

Abstract

This project visualizes freshwater consumption associated with large language model (LLM) inference and training by mapping water use to a $10 \times 10 \times 10$ m pool (volume = $1,000 \text{ m}^3 = 1,000,000 \text{ L}$). Using benchmark estimates (see “How Hungry is AI? Benchmarking Energy, Water, and Carbon Footprint of LLM Inference” and reasonable assumptions where the paper does not give explicit per-inference liters), the visualization animates how quickly the pool would fill under different usage scenarios. The aim is to provide an intuitive, time-based visual metaphor that helps non-experts and peers grasp the water footprint of current AI workloads, and to guide improvements in measurement and communication.

1. Introduction

Concern over the environmental footprint of AI has mostly focused on energy and carbon. Water is a less-visible but important resource impacted by data center cooling and other infrastructure. To make water use more tangible, we built a WebGL visualization that simulates filling a $10 \times 10 \times 10$ m pool and exposes controls for scenario parameters (model size, request rate, water-per-inference, and time scaling). The design goal is clarity: viewers should instantly see how quickly a familiar volume of water (a small pool) is consumed.

The goals as of now are:

- Describe the mapping from reported water-use metrics to visual fill rate.
- Provide working example scenarios and step-by-step calculations.
- Explain implementation and how to run the visualization locally.

2. Background and references

The project uses the arXiv paper “How Hungry is AI? Benchmarking Energy, Water, and Carbon Footprint of LLM Inference” as its empirical starting point. That work benchmarks inference water use across models and datacenter configurations and highlights that water use depends strongly on cooling architecture, PUE/WUE (Power/Water Usage Effectiveness), and geography. Because reported values vary and are often aggregated (e.g., liters per kWh or liters per inference averaged), we make explicit assumptions in the scenarios below and invite reviewers to suggest better parameter values or to supply raw model traces.

3. Visualization concept and data mapping

- Pool dimensions: $10 \text{ m} \times 10 \text{ m} \times 10 \text{ m} \rightarrow$ pool volume: