NO PREP

The Cost of AI

Energy, Water, and Data Centers

Includes: Student reading, case study, differentiated questions, teacher key

Name: Period: Date:

The Cost of AI: Energy, Water, and Data Centers

Every Query Has a Price

When you type a question into ChatGPT and get an answer in three seconds, it feels weightless and free. But that three-second response required a real physical process involving hundreds of specialized computer chips, drawing significant amounts of electricity, in a large building that may be cooled by millions of gallons of water per year. AI is a physical technology with physical costs, and those costs are growing rapidly as AI adoption accelerates. A 2023 analysis estimated that a single ChatGPT query uses approximately ten times as much energy as a Google search query. Google processes about 8.5 billion search queries per day. ChatGPT was estimated to process over 10 million queries per day in early 2023, a number that has grown dramatically

Name: **Period:** **Date:**

The Cost of AI: Energy, Water, and Data Centers

Every Query Has a Price

When you type a question into ChatGPT and get an answer in three seconds, it feels weightless and free. But that three-second response required a real physical process involving hundreds of specialized computer chips, drawing significant amounts of electricity, in a large building that may be cooled by millions of gallons of water per year. AI is a physical technology with physical costs, and those costs are growing rapidly as AI adoption accelerates. A 2023 analysis estimated that a single ChatGPT query uses approximately ten times as much energy as a Google search query. Google processes about 8.5 billion search queries per day. ChatGPT was estimated to process over 10 million queries per day in early 2023, a number that has grown dramatically since. When you multiply these numbers across millions of users making thousands of queries each, the energy and infrastructure costs become enormous.

Data Centers: The Physical Reality of AI

AI computation happens in data centers, massive buildings filled with rows of servers and specialized chips called GPUs (graphics processing units) or TPUs (tensor processing units). A single modern data center can span 100,000 to 200,000 square feet, consume as much electricity as a small city, and cost hundreds of millions of dollars to build. The largest hyperscale data centers operated by companies like Microsoft, Google, and Amazon are even larger. AI has driven a major acceleration in data center construction. Microsoft alone announced plans to invest $80 billion in data center infrastructure in 2025. Google announced over $75 billion in capital expenditure for the same year, much of it directed toward AI infrastructure. Amazon Web Services, the world's largest cloud provider, announced

similarly massive investments. This buildout is happening globally, with new data centers under construction across the United States, Europe, and Asia.

The Energy Demand Problem

The energy demand from AI training and inference is significant and growing. Training a large language model uses enormous amounts of electricity: researchers estimated that training GPT-3 used approximately 1,287 megawatt-hours of electricity, equivalent to the annual electricity consumption of roughly 120 average American homes. Training more recent, larger models likely requires several times more energy. Beyond training, every query processed by an AI chatbot requires energy for inference, which is the process of running a trained model to generate a response. At scale, inference energy demand adds up. Goldman Sachs estimated in 2024 that data centers would account for 8% of US electricity consumption by 2030, up from 3% in 2022, with AI being a major driver of that increase. Multiple power grid operators have warned that AI data center growth could strain electrical infrastructure in ways not seen since the industrial revolution.

Water: The Cooling Crisis

Computer chips generate heat, and that heat must be removed to keep the chips running efficiently. Data centers use water cooling systems to manage heat, either evaporating water to cool air or using water in direct-contact cooling loops. The water consumption of AI data centers has become a significant environmental concern. A 2023 study from UC Riverside found that training GPT-3 required approximately 700,000 liters of fresh water for cooling. Microsoft disclosed that its global water consumption increased 34% between 2021 and 2022, attributing part of the increase to AI development. In regions already experiencing water stress, such as Arizona, where major data center clusters are being built despite desert conditions, the water demands of AI infrastructure raise serious sustainability questions.

Carbon Footprint and Climate Commitments

Every kilowatt-hour of electricity used by an AI data center has a carbon footprint that depends on how that electricity was generated. If the electricity comes from coal or natural gas plants, AI computation contributes to carbon emissions. If it comes from solar, wind, or nuclear power, the carbon footprint is much lower. This

makes the energy mix in a data center's location extremely important for its environmental impact. Major AI companies have made net-zero or carbon-negative commitments. Google committed to running on carbon-free energy by 2030. Microsoft committed to being carbon-negative by 2030. But AI's energy growth is complicating these commitments. Google reported in 2024 that its greenhouse gas emissions had increased 48% since 2019, citing AI infrastructure growth as a factor. The gap between stated climate goals and actual emissions is a growing tension in the AI industry.

