LA&S Senior Seminar
Literature Review
23 September, 2025

Literature Review #2: A review of clean energy progress, and its relevance to AI-driven date center development.

Clean energy has, since the 1970s, been an issue in the broader public sphere, thanks to changes in global perception of the planet and the attention shone upon disastrous ecological events. Much progress in attacking environmental problems in equal effort to conventional ones has been made in the half-century since. Yet, there are many emerging environmental issues which demand a clear understanding still. Specifically, we have witnessed AI efforts rise to the top of the political and business agenda, which, coupled with national moves to consolidate hardware availability, has led to a widespread effort to establish more and more data centers to capitalize on the trillions of dollars at stake. As even a modest data center can consume energy equivalent to thousdands of homes, it is urgent that we understand our present energy situation, and how things will change going into the future.

Data centers rarely, if ever rely on isolated power generation (i.e. their own generators). Instead, data centers traditionally rely on the wider grid to keep the lights on. This can be problematic, as data center spaces can consume many times as much energy as a traditional office space. According to the International Energy Agency (IEA), data centers accounted for about 1.5% of the global electricity consumption in 2024. Those in charge have historically looked at this optimistically, postulating that data centers therefore have tremendous power to shape the future of the energy sector and barter for greener solutions. In this respect, there are many important statistics, but this article will emphasize two areas: greenhouse gas emissions, air quality & human health. For the first, since 2010, “efficiency measures avoided energyrelated carbon dioxide (CO_2) emissions equivalent to nearly 20% of the global total in 2023", stated to be much more than the entire European Union and India combined (IEA, 2025). Improvements in energy efficiency reduce overall fuel consumption, which lowers emissions of pollutants such as sulfur dioxide and fine particulates (IEA, 2025). At the same time, the increased use of natural gas and its by-products, natural gas plant liquids (NPGLs), has further improved air quality by displacing coal in power generation and industrial applications. NGPLs and natural gas emit far less sulfur, mercury, and particulate matter than coal, leading to measurable gains in regional air quality, though careful management of methane leakage remains important.

Despite these improvements, which come as a result of greater efficiency in the energy sector, non-renewable energy sources still dominate in the United States. Coal is the sole non-renewable energy source to have reduced since 2010, falling from >5 PWh to almost 3 in 2024. Natural gas, crude oil, and NPGL have all risen to at least double, representing a combined 22 PWh annually. Renewables, on the other hand, have not been expanded much, with nuclear, hydro-electric, geo-thermal, solar, wind, and biomass sources representing <5 PWh of production in 2024 (EIA 2025). This serves to contrast the argument that data centers offer a strongarm to barter for greener solutions alone. This is a rather troubling conclusion, as breakthroughs in AI have encouraged trillions in industry investment, largely being put towards data center construction. It is predicted that by 2030 data center power consumption will double globally, up to 3% of worldwide energy consumption. This represents an exponential growth, meaning that energy problems will become increasingly drastic if not addressed. Given insufficient growth in green energy capacity in the last decade, and a move towards non-renewable sources incentivized via tax breaks by recent U.S. tax law, it appears that there is a much more promising to encourage efficient data center design, and to consider areas for improvement in this (One Big Beautiful Bill Act, Pub. L. No. 119-21).

Data centers are highly nontrivial spaces to design. The critical nature of their operation necessitates reliability and power density capacity above all else, often weighted higher than efficiency is. Moreover, short design cycles can lead to strategies which merely work well enough without regard for environmental issues. This has prompted the U.S. Department of Energy's Federal Energy Management Program (FEMP) and the National Renewable Energy Laboratory (NREL) to set forth non-regulatory guidelines for the efficient design of data centers. They list the following parameters as most significant efficiency concerns, in order of importance: optimization of "Power Usage Effectiveness" (PUE), Reduce head to achieve the lowest "Energy Reuse Effectiveness (ERE) metric possible, and reject as much remaining heat to dry coolers as possible to optimize "Water Usage Effectiveness" (WUE). PUE measures total facility power divided by IT equipment power A score of 1.0 indicates zero overhead, while typical data centers score around 1.41 but varies heavily depending on the data center. For instance, hyperscale-class data centers typically score just under 1.15, but enterprise-class data centers often score around 2.0 (IEA 2025). Improvements to PUE simultaneously results in improvements to WUE, as higher compute efficiency directly reduces thermal loss. As data centers are built en masse, there is hope that these non-regulatory guidelines are more comprehensively enforced. Internationally, we can see this happening, with Bejing banning data centers with a PUE of 1.5 or less, meanwhile the Korean Government offers a 50% discount on the electricity facility levy to data centers built outside the Seoul metropolitan area. PUE is anticipated to increase as more technologies for ERE are made available, and projections expect PUE to decrease to 1.29 on average by 2030.

