www.sane-euratom.eu

SANE

Safety Assessment of Non-Electric uses of nuclear energy

The Potential Impact of Nuclear Energy on Data Centers

12.08.2025

Jussi-Pekka Ikonen, VTT

Disclaimer
Funded by Euratom Research and Training Programme.
The Associated Partner PSI is funded by the Swiss State Secretariat for Education, Research and Innovation (SERI).
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, Euratom Research and Training Programme.

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Funded by the European Union

Table of Contents

Abstract

The yet increasing electricity demand of data centers, which underpin the infrastructure of our digital society, are driving the search for carbon free energy solutions. Advances in artificial intelligence, cloud computing and cryptocurrencies may accelerate the growth even further. Data centers alone accounted for approximately 2% of global electricity demand in 2022, with projections suggesting this figure could rise to 5% by 2030. This increasing demand for data centers presents significant challenges for companies and governments aiming to meet ambitious decarbonization targets and climate goals.

Renewable energy has been expected to decarbonize and satisfy growing energy demand of data center industry. However, due to their intermittent nature, backup power with fossil fuels such as natural gas is required. Recent geopolitical events have caused drastic volatility and price hikes in oil and natural gas markets, exposing vulnerabilities in energy security, and thus, sparking interest in nuclear power and especially the novel small modular reactors. Nuclear power offers a stable, carbon-free baseload energy supply, making it an attractive prospect for the energy-intensive data center sector.

This study, based on a literature review, examines the evolving relationship between data centers and energy systems, with a particular focus on the integration of nuclear power. It explores the challenges of expanding data center capacity sustainably and evaluates how advanced nuclear technologies could support future growth while advancing global decarbonization goals and enhancing energy resilience.

1. Introduction

In recent years, data centers have become a critical part of the infrastructure that supports digitalization along with the electricity infrastructure that powers them, forming the backbone of our information society. Data centers house the servers and computing equipment required to store and process digital information, ensuring seamless operation of everything from cloud computing to social media. As the demand for storing and processing data increases, so does the need for sustainable and reliable energy sources. Nuclear energy, particularly with the arrival of Small Modular Reactors (SMRs), emerges as a promising solution to meet these growing energy demands while addressing environmental concerns.

Data centers are known for their huge energy consumption which is expected to grow in the future. Development in artificial intelligence (AI), internet of things devices and technologies related to cryptocurrencies can accelerate the growth even further. According to IEA, data centers, cryptocurrencies and AI consumed about 460 TWh of electricity in 2022 which is 2% of total global electricity demand. Furthermore, IEA has estimated that the electricity consumption could double towards 2026. (IEA 2024)

The rising energy demand and decarbonization targets are pushing major technology companies in the data center industry to explore nuclear power, especially novel SMR technologies. Sources of clean and reliable baseload power are greatly needed, making nuclear power with its dispatchable carbon free electricity a highly valued asset. Adaptation of nuclear power in the data center industry could work as a launchpad, accelerating commercialization of novel nuclear technologies, enabling its use in other areas also.

The power demand as well as emission reduction form critical constraints for data center expansion. Renewable energy has been expected to decarbonize and satisfy growing energy demand of data center industry. However, due to their intermittent nature, backup power with fossil fuels, most often natural gas is required. In 2022, drastic volatility and price hikes in oil and natural gas were introduced when the world’s second largest producer of natural gas, Russia, invaded Ukraine (Disavino 2022). During these times, the importance of energy security and dispatchable energy production were highlighted, directing countries and companies to study nuclear power in a new light. This study investigates the future prospect of data centers and how nuclear power could satisfy their needs through literature review.

2. Data centers and their energy consumption

Data centers play a central role in our digital age, supporting various online services, such as e-commerce websites, mobile phone applications, online banking systems, cloud services such as email and social media platforms amongst others. Data centers house computer systems that store, manage and distribute vast amounts of data, allowing users to access information quickly and securely through the internet. For example, when you upload an image to a social media platform, the image is stored in a data center, which allows it to be retrieved and displayed whenever someone views your social media profile. To run these operations in huge volumes, data centers require significant resources for computing power and storage capabilities. (Palo Alto Networks, 2025)

In data centers, both Central Processing Units (CPUs) and Graphics Processing Units (GPUs) play crucial roles in handling various computational tasks. While CPUs and GPUs share similarities, such as handing data and being silicon-based microprocessors, they have different architectures built for different tasks. CPUs are essential to all modern computing systems as they execute the commands and processes needed for your computer and operating system. CPUs are constructed from billions of transistors and can have multiple processing cores. A CPU is the primary processor of computer systems that focuses its fewer cores on individual tasks where latency or per-core performance is important, such as serial computing and running databases. (Intel n.d.)

As technology has advanced, GPUs have become essential components in data centers and computer systems in general. GPUs have many smaller and specialized cores, often in numbers of thousands. These cores are designed for parallel processing, delivering huge performance by dividing processing tasks across many cores simultaneously. This design approach makes GPUs more effective in tasks that involve processing large blocks of data in parallel such as solving cryptographic hash functions or more familiarly known as crypto mining, machine learning and running physics engines in various fields from video games to movies and scientific research. (Intel n.d.)

GPUs are often used in conjunction with CPUs to accelerate specific workloads and improve overall performance. For example, in AI model training, GPUs handle the intensive computation required for processing vast datasets and running complex algorithms, while CPUs manage the orchestration and coordination of these tasks. This combination of CPUs and GPUs enables data centers to achieve higher efficiency and faster processing times for demanding applications.

Ss GPUs and CPUs perform calculations and processing tasks, the electricity consumption for computing alone can be substantial in data centers. Computing not only consumes vast amounts of electricity but also generates significant amounts of heat, requiring cooling systems to maintain optimal operating temperatures. Consequently, cooling consumes another large portion of the electricity used in data centers. (IEA 2025)

GPUs and CPUs within data centers need to be connected and communicate with each other and other components in the data center. This networking infrastructure requires additional power to operate, contributing to the overall electricity consumption. As a result, the electricity consumption of data centers can be split into three categories: computing, which accounts for 40 % of the electricity consumed; cooling requirements, accounting for another 40 %; and the remaining 20% which is consumed by the networking infrastructure (IEA 2025).

