EVERY CHATGPT QUERY COSTS MORE THAN YOU THINK: QUANTIFYING THE HIDDEN WATER FOOTPRINT OF AI

Usmanova Shokhsanam Avazovna
Teacher at the Department of
The English language,
The University of World Economy
and Diplomacy, Tashkent, Uzbekistan
usmanova.sh@uwed.uz
Aziza Asrarova
First-year student of international relations
The University of World Economy and Diplomacy
Tashkent, Uzbekistan,
aziza.asrarova@gmail.com

ABSTRACT

Artificial intelligence (AI) has environmental impacts that are often overlooked. This paper examines the water footprint of generative AI models by analyzing water consumption associated with data center cooling and electricity generation. It is estimated that approximately 5.4 million liters of water were consumed during the training of GPT-3. In addition, a single ChatGPT query requires about 500 ml of water for every 10–50 responses. Water consumption varies significantly depending on the training location: approximately 15.29 million liters would be required in Washington compared with 3.68 million liters in Virginia. Temporal analysis showed no significant correlation between carbon and water efficiency (Pearson r = 0.06). Potential technical solutions include geographic load balancing, advanced cooling

technologies, and water recycling. Reducing the water footprint of AI requires standardized measurement methods and integrated technical strategies.

Keywords: Artificial Intelligence; Water Footprint; Data Centers; ChatGPT; Generative AI; Cooling Technologies; Geographic Load Balancing; Scope-1 Water Usage; Scope-2 Water Usage; Carbon-Water Trade-off; Sustainability Metrics.

1. INTRODUCTION

Since the introduction of ChatGPT in November 2022, generative AI has been adopted by over 200 million users globally in education, healthcare, business, and the arts (Garcia, 2025; Li et al., 2023). Although AI is capable to aid climate modeling and resource optimization, its functioning relies on a considerable amount of freshwater supply. Earlier studies have mostly concentrated on the energy consumption and carbon emissions while water use is hardly factored into the evaluation of AI models (Dhar, 2020; Schwartz et al., 2020; Li et al., 2023).

In data centers, water is used mainly for two purposes: cooling servers onsite and producing electricity offsite (Mytton, 2021). Cooling towers release the heat coming from server units while thermoelectric power plants draw water for making steam and cooling. Altogether, these two activities define the day-to-day water footprint of AI (Jegham et al., 2025; Reig et al., 2023).

Recent benchmarking of 30 state-of-the-art language models found that energy-intensive systems require over 29 Wh per long prompt up to 65 times more than efficient models (Jegham et al., 2025). Scaled to 700 million daily queries, even short prompts (0.42 Wh) can result in annual electricity consumption equivalent to that of 35,000 U.S. households and freshwater use comparable to the annual drinking water requirements of 1.2 million people (Li et al., 2023; Mytton, 2021; Nerini et al., 2025). This study addresses three questions: what is the water footprint of large language models during training and inference, how does it vary by location and time, and what technical solutions can reduce water intensity?

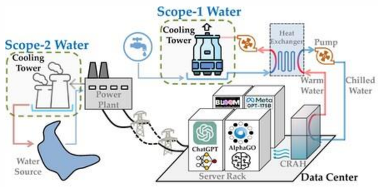
2. METHODS

Following the framework of Li et al. (2023), three scopes of AI water usage are distinguished

Scope- 1 Onsite Water Consumption: Water directly used for cooling in data center facilities This was calculated using reported water usage effectiveness (WUE) metrics, defined as the ratio of annual site water usage to ICT equipment energy consumption (L/kWh).

Scope-2 Offsite Water Consumption: Water used to generate electricity for AI operation. This was estimated using electricity water intensity factors (EWIF) for regional grid mixes, obtained from databases of the U.S. National Renewable Energy Laboratory and international energy agencies.

Scope-3 Water Consumption: Water used in manufacturing AI hardware. Due to limited public data, scope-3 estimates remain conservative because of the lack of publicly available data To estimate the water footprint of GPT-3 training, location-specific Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) data from Microsoft were applied to the model's reported training energy consumption of 1,287 MWh. It was estimated that per-request energy consumption was 0.004 kWh, which was used in inference calculations.



An example of datacenter's operational water usage: onsite scope-1 water usage for datacenter cooling (via cooling towers in the example), and offsite scope-2 water usage for electricity generation. The icons for AI models are only for illustration purposes.

In order to evaluate the spatial variation, the data center locations of Microsoft were evaluated with the region-based EWIF values (Reig et al., 2023). The temporal variation was investigated based on hourly grid mix data on Virginia comparing rates of carbon emissions with the intensity of water consumption among five days.

3. RESULTS

3.1 Water Consumption in AI Training

The act of training GPT-3 in U.S. data centers used by Microsoft took about 5.4 million liters of water, of which 700.000 liters was used on-site (Li et al., 2023). This however differs significantly depending on location as indicated in Table 1.

