+357 96 030922
info@marketpublishers.com

AI Infrastructure Market by Offerings (Compute (GPU, CPU, FPGA), Memory (DDR, HBM), Network (NIC/Network Adapters, Interconnect), Storage, Software), Function (Training, Inference), Deployment (On-premises, Cloud, Hybrid) – Global Forecast to 2030

https://marketpublishers.com/r/A03F513BD83EN.html
Date: December 2024
Pages: 338
Price: US$ 4,950.00 (Single User License)
ID: A03F513BD83EN

Abstracts

The AI Infrastructure market is expected to be worth USD 135.81 billion in 2024 and is estimated to reach USD 394.46 billion by 2030, growing at a CAGR of 19.4% between 2024 and 2030. The AI infrastructure market is being driven by the rapid growth in data generation due to digital transformation, IoT, social media, and e-commerce, which requires advanced computing and storage to manage vast datasets for AI and machine learning models. Additionally, the increasing need for cloud-based AI infrastructure in data centers is reshaping how companies manage complex AI workloads, with major cloud providers investing heavily in AI-ready infrastructure to meet growing global and industry-specific demands. These factors collectively serve as key drivers propelling the growth of the AI infrastructure market.

“Generative AI segment will hold highest CAGR in the forecast period.”

Generative AI is expected to exhibit high growth rate due to a rise in demand for advanced AI applications across industries. Generative AI powers capabilities like content creation, language models, and image synthesis, all of which depends heavily on substantial computational power to train and run large neural networks. This demand requires high infrastructure investments, with an emphasis on high-performance GPU and DPU capabilities able to support intensive processing requirements. As more enterprises seek to capitalize on generative AI's potential, the market for AI

+357 96 030922
info@marketpublishers.com

infrastructure will rise. In November 2024, GMO Internet Group, Inc. (Japan) launched the GMO GPU Cloud, powered by NVIDIA Corporation's (US) H200 Tensor Core GPUs, Spectrum-X Ethernet, BlueField-3 DPUs, and NVIDIA AI Enterprise software, exemplifies the kind of specialized infrastructure needed to support generative AI at scale. This infrastructure, developed on Dell PowerEdge servers, is aligned to support the needs of production-grade generative AI applications, with reduced latency and high-bandwidth capability. Such innovations underline industry momentum in wide-scale deployment of infrastructure optimized for generative AI workloads. Further, this makes the local high-performance cloud solutions to be drivers for growth of this segment.

”Enterprises is projected to grow at a high CAGR of AI Infrastructure market during the forecasted timeline”

The enterprise segment is expected to record high growth in AI Infrastructure Market. Enterprises are increasingly adopting AI infrastructure as they use AI to facilitate digital transformation, enhance operational efficiency, and enhance customer experience. Manufacturing, finance, and retail enterprises are increasing their investments in AI powered predictive analysis, process automation, and customer insight tools that require robust and scalable AI infrastructure. Enterprises are also expanding their AI capabilities by investing in private cloud infrastructure and hybrid cloud models, especially as they seek to protect sensitive data while benefiting from cloud-based AI's flexibility and innovation. Companies are providing cloud and on-premises flexibility along with comprehensive support that aligns with enterprise needs for adaptable, high-performance AI resources that can integrate with existing IT environments. In February 2024, Cisco and NVIDIA announced a partnership to bring AI infrastructure solutions tailored for data centers, designed to streamline deployment and management while offering high computing power necessary for enterprise AI. Its joint offering includes flexible AI infrastructure options for cloud-based and on-premises enterprises, as well as robust networking, security, and end-to-end observability features. Such collaborations support the growth of enterprise AI infrastructure market because companies demand scalable, secure, and manageable infrastructure for the deployment of AI solutions.

“Asia Pacific is expected to hold high CAGR in during the forecast period.”

AI infrastructure market in Asia Pacific will grow at a high CAGR during the forecast period. Countries like China, Japan, South Korea, and India are at the forefront of AI innovation and governments and private sectors in the region are making high

+357 96 030922
info@marketpublishers.com

investments in AI research, infrastructure, and talent development. In September 2024 Lenovo (Hong Kong) announced its mass manufacturing operations for high-performance AI servers in India, and it also unveiled a cutting-edge research & development (R&D) lab, adding to the advancement of Lenovo's Infrastructure Solutions. These are the significant efforts that Lenovo has taken towards the significant positioning of India as a hub in innovation and manufacturing, while supporting 'Make in India' and 'AI for All' vision by the Indian government. Initiatives like these will speed up region's leadership in AI technology and will lead to significant growth in AI infrastructure deployment across Asia Pacific industries. Moreover, as enterprises and governments are driving digital transformation and cloud adoption, the requirement for high-performance AI offerings is increasing, making Asia Pacific one of the fastest-growing markets for AI Infrastructure globally.