Usually, data center cooling is based on combination of terminal cooling sub-system (TCSS) and mechanical refrigeration sub-system (MRSS) which includes chillers, pumps and cooling towers. The TCSS absorbs the heat generated during data center operation and transfers it to the MRSS where the heat dissipates to the environment. Depending on the chosen cooling method, air cooling, liquid cooling or free cooling technology can be used. Air cooling with air conditioning units and fans are used widely in existing data centers, however, liquid cooling and free cooling offer increased cooling effectiveness and higher energy efficiencies. Liquid cooling can be divided further into direct liquid cooling and indirect liquid cooling. In the direct liquid cooling, the computer system components are immersed in liquid for heat dissipation. In indirect liquid cooling, a coolant distributor (CD) working as the MRSS is connected to an external cooling source and the heat source. Cold liquid coolant flows from the cooling source to the CD, which delivers the coolant to the heat source and back to the CD, forming a loop. In free cooling, the cooling system uses cold external sources like air or water to enhance cooling, generally used from late fall to early spring in temperate climate zones. (Zhang, Meng, Hong et al. 2021)

Implementation of high-efficiency cooling systems could reduce about 10% of the data center electricity demand. Implementation of liquid cooling systems instead of air cooling could reduce the electricity demand even further. In a study by Hnayno et al. the electricity consumption of data center was decreased by 20% through direct liquid cooling technology: direct-to-chip water cooling system with passive single immersion cooling technology. (Hnayno, Chehade, et al. 2023)

It's also possible to utilize the waste heat from data centers. In Finland, data center operators pay lesser electricity tax if enough waste heat is extracted from data centers. This approach has decreased the costs to run a data center significantly as well as reduced use of fossil fuel in district heat generation. The Utility company operating in capital city of Finland, Helen, states that waste heat is one of the most effective ways to produce heat throughout a year (Juuti 2023). The capital city of Finland, Helsinki, has ambitious goal to prohibit burning from its energy production by year 2030. The European Commission has also revised Energy Efficiency Directive to mandate waste hear recovery applications in data centers when technically and economically feasible. (IEA 2024)

2.1 Artificial Intelligence

In recent years, the use of data centers has expanded into more complex and demanding applications, especially in the field of artificial intelligence. AI model training and inference have become central to the operations of modern data centers. Training AI models, also known as large language models (LLMs) like ChatGPT, require enormous computational resources. These models need to process and learn from vast datasets, which involve running complex algorithms on powerful GPUs. Once AI models are trained, they can perform a variety of tasks, from natural language processing to image recognition. Recent data from Meta and Google reveal that the training phase consumes roughly 20-40% of energy use while inference phase consumes proportionally much higher, around 60-70% of energy use (Kamiya & Bertoldi 2024).

Training advanced AI models like GPT-4 involves the use of millions of GPU hours. Training the GPT-4 model required over 50 GWh of energy, which is fifty times more than what was needed for training its predecessor, GPT-3 (Cohen 2024). Additionally, water consumption has been found alarmingly high for AI model training. A study by Li et al. reveals that the training of GPT-3 in Microsoft's advanced US data centers evaporated 700,000 liters of clean freshwater.

In the inference phase, trained AI models generate responses or make predictions, which also demands significant computational power. For example, a single Google search using an LLM like the ChatGPT can consume nearly ten times the electricity of a traditional search engine query. In numbers, a Google search using ChatGPT consumes 2.19 Wh of electricity while a traditional search consumes 0.3 Wh. (Singer, Bingham et al. 2024). As a conclusion, LLMs and AI-related applications require more and more computing power, consuming increasing amounts of electricity and furthermore, the use of AI across industries is expected to increase significantly in the future.

However, it's quite possible that efficiency of AI model training and operation will increase in future. Recently, the Chinese company DeepSeek has shocked datacenter industry with claims of training a LLM on par with ChatGPT with a fraction of costs and computing power. Chinese AI companies have restricted access to chips manufactured in the US. The limited access to computing power has led Chinese companies to new innovations in the field, prioritizing efficiency and embracing open-source principles.

In a technical report, DeepSeek revealed that full training of the model DeepSeek-V3 required only 2.788M H800 GPU hours. The difference is huge when compared to training other LLM. For example, full training of OpenAI's GPT-4 required about 60 million GPU hours with Nvidia A100 GPUs which are relatively stronger. (Deepseek-AI, Liu, Feng et al. 2024)

DeepSeek finetuned the V3 model with reinforcement learning. DeepSeek-R1-Zero employs a "chain of thought" approach similar to OpenAI which lets the LLM to solve problems by processing queries step by step. Upon an input given by user, the DeepSeek-R1-Zero model first produces a reasoning process, followed by the final answer. DeepSeek's approach resulted in a huge reduction in the training process of the LLM while keeping up with just as good or even better performance when compared to OpenAI's LLMs. See Figure 1 below. (DeepSeek Ai, Guo, Yang et al. 2025) (Chen 2025)

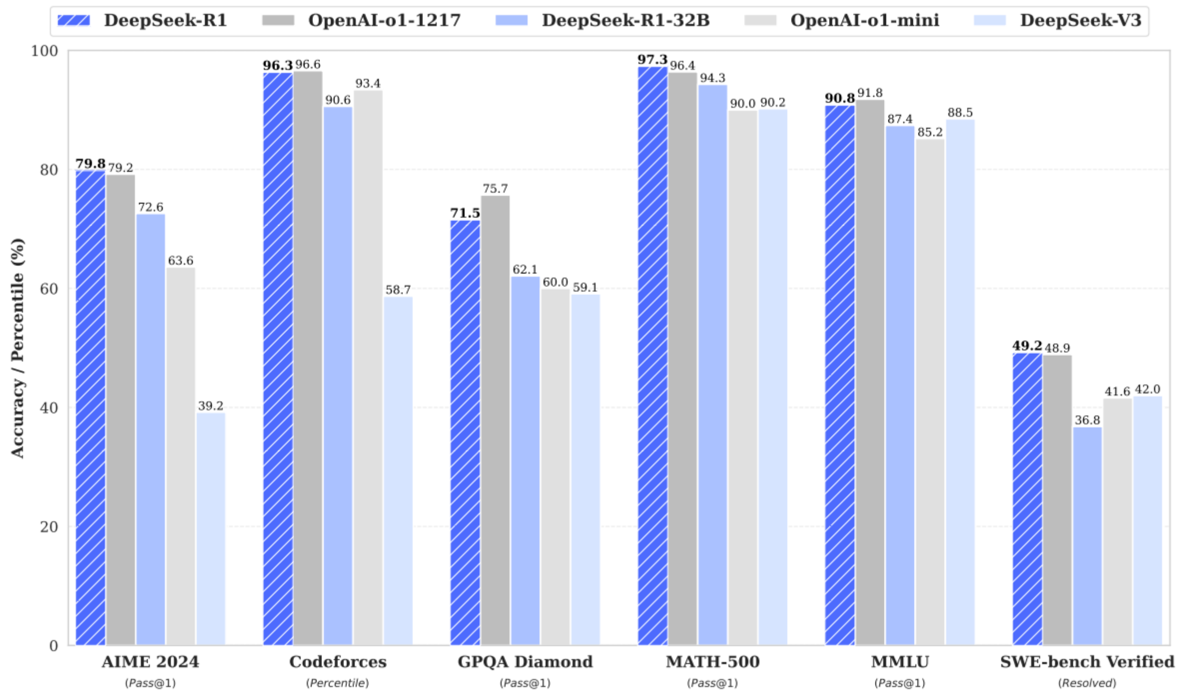

Figure 1. Comparison of DeekSeep-R1 performance to other LLMs. (DeepSeek Ai, Guo, Yang et al. 2025)

Improvements in the efficiency of training and using LLMs could potentially reduce global electricity consumption. Additionally, the entry barrier into the AI market might be lower than expected, possibly resulting in more companies entering the industry and further development of AI models. The key question is whether demand for AI development and usage is constrained by budget, computing demand, or remains unconstrained.

