



NVIDIA AI Workbench

Develop, collaborate on, and accelerate AI and machine learning workflows—from anywhere.



Challenges for AI and Machine Learning (ML) Developers

The rapid growth of generative AI, starting with the release of ChatGPT and image generation models and bolstered by the proliferation of API services and open models, has inspired enterprises to look into adopting this technology across their business. Data science workstations have the potential to provide AI and ML developers a way to work with data privately and avoid the cost and latency of cloud resources.

Unfortunately, AI and ML developers encounter many challenges to productivity, flexibility, and collaboration. Setting up a development environment on a GPU system isn't always straightforward. Migrating development projects and workloads to different systems can require significant effort and technical skill. And identifying, installing, and managing elements of AI and ML workflows requires time and technical expertise.

What Is AI Workbench?

NVIDIA AI Workbench is a free developer tool that simplifies interactive development workflows for AI, data science, and machine learning development. It automates technical tasks that halt beginners and derail experts. Developers can experience greater productivity, more flexibility, and smoother collaboration.

AI Workbench provides AI and ML developers with:

- > **Streamlined setup and configuration:** Automate tasks such as setting up the target GPU system and configuring a GPU-enabled container in your preferred developer environment.
- > **Portability and reproducibility:** Migrate workloads between different systems and locations without worrying about things breaking. Easily move workloads to the best platform for collaboration, speed, scale, and cost.
- > **Managed AI and ML workflows:** By letting AI Workbench handle file versioning, location changes, and tracking of project dependencies, novice and skilled developers can focus on execution without worrying about configuration and management challenges.

NVIDIA AI Workbench Example Projects

Chat with your documents:

- > Query your documents using retrieval-augmented generation (RAG).
- > Run locally with an NVIDIA AI API or a microservice.

Customize large language models (LLMs) for any scale:

- > Choose from a variety of models, frameworks, and fine-tuning methods.
- > Tailor and deploy on GPU workstations or scale up to data center servers.

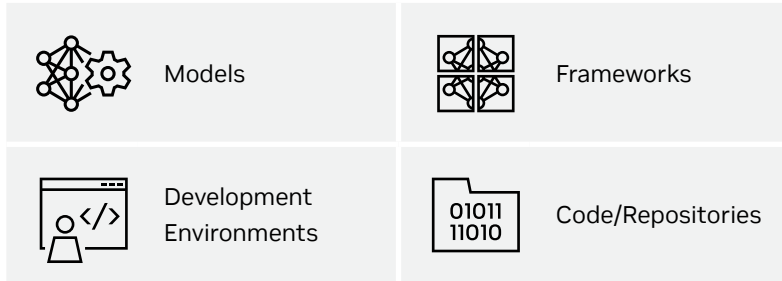
Generate custom images:

- > Create custom images from text prompts by running Stable Diffusion XL locally.
- > Easily reproduce in the cloud to fine-tune the model with your own images.

Workbench Projects

NVIDIA makes it easy to get started with example projects in generative AI using the latest models from Hugging Face and NVIDIA, which can be adapted and optimized for different datasets, GPUs, and use cases.

Workbench Projects are Git repositories that contain simple configuration files that define a containerized development environment. Workbench Projects contain or reference everything needed for interactive development including:



AI Workbench Projects include all resources in a development workflow.

NVIDIA AI APIs

The NVIDIA API Catalog contains endpoints for a wide range of generative AI use cases, including LLMs, semantic search, image generation, and speech. AI Workbench automatically manages API access, making it seamless to include these as a part of your AI workflows. Many of the APIs come with example projects, which can be downloaded and run locally in AI Workbench on your own GPU system.

Availability

NVIDIA AI Workbench is available as a free download for Windows, macOS, and Ubuntu Linux. Users can access a range of example projects for free on GitHub or create their own. Enterprise support for AI Workbench is available for customers who purchase a license for NVIDIA AI Enterprise.

Features

Interface:

- > Command-line interface (CLI)
- > Desktop app

Containers:

- > Docker or Podman
- > Containers from the NVIDIA® NGC™ catalog
- > Custom containers

Model repositories:

- > Hugging Face
- > NVIDIA

Development environments:

- > JupyterLab
- > Visual Studio Code

Operating systems:

- > Windows
- > Ubuntu
- > macOS

Platforms:

- > NVIDIA RTX™ laptops and desktops
- > Data center servers
- > Public cloud

Workbench Benefits

Benefit	Description
Greater Productivity	<ul style="list-style-type: none">> Streamlines AI and ML development workflows> Provides easy access to GPU-accelerated infrastructure> Shortens development cycles and time to value
More Flexibility	<ul style="list-style-type: none">> Provides easy access to a wide variety of AI and data science models and frameworks from industry-leading repositories> Works with your preferred development environment> Scales up or down to almost any compute resource—workstation, data center, or cloud
Seamless Collaboration	<ul style="list-style-type: none">> Offers ability to effortlessly share resources for collaboration on projects> Tracks project dependencies, maintains version control, and reproduces projects faithfully across platforms to enable effective collaboration> Automates tasks to eliminate barriers between developer roles and skill levels—resulting in well-coordinated team projects

Ready to Get Started?

To download and learn more about AI Workbench, visit:
nvidia.com/workbench

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, NGC, and RTX are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 3147703. MAR24

Partner
Logo