Location	Onsite WUE (L/kWh)	Offsite EWIF (L/kWh)	Total Water (million L)
U.S. Average	0.55	3.14	5.44
Arizona	1.63	4.96	9.63
Washington	0.95	9.50	15.29
Virginia	0.14	2.39	3.68
Ireland	0.02	1.48	2.29
Singapore*	0.01	N/A	N/A

Table 1. GPT-3 Training Water Consumption by Location

Note: EWIF data unavailable for Singapore. Adapted from Li et al. (2023) and Microsoft sustainability reports.

Training the same model in Washington requires over four times as much water as in Virginia for model training, cooling, and electricity generation.High consumption in Arizona is an indicator of high cooling and intensity of grid water.

3.2 Inference: Water Consumption per Query

GPT-3 also uses about 500ml of water per 10-50 responses, depending on where it is deployed, and at what time (Li et al., 2023). Table 2 gives estimates that are location-specific.

Location	Onsite Water (ml)	Offsite Water (ml)	Total Water (ml)	Queries per 500ml
U.S. Average	2.20	14.70	16.90	29.6
Arizona	6.52	23.41	29.93	16.7
Virginia	0.56	10.88	11.44	43.7
Iowa	0.56	14.40	14.96	33.4
Netherlands	0.24	15.71	15.95	31.4

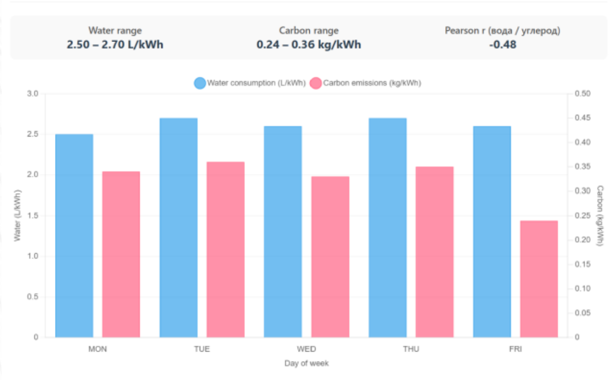
Table 2. Water Consumption per ChatGPT Query

Source: Calculated from Li et al. (2023) using location-specific WUE and EWIF.

One ChatGPT query uses about 30ml of water in Arizona, which is almost twice as much as Virginia needs. An end user who uses AI to make 10 queries every day would then use about 60 liters of water per year indirectly.

3.3 Temporal Variation and Grid Dynamics

The water consumption in scope-2 is time dependent, that is, the water consumption depends on the generation mix of electricity. Analysis shows that carbon and water efficiency are not aligned. The carbon emissions are reduced and the water consumption can rise during the high renewable penetration when the hydropower becomes a bigger portion. Pearson correlation coefficient between carbon and water efficiency on an hourly basis, in Virginia was 0.06 (Li et al., 2023), which means that optimization to one metric does not favor the other.



Hourly carbon emission rate and water consumption intensity in Virginia, April 2022

3.4 Projected Growth

By 2027, global AI water consumption is projected to reach 4.2–6.6 billion cubic meters annually, exceeding the total water withdrawal of Denmark or half the United Kingdom (Li et al., 2023). By 2028, U.S. data center water use may be more than 150280 billion liters, more than two to four times what it is in 2023 (Mytton, 2021).

4. DISCUSSION

4.1 Comparison with Recent Studies

The findings align with prior research. de Vries-Gao (2025) estimated annual AI water consumption at 312.5–764.6 billion liters, equivalent to global bottled water use. Indirect water use was found to be underestimated by a factor of three to four. A Nature Sustainability study by Xiao et al. (2025) projected U.S. AI servers could consume 731–1,125 million cubic meters annually by 2030, equivalent to household water use of 6–10 million Americans. Geographic optimization and efficient cooling were shown to reduce water impacts by approximately 76–86%, consistent with the present study.

4.2 Implications for the AI Industry

1. **Transparency gaps hinder mitigation.** Major technology companies do not disclose AI-specific water figures. Google recently declined to report indirect water use for its Gemini model, citing limited control over power plants. These practices conflict with Greenhouse Gas Protocol standards. The IEEE P7100 Working Group is developing technical standards to fill these gaps.

2. **Carbon-water trade-offs need to be managed with integrated strategies.** The non-correlation of carbon and water efficiency (r = 0.06) suggests that efforts to achieve carbon neutrality through the use of renewable energy might lead to an increase in water consumption if hydropower is the dominant source of electricity.

3. **The current sustainability commitments may not be enough**. Microsoft and Google have promised to become water positive by 2030, but the expected growth of AI, more than 1 billion cubic meters per year in the U.S. by 2030, will make it more and more difficult to meet these commitments.

4. **Innovation in cooling technology can separate growth from water usage.** Technologies such as advanced liquid cooling, immersion cooling, and closed-loop systems have the potential to completely remove onsite water consumption. Microsoft has made a pledge that its new data centers will "consume zero water for cooling."

5. **Lifecycle impacts matter.** From chip manufacturing to disposal, AI's resource footprint requires comprehensive accounting (Yang & Du, 2025).

