National Real Estate Advisors

The AI revolution: Data center industry impacts

How supply, demand and technological advancements driven by AI are affecting data center infrastructure and investment

As the technology landscape continues to evolve, the impact of artificial intelligence (AI) on almost all industries cannot be overstated, including the data center industry. Cloud creation and migration have driven significant changes in the data center industry for many years, and now the advent of widely accessible AI tools promises to reshape the data center landscape even more dramatically. In this article, we examine two crucial aspects of AI's impact on the data center industry: supply-and-demand dynamics and the evolution of new technologies.

1. Supply-and-demand dynamics

The amount of data we produce is growing at an astonishing rate, doubling approximately every two to three years. This growth rate equates to an annual compound increase of 50 percent to 60 percent. As AI continues to revolutionize various industries, it demands substantial storage and computational resources. With the progression of AI, many tasks previously handled by humans are becoming automated, leading to the generation of even more data through technologies such as the internet of things (IoT) and autonomous vehicles.

This surge in data and reliance on AI poses significant challenges for our computing infrastructure. During the past decade, the energy required to train the largest AI models has increased by a factor of 300,000, with the demand roughly doubling every three months. This rate of increase is six times faster than Moore's law, which historically predicted the doubling of computing power approximately every two years.[1] Such exponential growth highlights the urgent need for data center solutions that are not only scalable but also highly efficient, to meet the escalating demands of AI development and deployment.

Infrastructure demand, minimal availability

Major technology corporations, such as Microsoft, Google, Amazon, Meta, Oracle and IBM, all have booming cloud businesses but are now channeling billions of dollars into the infrastructure and applied research necessary to gain a competitive advantage in deploying AI across their product offerings. Despite these efforts, the challenge of finding sufficient data center space persists, as vacancy rates in key markets are at near-record lows. For instance, Northern Virginia has a vacancy rate of just 0.94 percent, making it the tightest data center market in the United States. In response to this scarcity, hyperscale companies are strategically preleasing space up to three years before the completion of construction, highlighting the fierce competition for these increasingly scarce resources.[2]

The apparent preference of hyperscale companies to lease a significant amount of data center space rather than exclusively self-build highlights their interest in partnering with reputable third-party operators and underscores the substantial challenges presented by supply-chain disruptions and power constraints. Simply increasing the physical infrastructure is insufficient to overcome these obstacles; instead, creative solutions and strategic foresight are crucial. Consequently, skilled developers are becoming increasingly valued by hyperscale's for their ability to apply their expertise in choosing locations; obtaining necessary approvals; and handling design, construction and management. This strategy emphasizes the critical role of specialized expertise in managing the intricacies of enhancing data center capacity under limiting conditions.

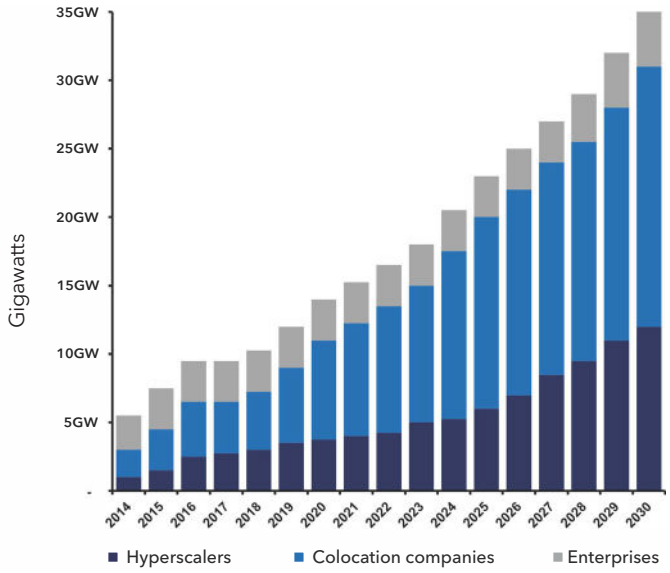
Rising energy consumption, power challenges

The increasing energy demands of AI workloads introduce further complications. As AI chips grow more powerful, their energy requirements surge, dwarfing the power consumption typical of conventional data center operations. Whereas traditional data center activities, such as cloud storage and processing, consume 10 kilowatts (kW) per rack to 14 kW per rack, AI-related tasks, including algorithm and model training, demand substantially more, ranging from 40 kW to 60 kW. Consequently, the overall power consumption of data centers has been surging, with an annual growth rate of 12 percent since 2021. This trend is expected to persist, with projections indicating a continued increase of about 10 percent per year during the coming decade.[3]

Data center operators are facing critical challenges in meeting the escalating demand for power and storage due to power constraints. Many markets are reaching or exceeding their power-capacity limits, making the construction of new power infrastructure expensive and time-consuming.

