AI Infrastructure - Market Share Analysis, Industry Trends & Statistics, Growth Forecasts (2024 - 2029)

https://marketpublishers.com/r/A46DC674FD25EN.html

Date: July 2024
Pages: 200
Price: US$ 4,750.00 (Single User License)
ID: A46DC674FD25EN

Abstracts

The AI Infrastructure Market size is estimated at USD 68.46 billion in 2024, and is expected to reach USD 171.21 billion by 2029, growing at a CAGR of 20.12% during the forecast period (2024-2029).

Key Highlights

An AI infrastructure involves the technology that aids in machine learning (ML), merging machine learning and artificial intelligence solutions to produce scalable, dependable, and specialized data solutions. It comprises the hardware, software, and networking aspects that empower organizations to efficiently develop, deploy, and oversee artificial intelligence (AI) projects. The increasing advancements in AI among diverse end users are anticipated to propel the market.

The global AI infrastructure market has experienced significant growth in recent years, thanks to the increasing utilization of artificial intelligence (AI) technologies in various industries. Hardware, software, and networking components play a crucial role in supporting AI workloads such as data processing, training, and inference. With the surge in data volume from diverse sources and enhancements in AI algorithms, businesses are increasingly investing in advanced AI infrastructure to maximize the benefits of AI-driven solutions.

In recent years, the AI industry has witnessed an unparalleled growth. With the increasing adoption and utilization of generative AI in various fields, more and more companies are transitioning from research and evaluation to actual production. This shift is anticipated to create a significant demand for robust computational infrastructure across different sectors, particularly in enterprises. One advantage of this trend is the

scalability it offers, enabling businesses to easily adjust their operations based on demand, particularly with the utilization of cloud-based AI/ML solutions. Additionally, automation plays a crucial role by reducing errors and enhancing the speed of delivering results through the elimination of repetitive tasks.

Companies are now realizing the benefits of integrating artificial intelligence (AI) into their operations to enhance efficiency and cut costs through automated processes. This has led to the adoption of autonomous systems to streamline operations and revolutionize customer service, such as AI-driven chatbots, driving innovation to unprecedented levels. AI comprises algorithms designed to address particular issues and is most effective when dealing with large amounts of top-notch Big Data. Chatbots have the potential to reduce business operational expenses by as much as 30%.

Moreover, with the shift of IT strategy emphasis from data management to intelligent action, businesses are increasingly acknowledging the significance of AI in assisting individuals in problem-solving, decision-making, and creative pursuits. Businesses understand that the integration and utilization of AI are essential for their ongoing expansion in the competitive landscape, presenting numerous possibilities for leveraging AI to spur innovation, establish connections, and pinpoint and nurture new advancements. For instance, AI apps like ChatGPT and DALL-E from OpenAI have been some significant AI developments in recent times.

The rapid integration of artificial intelligence in various sectors is fueling the significant growth of the worldwide AI infrastructure market. Nevertheless, the need for more skilled professionals proficient in AI technology remains a major challenge for this industry. The necessity for adept AI experts has surged as companies and institutions strive to leverage AI's capabilities to drive innovation and gain a competitive advantage.

Significant growth and advancements are projected for the global AI infrastructure market in the coming years. Businesses are expected to persist in adopting AI technology, resulting in a continued high demand for AI infrastructure across various sectors. Cloud-based AI solutions, edge AI infrastructure, and AI as a Service (AIaaS) services are expected to dominate the industry. Amid the COVID-19 pandemic, cloud computing played a crucial part in democratizing AI by providing AIaaS solutions that allow businesses of all sizes to launch AI applications without making major upfront infrastructure investments.

AI Infrastructure Market Trends

Enterprises to be the Fastest Growing End User

AI technology has introduced higher levels of automation to various areas, including vehicles, kiosks, utility grids, and financial networks. Within the enterprise segment of the AI infrastructure market, businesses from different sectors leverage AI to outperform competitors, drive innovation, and enhance operational effectiveness. AI is transforming the way corporate operations and strategic decisions are made in all industries, catering to both large multinational corporations and small to medium-sized enterprises (SMEs).

Enterprise IT departments require robust infrastructure capable of managing high resource requirements, intricate workflows, and operational burdens stemming from new applications and data silos. As the use of generative AI continues to rise, infrastructure needs to satisfy the requirements for securely creating and implementing models are also increasing. Among other companies, NVIDIA is consistently striving to expedite AI infrastructure deployment from cloud environments to on-premise data centers to facilitate quicker delivery of enterprise applications at reduced costs and faster returns. In addition, the investments made by other firms to enhance enterprise capabilities in adopting AI are anticipated to propel the market further.

