The Environmental Cost of DATA

It’s time to focus on Data Pollution

- **Environmental impact of data processes.** Data processes have significant environmental costs, including energy, water usage, and the reproduction of toxic chemicals. These impacts are expected to grow as data usage expands and climate risks increase.

- **From individual to industry responsibility:** Seemingly small actions like conducting a Google search contribute to data pollution when multiplied across billions of daily users. Individual actions, when combined, are dwarfed by the massive carbon footprint of large corporations like Amazon and Google, making it crucial to address both personal awareness and corporate accountability.

- **Underreported emissions by Big Tech.** A recent report by The Guardian reveals that Big Tech’s in-house data center emissions are 662% higher than officially reported. Given this, would you still trust the way they measure their pollution?

- **Lack of standardized carbon footprint calculation.** There is no universal method to calculate the carbon footprint of data processes, making accountability and transparency difficult.

- **Importance of auditing and accountability.** Auditing systems for their environmental impact and creating clear metrics are essential to holding corporations accountable. Solutions like taxation, energy-efficient protocols, and community-led movements can help mitigate the effects of environmental data pollution.

Modern society runs on a constant exchange of data, with an ever-growing demand for the network services that allow this exchange. Only in recent decades have researchers, policymakers, and advocates begun to confront **the environmental costs of such data processes. "Environmental data pollution" refers to the cumulative effect, including not only energy and water usage but also forms of pollution such as the production of wastewater and toxic chemicals.** As technology evolves, at least two trends seem certain: data processes will expand, and climate risks will deepen. Calculating data pollution costs and impacts is key to empowering society and policymakers to make decisions on the costs of innovation.

The challenge of measuring data's carbon footprint is complex, with no universal agreement on how to assess the impact accurately. Various studies have attempted to estimate the carbon footprint of specific processes, but factors such as location, devices, and energy sources complicate the calculations. **Without numbers, there is no accountability for this source of pollution.** The rise of artificial intelligence (AI) further exacerbates this issue, as training AI models requires immense computational power, leading to increased energy consumption and emissions from the data centers that power the digital economy.

While big tech companies have made promises to reduce emissions, the rapid expansion of AI and data services continues to drive up energy use. Fortunately, there are ways to **mitigate** these impacts, including data minimization, optimizing AI training, and switching to more energy-efficient protocols. But **the incentives to mitigate Data Pollution are not there if no one is counting.** Only with accurate oversight, and the willingness to make tough choices, can we curb the environmental costs of data, ensuring that technological progress does not come at the expense of the planet.

Advocating for AI auditing has been a core focus of our work at Eticas. As our work on community-led AI audits and building software solutions to measure and audit bias in AI systems, we continue to be committed to advancing the field of Responsible AI, as we've been doing for the last 15 years.

Below you will find references and links to some work we have been conducting in the last 3 years on data processes and data centers and their impact on the environment and on the possibility of taxing Environmental Data Pollution like we tax sales or property.

On this Climate Week, we are launching our **Data Pollution Auditing Project**, to build data and momentum around the need to audit AI systems and processes for their environmental impact.

Interested? Reach out and get involved!
mireia.orra@eticas.ai

What causes data pollution?

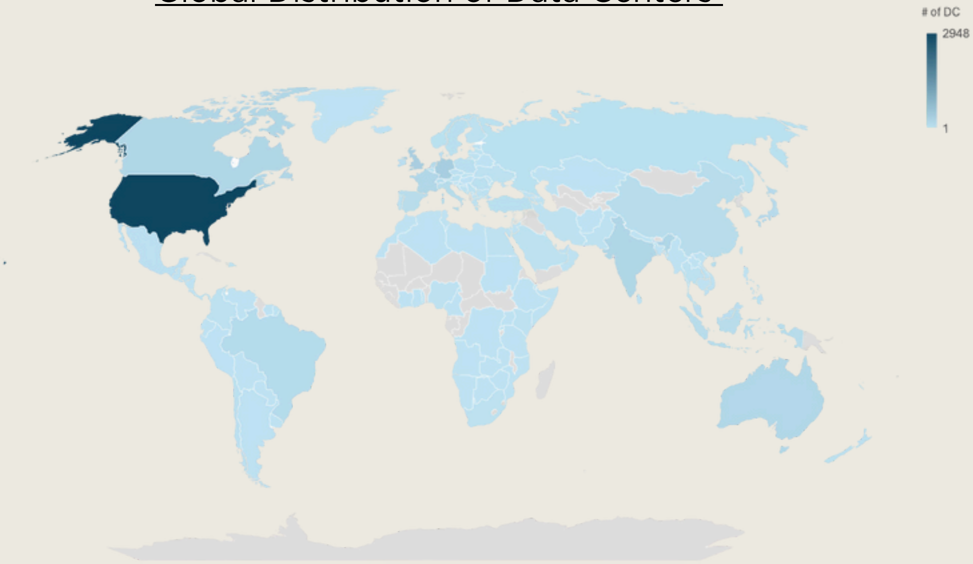
Every single exchange of data, or its storage, requires resources. Streaming content, refreshing web pages, asking ChatGPT to write a poem or sending a thank you e-mail requires power and water. Even a simple data process can result from multiple imperceptible steps, each with unique energy demands.

One of the top uses of energy is **machine learning**, which can require immense computing power and is at the heart of every new feature of the AI revolution. A 2019 study by the University of Massachusetts calculated that the energy needed to develop a standard AI model for natural language processing equaled the emissions of five cars over their entire lifetimes. As models grow, so does energy usage, with OpenAI estimating that the power needed to run large AI models doubles every three-and-a-half months. In fact, as reported by The Guardian, it is expected that the growth of data center power will increase by 160% by 2030 (O'Brien, 2024).

Bitcoin mining is another major consumer of energy. Proof of Work systems validate actions by requiring computers to solve a cryptographic puzzle—a secure but demanding task. For the Bitcoin blockchain, estimates of annual energy use run as high as 205 terawatt-hours, or the power consumption of Thailand. On top of that, cookies, video conferencing, and other energy-intensive tasks add to the overall tally.

Around the world, **data centers power the global flow of data**, housing the infrastructure and computer components that make our digital world possible. There are nearly **8,000 data centers distributed around the globe**. As of 2024, there are approximately 133 in Africa, 813 in Asia, 430 in Oceania, 3188 in North America –with more than 2000 in the United States alone-, 98 in Central America, 310 in South America, 566 in Eastern Europe, and 1962 in Western Europe ("Worldwide Data Centers - Colocation and Cloud," n.d.). Yet the environmental cost is steep, with data centers using two per cent of the world's electricity and enormous amounts of water.

Global Distribution of Data Centers



Source: Worldwide Data Centers - Colocation and Cloud. Own elaboration.

Although data centers bring with them a promise of better infrastructure for development, this coin has two faces. The **environmental impact extends beyond energy use.** Data centers also affect limited resources, such as water. For instance, in Uruguay in 2024, amid the worst water drought in 74 years, the announcement of a new data center was met with concern, particularly when it was revealed that the facility would draw 7.6 million liters of potable water daily which is enough to cover the domestic needs of 55,000 people (Fuentes, 2023).

This **history of development at the expense of the tangible conditions of entire communities** is an experience that knows no geographical boundaries. It is a story told in parched farms in Mexico (Baptista & McDonnell, 2024), in communities in India affected by waste produced by tech centers (Butler, 2024), or in Oregon, where data centers devour millions of liters of water annually (Rogoway, 2022). Even when we see the most rudimentary elements, lithium mining, for instance, can impinge on natural preserves, reduce biodiversity, and pollute water sources—not to mention involve human rights violations resulting from the exploitative labor. The variety of EDP makes for a **measurement challenge**, as scientists lack a simple metric to capture the impact of data processes on the climate.

As long as the **volume of data continues to increase daily and public awareness of its environmental impact remains low**, the technological revolution is quietly putting our planet in check. Hence, it is vital to recognize the cumulative effects of various data-intensive activities, not just within corporations but also in everyday activities we all take part in.

Small, actions like sending emails, streaming music, checking social media, seem minor, however they add up to an environmental burden. While these effects might seem abstract, data centers are working constantly to keep up with the demands of ordinary people, like any of us.

Ali's Daily Activities. Own elaboration.

Consider a typical day of a regular person, this time we can name them Ali. She wakes up and starts her morning with a podcast playing in the background while taking a shower. After, she heads to the kitchen, since she's just cooking breakfast, she catches up with the newest chapter of a new series she has been following. Although it doesn't seem like much, just half an hour of watching Netflix is equivalent to driving 100 meters in a conventional car (Kamiya, 2020). If Ali watches 2 hours of Netflix daily for a year, it would be equivalent to driving 146 kilometers, which is roughly the distance from Málaga to Granada or from New York to Philadelphia! All of this without leaving her house, just from watching shows.