4.3 Technical Solutions

Geographic and temporal load balancing can reduce water footprint by 76% (Table 1) through strategic model placement and inference scheduling during low water-intensity periods, optimizing both water and carbon efficiency without performance loss.

Innovative cooling technologies such as dry cooling, direct-to-chip liquid cooling, and immersion cooling can eliminate onsite water consumption. Microsoft has already committed to zero-water cooling in new data centers (Garcia, 2025), although the lifecycle impacts still need to be evaluated.

The reuse of water leads to a reduction in the demand for freshwater resources. According to Google, about one-third of its campuses are using reclaimed water, whereas most U.S. facilities continue to use potable water sources (Garcia, 2025). Therefore, the use of treated wastewater and rainwater can help alleviate the burden on local water resources.

Mandatory disclosure of water use should be implemented. Regulations should require annual reporting of water withdrawal, consumption, sources, and discharge quality using standardized procedures. Environmental

Protection Agency (EPA) should be allowed to audit these reports, and appropriate sanctions should be handed down in cases of violations.

The siting of data centers in water-stressed regions should undergo prior environmental impact assessment and include mitigation strategies. If significant impacts are expected, alternative locations should be considered.

4.4 Standardization Gaps

Standardized measurement methods are lacking. Although the EU AI Act mandates reporting of computational resources and energy, water reporting is not required. Benchmarking frameworks (Jegham et al., 2025) provide a model. OECD and IEEE are developing technical standards through the P7100 Working Group.

4.5 Limitations

Scope-3 water consumption is poorly quantified. Manufacturing AI hardware consumes significant water for semiconductor production, yet public data are scarce. Additional research is needed, particularly in developing countries where new data centers are being constructed.

5. CONCLUSION

It takes millions of liters of freshwater to train one large language model. Every ChatGPT query has a quantifiable water price that depends significantly on location and time. Recent research affirms that AI water footprint may grow to 1.1 billion cubic meters a year in the U.S. by 2030, the equivalent of the household used water of 10 million Americans. There are technical options: geographic load balancing is capable of cutting down on consumption by up to 86 percent; advanced cooling can completely do away with onsite use. But, it will be necessary to have standardized measurement procedures and corporate transparency to implement. The need of integrated optimization is highlighted by the carbon-water trade-off (r = 0.06). Ultimately, the sustainable evolution of artificial intelligence hinges not only on technological innovation but on a fundamental commitment to transparency and integrated resource management. The cost of every ChatGPT query is paid not just in electricity, but in a shared, finite resource—freshwater. Recognizing and acting upon this reality is the first step toward ensuring that AI's promise does not come at the expense of the planet's most essential needs.

REFERENCES

1. de Vries-Gao, A. (2025). AI's hidden carbon and water footprint. *Patterns*, published in conjunction with Vrije Universiteit Amsterdam. https://vu.nl/en/news/2025/ai-s-hidden-carbon-and-water-footprint?utm.com=

2. Xiao, T., You, F., et al. (2025). Environmental impact and net-zero pathways for sustainable artificial intelligence servers in the USA. *Nature Sustainability*. https://gizmodo.com/training-ai-could-soon-use-as-much-water-as-10-million-americans-2000683679

3. IEEE/OECD. (2025). The hidden costs of AI: Unpacking its energy and water footprint. *OECD AI Policy Observatory / IEEE P7100 Working Group*. https://oecd.ai/en/hidden-costs-ai

4. Jegham, N., et al. (2025). How hungry is AI? Benchmarking energy, water, and carbon footprint of LLM inference. https://arxiv.org/abs/2505.09598v5

5. Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). *Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models*. https://arxiv.org/html/2304.03271v5

6. Garcia, M. (2025). *AI Uses How Much Water? Navigating Regulation of AI Data Centers' Water Footprint Post-Watershed Loper Bright Decision.* Texas Tech Law Review, 57, 517-554. https://ttu-ir.tdl.org/items/3bb26787-9659-4867-9d5b-b853333f7693/full

7. Mytton, D. (2021). *Data centre water consumption.* npj Clean Water, 4(11), https://www.nature.com/articles/s41545-021-00101-w

8. **Dhar, P. (2020).** *The carbon impact of artificial intelligence.* Nature Machine Intelligence, 2(8), 423-425.https://www.nature.com/articles/s42256-020-0219-9

9. Nerini, F. F., et al. (2025). AI could consume 10% of US electricity by 2030. *Environmental Change Institute, University of Oxford / Nature Sustainability*. https://www.eci.ox.ac.uk/news/ai-could-consume-10-us-electricity-2030-researchers-outline-path-sustainable-growth

10. Responsible AI Foundation. (2025). AI companies are draining water from regions that are already running dry. https://www.responsibleaifoundation.com/post/ai-companies-are-draining-water-from-regions-that-are-already-running-dry

11. **Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020).** *Green AI.* Communications of the ACM, 63(12), 54-63. https://cris.huji.ac.il/en/publications/green-ai