This trend underscores the need for energy-efficient solutions and sustainable practices to mitigate the environmental impact of data

U.S. data center power consumption 2019–2030



Source: McKinsey & Co.

Image created with generative artificial intelligence

center operations. Transitioning to cleaner energy sources and implementing innovative cooling technologies will be essential in curbing energy consumption and reducing carbon footprints.

2. New technologies and processes

As this digital landscape evolves at an unprecedented rate, data centers stand at the forefront of technological advancement. The integration of AI has revolutionized the way data centers operate, creating new technologies that optimize efficiency, enhance performance and shape the future of digital infrastructure.

Among several recent advancements, the introduction of the transformer model in the paper "Attention Is All You Need"[4] in 2017 revolutionized natural language processing (NLP) and laid the foundation for many of the rapid advancements we are experiencing today, including the development of generative pretrained transformers (GPT). The key technological innovation of the transformer model is its self-attention mechanism, which allows it to weigh the importance of different words in a sentence irrespective of their positional distance.

The computational innovations brought forth by transformer models, such as GPT, have significantly increased the efficiency and capabilities of natural language understanding and generation, but they also come with substantial requirements for computational resources.

Two model processes currently driving data center operations – training and inference – reflect these demands. Inference AI demands low latency and high connectivity, making enterprise colocation data centers the ideal choice due to their proximity to end-users. Conversely, the training of AI models, especially those as large and complex as transformer-based architectures, requires immense computational power and energy, making hyperscale models and large campuses with low energy costs preferable. These differing requirements underscore the need for versatile data center solutions tailored to the specific tasks of the digital age, highlighting how the evolution of AI technologies such as the transformer model directly impacts and shapes the infrastructure of data center operations.

GPU technology and supply

The exponential growth of AI has also led to demand for graphical processing units (GPUs), which are specialized processors designed for accelerated rendering of images and videos, and efficient execution of complex mathematical and geometric calculations, often used in gaming, graphics design and AI tasks.[5] Both training and inference AI tasks necessitate substantial computational power, driving the need for advanced GPU technology. Recent advancements have resulted in GPUs capable of completing a greater number of calculations, increasing efficiency while lowering overall costs. Persistent shortages of GPUs, however, pose challenges for data center operators, prompting innovative solutions such as specialized cloud providers, such as CoreWeave, which offer access to AI-level GPUs tailored to the needs of smaller AI startups. By leasing space from data center owners and acting as cloud operators, CoreWeave provides affordable GPU resources, contrasting with higher prices charged by major cloud providers, such as AWS and Azure.

New cooling solutions

As AI workloads become increasingly intensive, traditional air-cooling methods struggle to meet the demands of high-power density applications. Liquid cooling has emerged as a solution, offering greater efficiency in transferring heat compared with air. Newer GPUs, with their enhanced power use, may necessitate liquid-cooling solutions to maintain optimal temperatures. While liquid cooling presents challenges in deployment and servicing, its sustainability benefits – including reduced power and water usage – make it a compelling option for data center operators looking to minimize environmental impact.

Conclusion

The growth potential for data center companies is substantial, driven by the ever-expanding AI ecosystem. While it faces significant challenges, as AI continues to shape the future of technology and business, the data center industry remains a crucial enabler, which we believe makes it a compelling area for investment and strategic consideration for institutional investors.

Notes: [1] Bank of America Global Research, "Me, Myself and AI – Artificial Intelligence Primer," March 8, 2023, Artificial Intelligence…Is Intelligent! (bofa.com); [2] CBRE, "North America Data Center Trends H1 2023," Sept. 6, 2023; [3] McKinsey & Co., "Investing in the Rising Data Center Economy," Jan. 17, 2024; [4] Ashish Vaswani et. al., "Attention Is All You Need," 2017; [5]Andreessen Horowitz, "Navigating the High Cost of AI Compute," April 27, 2023

CONTRIBUTORS

Jeff Kanne
President & CEO
National Real Estate Advisors

TJ Hofheimer
Managing Director, Portfolio and Asset Management
National Real Estate Advisors

COMPANY PROFILE

An investment manager developing, operating and managing commercial real estate projects across the United States, **National Real Estate Advisors** manages separate accounts and multiple commingled investment vehicles, including a data center fund.

CORPORATE CONTACT

Carl Boswell | Director, Investor Relations
cboswell@natadvisors.com | +1 202-378-1793