



How Can Generative AI Transform Your Business?

The key to leveraging the groundbreaking technology is starting with a foundation model.

Opportunities of Generative AI Adoption

Whether it's transforming customer service, improving the design, build, and operation of networks, enhancing the shopping experience, or helping public employees perform their jobs more efficiently, generative AI has the power to deliver cost-savings, efficiency benefits, and new revenue opportunities for businesses.

Rather than embark on the creation of a new generative AI foundation model, most businesses begin with a generic foundation model appropriate for their use case and customize it to become savvy about their business, vocabulary, customers, environment, and particular skills, by training it with the organization's proprietary data.

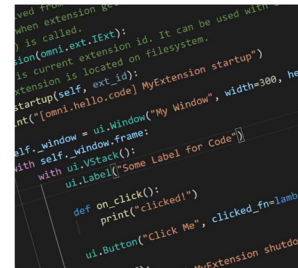
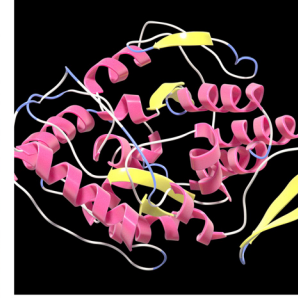
Data must be carefully curated so it can be used to train the foundation model on business specifics or be used in retrieval augmented generation (RAG) to make the trained model highly proficient. All the while, it is vital to take advantage of scalable accelerated infrastructure for training and inference.

Deployment with guardrails ensures safety when the trained model is being used. One can have guardrails specific to their use case, local customs, and sensibilities.

With **NVIDIA AI**, enterprises get the complete platform needed to adopt and accelerate generative AI.

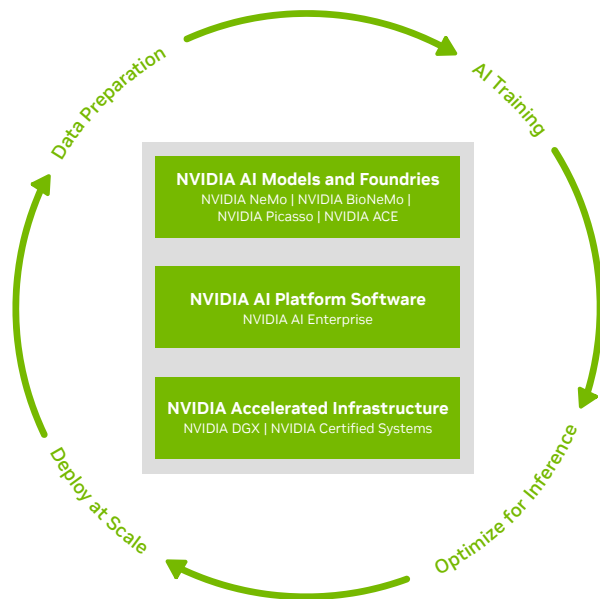
Introducing the NVIDIA AI Platform for Generative AI

As the world's most advanced platform for generative AI, NVIDIA's AI platform is designed to meet your application and business needs. With innovations at every layer of the stack—including accelerated computing, essential end-to-end AI software, pretrained models, and AI foundry services—you can build, customize, and run generative AI models for any application, anywhere. NVIDIA's **AI services**—for language, biology, visual media, and interactive avatars—are equipped with generative model architectures, tools, and workflows running on NVIDIA accelerated infrastructure for training, customizing, optimizing, and deploying generative AI.



Generative AI Challenges

- > **Training Requirements:** When building large language models (LLMs) from scratch, specialized knowledge, software, and hardware is required to be budgeted.
- > **Training Data:** Large, clean datasets must be curated to train the LLM to understand, predict, and be accurate.
- > **Data Storage:** Training data must be made accessible, collected, processed, and managed.
- > **Large-Scale Compute:** The infrastructure required to train an LLM with billions of parameters from scratch can require significant compute infrastructure, which can be expensive.



AI Models and Services

With NVIDIA pretrained models and AI foundries, you can build and customize generative AI models for any application, anywhere. NVIDIA AI foundries are equipped with generative model architectures, tools, and workflows, and they run on NVIDIA accelerated infrastructure for training, customizing, optimizing, and deploying generative AI.

AI Platform Software

To get started with building generative AI, all developers have access to NVIDIA AI frameworks, libraries, pretrained models, and tools and can deploy with NVIDIA AI Enterprise that provides enterprise-grade support, security, stability, and manageability, compatible with industry-leading infrastructure solutions.

Accelerated Infrastructure

The end-to-end NVIDIA AI platform, integrated across hardware and software, gives enterprises the blueprint to a robust, secure infrastructure that's optimized for generative AI.

Consulting and Service Delivery is available via our world-wide, vibrant partner ecosystem that puts solutions together for organizations in concert with their in-house personnel. They know the NVIDIA AI stack well and can provide implementation services and support that best leverage accelerated infrastructure anywhere.

> Training and Inference:

Generative AI models need to be chosen and customized for the business, then deployed for inference on distributed GPU-accelerated infrastructure.

> AI Expertise:

Generative AI requires an understanding of AI, machine learning, and data science principles. Businesses must add new technical expertise to their data science, AI, or application development teams.

NVIDIA AI Platform for Generative AI

AI Models and Foundries

> **NVIDIA NeMo™** enables organizations to build, customize, and deploy generative AI models anywhere. It includes training and inferencing frameworks, a guardrail toolkit, data curation tools, and pretrained models.

> **NVIDIA BioNeMo** rapidly predicts the structure and function of proteins and molecules, accelerating the creation of new drug candidates.

> **NVIDIA Picasso** runs optimized models to generate image, video, 3D, and 360 HDRi content from text or image prompts.

AI Platform Software

> **NVIDIA AI Enterprise** is an enterprise software platform that accelerates the development and deployment of production-grade generative AI. It includes over 100 frameworks, pretrained models, and open-source tools.

Get Enterprise-Grade Support, Stability, Manageability, and Security With NVIDIA AI Enterprise

NVIDIA AI Enterprise is an enterprise-grade, end-to-end software platform that delivers the performance, efficiency, and responsiveness critical to powering the next generation of AI—in the cloud, in the data center, at the network edge, on workstations, and in embedded devices. This includes over 100 frameworks, pretrained models, and open-source tools.

NVIDIA's full-stack architectural approach ensures AI-enabled applications deploy with optimal performance, fewer servers, and less power—resulting in faster insights with dramatically lower costs.

Key Benefits

- > **Time to Solution:** Quickly build custom enterprise-grade models with your own data and domain expertise.
- > **Ease of Use:** Simplify development with a suite of model-making services, pretrained models, cutting-edge frameworks, and APIs.
- > **Production-Ready:** Create enterprise-grade models that protect privacy, data security, and intellectual property.

Ready to Get Started?

To learn more about NVIDIA AI Enterprise, visit: www.nvidia.com/ai-enterprise-suite

To contact us about purchasing our AI software, visit: www.nvidia.com/ai-enterprise-sales/

To help your developers get hands-on training and earn a certificate on the latest generative AI technology, explore the [NVIDIA Generative AI and LLM Training](#).