Goldman Sachs, in their equity research, notes in their equity research that there is corporate optimism that AI innovations could drive significant efficiencies which would mitigate power demand and emission growth in the data center industry. However, the research reveals that efficiency gains would not lead to reduced need for data center infrastructure based on their recent corporate dialogue. The research continues saying that the more likely scenario would be in the unconstrained side until there is greater clarity on the impact and demand of AI solutions. (Singer, B., Corbett, B., Davenport, C., et al. 2025)

2.2 Cryptocurrencies

Cryptocurrencies, such as Bitcoin and Ethereum, are another new component that contributes to the growing energy demands of data centers. Cryptocurrencies are created and traded without the involvement of a central authority, using innovative blockchain technology. Blockchain is a decentralized digital database that stores "blocks" of data linked together to form a "chain." These blocks are stored in chronological order, with each block containing transaction data, a timestamp, and a cryptographic hash function of the previous block. (Campbell 2024)

A cryptographic hash function, or simply a hash, refers to mathematical functions solved by a network of computers, known as nodes, in the blockchain. Crypto miners, a subset of these nodes, are essential for the security of blockchain technologies, using the Proof of Work (PoW) mechanism. In PoW, crypto miners repeatedly solve hashes until other systems in the blockchain verify that the solution calculated by the crypto miner is correct and valid. Once verified, the crypto miner receives a payment in the form of new cryptocurrency. The payment is composed of block reward and a gas fee that the user pays. Crypto miners use advanced computer systems for high computing power, consuming significant amounts of electricity in the process and essentially converting it into money. As a note, cryptocurrency mining data centers host specialized mining hardware that is different to traditional and AI data centers. (Chandler & Campbell 2024)

There is an alternative to the PoW mechanism for creating new blocks in the blockchain. By adopting the Proof of Stake (PoS) mechanism, Ethereum, the second-largest cryptocurrency by market capitalization, has reduced its power demand by about 99.9%. In a blockchain operating under the PoS mechanism, the continuous operation of energy-intensive computer systems is not needed, and emissions from running the blockchain are almost completely eliminated. Instead, validators are selected based on the amount of cryptocurrency they own and are willing to "stake" as collateral in the staking process. In Ethereum, at least 32 units of the native cryptocurrency are required to participate in the staking process. The likelihood of being selected as a validator increases with the amount of cryptocurrency staked. (De Vries 2023)

2.3 Data center types and operation profile

Data centers are typically classified by their purpose or business model. Enterprise data centers are operated by companies to deliver and manage their services to the employees and customers. Colocation data centers can house multiple companies who own networks, servers and storage equipment in the facility. The colocation data center operator owns the space and manages power and cooling to the companies using the facility. Colocation data centers usually offer interconnection to Software as a Service (SAAS) products such as Salesforce or Microsoft Azure. Furthermore, cloud and hyperscale data center describe a large data center that is owned and operated by cloud service producers such as Amazon Web Services (AWS), Microsoft or other public cloud provider to solely deliver their SaaS or other cloud products to customers. Note, that the term hyperscale usually refers to the size of the data center rather than the ownership structure, thus, colocation data center can be defined as “hyperscale”. Power demand ranges from 2 MW for small data centers to up to 50 MW for colocation or cloud data centers while power range of hyperscale data centers can range between 50 MW and 300 MW. (Bilsen, Gröger, Devriendt et al. 2022)

Power demand of data center is usually quite steady and predictable throughout the year. Some load fluctuations may occur during work hours if the data center is connected to other company activities such as office spaces. Seasonal variations in data center power demand may be experienced as the cooling systems in data centers have higher efficiency during colder periods, resulting in slightly higher power demand during summer. In most cases, data center operation requires continuous, baseload electricity supply 24 hours per day and 365 days per year. Historically, data centers have been equipped with backup diesel generators to secure power reliability and be fully redundant. The requirement of never losing power is even higher in hyperscale and AI data center since they are so capital-intensive facilities requiring high utilization rate to maximize cost-effectiveness and profit for the data center operator. (Nøland, Hjelmeland & Korpås 2024a).

The importance of reliable and continuous baseload power for data centers is reinforced in a study by Nøland et al. The study points out that the cost of AI training is particularly sensitive to the load factor, which limits the incentives for load shifting. The levelized cost of computing (LCOC) of $1.01/EFLOP is achieved with a data center operating with a baseload electricity price of $100/MWH and a 100% load factor. Formulation of LCOC includes capital expenditure, operating expenditure, power consumption, and the total amount of computation during the lifetime of the infrastructure Alternatively, the same LCOC is achieved with a 69.10% load factor and a baseload electricity price of $0/MWH. As a result, continuous and reliable power production is extremely valuable for the competitiveness of data center operations, contrary to other power-intensive industries that are sensitive to electricity prices. The study findings reveal that AI and data center expansion could drive significant baseload power demand in the future energy systems. (Nøland, Hjelmeland & Korpås 2024a)

However, there are different views on data center flexibility. Especially regarding the cryptocurrency mining data centers which may exhibit a more flexible operation profile. Unlike latency-sensitive applications such as social media hosting or services related to AI and LLMs, cryptocurrency mining operations can be interrupted abruptly. Their operation is primarily economically driven; mining is prioritized when the electricity prices are low and paused when the electricity price is high.

A recent study by Hajiaghapour-Moghimi et al. introduces cryptocurrency mining loads as a novel form of virtual energy storage system, termed cryptocurrency energy storage systems (CESS). By converting surplus renewable electricity into cryptocurrency units such as Bitcoin, CESS enables long-duration, cost-effective energy storage. The study based on islanded and grid-connected microgrids in Finland, a region with high seasonal variability and significant renewable electricity production, reveals that integrating CESS reduced operating costs of the microgrid operator by 7.4% in islanded and 46.5% in grid-connected scenarios. The return on investment for the CESS reached 44% in the islanded and 51% in the grid-connected scenarios. These findings underscore that data center flexibility is highly dependent on the nature of services provided, and that economic incentives can enable deviation from a strictly continuous power load profile. Note that profitability of cryptocurrency mining is sensitive to both electricity price and the value of cryptocurrency. (Hajiaghapour-Moghimi et al., 2024)

3. State of Data Center deployment

Before the inception of AI, energy demand growth in data centers was quite flat. From 2015 to 2019, the workload of data centers tripled while the power demand remained the same. This was because data centers gained significant efficiency improvements, about 15% annually, due to expansion of cloud and hyperscale data centers. See Figure 2 for clarity.

