Environmental Considerations of Generative Artificial Intelligence

August 19, 2026

Understanding the true environmental footprint of generative artificial intelligence (GenAI) is notoriously difficult because tech providers are rarely required or incentivized to disclose their cost or energy metrics. This creates a black box where tracing carbon and water impacts is nearly impossible, as resource efficiency varies wildly across different regions and facilities (Xiao et al., 2025). This document has been developed to provide information about the impacts of GenAI on the environment and how daily use may be managed thoughtfully for those who use GenAI by requirement or preference. See ***Appendix A*** for a glossary of terms.

Resource Consumption and Components

Historically, data centers were built to handle predictable workloads like general computing, web traffic, and cloud storage. Today, however, we are witnessing a massive structural shift: by 2028, over 50% of all data center electricity in the United States will be driven exclusively by resource-heavy AI training and inference (Shehabi et al., 2024). This shift has led to downstream environmental effects that extend across the full physical GenAI lifecycle (Figure 1), from raw material inputs and manufacturing of hardware to the design and operation of data centers to support training and inference and finally the end of life of the hardware.

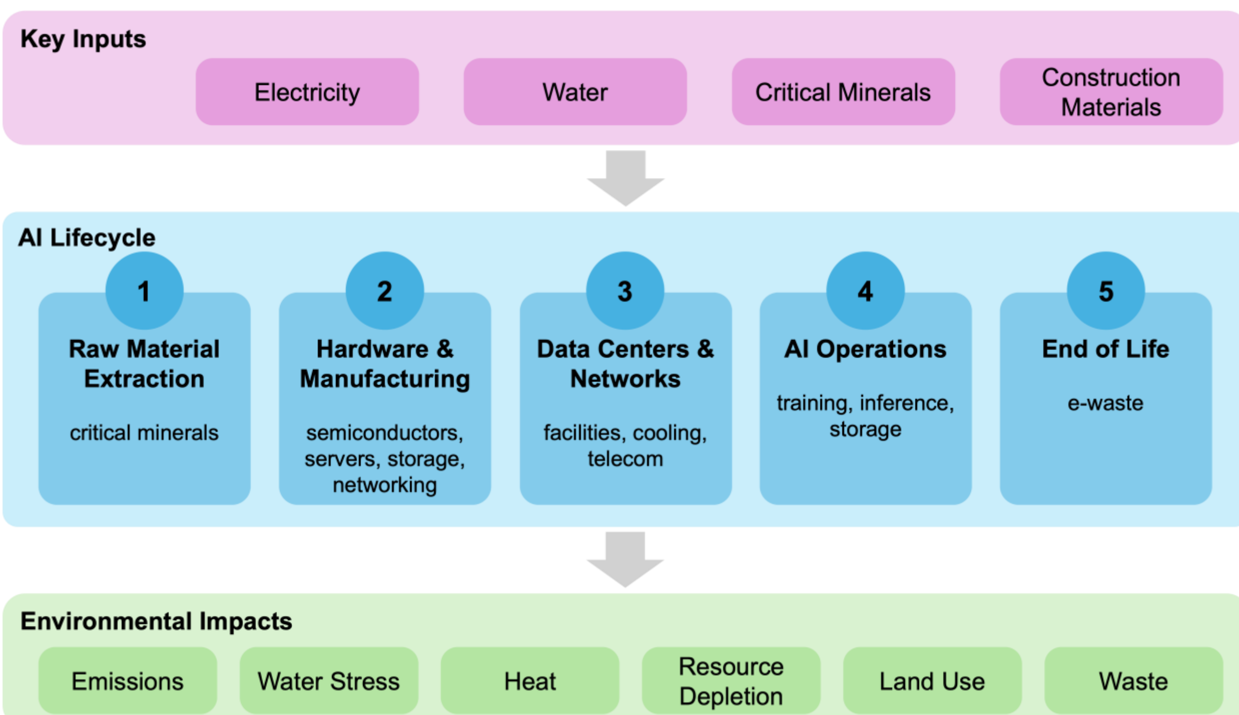

Figure 1. A schematic overview of the GenAI lifecycle including relevant environmental inputs and impacts.

