INTELLIGENCE UPDATE

Uptime's predictions for 2025: an overview

Douglas Donnellan 14 Jan 2025

At the start of each calendar year, Uptime Intelligence compiles a short list of trends or predictions relevant to the digital infrastructure sector for the coming year (and years) ahead. These aim to look beyond the more obvious trends and examine some of the latest developments and challenges shaping the data center industry.

Each list has been largely accurate, and many trends remain relevant heading into 2025 (see *Uptime's predictions 2022 to 2024 — relevant and actionable?*). Over the years, many predictions have followed similar themes, particularly around sustainability, power, and adopting new technologies. These issues are present again in 2025 — however, this year is also distinct in that AI forms a core element in each prediction.

The global spotlight on AI, which intensified in 2024, has placed the data center industry under greater scrutiny than in previous years. Massive investments in AI infrastructure and the proliferation of generative AI have garnered significant attention from major media outlets, prompting widespread speculation about the technology's potential benefits and the environmental and social impacts.

The rise of AI brings both opportunities and uncertainties for data centers. Organizations are grappling with challenges around capacity planning, power availability, and meeting the performance demands of increasingly dense IT environments. In 2025, these uncertainties will shape the industry's relationship with AI as significant investments materialize and diverse strategies are put to the test. Below is a summary of Uptime Intelligence's predictions for 2025.

Summary of the data center predictions for 2025

1. Data center resource use will raise deep questions — and opposition

Data center developments will become increasingly politicized in the coming years. Despite rising public opposition over environmental concerns and unmet promises of job creation,

governments will support rapid expansion for the perceived economic and AI-driven benefits. As a result, climate commitments will be downgraded or postponed as optimism around AI continues to drive growth.

2. Most AI models will be trained in the cloud

Most investments in AI infrastructure for large-scale training will come from hyperscalers and cloud providers, as enterprises avoid the cost and complexity of on-premises GPU clusters. Enterprises will rely on public cloud services and pre-trained foundation models, fine-tuning them to reduce computational overhead.

3. Grid demand will require active participation from data centers

New and expanded data centers will increasingly be expected to provide or store power and possibly even shed loads to support grids. Data center operators running non-latency-sensitive workloads, such as specific AI training tasks, could be financially incentivized or mandated to reduce power use when required.

4. AI to trigger radical overhaul of data center electrification

Infrastructure requirements for next-generation AI will force operators to explore new power architectures. As a result, innovations in data center power delivery, such as deploying medium-voltage distribution to the IT space and solid-state transformers, will begin to emerge.

5. Nvidia's vision for data centers is not without alternatives

The performance of Nvidia's GPUs has led to the company's near-monopoly in the enterprise GPU market, but they are costly, scarce, and challenging to deploy. Some organizations will seek alternatives to these power-hungry GPUs, especially for inference tasks that require fewer computing resources. At the same time, there are signs that AI hardware will become more diverse in 2025.

The full report, *Five data center predictions for 2025*, is available to subscribers here.
The report from the previous year, *Five data center predictions for 2024*, is available here.

ABOUT THE AUTHOR

Douglas Donnellan

16 Feb 2026

Douglas is a Research Analyst at Uptime Institute covering sustainability in data centers. His background includes environmental research and communications, with a strong focus on education.

ddonnellan@uptimeinstitute.com

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. With over 4,000 awards issued in over 122 countries around the globe, and over 1,100 currently active projects in 80+ countries, Uptime has helped tens of thousands of companies optimize critical IT assets while managing costs, resources, and efficiency. For over 30 years, the company has established industry-leading benchmarks for data center performance, resilience, sustainability, and efficiency, which provide customers assurance that their digital infrastructure can perform across a wide array of operating conditions at a level consistent with their individual business needs. Uptime's Tier Standard is the IT industry's most trusted and adopted global standard for the design, construction, and operation of data centers.

Offerings include the organization's Tier Standard and Certifications, Management & Operations reviews and assessments including SCIRA-FSI financial sector risk assessment, the Sustainability Assessment, and a broad range of additional risk management, performance, availability, and related offerings. Uptime Education training programs have been successfully completed by over 100,000 data center professionals, such as the much-valued ATD (Accredited Tier Designer) and AOS (Accredited Operations Specialist). The Uptime Education curriculum has been expanded by the acquisition of CNet Training Ltd. In 2023.

Uptime Institute is headquartered in New York, NY, with offices in London, Sao Paulo, Dubai, Riyadh, and Singapore, and full-time Uptime professionals based in over thirty-four countries around the world.

For more information, visit www.uptimeinstitute.com