Hyperscale data centers can run large-scale operations without a significant increase in electricity consumption due to economies of scale and specialized engineering. In 2022, cloud and hyperscale data centers handled more than 90% of the workload demand. As a result, the efficiency improvements from the shift to the cloud and hyperscale data centers have narrowed down to about 2%. Nonetheless, hyperscale data centers are expected to grow, and their global market is expected to double from 2023 to 2026 (Zoting 2024). (Singer, Davenport, Chang et al. 2024)

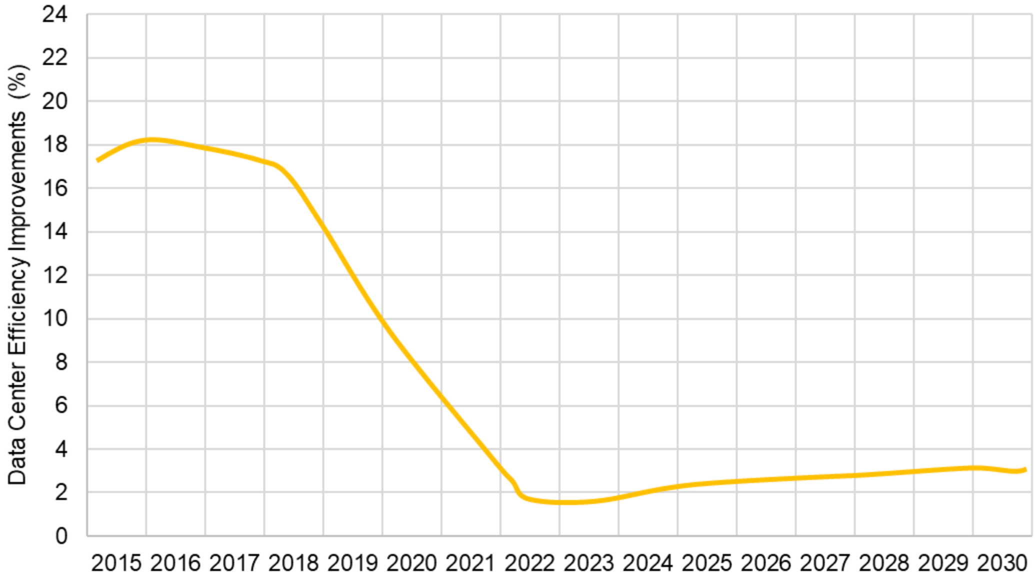

Figure 2. Data center efficiency gains as a function of time. (Singer, Davenport, Chang et al. 2024)

Technology used in the data centers and digitalization of industries and societies are evolving rapidly, so it's quite complex to estimate future trends of data centers. When considering the combination of rapid growth in computing demand for AI models as well as increases in the model sizes, it's very likely that the power demand outpaces efficiency improvements. As a result, the total AI-related energy consumption is expected to rise in the future. (Kamiya & Bertoldi 2024).

As of now, there are more than 8000 data centers globally with most of them, 33%, located in the United States, 16 % in Europe and almost 10% in China (IAEA 2024). The major technology companies: Microsoft, Amazon, Google and Meta have been expanding data center capacity rapidly. The combined electricity consumption of aforementioned companies more than doubled between 2017 and 2021 to about 72 TWh. In July 2024, Google reported that its greenhouse gas emissions rose 48% from the year 2019 due to increased electricity demand from its data centers and growth in AI (Siciliano 2025). Likewise, Microsoft reported that its emissions increased 30% since 2020 because of the indirect emissions from the construction of data centers (Kimball 2024a).

Artificial intelligence and cloud computing are anticipated to be the biggest growth drivers for data center expansion. Globally, IEA estimates the electricity consumption of data centers to grow from 460 TWh in 2022 to a range between 620 and 1050 TWh in 2026. The estimated range is quite wide due to uncertainties in pace of data center deployment, efficiency improvements of the technology as well as AI and cryptocurrency trends. The low and high case scenarios seen in Figure 3 illustrate these uncertainties. The huge energy demand in the data center industry creates a bottleneck for capacity increases as well as creating environmental challenges as there aren't easy solutions for carbon free electricity adapted to the operation profiles of data centers. (IEA 2024)

Equity research by Goldman Sachs on global power surge of AI and data centers estimates that the global data center power could grow more than 160% from 2023 to 2030, equivalent to top 10 power-consuming country. These estimates are slightly higher than the IEA estimations. The growth in data center demand is expected to be driven in part by AI and in part from deacceleration in efficiency gains of non-AI data centers. See Figure 4. (Singer, Corbett, Davenport, et al. 2025)

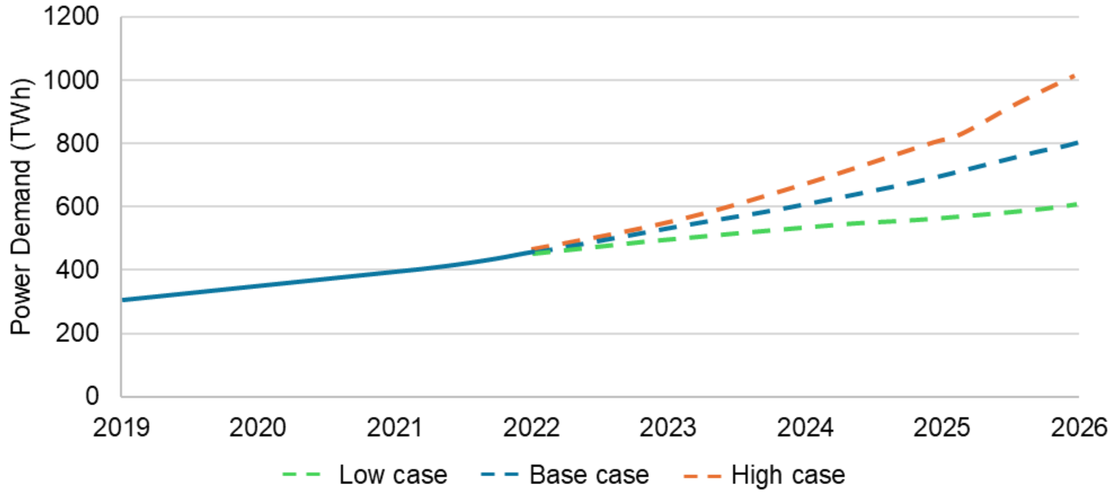

Figure 3. Global electricity demand from data centers, AI, and cryptocurrencies. (IEA 2024)

