1

1. Energy Consumption
• Training and Inference Costs: AI models, particularly large-scale deep learning models, require substantial computational resources, especially during training. Studies have measured the energy required to train different models, such as large language models (LLMs). For example, GPT-3 and similar models consume significant amounts of energy due to their complexity and the number of parameters.
• Carbon Emissions: The energy consumed by AI systems is often converted into carbon footprint measurements. This is done by considering the carbon intensity of the electricity mix (renewable vs. non-renewable sources) used to power the data centers. Tools like Carbontracker help estimate the carbon footprint of AI models by analyzing the energy usage during training and inference.

2. Model Size and Computational Intensity
• Larger AI models typically have a larger ecological footprint because they require more computational resources, both during training and inference. Researchers have compared the carbon emissions of different models, showing that larger models generally produce more emissions unless offset by more efficient hardware or renewable energy use.

3. Hardware and Data Center Efficiency
• The ecological footprint is also influenced by the efficiency of the hardware (GPUs, TPUs, CPUs) used to run AI models. Energy-efficient hardware and more efficient data center cooling and power systems can significantly reduce the overall footprint.

4. Training Location and Data Center Type
• The geographic location of data centers impacts their carbon footprint. Some regions have greener energy grids, relying more on renewable energy, while others depend heavily on fossil fuels. AI systems run in regions with greener grids or in carbon-neutral data centers have a smaller ecological footprint.

5. Frameworks and Tools
• There are efforts to standardize and provide frameworks for measuring AI’s ecological impact. Examples include:
• ML CO2 Impact: A tool developed to help practitioners estimate the

carbon emissions of their AI experiments.
• Green AI: A movement advocating for the reporting of energy consumption and carbon emissions of AI systems in research papers.
• CodeCarbon: A Python library that helps track the energy usage and carbon emissions of code execution, particularly useful for AI development.

6. Sustainability Reports and AI Companies
• Major tech companies that develop AI systems (e.g., Google, Microsoft, Amazon) publish sustainability reports detailing their efforts to minimize the environmental impact of their data centers and AI infrastructure. They often invest in renewable energy and carbon offset programs to mitigate the footprint of AI systems.

Several AI systems and tools have already been developed to measure energy consumption, carbon emissions, and other environmental metrics, providing users and developers insights into their ecological footprint. Here are some of the most commonly used tools and resources, along with where you can find them:

1. Carbontracker Carbontracker on GitHub
• Carbontracker is an open-source Python package that helps estimate the energy consumption and carbon emissions of training deep learning models.
• Features: Provides energy usage and CO_2 emissions estimates based on the model type, number of epochs, and geographic location.

2. CodeCarbon CodeCarbon
• Description: CodeCarbon is a lightweight Python library designed to track carbon emissions based on energy consumption from cloud and on-premises hardware.
• Features: Tracks emissions during code execution, provides emissions data based on the energy grid of the region, and integrates with popular cloud providers.

3. Experiment Impact Tracker Experiment Impact Tracker on GitHub
• Description: Created by the Allen Institute for AI, Experiment Impact Tracker is a tool that logs GPU energy consumption and estimates carbon footprint during deep learning experiments.
• Features: Tracks electricity usage and CO_2 emissions based on the hardware, allowing researchers to evaluate the ecological impact of their experiments.

4. ML CO2 Impact ML CO2 Impact on GitHub
• Description: ML CO2 Impact is a tool that lets users calculate the CO_2 emissions of

machine learning models based on their computational intensity and hardware specifications.
• Features: It factors in the carbon intensity of the data center’s location and provides comparisons with standard baselines.

5. Eco2AI Eco2AI on GitHub
• Description: Eco2AI is a tool that measures the ecological footprint of AI systems by tracking carbon emissions and electricity usage. It provides real-time insights into the environmental cost of running ML models.
• Features: It’s tailored for AI applications, giving accurate carbon footprint estimates and allowing users to compare the environmental impact of different configurations.

6. Green Algorithms Green Algorithms Calculator
• Description: Green Algorithms is an online calculator for estimating the energy usage and CO_2 emissions of various computational tasks, especially in scientific research, including AI workloads.
• Features: Takes input on hardware type, task duration, and geographic location to estimate energy and emissions.

7. Energy Efficiency Reporting in AI Papers
• Description: Some initiatives, like Green AI and NeurIPS’ energy efficiency standards, are encouraging researchers to report energy consumption and carbon emissions of their AI experiments in papers. While not a tool, it’s a reporting standard that drives transparency in the AI research community.
• Access: Publications adhering to these standards, often listed in AI research journals (e.g., NeurIPS, ACL).

Currently, there’s limited public data on the carbon footprint of major AI systems like ChatGPT, Claude, and others. While companies like OpenAI, Google, and Anthropic have expressed commitments to reducing their environmental impact, detailed and model-specific information on the carbon footprint of systems like ChatGPT or Claude is not fully disclosed. However, here’s what’s known:

1. High-Level Emission Estimates
• For some large language models (LLMs), researchers have made rough estimates based on known factors like hardware requirements, typical training durations, and energy use. For instance, training GPT-3 was estimated to emit hundreds of metric tons of CO_2, but OpenAI has not published exact figures.

2. Corporate Sustainability Reports

• Many AI companies, including OpenAI (partnered with Microsoft), Google, and Meta, have issued broader sustainability commitments. These reports outline overall carbon reduction strategies, investments in renewable energy, and goals for carbon neutrality. However, they typically do not provide specific carbon footprints for individual AI models or services.
• For example, Microsoft, which hosts OpenAI models on Azure, reports carbon neutrality and offers information about the energy sources for its data centers.