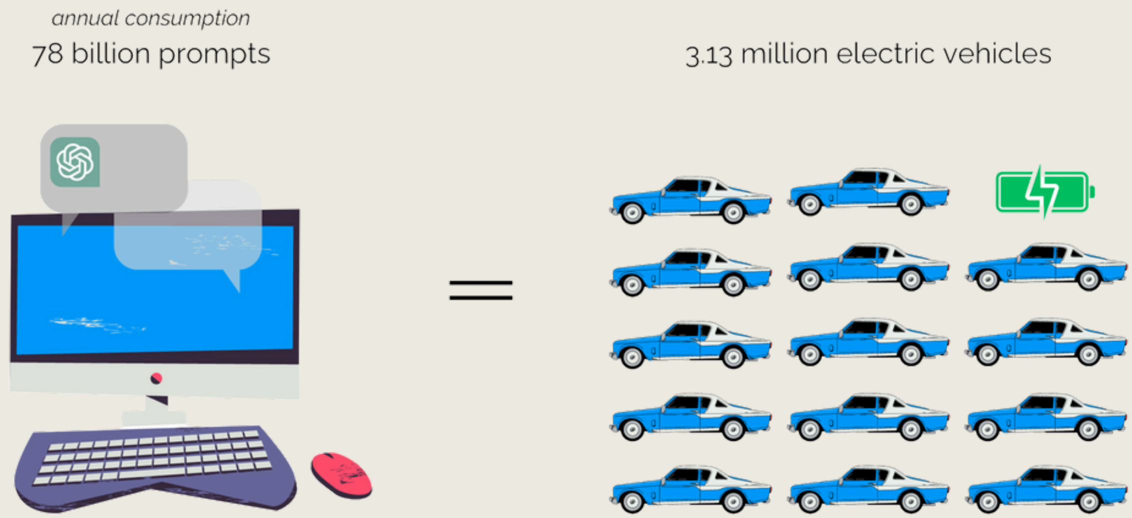
Throughout the day, Ali might send 40 emails, several short updates and other lengthy walkthroughs of her new music project. Yet, each email, depending on its size, emits between 0.3 grams and 17 grams of CO2 emissions. In a year, this routine could lead to 3 to 40 kilograms of CO2 (Chu, n.d.). which is the equivalent of two heavily packed suitcases, just from emails. Even her quick Google searches throughout the day add up to 0.2 grams per query. Although some questions might not seem like much, there is a hidden cost to not knowing.

As the day progressed, she joined a virtual meeting. However, one hour of videoconferencing can emit between 150 and 1000 grams (about the weight of a liter of water) of CO2, depending on the service (Travers, 2021). By the time the meeting ends, the environmental impact of these tasks keeps building up. After work, she started an insightful conversation with ChatGPT about the fundamentals of project management, and where exactly does wind come from. A conversation of 20 to 50 questions could consume enough water to fill a 500ml bottle (Mclean, 2023).

By night, if she scrolls on TikTok for 10 minutes before bed, that alone results in 26.3 grams of CO2 emissions (Gibson, 2023). By the end of the day, Ali's ordinary day full of podcasts, streaming a couple of episodes, sending emails, attending meetings, searching on Google and searching again on ChatGPT, have added to a growing environmental tally.

Ali might be fictional, but she could easily represent any of us. **The accumulation of small individual actions ends up fueling a heavy environmental contribution.** For example, a single Google query might emit only 0.2 grams of CO2, but when multiplied by the 8.5 billion searches happening daily, the impact is far from insignificant. Additionally, when we take a step back and look at the processes that make those results possible, we see a heavy load. For example, it is estimated that ChatGPT consumes 226.8 GWh annually to process 78 billion prompts. This is comparable to enough energy to fully charge around 3.13 million electric vehicles (Igor Todorović 2024).

<u>Energy Consumption of ChatGPT: Annual Usage Compared to Electric Vehicle Charging</u>



Source: Igor Todorović. Own elaboration.

This brings us to **two urgent demands** that work in parallel. First, we must recognize our individual responsibility in the rise of EDP. When we understand our contribution, not as an abstract but in concrete terms, **we see that to some degree we are part of the problem, but also a critical part of the solution.** Any meaningful movement needs informed, conscious individuals, and we have the power to be exactly that.

Our second, more urgent and stronger, demand: **industries must bear a larger responsibility**. Although individuals contribute to environmental costs, the burden is not carried equally. When put in comparison, Ali's 0.2 grams per query on Google is nothing compared to the 21 billion liters of potable water Google used in 2022 alone (Bosch, Gupta, & Vliet, 2024). As a point of comparison, 2.1 billion people worldwide lack access to safe drinking water, meaning Google used 10 times the amount needed to address global water insecurity.