Within data centers, not all components contribute equally to efficiency (IEA, 2025). Rack servers are known to be the main factor for energy waste, and are the primary load, physically and energetically, in a typical data center. Server efficiency is defined as transactions per second per watt. Therein, an increase in performance by 50% means a server is 50% more efficient, and vise versa. As a consequence, hyperscale data centers are often more efficient due to this load optimization. Storage devices are the second most significant source of power consumption as far as electronics are concerned and is roughly linear to the number of storage modules used.

One of the largest potential areas for improvement in most data centers is air management and cooling. Air management refers to the design and configuration in minimizing the mixing of cooling air and rejected hot air. This is an under-reported metric, and both NREL and FEMP emphasize more data centers index this. Optimization of air management is a key factor in efficiency, and a leading reason for enterprise-class data centers wasting, on average, 3 times as much energy cooling than hyperscale data centers. Air cooling alone is often insufficient, even in the best case, and air conditioners are typically required to provide during high loads or poor weather. Retrofitted spaces largely make up the enterprise-class data centers, and may not have properly sized conditioning, which can lead to inefficient cycling and is another contributing factor to the large cooling costs. In dry regions, evaporative cooling has also been incorporated into data center operation. This offers a trade-off in huge efficiency gains (~80%) compared to air cooling but may worsen water availability in such dry regions.

The main conclusion to draw from data center design are that efficiency is largely tied to server efficiency and cooling efficiency. Improvements which improve one can also be found to improve the other, resulting in multiplicative gains in environmental benefit. This also grants some optimism towards enterprise-class data centers, which largely utilize repurposed office spaces and are highly prone to thermal losses, and pose to have the largest environmental improvements in this respect. Another reason to be optimistic is that this data-center scale improvements to heating and efficiency can be deployed much faster than grid-scale changes.

"Monthly Energy Review – August 2025." *U.S. Energy Information Administration*, Aug. 2025, https://www.eia.gov/totalenergy/data/monthly/.

Multiple Benefits of Energy Efficiency - Emission Reductions. International Energy Agency, 2014, https://iea.blob.core.windows.net/assets/7d547730-44ff-404b-af52-0812a340aee5/MultipleBenefitsofEnergyEfficiency-EmissionReductions.pdf. Accessed 20 Sep. 2025.

United States, Congress, House. "H.R.1 - One Big Beautiful Bill Act." *Congress.gov*, 4 Jul. 2025, https://www.congress.gov/bill/119th-congress/house-bill/1. 119th Congress.

Shehabi, Arman, et al. "United States Data Center Energy Usage Report." *Lawrence Berkeley National Laboratory*, 1 Jun. 2024, https://doi.org/10.2172/1372902.

"Total Energy Monthly Data." *U.S. Energy Information Administration*, https://www.eia.gov/totalenergy/data/monthly/. Accessed 9 Sep. 2025.

"Tracking Clean Energy Progress 2023." *International Energy Agency*, 12 Jul. 2023, https://www.iea.org/reports/tracking-clean-energy-progress-2023.

Van Geet, Otto, and David Sickinger. "Best Practices Guide for Energy-Efficient Data Center Design." *National Renewable Energy Laboratory*, 25 Jul. 2024, https://doi.org/10.2172/2417618.