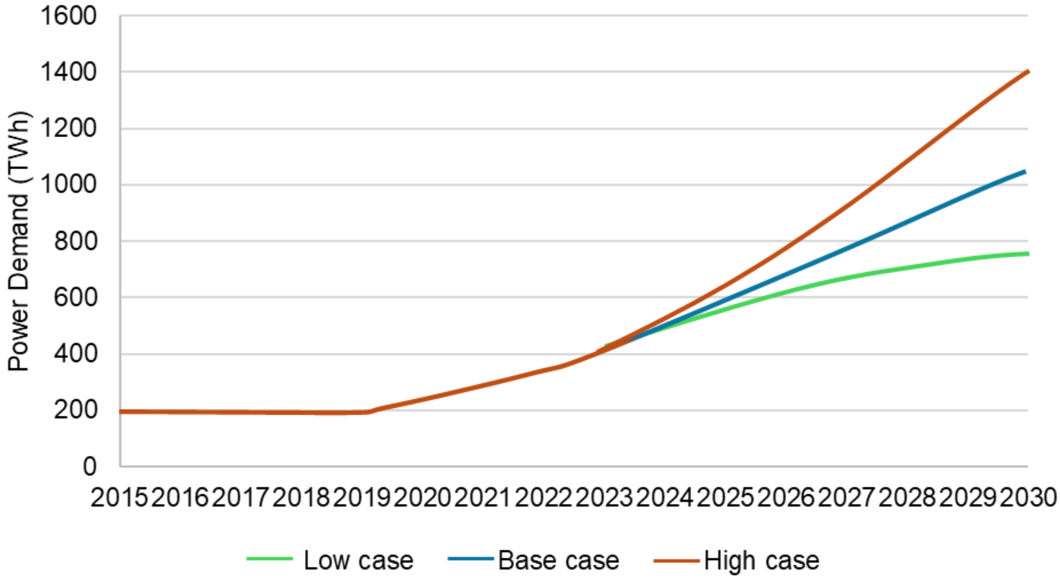

Figure 4. Goldman Sachs estimation of global electricity demand of data centers. Note that years after 2022 are projections of power demand. (Singer et al. 2025)

In the US, Electricity consumption of data centers is estimated to increase from 200 TWh (4% of national electricity demand) in 2022 to almost 260 TWh (6% of national electricity demand) in 2026. State Grid Energy Research Institute in China estimates electricity demand of data center sector to double to 400 TWh by 20230. In the European Union, electricity consumption of data centers was estimated to be slightly below 100 TWh (4% of total EU electricity demand) in 2022 which is expected to increase to 150 TWh by 2026. (IEA 2024)

Study by Kamiya & Bertoldi researches existing literature and public sources to estimate current and future data center power demand in Europe. There are several studies on data center power demand in Europe, but the results can get quite different because the data sources, assumptions and methodologies are not uniform. As seen from Figure 5, estimations of data center power demand range from 39.5 TWh to 104 TWh for year 2020. For the year 2030, the electricity demand is projected to grow between 98.5 TWh and 160 TWh. (Kamiya & Bertoldi 2024)

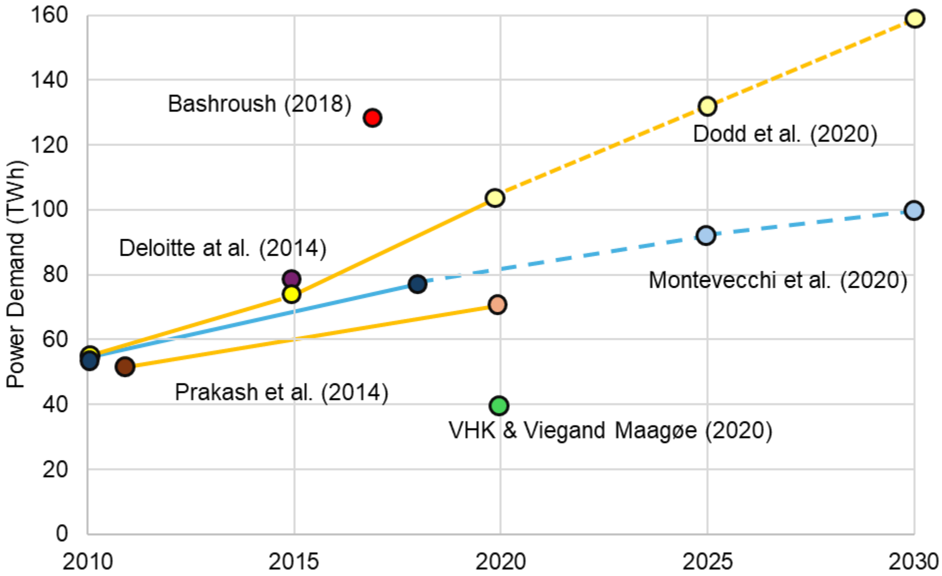

Figure 5. Summary of studies on electricity demand growth in European data centers. Note that the darker circles indicate estimates while lighter circles and dotted lines indicate projections. (Kamiya & Bertoldi 2024)

4. Integration of nuclear power and data centers

Sentiment towards nuclear power has improved in recent years. Renewables have been expected to satisfy future energy demand while reducing CO_2 emissions at the same time. Due to the intermittent nature of renewables, backup power with fossil fuels such as natural gas is needed to satisfy the energy demand continuously. In 2022, Russia's invasion of Ukraine caused volatile hikes prices of oil and natural gas (Disavino 2022). Issues with reliability of energy systems brought attention to the importance of energy security and dispatchable energy, motivating countries and companies to reconsider nuclear power. This sentiment culminated during COP28 in 2023, where more than 20 countries launched a declaration to triple global nuclear energy capacity by 2050, recognizing crucial role of nuclear power in achieving net-zero greenhouse gas emissions (U.S. Department of Energy 2023).

The power demand as well as emission reduction form critical constraints for data center expansion. Companies in data center industry are under huge pressure to find carbon free and reliable electricity production for their growing energy demand. Traditional nuclear power plants and especially the novel small modular reactors are gathering interest in data center industry. This chapter will explore potential of nuclear power for data centers while gathering information on the current expectations of the industry.

The equity research conducted at Goldman Sachs anticipates that renewables paired with energy storage will fulfil approximately 80% of data center power demand. However, the research points out that additional baseload generation is needed to meet the 24/7 data center demand. While nuclear power is preferred, Goldman Sachs sees natural gas power plants as a more realistic option for most data centers. The long lead times for building new nuclear power plants create challenges; however, similar issues might be present in new natural gas plants. Many US utilities estimate that a large gas turbine could be built as early as 2029 if the buyer enters the queue now. On the other hand, smaller peaking gas power plants can be built quickly, but they are less efficient and more expensive than the larger variants. The lack of baseload electricity generation could create a bottleneck for data center expansion. Natural gas power plants could alleviate this issue, but in the long term, they need to be replaced with carbon-free sources. Additionally, the competitiveness of natural gas power plants is limited to regions with cheap natural gas prices. In the future, rising carbon tax prices and the depletion of fossil fuel resources will further reduce the viability of natural gas plants. (Singer et al. 2025)

Even though the research estimates that most of the demand is fulfilled with renewables and energy storages, lots of new electricity capacity needs to be built to meet the data center power demand in the US. Roughly half of the power demand needs to be met with new capacity and natural gas plants are expected to be needed the most. In numbers, the research estimates that the added generation capacity will be around 40% natural gas combined-cycle gas turbines, 20% natural gas peakers, 25% solar and 15% wind. This would result in about 47 GW of incremental generation capacity in the US through 2030. On the long term, nuclear capacity is expected to increase in the 2030s. (Singer et al. 2025)

The research points out that the electricity grid will also require significant investments. Electricity transmission and distribution systems need to be developed to accommodate new power sources and data centers. The lack of development in this area could create another potential bottleneck for data center expansion if grid development is not handled proactively. Goldman Sachs expects transmission development projects to take several years just to permit, and another several years to build out (Singer et al. 2025).

Research by McKinsey points out similar issues. Power supply is becoming an issue in markets that have attracted data centers such as Northern Virginia. The report continues, saying that many utilities haven't been able to build out transmission infrastructure quickly enough, rising concerns about insufficient power generation. Some countries have brought a complete halt to data center expansion because of the strain on the electricity grid and the impact on national climate targets. For example, Ireland has stopped issuing new grid connections to data centers in the Dublin area until 2028. (Srivathsan, Sorel, Bhan et al. 2024)

Traditionally, data centers have been built close to population centers to have better connectivity and latency. Training of AI models changes this dynamic since low latency and network redundancy are not important. Thus, data centers specialized in AI model training can be built in remote locations, where the electricity grid is less strained, and power is abundant. However, McKinsey research concludes that these remote locations usually lack power transmission infrastructure, thus, power supply might become an issue as demand grows. (Srivathsan et al. 2024)

The McKinsey research has noticed that some data center operators are acquiring facilities built close to power plants to help overcome transmission issues. Furthermore, some companies have started generating their own off-grid power with fuel cells, batteries, and renewables. The research hints that SMRs could be an option in the long term. (Srivathsan et al. 2024)

4.1 Small modular reactors

Small Modular Reactors represent a novel technology in nuclear energy, attempting to mitigate the challenges such as huge budgets, long lead times and delays the traditional large-scale nuclear power plants are notorious for. SMRs are designed to be compact, modular, and versatile, making them ideal for providing off-grid baseload power. The compactness and modularity of SMRs allows them to be built faster and economies of serial production. Additionally, the smaller size of SMRs allows them to be built closer to industrial clusters and population. SMRs can operate independently from the main electricity grid ensuring continuous and reliable power generation, which is especially crucial for industries such as data centers that require uninterrupted energy supply. (IAEA 2024b)

One of the main advantages of SMRs is their scalability. Unlike conventional nuclear power plants, SMRs can be tailored to the specific power demands of niche markets such as data centers. The increased flexibility allows data center operator to deploy SMR modules proportionally to their energy needs, reducing the risk of overcapacity. If the data center is expanded, additional SMR modules can be added incrementally to align power generation with power demand of the facility. (IAEA 2024b)

Transportable SMRs are also being created for remote areas like the floating nuclear power plant (FNNP) Akademik Lomonosov developed in Russia. Launched in 2019, the Akademik Lomonosov serves as an example of how SMRs can be utilized to deliver off-grid baseload power. The FNNP is equipped with two KLT-40S reactors, capable of generating up to 70 MW of electricity and 300 MW of thermal energy. It primarily provides power to remote areas in Russia's Arctic region, demonstrating the potential of SMRs to supply energy to locations where conventional power infrastructure is either lacking or impractical. Floating nuclear power plants can be towed to the desired location, providing a flexible solution for power generation needs. This approach simplifies the transportation, installation and maintenance of the nuclear facility. Other FNNP designs are being developed in South Korea, United States and China. For example, Bandi-60 by Korea Electric Power Corporation and Compact Molten Salt Reactor by Samsung and UK-based Core Power. (IAEA 2024c)

The term SMR comprises numerous reactor designs which differ in their unit size, operation mode and maturity of the technology. Maturity in light water reactor (LWR) SMR designs range from conceptual designs (NuWard, BWRX-300, LDR-50) to certified design in the USA (NuScale) and operational SMRs such as the aforementioned KLT-40s reactors. New LWR SMRs are expected to generate electricity to grid by 2026 in China (ACP100) and by 2028 in Argentina (CAREM-25). China has also a modular pebble bed high-temperature gas-cooled reactor (HTR-PM) already in commercial operation since 2023. In the USA, Kairos Power began building fluoride salt-cooled high-temperature reactor (KP-FHR) for heat generation in 2024. The KP-FHR represents the first and only Generation IV reactor to receive a construction license from the U.S. Nuclear Regulatory Commission. (IAEA 2024b)

Large-scale nuclear power plants are notorious for their huge budgets and delays in construction times. Despite the advances in SMR technologies, it is still uncertain when SMRs are expected to be ready for extensive commercial operation as most technologies are under development. Most SMR designs use novel passive safety systems which potentially decrease expenses but also introduce regulatory uncertainties. Conventional nuclear power plants use active safety systems which rely on active driving devices such as electrical or diesel motors while the passive safety systems are based on natural forces such as gravity and natural heat convection. Safety demonstrations of passive safety system are still under development, and nuclear regulators worldwide might be cautious of licensing these systems. This is a national issue, each country has different legislation and regulations, adding an extra challenge. However, there are positive signs in Finland already: the Radiation and Nuclear Safety Authority of Finland (STUK) has begun familiarizing itself with SMR technology, recognizing that the safety of SMRs with low power levels could be fulfilled through passive safety systems assessed on a case-by-case basis (STUK 2024).

SMR designs are claimed to be more cost-efficient than large nuclear power plants. However, the true impact coming from modularization and serial production has not been proven yet until the first SMRs are being operated. A study by Nøland et al. about SMR economics through a model-based analysis reveal that SMRs may initially face higher costs, however, if enough SMR units are deployed, they can match and even exceed cost-efficiency of large nuclear power plants due to reduced construction time and learning effect. (Nøland, Hjelmeland & Korpås 2024b)

4.2 Nuclear power deals in data center industry

As data centers and other power-intensive industries continue to expand, the demand for reliable and sustainable energy sources grows. SMRs, with their high-capacity factor and ability to operate off-grid, present a promising option for meeting this demand while contributing to the reduction of carbon emissions. Despite the challenges nuclear power industry is known for, the main actors in the data center industry have identified the potential of nuclear power. Google, Microsoft and Amazon have all announced plans to invest in nuclear power to fuel the growing electricity demand of data centers.