The Nuclear Option

Faced with energy demands that renewable sources may struggle to meet quickly enough, some AI companies have turned to nuclear power. Microsoft signed a deal in 2024 to restart the Three Mile Island nuclear plant in Pennsylvania specifically to power its data centers. Google struck a deal with Kairos Power in 2024 to purchase power from small modular nuclear reactors being developed for deployment in the late 2020s. Amazon also announced agreements to purchase nuclear power for its data centers. Nuclear power produces no carbon emissions during operation and can provide a consistent, reliable power supply unlike solar or wind. Proponents argue it is the only energy source that can scale quickly enough to meet AI's growing demands without increasing carbon emissions. Critics note the challenges of nuclear waste disposal and the long timelines for building new nuclear plants. The emergence of AI as a driver of nuclear energy demand is a significant and unexpected development in the energy sector. The geographic concentration of AI data centers is creating regional energy and water stress that does not show up in national-level statistics. Northern Virginia, known as Data Center Alley, hosts more data center capacity than any other location on Earth and accounts for roughly 70 percent of the world's internet traffic flowing through its infrastructure. The region's electricity grid has been under significant strain, and local utilities have warned that they may struggle to meet projected data center demand within the next few years. Water-stressed states like Arizona, Nevada, and Texas are nevertheless attracting data centers with tax incentives, creating tensions between economic development goals and environmental sustainability. Waste heat recovery is one approach being explored to reduce data centers' net environmental impact. The computing equipment in a data center converts almost all electrical energy input into heat, which is then typically

exhausted into the surrounding environment. Some data center operators are exploring ways to capture this waste heat and use it productively: heating nearby buildings, greenhouses, or industrial processes. A few European data centers have already implemented waste heat recovery systems that supply hot water to district heating networks in surrounding communities. While waste heat recovery does not reduce data centers' energy consumption, it improves their overall energy efficiency by putting the heat to useful work rather than simply releasing it.

Key Information: AI Energy and Environmental Costs

Detail	Information
Energy per query	ChatGPT query uses ~10x the energy of a Google search
GPT-3 training energy	~1,287 MWh: equivalent to 120 homes' annual electricity use
GPT-3 training water	~700,000 liters for cooling (UC Riverside, 2023)
Microsoft data center investment	$80 billion planned for 2025
Google investment	$75+ billion capital expenditure planned for 2025
Projected data center share	8% of US electricity by 2030 (Goldman Sachs, 2024)
Microsoft water disclosure	34% increase in water consumption 2021-2022
Google emissions	48% increase since 2019, partly attributed to AI
Three Mile Island deal	Microsoft restarted the nuclear plant to power AI data centers (2024)
Nuclear appeal	No carbon during operation; consistent output unlike renewables

Case Study: The Data Center and the Desert Town

The town of Mesa View sits in a dry valley in the American Southwest. For years, the town council has been managing water allocation carefully: agriculture, residential use, and a small industrial zone all share a limited aquifer that hydrologists say is already being drawn down faster than it recharges. In the spring, a major tech company announced plans to build a large AI data center campus on the edge of town. The company promised 300 permanent jobs and significant tax revenue. The data center would also consume an estimated 3.5 million gallons of water per day for cooling, more than the entire town's current residential water use. The town council faced a genuine dilemma. The jobs and tax base were real and significant benefits for a struggling economy. But the water consumption would accelerate the aquifer depletion. A local rancher summarized the conflict at a public meeting: 'They're asking us to trade the long-term water security of this valley for jobs that might last twenty years before they automate those too. I'm not sure that's a trade I want to make for my grandchildren.'

1. Using what you learned about data center water consumption, explain why the town's water concerns about the AI data center are legitimate.

2. What factors should communities consider when deciding whether to approve large AI data center projects in water-stressed regions?

Multiple Choice

1. Approximately how much more energy does a ChatGPT query use compared to a Google search query?

a) About ten times as much

b) About twice as much

c) About five times as much

d) About the same amount

2. Why have major AI companies like Microsoft and Google made deals to purchase nuclear power?

a) Government regulations require all major data centers to exclusively use nuclear energy power sources

b) Nuclear plants are the only facilities large enough to physically house AI servers

c) Nuclear power is significantly cheaper than all other energy sources

d) Nuclear produces no carbon during operation and can provide consistent power at the scale AI demands

3. How does the energy source of a data center affect its carbon footprint?

a) The energy source makes no measurable difference, because every kind of AI computation emits exactly the same carbon

b) Data centers powered by renewable or nuclear energy have much lower carbon footprints than those using fossil fuels

c) All data centers use the same national energy grid, so their individual carbon footprints are identical

d) Carbon footprint only depends on the number of servers, not the energy source

Matching

Write the letter of the correct definition on the line next to each term.

Term	Definition
1. ___ GPU	A) A very large data center operated by major tech companies to handle massive computing loads
2. ___ Inference	B) A nuclear plant restarted by Microsoft in 2024 specifically to power AI data centers
3. ___ Hyperscale data center	C) The process of running a trained AI model to generate responses, consuming energy at scale
4. ___ Net-zero commitment	D) Specialized computer chip used for AI computation in data centers
5. ___ Three Mile Island	E) A corporate pledge to balance carbon emissions with equivalent carbon removal

Fill-in-the-Blank

1. Training GPT-3 required approximately _____ liters of fresh water for cooling data center equipment.

2. Goldman Sachs estimated that data centers would account for _____ percent of US electricity consumption by 2030.

3. Microsoft's global water consumption increased _____ percent between 2021 and 2022, partly due to AI development.

Short Answer

Answer each question in 2-4 complete sentences.

1. Explain why every AI query has a physical environmental cost even though it happens entirely through a screen.

2. How does the energy mix used by an AI company's data centers affect its overall climate impact, and why is this complicated for companies with net-zero commitments?

3. What responsibility, if any, do individual AI users have for the environmental costs of the tools they use? Defend your position.