Extensive primary interviews were conducted with key industry experts in the AI Infrastructure market space to determine and verify the market size for various segments and subsegments gathered through secondary research. The break-up of primary participants for the report has been shown below: The study contains insights from various industry experts, from component suppliers to Tier 1 companies and OEMs. The break-up of the primaries is as follows:

By Company Type: Tier 1 – 50%, Tier 2 – 20%, and Tier 3 – 30%

By Designation: C-level Executives – 20%, Directors – 30%, and Others – 50%

By Region: North America – 30%, Europe – 20%, Asia Pacific – 40%, and RoW – 10%

The report profiles key players in the AI Infrastructure market with their respective market ranking analysis. Prominent players profiled in this report are NVIDIA Corporation (US), Advanced Micro Devices, Inc. (US), SK HYNIX INC. (South Korea), SAMSUNG (South Korea), Micron Technology, Inc. (US), Intel Corporation (US), Google (US), Amazon Web Services, Inc. (US), Tesla (US), Microsoft (US), Meta (US), Graphcore (UK), and Cerebras (US), among others.

Apart from this, KIOXIA Holdings Corporation (Japan), Western Digital Corporation (US), Mythic (US), Blaize (US), Groq, Inc. (US), HAILO TECHNOLOGIES LTD (Israel), SiMa Technologies, Inc. (US), Kneron, Inc. (US), Rain Neuromorphics Inc. (US), Tenstorrent (Canada), SambaNova Systems, Inc. (US), Taalas (Canada), SAPEON Inc.

+357 96 030922
info@marketpublishers.com

(US), Rebellions Inc. (South Korea), Rivos Inc. (US), and Shanghai BiRen Technology Co., Ltd. (China) are among a few emerging companies in the AI Infrastructure market.

Research Coverage: This research report categorizes the AI infrastructure market based on offerings, function, deployment, application, end user, and region. The report describes the major drivers, restraints, challenges, and opportunities pertaining to the AI infrastructure market and forecasts the same till 2030. Apart from these, the report also consists of leadership mapping and analysis of all the companies included in the AI infrastructure ecosystem.

Key Benefits of Buying the Report The report will help the market leaders/new entrants in this market with information on the closest approximations of the revenue numbers for the overall AI infrastructure market and the subsegments. This report will help stakeholders understand the competitive landscape and gain more insights to position their businesses better and plan suitable go-to-market strategies. The report also helps stakeholders understand the pulse of the market and provides them with information on key market drivers, restraints, challenges, and opportunities.

The report provides insights on the following pointers:

Analysis of key drivers (Rising demand for high-performance computing in AI workloads, government initiatives and investments in AI research and development, and growing implementation of AI and ML solutions across enterprises) influencing the growth of the AI infrastructure market.

Product Development/Innovation: Detailed insights on upcoming technologies, research & development activities, and new product & service launches in the AI infrastructure market.

Market Development: Comprehensive information about lucrative markets – the report analysis the AI infrastructure market across varied regions

Market Diversification: Exhaustive information about new products & services, untapped geographies, recent developments, and investments in the AI infrastructure market

Competitive Assessment: In-depth assessment of market shares, growth strategies, and service offerings of leading players like NVIDIA Corporation (US), Advanced Micro Devices, Inc. (US), SK HYNIX INC. (South Korea),

+357 96 030922
info@marketpublishers.com

SAMSUNG (South Korea), Micron Technology, Inc. (US) among others in the AI infrastructure market.

+357 96 030922
info@marketpublishers.com

Contents

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

13 COMPETITIVE LANDSCAPE

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

+357 96 030922
info@marketpublishers.com

I would like to order

Product name: AI Infrastructure Market by Offerings (Compute (GPU, CPU, FPGA), Memory (DDR, HBM), Network (NIC/Network Adapters, Interconnect), Storage, Software), Function (Training, Inference), Deployment (On-premises, Cloud, Hybrid) – Global Forecast to 2030

Product link: https://marketpublishers.com/r/A03F513BD83EN.html

Price: US$ 4,950.00 (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/A03F513BD83EN.html