GenAI infrastructure begins with **Raw Material Extraction** through mining of critical minerals and other raw materials required for the development of computing hardware and supporting infrastructure. The US relies on imports from other countries for many of these minerals (USGS, 2025). Extraction and upstream production require large amounts of energy, water, and finite resources, contributing to greenhouse gas emissions and resource depletion (Clemm et al., 2024; Dokic et al., 2024; Lambert & Luccioni, 2026).

The **Hardware and Manufacturing**, including that of specialized processors, memory, storage, servers, and networking equipment, is energy- and material-intensive. These embodied impacts include greenhouse gas emissions and resource consumption (Clemm et al., 2024; Dokic et al., 2024; Lambert & Luccioni, 2026; United Nations Environment Programme, 2024).

The **Data Centers and Networks** provide the facilities, cooling, power systems, and connectivity needed to house computing and storage hardware and the networks that enable the movement and processing of data. The specific needs differ widely across data centers depending on their construction and operation, but broadly require substantial physical space, materials, electricity, and water, contributing to land-use pressures, emissions, water stress, and heat generation (Clemm et al., 2024; Mytton, 2021; Rojahn & Grum, 2026; Shehabi et al., 2024):

- **Building Infrastructure:** Closed loop water cooling requires a single input of water initially, but much greater demands on energy to cool that water, while evaporative cooling systems result in larger water demands but require less energy, while leaving behind concentrated dissolved solids and treatment chemicals discharged as cooling-tower blowdown**.** This may require wastewater treatment before being released to the environment. (Baird, 2026; EPA, 2017; US EPA, 2023)
- **Hardware:** Modern cyberinfrastructure is more efficient than previous generations, leading to reduced energy and cooling demands for the same computation, however frequent replacement leads to generation of e-waste (detailed below).
- **Geography and Regional Factors:** The availability of renewable energy and water drive the overarching impact of a data center; facilities built in arid or fossil-fuel-dependent states face especially severe environmental viability challenges, but many data centers require additional energy resources beyond what is currently available through the grid (Xiao et al., 2025)

Once the physical infrastructure and data are available, **AI Operations**, which include training, inference, and data storage, consume electricity and generate heat that must be managed through cooling systems. Their impacts include greenhouse gas emissions, heat, and water use, with severity shaped by workload scale, facility efficiency, location, and the regional electricity mix (International Energy Agency, 2025; Mytton, 2021; Oliveira et al., 2026; Rojahn & Grum, 2026; Shehabi et al., 2024).

- **Training:** This initial phase is where the GenAI model is fed enormous amounts of data that must be ingested and stored prior to being processed for months to build the foundation of the model. This is a highly time- and energy-intensive process. It is estimated that training of OpenAI's GPT-4 required more than $100M and 50 gigawatthours (O’Donnell & Crownheart, 2025), which is equivalent to the annual energy consumption of more than 4,500 American households.

- **Inference:** This is where most of the day-to-day interactions with GenAI exist. Each large language model (LLM) prompt leverages the training to generate a response. The impact of each query is minimal, however billions of queries are run globally each day. On a micro-level, a simple text query draws a very small amount of energy (approximately 114 Joules), roughly equivalent to briefly running a microwave for a second, watching a few seconds of TV, or riding an e-bike for a few feet (O’Donnell & Crownheart, 2025). But generating an image spikes energy use by 20 times, a large reasoning model query by 60 times, and producing an AI video demands by up to 14,000 times that amount (O’Donnell & Crownheart, 2025).