This is not just an isolated, extremely thirsty, incident. For example, in 2023, Amazon's Sustainability Report highlighted that the company's emissions reached 68.82 million metric tons of CO_2. This amount is more than the annual emissions of countries like Singapore, Austria, Portugal, or Bolivia. In fact, according to the emissions by country reported by Worldometer (2016), **if we combine the emissions of Cameroon, Honduras, Estonia, Zimbabwe, and Kyrgyzstan, they still wouldn't match Amazon's carbon footprint.** Similarly, Microsoft's global water consumption is another alarming case, the company's consumption of 6.4 million cubic meters in 2022 is equivalent to 2,500 Olympic-sized swimming pools. (Sreedhar, 2023). To further illustrate these numbers, if we were to combine the emissions of Amazon, Apple, Microsoft, Google, and Meta, their collective carbon output would rank them as the 33rd largest emitter globally, just behind the Philippines (O'Brien, 2024).

Amazon's Carbon Footprint: Surpassing the Combined Emissions of Cameroon, Honduras, Estonia, Zimbabwe, and Kyrgyzstan

Source: Amazon's Sustainability Report. Own elaboration.

These are not just abstract numbers; they highlight the **importance of measuring environmental impacts accurately.** A shared standard is essential. As The Guardian reported (O'Brien, 2024), companies often use renewable energy certificates (Recs) to make their data centers appear greener than they actually are. Recs allow firms to claim they're using renewable energy, even if that energy is generated far from where their data centers operate. This leads to a **significant gap between market-based emissions** (what's officially reported) **and location-based emissions** (the actual environmental impact). In reality, local emissions can be much higher. For instance, when comparing official versus location-based emissions, Amazon's emissions are 5 times higher than reported, Google's are 2.7 times higher, Microsoft's are 21 times higher, Meta's are a shocking 3,100 times higher, and Apple's are 402 times higher. This discrepancy shows how Recs help underplay the true carbon footprint of these companies.

Without clear metrics, it's nearly impossible to rationalize the amount of carbon emissions or water usage. This is not just a problem about standardization, **measurements help mobilize the support needed to hold companies accountable for their environmental toll.** Having examined the context and causes of EDP, we now turn to solutions, specifically focusing on the introduction of auditing practices to effectively measure and mitigate its impacts.

Auditing Data Pollution

As seen above, **data processes eat into the natural world and fuel climate change**. Possible solutions are as diverse as the problems. Taxation can shift incentives, especially if focused on data centers, and other measures can mitigate environmental costs.

But the incentives to do better won't emerge unless someone counts and holds environmental data pollution and data polluters accountable.

As we begin to test and implement ways to measure EDP and audit systems for their environmental impact, we want to contribute to building a global movement and debate on EDP. This movement seeks to track the actions of major corporations and, collectively, uncover the true environmental costs of their operations. By shining a light on these practices, we can help hold companies accountable for their environmental footprint and push for greater transparency and responsibility. We envision this movement to involve **academics**, so that the strategies for mitigating data pollution are informed by the latest scientific and technological advancements, **advocates**, to translate research into actionable plans and campaigns, and build EDP into existing initiatives to combat Climate Change, and **policymakers**, who can then bridge the gap between concern and regulation, creating the incentives for industry to do better.

Explore our work

Explore more here.

Our reports:

1. Accounting for the Impact of Data Processes, and the history of environmental taxation in the US and Europe.

2. Data Centers as Taxable Property: A feasibility report on taxation schemes on the amount of data kept in data centers.

3. Data Brokerage Tax: A feasibility report on applying a sales tax to online transactions of personal data.

Get involved!

If EDP is something you have worked on and would like to work with us, e-mail us at mireia.orra@eticas.ai

References

Baptista, D., & McDonnell, F. (2024, September 6). Thirsty data centres spring up in water-poor Mexican town. Retrieved September 17, 2024, from Context.news website: https://www.context.news/ai/thirsty-data-centres-spring-up-in-water-poor-mexican-town

Butler, G. (2024). Microsoft's Telangana data center allegedly illegally dumping waste in local lake. Retrieved from DCD website: https://www.datacenterdynamics.com/en/news/microsofts-telangana-data-center-allegedly-illegally-dumping-waste-in-local-lake/

Amazon. (2023). Amazon Sustainability Report. Retrieved from https://sustainability.aboutamazon.com/content/dam/sustainability-marketing-site/pdfs/reports-docs/2023-amazon-sustainability-report.pdf