In July 2023, NVIDIA announced that NVIDIA DGX Cloud is now widely accessible, offering powerful tools to transform almost any business into an AI-driven organization. This service is readily available, featuring thousands of NVIDIA GPUs accessible through Oracle Cloud Infrastructure. NVIDIA's infrastructure is also present in the United States and the United Kingdom, providing further accessibility and support for AI initiatives.

The move from paper-based information to digital data was the second trend facilitating the development of enterprise AI systems. The success of enterprise AI depends on the amount, accuracy, and range of data. More varied and extensive data enables companies to reach more detailed and precise conclusions. AI has become an essential element of business strategy and operations in this digital transformation era. Its capability to analyze large volumes of data and provide actionable insights helps businesses enhance productivity and operational efficiencies. To remain competitive, it is crucial for enterprises to integrate AI into a majority of their business operations effectively.

AI hardware implemented in enterprise data centers has the potential to greatly enhance the effectiveness of data operations. Rising investments in data centers are projected to greatly boost the utilization of AI infrastructure. Cloudscene reported that

there were 2,701 data centers in the United States by January of the previous year, with an additional 487 in Germany. The United Kingdom ranked third among countries with 456 data centers, closely followed by China with 443. The substantial quantity of data centers presents a promising opportunity for expanding the research market.

Furthermore, it is imperative for the organization that seeks to automate the global landscape to first automate its operations. With the increasing size and complexity of data loads, as well as the expansion of infrastructure beyond traditional data centers to encompass the cloud and edge, the speed at which these new environments are set up, optimized, and dismantled will soon surpass the capabilities of human operators. Consequently, the presence of AI at the operational level will be indispensable in meeting the demands of AI-driven initiatives.

According to a report by Flexera Software, 75% of enterprise respondents are planning to adopt Microsoft Azure for public cloud usage by 2023. The top cloud computing platform providers globally, including AWS, Microsoft Azure, and Google Cloud, are known as hyper scalers. The increasing demand for AI infrastructure in businesses is a significant factor driving the widespread adoption of AI technology. Various industries leverage AI to enhance productivity, improve customer service, and achieve strategic objectives. Enterprises will play a crucial role in shaping the future development of AI-driven innovations, reinforcing AI's position as a transformative technology that empowers companies to thrive in a data-driven landscape.

For instance, Microsoft continues to invest substantially in various GenAI vendors worldwide, demonstrating its strong presence in the generative AI market. In April 2024, the tech giant announced its intention to invest a significant amount of USD 1.5 billion in G42, an AI vendor in the United Arab Emirates (UAE). G42 will leverage the Microsoft Azure cloud to deploy its AI applications and services. At the same time, both companies collaborate to establish digital infrastructure in the regions where G42 operates, namely the Middle East, Central Asia, and Africa. This strategic partnership highlights Microsoft's commitment to advancing AI technologies globally.

North America to Hold Major Market Share

North America is contributing a significant portion to the market share due to the adoption of AI infrastructure in various industries in the United States, supported by the development of AI programs and fueling the market's growth. The AI infrastructure

market in the United States is characterized by a broad set of respondents, including tech giants or big techs, startups, and cloud service providers competing to provide creative and advanced solutions that address the specific demands of enterprises looking to deploy AI technologies.

Recently, AI infrastructure has significantly increased in popularity, changing numerous fields and applications. Cloud service providers such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, and IBM, among others, contribute to the AI infrastructure market. Such big tech firms provide a wide range of AI-focused hardware acceleration. Their extensive worldwide infrastructure allows businesses of all sizes to access AI resources on a pay-as-you-go-basis, making AI adoption more accessible.

Furthermore, government investment and initiatives, tech startups, and infrastructure development in multiple sectors, such as healthcare, enterprises, and others, further fuel the growth of the AI infrastructure market. AI development and applications are shaping the US technological landscape's future and are expected to boost the market during the forecast period. For instance, in November 2023, AI strategies were released by the Department of Defense (DoD) and the Department of State (DoS), indicating that policy is beginning to align with and potentially influence expenditures.