Finally, given the accelerated development of new hardware to support GenAI, the fast technology **End-of-Life** timeline creates a major environmental impact. Retired processors, servers, storage devices, and networking equipment enter the e-waste stream. Their primary impacts include e-waste generation, loss of recoverable materials, and environmental risks associated with disposal, making responsible reuse and recycling important mitigation strategies (Lambert & Luccioni, 2026; United Nations Environment Programme, 2024).

Action

Our collective digital choices shape Utah's environmental landscape:

- **Choose GenAI Intentionally:** GenAI is not an effective solution for all tasks, so use GenAI only where it is genuinely provides value, rather than treating it as a default assistant for everyday tasks and entertainment.
- **Be a Critical Consumer:** Educate yourself on the environmental costs of different models. The footprint of some providers' infrastructure is far worse than others who utilize more renewable energy sources.
- **Advocate Locally:** Advocate for university, local, and state policies that mandate strict water-minimums and clean-energy requirements for any data centers operating within Utah's borders.
- **Establish Personal GenAI Principles:** Proactively draft a personal code of conduct consisting of several succinct, actionable statements to govern your daily engagement with GenAI. Use these commitments to actively weigh factors like data integrity alongside broader societal impacts, such as the tool's environmental footprint and human labor costs, creating a concrete benchmark to guide your choices and anchor collaborative conversations around responsible technology use. See ***Appendix B***.

Compiled by Penny Atkins, Jennifer Macali, Seneca Moore, Brandon Patterson, with input from colleagues across the University of Utah. Please direct inquiries to seneca.moore@nurs.utah.edu.

Appendix A

Glossary of Terms

- **Electronic Waste (E-Waste):** Discarded electronic equipment, servers, and processing components. Rapid hardware turnover in AI creates risks around managing hazardous components and recovering critical minerals at the end of equipment lifecycles.
- **Evaporative Cooling System:** A cooling method that relies on the evaporation of water to cool data center air, resulting in high overall water usage but lower direct electrical demand for cooling.
- **Closed-Loop Water Cooling:** A cooling infrastructure setup that recycles an initial volume of water internally, reducing ongoing water consumption but requiring significant electricity to chill the recirculating water.
- **Data Center:** A dedicated facility designed to house computing equipment, servers, and storage systems with continuous access to high-voltage electricity, thermal cooling, physical security, and high-capacity networking.
- **Generative Artificial Intelligence (GenAI):** Machines simulating human intelligence to perform tasks like problem-solving, speech recognition, and decision-making.
- **Large Language Model (LLM):** A type of AI trained on vast amounts of text data, capable of understanding and generating human-like language responses.

Appendix B

Establishing Personal AI Ethics

Objective
The goal of this activity is to establish personal principles for your engagement with Generative Artificial Intelligence (GenAI). By articulating a personal code of conduct, you will prepare to balance the convenience of AI tools with their real-world costs, focusing particularly on environmental efficiency, resource management, and digital citizenship.

Overview
Artificial Intelligence (AI) is a transformative force across modern society, but its rapid adoption comes with hidden physical costs. Every query, image generation, and automated workflow relies on intensive infrastructure that consumes significant energy and water resources.

While AI can streamline our tasks, users remain personally accountable for *how* and *when* they deploy these tools. This exercise asks you to define how you will navigate the trade-offs between technological utility, personal efficiency, and environmental sustainability.

Task Instructions
Develop a few principles that will govern your choices when using AI tools in your work, study, or daily life. These principles should reflect your commitment to intentional technology use, environmental awareness, and digital accountability. Consider how factors such as resource intensity, computing scale, data integrity, and human labor guide your decision to use (or skip) an AI tool.

To assist you, example principles are provided below. You may use these as models or develop your own standards as they apply to your specific knowledge and expertise.

