knowledge* project.
- LLM usage for curating thematic sub-collections.
- Keyword matching fails for complex environmental terms.
- LLMs employed for pre-sorting interviews by relevance.
- Human-in-the-loop workflow for final selection and validation.
- Comparison of human effort vs. LLM results.



[home](#) [about](#) [contact us](#) | [sign up](#) [login](#)

CORAL: Commoning Oral Histories of Knowledge is a digital platform that facilitates research across dispersed oral history collections by indexing their transcripts and metadata. It enables researchers to search across collections held at different institutions, visualize and analyze connections between interviews, and explore themes, places, and people, while acting as a meta-platform that increases the visibility of underused archives and guides users back to the institutions that preserve them. The platform features catalogues of collections relevant to specific themes in the history of science, technology, medicine, and environment.

ComBio

Commoning Biomedicine (ComBio) aims to bring together numerous oral history repositories that can already be found online in a manner that makes them more visible and accessible to historians of biomedicine. These repositories are diverse and range from large institutional archives to small and thematically focused collections.

ComBio currently makes fully searchable a total of **1637** records from **15** collections.

[Search](#) [Statistics](#) [Analytical tools](#)

SEES

The Storying the Earth and Environmental Sciences (SEES) catalogue indexes diverse collections of oral histories of environmental knowledge, making them more discoverable and accessible to historians, practitioners, community members, policymakers, and environmental scientists. The SEES catalogue is under active development and will continue to grow as relevant collections are identified and our community of collaborators expands.

SEES currently makes fully searchable a total of **332** records from **8** collections.

[Search](#) [Statistics](#) [Analytical tools](#)

See the platform

<https://coral.mpiwg-berlin.mpg.de/>

Conclusions and Reactions

A bright future?

- Surprise at limited interest in CHR but engagement from this audience and on Mastodon is appreciated.
- Discussion continued on Mastodon. **Criticism:** Estimates likely underestimate actual environmental impact. Suggestion to consider both input and output tokens for better accuracy.

1) we only used completion, i.e. output, tokens as that is what the tools we used do, too, although I'm not 100% sure how this is justified.

1

 **Osma Suominen** @osma@sigmoid.social Jan 7

@sarahalang Thanks for confirming! I think that prompt (input) processing is important too, especially for tasks where the input size is large relative to the output (e.g. in document classification, where the input can be long and the output is just a few tokens). PP is typically much faster than generation, but can still be a significant factor in response time and thus energy use.

2

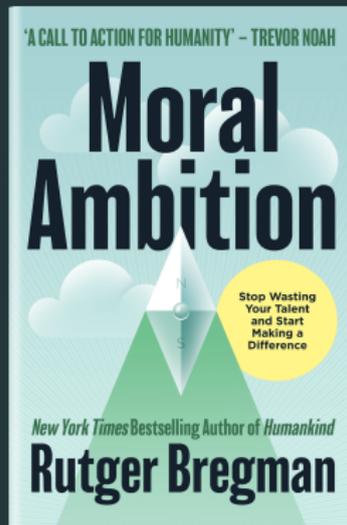
 **Osma Suominen** @osma@sigmoid.social

@sarahalang It seems to me that your LLM usage was quite prompt-heavy since the documents were quite long (IIRC you didn't state how many tokens average/median) and the output was just a classification (relevant/irrelevant) plus a few bits of structured information. Thus your energy use was likely severely underestimated. 40 kWh is not that much.

Reference

<https://fedihum.org/@osma@sigmoid.social/115854498900953842>

Carbon Reporting helps us raise awareness &
What gets measured gets managed. (P. Drucker)



But are more awareness and visibility enough?

References i

- [1] Anne Baillot et al. *Digital Humanities and the Climate Crisis: A Manifesto*. <https://dhc-barnard.github.io/dhclimate/>. Online document. 2021.
- [2] Emily M. Bender et al. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21. Virtual Event, Canada: Association for Computing Machinery, 2021, pp. 610–623. ISBN: 9781450383097. DOI: [10.1145/3442188.3445922](https://doi.org/10.1145/3442188.3445922). URL: <https://doi.org/10.1145/3442188.3445922>.
- [3] Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat. “Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model”. In: *Journal of Machine Learning Research* 24 (2023), pp. 1–15. URL: <https://www.jmlr.org/papers/volume24/23-0069/23-0069.pdf>.
- [4] Andy Masley. *Why using ChatGPT is not bad for the environment – a cheat sheet*. <https://andymasley.substack.com/p/a-cheat-sheet-for-conversations-about>. Substack blog post. Apr. 2025.

- [5] Arvind Narayanan and Sayash Kapoor. *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*. Princeton, NJ: Princeton University Press, 2024.
- [6] Keith Pendergrass et al. “**Toward Environmentally Sustainable Digital Preservation**”. In: *The American Archivist* 82.1 (June 2019), pp. 165–206. DOI: [10.17723/0360-9081-82.1.165](https://doi.org/10.17723/0360-9081-82.1.165). URL: <https://doi.org/10.17723/0360-9081-82.1.165>.