3. Indirect Tools
• While direct data is limited, some tools like CodeCarbon and Carbontracker allow users to estimate carbon emissions for general AI tasks on cloud platforms, which could provide rough estimates based on hardware and runtime assumptions.

4. AI-Specific Research
• Research initiatives, like those by the Allen Institute for AI or Green Algorithms, publish benchmarks and footprint estimates for various model types, though these aren’t specific to proprietary systems like ChatGPT or Claude.
However there are several documented cases where researchers and organizations have measured the carbon footprint of training AI models. These studies typically analyze large-scale models like deep learning systems, where training can involve substantial computational resources. Here are some notable examples:

1, ChatGPT
• Description: GPT-3, one of the largest language models, contains 175 billion parameters and required massive computational resources to train.
• Estimated Carbon Footprint:
• Training GPT-3 is estimated to have consumed 1,287 MWh of energy, resulting in 552 metric tons of CO_2 emissions. This estimate assumes training occurred in a location with an average grid carbon intensity of 0.43 kg CO_2 per kWh.

2. BERT by Google
• Description: BERT (Bidirectional Encoder Representations from Transformers) is a widely-used natural language processing model.
• Measured Carbon Footprint:
• A 2019 study from the University of Massachusetts Amherst estimated that training a large BERT model (base version) emitted 1,438 pounds (650 kg) of CO_2. For a larger BERT variant, emissions rose to over 6,000 pounds (2,700 kg).
• Training these models was equivalent to the carbon footprint of a roundtrip transcontinental flight.

3. DeepMind’s AlphaFold
• Description: AlphaFold is an AI system developed by DeepMind to predict protein structures, requiring substantial computational power.
• Measured Carbon Footprint:
• DeepMind reported that training AlphaFold 2 consumed 210 MWh of energy, resulting in around 96 metric tons of CO_2 emissions. However, DeepMind offset these emissions by using renewable energy.

4. Large Model Pretraining Analysis by Microsoft
• Description: Microsoft trained a large transformer model similar to GPT-3 using Azure cloud infrastructure.
• Measured Carbon Footprint:
• The energy usage was estimated at 1,120 MWh, resulting in approximately 450 metric tons of CO_2. Microsoft noted that the footprint was reduced by utilizing its carbon-neutral cloud services and renewable energy initiatives.

5. General Study on AI Training
• A widely-cited 2019 study by Emma Strubell et al. measured the carbon footprint of training several popular NLP models and found:
• Training a single Transformer (big) model emitted around 284 metric tons of CO_2, equivalent to the lifetime emissions of five cars.
• Hyperparameter tuning (repeated training) could increase the emissions by up to 5 times.

Tools Used in These Measurements
• Most of these studies rely on:

1. Carbon Intensity Data: Regional carbon intensity values (kg CO_2 per kWh) are obtained from sources like ElectricityMap.

2. Energy Usage: Metrics from cloud providers or power consumption of GPUs (e.g., Nvidia or Google TPUs).

3. Software Tools: Tools like Carbontracker or Experiment Impact Tracker have been utilized in some of these cases.

To estimate the water usage associated with CO_2 emissions, requires multiple factors and assumptions. Generally, this involves understanding the water-energy-carbon nexus—the interrelated nature of water use, energy consumption, and carbon emissions.

1. Energy-Water Relationship

• Electricity generation often requires water for cooling, especially in traditional power plants (e.g., coal, natural gas, nuclear). The water usage per unit of energy generated can vary significantly depending on the energy source.

• For renewable energy sources like solar and wind, water use is relatively minimal compared to fossil-fuel-based power. For non-renewable sources, water use may be calculated in liters or gallons per kilowatt-hour (kWh) produced.

2. CO_2 Emissions and Energy Consumption

• By measuring the CO_2 emissions associated with AI systems or other processes, one can estimate the energy consumed, often reported as kWh. This energy estimate can then be translated into water usage, using benchmarks for how much water is needed to produce that specific amount of energy.

• For example, natural gas power plants may require about 3 liters (0.8 gallons) of water per kWh, while coal plants can require up to 4 liters (1.05 gallons) per kWh.

3. Conversion Process

• Step 1: Measure CO_2 emissions from a process, like training an AI model.

• Step 2: Determine the carbon intensity of the local electricity grid, which converts CO_2 emissions to energy used.

• Step 3: Multiply energy usage by the water-intensity factor for the specific electricity source, estimating water usage.

4. Example Calculation

• Suppose a process generates 100 kg of CO_2, and the local grid emits about 0.4 kg CO_2 per kWh (common in gas-powered grids). This means 250 kWh of energy was consumed.

• If the power source is a natural gas plant with a water usage rate of 3 liters per kWh, the total water usage would be: $250\text{kWh}\times 3\text{liters per kWh}=750\text{liters}$

5. Challenges

• The accuracy of these estimates depends on precise knowledge of the local grid's carbon intensity and the water intensity of specific energy sources, which can vary by location and power plant type.

Tools for Calculations

• There aren’t yet direct tools for converting CO_2 emissions to water usage in AI.

However, energy tracking and estimation tools like Carbontracker or CodeCarbon can provide a basis by first estimating the energy use, and then you can apply location-specific conversion rates for water consumption.

Revision #1
Created 29 December 2025 11:12:03 by Marklar
Updated 29 December 2025 11:18:05 by Marklar