Example Principles

- *"I will treat AI as a specialized tool rather than a default assistant. I will choose traditional, low-energy methods (like standard web searches or manual drafting) for basic tasks and reserve AI tools only for complex problems where they offer distinct value."*
- *"I will weigh the resource intensity of my prompt before generating output. I will opt for efficient, text-based queries over high-consumption media outputs—such as image or video generation—unless visually necessary."*
- *"I will prioritize using AI tools and providers that prioritize transparency around their energy sources, infrastructure, and carbon footprint. I will include environmental impact as a explicit benchmark when selecting software for my personal or professional workflows."*

References

Baird, L. L., Addy. (2026, January 21). Can you build data centers in a desert without draining the water supply? Utah is finding out. *Grist*. https://grist.org/technology/utah-data-center-water-supply-meta-novva/

Clemm, C., Stobbe, L., Wimalawarne, K., & Druschke, J. (2024). Towards Green AI: Current Status and Future Research. *2024 Electronics Goes Green 2024+ (EGG)*, 1–11. https://doi.org/10.23919/EGG62010.2024.10631247

Dokic, D., Groen In'T Woud, F., & Maass, W. (2024). *Towards Sustainability of AI: A Systematic Review of Exisiting Life Cycle Assessment Approaches and Key Environmental Impact Parameters of Artificial Intelligence*. https://doi.org/10.24251/HICSS.2024.518

EPA. (2017). Water Efficiency Management Guide Mechanical Systems. *Mechanical Systems*.

International Energy Agency. (2025). *Energy and AI* [Special Report]. International Energy Agency. https://iea.blob.core.windows.net/assets/de9dea13-b07d-42c5-a398-d1b3ae17d866/EnergyandAI.pdf

Lambert, K., & Luccioni, S. (2026). From Cradle to Cloud: A Life Cycle Review of AI's Environmental Footprint. In *Proceedings of the 2026 ACM Conference on Fairness, Accountability, and Transparency* (pp. 518–540). Association for Computing Machinery. https://dl.acm.org/doi/10.1145/3805689.3812401

Mytton, D. (2021). Data centre water consumption. *Npj Clean Water*, *4*(1), 11. https://doi.org/10.1038/s41545-021-00101-w

O'Donnell, J., & Crownheart, C. (2025, May 20). *We did the math on AI's energy footprint. Here's the story you haven't heard.* MIT Technology Review. https://www.technologyreview.com/2025/05/20/1116327/ai-energy-usage-climate-footprint-big-tech/

Oliveira, A. P., Carraquico, T., & Martinez-Perez, C. (2026). Beyond Efficiency: A Systematic Review of Energy Consumption and Carbon Footprint Across the AI Lifecycle. *Sustainability*, *18*(3), 1359. https://doi.org/10.3390/su18031359

Rojahn, M., & Grum, M. (2026). Green artificial intelligence and its lifecycle, hardware, and measurement dimensions. *Information and Software Technology*, *198*, 108186. https://doi.org/10.1016/j.infsof.2026.108186

Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2024). *United States Data Center Energy Usage Report* (LBNL--1005775, 1372902; p. LBNL--1005775, 1372902). https://doi.org/10.2172/1372902

United Nations Environment Programme. (2024). *Artificial Intelligence (AI) End-to-End: The Environmental Impact of the Full AI Lifecycle Needs to Be Comprehensively Assessed* [Issue Note]. United Nations Environment Programme.

US EPA, O. (2023, October 23). *Water Reuse Case Study: Quincy, Washington* [Overviews and Factsheets]. https://www.epa.gov/waterreuse/water-reuse-case-study-quincy-washington

USGS. (2025, July). *Key Minerals in Data Centers Infographic | U.S. Geological Survey*. Mineral Resource Program. https://www.usgs.gov/media/images/key-minerals-data-centers-infographic

Xiao, T., Nerini, F. F., Matthews, H. D., Tavoni, M., & You, F. (2025). Environmental impact and net-zero pathways for sustainable artificial intelligence servers in the USA. *Nature Sustainability*, *8*(12), 1541–1553. https://doi.org/10.1038/s41893-025-01681-y