Starting with the deal between Microsoft and Constellation Energy, Three Mile Island nuclear power plant is planned to restart through a 20-year agreement to power Microsoft's data centers. Constellation Energy plans to invest 1.6 billion dollars in restarting the plant through 2027, stating that the agreement with Microsoft as largest power purchase agreement that nuclear power plant operator has ever signed (Kimball 2024b). Then, Google made an agreement with the SMR developer Kairos Power to purchase power from a fleet of SMRs which is estimated to add 500 MW to the grid (Stevens 2024). The first reactor is expected to come online by 2030, with more reactors going live through 2035.

Continuing with Amazon, its subsidiary specializing in cloud computing, Amazon Web Services (AWS), is investing more than 500 million dollars in nuclear power, announcing three projects from Virginia to Washington state. AWS has signed an agreement with Virginia's utility companion, Dominion Energy, to explore the development of SMRs. Virginia is known for the huge data center hub known as Data Center Alley, housing nearly half of all the data centers in the US. New SMRs are expected to add 300 MW of power to Virginia. Also, Amazon made an agreement with utility company Energy Northwest to fund the development, licensing and construction of four SMRs in Washington State. Amazon has the right to purchase electricity from the first four, and Energy Northwest has the option to build up to eight additional modules. Furthermore, Amazon's Climate Pledge Fund is the lead anchor in a 500-million-dollar financing round for SMR and nuclear fuel developer X-energy. (Olick 2024)

5. Conclusions

Small modular reactors are promising a novel way of producing carbon free dispatchable energy production for power-intensive sectors such as data center industry. SMRs offer increased scalability, allowing tailoring of energy supply to evolving demand. Additionally, the novel SMR technology promises faster construction time compared to the traditional nuclear power plants. Economically, the promise of modularization and serial production suggests that, with sufficient scale, SMRs could rival or surpass traditional large nuclear plants in cost-effectiveness. Recent studies reveal that initial costs of SMRs may be higher, but learning effects and streamlined construction could yield more competitive pricing as more SMR units are deployed.

Nevertheless, the path to widespread commercial adoption of SMRs is not without challenges. While some SMR types have reached operational or near-operational status, regulatory acceptance may cause an obstacle, especially given the shift novel design features such as passive safety systems, which, despite their potential for enhanced safety and cost-efficiency, introduce new uncertainties for nuclear safety authorities worldwide.

The data center industry, representing one of the fastest-growing energy consumers, is at the forefront of embracing nuclear innovations. Major companies such as Microsoft, Google, and Amazon are entering long-term agreements and investing significant capital into both traditional nuclear power plant and SMR projects. These partnerships aim not only to secure reliable baseload power but also to advance broader sustainability goals by reducing carbon emissions.

This trend coincides with the studies presented in this paper, indicating that data centers are more sensitive to load shifting rather than by fluctuations in electricity prices. Traditional and AI data centers typically operate on a 24/7 basis, requiring continuous uninterrupted electricity to ensure the services it provides. This aligns well with the operational characteristics of nuclear power plants, both SMRs and large reactors, which are designed to supply reliable carbon free baseload power over long periods.

In conclusion, SMRs and nuclear power in general are expected to play a crucial role in meeting the energy needs of modern data centers. Their potential to provide clean, reliable, and scalable power makes them an essential component of the global transition towards carbon free energy systems and infrastructures which build our digital spaces. The coming decade will be pivotal in demonstrating the feasibility of SMRs, and, by extension, their impact on shaping the sustainable energy landscape for years to come.

6. References

Bilsen, V., Gröger, J., Devriendt, W., et al. (2022) "Study on Greening Cloud Computing and Electronic Communications Services and Networks: Towards Climate Neutrality by 2050", European Commission. Available at: https://digital-strategy.ec.europa.eu/en/library/study-greening-cloud-computing-and-electronic-communications-services-and-networks-towards-climate (Accessed 20 March 2025).

Campbell, T., (2024) "What is blockchain?", Business Insider [Online]. Available at: https://www.businessinsider.com/personal-finance/investing/what-is-blockchain (Accessed 17 February 2025)

Chandler, S. and Campbell, T., (2024) "What is Proof of Work (PoW) in cryptocurrency?", Business Insider [Online]. Available at: https://www.businessinsider.com/personal-finance/investing/proof-of-work (Accessed 17 February 2025)

Chen, C., (2025) "How a top Chinese AI model overcame US sanctions", MIT Technology Review [Online]. Available at: https://www.technologyreview.com/2025/01/24/1110526/china-deepseek-top-ai-despite-sanctions/ (Accessed 14 February 2025)

Cohen, A., (2024) "AI Is Pushing The World Toward An Energy Crisis", Forbes [Online]. Available at: https://www.forbes.com/sites/arielcohen/2024/05/23/ai-is-pushing-the-world-towards-an-energy-crisis/ (Accessed 5 February 2025).

De Vries, A., (2023) "Cryptocurrencies on the road to sustainability: Ethereum paving the way for Bitcoin", Patterns, Volume 4, Issue 1, 100633, ISSN 2666-3899, [Online]. Available at: https://doi.org/10.1016/j.patter.2022.100633 (Accessed 17 February 2025).

DeepSeek-AI, Guo, D., Yang, D., et al., (2025) "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning", arXiv [Online]. Available at: https://arxiv.org/abs/2501.12948 (Accessed 14 February 2025).

DeepSeek-AI, Liu, A., Feng, B., et al., (2024) "DeepSeek-V3 Technical Report", arXiv [Online]. Available at: https://arxiv.org/abs/2412.19437 (Accessed 14 February 2025).

Disavino, S., (2022) "Oil surges to multi-year highs as Russian supply shortfall looms", Reuters. Available at: https://www.reuters.com/business/energy/oil-rebounds-escalating-ukraine-conflict-raises-supply-concerns-2022-03-04/ (Accessed 4 February 2025).

Evans, R. and Gao, J., (2016) "DeepMind AI Reduces Google Data Centre Cooling Bill by 40%", Google DeepMind [Online]. Available at: https://deepmind.google/discover/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40 (Accessed: 18 February 2025).

Fellous-Asiani, M., (2022) "Could energy efficiency be quantum computers' greatest strength yet?", The Conversation [Online]. Available at: https://theconversation.com/could-energy-efficiency-be-quantum-computers-greatest-strength-yet-191989 (Accessed 18 February 2025)

Hajiaghapour-Moghimi, M., Hajipour, E., Azimi Hosseini, K., Vakilian, M., and Lehtonen, M. (2024) "Cryptocurrency mining as a novel virtual energy storage system in islanded and grid-connected microgrids," *International Journal of Electrical Power & Energy Systems*, 158, 109915. Elsevier. https://doi.org/10.1016/j.ijepes.2024.109915

Hnayno, M., Chehade, A., Klaba, H., Polidori, G., and Maalouf, C. (2023). "Experimental investigation of a data-centre cooling system using a new single-phase immersion/liquid technique." Case Studies in Thermal Engineering, 45, 102925. Available at: https://doi.org/10.1016/j.csite.2023.102925 (Accessed 18 February 2025).