In fiscal year 2022, the US government's spending on AI contracts reached USD 3.3 billion, as reported in a recent study released by Stanford University. Federal agency expenditure on AI technology rose by over USD 600 million compared to the previous year, climbing from USD 2.7 billion in 2021. Most funding was allocated to the decision science, computer vision, and autonomy sectors. In April 2023, the US National Science Foundation also disclosed a USD 16.1 million commitment to back collaborative research infrastructure for AI researchers.

There has been an increase in investments in AI infrastructure projects by both public and commercial sectors to meet the increasing demand. These projects involve data centers, cloud computing services, and high-performance computing (HPC) facilities. Several AI enterprises are emerging in North America with the objective of creating cutting-edge AI platforms and solutions. In April 2024, the US Department of Energy (DOE) detailed multiple efforts to hasten the responsible deployment of AI technologies to drive innovation, bolster America's energy and national security, and address the climate crisis.

The AI infrastructure market in Canada has grown significantly in recent years due to technological advances. The use of AI is changing the way businesses function across

multiple industries across the country. As businesses grasp AI's enormous potential to improve efficiency, productivity, and customer experience, the demand for robust AI infrastructure has increased. Enterprises, by end users, are driving the growth of the market studied in Canada.

Canadian banks were the initial adopters of AI technologies in the financial sector to optimize procedures and improve client experiences. AI-powered applications, virtual assistance, and other solutions manage customer inquiries. As these financial institutions ramp up their AI activities, the demand for AI infrastructure, such as high-performance computing and data storage capabilities, grows.

AI Infrastructure Industry Overview

The AI infrastructure market comprises global players focusing on gaining a point of difference in the contested market space. In addition, startups in the AI infrastructure market, including Cerebras Systems and Graphcore, provide specialized AI accelerators to take on well-established giants like NVIDIA and Google Inc. This market is characterized by moderately high product differentiation, growing product penetration levels, and high levels of competition, with a high level of innovation. Some major market players are Intel Corporation, NVIDIA Corporation, Samsung Electronics Co. Ltd, Micron Technology Inc., and Sensetime Group Inc.

In May 2024, AWS revealed plans to invest USD 12 billion in Singapore over the next four years, focusing on digital infrastructure, cloud, and AI projects. Additionally, the company will expand its operations in the country and conduct workshops with over 100 enterprises to promote AI adoption. This investment is also expected to position AWS to meet the increasing demand for generative AI tools that rely on significant computational power.

In April 2024, Mytek Innovations, an infrastructure technology company, introduced a groundbreaking digital platform that utilizes open artificial intelligence tools to assist contractors in efficiently executing infrastructure projects. This platform, empowered by AI, can evaluate the necessary input cost for project completion and automatically generate a comprehensive plan outlining milestones to ensure timely project delivery. Moreover, it can potentially reduce completion costs by 10-12%.

Additional Benefits:

AI Infrastructure - Market Share Analysis, Industry Trends & Statistics, Growth Forecasts (2024 - 2029)

The market estimate (ME) sheet in Excel format

3 months of analyst support

Contents

7 COMPETITIVE LANDSCAPE

7.1 Company Profiles
  7.1.1 Intel Corporation
  7.1.2 Nvidia Corporation
  7.1.3 Samsung Electronics Co. Ltd
  7.1.4 Micron Technology Inc.
  7.1.5 Sensetime Group Inc.
  7.1.6 IBM Corporation
  7.1.7 Google LLC
  7.1.8 Microsoft Corporation
  7.1.9 Amazon Web Services Inc.
  7.1.10 Cisco Systems Inc.
  7.1.11 Arm Holdings
  7.1.12 Dell Inc.
  7.1.13 Hewlett Packard Enterprise Development LP
  7.1.14 Advanced Micro Devices
  7.1.15 Synopsys Inc.

8 INVESTMENT ANALYSIS

9 FUTURE OF THE MARKET

I would like to order

Product name:	AI Infrastructure - Market Share Analysis, Industry Trends & Statistics, Growth Forecasts (2024 - 2029)
Product link:	https://marketpublishers.com/r/A46DC674FD25EN.html
Price:	US$ 4,750.00 (Single User License / Electronic Delivery)
	If you want to order Corporate License or Hard Copy, please, contact our Customer Service:
	info@marketpublishers.com

Payment

To pay by Credit Card (Visa, MasterCard, American Express, PayPal), please, click button on product page https://marketpublishers.com/r/A46DC674FD25EN.html