NVIDIA Collaborates With Microsoft to Accelerate Enterprise-Ready Generative AI

NVIDIA AI Enterprise Integration With Azure Machine Learning Provides End-to-End Cloud Platform for Developers to Build, Deploy and Manage AI Applications for Large Language Models

Microsoft Build—NVIDIA today announced that it is integrating its NVIDIA AI Enterprise software into Microsoft's Azure Machine Learning to help enterprises accelerate their AI initiatives.

The integration will create a secure, enterprise-ready platform that enables Azure customers worldwide to quickly build, deploy and manage customized applications using the more than 100 NVIDIA AI frameworks and tools that come fully supported in NVIDIA AI Enterprise, the software layer of NVIDIA's AI platform.

"With the coming wave of generative AI applications, enterprises are seeking secure accelerated tools and services that drive innovation," said Manuvir Das, vice president of enterprise computing at NVIDIA. "The combination of NVIDIA AI Enterprise software and Azure Machine Learning will help enterprises speed up their AI initiatives with a straight, efficient path from development to production."

NVIDIA AI Enterprise on Azure Machine Learning will also provide access to the highest-performance NVIDIA accelerated computing resources to speed the training and inference of AI models.

"Microsoft Azure Machine Learning users come to the platform expecting the highest performing, most secure development platform available," said John Montgomery, corporate vice president of AI platform at Microsoft. "Our integration with NVIDIA AI Enterprise software allows us to meet that expectation, enabling enterprises and developers to easily access everything they need to train and deploy custom, secure large language models."

With Azure Machine Learning, developers can easily scale applications, from tests to massive deployments, while using Azure Machine Learning data encryption, access control and compliance certifications to meet security and compliance with their organizational policies requirements. NVIDIA AI Enterprise complements Azure Machine Learning with secure, production-ready AI capabilities and includes access to NVIDIA experts and support.

NVIDIA AI Enterprise includes over 100 frameworks, pretrained models and development tools, such as NVIDIA RAPIDS™ for accelerating data science workloads. NVIDIA Metropolis accelerates vision AI model development, and NVIDIA Triton Inference Server™ supports enterprises in standardizing model deployment and execution.

Availability
The NVIDIA AI Enterprise integration with Azure Machine Learning is available in a limited technical preview.

NVIDIA AI Enterprise is also available on Azure Marketplace, providing businesses worldwide with expanded options for fully secure and supported AI development and deployment.

Additionally, the NVIDIA Omniverse Cloud™ platform-as-a-service is now available on Microsoft Azure as a private offer for enterprises. Omniverse Cloud provides developers and enterprises with a full-stack cloud environment to design, develop, deploy and manage industrial metaverse applications at scale.

About NVIDIA
Since its founding in 1993, NVIDIA (NASDAQ: NVDA) has been a pioneer in accelerated computing. The company's invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined computer graphics, ignited the era of modern AI and is fueling the creation of the industrial metaverse. NVIDIA is now a full-stack computing company with data-center-scale offerings that are reshaping industry. More information at https://nvidianews.nvidia.com/.

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact and availability of NVIDIA's products and technologies, including NVIDIA AI Enterprise, NVIDIA RAPIDS, NVIDIA Metropolis, NVIDIA Triton Inference Server and NVIDIA Omniverse Cloud; NVIDIA's collaboration with Microsoft, including the benefits and impact thereof; and the coming wave of generative AI applications and enterprises seeking secure accelerated tools and services that drive innovation are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of

performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the most recent reports NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its annual report on Form 10-K and quarterly reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

Anna Kiachian
PR Manager
NVIDIA Corporation
+1-650-224-9820
akiachian@nvidia.com