Bosch, H., Gupta, J., Vliet, L. van. (2024, March 21). AI's excessive water consumption threatens to drown out its environmental contributions. Retrieved from The Conversation website: https://theconversation.com/ais-excessive-water-consumption-threatens-to-drown-out-its-environmental-contributions-225854

Chu, D. (n.d.). What's the carbon footprint of an email? Retrieved from www.pawprint.eco website: https://www.pawprint.eco/eco-blog/carbon-footprint-email#:~:text=Over%20the%20course%20of%20a

Fuentes, F. (2023, July 12). El polémico data center que Google planea construir en Uruguay en medio de la peor sequía en 74 años. Retrieved September 17, 2024, from La Tercera website: https://www.latercera.com/la-tercera-pm/noticia/el-polemico-data-center-que-google-planea-construir-en-uruguay-en-medio-de-la-peor-sequia-en-74-anos/PO7W4MF4J5G3JKXJYLDFFOLGTY/

Galdon Clavell, G., & Mastracci, M. (2023a). Accounting for the Environmental Impact of Data Processes. Retrieved from https://www.next-now.org/sites/default/files/2023-03/Accounting%20for%20the%20enviromental%20impact%20of%20data%20process_0.pdf

Clavell, G., & Mastracci, M. (2023b). Data Brokerage Tax. Retrieved from https://www.next-now.org/sites/default/files/2023-03/Data%20brokerage%20tax_0.pdf

Galdon Clavell, G., & Mastracci, M. (2023c). Data Centers as Taxable Property. Retrieved from https://www.next-now.org/sites/default/files/2023.03/Data%20centers%20as%20taxable%20property_0.pdf

Gibson, N. (2023, June 13). TikTok is the most polluting social media. Retrieved from Net Zero Professional website: https://netzeroprofessional.com/tiktok-is-the-most-polluting-social-media/

Igor Todorović. (2024, September 4). ChatGPT consumes enough power in one year to charge over three million electric cars. Retrieved September 18, 2024, from Balkan Green Energy News website: https://balkangreenenergynews.com/chatgpt-consumes-enough-power-in-one-year-to-charge-over-three-million-electric-cars/#:~:text=ChatGPT%20consumes%20226.8%20GWh%20each

Kamiya, G. (2020, December 11). The carbon footprint of streaming video: fact-checking the headlines – Analysis. Retrieved from IEA website: https://www.iea.org/commentaries/the-carbon-footprint-of-streaming-video-fact-checking-the-headlines

Mclean, S. (2023, April 28). The Environmental Impact of ChatGPT: A Call for Sustainable Practices In AI Development. Retrieved from Earth.org website: https://earth.org/environmental-impact-chatgpt/

O'Brien, I. (2024, September 15). Data center emissions probably 662% higher than big tech claims. Can it keep up the ruse? Retrieved from the Guardian website: https://www.theguardian.com/technology/2024/sep/15/data-center-gas-emissions-tech

Pilapitiya, P. G. C. N. T., & Ratnayake, A. S. (2024). The world of plastic waste: A review. Cleaner Materials, 11(100220). https://doi.org/10.1016/j.clema.2024.100220

Riseup. (2024). Riseup Pad. Retrieved September 17, 2024, from Riseup.net website: https://pad.riseup.net/p/Against_AI_and_Its_Environmental_Harms-keep

Rogoway, M. (2022, December 17). Google's water use is soaring in The Dalles, records show, with two more data centers to come. Retrieved from oregonlive website: https://www.oregonlive.com/silicon-forest/2022/12/googles-water-use-is-soaring-in-the-dalles-records-show-with-two-more-data-centers-to-come.html

Sreedhar, N. (2023, October 22). AI and its carbon footprint: How much water does ChatGPT consume? Retrieved from Mintlounge website: https://lifestyle.livemint.com/news/big-story/ai-carbon-footprint-openai-chatgpt-water-google-microsoft-111697802189371.html

Travers, K. (2021, March 4). How to reduce the environmental impact of your next virtual meeting. Retrieved from MIT News | Massachusetts Institute of Technology website: https://news.mit.edu/2021/how-to-reduce-environmental-impact-next-virtual-meeting-0304

Worldometer. (2016). CO2 Emissions by Country . Retrieved from Worldometer website: https://www.worldometers.info/co2-emissions/co2-emissions-by-country/

Worldwide Data Centers - Colocation and Cloud. (n.d.). Retrieved from www.datacentermap.com website: https://www.datacentermap.com/datacenters/

The Environmental Cost of DATA

Contact mireia.orra@eticas.ai