IEA, (2024) "Electricity 2024, Analysis and forecast to 2026", International Energy Agency [Online]. Available at: https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf (Accessed 4 February 2025).

IAEA, (2024a) "Data Centres, Artificial Intelligence and Cryptocurrencies Eye Advanced Nuclear to Meet Growing Power Needs", International Atomic Energy Agency, Vienna. [Online]. Available at: https://www.iaea.org/bulletin/data-centres-artificial-intelligence-and-cryptocurrencies-eye-advanced-nuclear-to-meet-growing-power-needs (Accessed 4 February 2025).

IAEA, (2024b) "Small Modular Reactors: Advances in SMR Developments 2024", Non-serial Publications, International Atomic Energy Agency, Vienna. Available at: https://doi.org/10.61092/iaea.3o4h-svum (Accessed 25 March 2025).

IAEA, (2024c) "Small Modular Reactor Technology Catalogue 2024", International Atomic Energy Agency, Vienna. Available at: https://aris.iaea.org/publications/SMR_Catalogue_2024.pdf (Accessed: 25 March 2025).

Intel. (n.d) "CPU vs. GPU: What's the Difference?", Available at: https://www.intel.com/content/www/us/en/products/docs/processors/cpu-vs-gpu.html (Accessed 5 February 2025).

Juuti, P., (2023) "Datakeskuksen hukkalämpö ei aluksi kelvannut Helsingin kaukolämpöyhtiölle – sitten maailma muuttui ja nyt lämpö virtaa Pitäjänmäessä", Yle [Online]. Available at: https://yle.fi/a/74-20035644 (Accessed 18 February 2025)

Kamiya, G. and Bertoldi, P., (2024) "Energy Consumption in Data Centres and Broadband Communication Networks in the EU, Publications Office of the European Union", Luxembourg, doi:10.2760/706491, JRC135926.

Kimball, S., (2024a) "Microsoft's carbon emissions have risen 30% since 2020 due to data center expansion", CNBC [Online]. Available at: https://www.cnbc.com/2024/05/15/microsofts-carbon-emissions-have-risen-30percent-since-2020-due-to-data-center-expansion.html (Accessed 5 February 2025).

Kimball, S., (2024b) "Constellation Energy to restart Three Mile Island nuclear plant, sell the power to Microsoft for AI", CNBC [Online]. Available at: https://www.cnbc.com/2024/09/20/constellation-energy-to-restart-three-mile-island-and-sell-the-power-to-microsoft.html (Accessed 5 February 2025).

Li, P., Yang, J., Islam, M. A., and Ren, S., (2025) "Making AI Less 'Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models", arXiv [Online]. Available at: https://arxiv.org/abs/2304.03271 (Accessed 5 February 2025).

Nøland, J.K, & Hjelmeland M., & Korpås M., (2024a) "Will Energy-Hungry AI Create a Baseload Power Demand Boom?" IEEE Access. PP. 1-1. 10.1109/ACCESS.2024.3440217. (Accessed 21 March 2025).

Nøland, J.K, & Hjelmeland M., & Korpås M., (2024b) "Cost Projections of Small Modular Reactors: A Model-Based Analysis", 2024, 10.13140/RG.2.2.15471.04004.

Olick, D., (2024) “Amazon goes nuclear, to invest more than $500 million to develop small modular reactors”, CNBC [Online]. Available at: https://www.cnbc.com/2024/10/16/amazon-goes-nuclear-investing-more-than-500-million-to-develop-small-module-reactors.html (Accessed 5 February 2025).

Palo Alto Networks. (2025) “What is a Data Center?”. Available at: https://www.paloaltonetworks.com/cyberpedia/what-is-a-data-center (Accessed 5 February 2025).

Siciliano, J., (2025) “Google says datacenters, AI cause its carbon emissions to rise sharply”, S&P Global [Online]. Available at: https://www.spglobal.com/commodity-insights/en/news-research/latest-news/electric-power/070924-google-says-datacenters-ai-cause-its-carbon-emissions-to-rise-sharply (Accessed 5 February 2025).

Singer, B., Bingham D, et al., and (2024) “AI/data centers' global power surge and the Sustainability impact”, Goldman Sachs [Online]. Available at: https://www.goldmansachs.com/images/migrated/insights/pages/gs-research/gs-sustain-generational-growth-ai-data-centers-global-power-surge-and-the-sustainability-impact/sustain-data-center-redaction.pdf (Accessed 5 February 2025)

Singer, B., Corbett, B., Davenport, C., et al., (2025) “AI/data centers’ global power surge and the Sustainability impact”. [Online]. Available at: https://www.goldmansachs.com/pdfs/insights/goldman-sachs-research/five-drivers/FinalReport.pdf (Accessed 26 February 2025).

Srivathsan, B., Sorel, M., Bhan, A., Sachdeva, P., Batra, H., Sharma, R., Gupta, R., and Choudhary, S., (2024) "AI power: Expanding data center capacity to meet growing demand", McKinsey [Online]. Available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand (Accessed: 24 March 2025).

Stevens, P., (2024) “Google signs deal with nuclear company as data center power demand surges”, CNBC [Online]. Available at: https://www.cnbc.com/2024/10/14/google-inks-deal-with-nuclear-company-as-data-center-power-demand-surges.html (Accessed: 5 February 2025).

STUK, (2024) “Small modular reactors”, [Online]. Available at: https://stuk.fi/en/small-modular-reactors (Accessed 24 March 2025).

U.S. Department of Energy, (2023) “At COP28, Countries Launch Declaration to Triple Nuclear Energy Capacity by 2050, Recognizing the Key Role of Nuclear Energy in Reaching Net Zero“, [Online]. Available at: https://www.energy.gov/articles/cop28-countries-launch-declaration-triple-nuclear-energy-capacity-2050-recognizing-key (Accessed 26 February 2025).

Zhang, Q., Meng, Z., Hong, X., Zhan, Y., Liu, J., Dong, J., Bai, T., Niu, J., Deen, M.J., (2021) “A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization”. Journal of Systems Architecture, Volume 119, 102253. [Online]. Available at: https://www.sciencedirect.com/science/article/pii/S1383762121001739 (Accessed 20 March 2025).

Zoting, S., (2024) “Hyperscale Data Center Market Size, Share, and Trends 2024 to 2034”, Precedence Research [Online]. Available at: https://www.precedenceresearch.com/hyperscale-data-center-market (Accessed 18 February 2025).